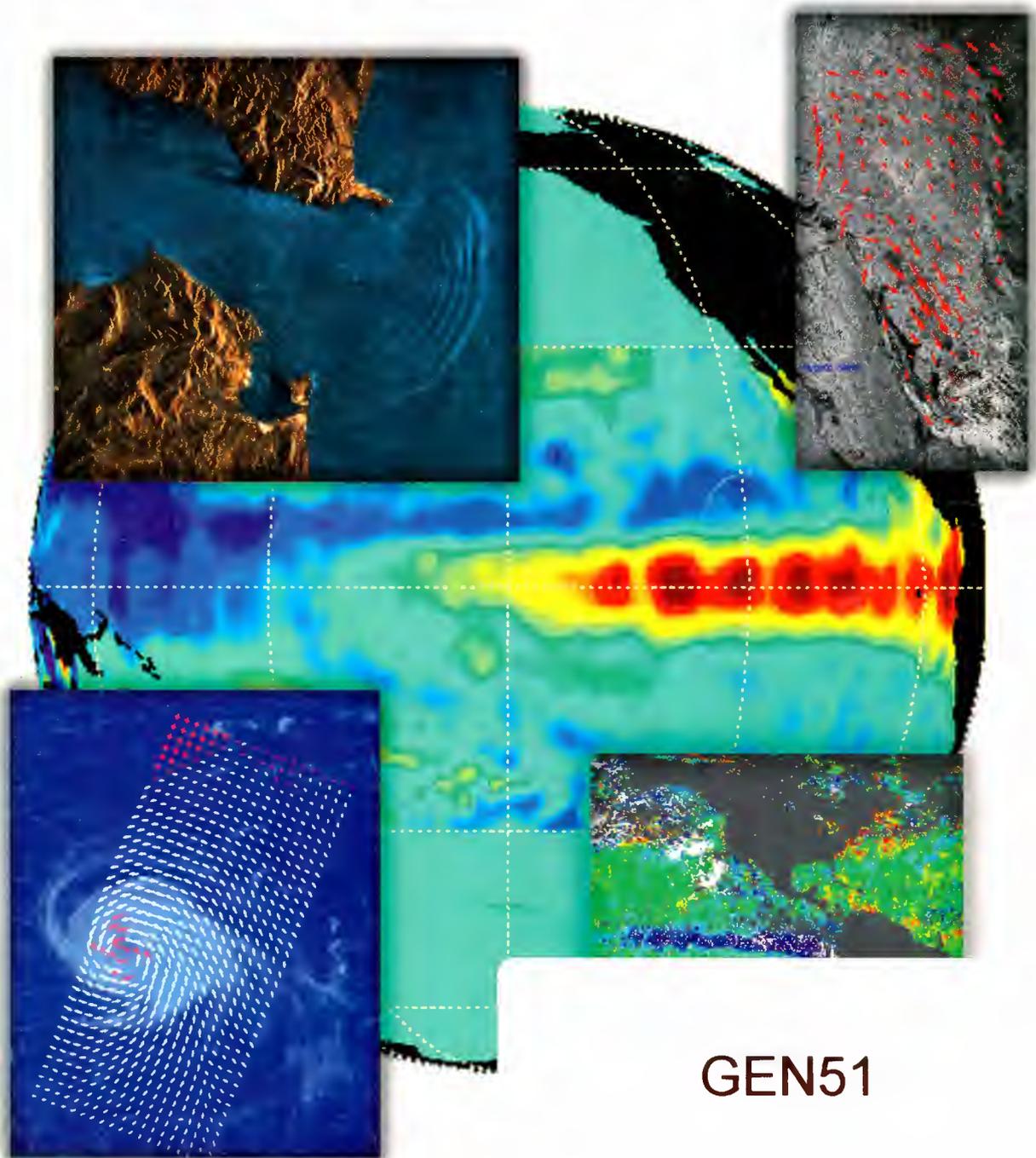


Advances in Oceanography and Sea Ice Research Using ERS Observations



GEN51

Reprinted from the *Journal of Geophysical Research*

**Advances in Oceanography and Sea Ice Research
Using ERS Observation**

Reprinted from the *Journal of Geophysical Research*

Volume 103, Number C4, April 15, 1998



Published by the American Geophysical Union

This cover picture is a collage of ERS satellite images. In the background we show the sea surface topography anomaly (red-to-yellow colours) of the 1997 El Niño derived from the ERS Radar Altimeter (RA) (Acknowledgement Delft Univ. of Technology/ESA-ESRIN). Overlaid this are four images which in a clockwise direction starting from top left include: a) An ERS synthetic aperture radar (SAR) image from the Strait of Gibraltar showing expression of internal waves which are propagating into the Mediterranean Sea (Copyright ESA); b) Ice velocity field (red arrows) in the Kara Sea as derived from two consecutive ERS SAR images taken with three days interval (Acknowledgement NERSC); c) The sea surface temperature anomaly ($> 4^{\circ}$ C in blue colour) of the 1997 El Niño derived from the monthly mean ERS Along Track Scanning Radiometer (ATSR) observations of July 1995 subtracted from July 1997 (Acknowledgement ESA-ESRIN); and d) Combined Meteosat-3 weather image (white cloud pattern) and ERS-1 scatterometer data (arrows) of Hurricane Emily taken on 30 August 1993. The arrows show wind direction, and their length corresponds to wind speed. The red arrows near the eye of the hurricane indicate high wind speeds exceeding 15 m/sec (Copyright ESA).

Advances in Oceanography and Sea Ice Research Using ERS Observations

- Preface (Paper 97JC03412)
J. A. Johannessen, P. Janssen, J. F. Minster, I. Robinson, D. Rothrock, and P. W. Vachon 7753
- Toward the true near-surface wind speed: Error modeling and calibration using triple collocation
(Paper 97JC03180) *Ad Stoffelen* 7755
- Observation of tropical cyclones by high-resolution scatterometry (Paper 97JC01911)
Y. Quilfen, B. Chapron, T. Elfouhaily, K. Katsaros, and J. Tournadre 7767
- Directivity of wind vectors derived from the ERS-1/AMI scatterometer (Paper 97JC02727)
Naoto Ebuchi and Hans C. Graber 7787
- Arabian sea surface winds and ocean transports determined from ERS-1 scatterometer
(Paper 97JC02572) *David Halpern, Michael H. Freilich, and Robert A. Weller* 7799
- VIERS-1 scatterometer model (Paper 97JC02911)
P. A. E. M. Janssen, H. Wallbrink, C. J. Calkoen, D. van Halsema, W. A. Oost, and P. Snoeij 7807
- Analysis of ERS-1/2 synthetic aperture radar wave mode imagettes (Paper 97JC01579)
Vincent Kerbaol, Bertrand Chapron, and Paris W. Vachon 7833
- Mesoscale wind measurements using recalibrated ERS SAR images (Paper 97JC02726)
S. Lehner, J. Horstmann, W. Koch, and W. Rosenthal 7847
- Coastal wind field retrievals from ERS synthetic aperture radar images (Paper 97JC02580)
Erik Korsbakken, J. A. Johannessen, and O. M. Johannessen 7857
- Katabatic wind fields in coastal areas studied by ERS-1 synthetic aperture radar imagery and numerical
modeling (Paper 97JC01774) *Werner Alpers, Ulrike Pahl, and Günther Gross* 7875
- Assimilation of ERS SAR wave spectra in an operational wave model (Paper 97JC02728)
*Lars-Anders Brevik, Magnar Reistad, Harald Schyberg, Jens Sunde,
Harald E. Krogstad, and Harald Johnsen* 7887
- The effect of assimilating ERS-1 fast delivery wave data into the North Atlantic WAM model
(Paper 97JC02570) *E. M. Dunlap, R. B. Olsen, L. Wilson, S. De Margerie, and R. Lalbeharry* 7901
- Tracking storm-generated waves in the northeast Pacific Ocean with ERS-1 synthetic aperture radar
imagery and buoys (Paper 97JC02567)
Benjamin Holt, Antony K. Liu, David W. Wang, Anand Gnanadesikan, and H. S. Chen 7917
- Statistical analysis and intercomparison of WAM model data with global ERS-1 SAR wave mode
spectral retrievals over 3 years (Paper 97JC03203) *P. Heimbach, S. Hasselmann, and K. Hasselmann* 7931
- Coastal lee waves on ERS-1 SAR images (Paper 97JC02176)
Quanan Zheng, Xiao-Hai Yan, Vic Klemas, Chung-Ru Ho, Nan-Jung Kuo, and Zongming Wang 7979
- Evolution of nonlinear internal waves in the East and South China Seas (Paper 97JC01918)
Antony K. Liu, Y. Steve Chang, Ming-K. Hsu, and Nai K. Liang 7995
- Role of surface films in ERS SAR signatures of internal waves on the shelf, 1, Short-period internal
waves (Paper 97JC02725) *J. C. B. da Silva, S. A. Ermakov, I. S. Robinson, D. R. G. Jeans, and S. V. Kijashko* 8009
- Role of surface films in ERS SAR signatures of internal waves on the shelf, 2, Internal tidal waves
(Paper 97JC02724) *S. A. Ermakov, J. C. B. da Silva, and I. S. Robinson* 8033

- ERS-1/2 orbit improvement using TOPEX/POSEIDON: The 2 cm challenge (Paper 97JC01917)
P.-Y. Le Traon and F. Ogor 8045
- Complex empirical orthogonal functions analysis of ERS-1 and TOPEX/POSEIDON combined
 altimetric data in the region of the Algerian current (Paper 97JC02909)
Catherine Bouzinac, Jorge Vazquez, and Jordi Font 8059
- Correlations between altimetric sea surface height and radiometric sea surface temperature in the South
 Atlantic (Paper 97JC02177) *Matthew S. Jones, Myles Allen, Trevor Guymer, and Mark Saunders* 8073
- Fully reprocessed ERS-1 altimeter data from 1992 to 1995: Feasibility of the detection of long term sea
 level change (Paper 97JC02566) *M. Anzenhofer and T. Gruber* 8089
- Precise orbit determination and gravity field improvement for the ERS satellites (Paper 97JC03179)
Remko Scharroo and Pieter Visser 8113
- Global marine gravity field from the ERS-1 and Geosat geodetic mission altimetry (Paper 97JC02198)
Ole Baltazar Andersen and Per Knudsen 8129
- Improved techniques for the derivation of sea surface temperatures from ATSR data
 (Paper 97JC02569) *Ian J. Barton* 8139
- Actual and potential information in dual-view radiometric observations of sea surface temperature from
 ATSR (Paper 97JC02180)
M. J. Murray, M. R. Allen, C. T. Mutlow, A. M. Závody, M. S. Jones, and T. N. Forrester 8153
- Regime shifts in mesoscale deformations of sea ice during the winter (Paper 97JC03332)
Reinert Korsnes 8167
- Evolution of the passive and active microwave signatures of a large sea ice feature during its 2½-year
 drift through the Arctic Ocean (Paper 97JC03333) *Francis Gohin, Alain Cavanié, and Robert Ezraty* 8177
- Sea ice motion from satellite passive microwave imagery assessed with ERS SAR and buoy motions
 (Paper 97JC03334) *R. Kwok, A. Schweiger, D. A. Rothrock, S. Pang, and C. Kottmeier* 8191

Preface

This special section of the *Journal of Geophysical Research* is devoted to recent advances in oceanography and sea ice research using observations from the European Remote Sensing Satellites (ERS-1/2). Individual papers address a wide range of topics on algorithm development, geophysical parameter retrieval, and validation studies of the following: (1) near-surface winds, (2) surface waves, (3) oceanic features, (4) ocean circulation, and (5) polar sea ice. The papers provide clear and convincing evidence that ERS satellite data, used in conjunction with in situ measurements; other satellite data, such as from the TOPEX/POSEIDON radar altimeters; and numerical models have enabled significant advances to be achieved in physical oceanography and polar sea ice research.

The European Space Agency (ESA) launched ERS-1 in July 1991. ERS-2 was launched in April 1995 to ensure the long-term continuity of ERS-1 data collection. For 12 months in 1995–1996, during the Tandem mission, ERS-1 and ERS-2 were operated in the same 35-day repeat orbit with a revisit time of 1 day. ERS-1 had also been operated in 3- and 168-day repeat cycles to cater to a wide range of scientific and operational requirements. Since June 1996, ERS-2 has been the primary operating satellite with ERS-1 being kept in hibernation. Tandem operation of limited duration is still performed on a campaign basis. Both satellites carry an infrared visible along-track scanning radiometer (ATSR) and three active microwave instruments: a synthetic aperture radar (SAR), a wind scatterometer, and a radar altimeter (RA) supported by a microwave radiometer (MWR) and precise range and range rate equipment (PRARE). The RA and ATSR instruments collect data continuously while the SAR and the wind scatterometer are configured in an alternating mode of operation such that the SAR wave mode runs jointly with the wind scatterometer, while the SAR image mode only runs with the scatterometer turned off.

In this special section the 28 papers represent a compendium of ongoing research activities carried out with ERS. It complements the series of ERS symposia proceedings, application workshop reports, and scientific and application achievement documents published during the last 5 years. In addition, it adds to the numerous papers, relying on the use of ERS observations, which over the years have been published in review journals or special issues. More than half of the papers in this special section focus on topics in which the use and interpretation of SAR data are essential. These include wave modeling and data assimilation, new methods to obtain high-resolution mesoscale wind field patterns, improved techniques to interpret and analyze signatures of internal waves, and sea ice pattern and motion recognition. The remaining papers are predominantly focused on the use of (1) the RA for retrieval of the marine gravity field, studies of sea level change, and variability in upper ocean circulation; (2) the wind scatterometer for the determination of the vector wind field and the estimation of retrieval accuracies as well as studies of wind stress and upper ocean transports; and (3) the ATSR for accurate derivation of sea surface temperatures.

ERS-2 is expected to operate for at least 2 more years. Between the two satellites, nearly 10 years of regular active microwave observations for oceanography and polar sea ice research are available.

The help and support from the ESA Mission Management Office (G. Duchossois, European Space Agency Headquarters, Paris) are highly acknowledged.

Guest Editors

J. A. Johannessen

ESA European Space Research and Technology Centre
Noordwijk, Netherlands

P. Janssen

European Centre for Medium-Range Weather Forecasts
Reading, England, United Kingdom

J. F. Minster

Institut National des Sciences de l'Univers, Paris

I. Robinson

Southampton Oceanography Centre
University of Southampton
Southampton, England, United Kingdom

D. Rothrock

Applied Physics Laboratory
University of Washington, Seattle

P. W. Vachon

Canada Centre for Remote Sensing
Ottawa, Ontario, Canada

Toward the true near-surface wind speed: Error modeling and calibration using triple collocation

Ad Stoffelen

Royal Netherlands Meteorological Institute, de Bilt, Netherlands

Abstract. Wind is a very important geophysical variable to accurately measure. However, a statistical phenomenon important for the validation or calibration of winds is the small dynamic range relative to the typical measurement uncertainty, i.e., the generally small signal-to-noise ratio. In such cases, pseudobiases may occur when standard validation or calibration methods are applied, such as regression or bin-average analyses. Moreover, nonlinear transformation of random error, for instance, between wind components and speed and direction, may give rise to substantial pseudobiases. In fact, validation or calibration can only be done properly when the full error characteristics of the data are known. In practice, the problem is that prior knowledge on the error characteristics is seldom available. In this paper we show that simultaneous error modeling and calibration can be achieved by using triple collocations. This is a fundamental finding that is generally relevant to all geophysical validation. To illustrate the statistical analysis using triple collocations, in situ, ERS scatterometer, and forecast model winds are used. Wind component error analysis is shown to be more convenient than wind speed and direction error analysis. The anemometer winds from the National Oceanic and Atmospheric Administration (NOAA) buoys are shown to have the largest error variance, followed by the scatterometer and the National Centers for Environmental Prediction (NCEP) forecast model winds proved the most accurate. When using the in situ winds as a reference, the scatterometer wind components are biased low by $\sim 4\%$. The NCEP forecast model winds are found to be biased high by $\sim 6\%$. After applying a higher-order calibration procedure an improved ERS scatterometer wind retrieval is proposed. The systematic and random error analysis is relevant for the use of near-surface winds to compute fluxes of momentum, humidity, or heat or to drive ocean wave or circulation models.

1. Introduction

Surface truth is very important for the computation of fluxes of momentum, humidity, or heat, as these are relevant for climate studies on the seasonal scale (El Niño Southern Oscillation) and the interannual scale. Ocean circulation models are driven by the near-surface wind. Surface-based anemometer winds cover the spatial and temporal domains poorly. On the other hand, meteorological analyses and scatterometers provide a wealth of information, but an absolute calibration of these is lacking. In this paper we attempt to solve this problem.

Wind errors are generally large, such that the signal-to-noise ratio (SNR) is not large with respect to 1. In such cases, standard regression or bin-average (BA) analyses could easily lead to pseudobias effects. It is shown in this paper that only by using triple collocations and a profound error analysis such effects may be avoided. We use a 1-month data set of triple collocations of anemometer winds from the

National Oceanic Atmospheric Administration (NOAA) buoys, ERS scatterometer winds, and National Centers for Environmental Prediction (NCEP) forecast model winds in a three-way comparison. Thus we readdress the wind calibration of CMOD4.

The current operational ERS scatterometer processing uses the transfer function CMOD4 to derive winds from the backscatter measurements. *Stoffelen* (1998) discusses the effect of backscatter calibration on the wind processing. Both backscatter and wind calibration may be performed over the ocean but in an almost independent way [*Stoffelen*, 1998]. This is not to say that the backscatter calibration has no effect in the wind domain. For instance, following an onboard hardware problem in mid-1996, the European Space Agency (ESA) switched to a redundant hardware module that slightly affected the radar backscatter calibration. A bias of -0.2 m s^{-1} in the ESA "fast delivery" was the known consequence. In mid-1997 this bias was corrected. The results in this paper apply to ERS scatterometer winds processed from calibrated backscatter measurements using CMOD4 [*Stoffelen and Anderson*, 1997 a, b, c].

CMOD4 was derived with a maximum likelihood estimation (MLE) procedure using ERS measurements and

Copyright 1998 by the American Geophysical Union.

Paper number 97JC03180.
0148-0227/98/97JC-03180\$09.00

European Centre for Medium-range Weather Forecasts (ECMWF) analysis winds (operational winds in November 1991) as input [Stoffelen and Anderson, 1997a, b, c]. Amongst some other proposals this transfer function was verified against winds from the ESA-led Haltenbanken field campaign and winds from the global forecast model of the United Kingdom Meteorological Office (UKMO) [Offiler, 1994] and selected as the preferable function. Winds from numerical weather prediction (NWP) models are only a good reference when they in turn are calibrated against in situ winds from conventional platforms. Also, in a calibration exercise it is important that a representative sample of the day-to-day weather events is present. With hindsight, the Haltenbanken campaign was perhaps too limited in extent to guarantee this.

We study the error characteristics of in situ, scatterometer, and forecast model winds in section 2. We discuss the selection of a measurement domain where the errors can be described simply by a systematic and random part. Pseudobias after nonlinear transformation of such errors are also discussed. The wind components rather than speed and direction are shown to be the most convenient to provide an accurate description of observation errors. Mean wind components of different measurement systems are generally close to each other. Therefore we first focus on the first-order calibration, which is a multiplication factor or scaling (a linear regression with zero bias term) that would for a particular true value t provide calibrated measurements x with expectation value $\langle x \rangle = t$. In section 3 it is stated that without prior knowledge it is in the case of intercomparison of two noisy systems not possible to resolve both the observation system error characteristics and the calibration. In section 4 it is shown that with three noisy systems it is possible to calibrate two of the systems with respect to the third and, at the same time, provide an error characterization for all three systems. We have used the in situ winds as a reference and scaled the scatterometer and forecast model winds to have the same average strength. In section 5 a higher-order, or more detailed, calibration is considered. Section 6 provides the error model parameters and calibration factors for the three collocation data sets. Section 7 discusses the implications of this study for scatterometer data processing and wind data interpretation and application.

2. Observation Errors and Error Domain

In order to calibrate an observing system we need to have a good notion of what parameter we want to measure, i.e., what variable, but also on what temporal and spatial scales. The variable that we deal with here is the vector wind at a height of 10 m above the ocean surface. We ignore temporal effects and assume that all observing systems involved represent the same temporal scale (10-min averages). We do not consider temporal averaging of the buoy winds, since in our analysis this would not affect the calibrations and would just reduce the random observation error of the buoys. In the spatial domain the in situ data represent a local estimate and therefore include the wind variability on all scales. On the

other hand, the scatterometer with a footprint of 50 km does not measure the variability on scales smaller than 50 km. The variability measured by an anemometer and not by the scatterometer is more generally indicated as representativeness error [see, e.g., Lorenc, 1986].

For a detailed calibration a good notion of the accuracy of the observation systems is necessary, that is, we need to know what may be called "the cloud of doubt" around the measurement. The nature and amount of system error has to be taken into account. Usually, errors are classified as systematic (bias) or random (by their standard error (SE)). For most observing systems one could distinguish detection errors and interpretation errors. The detection error includes measurement accuracy and digitizing effects, whereas the interpretation error is made when transforming the measurement(s) to the required variable(s).

For example, for buoys the detection error is determined by anemometer characteristics and buoy motion. The interpretation error for anemometer winds has only to do with the correction of the measurements to a height of 10 m and with the collocation time and space window (see, e.g., Wilkerson and Earle [1990] for a more detailed discussion).

For the scatterometer the detection error is fairly small and expressed in vector wind root-mean-square (rms roughly 0.5 m s^{-1} [Stoffelen and Anderson, 1997b, c]). On the other hand, the interpretation error is larger and depends mainly on the accuracy of CMOD4 since the inversion error is small. CMOD4 does contain effects for instance of stability, surface slicks, and waves but only as far as they are correlated with the area-averaged 10 m vector wind. However, backscatter effects that are not correlated with the 10 m wind will contribute to the random error of CMOD4.

Although a weather forecast model wind is not a measurement, it may be treated as if it was an observation, since it contains information from all tropospheric observations of mass and wind that were assimilated in the past. Its error will be largely independent from the errors of the current observations. The lowest model level is generally just below the top of the surface layer, which is roughly at 50 m. In a post-processing step, 10-m winds are derived from model variables. Errors here are caused by errors in the model state (dynamics) and by errors in the extrapolation module for the atmospheric boundary layer.

When trying to characterize measurement errors, it is practical to select a parameter domain where the cloud of doubt is simple to describe. When it is symmetric, then first- and second-order statistical moments may be sufficient to describe the errors. Although we need not limit ourselves to these, for wind the two physical choices are either wind components (u, v) or wind speed and direction (f, ϕ). These sets are nonlinearly related, and random errors in the one domain may generate a serious pseudobias in the other domain, as will be shown later.

One way to approach error characterization is to look in detail at the above error sources. The anemometer characteristics for in situ winds will vary but will generally not be the dominant error source. Interpretation errors, including

height correction to 10 m and platform motion correction errors, may be more substantial for the conventional winds. Some components of it may be well characterized in the (f, ϕ) domain, while other components are better characterized in the (u, v) domain. A major contribution to the observation error for conventional winds when comparing to scatterometer data or forecast model winds will be the spatial representativeness error. This part of the total observation error is well characterized in the wind component domain. Scatterometer winds are empirically derived, and opinions will differ as to which geophysical elements (e.g., waves, stability, rain, or sea surface temperature) determine the interpretation error. The error sources in the forecast model that project onto the surface wind are even more difficult to elaborate on. It may be clear that a characterization of the total observation error from a quantification of all the error sources contributing to it will be undoable. Therefore an empirical approach is needed.

In Figure 1 the distribution of scatterometer winds for a fixed forecast model wind subdomain is shown in both physical spaces. Since the forecast model is not perfect, the subdomain of true winds will be larger than the subdomain of the model winds, and as such, it may be clear that the distribution shown is affected by errors in both the forecast model and the scatterometer. We can see that the combined component errors are well captured by a symmetric distribution, and one may assume that both scatterometer and forecast model error distributions are symmetric (Gaussian) as well. On the other hand, the wind direction random error

clearly depends on wind speed, i.e., the lighter the wind the larger the wind direction error. The wind speed error is not symmetrically distributed for light winds but skew; that is, large positive errors are more likely than large negative errors [Hinton and Wylie, 1985]. This is related to the fact that measured negative wind speeds cannot occur.

Moreover, the cloud of doubt in the (f, ϕ) space is quite complicated and cannot be described by second-order statistics, whereas in (u, v) space the cloud of doubt seems much simpler to describe. Therefore, as is common practice in meteorological data assimilation, we define an error model in the wind components.

In practice, it is found that the random error on both the u and v components is similar, as one may expect (see, e.g., Figure 1a). By verifying the error distributions at higher speeds we found little evidence of a speed dependence of the component errors in the observation systems studied (see, e.g., Figure 2). As such an error model with constant and normal component errors appears appropriate. It implies for speed and direction that the expected rms wind speed difference $\langle (f_x - f_y)^2 \rangle$ of two noisy systems X and Y increases monotonically with wind speed, and the wind direction rms $\langle (\phi_x - \phi_y)^2 \rangle$ increases monotonically to a value of 104° for decreasing wind speed. Hinton and Wylie [1985] used a truncated Gaussian error distribution that did not allow negative speeds, to correct for the low-speed pseudobias. This procedure is rather unsatisfactory since it is not likely that the true error distribution contains discontinuities. By assuming Gaussian error distributions in the wind compo-

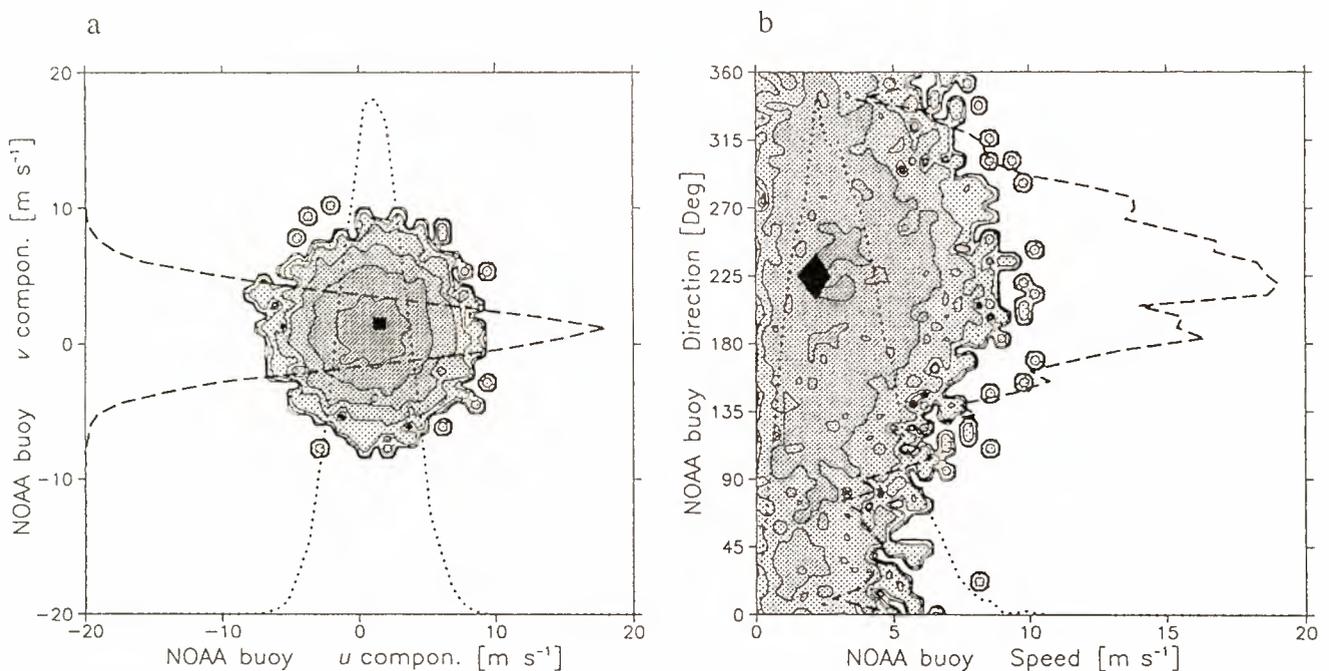


Figure 1. The distribution of scatterometer winds for forecast model winds with component values in between 1.1 and 1.9 m s^{-1} for 6,738 northern hemisphere high-latitude cases in March 1995. The distribution as a function of (a) the components and (b) as a function of speed and direction is shown. The forecast wind subdomain is indicated by a black box. The relative distribution of points (probability density function) along the horizontal and vertical parameter axes are represented by the dotted and dashed lines, respectively. Component errors are simpler to describe than speed and direction errors.

nents the cutoff speed effect is naturally simulated, thereby avoiding an ad hoc correction.

A good way to verify our approach is to simulate the wind speed and direction difference statistics with the error model we have obtained for the wind components. Figure 2 shows such a comparison (compare errors to Table 2). We can see that the average wind speed difference (pseudobias) indeed varies as a function of wind speed and that it can be as large as 1 m s^{-1} for realistic errors. The standard deviation of the wind speed difference and the vector rms difference go to a small value for low wind speed, as is observed for the real data as well. The wind direction standard deviation increases for decreasing wind speed and goes to a value of a hundred odd, as expected (random direction). Thus our error model set up to describe the observed difference statistics in the wind components also qualitatively describes the observed difference statistics in wind speed and direction very well, thereby confirming its adequacy. A quantitative validation of the error model can be made when the wind component errors are known.

3. Error Modeling and Calibration With Two Systems

Unfortunately, all observation systems contain error. This means that we cannot assume that one measurement represents the true state and calibrate the other against it, as is illustrated here. Assume we have a distribution of "true" states, indicated by the variable t , with expected variance $\langle t^2 \rangle = \sigma^2$, and two independent measurement systems X and Y , indicated by the variables x and y with error variances of $\langle (t-x)^2 \rangle = \varepsilon_x^2$ and $\langle (t-y)^2 \rangle = \varepsilon_y^2$ respectively. If the distribution of true values and the error distributions are normal, one can show that for fixed x the average of y does not lie at $\langle y \rangle = x$ but at $\langle y \rangle = \sigma^2 (\sigma^2 + \varepsilon_x^2)^{-1} x$ (see appendix A); that is for $\sigma = 5 \text{ m s}^{-1}$ and a typical wind error of $\varepsilon_x = 2 \text{ m s}^{-1}$ we find $\langle y \rangle = 0.84 x$, which implies a 1.6 m s^{-1} difference at 10 m s^{-1} . So, for unbiased Gaussian error distributions, computing the mean of y for a fixed subrange or bin of x (bin-average analysis) does, in general, reveal a pseudobias that depends on the error characteristics of system X .

Scatterometer data are often verified against buoy data, where the buoy data are assumed to be "surface truth" [see, e.g., Rufenach, 1995]. However, the representativeness error (see sections 1 and 4) for anemometer winds is substantial, and therefore this assumption does not hold. As such, in this work the observation error of in situ winds will be accounted for in the interpretation in order to be able to draw valid conclusions on the scatterometer and forecast model bias.

A better assumption often used either implicitly (e.g., in "geometric mean" linear regression) or explicitly is $\varepsilon_Y = \varepsilon_X = \varepsilon$, leading to the expectation $\langle x^2 \rangle = \langle y^2 \rangle = \sigma^2 + \varepsilon^2$. Again, for $\sigma = 5 \text{ m s}^{-1}$ but now for the common wind errors of $\varepsilon_X = 3 \text{ m s}^{-1}$ and $\varepsilon_Y = 1 \text{ m s}^{-1}$, we find a ratio of total variances of $\langle x^2 \rangle / \langle y^2 \rangle = 1.32$, which would lead after linear regression to the conclusion that system Y is biased low by 16% if system X is assumed to be bias free.

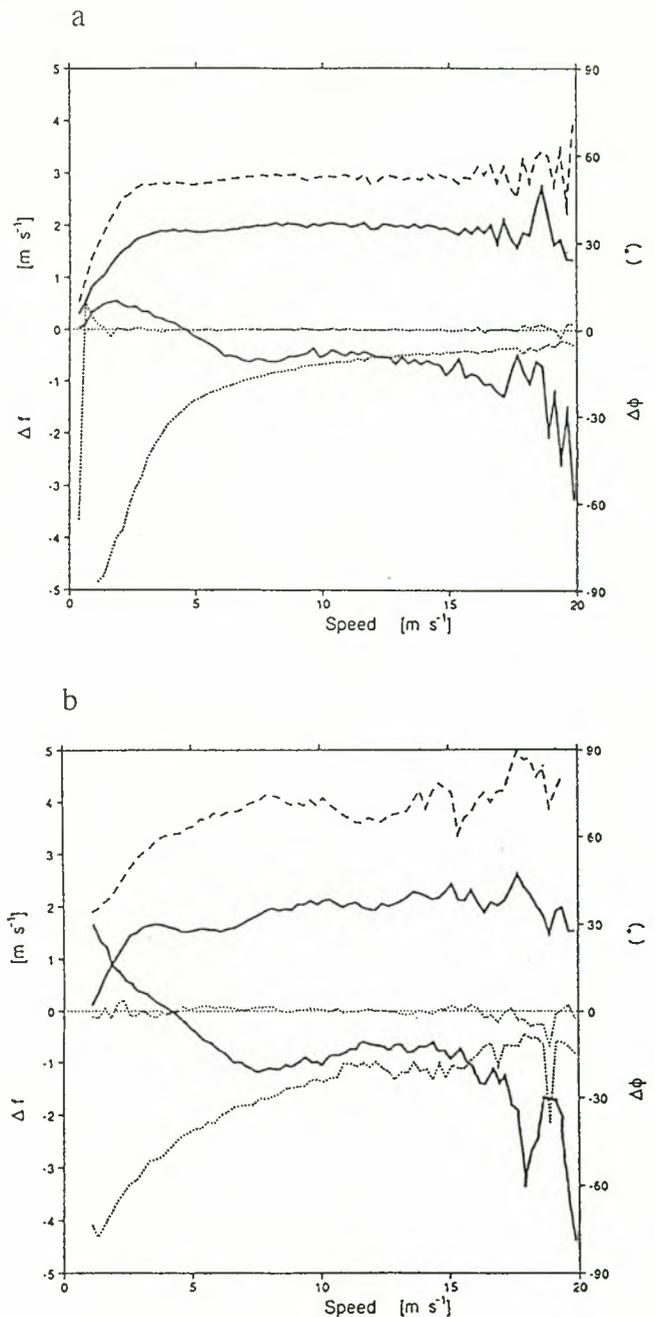


Figure 2. (a) Simulated and (b) true wind speed and direction difference statistics of ECMWF forecast model minus scatterometer as a function of average wind speed for all global collocations from February 13-16 1994. Speed bias (thin solid line), standard deviation (thick solid line), direction bias (thin dotted line), standard deviation (thick dotted line), and vector root-mean-square (dashed line) of differences are shown. The simulation (Figure 2a) is done with the scatterometer wind distribution as "truth" and wind component standard errors of 1.0 and 1.8 m s^{-1} for the forecast model and scatterometer, respectively (compare Table 2). Figure 2b is for the first node at the inner swath, which is the noisiest of all nodes. Although Figure 2b is noisier, the general speed and direction error characteristics are qualitatively well simulated in Figure 2a by the wind component error model.

Thus the assumption may imply a 1.6 m s^{-1} pseudobias at 10 m s^{-1} for, in reality, unbiased Gaussian error distributions. The examples illustrate that when a noisy system is used as a reference for calibration in a dual collocation, we will need to know the error characteristics of that system. Further proof of this is given in Appendix A.

Another possibility of generating pseudobias is by nonlinear transformation. An unbiased symmetric error distribution may then be transformed into a skew and biased error distribution. For example, Gaussian errors on the wind components u and v for system X or Y will not correspond to Gaussian error distributions on wind speed f and direction ϕ . Using the same notation and assumptions common to the previous two examples, we show in Appendix B that $\langle f_X \rangle = \sqrt{\pi/2} \sigma_X$, where $\sigma_X^2 = \sigma^2 + \varepsilon_X^2$, with σ the standard deviation (SD) of the true wind component distributions, ε_X the SE of system X , and where the errors are assumed identical for the u and v components. A similar expression can be derived for system Y . The expected mean wind speed difference is approximated as $\langle f_X - f_Y \rangle = \sqrt{\pi/2} (\sigma_X - \sigma_Y) \approx \sqrt{\pi/2} (\varepsilon_X^2 - \varepsilon_Y^2) \sigma^{-1}$. For example, the typical values of $(\sigma, \varepsilon_X, \varepsilon_Y) = (5, 3, 1) \text{ m s}^{-1}$ will lead to an average wind speed bias of 0.4 m s^{-1} . This is a pseudobias since the error in the wind components is unbiased. One can show that the pseudobias generally oscillates from positive to negative as a function of wind speed and is largest in a relative sense for low speeds (as in Figure 2).

It may be clear from the above that it is desirable that the error characteristics of measurement systems are well described, thereby avoiding pseudobias effects only caused by inaccurate assumptions on the errors. Moreover, in order to provide a calibration and error model of observing systems it is desirable that a domain is chosen where the errors are simple to describe, preferably avoiding statistical moments of order higher than 2 (see section 2). Also, in geophysical applications the use of data with complex error characteristics may lead to biased results when the error characteristics are not properly accounted for.

4. Error Modeling and Calibration With Three Systems

In the previous section it was indicated that calibration of one noisy system against another is not possible without fundamental assumptions on the noise characteristics of at least one system. It was shown that these assumptions may lead to substantial pseudobias problems. This is further elaborated in appendix A. Here a method is introduced to perform calibration and error modeling using triple collocations. The method is quite general and is introduced as such. Later on, the method is applied on in situ, scatterometer, and forecast model wind components. Now suppose three measurement systems X , Y and Z measuring a true variable t . Let us define

$$\begin{aligned} x &= t + \delta_x, & \varepsilon_x^2 &= \langle \delta_x^2 \rangle \\ y &= s_Y(t + \delta_Y), & \varepsilon_Y^2 &= \langle \delta_Y^2 \rangle \\ z &= s_Z(t + \delta_Z), & \varepsilon_Z^2 &= \langle \delta_Z^2 \rangle \end{aligned} \quad (1)$$

with as before, $\sigma^2 = \langle t^2 \rangle$, and now δ_x , δ_Y , and δ_Z are the random observation errors in the measurements x , y , and z respectively. Here s_Y and s_Z are the calibration (scaling) constants. We have assumed no bias such that $\langle \delta_x \rangle = \langle \delta_Y \rangle = \langle \delta_Z \rangle = 0$. For marine winds this is valid to good approximation (see section 5), but otherwise, bias may be easily removed.

It is unlikely that the three systems represent the same spatial scales. Therefore we will arbitrarily assume that observation systems X and Y can resolve smaller scales than system Z by taking $\langle \delta_x \delta_Y \rangle = r^2$. Here r^2 is the variance common to these smaller scales and taken as part of the observation errors δ_x and δ_Y and t only represents the spatial scales resolved by z . By definition r^2 is the correlated part of the representativeness errors of X and Y . The choice for t to only resolve the coarsest scale measurement allows the approximation $\langle \delta_x t \rangle = \langle \delta_Y t \rangle = \langle \delta_Z t \rangle = 0$, that is, the observation errors are assumed uncorrelated with t . Furthermore, since z does not include the smaller scales, the observation error of system Z is independent of the errors of X and Y , that is, $\langle \delta_x \delta_Z \rangle = \langle \delta_Y \delta_Z \rangle = 0$. The assumption that the wind component errors of the different observation systems are uncorrelated, except for the representativeness error, is essential to determine the calibration. Now the calibration coefficients can be derived from the different covariances

$$\begin{aligned} s_Y &= \langle yz \rangle \langle zx \rangle^{-1} \\ s_Z &= \langle yz \rangle (\langle xy \rangle - r^2 s_Y)^{-1} \end{aligned} \quad (2)$$

These coefficients can be used to create

$$\begin{aligned} y^* &= s_Y^{-1} y \\ z^* &= s_Z^{-1} z \end{aligned} \quad (3)$$

which are the calibrated data. Subsequently, all random error parameters of the observation systems X , Y , and Z can be resolved pairwise from the different covariances, as illustrated in Appendix A by (A5). Here we used observation system X as a reference system. This preference can be easily altered by scaling all parameters to one of the other systems.

So under the premise that we find an estimate for r^2 we have found a way to perform a first-order calibration. In work by Stoffelen [1996] the spatial representativeness error of the scatterometer with respect to the ECMWF model is estimated to be $r^2 = 0.75 \text{ m}^2 \text{ s}^{-2}$. We use this as a baseline assumption here as well. The sensitivity of the results to this choice is discussed later on and shown to be small (section 6).

5. Higher-Order Calibration

After the first-order calibration the three systems should be largely unbiased. However, in this section we consider a more detailed calibration of the systems by pairwise comparison. For the triple-located data the procedure is run comparing X and Y , Y and Z , and Z and X so that consistency can be checked between the results.

Now first consider X and Y . After obtaining ε_X and ε_Y we decide which system is the least noisy; for example, suppose $\varepsilon_X > \varepsilon_Y$. Then system Y may be convoluted with a Gaussian distribution with width $\sqrt{\varepsilon_X^2 - \varepsilon_Y^2}$ to obtain a distribution (and a system Y') that has the same error properties as that of system X , i.e., $\varepsilon_{Y'} = \varepsilon_X$. In fact, since they represent the same true distribution, the resulting distributions of X and Y' should be identical in case of a large sample size. When dealing with winds, the errors of X and Y may be matched for both components, such that either components or speed and direction distributions should be identical. By comparing the cumulative distributions of these two systems, $f(x)$ and $g(y')$, that are monotonically increasing functions, we can easily compute a mapping $y'' = \mu(y')$ that results in identical distributions $f(x)$ and $g(y'')$. The higher-order correction would then be $\mu(y) - y$, which may be plotted versus y .

The higher-order calibration is only reliable when the errors of system X and Y are well characterized by a normal distribution. However, if substantial errors of order higher than 2 would be present in either system X or Y , then the detailed bias computation method set out above may provide large pseudobias. When more than two systems are involved, such as in our triple collocation data set, a consistency check between the different comparisons may reveal such problems. In the extreme parts of the domain, where the error distributions of either system are insufficiently sampled, further pseudobias effects may occur. To gain confidence, the computed corrections may be verified by Monte Carlo simulation to test such sampling problems.

6. Calibration and Error Model Results

In this section we describe the implementation of the method given earlier and show the results of the calibration of scatterometer and forecast winds relative to the anemometer winds. A 1-month data set kindly provided by NCEP from March 1995 with off-line anemometer winds from the NOAA buoys corrected to 10-m height [Wilkerson and Earle, 1990] was used collocated with the ESA-processed ERS-1 scatterometer and NCEP forecast model winds. The NCEP forecast wind is valid at the time and location of the buoy measurement. All scatterometer data within 200 km and 3 hours of a buoy were selected and as a result 40,091 triple collocations were available resulting from 1465 independent buoy measurements.

The average wind components of the in situ, scatterometer, and forecast winds are very close (within a few tenths of a m s^{-1}). As such, the assumption just below (1) that the systems have no absolute bias, i.e., $\langle \delta_i \rangle = 0$ for $i = (X, Y, Z)$ proves valid.

A quality control procedure is applied to exclude gross errors. An iterative scheme is followed, where in each trial a calibrated (equation (3)) collocation triplet is rejected when for any pair the following condition holds, illustrated here for pair (x, y^*)

$$|x - y^*| > 3 \sqrt{\varepsilon_X^2 + \varepsilon_Y^2} \quad (4)$$

where the calibrations and the errors are taken from the previous trial. In the first trial we start with calibration factors equal to 1 and errors equal to 2 m s^{-1} . The calibration factors are within a percent of their final value after one iteration, and also, the rejection rates remain constant (at $\sim 1\%$) after 1 or 2 trials; 6 trials were used. A wind direction bias correction is repeated after every trial, and the resulting corrections are 4° - 5° for the scatterometer and 1° - 2° for the forecast model, where the incremental correction in the sixth trial was very small ($< 0.01^\circ$). Wind direction corrections of this size do not substantially impact the wind component statistics, but the scatterometer bias may need further attention. Stoffelen [1996] shows that the bias reverses in the southern hemisphere and is difficult to explain from existing theories on air-sea interaction.

6.1. First-Order Calibration

The calibration factors resulting from the above procedure are shown in Table 1. The NCEP forecast model appears biased high by roughly 6%. The representativeness error estimate modestly influences the calibration coefficient of the forecast model as we would expect from (2). In contrast, the scatterometer calibration enforces the winds. Remarkably, the scatterometer along-track component is biased less than the across-track component. This difference, however, may be explained by the relatively small number of independent collocations used.

Figure 3 shows the joint distributions of the wind components of in situ and scatterometer, scatterometer and NCEP, and NCEP and in situ data. It is evident that the scatter in the scatterometer and NCEP plot is smallest. This means that the in situ winds have the largest error. The in situ and scatterometer plot shows the largest scatter, which indicates that the NCEP winds are the most accurate. Table 2 shows the results of our random error estimates from (A5) in Appendix A, which confirm our subjective analysis. The error estimates for the u and v components are quite similar for NCEP and conventional winds, but for the scatterometer the along-track component seems slightly worse than the across-

Table 1. Calibration Scaling Factors Against Buoy Winds for Wind Components From the Scatterometer and Forecast Model for Different Representativeness Errors

	Component Scaling		
	u	v	$r^2, \text{m}^2 \text{s}^{-2}$
Scatterometer	0.97	0.95	all
NCEP model	1.06	1.06	0.50
NCEP model	1.06	1.07	0.75
NCEP model	1.07	1.08	1.00

Here u is the along-track, and v is the across-track wind component; r^2 is the representativeness error (due to scatterometer and model spatial resolution difference).

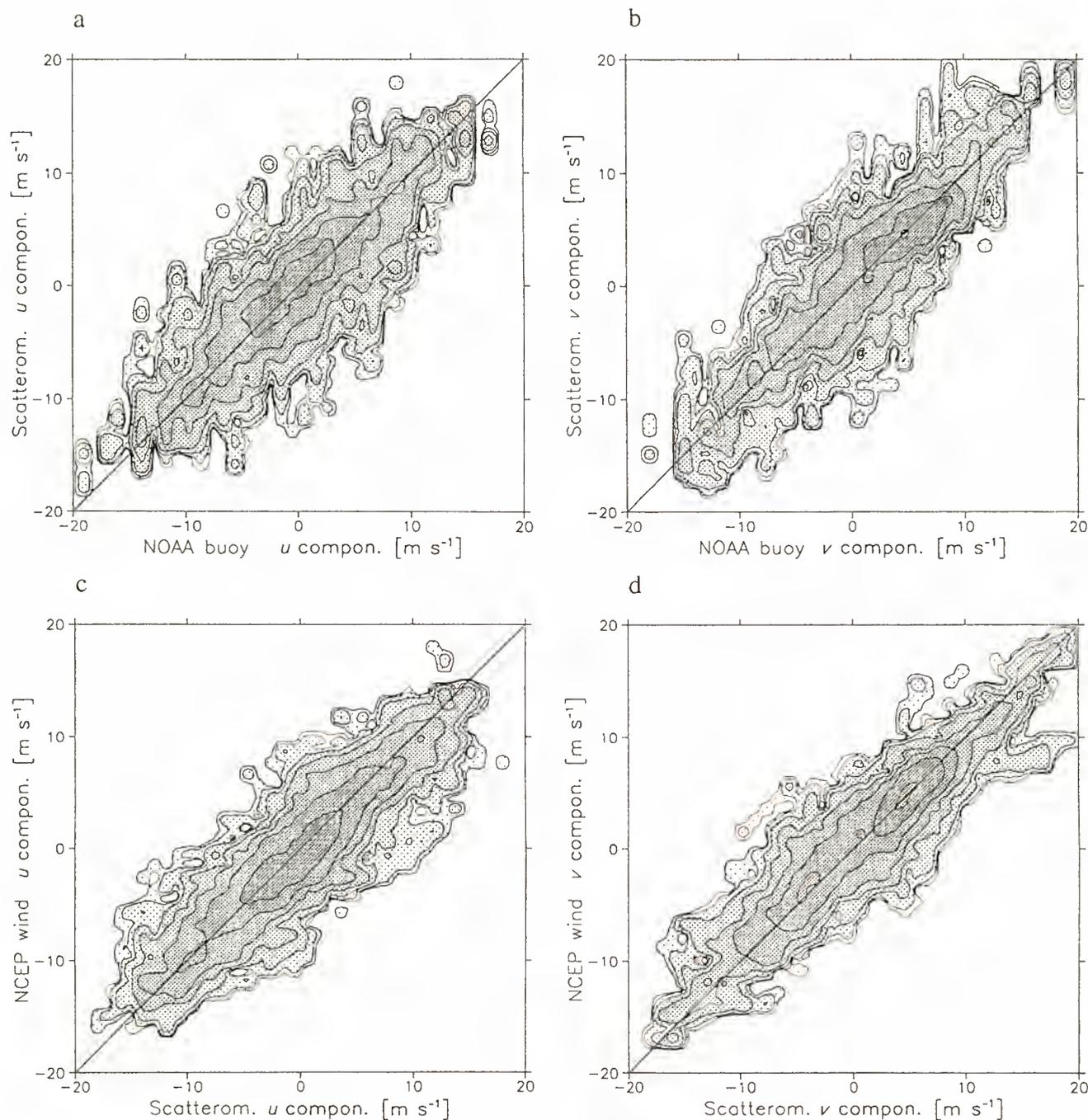


Figure 3. Joint distributions for the along-track (u) and across-track (v) ERS-1 wind components, respectively, for the (a, b) NOAA buoy anemometer and ERS-1 scatterometer, (c, d) scatterometer and NCEP forecast model, and (e, f) forecast model and anemometer winds. These plots show the full characteristics of the triple collocation database in the wind domain.

track component. Given the fact that the wind direction was predominantly across-track in this data set, it may indicate some anisotropy in the scatterometer error distribution.

From a climatological wind spectrum one may estimate the representativeness error variance of the conventional winds with respect to NCEP winds to be $2.1 \text{ m}^2 \text{ s}^{-2}$ [Stoffelen, 1996]. As such, the (local) error of the in situ winds may be estimated as 1.41 m s^{-1} for the u and 1.21 m s^{-1} for the v component. However, for many applications the local wind

is not as relevant as an area-averaged quantity such as provided by the scatterometer.

Since the scatterometer winds are not exactly collocated in space and time with the in situ and NCEP winds, a collocation error may be subtracted from the former. Wind measurements separated by in between 2 and 3 hours will have an additional error component of $\sim 1.5 \text{ m}^2 \text{ s}^{-2}$ [Stoffelen and Anderson, 1997c]. When we assume that the collocations are randomly distributed in time (over 6 hours), the average er-

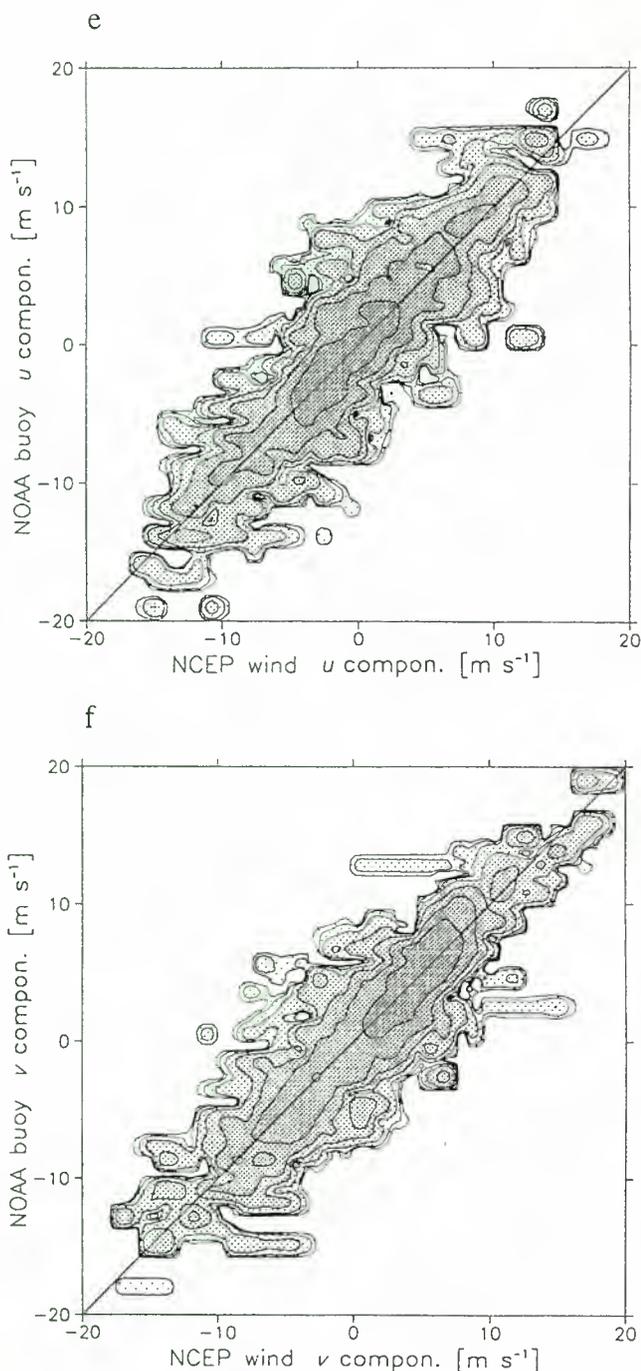


Figure 3. (continued)

ror contribution would be $\sim 0.75 \text{ m}^2 \text{ s}^{-2}$. Such a collocation error would reduce the scatterometer error estimates by $\sim 0.2 \text{ m s}^{-1}$.

Stoffelen [1996] studied the spatial representation of the ECMWF forecast model and the ERS-1 scatterometer winds on scales between 50 and 250 km. On these scales the scatterometer contained substantially (20-40%) more variance. It was very likely that part of this additional variance verifies with the true wind and that part of it contributes to the scatterometer wind error. From the study it follows that the sum of the scatterometer error on scales between 50 and 250 km and the representativeness error would be 1.4 m s^{-1} , which

verifies reasonably well with the error estimates obtained in this work. The ECMWF and NCEP model errors would be comparable, since the ECMWF wind error on scales between 50 and 250 km was estimated to be 1.1 m s^{-1} in Stoffelen [1996].

6.2. Higher-Order Calibration

The higher-order calibration procedure of section 5 requires a convolution of the most accurate system with a Gaussian error distribution with a width determined by the difference in error variance of the two systems. Figure 4 shows the higher-order calibration by the cumulative distribution mapping method. The wind component biases appear to be rather random in nature. The smallest biases are generally present at small wind component values, and the highest are present at high component speeds. This is due to the fact that the largest number of data is present at low component speeds, and the lowest number is present at high speeds. In other words, the higher the speeds the less accurate the results are. In an attempt to fit the bias with a smooth function this would have to be taken into account, and changes at the tail of the distribution should be kept small.

Inconsistency of the wind component biases between the calibration pairs is most noticeable at the tail of the distributions; that is, for example, the bias of the in situ winds with respect to the scatterometer is not equal to the sum of the biases of the conventional winds with respect to NCEP and the bias of NCEP with respect to the scatterometer. Moreover, the plot suggests that the scatterometer has a generally negative bias with respect to the anemometer winds and NCEP for the along-track component. However, this effect is not confirmed by the differences in mean value of the scatterometer along-track component. Closer inspection reveals more (but smaller) inconsistencies. These indicate insufficient sample size or higher-order statistical moments in the error distributions than those accounted for.

The cloud of doubt is assumed to be Gaussian with a fixed standard error. However, in reality, also higher-order moments may be relevant here. Furthermore, the standard error may depend on the geophysical condition, for instance on

Table 2. Estimates of the Wind Component Standard Deviation of the True Distribution and the Standard Errors of the Buoy, Scatterometer, and Forecast Error Distributions

	Component, ms^{-1}	
	u	v
True variance	4.68	5.24
In situ error	2.02	1.89
Scatterometer error	1.89	1.62
NCEP error	1.11	1.15

Here u is the along-track and v the across-track wind component. The values are computed at the spatial representativeness of the forecast winds with $r^2 = 0.75 \text{ m s}^{-1}$ (see text).

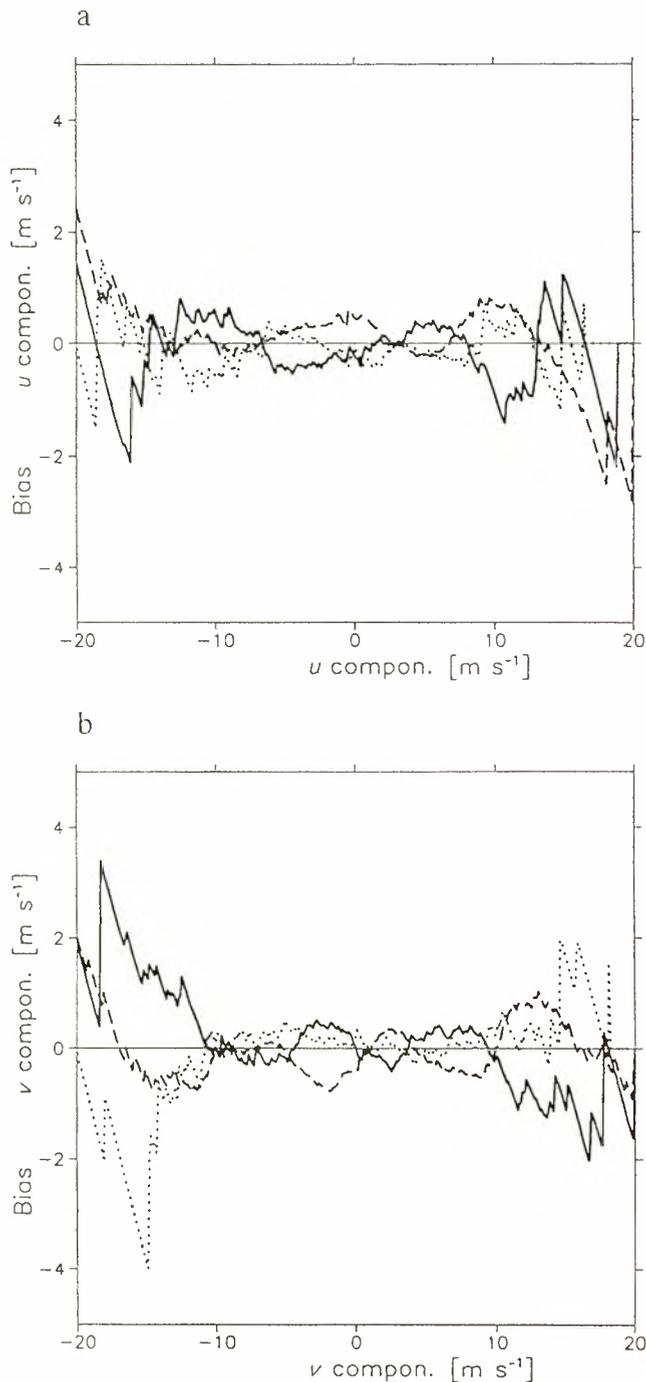


Figure 4. Higher-order calibration by cumulative distribution mapping (see text) for (a) the along-track and (b) the across-track ERS-1 wind components. The biases of scatterometer with respect to NOAA buoy anemometer (solid line), NCEP forecast model with respect to scatterometer (dotted line), and anemometer with respect to forecast model (dashed line) winds are shown. Although large biases are present, there is no consistent pattern apparent in the data.

stability, and we would need a representative sample of all these conditions to perform the higher-order calibration accurately. As such, the number of samples we need to determine the calibration and error model may be quite large.

Since there is no mechanism to remove the pseudobiases due to sampling and higher-order moments, we believe that the methodology has to be used with restraint.

Physically, one may expect systematic errors to scale with f and ϕ , rather than with u and v . If the errors in u and v of systems X and Y are identical, then their f and ϕ distributions should ideally also be identical. The wind direction cumulative mapping also does not result in substantial and consistent systematic effects, but for wind speed it does, as shown in Figure 5. The scatterometer is biased high for high and very low wind speed and biased low for moderate speeds with respect to the buoys. The NCEP model shows similar differences with respect to the buoys but roughly half in size. However, at the very low wind speeds, NCEP and scatterometer are more consistent. The comparison of NCEP and scatterometer winds generally confirms these results, except at the poorly sampled high speeds. After applying to the scatterometer and NCEP winds the first-order corrections and the second-order corrections up to 17 m s^{-1} , a repeated application of the calibration method does result in no further substantial corrections and in errors that are very similar to the ones of Table 2. As such, up to 17 m s^{-1} the higher-order correction of the scatterometer against the buoys appears sensible, and we suggest it, together with a 4% linear correction, as a modification to CMOD4.

It is worth noting at this point that by taking the buoys as a reference for the scatterometer calibration, we assume that the buoy generally provides an unbiased estimate of the scatterometer footprint area-averaged wind vector. Especially for the very low wind speeds, one may want to further investigate this.

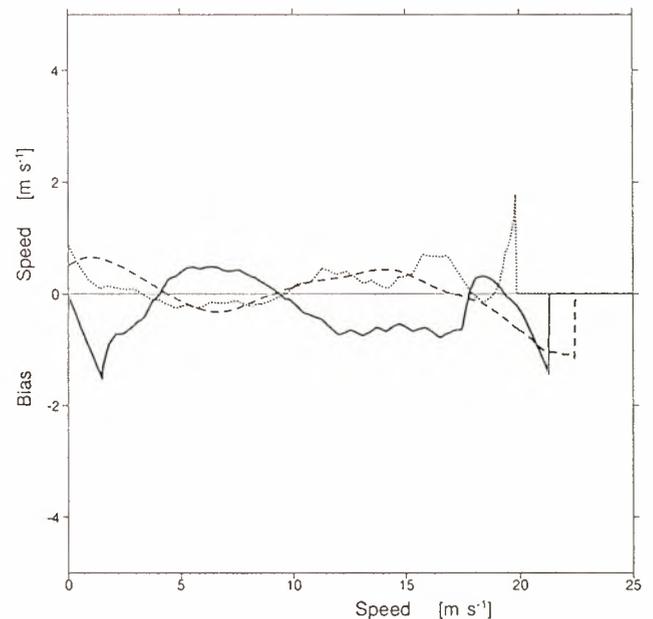


Figure 5. As Figure 4, but for wind speed. A consistent speed correction for the NCEP model and scatterometer is present at wind speeds up to 17 m s^{-1} . For these speeds the solid line shows the suggested higher-order correction to CMOD4 (in addition to a 4% linear correction).

7. Conclusions

7.1. Résumé

In this paper, calibration and error modeling are discussed, and a methodology is provided to obtain the absolute calibration and accuracy of observing systems. In particular, the focus is put on ocean surface wind speed (or stress) biases that are detrimental for the computation of fluxes of momentum, humidity, and energy through the air-sea interface. An improved ERS scatterometer wind processing is proposed.

It was demonstrated that the selection of a simple measurement domain where second-order statistics are sufficient to describe the uncertainty of the measurements is preferred. More specifically, we have shown that wind error modeling using component error distributions that are Gaussian and constant represents a simple method to describe the more complex errors in speed and direction. We have shown that substantial pseudowind speed biases can occur through the nonlinear transformation of unbiased wind component errors to the wind speed and direction domain. In a direct wind speed calibration, where usually unjustly symmetric error distributions are assumed, the pseudobiases would be taken out, leading to biased wind components [see also *Hinton and Wylie, 1985*]. Wind component error modeling as proposed here elegantly solves this problem.

In order to calibrate one observing system with respect to the other, one may use, either explicitly or implicitly, a simplifying assumption on the errors of the two systems. For instance, it is common practice to assume that the errors of two systems that are compared are equal or to assume that one system is much more accurate (i.e., is "truth") than the other. Given our results in Table 2 and Figure 3, it is obvious that both of these choices would have been crude for any of the observation systems dealt with in this paper. We have shown that such assumptions may lead to substantial pseudobias effects. Furthermore, in Appendix A it is shown that it is impossible to calibrate one noisy system against another without such assumptions or other prior knowledge on the error characteristics of one or both systems.

For a proper calibration of an observing system a reference system is necessary and at least one other observation system. Using triple collocations, a method to calibrate noisy systems has been developed. Subsequently, in a pairwise comparison of the calibrated observation systems the covariances were used to estimate the true variance resolved by both systems and the error variance of the observations. To complement the linear calibration, a more refined bias estimation procedure was adopted.

We used the NOAA buoy anemometer winds as a reference, although they turned out to be the least accurate amongst the scatterometer and NCEP forecast model winds. The spatial representativeness error is the main reason for the low accuracy of the buoys in our triple comparison. One would expect this error to be of a random nature and not lead to biases in the results. We found that the CMOD4-derived scatterometer wind components are biased low by 4%. The

NCEP forecast model appears to be very accurate but biased high by 6% for the period we examined (March 1995). In another study, using triple collocations of the ECMWF forecast model, ERS scatterometer winds, and real-time available anemometer winds, similar results were obtained [*Stoffelen, 1996*]. The higher-order scatterometer calibration with respect to the buoys resulted in a correction additional to the 4% mentioned above. The total correction is recommended for operational implementation in the ESA fast delivery processing chain.

7.2. Application

Our statistical analysis on surface winds has direct implications in the area of data assimilation in numerical weather prediction (NWP) models and in ocean circulation and wave model forcing. However, the methodology may be applied for the interpretation of any geophysical variable with a high variability on the smaller scales or high signal-to-noise ratio. It provides a way to compare data with different amounts of noise or different spatial and/or temporal resolution.

It is essential that NWP models assimilate unbiased data. It has been observed by ECMWF that the scatterometer bias with respect to their forecast model (~10%) had the tendency to slow down the forecast winds in the analysis and subsequently fill in low-pressure systems. This effect can be circumvented by a model wind correction to match the mean observed wind [*Roquet and Gaffard, 1995*]. The 6% bias we found between the buoys and the NCEP model may result from physical parameterizations that also control atmospheric boundary layer humidity, depth, and temperature and that require careful tuning. Such biases, when detected, are therefore not easily corrected, and a short-term solution such as adopted by Roquet and Gaffard will be beneficial. However, in the long term the forecast model bias correction should be replaced by forecast model improvements.

Roquet and Gaffard [1995] computed a bias correction in the wind speed domain, rather than in the wind components. The wind components, however, are used for data assimilation and should be unbiased. Figure 2 clearly shows that because of the nonlinear transformation, unbiased component errors will lead to biased speeds and vice versa. It is therefore essential to compute biases in the domain of the analysis variables, i.e., the wind components [see also *Le Meur et al., 1997*]. For other analysis variables the same strategy may be applied; that is, observations and forecast data are compared in that domain where the errors are best described, and bias corrections are computed after the transformation of the random errors to the analysis variable domain. For noisy data a careful statistical error analysis as described in this work will be essential to arrive at an optimal bias correction and assimilation.

Wave models directly rely on NWP model winds. Here it was shown that NWP forecast model winds are very accurate in describing the synoptic scale flow but may be biased. However, since we quantified the bias, it is easily corrected for when the winds are used in wave models. We note that it is more problematic to take account of the error in the forc-

ing since this error is nonlinearly related to the random wind error. Also, it is relevant to be aware of the fact that a NWP model provides an area-averaged vector wind and that the additional forcing due to the wind on the unresolved scales needs be parameterized.

Tropical wind analysis is very important for ocean circulation models and as such for the study of the Earth's climate. In ocean circulation experiments, similar arguments apply as in wave forecasting concerning the forcing problem. Stoffelen [1996] and Bryan *et al.* [1995] noted a large directional inconsistency between scatterometer and ECMWF wind direction at the equator. The tropical array of (TAO) buoys may provide a relevant data set to enhance our knowledge on the utility of scatterometer and forecast winds in this area using the methodology described here.

Appendix A: Necessity of Error Modeling Before Calibration or Validation

In this appendix the problem of calibration and validation of one noisy system with respect to another one will be discussed. Usually, scatterplots are used to compare the data followed by a regression analysis to compute a calibration coefficient or to validate the system(s). First, the interpretation of scatterplots and associated regression and bin average (BA) analyses are discussed, and it is illustrated that calibration or validation, without knowing the error characteristics of one or both systems, can easily lead to pseudobias. In the second part it is shown that calibration or validation of one noisy system against another, without knowing the error characteristics of the observing systems, is generally not possible.

A1. Scatterplots and Regression

Usually, a scatterplot is used to determine the error characteristics of a measurement system (see, for example, Figure 4). In this section we quantify the properties of the scatterplot. If enough collocation data are available, then the density of points in the scatterplot is proportional to the joint probability density of x and y given by

$$p(x, y) = \int p(x|t) p(y|t) p(t) dt \quad (A1)$$

The integration is over the distribution of true states $p(t)$ and over the distributions of error. Here $p(x|t)$ is the conditional probability density of x given t , which includes all measurement and error characteristics of the measurements x . (It is closely related to what was introduced as the cloud of doubt around observation x , which formally reads $p(t|x)$ and $p(t|x) = p(x|t) p(t)$.) We can see that the joint distribution of x and y is not only determined by the error characteristics of both systems but also by the distribution of true states. In the simple case of unbiased Gaussian errors with standard error (SE) equal to ε_x or ε_y and a Gaussian true distribution with zero mean and RMS σ , the joint probability of x and y can be written as

$$p(x, y) \propto \exp\left[-\frac{(\sigma^2 + \varepsilon_y^2)x^2 + (\sigma^2 + \varepsilon_x^2)y^2 - \sigma^2 xy}{2(\sigma^2 \varepsilon_x^2 + \sigma^2 \varepsilon_y^2 + \varepsilon_x^2 \varepsilon_y^2)}\right] \quad (A2)$$

For given x the mean y value of this distribution does not lie at $y = x$ but at $y = \sigma^2 (\sigma^2 + \varepsilon_x^2)^{-1} x$. So, even for unbiased Gaussian error distributions, computing the mean of y for a fixed subrange or bin of x (bin average) does, in general, reveal a pseudobias that depends on the error characteristics of x . This needs to be accounted for in the interpretation, and therefore, when correcting for bin-average biases with respect to a reference system, the error characteristics of that system, in this case ε_x , need to be known well.

For calibration or validation, often linear regression is used as a tool. For a well-calibrated system and in the case of $\varepsilon_x = \varepsilon_y$ a geometric mean linear regression on the joint distribution would indeed result in the line $y = x$, but in the more general case of $\varepsilon_x \neq \varepsilon_y$ it would result in $y = (\sigma^2 + \varepsilon_y^2) (\sigma^2 + \varepsilon_x^2)^{-1} x$, where it would again result in a pseudobias. When the error characteristics of x and y are known, a weighted fit may result in a proper calibration. However, most often the errors are unknown.

A2. Necessity of Error Modeling Before Calibration

Then the question emerges; Is it possible to perform a calibration or validation of one noisy system against the other without prior information on the error characteristics of one or both of the systems? Below we illustrate that this is generally not possible.

Suppose we have a set of true states, indicated by variable t , measured by systems X and Y resulting in measurements x and y . We define $\sigma^2 = \langle t^2 \rangle$, where $\langle \rangle$ denotes the expected mean, and introduce the error model

$$\begin{aligned} x &= t + \delta_x & \varepsilon_x^2 &= \langle \delta_x^2 \rangle \\ y &= t + \delta_y & \varepsilon_y^2 &= \langle \delta_y^2 \rangle \end{aligned} \quad (A3)$$

where x , y , and t are as defined before and δ_x and δ_y are the independent observation errors on x and y respectively, i.e., $\langle \delta_x \delta_y \rangle \approx 0$. The observation errors are random and uncorrelated with t , i.e., $\langle \delta_x t \rangle \approx 0$ and $\langle \delta_y t \rangle \approx 0$. For simplicity we have removed the true distribution's mean and the systematic errors, i.e., $\langle t \rangle = 0$, $\langle \delta_x \rangle = 0$, and $\langle \delta_y \rangle = 0$. We now find

$$\begin{aligned} \langle x^2 \rangle &= \sigma^2 + \varepsilon_x^2 \\ \langle y^2 \rangle &= \sigma^2 + \varepsilon_y^2 \\ \langle xy \rangle &= \sigma^2 \end{aligned} \quad (A4)$$

which are three equations with three unknowns that are easily resolved

$$\begin{aligned} \varepsilon_x^2 &= \langle x^2 \rangle - \langle xy \rangle \\ \varepsilon_y^2 &= \langle y^2 \rangle - \langle xy \rangle \\ \sigma^2 &= \langle xy \rangle \end{aligned} \quad (A5)$$

Thus from the covariances we can resolve the true variance and the standard errors of systems X and Y .

Now suppose that one system is not calibrated, for example, we change (A3) to

$$\begin{aligned} x &= t + \delta_x & \varepsilon_x^2 &= \langle \delta_x^2 \rangle \\ y &= s_Y(t + \delta_Y) & \varepsilon_Y^2 &= \langle \delta_Y^2 \rangle \end{aligned} \quad (\text{A6})$$

where s_Y is the calibration (scaling) constant. We now find

$$\begin{aligned} \varepsilon_x^{*2} &= \langle x^2 \rangle - \langle xy \rangle = \varepsilon_x^2 + (1 - s_Y) \sigma^2 \\ \varepsilon_Y^{*2} &= \langle y^2 \rangle - \langle xy \rangle = s_Y^2 \varepsilon_Y^2 + s_Y(s_Y - 1) \sigma^2 \\ \sigma^{*2} &= \langle xy \rangle = s_Y \sigma^2 \end{aligned} \quad (\text{A7})$$

such that for $\varepsilon_x^{*2} \geq 0$ and $\varepsilon_Y^{*2} \geq 0$ the transformed distribution of $t^* = \sqrt{s_Y} t$ in combination with error model

$$\begin{aligned} x^* &= t^* + \delta_x^* & \langle \delta_x^{*2} \rangle &= \varepsilon_x^{*2} \\ y^* &= t^* + \delta_Y^* & \langle \delta_Y^{*2} \rangle &= \varepsilon_Y^{*2} \end{aligned} \quad (\text{A8})$$

is statistically identical to the system defined in (A3). Consequently, one can show that substituting the transformed quantities defined in (A7-8) into (A2) rather than the real quantities of (A3) leads to an identical joint distribution. As such, the calibration coefficient cannot be resolved unambiguously without further information on the errors in this general case.

This is due to the fact that we have three independent statistical variables, i.e., $\langle x^2 \rangle$, $\langle y^2 \rangle$, and $\langle xy \rangle$, and four unknowns that are s_Y , σ , ε_x , and ε_Y , which leaves one degree of freedom. The constraint that the error variances need to be larger than zero allows the following values for the scaling constant s_Y^* in the case of $s_Y = 1$: $-\varepsilon_Y^2 (\sigma^2 + \varepsilon_Y^2)^{-1} \leq s_Y^* - 1 \leq \varepsilon_X^2 \sigma^{-2}$. Only the addition of an independent third system can help determine the unique and correct value of s_Y .

Appendix B: Pseudobias Through Nonlinear Transformation

A statistical problem that is relevant for considering the optimal error domain may occur through nonlinear transformation. This is easily shown for wind speed f that depends in a nonlinear manner on the wind components (u, v). As an alternative to Hinton and Wylie [1985], we assume normal distributed errors on the wind components u and v for measurement systems X and Y , i.e., $u_X = N(u, \varepsilon_X)$, $v_X = N(v, \varepsilon_X)$, $u_Y = N(u, \varepsilon_Y)$, and $v_Y = N(v, \varepsilon_Y)$ and, in addition, normal distributed true components, $u_t = N(0, \sigma)$ and $v_t = N(0, \sigma)$, where $N(0, \sigma)$ indicates the normal distribution with zero mean and standard deviation σ . The wind speed distribution of system X , $p(f_X) df_X$, then becomes

$$p(f_X) df_X = \frac{f_X}{\sigma_X^2} \exp\left[-\frac{f_X^2}{2\sigma_X^2}\right] df_X \quad (\text{B1})$$

where $\sigma_X^2 = \sigma^2 + \varepsilon_X^2$. A similar expression can be derived for system Y . The expected mean value of f_X will be $\langle f_X \rangle = \sqrt{\pi/2} \sigma_X$, and $\langle f_X - f_Y \rangle = \sqrt{\pi/2} (\sigma_X - \sigma_Y) \approx \sqrt{\pi/2} (\varepsilon_X^2$

$-\varepsilon_Y^2) \sigma^{-1}$. Thus the difference in mean wind speed is generally nonzero, and a pseudobias occurs, since the component error distributions were not biased. More detailed analysis shows that the pseudobias results from the fact that the wind speed error distribution is asymmetric in the case of symmetric wind component error distributions, especially for low wind speeds. The magnitude of the pseudobias depends on wind speed (see, e.g., Figure 2).

Acknowledgments. This work has been greatly stimulated by discussions in the ESA ASCAT Science Advisory Group and in particular by discussions with Didier Le Meur from ECMWF. Tsan Wang Yu from NCEP kindly provided me with the collocation data set and is especially acknowledged.

References

- Bryan, F. O., I. Wainer, and W. R. Holland, Sensitivity of the tropical Atlantic circulation to specification of wind stress climatology. *J. Geophys. Res.*, 24, 729-744, 1995.
- Hinton, B. B., and D. P. Wylie, A correction for the errors in ship reports of light wind, *J. Atmos. Oceanic Technol.*, 2, 353-356, 1985.
- Holton, J. R., An Introduction to Dynamic Meteorology, 2nd ed., *Int. Geophys. Ser.* - vol. 23, Academic, San Diego, Calif., 1979.
- Le Meur, D., L. Isaksen, and A. Stoffelen, Wind calibration of the ERS scatterometer for data assimilation purposes at ECMWF, paper presented at the Wind and Wave Validation Workshop, Comm. on Earth Obs. Satell., Eur. Space Agency, Noordwijk, Netherlands, June 3-5, 1997.
- Lorenz, A. C., Analysis methods for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, 112, 1177-1194, 1986.
- Offiler, D., The calibration of ERS-1 satellite scatterometer winds, *J. Atmos. and Oceanic Technol.*, 11, 1002-1017, 1994.
- Roquet, H. and C. Gaffard, Proposal for a new version of CMOD4, memorandum Forecast Res. Dep., Eur. Centre Medium-range Weather Forecasts, Reading, UK, 1995.
- Rufenach, C., A new relationship between radar cross-section and ocean surface wind speed using ERS-1 scatterometer and buoy measurements, *Int. J. Remote Sens.*, 16(18), 3629-3647, 1995.
- Stoffelen, A., Error modelling of scatterometer, in-situ, and ECMWF model winds: A calibration refinement, *Tech. rep. 193*, R. Neth. Meteorol. Inst., de Bilt, Netherlands, 1996.
- Stoffelen, A., A simple method for calibration of a scatterometer over the ocean, *J. Atmos. Oceanic Technol.*, in press, 1998.
- Stoffelen, A., and D. Anderson, Scatterometer data interpretation: Measurement space and inversion, *J. Atmos. Oceanic Technol.*, 14(6), 1298-1313, 1997a.
- Stoffelen, A., and D. Anderson, Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4, *J. Geophys. Res.*, 102(C3), 5767-5780, 1997b.
- Stoffelen, A., and D. Anderson, Ambiguity removal and assimilation of scatterometer data, *Q. J. R. Meteorol. Soc.*, 123, 491-518, 1997c.
- Wilkerson, J. C., and M. D. Earle, A study of differences between environmental reports by ships in the voluntary observing program and measurements from buoys, *J. Geophys. Res.* 95(C3), 3373-3385, 1990.

A. Stoffelen, KMNI, Postbus 201, 3730 AE de Bilt, Netherlands. (e-mail: stoffelen@knmi.nl)

(Received December 18, 1996; revised June 26, 1997; accepted November 5, 1997.)

Observation of tropical cyclones by high-resolution scatterometry

Y. Quilfen, B. Chapron, T. Elfouhaily, K. Katsaros,¹ and J. Tournadre

Laboratoire d'Océanographie Spatiale, Institut Français de Recherche pour l'Exploitation de la Mer
Plouzané, France

Abstract. Unprecedented views of surface wind fields in tropical cyclones (hereafter TCs) are provided by the European Remote Sensing Satellite (ERS) C band scatterometer. Scatterometer measurements at C band are able to penetrate convective storms clouds, observing the surface wind fields with good accuracy. However the resolution of the measurements ($50 \times 50 \text{ km}^2$) limits the interpretation of the scatterometer signals in such mesoscale events. The strong gradients of the surface wind existing at scales of a few kms are smoothed in the measured features such as the intensity and location of the wind maxima, and the position of the center. Beyond the ERS systems, the scatterometers on-board the ADEOS and METOP satellites, designed by the Jet Propulsion Laboratory and by the European Space Agency, respectively, will be able to produce measurements of the backscattering coefficient at about $25 \times 25 \text{ km}^2$ resolution. A few sets of ERS-1 orbits sampling TC events were produced with an experimental $25 \times 25 \text{ km}^2$ resolution. Enhancing the resolution by a factor of 2 allows location of the wind maxima and minima in a TC with a much better accuracy than at 50 km resolution. In addition, a better resolution reduces the geophysical noise (variability of wind speed within the cell and effect of rain) that dominates the radiometric noise and hence improves the definition of the backscattering measurements. A comprehensive analysis of the backscattering measurements in the case of high winds and high sea states obtained within TCs is proposed in order to refine the interpretation of the wind vector derived from a backscattering model that is currently only calibrated up to moderate winds ($< 20 \text{ m/s}$) in neutral conditions. Observations of the TOPEX-POSEIDON dual-frequency altimeter are also used for that purpose. Patterns of the surface winds in TCs are described and characteristic features concerning asymmetries in the maximum winds and in the divergence field are discussed.

1. Introduction

The surface wind field is one of the most important parameters for estimating the surface heat fluxes that drive the tropical cyclones (TCs) [Emanuel, 1988]. Surface data are very scarce, and satellite-borne radiometers in visible and infrared channels are limited to observing the upper regions of the storms due to the widespread cloud cover. These data are currently the primary source for inferring the maximum surface wind speed in TC forecast centers around the world [Dvorak, 1976]. The polar orbiting European Remote-Sensing Satellites (ERS-1 and ERS-2), launched by the European Space Agency (ESA) on July 17, 1991, and April 21, 1995, respectively, carry the C band active microwave instrument (AMI). The AMI is operated as a synthetic aperture radar for surface imaging and as a scatterometer to

measure the surface wind vector with a nominal resolution of $50 \times 50 \text{ km}^2$. A limited data set of experimental products were also generated by the french space agency (Centre National d'Etudes Spatiales) with a resolution of $25 \times 25 \text{ km}^2$ to investigate its feasibility and its interest for mesoscale meteorology. Both wind speed and direction can be retrieved by inverting an empirical model relating the radar backscatter signals to the 10 m neutral wind vector. The model validity has been verified in a statistical sense [Quilfen and Bentamy, 1994; Graber *et al.*, 1996], but very little is known about the scatterometer signals in TCs, where very high winds and sea states are encountered. A TC case has been selected to illustrate the potentialities of the scatterometer measurements and the advantages of the higher resolution. This case study outlines the great improvement to be expected in TC analysis and forecasting with future scatterometers. Moreover, it shows the need to further analyze the microwave signature in these extreme events, the scatterometer wind speed being underestimated as it was found to be with data from the Seasat scatterometer [Jones *et al.*, 1982]. Furthermore, the heavy rains occurring in TCs are likely to influence the signals at C band. To verify the scatterometer inferred winds in TCs, other microwave data sources providing environmental parameters are used. The

¹Now at Atlantic Oceanographic and Meteorological Laboratory, National Oceanic and Atmospheric Administration, Miami, Florida.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC01911.
0148-0227/98/97JC-01911\$09.00

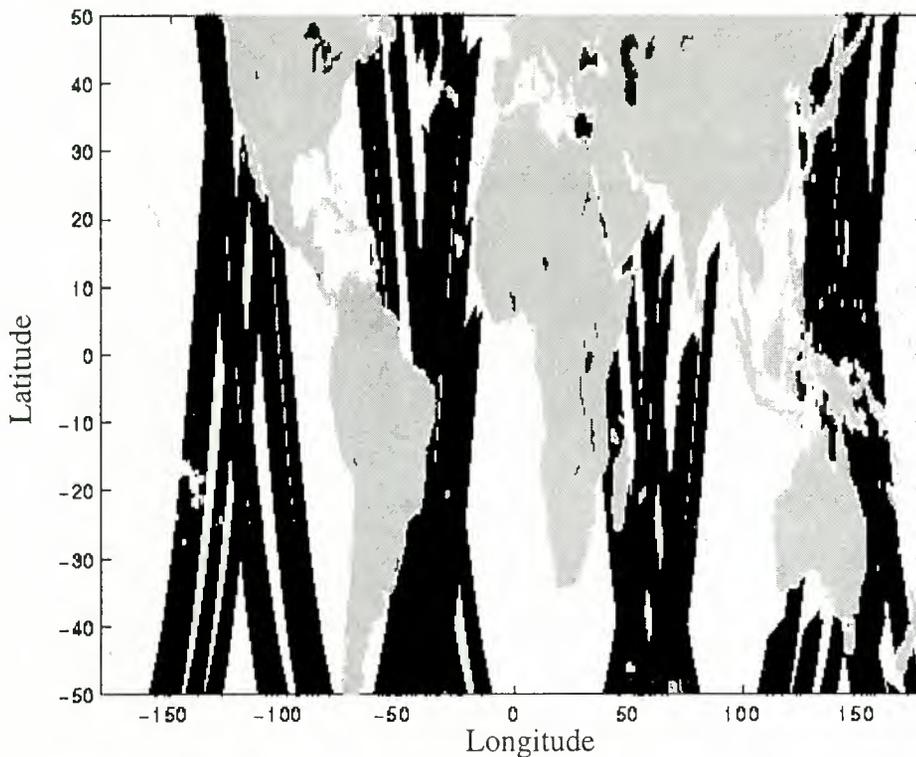


Figure 1. ERS-1 scatterometer coverage for the 29 orbits processed in the HR mode.

altimeter on board the TOPEX-POSEIDON satellite (hereafter T/P) provides high resolution measurements of the wind speed and sea state and can be used to provide an estimate of the rain events [Tournadre and Morland, 1997]. Surface wind and wave parameters at various scales will thus help to better describe the wind and wave patterns within TCs as well as to study the associated air-sea interaction mechanisms. However, altimeters only sample a narrow swath at nadir. The two DMSP satellites F11 and F12 carrying the special sensor microwave imager (SSM/I) overlap with ERS and provide a unique opportunity to relate convective patterns of clouds and precipitation to surface winds within TCs. The SSM/I instruments have a broad swath of 1400 km width and are in a highly inclined Sun-synchronous orbit with local overpass times near dawn and dusk. The resolution of the SSM/I varies from 15 to 50 km as frequency varies from 85 to 19 GHz.

Section 2 presents the data. Section 3 is devoted to a comparative analysis of the normalized radar cross sections (NRCS) for the high-resolution and low-resolution modes of the scatterometer. It shows that measurements at high resolution lead to a better evaluation of the radar measurements in cases of high wind gradients. Improved estimates of maximum wind radii and storm-surge forecasting may then be expected for operational purposes. Section 4 is dedicated to the investigation of the backscattering signal for high wind and sea state conditions as measured by the different microwave instruments. Section 5 describes the wind patterns and their evolution for TC Elsie, a western Pacific typhoon that was well sampled by the ERS-1 scatterometer.

2. Data

The ERS scatterometers use three antennae pointing in azimuths of 45°, 90° and 135° with respect to the satellite ground track. It obtains thus for each node three independent measurements of the NRCS. Each antenna illuminates a 500-km-wide swath, covering a range of incidence angles from 18° to 60°.

ERS-1 scatterometer data used to compute winds are the NRCS (σ_0) provided by the Centre National d'Etudes Spatiales [Malardé, 1992]. Twenty-nine orbits covering TCs occurring between July 1992 and November 1992 were processed. The data coverage is shown in Figure 1. Two grids are available for the NRCS measurements sampled over the 500-km-wide swath: a 25 km grid for the 50 km low-resolution mode (LR) and a 12.5 km grid for the 25 km high-resolution mode (HR). For each node a noise estimation from two parameters, the radiometric resolution K_p and the NRCS standard deviation STD, can be associated with each of the three antennae measurements.

The K_p is the classical parameter used to characterize the noise. It is defined as the standard deviation of the raw NRCS measurements used to evaluate the mean NRCS value characterizing the target over a given area [Fischer, 1972]. The theoretical developments presented by Fisher [1972] lead to the following formulation for K_p :

$$K_p = \left[\frac{1}{N} \times \left(1 + \frac{2}{\text{SNR}} + \frac{1}{\text{SNR}^2} \times \left(1 + \frac{T_{sn}}{T_N} \right) \right) \right]^{1/2} \quad (1)$$

where N is the number of independent measurements used to estimate the NRCS (about 300 for the HR and 1200 for the LR), SNR is the signal to noise ratio, T_{sn} is the signal integration time and T_N is the noise integration time.

As this computation makes some assumption about the target characteristics, one being its stationarity, we also use another parameter to characterize the noise, STD, defined as

$$\text{STD} = \left[\frac{1}{N} \times \sum_{i=1}^N \left(\sigma_{0i}^2 - \sigma_0^2 \right) \right]^{1/2} / \sigma_0 \quad (2)$$

where N is the number of independent measurements (σ_{0i}) used to estimate the mean NRCS (σ_0).

The NRCS produced by the CNES were calibrated and validated against those delivered by ESA, showing the consistency of the two NRCS sources [Malardé, 1992].

To compute the wind vector, NRCS data were processed using the C band model described below (referenced as CMOD_IFR2) and with the algorithms described by Quilfen and Bentamy [1994]. Following Long [1992], the C band model is expressed as

$$\sigma_0 = b_0(1 + b_1 \times \cos\varnothing + b_2 \times \cos 2\varnothing) \quad (3)$$

where $b_0 = 10^{a+b\sqrt{V}}$ $\varnothing = \Phi - \Psi$

a and b are arrangements of polynomials of order up to 3 in θ , and b_1 and b_2 are arrangements of polynomials of order up to 2 in θ and V ; θ is the incidence angle, V is the wind speed, Φ is the wind direction, and Ψ is the antenna azimuth angle. The model is fully described in appendix.

Parameters a , b , b_1 , and b_2 were calibrated with a method minimizing a maximum likelihood estimator between the σ_0 measured values and the C band model values [Bentamy et al., 1994]. Data from the U.S. National Oceanic and Atmospheric Administration (NOAA), collected by 24 buoys considered to be moored far enough from the shore during the time period March-June 1992, were used to calibrate a and b . Selected data from the European Centre for Medium-Range Weather Forecasts (ECMWF) wind analysis were used to calibrate b_1 and b_2 [Quilfen and Bentamy, 1994]. As there were not enough reference data of high wind speed to define and to calibrate the C band model shape for these conditions, a mean wind speed dependent bias was estimated and is used to correct the wind speed computed by the model inversion process beyond 10 m/s. Indeed, the calibration process is a global one-step process and would have required many high-wind measurements to converge, while much less measurements are required to estimate a bias correction. An extended buoy data set was used to estimate this bias correction (March 1992 to October 1993). The bias correction is described in the appendix and its adequacy is discussed in section 4.

For the validation, another extended NOAA buoy data set is used for the time period July 1992 to February 1994. It is independent of the one used to calibrate the C band model, with the exception of the high winds used to estimate the bias term between March 1992 to October 1993.

We also use in this study one track of the T/P altimeter delivered by the Archiving, Validation, and Interpretation of Satellite Data in Oceanography (AVISO) processing center [AVISO, 1992]. This track intersects TC Elsie close to its center on November 4, 1992.

In order to identify convective regions, the Petty and Katzaros [1990] algorithms have been used to infer the following parameters from SSM/I brightness temperature data on board the F11 DMSP satellite: the integrated atmospheric water vapor, the integrated cloud liquid water, and an index of scattering by large ice particles (based on a combination of brightness temperatures at 85 GHz). Large values of the latter parameter are indicative of convection.

3. Comparison of the NRCS Distributions for the High- and Low-Resolution Data

3.1. Characterization of the Noise

The accuracy of NRCS measurements is generally defined in terms of radiometric resolution Kp (defined in section 2). Kp provides an estimation of the noise contribution from the receiver and due to the speckle (the random nature of the target). This calculation relies on the basic assumption of stationarity of the target, which is not always true, as is outlined below.

In comparing the noise for the HR and LR measurements, we are also interested in the non stationarity of the target. For this purpose, a measure of the NRCS standard deviation (STD) defined in section 2 is performed for each NRCS measurement, in addition to the Kp estimation, which is a more theoretical and statistical estimation of the noise. Non stationary effects are mainly due to the mesoscale geophysical variability (a few kms to a few hundred kms in scale), which is of great interest for the dynamics of the atmosphere. In particular, it will be shown that the mesoscale wind patterns are better captured with the HR measurements.

Figure 2 shows the mean values of the NRCS, the radiometric resolution Kp and the standard deviation STD, for the LR and HR respectively, over the 29 orbits (about 300,000 measurements, land points being discarded by using the same mask for the two resolutions). Data collected to compute the distributions are sampled every 50 km. The differences between the mean NRCS for the HR and LR never exceed 0.15 dB, which is under the noise level of about 0.3 dB. Averaging the NRCS over different cell sizes does not introduce any significant bias (at the 95% confidence level, not shown).

The mean Kp varies from 2.8% to 4% for the LR and from 5.5% to 7.2% for the HR. Values of Kp for the HR re-

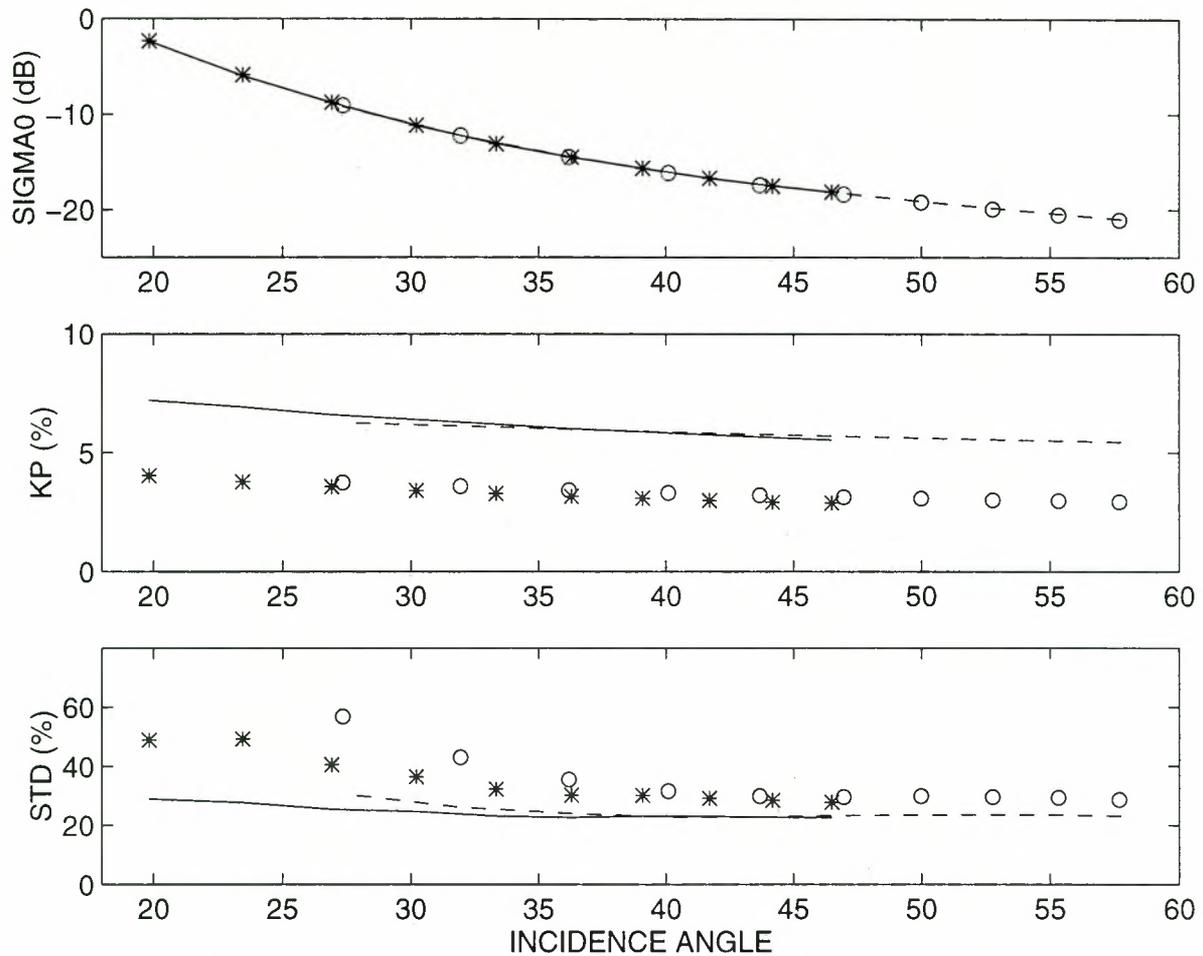


Figure 2. Mean values as a function of the incidence angle (in degrees): (top) NRCS σ_0 in decibels, (middle) radiometric resolution Kp in percentage, and (bottom) NRCS standard deviations STD in percentage. Solid and dashed lines are for the HR measurements of the central and lateral antennae, respectively; stars and open circles are for the LR measurements of the central and lateral antennae, respectively.

main lower than the ESA initial specifications (of the order of 8%). As the Kp depends upon the signal to noise ratio, it depends on the incidence angle but also on the wind conditions. It is higher for low wind speeds and for cross-wind measurements. It is rather difficult to compare the LR and HR measurements in terms of the accuracy of the estimated NRCS because the computed Kp includes only a part of the total noise, i.e., the receiver thermal noise and the speckle.

The STDs also include the geophysical effects, mainly due to the presence of wind gradients or to attenuation variations due to rain, and the variability due to the NRCS cross-track gradients. The gradient being higher at low incidence angles, the mean STDs are also higher. Thus both effects, the geophysical variability that is reduced for the HR and the NRCS cross-track gradients, can explain why the mean STDs are higher for the LR although the Kp are lower.

In order to better characterize the NRCS noise, the distribution of STDs over the 29 orbits was computed and is displayed in Figures 3a and 3b for incidence angles of 27° and 57°, respectively. As was noted above, it appears that the mean STDs are higher for the LR, mainly on account of the cross-track NRCS gradient. The most interesting feature is

that the dispersion is much lower for the HR, meaning that the non stationary part of the target, i.e., the geophysical effects (wind and sea state variability within the cell or rain effects), is greatly reduced. The NRCS estimation is thus improved in many cases.

3.2. Example of a Geophysical Phenomenon

Figure 4 gives a typical example summarizing the features discussed above. It shows the NRCS standard deviations for the case of TC Elsie (November 3, 1992) which will be analyzed in more detail in the following sections. There is much more structure in the LR measurement noise (Figure 4b), while STDs are more uniform for the HR mode (Figure 4c). STDs are much higher for the low incidence angles of the LR measurements, while this effect is much reduced in the case of the HR. Moreover, non stationary effects in the 50 km cells due to the geophysical variability are clearly highlighted in the area near 12°N, 142°E. This region corresponds to high incidence angle and high wind measurements. As shown in the wind field (Figure 4a), strong gradients occur with low level convergence of the flow in the western eyewall of the TC. The maximum STDs

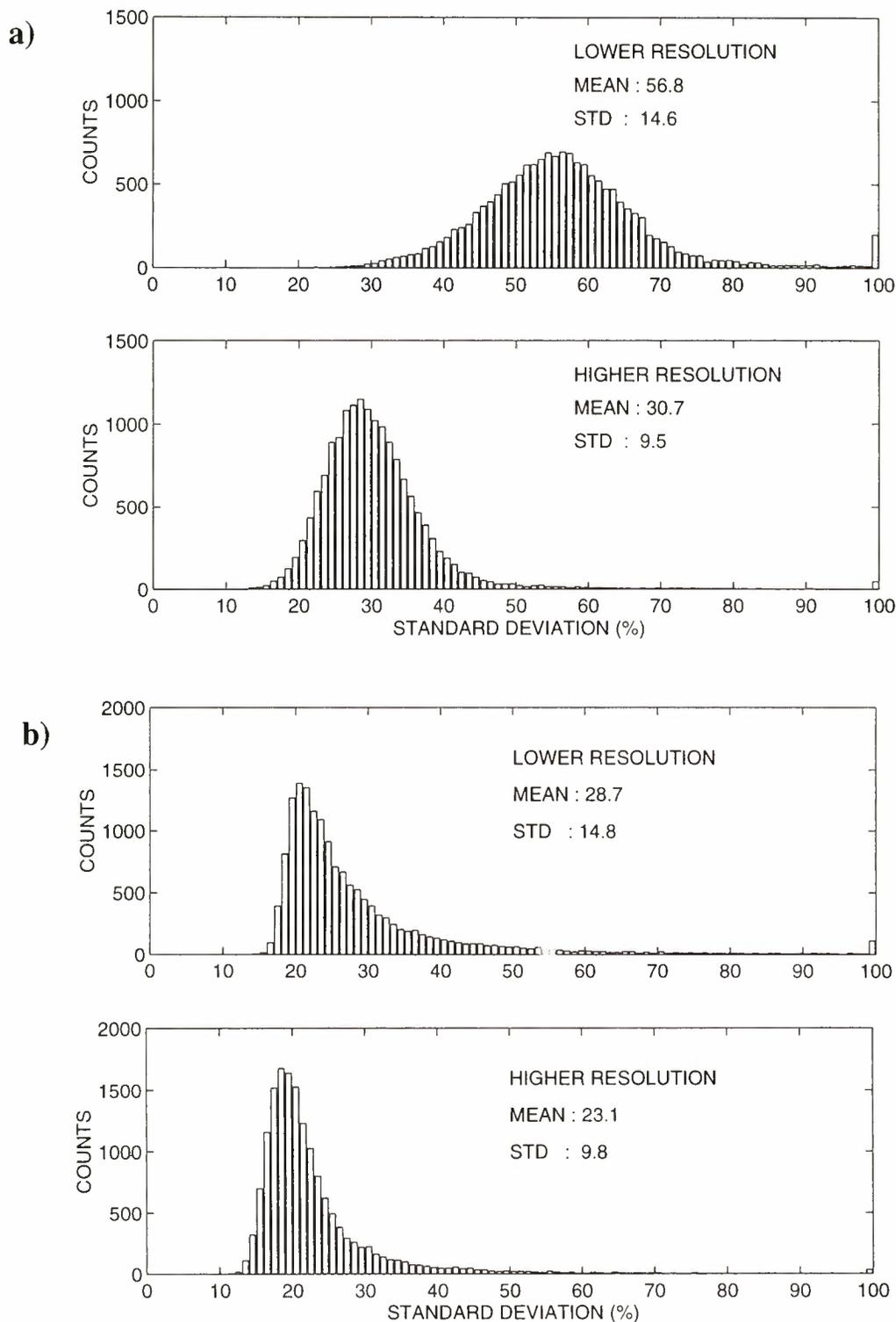


Figure 3. Distribution of the NRCS standard deviations (in percentage) for the HR and LR NRCS measurements for incidence angles of (a) 27° and (b) 57°.

in this area are of the order of 50% for the LR measurements and of the order of 30% for the HR measurements. Such a result is much closer to the mean value of 23% (Figure 3b), which confirms that the HR measurements may provide a better estimate of the NRCS field and thus a better representation of the wind structures.

Another maximum of STD occurs in the area centered on the island of Guam (approximately 145°E, 13.4°N). The

mean level of the backscattered signal is different for the sea and for the land. This discontinuity is characterized by high STDs values.

It appears that the ERS scatterometer high resolution does not deteriorate the radar cross-section measurement. Further, it reduces the non stationary part of the wind variability within the cell to improve the definition of the wind structures.

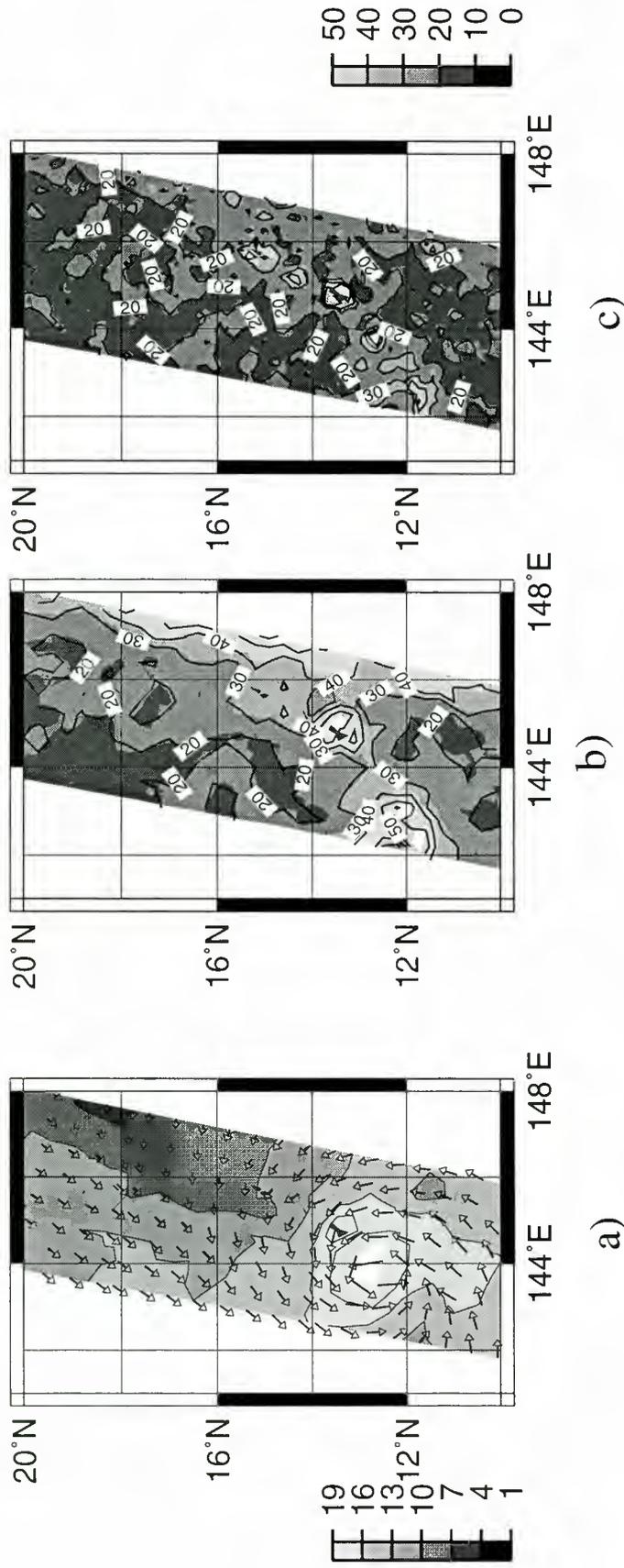


Figure 4. (a) Surface wind field (left scale in m/s); (b) LR NRCS standard deviations; (c) HR NRCS standard deviations (right scale in percentage). Case of tropical cyclone Elsie (November, 3, 1992).

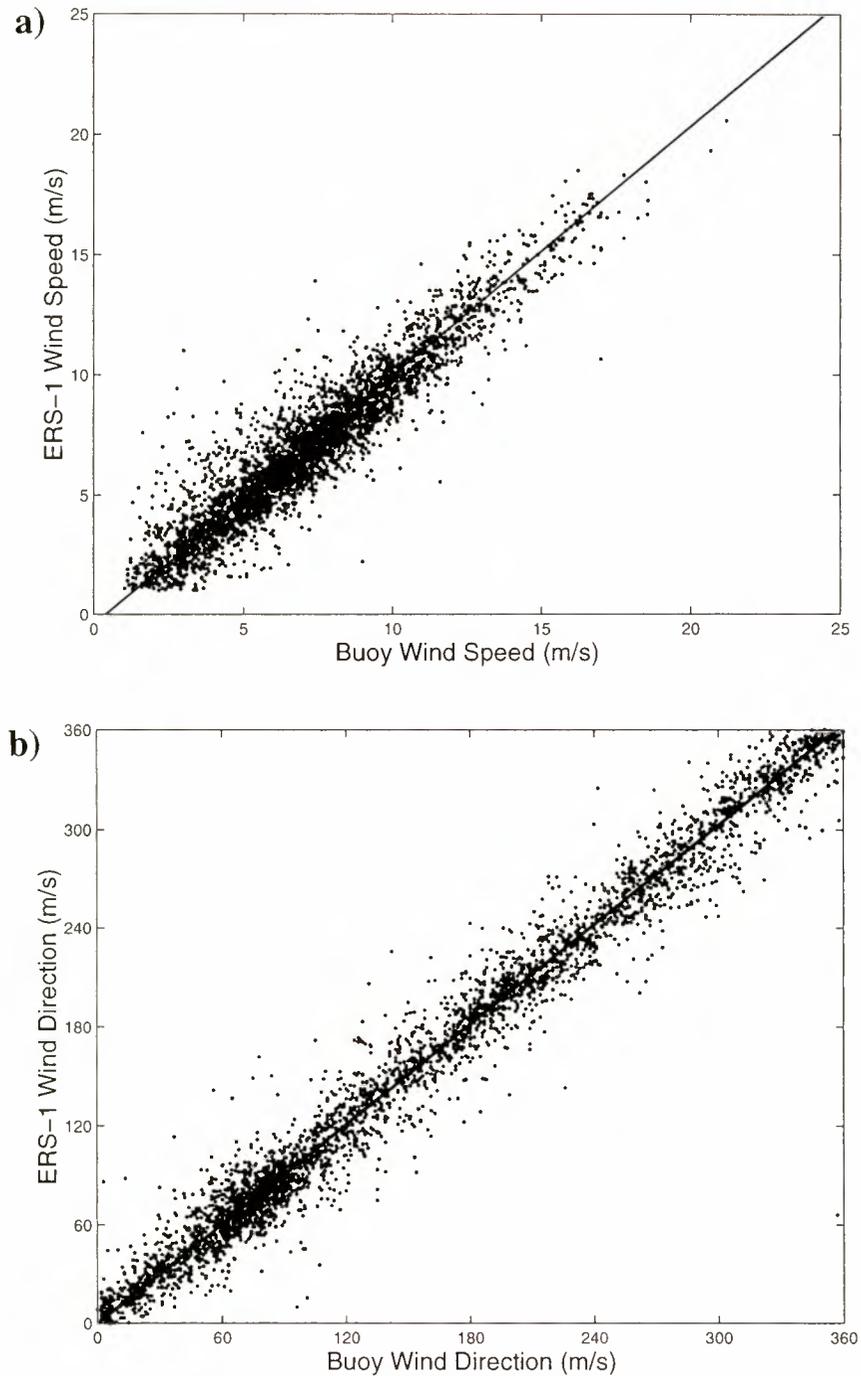


Figure 5. Comparison of (a) the wind speeds (m/s), and (b) the wind directions (degrees) observed by the ERS-1 scatterometer with those measured by the NOAA buoys. The symmetrical regression lines are shown.

4. The Microwave Signature at High Wind Speeds

4.1. Evaluation of the C Band Model

High wind speed (above 15 m/s) measurements at the sea surface by means of anemometers are very scarce, and their accuracy is being questioned for high sea-state conditions because of the buoy movements and the anemometers being shadowed by the waves [Large et al., 1995]. For this reason and also because storms are often inaccurately represented

in the numerical weather prediction models at the scale of the scatterometer footprint, there is no good reference high-wind data set, except for airborne flight level measurements in atlantic hurricanes, against which microwave measurements may be calibrated. Consequently, the scatterometer C band model CMOD_IFR2 uses very few high-wind measurements, and the model shape for high winds (above 15 m/s) must be considered as an extrapolation of the shape for moderate to high winds. Since predicted NRCS values were

found to increase too fast with the wind speed [Bentamy *et al.*, 1994], high wind speed is likely to be underestimated. Thus a bias correction of the scatterometer winds has been estimated (see appendix) and is applied to better estimate high winds. Scatterometer winds at 50 km resolution have then been validated against other data sources and intensively compared with collocated measurements (within 25 km and 10 mn) of the NOAA buoys. Buoy measurements are converted to 10 m neutral winds. Between July 1992 and February 1994, 3433 collocated pairs were thus available to perform such comparisons. For each ERS-1 cell the wind vector considered, among the four possible solutions, is the one whose direction is the closest to the buoy wind direction [Quilfen and Bentamy, 1994]. Figures 5a and 5b show an overall reasonable agreement, also shown by Graber *et al.*, [1996]. A symmetrical regression is performed to take into account the fact that each data set has its own intrinsic errors, the regression coefficients thus being invariant with respect to the interchange of buoy and scatterometer data. The mean biases are negligible at a significance level of 95%. The root-mean-square (rms) errors are 1.38 m/s and 18.6° for the wind speed and direction, respectively. Correlation coefficients between buoy and ERS-1 winds are 0.91 and 0.98 for speed and direction, respectively. The symmetrical regression coefficients are 1.04 and 1.01 for speed and direction, respectively. However, as was mentioned above, the bias is likely to be wind speed dependent, and although a bias correction is applied, differences are still significant for winds greater than 15 m/s [Graber *et al.*, 1996]. It must be noted that these biases are difficult to quantify because the high-wind buoy measurements are also of lower accuracy and because of their large temporal and spatial variability. Graber *et al.* also showed that at high winds the estimated bias is even greater for the ESA model function CMOD4 [Stoffelen and Anderson, 1997], while it is close to 0 for the model function developed by the Jet Propulsion Laboratory [Freilich and Dunbar, 1993]. The latter model is defined with a tabulated model function and will not depend on a mathematical expression. It can then be hypothesized that biases at high winds in CMOD4 and CMOD_IFR2 are related to an improper extrapolation by both models: the NRCS square root dependency as a function of the wind speed or the model coefficients do not seem to hold at high wind speed (above 15 m/s).

To better illustrate such a conclusion, the measured and predicted by CMOD_IFR2 NRCS have been compared

(Figure 6). Six weeks of scatterometer measurements collocated with the ECMWF analyses have been considered because there are not enough buoy data to do this analysis. For each week the model parameters b_0 , b_1 , and b_2 (see section 2) have been estimated. Mean values and standard

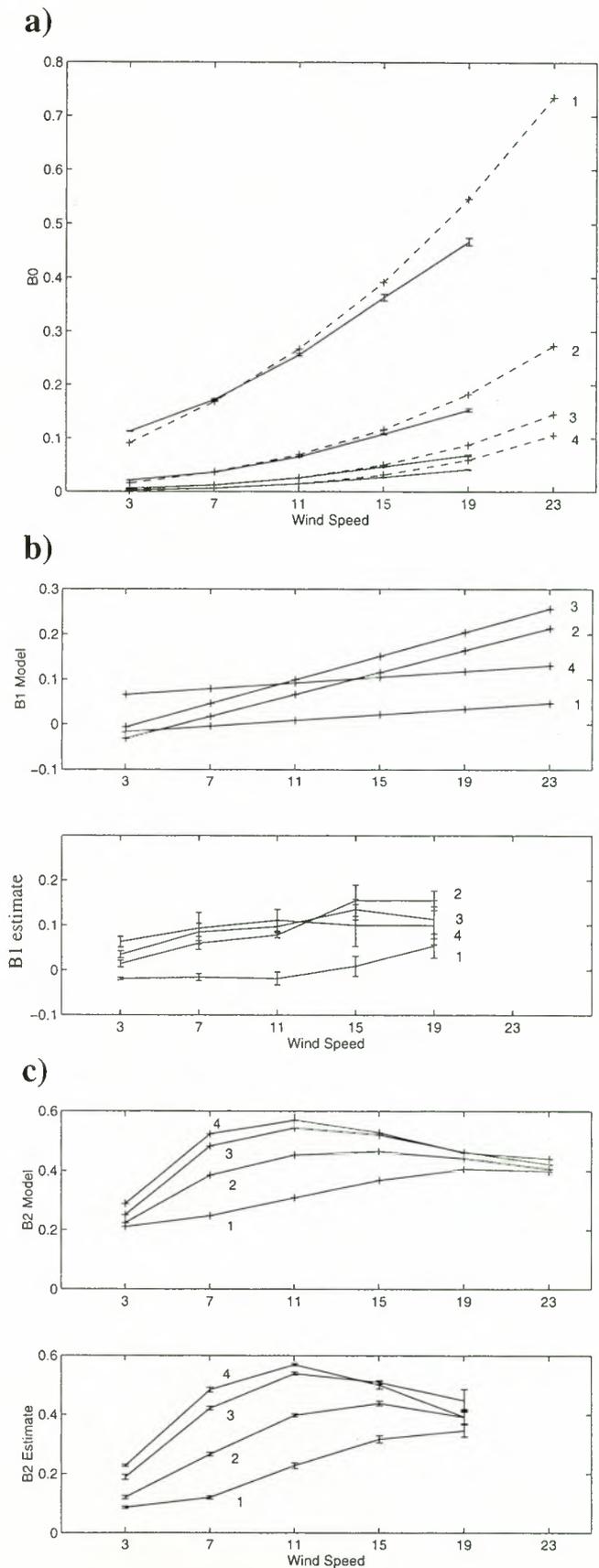


Figure 6. (a) Dependence of the term b_0 on the wind speed (m/s) as provided by CMOD_IFR2 (dashed line) and as estimated from measured NRCS (solid line). (b) Dependence of the term b_1 on the wind speed (m/s), (top) as provided by CMOD_IFR2 and (bottom) as estimated from measured NRCS. (c) Dependence of the term b_2 on the wind speed (m/s), (top) as provided by CMOD_IFR2 and (bottom) as estimated from measured NRCS. Computations are done for 25° (curve 1), 35° (curve 2), 45° (curve 3), and 55° (curve 4) of incidence angle.

deviations are then compared with the CMOD_IFR2 values. Computations are made as follows:

$$\begin{aligned}
 b_0 &= (S_u + S_d + 2xS_c)/4 \\
 b_1 &= (S_u - S_d)/(2b_0) \\
 b_2 &= 1 - (S_c)/b_0
 \end{aligned}
 \tag{4}$$

where S_u , S_d , and S_c are the NRCS sampled in 20° degrees wide sectors in the upwind, downwind, and crosswind directions, respectively. The ECMWF wind speed and direction are taken as references for the binning process. The same computations have been done by using as a reference the scatterometer wind speed and direction in order to avoid as much as possible the effects of the reference wind errors. Computing the estimates in this manner makes no significant changes in the results (the two different estimates remain within their error bars).

The b_0 term (Figure 6a) is indicative of the NRCS wind speed dependency and shows a slower growth than is predicted by CMOD_IFR2. It confirms thus the tendency of the CMOD_IFR2 to underestimate the winds above 10 m/s. This effect is partly compensated by adding the bias term described in appendix. There is no indication on how it evolves above 20 m/s.

The b_1 and b_2 terms (Figures 6b and 6c) are indicative of the azimuthal modulation of the backscattering signal,

with b_1 related to the upwind/downwind asymmetry and b_2 related to the upwind/crosswind anisotropy. Like the b_0 term, the b_1 term reproduces the observations for moderate winds well and strongly overestimates the upwind/downwind asymmetry for high winds. As depicted by the theoretical models, upwind/downwind asymmetries arise mainly by the modulation of the short resonant Bragg waves by the long waves [Plant, 1986; Elfouhaily, 1997]. The degree of sea state development and long-wave-induced wind stress modulations are thus likely to influence the b_1 term. These effects are not well known and the observations do not provide any indication above 20 m/s.

The behavior of the b_2 term is well reproduced by CMOD_IFR2 for a broad range of wind speeds, but it is overestimated for the lowest incidence angles. Its decrease at high winds is reproduced, but there is still no indication of how it behaves above 20 m/s. There were not enough observations to calibrate CMOD_IFR2 in this wind speed range, and the b_2 term is thus constrained by the normalization of the wind speed applied in CMOD_IFR2 (see appendix). The upwind/cross-wind anisotropy is associated with the directionality of the Bragg waves. We thus do not know if there are sufficient azimuthal NRCS modulations for wind speeds higher than 20 m/s to retrieve the wind direction with good accuracy.

At this point the question remains, whether there is sufficient modulation of the mean backscattering signal as a function of the wind speed and azimuth angle to estimate

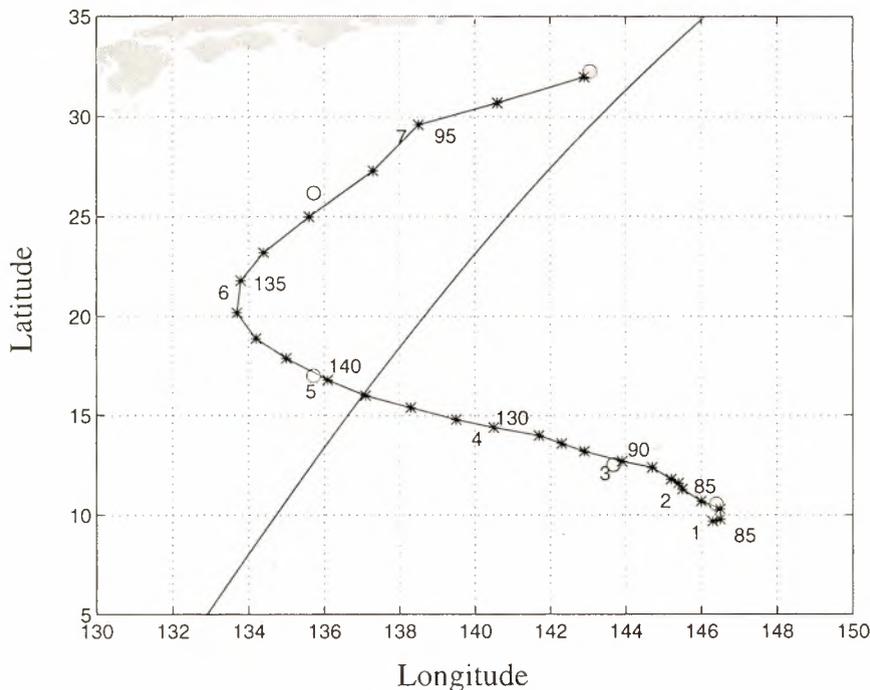


Figure 7. TC Elsie track as provided by the Joint Typhoon Warning Center (JTWC) every 6 hours (stars) from November 1 to November 7, 1992. The day numbers are indicated on the left, and the peak intensities in knots (1 knot = 0.5 m/s) are indicated on the right. The open circles indicate location of the TC center as derived from the ERS-1 scatterometer. The TOPEX-POSEIDON altimeter track intersects the TC track on November 4 at 1730 UTC.

the wind characteristics above 20 m/s. To investigate this topic, we propose in the following section to analyze qualitatively the surface microwave signature provided by different satellite sensors for the case of a tropical cyclone.

4.2. The Microwave Signature in Tropical Cyclone Elsie

4.2.1. Microwave signature of the ERS-1 scatterometer. The case of TC Elsie (November 1992) was chosen because it was an average-sized, intense typhoon that was well sampled by ERS-1 and for which there was one track of the T/P altimeter near the center of the storm. There were also several SSM/I passes available from the F11 satellite. Figure 7 displays the location of TC Elsie every 6 hours, as determined by the Joint Typhoon Warning Center (JTWC) on Guam, together with the estimated maximum 1 min average sustained winds derived from the *Dvorak* [1976] satellite estimation method. The T/P altimeter swath and the positions of the TC center as determined from the ERS-1 scatterometer data are also shown. The TC locations from ERS-1 are found by locating the minimum high resolution wind speed in the TC center.

Elsie's track was directed west-northwest while the storm was intensifying until November 5 with a peak intensity of 150 knots (75 m/s). The recurvature occurred on November 6, after which Elsie accelerated toward the northeast and became an extra tropical storm on November 7. Between November 1 and 7, five ERS-1 scatterometer swaths sampled the cyclone.

As anticipated, ERS-1 computed wind speeds are much lower than the JTWC estimates (see Table 1). The following reasons can be given:

1. The scatterometer measurements are spatial averages over a footprint of about 25 km where strong gradients occur, while the maximum wind occurs a few tens of kms from the TC center, according to the Holland or Rankine profiles [Franklin *et al.*, 1993]. The two estimates of the wind speed maximum are not strictly comparable because JTWC provides 1 mn local averages and due to their space/time separation. Nevertheless, they depict the same evolution of the TC intensity with a peak on November 5.

2. As outlined in section 4.1, the C band model overestimates the NRCS corresponding to high winds. The wind speed is therefore underestimated, and the sensitivity to increasing wind speed is too low for the very high wind speeds. This is illustrated in Table 1 where the maximum wind speed only increases by 2 m/s between November 3 and 5 when the 1 min average estimated by JTWC increases by 25 m/s.

3. Even though C band signals are not too strongly affected by atmospheric conditions, heavy precipitation and ice particles in the clouds may attenuate the NRCS signals and the rain droplets may also affect the sea surface roughness (wave damping). Such effects are hardly detectable without coincident measurements of the atmospheric parameters and because rainbands in TCs are generally also related to modification of the surface wind patterns [Anthes, 1982].

Table 1. Estimated Maximum Wind Speed According to ERS-1 Scatterometer HR Measurements for Five ERS-1 Passes and the Concurrent JTWC Wind Estimates

Date	Time, UTC	Peak Wind Speed, m/s	
		ERS	JTWC
Jan. 11, 1992	12 40	20.8	42.5
March 11, 1992	00 40	26.4	45.0
May 11, 1992	01 20	28.8	70.0
June 11, 1992	13 30	28.5	67.5
July 11, 1992	13 05	27.1	47.5

JTWC, Joint Typhoon Warning Center (Guam).

We will show that the dual-frequency altimeter T/P can be used to qualitatively estimate the influence of rain on the NRCS signal at C band.

The microwave signature of the sea surface in TCs is illustrated by Figures 8 and 9 with along-swath sections of the NRCS and wind speed at constant incidence angles as a function of the latitude, through the wind maximum and through the TC center. Two ERS-1 passes are shown occurring on November 3 (Figure 8) and on November 5 (Figure 9). These two passes correspond respectively to the intensifying and to the mature phases of Elsie, stages where the wind gradients and intensities are the highest. On November 3, the incidence angles of the NRCS sections are 36.4° and 38.5°, respectively, through the TC maximum and through the TC center. On November 5, there is only one NRCS section at 35.7° incidence angle crossing the TC center as well as the TC maximum (see Plate 2 to locate the sections). To meaningfully describe the microwave signatures, the following quantities S_1 and S_2 have been used rather than the individual NRCS for each antenna:

$$\begin{aligned} S_1 &= (\sigma_1 + \sigma_3)/2 \approx b_0 \\ S_2 &= (\sigma_1 - \sigma_3)/(\sigma_1 + \sigma_3) \approx b_2 \times \cos 2\phi \end{aligned} \quad (5)$$

where σ_1 and σ_3 are the NRCS for the lateral antennae, ϕ is the wind direction relatively to the central antenna and b_0 and b_2 are the model parameters as already described in section 2.

These quantities are first-order approximations, assuming a negligible upwind/downwind difference, i.e., that $b_1 \ll 1$. S_1 does not depend on the wind direction, the two lateral antennae being separated by 90°, and is thus indicative of the behavior of the NRCS as a function of the wind speed. S_2 , designed as the anisotropy coefficient, characterizes the NRCS azimuthal anisotropy. It depends mainly on the wind direction and depends slightly on the wind speed only through b_2 as shown in Figure 6c. A maximum in S_2 corresponds to an upwind or downwind (crosswind) case for the forebeam (aft beam), and a minimum corresponds to a crosswind (upwind or downwind) case for the forebeam (aft beam).

Computations are carried out for the HR as well as for the LR scatterometer measurements. One can make the following observations:

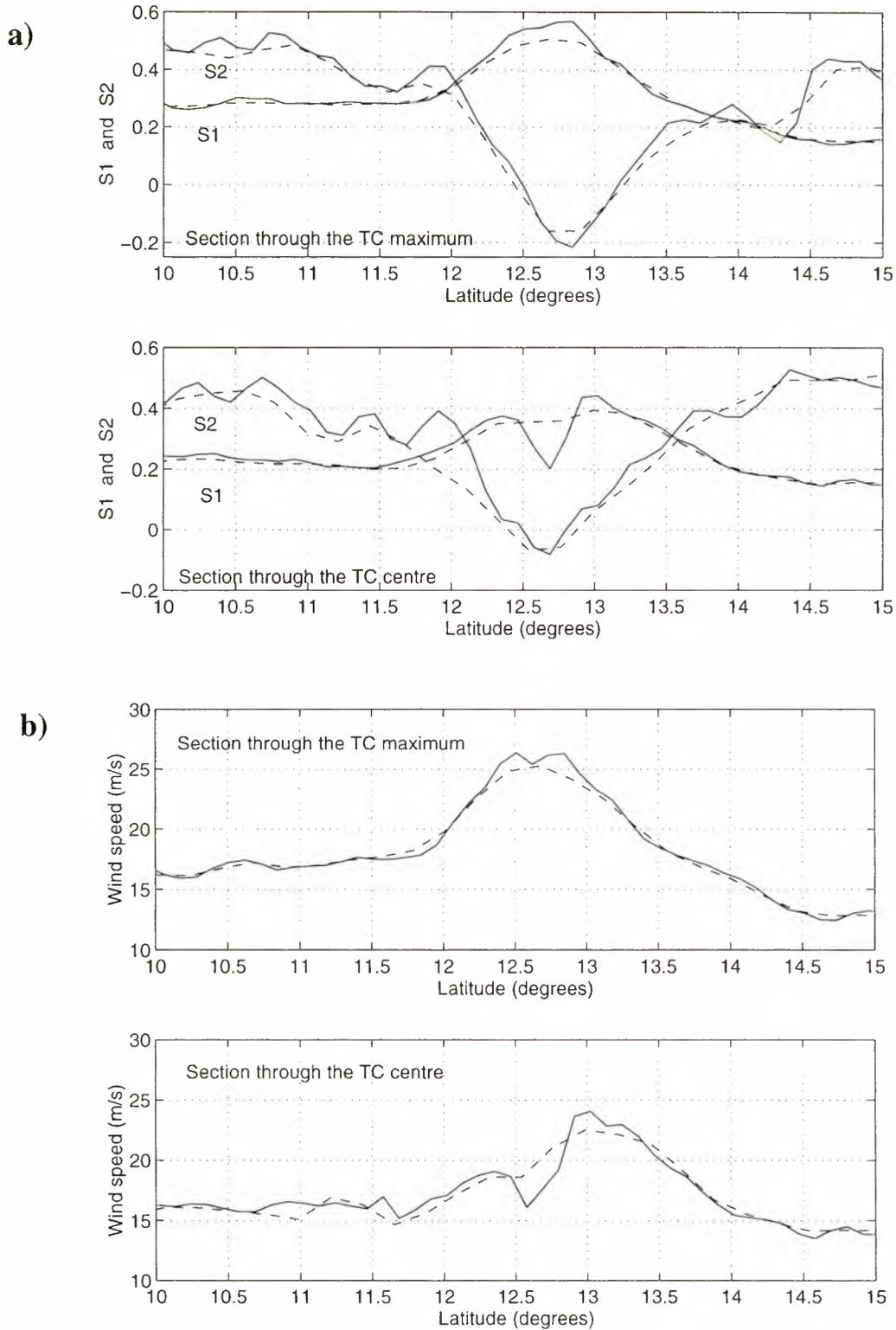


Figure 8. Sections of (a) S_1 and S_2 , (b) wind speed, (top) through the region of maximum winds (incidence 36.4°) and (bottom) through the TC center (incidence 38.5°) on November 3, 1992. Solid and dashed lines are for the high- and low-resolution modes, respectively.

1. In the sections crossing the TC wind maxima (top panel of Figure 8a, and Figure 9a), S_1 changes smoothly on both sides of the TC and does not exhibit a saturation at the maximum level. The NRCS are significantly greater for the HR, meaning that taking a spatial average is of concern at such

high wind speeds. The HR allows better representation of the sea surface roughness variability, but it is not accompanied by a great difference in the maximum wind speed between the HR and LR measurements owing to the lack of sensitivity of the C band model (Figure 8b and 9b).

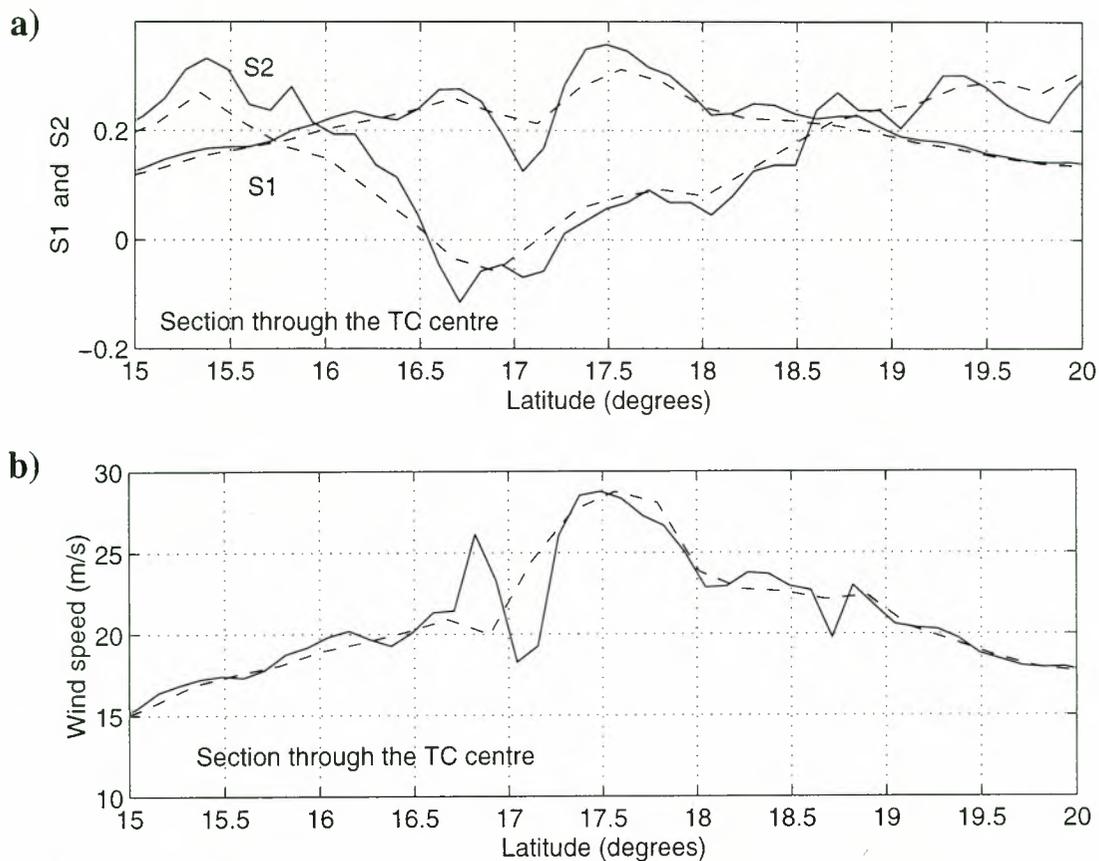


Figure 9. Sections of (a) S_1 and S_2 , (b) wind speed as a function of the latitude, through the TC center (also the region of maximum winds, incidence 35.7°) on November 5, 1992. Solid and dashed lines are for the high- and low-resolution modes, respectively.

The coefficient of anisotropy, S_2 , exhibits a variability mainly linked to the wind direction variability. On November 3, in the area of maximum winds (above 20 m/s between 12°N and 13.5°N), S_2 varies from a value of 0.4 (wind direction upwind for the forebeam or cross-wind for the aft beam) at 12°N to a value of -0.2 (wind direction upwind for the aft beam or cross-wind for the forebeam) at 12.8°N . It means that the upwind forebeam NRCS is roughly 75% or 2.4 dB higher than the cross-wind aft beam value at 12°N (wind speed about 20 m/s) and the upwind aft beam NRCS is 50% or 1.8 dB higher than the cross-wind forebeam value at 12.8°N (wind speed > 25 m/s). These values for a wind speed of 20 m/s are of the same order as those provided by CMOD_IFR2 for medium wind speeds, showing therefore still a sufficient azimuthal modulation. If the physical basis of the azimuthal anisotropy is relatively well understood for moderate winds, it is not obvious what scattering mechanism dominates at the sea surface under the extreme conditions that occur in TCs. The well-defined behavior of the NRCS measurements in TCs is very encouraging and can explain why patterns of retrieved wind fields fit to known wind patterns of TCs (Plate 2).

2. In the sections crossing the TC centers (bottom panel of Figure 8a, and Figure 9a), the previous remarks also hold. Moreover, a striking feature is the notable difference between the HR and LR measurements. Indeed if the TC cen-

ter is well delineated by a drop in the HR values of S_1 on November 3 (bottom panel of Figure 8a near 12.7°N) and on November 5 (Figure 9a near 17.1°N), it is drastically smoothed in the LR measurements. It is also shown in the wind speed patterns (Figures 8b and 9b) for which significant differences appear between the HR and LR patterns. It appears that the 25 km resolution is critical to retrieve the main TC wind pattern characteristics.

In summary, the scatterometer NRCS signal still exhibits a very well defined modulation in the regions of maximum winds and in the center of an intense typhoon, allowing retrieval of the wind vector patterns. Additionally, it is shown that the HR measurements allow a smaller scale microwave signature in TCs and thus improvement of the description of the TC structure. Nevertheless, the scatterometer transfer function is not really calibrated for such extreme conditions. Moreover, the influence of heavy rain or ice particles on the ERS-1 C band scatterometer signal is not known and may be responsible in part for the underestimation of the maximum winds.

In the following section we analyze the microwave signature of the T/P altimeter in order to provide another perspective on the problems mentioned above.

4.2.2. Microwave signature of the TOPEX-POSEIDON altimeter. The dual-frequency T/P altimeter obtains C band and Ku band NRCS measurements at nadir (No-

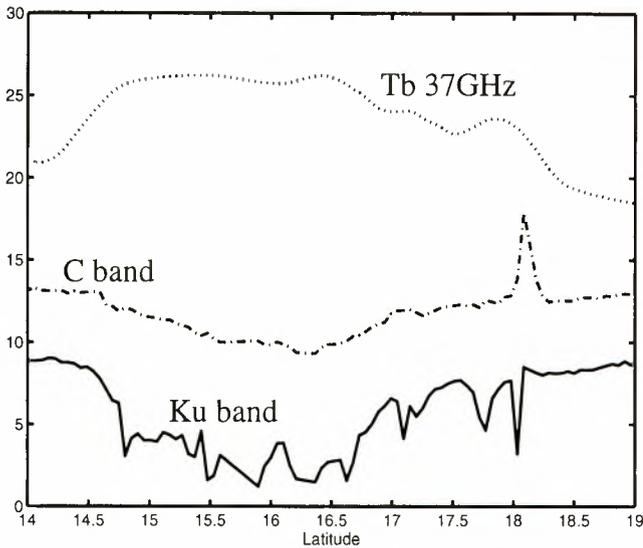


Figure 10. TOPEX-POSEIDON altimeter data near Tropical Cyclone Elsie: NRCS in Ku band (solid) and in C band (dashed-dotted), in decibels; TMR brightness temperature at 37 GHz (dotted), in units of $0.1 \times ^\circ\text{K}$. High modulation is observed in Ku band while C band is still smoother. The TMR brightness temperature increases near the cyclone for the latitudes 14° to 19°N .

November 4, 1992, 1730 UTC), just a few kms ahead of Elsie’s center (see Figure 7 and Plate 1).

4.2.2.1. Atmospheric effects: Figure 10 displays the NRCS values, together with the T/P Microwave Radiometer (TMR) brightness temperatures at 37 GHz, near the TC center (around 16°N). Altimeter NRCS values at C band and Ku band decrease towards the TC center and are related mainly to increasing wind speed. The Ku band σ_0 decreases more and faster than the σ_0 at C band, but it is not clear if this is due to a greater sensibility to the sea surface roughness or to greater interactions with the atmospheric precipitation and cloud particles. There is much more small-scale variability, of the order of a few tens of kms, in the Ku band signal. It can be hypothesized that this variability is due to large attenuation of the signal at Ku band by rain cells or ice particles. The altimeter has lost the signal near the outer convective band (Plate 1) at 18°N , and the σ_0 measurements are thus erroneous at this location. At this location both NRCS signals and also the significant wave height (not shown) have large deviation from the mean behavior, but they deviate in opposite directions. This behavior was observed earlier in T/P altimeter measurements by *Tournadre and Morland* [1997].

If the C band signal is also likely to be affected by atmospheric particles or by the effect of rain droplets on the sea surface roughness, a strong or systematic influence is not discernible as it is in the Ku band signal. With the exception of the anomaly described above, we observe that the TMR brightness temperatures increase with decreasing NRCS, but the TMR resolution of 35 km does not allow examination of the small scale variability in the atmospheric param-

eters. In order to better analyze the relationship between the NRCS and the rain events, we have collocated the T/P data in space and time with the 85 GHz brightness temperatures of the SSM/I for November 4, 1992, 2010 UTC (Plate 1). Estimation of the rain rate is not available because it requires an estimation of the wind speed and the SSM/I wind algorithms are no longer valid for high levels of liquid water content. Nevertheless effects of hydrometeors on the brightness temperatures at these frequencies result in a depolarization of the signal [Petty, 1994], and thus the vertically and horizontally polarized signals are used to compute an index of depolarization. After it has been normalized by the effects of the water vapor content and surface wind speed in the environment, it is defined as the normalized difference of the two channels. It can be used as an indicator of rain/no rain occurrence in the cells. These variables are displayed in Figure 11. The brightness temperatures decrease from a latitude of 14°N toward a minimum corresponding to the temperature at the top of the clouds in the eyewall region where the surface convergence and atmospheric water content are maximum. Then they increase in the vicinity of the TC center near 16°N and decrease again on the north side of the TC. Farther north (18°N) there is another minimum of temperature corresponding to an outer rainband and where the altimeter NRCS signal at Ku and C bands are obviously contaminated. The SSM/I index of depolarization is also displayed in Figure 11. It has two major maxima (signal totally depolarized) on both sides of the TC center, indicating the possible occurrence of heavy rains surrounded by large ice particles, and a third one in the rainband area. The index exhibits small scale variations elsewhere. To give more consistency to this analysis, we now relate this index to another rain index derived from the

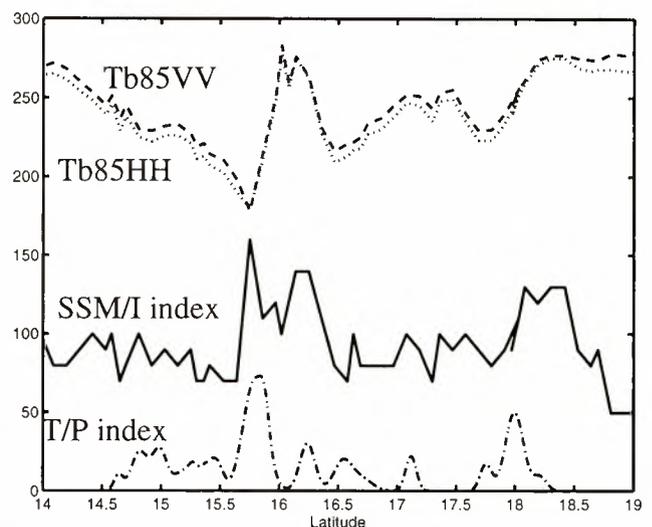


Figure 11. Collocated TOPEX-POSEIDON and SSM/I measurements. Brightness temperature at 85 GHz VV-pol (dashed) and HH-pol (dotted) in degrees Kelvin, SSM/I scaled polarization difference (solid), and TOPEX-POSEIDON rainfall index (dashed-dotted).

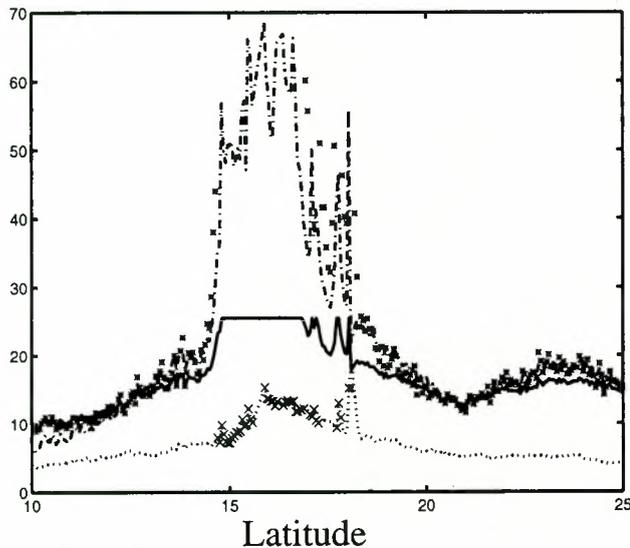


Figure 12. Wind speed (m/s) estimation from the single-frequency algorithms of *Witter and Chelton* [1991] (solid) and of *Young* [1993] (dashed-dotted) and the dual-frequency algorithm of *Elfouhaily* [1996] (asterisks). The Ku band significant wave height (SWH in meters) is also given for reference (dotted). The cross characters are data points for which *Elfouhaily's* algorithm did not converge. Between the latitudes of 15° and 18°N, convergent and divergent points are interlaced, which seems to indicate that rain cells induce differences from the Ku/C band relationship for no-rain conditions and therefore these situations result in numerically non converging calculations for the dual-frequency wind algorithm.

analysis of T/P echo waveforms and also displayed in Figure 11. This latter index is based on the Ku band NRCS departures from the mean Ku band/C band relationship [Chapron *et al.*, 1995]. *Tournadre and Morland* [1997] have thoroughly examined many cases of T/P altimeter measurements in the presence of rain and have concluded that the Ku band NRCS are lowered in nearly all cases when the C band signal exhibits a lowering or an increase only in a few cases. Thus they have defined a rain index, to be used as a rain flag, that is related to the departure of the Ku band signal from the mean Ku band/C band relationship, the later having been defined for no rain cases. As displayed in Figure 11, this index indicates a variability that is qualitatively well related to the SSM/I index, even if errors in the collocation, the sensor resolutions and sampling, and if geophysical effects induce differences in the exact location of the maxima that are seen by the two indices. It appears that there is a close relationship between the small scale drops in the NRCS signal at Ku band and the peaks of the SSM/I depolarization index, showing that the Ku band NRCS are strongly contaminated by the atmospheric particles.

We then note that the C band signal is not so much contaminated by atmospheric particles, unlike the Ku band sig-

nal, and that it is possible to retrieve the wind structures with a relatively good accuracy even in these extreme events. This is also illustrated by the examples of the retrieved wind fields shown in Plate 2. They are quite consistent, and no obvious sign of contamination is visible.

4.2.2.2. A high-wind branch for the C band model: As outlined in the previous sections, the C band model sensitivity is too low to display the high wind variability. This behavior was also observed on the altimeter winds. In most of the cyclone cases, the wind exceeds the upper limit of wind speed (20 m/s) imposed by altimeter inversion algorithms whose calibration is based upon buoy measurements. On the basis of parametric model predictions [Holland, 1980] of the surface winds in TCs, *Young* [1993] developed a hurricane algorithm for the Geosat altimeter (Ku band) whose validity ranges from 20 to 40 m/s. This algorithm is only applicable to altimeter measurements that are not highly attenuated by the presence of hydrometeors in the atmosphere. Unfortunately, hurricane winds are highly correlated with precipitation in tropical cyclones. Therefore use of *Young's* algorithm to estimate high winds must be restricted to cases where precipitation does not significantly affect the altimeter measurements. As was already shown by *Tournadre and Morland* [1997] the Ku/C band relationship for T/P is a good indicator of rain in tropical cyclones. In addition, a recent study conducted by *Elfouhaily* [1996] demonstrated the ability to estimate a sea state dependent local wind speed using the same Ku/C band relationship. The later was called local wind speed because of its dependence on the local sea state. Figure 12 shows that the local wind speed is comparable to the one from *Witter and Chelton* [1991] for winds lower than 15 m/s and with winds from *Young* [1993] otherwise. For latitudes between 15°N and 18°N the dual frequency based algorithm features the ability to discriminate between measures without significant atmospheric contamination and measures affected by a strong attenuation of the Ku band NRCS. As hypothesized, this attenuation is always attributed to rain cells present during the altimeter pass.

The inclusion of C band estimation in the dual-frequency algorithm thus helped both the estimation of the wind and the detection of an excessive attenuation of the Ku band in the presence of rain.

Direct application of altimeter findings to scatterometers for high-wind estimations is not a straightforward manipulation mainly because of the differences in range of incidence angles at which these instruments operate. However, when one looks more closely to the scattering mechanism involved in both configurations, possible transition may be envisaged. Indeed, differences between C and Ku band altimeter measurements may be associated with different cut-off wavelengths that separate long and short scales on the surface: the real surface appears smoother to larger microwave wavelength [Chapron *et al.*, 1995]. For C and Ku band radars, well-accepted cut off values are about 3 times the incident electromagnetic wavelength [Brown, 1990], 16 cm and 6 cm, respectively. Therefore differences between C and Ku band measurements are then mainly dominated by small scale slope and height variances covering the

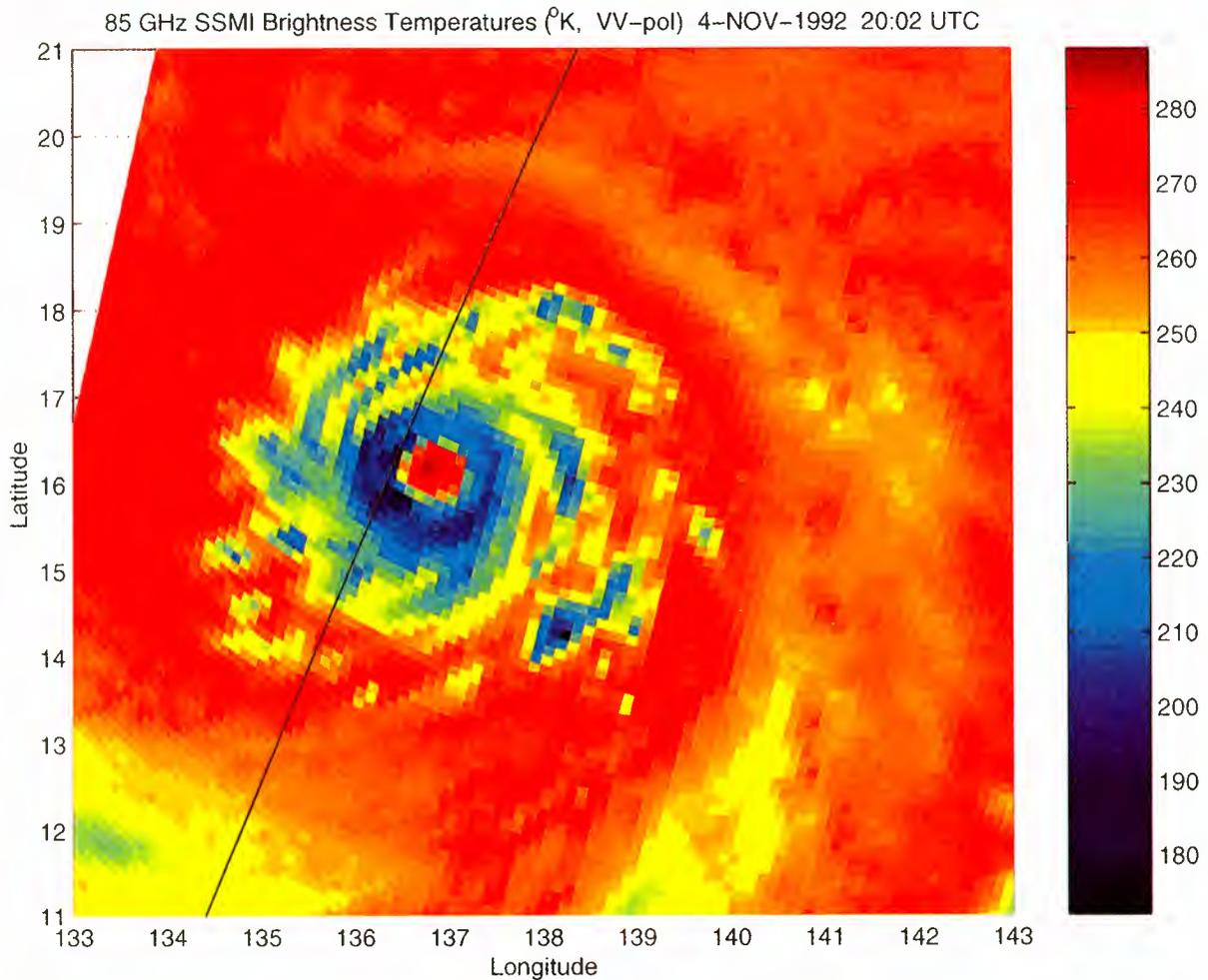


Plate 1. SSM/I 85 GHz brightness temperatures in degrees Kelvin (November 4, 1992, 2002 UTC), and the TOPEX-POSEIDON altimeter ground track (November 4, 1992, 1730 UTC). The altimeter positions have been shifted in order to take the time differences into account.

range of centimetric waves (from 6 to about 16 cm). Such a range is also preponderant for C band off nadir scatterometer measurements. We can then envisage to construct a high-wind branch for the C band scatterometers based on the high-wind branch of the altimeter algorithm. The possible methodology would be to collocate ERS scatterometer measurements with dual-frequency altimeter ones close to hurricane winds (after elimination of anomalous Ku band measurements).

5. Examples of Surface Wind Patterns in Tropical Cyclones

The knowledge of surface wind patterns in TCs is of great importance to understanding and forecasting the formation and the evolution of cyclones. If the physics of mature cyclones are rather well known, there remain many uncertainties regarding how TCs develop and what the atmospheric conditions are that enable or inhibit intensification of a tropical disturbance. Currently, there is a dramatic lack of three-dimensional measurements of most of the parameters of interest. Observations by scatterometers pro-

vide a unique view of the surface wind patterns in the outer core of TCs with a good resolution and of the surface wind of the environmental flow embedding the TC. We intend in this section to discuss the features delineated by the scatterometer in a few cases, in order to assess this new data source and its usefulness.

Plate 2 displays the HR surface winds in TC Elsie for four ERS-1 passes where the TC core was identified. Elsie's track determined by JTWC and location of its center determined with ERS-1 scatterometer are shown in Figure 7. Elsie was moving west-northwest away from Guam on November 3, 1992, while intensifying until November 5, when it reached its peak intensity of 150 knots (75 m/s). It turned north-northeast later that day and accelerated on November 6 while weakening. Then it turned east-northeast and became a typhoon-force extratropical low on November 7. During this period the region of maximum winds in the outer radius grew from about 100 km to several hundred of kilometers. One striking feature is strong asymmetries in the surface winds as displayed in Plate 2. Asymmetries in TCs have been previously documented [Holland, 1980; Anthes, 1982; Quilfen et al., 1994; Franklin et al, 1993]. They

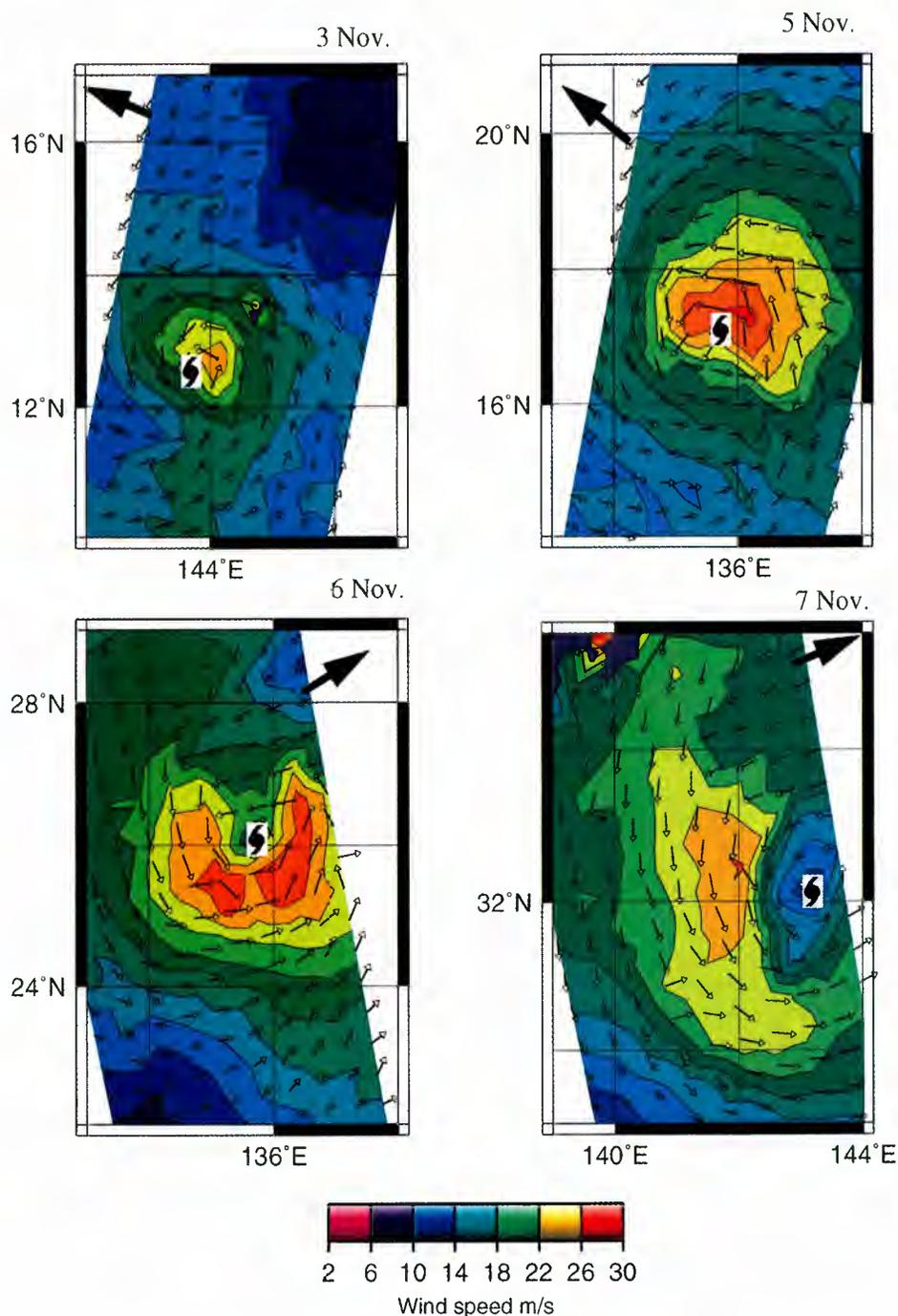


Plate 2. Surface wind field from the ERS-1 scatterometer at a resolution of 25 km in tropical cyclone Elsie: (a) November 3, (b) November 5, (c) November 6, and (d) November 7, 1992. Location of the TC center determined from the scatterometer measurements and the TC direction are indicated.

are related to the movement of the axisymmetric vortex in the environmental flow but also to many other mechanisms. They partly organize the distribution of convection and are a characteristic feature of the tropical cyclone evolution. The maximum winds are on the right side of the typhoon moving northwestward on November 3 and 5 and also on the right side of Elsie as it moved northeastward on November 6. The position of the maximum winds reversed after Elsie's recurvature. This distribution could be due to more prevailing easterlies in the south part and more prevailing

westerlies in the north part. This characteristic has been observed in many of the TCs sampled by the ERS-1 scatterometer, and is in agreement with previous studies. *Franklin et al.* [1993] outlined a similar distribution of surface winds by using nested analyses of Hurricane Gloria in the Atlantic Ocean.

Figure 13 displays the surface wind divergence fields for a few cases of mature or intensifying TCs. It features characteristic patterns of the wind divergence that are nearly the same in all cases. Strong convergence occurs in the south

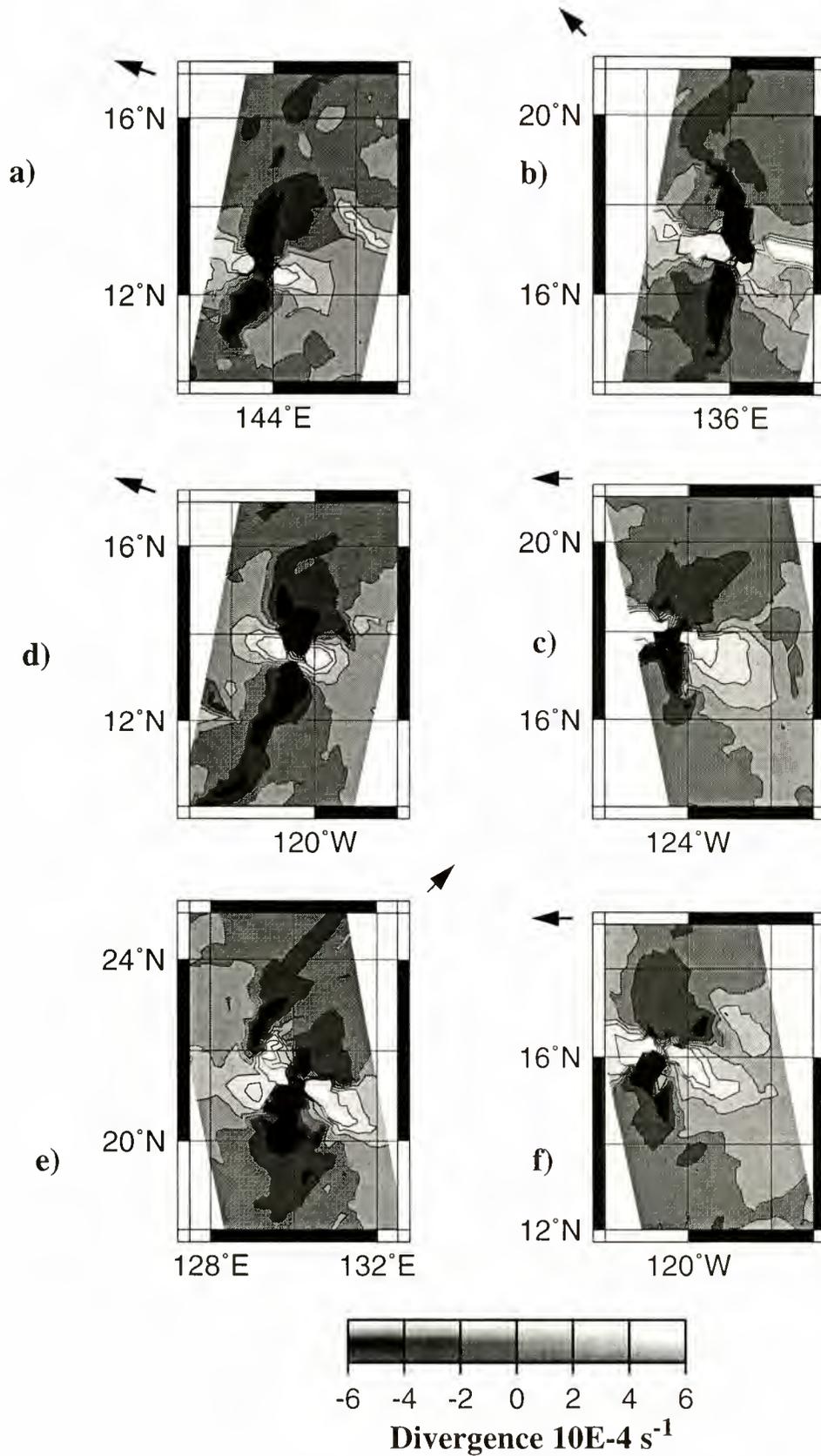


Figure 13. Surface wind divergence in (a) TC Elsie, November 3, 1992; (b) TC Elsie, November 5, 1992; (c) TC Tina, October 2, 1992; (d) TC Orlene, September 5, 1992; (e) TC Yvette, October 14, 1992; and (f) TC Frank, July 18, 1992. The arrows indicate the direction of the TC track.

and north parts of the TCs and the flow is divergent in the other quadrants. *Anthes* [1982] noted that surface wind convergence occurs in the vicinity of rainbands with values exceeding $4 \times 10^{-4} \text{ s}^{-1}$. It is also the order of magnitude of the ERS-1 wind divergence. Nevertheless, it was shown by *Ebuchi* [1994] that the scatterometer antenna geometry, imperfections in the backscattering models, and processing algorithms could produce wind direction trapping in favored directions aligned with the antennae. To verify whether or not these divergence structures are artifacts, we have simulated the NRCS as observed by the ERS scatterometer in a perfectly symmetric vortex. We have then retrieved the wind field and computed its divergence. Patterns similar to those shown in Figure 13 also occur, showing that the retrieved features are partly artifacts related to the antenna geometry. It is thus hardly possible to delineate the true geophysical divergence field from these artifacts.

6. Conclusion

The wind fields derived from 29 orbits of ERS-1 scatterometer measurements at high (25 km) and low (50 km) resolutions have been analyzed for a few tropical cyclone events. Increasing the resolution enables to greatly reduce the geophysical noise in the measurement cells, and thus to improve the NRCS estimates. Such an enhancement for the measurement definition is due to the high instrument signal to noise ratio: the radiometric resolution does not dominate the noise budget.

Examining the surface wind structures in tropical cyclones, we show that usable ERS scatterometer data are obtained, even through deep convective clouds. Three main shortcomings of scatterometry for remote sensing of such events were identified as follows:

1. The measurement resolution as discussed above is only 50 km for the operational ERS scatterometer activities. It is too coarse to measure the wind structures that vary typically over a few kms; NRCS measurements are dramatically smoothed in the vicinity of the TC centers. The knowledge of these small scale structures is of great interest for TC forecasting and physical modeling. It is shown that the higher resolution of 25 km providing measurements every 12.5 km retains some of the main characteristics of the surface wind patterns such as the location of the TC center with its area of low winds, the location of the outer core maximum winds and its radius to be featured. A 25 km (or even better) resolution is critical for measuring the surface wind structures with sufficient accuracy to help to understand the TC mechanisms and to improve the estimates of wind radii and storm surge forecasting.

2. Heavy rain is likely to interact with a microwave signal. The C band signal is theoretically only weakly affected by rain through its path in the atmosphere, but little is known about effects of heavy rains (both on the sea surface roughness and on the C band signal itself) that are permanent features of TCs. To provide a qualitative insight on this topic, we have analyzed the backscattered signals of the T/P dual-frequency altimeter for one track crossing *Elsie's* core. The

NRCS at Ku band exhibits drops that are strongly linked to the variability of a rain index derived from the SSM/I high-frequency brightness temperatures. At the same time, the C band signal does not exhibit such a variability at this scale, and it is thus confirmed that the C band NRCS are not too seriously contaminated by rain. Moreover, this result supports the usefulness of a rain index, derived from the T/P altimeter, based upon the deviation of the Ku signal from a mean C band/Ku band relationship.

3. Very little is known about the behavior of the scatterometer signal for very high winds, and the C band backscattering model is less valid for winds above 20m/s. Estimation of these winds by using the operational backscattering models gives very crude estimates of the mean wind level even if such estimates provide some information on the wind variability. It is clear that NRCS for high winds are overestimated by the existing models, leading to an underestimation of the wind speeds. An examination of the NRCS signals in TCs shows that the mean backscattered signal does not appear to saturate and exhibit sufficient azimuthal modulation in the most intense tropical cyclones, with a superior sensitivity for the highest resolution. Previous work developing altimeter wind algorithm for higher winds are illustrated for a T/P track analyzed in this paper. We propose that an analogous approach may be used to define a high-wind branch for the scatterometer C band model.

With these limitations in mind, the surface winds are analyzed in TC *Elsie*. The most prominent features are the asymmetries existing in the wind speed field and the well-defined patterns in the divergence field. The location of the wind maximum appears to occur mainly on the right side of the moving storm, a feature that was previously identified by different authors on the basis of observational and theoretical work. This is hypothesized to be mainly due, among other mechanisms, to the displacement of the storm in an environmental flow with a prevailing surface wind direction. Another characteristic feature is the distribution of the wind divergence, strong convergence patterns being identified on the south and north sides of most of the studied TC cases. However, a simulated test shows that the convergence patterns delineated by the scatterometer are partly artifacts of the antenna geometry, thus limiting their interpretation.

To conclude, the scatterometer data source of surface winds is very promising but much work is still to be done in order to interpret scatterometer signals at C band and Ku band. Such studies will help to better understand the characteristic wind patterns of tropical cyclones as measured by scatterometry. The forthcoming operational scatterometers on-board the ADEOS (Ku band) and METOP (C band) satellites will provide at least twice as many complementary measurements with a better resolution than the ERS operational one and should contribute strongly in the near future to better understanding and forecasting of tropical cyclones. Continued improvement of the interpretation of the scatterometer signals in tropical cyclones is therefore of great importance, especially because these data are already used by forecasters.

Appendix

A1. The Backscattering Model CMOD_IFR2

The backscattering model CMOD_IFR2 is formulated as follows:

$$\sigma_0 = 10^{\alpha + \beta\sqrt{V}} \times (1 + b_1 \times \cos\Phi + \tanh b_2 \times \cos 2\Phi)$$

(A1)

where tanh is the hyperbolic tangent, with

$$\alpha = c_1 + c_2 P_1 + c_3 P_2 + c_4 P_3$$

$$\beta = c_5 + c_6 P_1 + c_7 P_2$$

where

$$P_1 = x \quad P_2 = (3x^2 - 1)/2 \quad P_3 = x(5x^2 - 3)/2$$

$$x = (\theta - 36)/19$$

and

$$b_1 = c_8 + c_9 V_1 + c_{10} P_1 + c_{11} P_1 V_1 + c_{12} P_2 + c_{13} P_2 V_1$$

$$b_2 = c_{14} + c_{15} P_1 + c_{16} P_2 + (c_{17} + c_{18} P_1 + c_{19} P_2) V_1$$

$$+ (c_{20} + c_{21} P_1 + c_{22} P_2) V_2 + (c_{23} + c_{24} P_1 + c_{25} P_2) V_3$$

where

$$P_1 = x \quad P_2 = 2x^2 - 1 \quad x = (2\theta - 76)/40$$

$$V_1 = (2V - 28)/22 \quad V_2 = 2V_1^2 - 1 \quad V_3 = (2V_2 - 1)V_1$$

σ_0 is the normalized radar cross section, V is the wind speed in meters per second, ϕ the wind direction in degrees relative to the antenna, and θ the incidence angle in degrees. Parameters values and test values are given in Tables A1 and A2.

A2. High Wind Bias Correction

After the wind speed V_m has been computed from the C band model, the following bias correction is added:

Table A1. Parameter Values

Parameter	Value	Parameter	Value
c_1	-2.437597	c_{14}	0.412754
c_2	-1.567031	c_{15}	0.121785
c_3	0.370824	c_{16}	-0.024333
c_4	-0.040590	c_{17}	0.072163
c_5	0.404678	c_{18}	-0.062954
c_6	0.188397	c_{19}	0.015958
c_7	-0.027262	c_{20}	-0.069514
c_8	0.064650	c_{21}	-0.062945
c_9	0.054500	c_{22}	0.035538
c_{10}	0.086350	c_{23}	0.023049
c_{11}	0.055100	c_{24}	0.074654
c_{12}	-0.058450	c_{25}	-0.014713
c_{13}	-0.096100		

Table A2. Test Values

V , m/s	Φ , deg	θ , deg	σ_0 , dB
1	0	60	-31.84
1	90	60	-32.81
8	0	40	-14.45
8	90	40	-19.02
15	0	25	-2.57
15	180	25	-2.71
22	0	18	4.38
22	180	18	5.32
28	0	18	4.74
28	180	18	6.19

$$V_{scat} = V_m + V_{bias} \tag{A2}$$

with

$$V_{bias} = 0 \quad \text{if } V \leq 10 \text{ m/s}$$

$$V_{bias} = 0.0831 \times V_m - 0.0173 \times V_m^2 + 0.0009 \times V_m^3 \quad \text{if } 10 < V \leq 22 \text{ m/s}$$

$$V_{bias} = \arctan(V_m - 22) + 3.0382 \quad \text{if } V > 22 \text{ m/s}$$

The function arctan designs the arctangent.

The bias was evaluated by comparison of the C band model winds with the NOAA buoys winds for the time period March 1992 to October 1993.

Acknowledgments. The scatterometer orbits were processed at the Centre National d'Etudes Spatiales. We thank Jean-Pierre Mardardé for providing the material necessary to use these data, Chip Guard for providing the JTWC data, Abderrahim Bentamy, Peter Black, and Ad Stoffelen for their useful comments. The SSM/I data were provided by Frank Wentz of Remote Sensing Systems Inc.

References

Anthes, R.A., *Tropical Cyclones, Their Evolution, Structure and Effects, Meteorol. Monogr. Ser.*, Vol. 19, No. 41, 208 pp, Am. Meteorol. Soc., Boston, Mass., 1982.

Archiving, Validation and Interpretation of Satellite data in Oceanography (AVISO), *AVISO User Handbook: Merged TOPEX/POSEIDON Products, Rep. AVI-NT-02-101-CN*, edition 2.1, 208 pp, Centre National d'Etudes Spatiales, Toulouse, France, 1992.

Bentamy, A., Y. Quilfen, P. Queuiffelou, and A. Cavanie, Calibration of the ERS-1 scatterometer C-band model, *Tech. Rep. IF-REMER DRO/OS-94-01*, 72 pp, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Plouzané, France, 1994.

Brown, G.S., Quasi-specular scattering from the air-sea interface, in *Surface Waves and Fluxes*, vol. 2, edited by G.L. Geernaert and W.J. Plant, pp 1-39, Kluwer Acad., Norwell, Mass., 1990.

Chapron, B., K. Katsaros, T. Elfouhaily, and D. Vandemark, A note on relationships between sea surface roughness and altimeter backscatter, in *Air-Water Gas Transfer*, pp. 869-878, Heidelberg Univ., Heidelberg, Germany, 1995.

- Dvorak, V.F., Tropical cyclone intensity analysis and forecasting from satellite imagery, *Mon. Weather Rev.*, 103, 420-430, 1976.
- Ebuchi, N., Directional distribution of wind vectors observed by ERS-1 scatterometer, in *Proceedings 2nd ERS-1 workshop*, *Tech. Rep. IFREMER DRO/OS/94-06*, pp 89-92, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Plouzané, France, 1994.
- Elfouhaily, T., Modèle couplé vent/vagues et son application à la télédétection par micro-onde de la surface de la mer, 229 pp., Thèse de Doctorat, Univ. Pierre et Marie Curie, Paris, 1996.
- Elfouhaily, T., B. Chapron, K. Katsaros, and D. Vandemark, A unified directional spectrum for long and short wind-driven waves, *J. Geophys. Res.*, 102, 15781-15796, 1997.
- Emanuel, K. A., The maximum intensity of hurricanes. *J. Atmos. Sci.*, 45, 1143-1155, 1988.
- Fischer, R. E., Standard deviation of scatterometer measurements from space, *Proc. IEEE*, 10, 106-113, 1972.
- Franklin, J.L., S.J. Lord, S.E. Feuer, and F.D. Marks, The kinematic structure of Hurricane Gloria (1985) determined from nested analyses of dropwindsonde and Doppler radar data, *Mon. Weather Rev.*, 121, 2433-2451, 1993.
- Freilich, M. H., and R. S. Dunbar, A preliminary C-band scatterometer model function for the ERS-1 AMI instrument, in *Proceedings of First ERS-1 Symposium: Space at the Service of Our Environment*, *Eur. Space Agency Spec. Publ., ESA SP-359*, pp. 79-83, March 1993.
- Graber, H. C., N. Ebuchi, and R. Vakkayil, Evaluation of ERS-1 scatterometer winds with wind and wave ocean buoy observations, *Tech. Rep. RSMAS 96-003*, 69 pp, Rosentiel School of Mar. and Atmos. Sci., Univ. of Fla., Miami, 1996.
- Gray, W.M., Tropical cyclone formation and intensity change, paper presented at ICSU/WMO International Symposium on Tropical Cyclone Disasters, Int. Counc. of Sci. Unions, Beijing, 1993.
- Holland, G.J., An analytical model of the wind and pressure profiles in hurricanes, *Mon. Weather Rev.*, 108, 1212-1218, 1980.
- Jones, W.L., L.C. Schroeder, D.H. Boggs, E.M. Bracalente, R.A. Brown, G.J. Dome, W.J. Pierson, and F.J. Wentz, The Seasat-A satellite scatterometer: The geophysical evaluation of remotely sensed wind vectors over the ocean. *J. Geophys. Res.*, 87, 3297-3317, 1982.
- Large, W.G., J. Morzel, and G.B. Crawford, Accounting for surface wave distortion of the marine wind profile in low-level ocean storms wind measurements, *J. Phys. Oceanogr.*, 25, 2959-2971, 1995.
- Long, A. E., C-band V-polarized radar sea-echo model from ERS-1 Haltenbanken campaign, in *Proceedings of ERS-1 Geophysical Validation*, *Publ. WPP-36*, edited by M. Wooding, pp. 113-118, Eur. Space Agency, Paris, 1992.
- Malardé, J. P., Traitement géométrique et radiométrique du radar diffusiomètre de ERS-1, Thèse de Doctorat, Univ. Paris 7, Paris, 1992.
- Petty, G.W., Physical retrievals of over-ocean rain rate from multi-channel microwave imagery, 2; Algorithm implementation, *Meteorol. Atmos. Phys.*, 54, 101-121, 1994.
- Petty, G.W., and K.B. Katsaros, New geophysical algorithms for the special sensor microwave/imager, paper presented at Fifth International Conference on Satellite Meteorology and Oceanography, London, 1990.
- Plant, W.J., A two-scale model of short wind-generated waves and scatterometry, *J. Geophys. Res.*, 91, 10,735-10,749, 1986.
- Quilfen, Y., and A. Bentamy, Calibration/validation of ERS-1 scatterometer precision products, in *Proceedings of IG-ARSS'94*, pp. 945-947, IEEE Press, Piscataway, N.J., 1994.
- Quilfen, Y., K.B. Katsaros, and B. Chapron, Structures des champs de vent de surface dans les cyclones tropicaux observés par le diffusiomètre d'ERS-1 et relation avec les vagues et les précipitations, *Mem. Inst. Océanogr., Monaco*, 18, 13-22, 1994.
- Stoffelen, A., and D. Anderson, Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4, *J. Geophys. Res.*, 102, 5767-5780, 1997.
- Tournadre, J., and J.C. Morland, The effect of rain on TOPEX/POSEIDON altimeter data: A new rain flag based on Ku and C band backscatter coefficient, *IEEE Trans. Geosci. Remote Sens.*, in press, 1997.
- Witter, D.L., and D.B. Chelton, A Geosat altimeter wind speed algorithm and a method for altimeter wind speed algorithm development, *J. Geophys. Res.*, 96, 8853-8860, 1991.
- Young I.R., An estimate of the Geosat altimeter wind speed algorithm at high wind speeds, *J. Geophys. Res.*, 98, 20,275-20,285, 1993.
- B. Chapron, T. Elfouhaily, Y. Quilfen, and J. Tournadre, Laboratoire d'Océanographie Spatiale, IFREMER, BP70, 29280, Plouzané, France. (email: yquilfen@ifremer.fr)
- K. Katsaros, Atlantic Oceanographic and Meteorological Laboratory, NOAA, 4301 Rickenbacher Causeway, Miami, FL 33149. (email: katsaros@aoml.noaa.gov)

(Received December 30, 1996; revised June 9, 1997; accepted July 3, 1997.)

Directivity of wind vectors derived from the ERS-1/AMI scatterometer

Naoto Ebuchi

Center for Atmospheric and Oceanic Studies, Faculty of Science, Tohoku University, Sendai, Japan

Hans C. Graber

Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida

Abstract. Directional distribution of the wind vectors observed by the ERS-1 active microwave instrument (AMI) scatterometer over the global oceans is investigated to examine effects of errors in the geophysical model function used to retrieve the wind vectors from backscattering measurements. Three scatterometer data products, together with wind data derived from two operational meteorological analysis models, are analyzed and compared with each other for a period of 2 months. Frequency distributions of the wind directions relative to the midbeam looking direction are calculated in bins of 10° . It is found that scatterometer-derived wind vectors may exhibit systematic directional preference relative to the antenna beams in comparison with those of the model winds. This artificial directivity is considered to be caused by imperfections in the model function used to retrieve the wind vectors. Dependencies of the directional distributions on incidence angle and wind speed are also discussed for each data products. The distribution of wind directions is very sensitive to the errors in the model function and can be used as an effective measure of self-consistency to validate the model function and retrieved wind vectors.

1. Introduction

Satellite-borne scatterometers can measure the surface wind vectors with global coverage, high spatial resolution, and frequent sampling. The observed wind fields over the global ocean are of critical importance for various studies in fields of oceanography, meteorology, and climatology. However, the scatterometers do not directly measure the wind vectors but measure the radar signal backscattered from the sea surface. The surface wind vectors are obtained from the observed normalized radar cross sections (σ_0 values) by using a wind retrieval algorithm with a geophysical model function, which describes the relationship between the σ_0 and the incidence angle, wind speed, and azimuth angle of wind direction relative to the radar antenna beams. Validation of the model function and retrieved wind vectors is an important step in determining the quality of remotely sensed wind products and assessing the error structure of vector winds.

Numerous validation studies have been carried out by comparing the observed σ_0 values and retrieved wind vectors with in situ wind measurements by buoys and ships [e.g., Jones *et al.*, 1982; Bentamy *et al.*, 1994; Quilfen and Bentamy, 1994; Ebuchi *et al.*, 1996; Graber *et al.*, 1996]. In this method it is sometimes difficult to obtain a large number of collocated data, especially at initial stages after the satellite launch (enough data to assess the three-dimensional dependence of the σ_0 values on the incidence angle, wind speed, and direction), because reliable wind observations are limited in number.

Approaches to validate scatterometer winds using outputs

from numerical weather forecast and analysis models [e.g., Anderson *et al.*, 1991; Stoffelen and Anderson, 1995; Freilich and Dunbar, 1993; Offiler, 1994; Freilich *et al.*, 1997] have also been employed. This method totally depends on the accuracy of the numerical weather models. The differences of the spatial and temporal resolution between the scatterometer observations and the numerical models may also affect results of the validation.

A different validation approach without reference wind data was tried to find systematic errors in the scatterometer winds. By analyzing the wind speeds derived from the Seasat A satellite scatterometer (SASS) using the SASS 1 model function, Woiceshyn *et al.* [1986] found that the model function for the vertical and horizontal polarizations are inconsistent. This type of validation examining systematic errors and self-consistency of the retrieved wind vectors may also be effective to refine the retrieval algorithms and model functions, together with comparisons of in situ wind measurements and model outputs.

In the present study, directional distribution relative to the antenna beam direction of wind vectors observed by the ERS-1/active microwave instrument (AMI) scatterometer over the global oceans is examined as a measure of the self-consistency. If the scatterometer correctly measures the wind fields, the derived distribution of wind directions is expected to represent the nature of the global wind field such as the trades and westerlies and should not have any systematic directivity relative to the antenna beam looking directions. In comparison of the directions of the SASS 1 winds with in situ measurements presented by Jones *et al.* [1982, Figure 13] the distribution of data was not uniform but clearly showed four gaps. Anderson *et al.* [1991] reported some directional irregularities in the SASS 1 winds.

The directional distributions of three ERS-1/AMI scat-

terometer wind data products are investigated in order to examine differences due to geophysical model functions used to retrieve the wind vectors in the present study. Wind fields obtained from two operational meteorological analysis models are also analyzed in the same way and compared with scatterometer winds. Dependencies on incidence angle and wind speed are also discussed.

2. ERS-1/AMI Scatterometer and Wind Data Products

The first European Remote Sensing Satellite (ERS-1) was launched by the European Space Agency (ESA) on July 17, 1991, into a 785-km polar orbit at an inclination angle of 98.52°. The mission carried the active microwave instrument (AMI), along-track scanning radiometer/microwave sounder (ATSR/M), and radar altimeter (RA). The C band (5.3 GHz) active microwave instrumentation (AMI) comprises a side-looking system with three in-flight selectable modes for high-resolution imaging (image mode) as a synthetic aperture radar (SAR), measurement of directional wave spectra (wave mode), and wind scatterometer over the ocean surface (wind mode).

In the AMI wind mode (scatterometer) the sea surface is sequentially illuminated by C band (5.3 GHz) microwave pulses with vertical polarization, and the backscattered signal is measured from three different azimuth angles by three antennas looking 45° (forebeam), 90° (midbeam), and 135° (aft-beam) to the right of the spacecraft flight direction. The incidence angle of the radar beam varies from 24° to 57° for the fore and aft antenna beams and from 18° to 47° for the mid-beam. The three antennas form fan beams with a narrow azimuth pattern and a relatively wide elevation pattern in order to cover a wide swath of 500 km, starting 225 km from the subsatellite track. The 500-km-wide swath is sampled every 25 km, resulting in 19 measurement cells across the swath. The along-swath sampling interval is also equal to 25 km. The effective spatial resolution of the scatterometer is reduced, however, to ~50 km because adjacent measurement cells are not independent.

Surface wind vectors over the ocean are obtained from the observed normalized radar cross-section (σ_0) triplets using a wind retrieval algorithm with a geophysical model function. Several algorithms and model functions have been proposed for the ERS-1/AMI C band scatterometer and have been used to produce wind data products. Three data products such as the European Space Agency (ESA) fast delivery products, the Jet Propulsion Laboratory (JPL) value-added products, and the Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER) off-line products are utilized in the present study.

At the initial stage of data processing after the launch the ESA fast delivery products were processed using the pre-launched model function, CMOD-2 [Long, 1985]. After the calibration phase an intermediate transfer function, CMOD-3, was developed by tuning to European Centre for Medium-Range Weather Forecasts (ECMWF) analysis winds, using data for 1 month. Then the CMOD-4 model function was derived by Stoffelen and Anderson [1995] and has been used to generate the ESA fast delivery products since February 1993.

Freilich and Dunbar [1993] developed a preliminary fully empirical tabular model function, CMOD-FD, by collocating σ_0 values with surface wind data from operational global numerical weather prediction (NWP) analyses. The JPL value-

added products have been produced by the NASA/JPL using the JPL maximum likelihood estimating (MLE) wind retrieval algorithms with the CMOD-FD. Ambiguity removal has been achieved by a median filter technique with field initialization using the National Meteorological Center (NMC) wind analysis data.

The ERS-1 off-line data products have been generated at IFREMER. The first version of the wind products of 16 months' data was delivered in 1993 for test purposes. A model function called CMOD-IFR, which was developed by tuning coefficients for an assumed formula with 16 National Oceanic and Atmospheric Administration/National Data Buoy Center (NOAA/NDBC) buoy measurements and the Norwegian Institute of Meteorology (DNMI) wind analysis data, was used to retrieve wind vectors [Quilfen and Cavanie, 1991; Katsaros *et al.*, 1993; Bentamy, 1992; Bentamy *et al.*, 1994].

The second version was distributed in January 1995. The main differences are a new version of the C band model, CMOD-IFR2, and refinement of ice detection algorithm [Quilfen, 1995]. To tune the coefficients in the model function formulation, the NOAA/NDBC buoy data were used for the coefficients concerning the wind speed, and the ECMWF wind data were used for those concerning directions [Quilfen and Bentamy, 1994; Quilfen, 1995]. A bias correction, which was estimated by comparison with the buoy measurements, was applied to the model wind speed estimates at high winds. Ambiguity removal was only applied to wind vectors of wind speed higher than 3 m s^{-1} .

Ebuchi *et al.* [1996] and Graber *et al.* [1996] compared these three wind data products with the NOAA/NDBC, Tropical Ocean-Global Atmosphere/Tropical Atmosphere Ocean (TOGA-TAO) and Japan Meteorological Agency (JMA) buoy observations for a period of 2 years. It was shown that the IFREMER winds give the best agreement among the three data products. Systematic biases in wind speed were found for the JPL and ESA winds.

3. Directional Distribution of Wind Vectors

Three ERS-1/AMI scatterometer wind products such as the ESA fast delivery products, the JPL value-added products and the IFREMER off-line products, for a period of 61 days from March 1, 1993, to April 30, 1993, are analyzed in the present study. Only the wind vector cells between 45°S and 45°N are used to avoid contamination by sea ice. Number of the wind vector cells analyzed is about 8.7 million. Frequency distributions of the wind directions relative to the midbeam looking direction are calculated in bins of 10°.

In order to compare with the scatterometer winds, surface wind fields of the ECMWF analyses and the JMA global wind analysis (GANAL) data are also analyzed in the same way. The ECMWF wind data (6 hourly and 1.125° spatial resolution), which are collocated with the scatterometer observations by linear interpolation in time and space, have already been contained in the IFREMER data sets [Quilfen, 1995]. The JMA GANAL wind fields (12 hourly and 1.875° spatial resolution) are also linearly interpolated to the time and location of the scatterometer wind cells in the similar way.

Figure 1 shows polar plots of the frequency distributions of relative wind directions for the JPL value-added products (Figure 1a), the ESA fast delivery products (Figure 1b), the IFREMER off-line products (Figure 1c), the ECMWF winds (Figure 1d), and the JMA GANAL data (Figure 1e). Only

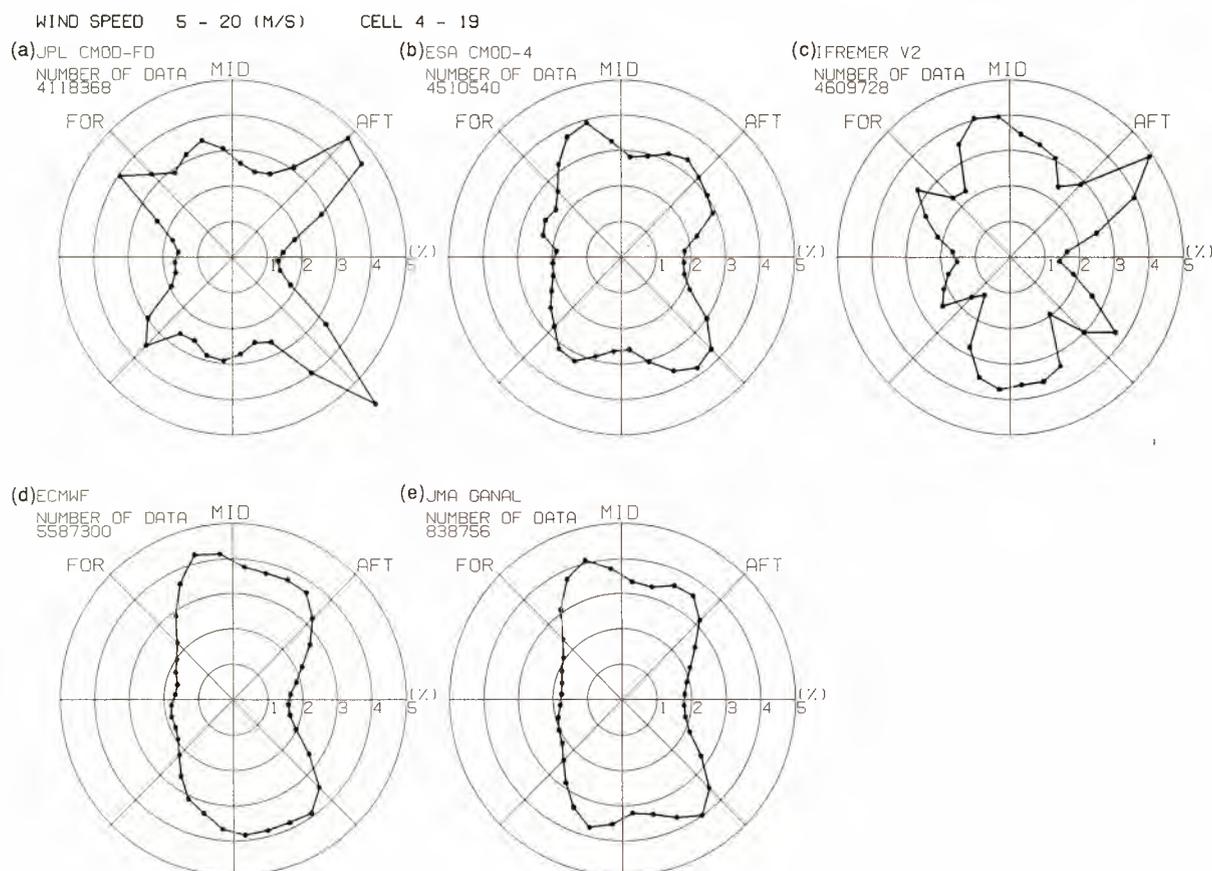


Figure 1. Frequency distributions of wind directions relative to the midbeam looking direction for (a) Jet Propulsion Laboratory (JPL) value-added products, (b) European Space Agency (ESA) fast delivery products, (c) Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER) off-line products, (d) European Centre for Medium-Range Weather Forecasts (ECMWF) winds and (e) Japan Meteorological Agency global wind analysis (JMA GANAL) data.

wind vector cells at middle to far ranges (cells 4–19, corresponding to incidence angles from 32° to 56° for the forebeam and aftbeam and 23° to 45° for the midbeam) within the swath and of wind speeds from 5 to 20 m s⁻¹ are used, because wind directions observed at the inner cells (i.e., low incidence angles) and low wind speeds are less accurate [Quilfen and Canvanie, 1991; Quilfen and Bentamy, 1994; Bentamy et al., 1994; Ebuchi et al., 1996; Graber et al., 1996].

The distributions for the two wind analyses, ECMWF and JMA GANAL, agree very well with each other. Also, the distribution of the ESA winds agrees well with those of the ECMWF and JMA wind analyses as should be expected. In contrast, the distributions of the JPL and IFREMER winds have artificial lobes relative to the antenna beam directions, though these two wind products show better agreement with in situ buoy measurement than the ESA fast delivery products [Ebuchi et al., 1996; Graber et al., 1996].

Similar frequency distributions of wind directions are calculated from the wind vectors retrieved by earlier versions of the model functions. In Figure 2 the distributions for wind vectors of the ESA prelaunch CMOD-2 model for a period of February 1992 (Figure 2a), the ESA CMOD-3 model for July 1992 (Figure 2b), and the first version of IFREMER off-line products for March and April 1993 (Figure 2d) are shown for comparison with the current versions in Figures 2c and 2e.

The directional distributions of the CMOD-2 and CMOD-3

winds are apparently distorted with reference to the antenna beams. There are very sharp gaps between lobes of distribution. These directivities are much reduced in the distribution for the CMOD-4 winds, reflecting improvement of the model function. The gaps in the wind direction distributions may strongly affect the retrieved wind fields. For a case where the wind direction is changing gradually, the scatterometer-derived wind field may be discontinuous since wind vectors in the direction of the gaps are not obtained.

It is surprising that the first version of the IFREMER wind products shows a better distribution than the second version, even though the second version has been much improved, especially in wind speed [Quilfen and Bentamy, 1994]. The refinement of the IFREMER model function was made mainly concerning the magnitude of wind velocity by turning to the NOAA/NDBC buoy measurements; especially, a bias correction was applied at high wind speed ranges. It is considered that these procedures have not improved (or made worse) the self-consistency of the retrieved wind vectors.

As shown in Figures 1 and 2, the scatterometer-derived wind vectors may show artificial directivity relative to the antenna beam directions. Anderson et al. [1991] reported similar directional irregularities in the Seasat A scatterometer winds retrieved by the SASS 1 model function. In the comparison of the SASS 1 wind directions with surface measurements by Jones et

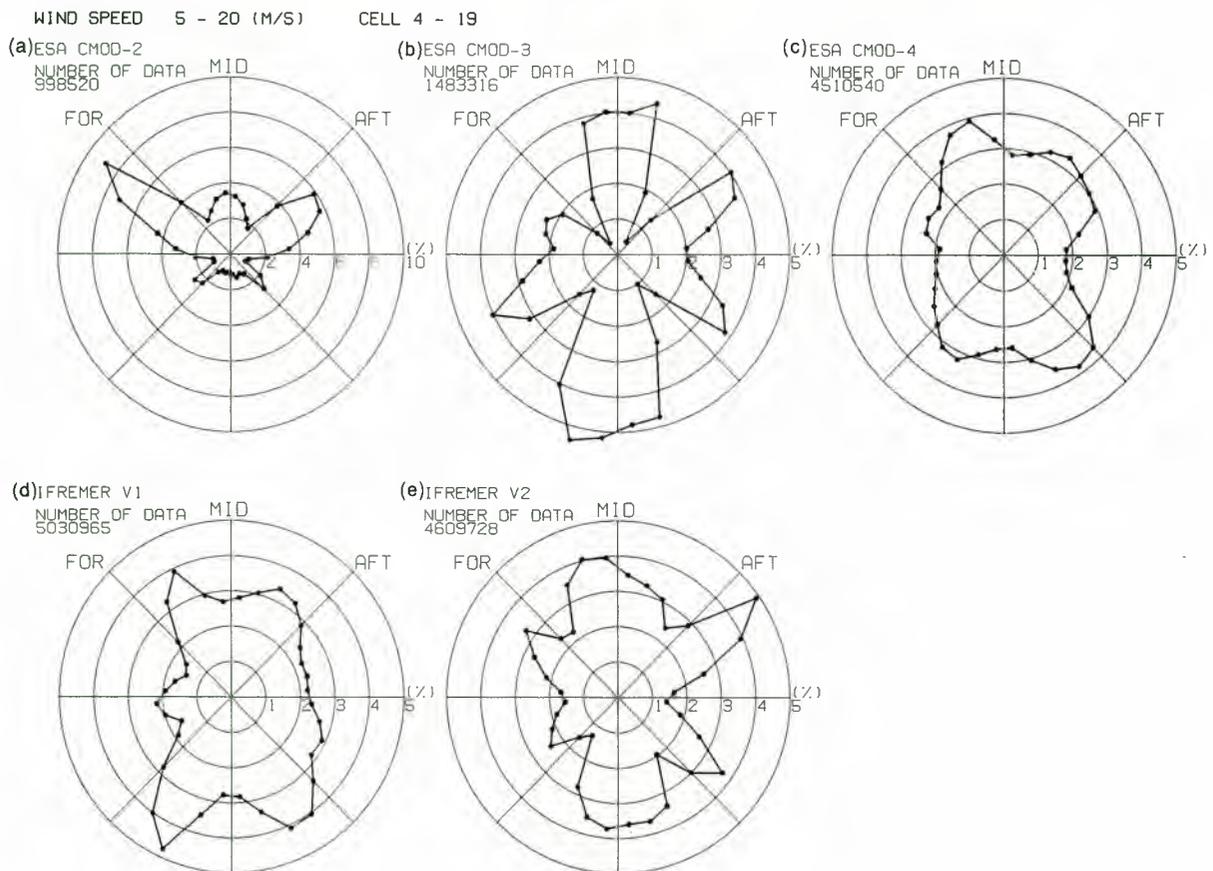


Figure 2. Directional distribution relative to the midbeam direction of wind vectors retrieved by earlier and current versions of the model functions. (a) The ESA CMOD-2 winds for February 1992, (b) the ESA CMOD-3 winds for July 1992, (c) the ESA CMOD-4 winds for March and April 1993 (same as Figure 1a), (d) the first version of IFREMER off-line products for March and April 1993, and (e) the second version of IFREMER off-line products for March and April 1993 (same as Figure 1c).

al. [1982, Figure 13] the distribution of data was not uniform but showed clearly four gaps.

In Figures 3 and 4, two examples for influence of the directivity on the observed wind fields are shown. As an extreme example, Figure 3 shows comparison of wind fields retrieved by the CMOD-2 and CMOD-FD. The wind field was observed on the west side of a cyclone which was located at about 29°N, 136°E. Wind was blowing counterclockwise around the cyclone. In the wind field retrieved by the CMOD-2, there exist several discontinuities of wind directions around 25°N. Most of wind vectors are aligned along one of antenna beams which have angles of 45°, 90°, and 135° with the satellite flight direction. This feature corresponds to what is expected from the result shown in Figure 2a. The discontinuity of wind directions is considered to be caused by the gaps of wind directions located between the antenna beam directions. In the wind field of CMOD-FD, wind direction changes continuously around the cyclone, and the cold front which extends across the swath can be recognized easily.

Figure 4 shows two wind fields retrieved using the CMOD-FD and CMOD-4. The wind field was observed on the east side of a cyclone located at about 41°N, 150°E. Though differences of the two wind fields are not very large, there exists discontinuity of wind direction around 34°N, 155°E in the CMOD-FD winds. In the region east of the discontinuity, wind vectors are aligned in the forebeam direction. In contrast, wind

direction changes continuously in this region for the CMOD-4 winds.

As shown in the examples in Figures 3 and 4, the directivity in the scatterometer-derived wind vectors may easily cause artificial discontinuity in the observed wind fields. This may largely affect the differential of the wind fields such as divergence/convergence field and wind stress curl, which are very important for meteorological and oceanographic studies. For example, *Chen et al.* [1996] considered numerical simulations on the path of the Kuroshio Current near Taiwan. They found that intrusions of the Kuroshio are sensitive and steered by the direction of the wind stress. *Yasuda and Hanawa* [1996] and *Özgökmen et al.* [1997] examined the impact of wind forcing on the separation of the western boundary currents such as the Gulf Stream and Kuroshio from the coast. These studies show that directional errors in the forcing (i.e., shift in the zero location of the wind stress curl) can have a dramatic effect on the response of ocean circulation. Further studies are needed to evaluate quantitative influence of the irregularity in the wind vectors on wind stress curl fields and ocean circulation. In a simple way the influence will be simulated by sampling the climatological wind stress field with the scatterometer geometry and then adding the systematic directivity as found in the present study.

According to validation studies comparing ERS-1 winds with buoy observations [*Bentamy et al.*, 1994; *Ebuchi et al.*, 1996;

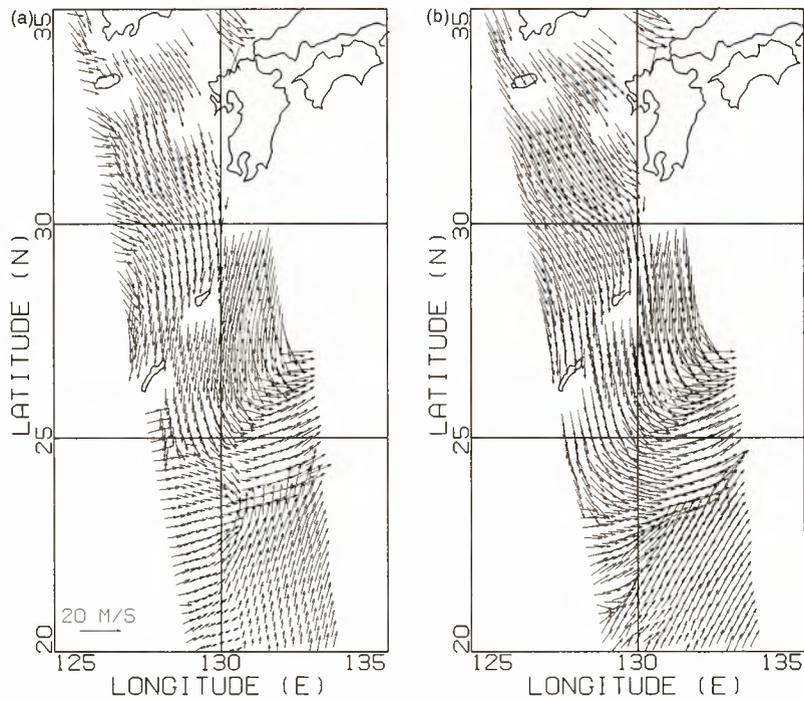


Figure 3. Difference of wind fields retrieved using (a) the CMOD-2 and (b) the CMOD-FD (1354 UT; February 15, 1992).

Graber *et al.*, 1996] the directions of wind vectors retrieved using CMOD-4, CMOD-FD, and CMOD-IFR2 agree well with buoy data with root-mean-square (rms) differences around 25°. No systematic errors in the scatterometer wind directions have been reported. In Figure 1, however, the dif-

ferences among the three wind products are significant. The directivity relative to the antenna beams is obvious in the distributions of the JPL and IFREMER winds. The directional distribution relative to the antenna beams investigated in the present study is considered to be sensitive to the difference of

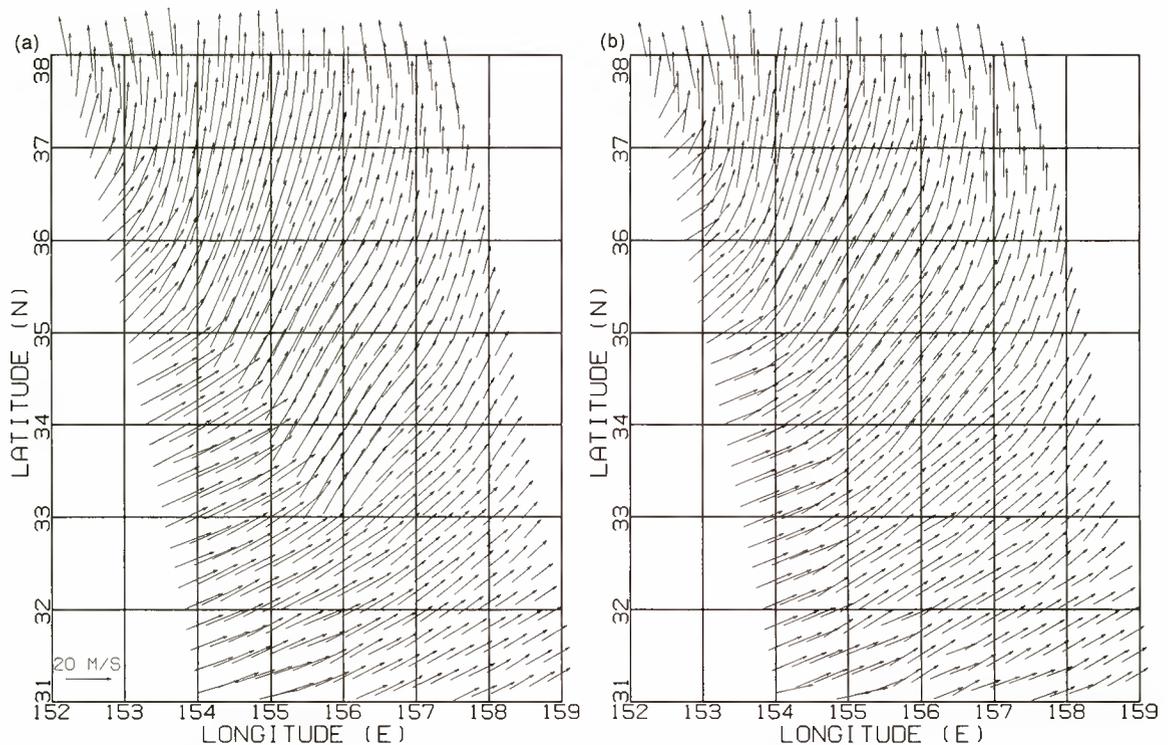


Figure 4. Difference of wind fields retrieved using (a) the CMOD-FD and (b) the CMOD-4 (1206 UT; March 12, 1993).

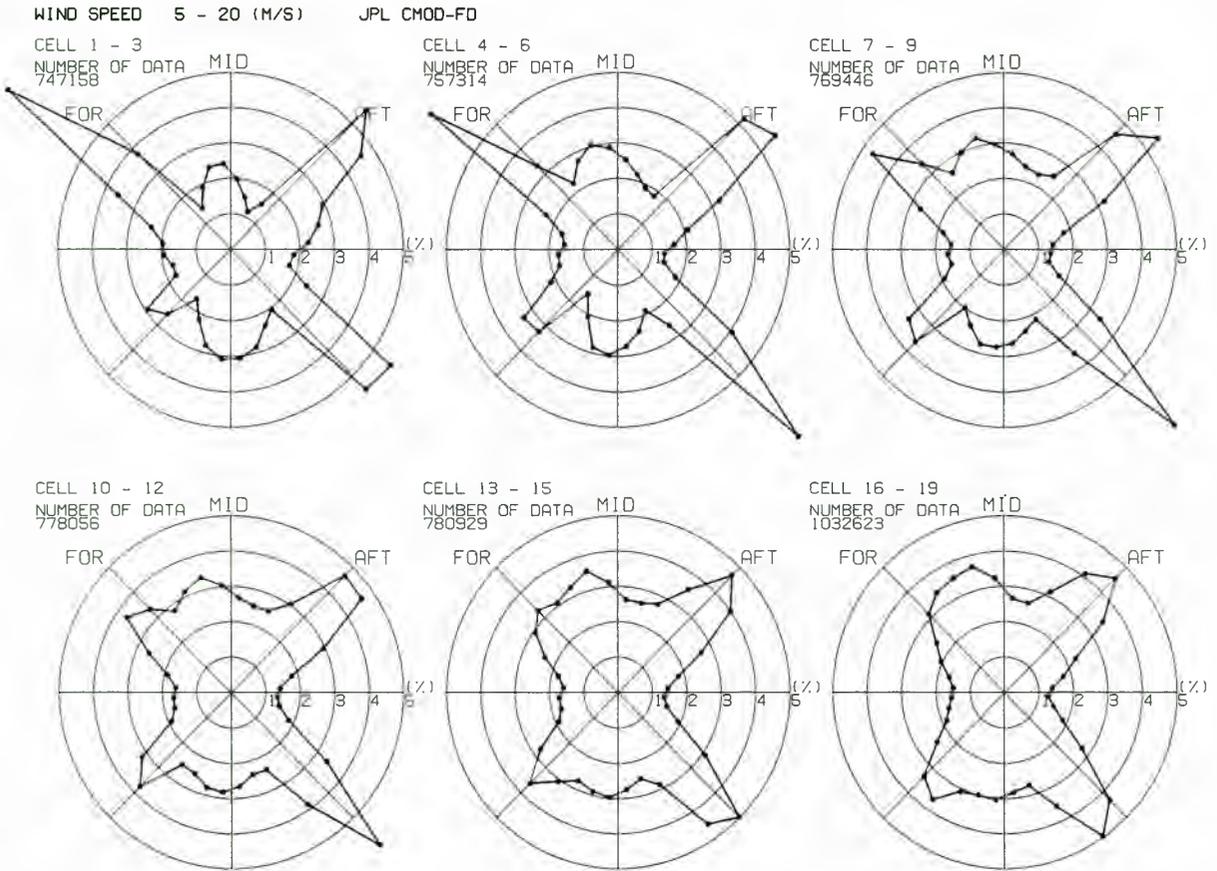


Figure 5. Dependence of the directional distributions on the location of wind vector cells for the JPL value-added products.

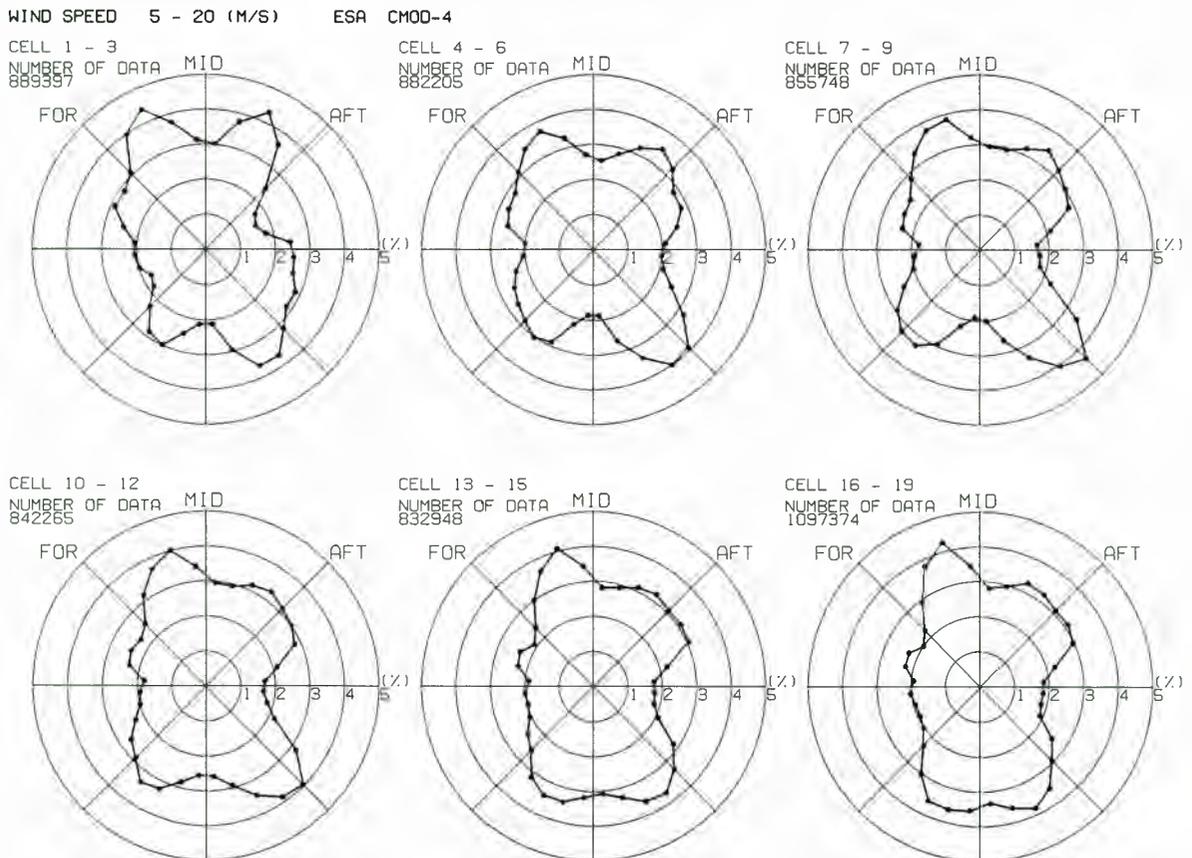


Figure 6. Same as Figure 5, except for the ESA fast delivery products.

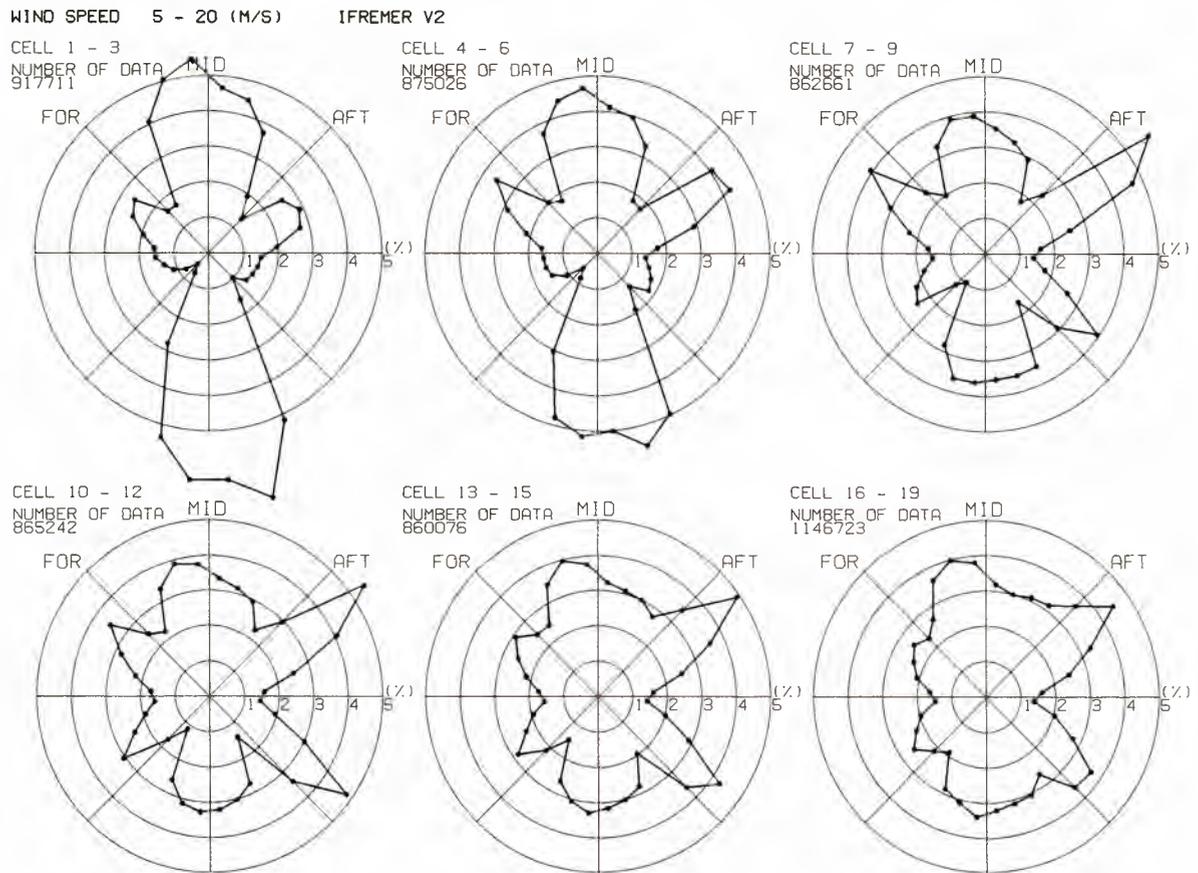


Figure 7. Same as Figure 5, except for the IFREMER off-line products.

model function used to retrieve wind vectors. Since these wind data sets are produced from the same sigma 0 observations by the ERS-1 scatterometer through similar MLE algorithms, the directivity is considered to be caused by errors in the model functions. Though errors in ambiguity removal procedure may also influence directional distribution to some extent, they may not cause directional preference. It can be easily expected that errors in the model function result in systematic errors in the retrieved winds. However, there is a very complex and nonlinear interaction between the model function and retrieved winds, which has not been studied systematically. W. Pierson (unpublished report distributed at the Advanced Earth Observing Satellite/NASA Scatterometer (ADEOS/NSCAT) Science Working Team Meeting in Kyoto, Japan, 1995) tried to make a simple qualitative explanation to the artificial directivity of scatterometer winds by modeling errors in a model function and analyzing properties of the MLE wind retrieval algorithm. Further studies are necessary to understand the nonlinear relationship between the errors in the model function and directional distribution of retrieved wind vectors.

If the irregularities of the directional distributions are caused by the errors in the model function, then the shape of the directional distribution may vary with the incidence angle and wind speed, which are two major parameters of the model function. Figures 5, 6, and 7 show the dependence of the directional distribution on the location of wind vector cells, which is associated with the incidence angle, for the JPL value-added products, the ESA fast delivery products, and the IFREMER off-line products, respectively. Only wind vector cells of wind speed from 5 to 20 m s⁻¹ are used. The true global

wind field should not depend on the cross-track distance or incidence angle and is expected to show distributions close to those of the global wind analyses shown in Figures 1d and 1e.

For each product the distributions at the inner cells (i.e., low incidence angles) show artificial directivity relative to the antenna beam directions. The distributions at the outer cells (i.e., large incidence angles) are smoother and closer in agreement with those of the ECMWF and JMA wind analyses shown in Figure 1. This difference with incidence angle might be due to the low upwind/crosswind modulation of the backscattering property of the C band microwave from the sea surface at low incidence angles [e.g., Freilich and Dunbar, 1993].

Figures 8–12 show the dependence of the directional distributions on wind speed for the JPL value-added products, the ESA fast delivery products, the IFREMER off-line Products, the ECMWF winds, and the JMA GANAL winds, respectively. Only wind vector cells at middle to far ranges (cells 4–19) are used. The distributions for the two wind analyses in Figures 11 and 12 agree well with each other except for those at very high winds. For each scatterometer wind product the distributions at wind speed ranges of 5–10 and 10–15 m s⁻¹ agree better with those of the wind analyses. At lower winds the distributions are mainly affected by the directivity and show features different from those of the wind analyses. At higher winds, very little is known about the wind distribution either for the scatterometer winds and meteorological analyses. Also, the data sampling is not the same for the all distributions. Further studies are needed to clarify characteristics of wind vectors at very high winds.

From the above analysis it is found that scatterometer-

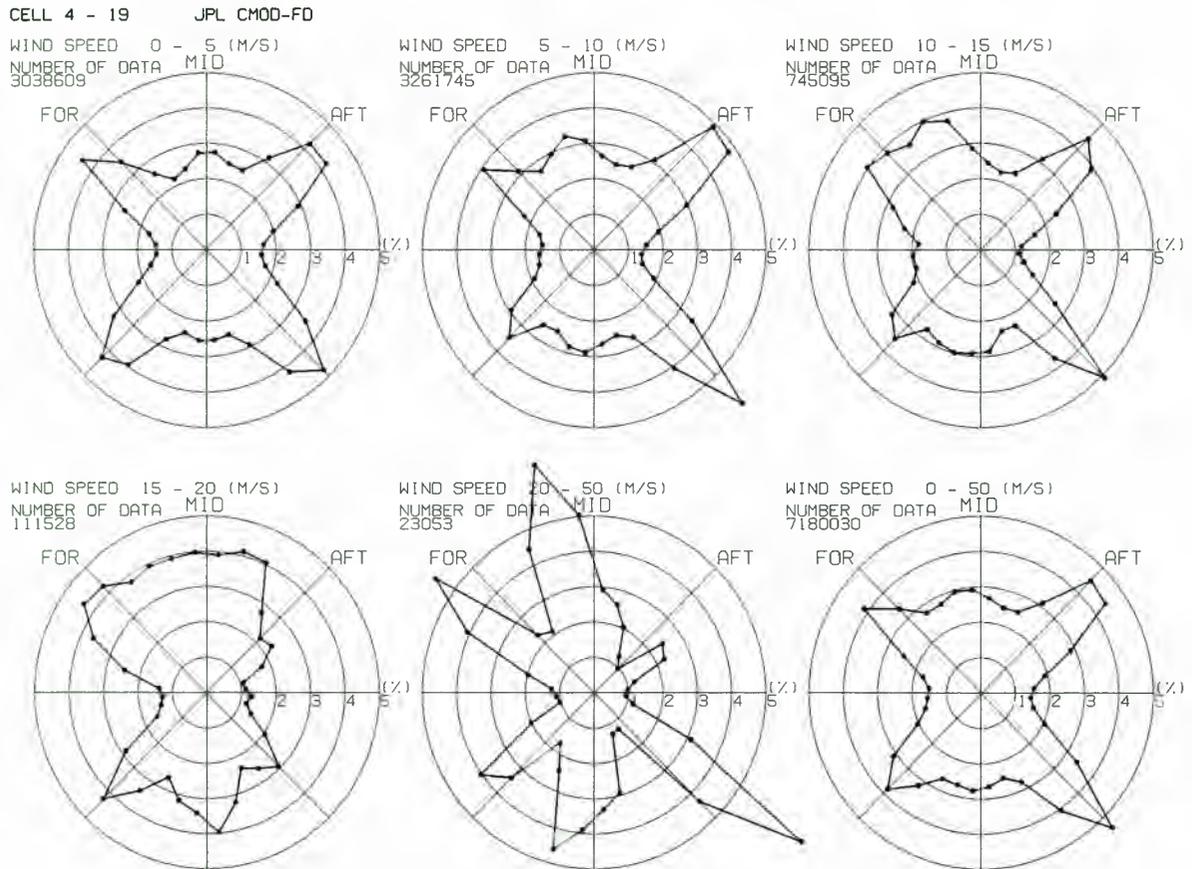


Figure 8. Dependence of the directional distributions on the wind speed for the JPL value-added products.

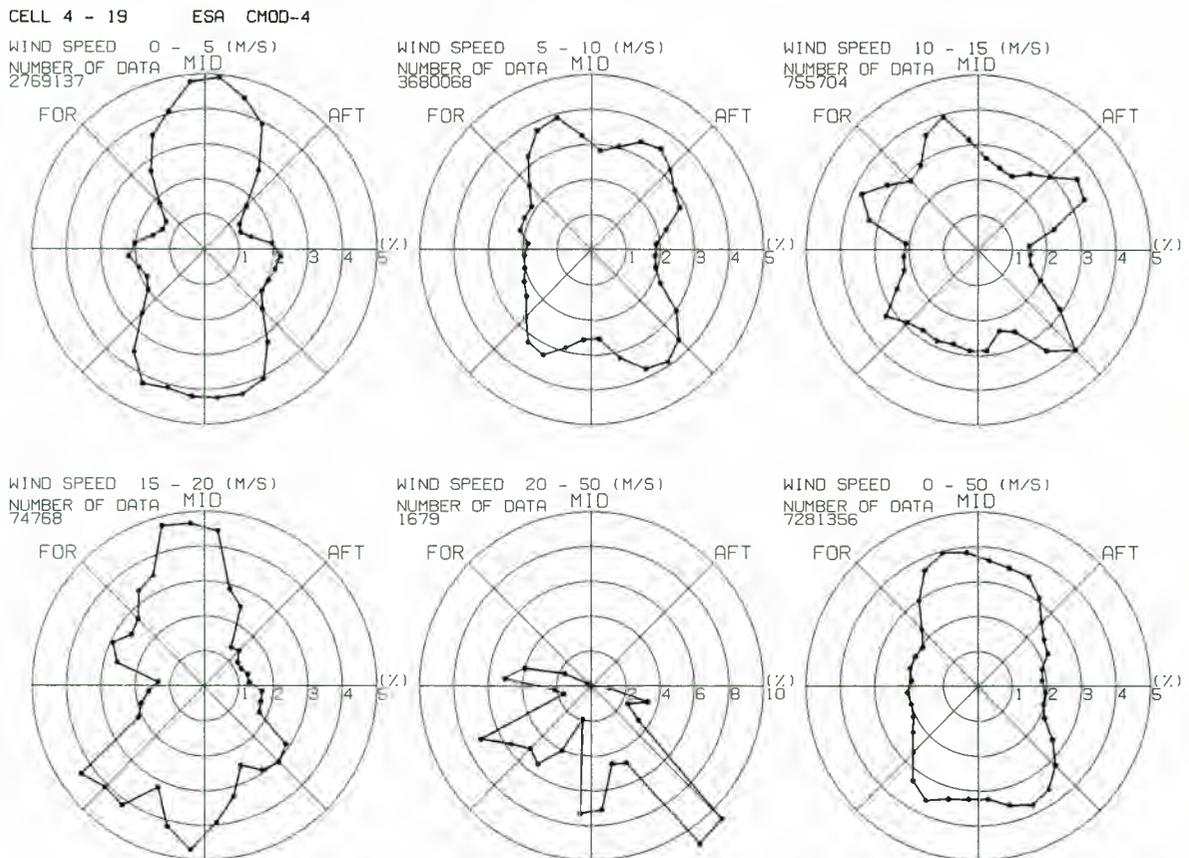
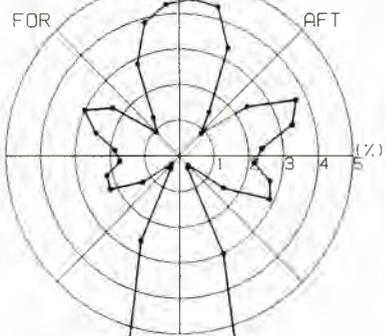


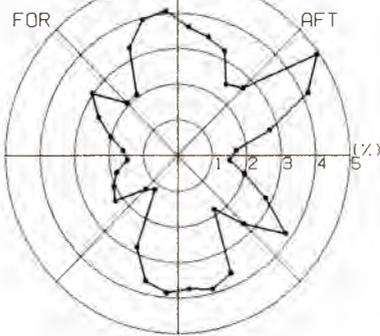
Figure 9. Same as Figure 8, except for the ESA fast delivery products.

CELL 4 - 19 IFREMER V2

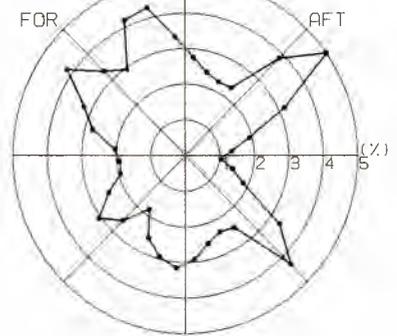
WIND SPEED 0 - 5 (M/S)
NUMBER OF DATA MID
1412852



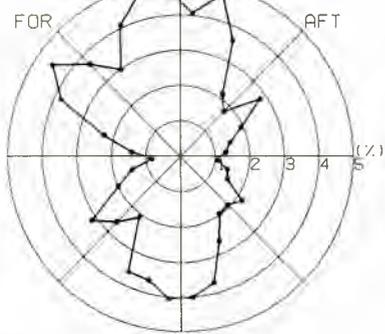
WIND SPEED 5 - 10 (M/S)
NUMBER OF DATA MID
3651871



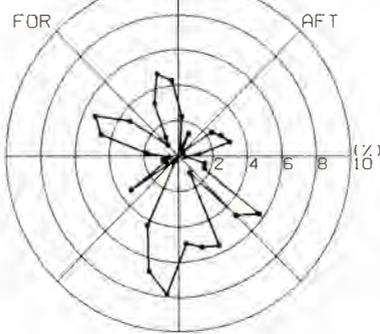
WIND SPEED 10 - 15 (M/S)
NUMBER OF DATA MID
866197



WIND SPEED 15 - 20 (M/S)
NUMBER OF DATA MID
91660



WIND SPEED 20 - 50 (M/S)
NUMBER OF DATA MID
3673



WIND SPEED 0 - 50 (M/S)
NUMBER OF DATA MID
6026253

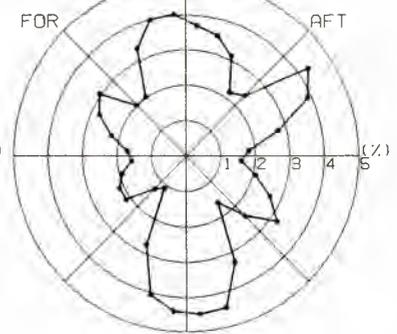
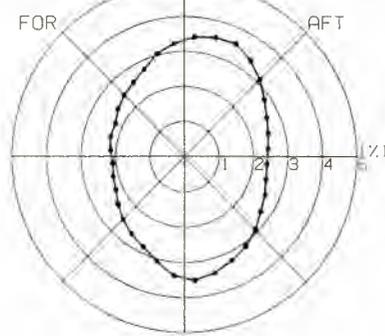


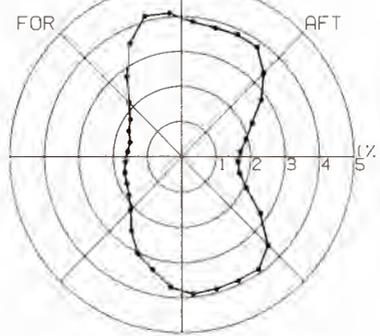
Figure 10. Same as Figure 8, except for the IFREMER off-line products.

ECMWF

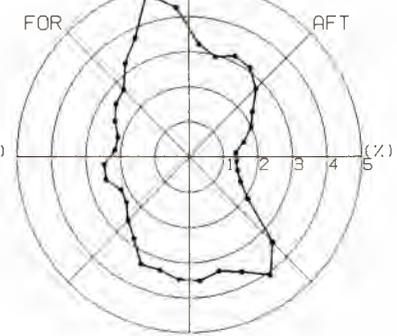
WIND SPEED 0 - 5 (M/S)
NUMBER OF DATA MID
3353012



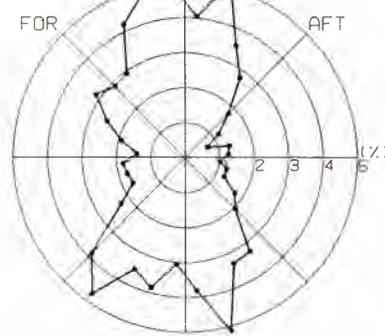
WIND SPEED 5 - 10 (M/S)
NUMBER OF DATA MID
4611686



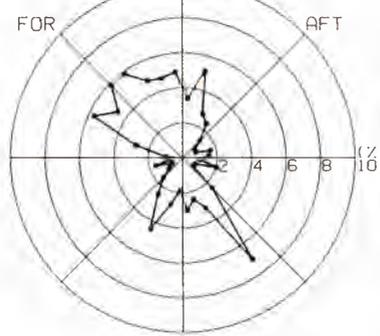
WIND SPEED 10 - 15 (M/S)
NUMBER OF DATA MID
879000



WIND SPEED 15 - 20 (M/S)
NUMBER OF DATA MID
96614



WIND SPEED 20 - 50 (M/S)
NUMBER OF DATA MID
9038



WIND SPEED 0 - 50 (M/S)
NUMBER OF DATA MID
8949350

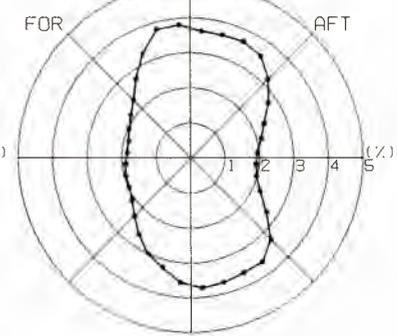


Figure 11. Same as Figure 8, except for the ECMWF winds.

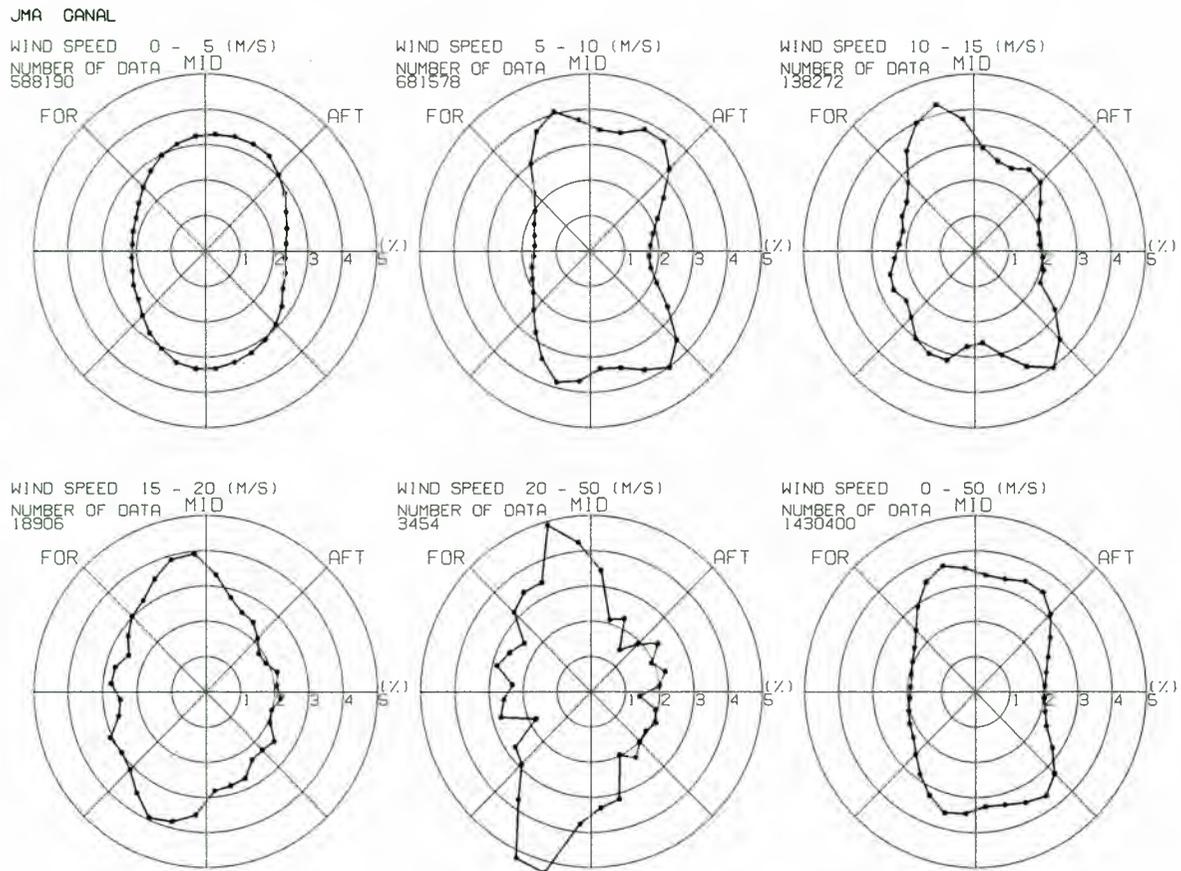


Figure 12. Same as Figure 8, except for the JMA GANAL winds.

derived wind vectors may show systematic directional preference relative to the antenna beams. This artificial directivity is considered to be caused by errors in the model function used to retrieve the wind vectors. The distribution of wind directions varies with the incidence angle and wind speed and is very sensitive to the errors in model function. It is concluded that the directional distribution relative to the antenna beam direction can be used as an effective measure of self-consistency to validate the model function.

4. Summary and Conclusion

In the present study the directional distribution of the wind vectors observed globally by the ERS-1/AMI scatterometer is investigated to examine effects of errors in the geophysical model function used to retrieve the wind vectors from back-scattering measurements. Three scatterometer wind data products such as the JPL value-added products, the ESA fast delivery products, and the IFREMER off-line products for a period from March 1, 1993, to April 30, 1993 (61 days), are analyzed. Only the wind vector cells between 45°S and 45°N are used to avoid contamination by sea ice. Number of the wind vector cells analyzed is about 8.7 million. Frequency distributions of the wind directions relative to the midbeam looking direction are calculated in bins of 10°. In order to compare with the scatterometer winds, surface wind fields of the ECMWF analysis and the JMA GANAL are also analyzed in the same way.

It is found that the scatterometer-derived wind vectors may exhibit systematic directional preference relative to the an-

tenna beam directions. The directional distributions for the JPL value-added products and the IFREMER off-line products obviously show systematic directivity in the direction of the antenna beams. In contrast, the ESA fast delivery products processed by the CMOD-4 model function show smooth distributions very close to those for the ECMWF and GANAL winds. It is shown that the directivity may cause discontinuous wind fields and may result in artificial divergence or convergence in the wind fields. Future investigations are needed to evaluate impacts of the directivity error on the meteorological and oceanographic utility of scatterometer wind data.

Since these three scatterometer wind data sets are produced from the same sigma 0 observations through similar retrieval algorithms, it is considered that these artificial directivities are caused by errors in the model functions. The directional distributions of wind directions varies with the incidence angle and wind speed. Further studies are necessary to understand the very complex and nonlinear interaction between the model function and retrieved winds. A simple examination retrieving wind vectors using a model function with known errors will be effective to investigate the relationship. The antennae geometry of the scatterometer may also affect the wind direction distribution. This will be tested by comparing data from different scatterometers such as SASS and NSCAT.

It is concluded that the directional distribution relative to the antenna beam direction can be used as an effective measure to validate the model function. This method is useful in choosing a better model function among some candidates, as far as for the directional component is concerned. In this sense

the CMOD-4 model provided a better result than the CMOD-FD and CMOD-IFR2. Though the present method is not quantitative, at least, scatterometer-derived wind vectors should not show any artificial directivities relative to the antenna beams.

Acknowledgments. The authors gratefully acknowledge funding support by the Japanese Ministry of Education, Culture and Science under a Grant-in-Aid for Scientific Research on Priority Areas (projects 08241206 and 09227207), the National Space Development Agency of Japan (NASDA), the National Aeronautics and Space Administration (NASA), and the Jet Propulsion Laboratory (JPL) under JPL contract 957652 (NSCAT Science Working Team). We also wish to thank the European Space Agency (ESA) for providing the fast delivery products to the NSCAT and NASDA teams; Michael Freilich (Oregon State University) and Scott Dunbar (JPL) and the NASA Physical Oceanography Distributed Active Archive Center (PO-DAAC) at JPL for making the JPL value-added products available; Kristina Katsaros, Yves Quilfen, Abderrahim Bentamy, and Silvie Pouliquen (IFREMER) and Department d'Océanographie Spatiale, IFREMER and Centre ERS d'Archivage et de Traitement (CERSAT) for providing the IFREMER off-line wind scatterometer data. We also appreciate the encouragement and comments from W. Timothy Liu (JPL), Harunobu Masuko (Communication Research Laboratory, Japan), and members of the NSCAT Science Working Team on carrying out this study. Especially, we thank Willard Pierson (City College of New York) for his valuable suggestions and comments.

References

- Anderson, D., A. Hollingsworth, S. Uppala, and P. Woiceshyn, A study of the use of scatterometer data in the European Centre for Medium-Range Weather Forecasts operational analysis-forecast model, 1, Quality assurance and validation, *J. Geophys. Res.*, **96**(C2), 2619–2634, 1991.
- Bentamy, A., Validation of ERS-1 scatterometer wind data during calibration/validation campaign RENE-91, in *Proceedings of the ERS-1 Geophysical Validation Workshop*, Penhors, France, ESA WPP-36, 77–82, 1992.
- Bentamy, A., Y. Quilfen, P. Queffeuilou, and A. Cavanie, Calibration of the ERS-1 scatterometer C-band model, *Tech. Rep. DRO/OS-94-01*, 72 pp., Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, 1994.
- Chen, H.-T., X.-H. Yan, P.-T. Shaw, and Q. Zheng, A numerical simulation of wind stress and topographic effects on the Kuroshio Current path near Taiwan, *J. Phys. Oceanogr.*, **26**, 1769–1802, 1996.
- Ebuchi, N., H. C. Graber, and R. Vakkayil, Evaluation of ERS-1 scatterometer winds with wind and wave ocean buoy observations, *Tech. Rep. CAOS 96-001*, 69 pp., Cent. for Atmos. and Oceanic Stud., Tohoku Univ., Sendai, Japan, 1996.
- Freilich, M. H., and R. S. Dunbar, A preliminary C-band scatterometer model function for the ERS-1 AMI instrument, in *Proceedings of the First ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-359*, 79–83, 1993.
- Freilich, M. H., H. Qi, and R. S. Dunbar, Scatterometer beam-balancing using open-ocean backscatter measurements, *J. Atmos. Ocean. Tech.*, in press, 1997.
- Graber, H. C., N. Ebuchi, and R. Vakkayil, Evaluation of ERS-1 scatterometer winds with wind and wave ocean buoy observations, *Tech. Rep. RSMAS 96-003*, 69 pp., Rosenstiel Sch. of Mar. and Atmos. Sci., Univ. of Miami, Miami, Fla., 1996.
- Jones, W. L., L. C. Schroeder, D. H. Boggs, E. M. Bracalante, R. A. Brown, G. J. Dome, W. J. Pierson, and F. J. Wentz, The relationship between wind vector and normalized radar cross section used to derive Seasat A satellite scatterometer winds, *J. Geophys. Res.*, **87**(C5), 3318–3336, 1982.
- Katsaros, K. B., A. Bentamy, A. Cavanie, B. Chapron, R. Ezraty, P. Farcy, F. Gohin, P. Queffeuilou, Y. Quilfen, and J. Tournadre, ERS-1 studies at the Department of Oceanography from space, in *Proceedings of the First ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-359*, 61–66, 1993.
- Long, A. E., Towards a C-band radar sea echo model for the ERS-1 scatterometer, in *Proceedings of the Third International Colloquium on Spectral Signatures of Objects in Remote Sensing, Eur. Space Agency Spec. Publ., ESA SP-247*, 29–34, 1985.
- Offiler, D., The calibration of ERS-1 satellite scatterometer winds, *J. Atmos. Oceanic Technol.*, **11**, 1002–1017, 1994.
- Özgökmen, T. M., E. P. Chassignet, and A. M. Paiva, Impact of wind forcing, bottom topography and inertia on mid-latitude jet separation in a quasi-geostrophic model, *J. Phys. Oceanogr.*, in press, 1997.
- Quilfen, Y., ERS-1 off-line wind scatterometer products, *Rep. C1-EX-MUT-CD0000-03-IF*, 59 pp., Cent. ERS d'Arch. et de Trait., Brest, France, 1995.
- Quilfen, Y., and A. Bentamy, Calibration/validation of ERS-1 scatterometer precision products, in *Proceedings of IGARSS*, pp. 945–947, Inst. of Elec. and Electron. Eng., New York, 1994.
- Quilfen, Y., and A. Cavanie, A high precision wind algorithm for the ERS1 scatterometer and its validation, in *Proceedings of IGARSS*, pp. 873–876, Inst. of Elec. and Electron. Eng., New York, 1991.
- Stoffelen, A., and D. Anderson, The ECMWF contribution to the characterization, interpretation, calibration and validation of ERS-1 scatterometer backscatter measurements and winds, and their use in numerical weather prediction models, *ESA Contract Rep. 9097/90/NL/BI*, 106 pp., Eur. Cent. for Medium-Range Weather Forecasts, Reading, England, 1995.
- Woiceshyn, P. M., M. G. Wurtele, D. H. Boggs, L. F. McGoldrick, and S. Peteheych, The necessity for a new parameterization of an empirical model for wind ocean scatterometry, *J. Geophys. Res.*, **91**(C2), 2273–2288, 1986.
- Yasuda, T., and K. Hanawa, Influence of asymmetric wind stress curl on the general ocean circulation using eddy-resolving quasi-geostrophic model, *J. Oceanogr.*, **52**, 189–206, 1996.
- N. Ebuchi, Center for Atmospheric and Oceanic Studies, Faculty of Science, Tohoku University, Aoba, Sendai 980-77, Japan. (e-mail: ebuchi@ocean.caos.tohoku.ac.jp)
- H. C. Graber, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149-1098.

(Received March 24, 1997; revised August 14, 1997; accepted September 24, 1997.)

Arabian sea surface winds and ocean transports determined from ERS-1 scatterometer

David Halpern

Earth and Space Sciences Division, Jet Propulsion Laboratory, California Institute of Technology, Pasadena

Michael H. Freilich

College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis

Robert A. Weller

Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

Abstract. Satellite scatterometer wind velocity measurements are a new source of data for studies of seasonal-to-interannual ocean-atmosphere interactions in the Arabian Sea, where the largest and steadiest wind speeds occur in northern summer. Three satellite wind velocity data products (named CMOD4, FD, and IFR2), all created from radar backscatter measured by the first European Remote Sensing Satellite (ERS-1), were compared to moored-buoy wind measurements recorded in the central Arabian Sea during October 1994 to October 1995. Orthogonal regression analysis of 74 collocations indicated IFR2 winds were more representative of buoy winds compared to CMOD4 and FD. IFR2 winds during 1995 are described, including the time of monsoon onset and the path of the Somali Jet. Monthly mean vertical transport across the bottom of the Ekman layer of the Arabian Sea north of 8°N, meridional Ekman transport along 8.5°N, and Sverdrup transport along 8.5°N were computed with IFR2 data from May 1992 to May 1996. During the 1992–1995 June–September southwest monsoon, upwelling into the Ekman layer of the Arabian Sea occurred at a mean rate of 3.4 Sv ($1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$), which was one-third smaller than the volume transport exiting within the Ekman layer across the southern boundary of the Arabian Sea, and the southward Sverdrup transport was 15 Sv. During the 1992–1996 December–March northeast monsoon, nearly 5 Sv sank below the bottom of the Ekman layer, 4 Sv entered the Arabian Sea in the Ekman layer, and the Sverdrup transport was negligible. Year-to-year transport variations were small.

1. Introduction

Knowledge of spatial variations of vector surface wind stress is critical to understand, model, and predict ocean circulation processes and biological-physical interactions in the Indian Ocean, where seasonally reversing winds or monsoons are caused by differential heating of ocean and land. The coupled atmosphere-land-ocean system of the monsoon is one of Earth's most complex phenomena. In northern hemisphere summer the wind over the Arabian Sea is dominated by the Somali Jet. Wind direction over the Arabian Sea is southwesterly (toward the northeast) during the June–September southwest monsoon and northeasterly (toward the southwest) during the December–March northeast monsoon. July wind stress in the Arabian Sea is 2–3 times larger than that in January. Arabian Sea currents are highly sensitive to surface wind stress. Knox [1987] provided a good primer on the richness and uniqueness of wind-driven ocean circulations in the North Indian Ocean.

Surface wind stress curl produces a vertical velocity at the bottom of the Ekman layer, which causes water to upwell and sink and which significantly influences large-scale ocean cir-

ulation, sea surface temperature, and near-surface phytoplankton biomass. Wind stress is extremely difficult to measure over the ocean. Therefore a parameterization of surface wind stress is normally formulated from near-surface wind velocity components [Trenberth *et al.*, 1990]. Near-surface wind speed and direction are routinely observed by ships and buoys, inferred from satellite scatterometer measurements, and simulated from atmospheric general circulation models. Of the three wind-measuring techniques, the satellite method is the newest. The coverage, resolution, all-weather capability, and ability to infer wind direction as well as speed make spaceborne scatterometers particularly appealing sources of surface wind velocity for air-sea interaction studies in monsoon regions. Except for 3 months in 1978, satellite measurements of surface wind velocity did not exist until July 1991, when the European Space Agency (ESA) first European Remote Sensing Satellite (ERS-1) was launched.

ERS-1 measures microwave radar backscatter, from which near-surface wind velocity is determined. Typically, four possible wind vectors are retrieved for the same location, and selection of a unique direction is difficult because of symmetries in the model function and the scatterometer measurement geometry [Naderi *et al.*, 1991]. Three ERS-1 surface wind velocity data products, all created from backscatter measurements provided by ESA, are compared to moored buoy wind observations in the central Arabian Sea to select a data set for

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02572.
0148-0227/98/97JC-02572\$09.00

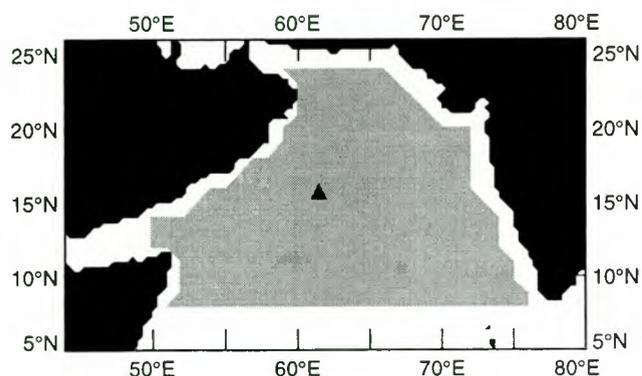


Figure 1. Location of WHOI moored buoy is represented by the triangle. The region defined to be the Arabian Sea is shaded.

analysis. The selected data product is used to compute vertical transport across the bottom of the Ekman layer over the Arabian Sea and Sverdrup and meridional Ekman transports along the southern boundary of the Arabian Sea for 1992–1996. Results demonstrate the utility of scatterometer data for monsoon ocean-atmosphere response studies.

2. Instrumentation

2.1. IMET

Weller [1996] deployed a surface mooring at 15.5°N, 61.5°E (Figure 1) with an Improved Meteorological (IMET) instrument to measure wind speed and direction at 1-min intervals at 2.9-m height. Accuracies of IMET wind speed and direction measurements were 5% and 10°, respectively [Weller and Anderson, 1996]. IMET was located near the axis of the Somali Jet. IMET data were referenced to 10-m height, assuming neutral atmospheric stratification, logarithmic wind profile, von Karman constant of 0.4, and drag coefficient dependent on wind speed according to that specified by Trenberth *et al.* [1990]. At low wind speeds an error may have been introduced in the IMET 10-m height wind estimate by assumption of neutral stratification and neglect of stratification-dependent corrections to the logarithmic profile.

For comparison with ERS-1 data, arithmetic means of sixty 1-min IMET values of east-west component (u , positive eastward), north-south component (v , positive northward), and scalar speed were computed. The 1-hour average wind direction was calculated from 1-hour u and v components. The midpoint of 1-hour averaged IMET wind speed and direction was centered at the time of ERS-1 overflight. The choice of 1 hour to create IMET collocated data was reasonable during the southwest monsoon, when the displacement of an air parcel in 1 hour was comparable to the 50 km diameter of the circular region associated with ERS-1 data.

2.2. ERS-1

The ERS-1 Active Microwave Instrument (AMI) [Attema, 1991] was designed to operate either as a synthetic aperture radar or as a wind scatterometer. In wind scatterometer mode the AMI C band (5.3 GHz), vertically polarized backscatter cross-section measurements were made in a single 500-km-wide swath using three fan-beam antennas oriented at 45°, 90°, and 135° relative to the spacecraft subtrack. Each raw backscatter measurement had a characteristic dimension of 50–70

km. Ground-based data processing produced normalized radar cross-section measurements from each antenna on an approximate 25 km × 25 km along-track/across-track grid throughout the swath. At the moderate incidence angles characteristic of the ERS-1 measurement swath, backscatter from the ocean surface results primarily from a two-scale resonant scattering mechanism. The calculated normalized radar cross section of the sea surface is thus a function of both the amplitudes and the directional distributions (relative to the antenna azimuth) of centrimetric-scale surface roughness elements. Wind created centrimetric surface waves, and wind and waves were in equilibrium. Collocated backscatter measurements were used to estimate both wind speed and wind direction [Naderi *et al.*, 1991].

Three ERS-1 10-m height wind velocity data sets were used: CMOD4 [Offiler, 1994; Stoffelen and Anderson, 1997], FD [Freilich and Dunbar, 1993; Freilich, 1994], and IFR2 [Quilfen and Bentamy, 1994]. All employed the same backscatter cross-section measurements provided by ESA. Each retrieved wind used a different empirical model function. A model or transfer function relates radar backscatter measurements, incidence angle, radar frequency, radar polarization, and wind speed and direction. A unique physically based model function cannot be constructed because of insufficient understanding of radar scattering from a wind-generated centimeter-scale sea surface. The CMOD4, FD, and IFR2 model functions used November 1991 European Center for Medium-Range Weather Forecasts (ECMWF) wind analyses, February 1992 National Centers for Environmental Prediction wind analyses, and March–June 1992 National Oceanic and Atmospheric Administration moored-buoy wind measurements, respectively. Each used a different algorithm to select a unique wind vector from the (typically four) possible solutions derived from the initial wind retrieval processing. More important, no model function or numerical scheme employed IMET data. Each of the data sets corresponded to estimates of 10 m equivalent neutral stability winds. Errors in the FD model function resulted in an anomalously large number of low speeds [Graber *et al.*, 1996], and thus all FD data with wind speeds less than 1.2 m s⁻¹ were discarded prior to comparisons with IMET measurements.

Comparisons between IMET measurements and spatially and temporally collocated ERS-1 wind velocity estimates were used to assess the accuracy and representativeness of the ERS-1 data sets and to select a single scatterometer product for further analysis. Only ERS-1 wind vectors located within 12.5 km of the buoy were included in the comparisons, yielding 74 collocations spanning the time interval October 24, 1994, to October 15, 1995. The range of IMET collocated wind speeds was 3–15 m s⁻¹. The comparison test is noteworthy because of the large percentage of buoy wind speeds greater than 10 m s⁻¹. The mean collocated IMET wind speed was 7.6 m s⁻¹, which is approximately the same as the mean global ocean wind speed. Forty-four (30) collocations were associated with IMET wind speeds below (above) 7.6 m s⁻¹.

IMET measurements are assumed to have fewer errors than ERS-1 data. Moored-buoy wind speed measurements are not error free, as illustrated by Large *et al.* [1995], who identified significant wave-induced biases in 10-m wind speeds computed from buoy measurements recorded within a few meters of the surface. They found altitude-corrected wind speeds to be systematically biased low for true 10-m wind speeds exceeding a transition speed of 7–10 m s⁻¹, depending on anemometer height, because surface waves caused a distortion in the wind

profile. Such biases in buoy data would lead to a significant negative slope in the linear regression analysis between IMET minus ERS-1 residual speed versus IMET wind speed if the ERS-1 data were, in fact, representative of the 10-m wind speed and the collocated data contained a high percentage of high speeds, as in the present case.

Mean IMET minus ERS-1 wind speed residuals were 0.0, 0.5, and 0.6 m s^{-1} for IFR2, FD, and CMOD4 data, respectively. Residuals were dependent on wind speed. IFR2 wind speeds were less (greater) than IMET for speeds below (above) 8 m s^{-1} (Figure 2a); this behavior was different from that reported by *Graber et al.* [1996], in which IFR (an early version of IFR2) underestimated buoy observations for wind speeds from 5 to 15 m s^{-1} . The dynamic range of each ERS-1 collocated data set exceeded that of the IMET data. Slopes of the orthogonal regression line computed between IMET minus ERS-1 wind speed residuals and IMET wind speed were -0.1 , -0.2 , and -0.4 for CMOD4, IFR2, and FD data, respectively. The slopes -0.1 for CMOD4 and -0.2 for IFR2 were not statistically different at the 95% confidence level. Root-mean-square (rms) difference between IMET and IFR2 wind speeds was 0.8 m s^{-1} , which was equal to that computed with CMOD4 data and which was 0.5 m s^{-1} smaller compared to FD. The 0.5 m s^{-1} rms difference was significant at the 95% confidence level.

There were 43 collocations with IMET directions between 0° and 100° , and 31 matchups between 210° and 310° (Figure 2b). Wind direction is defined to be the direction from which the wind was blowing. During the northeast monsoon, mean (standard deviation) values of the collocated IMET minus ERS-1 wind direction residuals were 2 (13), 2 (11), and 4 (24) for IFR2, CMOD4, and FD, respectively. During southwest monsoon conditions, mean (standard deviation) values of the collocated IMET minus ERS-1 wind direction residuals were 9° (8°), 9° (11°), and 10° (7°) for IFR2, CMOD4, and FD, respectively. It is noted that *Large et al.* [1995] found negligible biases in the directions of altitude-corrected wind directions computed from low-level buoy data.

It is not surprising that IFR2 data were more representative of IMET observations. The σ_0 to wind speed and direction model function for IFR2, unlike those for CMOD4 and FD, was created with moored buoy wind data in the North Atlantic and North Pacific during 1992. Biases such as that identified by *Large et al.* [1995] could therefore be inherent in the IFR2 model and would not be identified by additional buoy comparisons. The IFR2 data were the ERS-1 vector data most consistent with IMET measurements. However, we cannot unequivocally state that IFR2 data are the most accurate of the three ERS-1 data products.

3. Arabian Sea Surface Wind Stress Field

Wind stress components, τ_x (positive eastward) and τ_y (positive northward), are defined by

$$\tau_x = \rho_a C_D u (u^2 + v^2)^{1/2} \quad (1a)$$

and

$$\tau_y = \rho_a C_D v (u^2 + v^2)^{1/2}, \quad (1b)$$

where ρ_a is air density (1.225 kg m^{-3}) and C_D is a nondimensional drag coefficient. We used a 10-m height wind-speed-dependent C_D [Trenberth *et al.*, 1990]. The τ_x and τ_y were

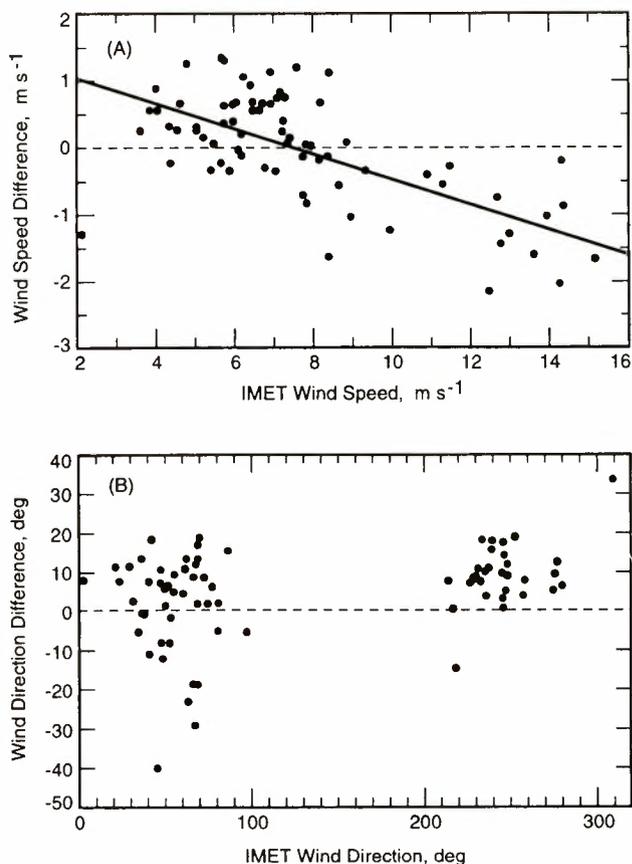


Figure 2. Scatter diagrams of IMET-IFR2 (a) wind speed residuals and (b) wind direction residuals. Abscissas are (a) IMET wind speed and (b) IMET wind direction. Solid line in Figure 2a is least squares orthogonal regression line.

computed from 1-day average u and v in $1/3^\circ \times 1/3^\circ$ areas. All $1/3^\circ \times 1/3^\circ$ daily τ_x and τ_y values in $1^\circ \times 1^\circ$ regions were averaged to monthly mean τ_x and τ_y , from which monthly mean wind stress magnitude, $\tau = (\tau_x^2 + \tau_y^2)^{1/2}$, and direction, $\phi = \tan^{-1}(\tau_x/\tau_y)$, were calculated.

3.1. Seasonal Cycle

In January 1995 (Figure 3a) the dominant wind directions were northerly off India and northeasterly in the western Arabian Sea. The directions were representative of northeast trade winds, which are found throughout the year at similar latitudes in the Atlantic and Pacific. Wind stress decreased from west to east, and near India, wind stresses were two thirds as large as those in the western Arabian Sea.

In July 1995 (Figure 3c) the wind stress field was remarkably different from January, with southwesterly (westerly) direction over the western (eastern) Arabian Sea and much larger wind stresses. In July, virtually the entire Arabian Sea had wind stresses greater than 0.1 N m^{-2} , which in January occurred only over a small region near Somalia. The Somali Jet entered the North Indian Ocean at the Somalia coast near 5°N , rapidly gaining strength as it penetrated toward the northeast. Maximum intensity (0.33 N m^{-2}) occurred about 200 km offshore of Somalia at 9.5°N , 53.5°E , where some of the strongest (July 1995 mean wind speed was 14 m s^{-1}) and steadiest winds occur in the world during northern summer. The area of $\tau > 0.2 \text{ N m}^{-2}$ extended downwind toward the northeast for more than 1000 km, covering a large portion of the basin. Assuming that

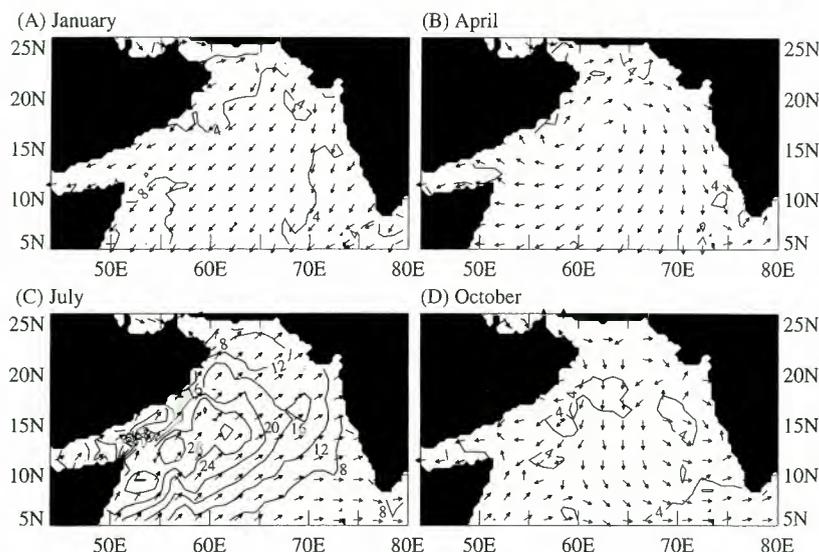


Figure 3. Monthly mean IFR2 wind stress magnitudes and directions during (a) January 1995, (b) April 1995, (c) July 1995, and (d) October 1995. Contour interval is $4 \times 10^{-2} \text{ N m}^{-2}$. Wind stress direction was computed from monthly mean $1^\circ \times 1^\circ \tau_x$ and τ_y values. Arrows point in the direction the wind is blowing. Arrow length is constant. For clarity, arrows are shown for alternate $1^\circ \times 1^\circ$ areas.

patches of $\tau > 0.3 \text{ N m}^{-2}$, of which three occurred in July 1995, are within the Somali Jet, then in the western Arabian Sea the jet trajectory is easily defined: from Somalia at 5°N to 10°N , 53°E to 12°N , 57°E to 13°N , 61°E . East of 61°E , the Somali Jet diminished in strength. Using directions of high wind stresses as a guide, the Somali Jet entered India at 20°N . This pattern resembled one branch of the split Somali Jet portrayed by Findlater [1971], who analyzed 1-km-height winds. A split Somali Jet did not appear in the IFR2 wind pattern for July 1992, 1993, and 1994. Findlater [1971] had inferred a split Somali Jet from very sparse wind observations at 1-km altitude with no data in the $5\text{--}15^\circ\text{N}$, $53\text{--}73^\circ\text{E}$ area. Whether the Somali Jet has one or two branches is a subject of continuing debate (P. Webster, personal communication, 1997).

Whether the number of intense wind stress patches along the Somali Jet was dependent on the ERS-1 sampling scheme was investigated with ECMWF winds, which are produced at 1.125° latitude \times 1.125° longitude resolution every 6 hours. ECMWF winds yielded a single patch near Somalia. ECMWF winds subsampled with the ERS-1 sampling scheme yielded two patches, indicating that ERS-1 sampling produced one or more spurious patches of intense wind stress. A future study of Somali Jet patchiness will be made with surface wind velocity data from the National Aeronautics and Space Administration (NASA) scatterometer (NSCAT), which was launched in August 1996 and which provides twice the quantity of data because it is a double-swath scatterometer.

April 1995 (Figure 3b) was associated with calm conditions. The October 1995 (Figure 3d) wind stress field was similar to that in April, both representing transitions between monsoons.

3.2. Monsoon Onset

Throughout India, the time of monsoon onset is easily defined as the day when it begins to rain and continues to rain heavily for 1–2 weeks, before a break in the monsoon appears. Rainfall is not a good indicator of monsoon conditions over the Arabian Sea because there is not much rainfall, especially in the western area, where hardly any rain occurs. IMET rainfall

measurements totaled 5 cm from October 1994 to October 1995 [Weller, 1996].

For ocean circulation studies, we define the start of the southwest monsoon at a specific location to be the beginning of the initial interval in May or June of sustained wind velocities having eastward and northward components and having a speed greater than 7.5 m s^{-1} . These conditions must persist for at least 6 consecutive days, which is three inertial periods at 15°N and two inertial periods at 10°N and which is considered to be a minimum time to generate Ekman currents. The choice of 7.5 m s^{-1} seems reasonable because the global mean wind speed over the ocean is about 7.5 m s^{-1} , and monsoon wind speeds are known to be above average. With this definition of monsoon onset time, the start of the 1995 monsoon at IMET was June 6–7 (Figure 4).

IFR2 wind velocities recorded within the $1^\circ \times 1^\circ$ area centered at IMET yielded June 6–7 (with no data recorded from May 29 to June 5) as the time of monsoon onset (Figure 4). The excellent agreement between IFR2 and IMET times of monsoon onset provides an opportunity to examine the monsoon onset at the IMET location in other years, when no IMET was there. For 1992, 1993, and 1994 the times of monsoon-

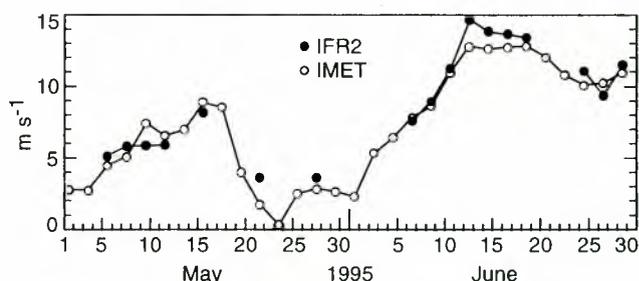


Figure 4. Time series of 2-day average wind speed computed from IMET (open circles) and IFR2 (solid circles) data. About 17 individual IFR2 wind vectors were averaged to form a $1^\circ \times 1^\circ$ 2-day value centered at the IMET location.

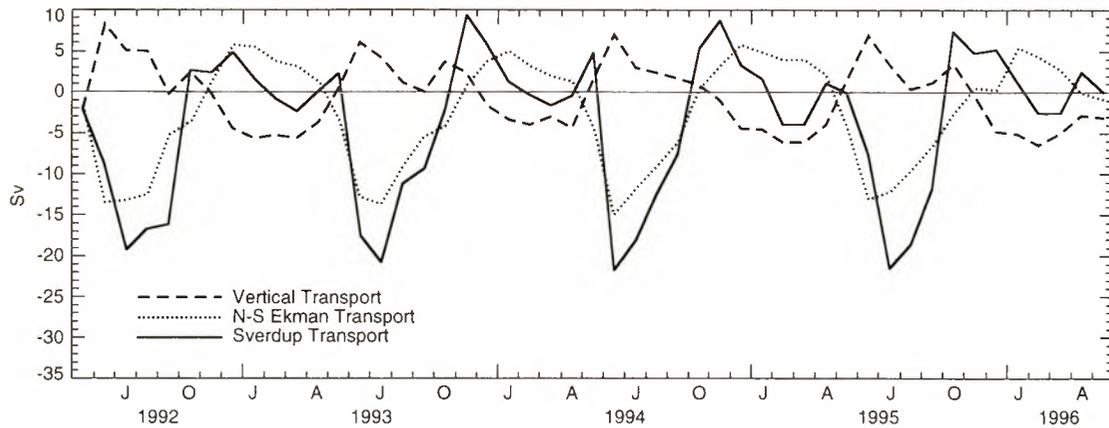


Figure 5. IFR2 monthly mean vertical transports into the Ekman layer of the Arabian Sea north of 8°N (dashed line), north-south component of Ekman transport along 8.5°N (dotted line), and Sverdrup transport along 8.5°N (solid line). ERS-1 data coverage in January, February, and March 1994 was about 50% less compared to other months.

wind onset were June 14–15, June 16–17 (with no data from June 8–15), and June 10–11 (with no data from June 6–9), respectively. Onset times were consistent with the June 11–12 (with a 30-day standard deviation) climatological-mean onset time of monsoon rainfall on the west coast of India near 20°N [Dhar *et al.*, 1980]. At IMET (15.5°N , 61.5°E) the range of monsoon onset time for 1992–1995 was 10 days.

Onset times at the 9.5°N , 53.5°E site of maximum wind stress, named site T, were June 14–15, 1992, May 29–30, 1993, May 31 to June 1, 1994, and June 8–9, 1995. The range was 16 days. Comparing onset times at site T and IMET indicated considerable geographical variation. The influence of ERS-1 sampling [Zeng and Levy, 1995] on spatial features of monsoon onset time will be examined with NSCAT data.

4. Wind-Driven Ocean Transports

4.1. Background

Assuming the monthly mean vertical velocity at the sea surface is zero, the horizontal divergence of the Ekman wind drift is equal to ρW_E , where ρ is seawater density (1025 kg m^{-3}) and W_E is the vertical velocity at the bottom of the Ekman layer. According to Stommel [1965],

$$W_E = (1/\rho f)(\text{curl}_z \tau + \beta \tau_x / f). \quad (2)$$

The Coriolis parameter f is equal to $2\Omega \sin \vartheta$, where ϑ is the latitude, Ω is the rotation rate of Earth ($7.29 \times 10^{-5} \text{ rad s}^{-1}$), and β , the rate of change of the Coriolis parameter with latitude, is equal to $2\Omega \cos \vartheta / R$, with R equal to the radius of Earth ($6.37 \times 10^6 \text{ m}$). $\text{curl}_z \tau$, the vertical component of wind-stress curl, is defined by

$$\text{curl}_z \tau = \partial \tau_y / \partial x - \partial \tau_x / \partial y = \Delta \tau_y / \Delta x - \Delta \tau_x / \Delta y. \quad (3)$$

$\text{curl}_z \tau$ was computed from monthly mean $1^{\circ} \times 1^{\circ}$ τ_x and τ_y values.

The meridional Ekman transport per unit zonal width is $-\tau_x / \rho f$ [Stommel, 1965]. The vertically integrated north-south geostrophic and Ekman transports per unit zonal width, named Sverdrup transport, is $\text{curl}_z \tau / \rho \beta$ [Stommel, 1965].

The Arabian Sea is considered to be a closed basin, with boundaries along 8°N , 24°N , 50°E , and 77°E (Figure 1). Employing conservation of mass, a lowest order approximation of

wind-driven circulation in the Arabian Sea during the southwest monsoon consists of water upwelling into the Ekman layer, water exiting the Arabian Sea along 8.5°N , and water entering the Arabian Sea by the Somali Current. In this scenario the Somali Current transport would equal the Sverdrup transport: this simple view is contrary to the idea that the Somali Current is largely the coastal manifestation of the Great Whirl. The total upward transport into the Ekman layer is assumed to exit the Arabian Sea in the Ekman layer. Over the interior of the Arabian Sea, the upward transport computed from wind stress curl is expected to be smaller than the Ekman transport across 8.5°N because two additional sources of water entering the Ekman layer are coastal upwelling and Somali Current. Because land contaminates the σ_0 from the ocean, no estimate of coastal upwelling is made. Hereafter, vertical transport across the bottom of the Ekman layer over the Arabian Sea excludes transport within a 50-km coastal zone.

4.2. Monthly Mean Ocean Transports 1992–1996

Monthly mean vertical transport across the bottom of the Ekman layer of the Arabian Sea north of 8°N , meridional Ekman transport along 8.5°N , and Sverdrup transport along 8.5°N were computed from May 1992 to May 1996 (Figure 5). Time series began in May 1992 because final calibration of σ_0 was made in March 1992 and wind coverage in April 1992 was poor because the ERS-1 orbit was dedicated to a nonoceanographic objective.

4.2.1. Comparison with in situ data. Chereskin *et al.* [1997] computed wind-driven upper ocean transports along 8.5°N from acoustic Doppler current profiler (ADCP) measurements recorded from research vessels *Malcolm Baldrige* and *Knorr* during June 4–12, 1995, and September 13–26, 1995, respectively. *Malcolm Baldrige* measurements did not extend east of 70°E . ADCP estimates of wind-driven southward transport in June and September were -18 and -8 Sv , respectively. IFR2 southward Ekman transports across 8.5°N for June 4–12, and September 13–26, 1995, were -15 and -6 Sv , respectively. The June 4–12 to September 13–26 ratio of ADCP wind-driven transports was nearly the same as that determined from IFR2 Ekman transports, which is an encouraging sign about reliability of IFR2 Ekman transport. The larger magnitudes

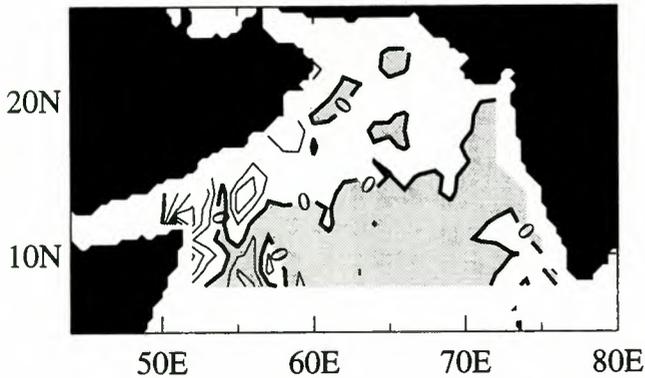


Figure 6. Vertical velocity at the bottom of the Ekman layer per 1° longitude \times 1° latitude regions computed from IFR2 winds for July 1995. Contour interval is $10 \times 10^{-6} \text{ m s}^{-1}$. Shaded region represents downwelling.

associated with the ADCP method could be related to the Great Whirl, a nonwind-driven feature which would create southward flow near 55°E .

A presumed equivalence between Somali Current transport and Sverdrup transport at the same latitude suggests a test about the representativeness of IFR2 Sverdrup transport. During June 3–4 and September 12–13, 1995, *Malcolm Baldrige* and *Knorr*, respectively, recorded ADCP measurements across the Somali Current near 9°N . D. Wilson (personal communication, 1996) and T. Chereskin (personal communication, 1996) computed Somali Current transports over the uppermost 450 m of 17 and 24 Sv in June and September, respectively, which included estimates of the current over the continental shelf where no ADCP data were recorded. IFR2 Sverdrup transports along the Arabian Sea southern boundary at 8.5°N were not calculated for June 3–4 and September 12–13, 1995, because of poor data coverage. IFR2 Sverdrup transports along 8.5°N for June 4–12 and September 13–26, 1995, were 3 and -8 Sv, respectively; however, these values are not reliable because of large variations of zonal averaged IFR2 transports at 1° latitude intervals. In addition to inadequate ERS-1 sampling during a 12-day interval, another reason for employing a longer averaging time to compute the Sverdrup transport is that the Sverdrup transport is applicable to steady state motion or for motion with long-period variations, such as 10 inertial periods. Monthly mean zonal averaged IFR2 Sverdrup transport varied smoothly with latitude. Thus monthly mean IFR2 Sverdrup transport values are compared to Somali Current transport. IFR2 Sverdrup transports along the Arabian Sea southern boundary at 8.5°N for June and September 1995 were -8 and -12 Sv, respectively (Figure 5), which were one half the magnitude of the measured Somali Current transport. The June-to-September ratios of ADCP Somali Current transports and IFR2 Sverdrup transports were almost the same. The difference between Somali Current transport and Sverdrup transport could be related to the Great Whirl, which would enhance Somali Current transport.

4.2.2. Seasonal cycle. Four-year (June 1992 to May 1996) mean values of vertical transport at the bottom of the Ekman layer over the Arabian Sea (excluding a 50-km wide coastal zone), meridional Ekman transport along the Arabian Sea southern boundary at 8.5°N , and Sverdrup transport along the Arabian Sea southern boundary at 8.5°N were -0.5 , -2.7 , and -4.0 Sv, respectively. For each variable, month-to-month vari-

ations were similar for each year (Figure 5), and the ranges of four 1-year mean values were smaller than 1 Sv.

Caution is advised in interpretation of vertical transport over the Arabian Sea to be representative of vertical motion throughout the Arabian Sea. The Somali Jet strongly influences the spatial distribution of wind-stress curl. During summer monsoon, water sinks (rises) in the southern (northern) portion of the Arabian Sea (Figure 6).

Transports during monsoons were substantially greater than the annual mean (Figure 5). At the beginning of the southwest monsoon, water is upwelled into the surface layer. A break in upwelling in August–September occurred in 3 of 4 years; the exception occurred in 1994. During the 1992–1995 June–September monsoon, the average volume of water upwelling into the Ekman layer (3.4 Sv) was about one-third smaller than that exiting in the Ekman layer across the Arabian Sea southern boundary at 8.5°N . Additional waters for the Ekman layer are the Somali Current, in which the surface current reaches 1 m s^{-1} , and coastal upwelling off Oman [Smith and Bottero, 1977]. The 1992–1995 June–September mean Sverdrup transport along the Arabian Sea southern boundary at 8.5°N was 15 Sv, and year-to-year variability was very small (0.3 Sv).

During the 1992–1996 December–March monsoon, the average amount of northward flowing water in the Ekman layer at the Arabian Sea southern boundary was 3.8 Sv (Figure 5), which was 1 Sv or 20% less than that downwelled through the bottom of the Ekman layer over the Arabian Sea. The 1 Sv imbalance is consistent with 5 Sv equatorward flow along east Africa [Bruce et al., 1994], 7 Sv poleward flow off west India [Shetye et al., 1991], and a 0.3 Sv Sverdrup transport along 8.5°N (Figure 5).

In 3 of 4 years the maximum southward Sverdrup transport occurred in July, which was about 1 month after the time of the maximum southward Ekman transport (Figure 5). This is consistent with different spin-up times for Ekman and Sverdrup transports. Spin-up times will be analyzed with NSCAT data because of the greatly increased daily coverage.

5. Concluding Remarks

Three ERS-1 surface wind velocity data products were examined over the Arabian Sea, where large wind speeds and large gradients of wind components exist. Comparison of IMET and ERS-1 data during October 1994 to October 1995 revealed that the IFR2 data product was more representative of IMET observations than CMOD4 and FD data. The three ERS-1 data sets differed in the formulation of the model function and in the ambiguity removal algorithm. Perfect agreement between IMET and IFR2 data cannot be achieved because a time-averaged buoy wind measurement is not compatible with a near-instantaneous areal-averaged satellite wind measurement, even if no errors were associated with each measurement. This fundamental incompatibility, as well as collocation errors and random errors in the scatterometer data can lead to apparent systematic biases, especially in wind speed comparisons, if the buoy measurements are assumed to be error-free [Freilich, 1997]. Reliability of IFR2 data was further enhanced by the good correspondence between IFR2 Ekman transports and ADCP wind-driven upper ocean transports. Agreement between variations of measured Somali Current transport and computed Sverdrup transport along the Arabian Sea southern boundary at 8.5°N in June and September 1995 suggests that the Sverdrup transport along the Arabian Sea

southern boundary has merit to be a Somali Current transport index, which requires further evaluation. IFR2 Sverdrup transport should be compared with that simulated from an ocean general circulation model (OGCM) and with OGCM simulations of Somali Current transport.

IFR2 data did not portray a split Somali Jet, unlike that described by Findlater [1971]. Wind vectors at the surface are not the same as at 1-km altitude because wind direction rotates with height in the tropical planetary boundary layer. The difference of the Somali Jet at the surface and aloft is a subject for further investigation.

A remarkable feature of the Arabian Sea is that sea surface temperature in July is 2°C lower than in April, and in July, sea surface temperature is almost as low as that in January. This atypical behavior for northern latitudes in summer is caused by extreme ocean-atmosphere interactions during the southwest monsoon. Understanding the relative importance of several physical processes, such as evaporation, entrainment, upwelling, and horizontal advection, requires accurate knowledge of surface winds. Aliasing will be substantially reduced with NSCAT data. Because no single data source can capture the rich spectrum of surface winds, research is warranted on integration of ERS-2 (ERS-1 follow-on) and NSCAT data with in situ data.

Acknowledgments. We are extremely grateful to the IFREMER Centre ERS d'Archivage et de Traitement (CERSAT) for producing and distributing IFR2 data. We are grateful to Teresa Chereskin (Scripps Institution of Oceanography) and Douglas Wilson (NOAA Atlantic Oceanographic and Meteorological Laboratories) for kindly sending Knorr and Malcolm Baldrige results in advance of their publication. Kenneth Brink (Woods Hole Oceanographic Institution) and Robert Knox (Scripps Institution of Oceanography) provided valuable comments on an early version of the manuscript. Computer-generated data processing and graphical/tabular analyses were made by Peter Woiceshyn (JPL) in an outstanding manner. The work described in this paper was supported by the NASA Physical Oceanography Program RTOP 622-47-09 (D.H.); NASA SeaWinds Project RTOP 229-18-41 (D.H.); NASA NSCAT Project via JPL contract 959351 (M.H.F.), and ONR grant N00014-94-1-0161 (R.W.). The research described in this paper was performed, in part, by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

- Attema, E. P. W., The active microwave instrument on board the ERS-1 satellite, *Proc. IEEE*, **79**, 791–799, 1991.
- Bruce, J. G., D. R. Johnson, and J. C. Kindle, Evidence for eddy formation in the eastern Arabian Sea during the northeast monsoon, *J. Geophys. Res.*, **99**, 7651–7664, 1994.
- Chereskin, T. K., W. D. Wilson, H. L. Bryden, A. Field, and J. Morrison, Observations of the Ekman balance at 8°30'N in the Arabian Sea during the 1995 southwest monsoon, *Geophys. Res. Lett.*, **24**, 2541–2544, 1997.
- Dhar, O. N., P. Rakhecha, and B. Mandal, Does early or late onset of monsoon provide any clue to subsequent rainfall during the monsoon season?, *Mon. Weather Rev.*, **108**, 1069–1072, 1980.
- Findlater, J., Mean monthly air flow at low levels over the western Indian Ocean, *Geophys. Mem.*, **115**, 53 pp., Her Majesty's Stn. Off., London, 1971.
- Freilich, M. H., ERS-1 scatterometer measurements over the Southern Ocean, in *Proceedings, Second ERS-1 Symposium, Space at the Service of our Environment, 11–14 October 1993, Hamburg*, pp. 1111–1116, Eur. Space Agency, Paris, 1994.
- Freilich, M. H., Validation of vector magnitude data sets: Effects of random component errors, *J. Atmos. Oceanic Technol.*, **14**, 695–703, 1997.
- Freilich, M. H., and R. S. Dunbar, A preliminary C-band scatterometer model function for the ERS-1 AMI instrument, *Proceedings, First ERS-1 Symposium: Space at the Service of our Environment, 4–6 November 1992, Cannes*, pp. 79–83, Eur. Space Agency, Paris, 1993.
- Graber, H. C., N. Ebuchi, and R. Vakkayil, Evaluation of ERS-1 scatterometer winds with wind and wave ocean buoy observations, *RSMAS Tech. Rep. 96-003*, 78 pp., Univ. of Miami, Miami, Fla., 1996.
- Knox, R. A., The Indian Ocean: Interaction with the monsoon, in *Monsoon*, edited by J. S. Fein and P. L. Stephens, pp. 365–397, John Wiley, New York, 1987.
- Large, W. G., J. Morzel, and G. B. Crawford, Accounting for surface wave distortion of the marine wind profile in low-level Ocean Storms wind measurements, *J. Phys. Oceanogr.*, **25**, 2959–2971, 1995.
- Naderi, F., M. H. Freilich, and D. Long, Spaceborne radar measurement of wind velocity over the ocean—An overview of the NSCAT scatterometer system, *Proc. IEEE*, **79**, 850–866, 1991.
- Offiler, D., The calibration of the ERS-1 satellite scatterometer winds, *J. Atmos. Oceanic Technol.*, **11**, 1002–1017, 1994.
- Quilfen, Y., and A. Bentamy, Calibration/validation of ERS-1 scatterometer precision products, in *Proceedings, IGARSS 1994, Pasadena, 8–12 August 1994*, pp. 945–947, IEEE, Piscataway, N. J., 1994.
- Shetye, S., A. Gouvea, S. Shenoi, G. Michael, D. Sundar, and K. Santanam, The coastal current off western India during the north-east monsoon, *Deep Sea Res.*, **38**, 1517–1529, 1991.
- Smith, R. L., and J. Bottero, On upwelling in the Arabian Sea, in *A Voyage of Discovery, George Deacon 70th Anniversary Volume*, edited by M. Angel, pp. 291–304, Pergamon, Tarrytown, N. Y., 1977.
- Stoffelen, A., and D. Anderson, Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4, *J. Geophys. Res.*, **102**, 5767–5780, 1997.
- Stommel, H. M., *The Gulf Stream*, 248 pp., Univ. of Calif. Press, Berkeley, 1965.
- Trenberth, K. E., W. G. Large, and J. G. Olson, The mean annual cycle of the global ocean wind stress, *J. Phys. Oceanogr.*, **20**, 1742–1760, 1990.
- Weller, R. A., In situ observations of surface meteorology and air-sea fluxes, *1996 U.S. WOCE Rep.*, pp. 17–20, Texas A&M Univ., College Station, 1996.
- Weller, R. A., and S. P. Anderson, Surface meteorology and air-sea fluxes in the western equatorial Pacific warm pool during the TOGA Coupled Ocean-Atmosphere Response Experiment, *J. Clim.*, **9**, 1959–1990, 1996.
- Zeng, L., and G. Levy, Space and time aliasing structure in monthly mean polar-orbiting satellite data, *J. Geophys. Res.*, **100**, 5133–5142, 1995.
- D. Halpern, Earth and Space Sciences Division, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109. (e-mail: halpern@pacific.jpl.nasa.gov)
- M. H. Freilich, College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331.
- R. A. Weller, Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, MA 02543.

(Received May 22, 1997; revised September 2, 1997; accepted September 8, 1997.)

VIERS-1 scatterometer model

P. A. E. M. Janssen,¹ H. Wallbrink, C. J. Calkoen,² D. van Halsema,³
W. A. Oost, and P. Snoeij⁴

Department of Oceanography, Royal Netherlands Meteorological Institute, De Bilt, Netherlands

Abstract. In this paper a description is given of a physically based theoretical ocean backscatter model (called the VIERS-1 model) for intermediate incidence angles, and a comparison of its performance against the CMOD4 empirical model is made. The VIERS-1 scatterometer algorithm is based on a two-scale composite surface model which includes both specular and Bragg scattering. Its short wave model is based on the energy balance equation and accounts for viscous damping, slicks, dissipation due to whitecapping, and nonlinear three- and four-wave interactions. A number of parameters in the model have been determined by means of laboratory data and analyzed European Centre for Medium-Range Weather Forecasts (ECMWF) winds. Because of the two-scale approach the wave number up to which Bragg scattering applies should be determined. This is done by means of laboratory data at X band. In addition, laboratory data of the wave spectrum have been utilized to validate the VIERS-1 short wave spectrum. An inverse of the algorithm is developed to derive wind speed and direction from the observed (ERS-1) backscatter and by comparison with ECMWF analyzed winds' three parameters for the short wave spectrum, namely, the Phillips parameter, the directional width of the spectrum, and the wave number boundary between gravity waves and short waves have been obtained. Comparisons between VIERS-1, C band model, version 4 (CMOD4), and ECMWF analyses are made. VIERS-1 performs better in the high wind speed range, and this feature is of importance when scatterometer winds are assimilated into an atmospheric model. However, in terms of backscatter rather than wind speed, CMOD4 shows better results. It is suggested that this is caused by the too simple directional distribution of the VIERS-1 short wave spectrum.

1. Introduction

Traditionally, the operational retrieval algorithms for the scatterometer, which relate the radar backscatter measurements to the surface wind vectors, have been empirical. A review of the history of this empirical relationship is given by, for example, *Moore and Fung* [1979], *Jones et al.* [1982], *Schroeder et al.* [1982], and *Barrick and Swift* [1980]. After *Moore and Pierson* [1967] proposed to use a satellite scatterometer's radar echo to determine the wind speed at sea, a variety of early scatterometer models appeared in the 1970s [e.g., *Valenzuela et al.*, 1971; *Guinard et al.*, 1971; *Jones et al.*, 1977; *Moore and Fung*, 1979; *Wentz et al.*, 1984]. The most successful scatterometer model of the early 1980s was the Seasat A scatterometer system (SASS 1) model. The SASS 1 model assumed a power law between the radar backscatter σ and wind speed U and was tuned to a subset of the available surface truth wind data from the Joint Air-Sea Interaction Experiment (JASIN). The wind speed data set was relatively small, 4–16 m/s. When

results of the tuned SASS 1 model were compared with the JASIN data not used in the tuning, a favorable agreement was found, giving confidence in the empirical approach. Nevertheless, *Woiceshyn et al.* [1986] and *Anderson et al.* [1987] pointed out several weak points of the SASS 1 algorithm. First of all, low wind speeds were systematically too high while high wind speeds were too low. Secondly, winds obtained from horizontal polarization were not consistent with vertical polarization, suggesting that a power law relationship between backscatter and wind is not adequate. Furthermore, it was also felt that other geophysical parameters such as atmospheric stability and water viscosity would have resulted in an improved wind field retrieval, in particular at the lower wind speeds.

Despite the shortcomings, the statistical fitting approach has resulted in a useful algorithm as follows from the work of *Stoffelen and Anderson* [1997], although a somewhat more sophisticated power law relationship needed to be introduced. The resulting backscatter algorithm, C band model, version 4 (CMOD4), showed a very good fit in backscatter space while in comparison with European Centre for Medium-Range Weather Forecasts (ECMWF) wind fields, the retrieved wind velocity had a small wind speed error of about 2 m/s and directional error of the order of 20–30°. However, CMOD4 showed for low and high wind speed similar problems as the SASS 1 algorithm. When using CMOD4 in ECMWF's analysis system, *Gaffard and Roquet* [1995] found that the underestimation of wind speed in the high wind speed range resulted in less deep lows (by as much as 8 hPa) and, as a consequence, the quality of the atmospheric forecast suffered. By applying a wind speed dependent bias correction to CMOD4 (which was

¹Now at European Centre for Medium-Range Weather Forecasts, Reading, England, United Kingdom.

²Now at Advisory and Research Group on Geo Observation Systems and Services, Vollenhove, Netherlands.

³Now at Physics and Electronics Laboratory, Netherlands Organization for Applied Scientific Research, Delft.

⁴Now at Delft University of Technology, Delft, Netherlands.

obtained by a comparison with buoy wind speed data) the scatterometer winds were found to have a favorable impact on atmospheric analysis and forecast and even on the ocean wave analysis and forecast [Andersson *et al.*, 1998].

From a scientific point of view the assumption that the backscatter only depends on the local wind field may be questioned, however, since the backscatter reflects to some extent the state of the high-frequency wind waves. The spectrum of wind-generated gravity-capillary waves not only depends on the local wind but is determined by a number of physical processes, namely, wind input, nonlinear three- and four-wave interaction, viscous dissipation, and dissipation due to slicks. Thus, when the waves are sufficiently steep, which may occur, for example, for young wind seas, nonlinear processes may be dominant so that the state of the short waves is, through the energy cascade, mainly determined by the longer gravity waves. The state of the longer waves depends on factors such as coastal geometry, duration of a storm system, currents, and bathymetry, and therefore the radar backscatter may depend on these environmental circumstances as well. On the other hand, for low wind speed, viscous dissipation and dissipation due to slicks may be relevant processes in determining the shape of the short wave spectrum, again suggesting that not only the local wind determines the backscatter. Furthermore, it should be pointed out that the radar backscatter shows an additional dependence on the state of the long gravity waves because the short waves which provide the backscatter are tilted by the longer gravity waves. This effect is especially relevant for small incidence angles.

The above considerations prompted an extensive investigation into the dependence of the radar backscatter on physical parameters such as wind speed, sea state, the presence of slicks, and the air-sea temperature difference. The work was supported by the Netherlands Remote Sensing Board (BCRS), and the Verification and Interpretation of ERS (VIERS) group emerged, which started an experimental study in the laboratory and at sea to address the above mentioned issues. The radar used in these studies operated at X band. Parallel to the experimental work, the VIERS group started the development of a scatterometer algorithm based on the present understanding of the radar backscatter process and of the relevant processes governing the shape of the gravity-capillary spectrum. The observed results on radar backscatter and the short wave spectrum were used as a guideline to tune a number of unknown parameters in the scatterometer algorithm. As a result, a backscatter algorithm based on physics rather than empirical fitting was obtained [van Halsema *et al.*, 1989; Calkoen *et al.*, 1990; Snoeij *et al.*, 1993; Janssen *et al.*, 1995].

Confidence in the performance of the VIERS algorithm increased when the simulated backscatter, obtained using the ECMWF wind and wave model (WAM) wave fields, was compared with the backscatter as observed by the ERS-1 satellite. The air-sea temperature difference was set to zero since neutral stability is the most common situation that occurs over the oceans. The period of interest was November 6, 1991, at 1200 UT. In Figure 1 we compare the simulated backscatter with the observed ones showing an overall good agreement, even for low wind speed. As a benchmark, we have also shown results of the present operational scatterometer algorithm CMOD4. We conclude from this comparison that the VIERS scatterometer algorithm performs well, even compared to CMOD4. We add to this that while the VIERS-1 two-scale composite surface model was tuned at X band, the ERS-1 satellite operates at C

band, thus giving some confidence in the validity of the VIERS-1 model over a wider range of radar frequencies. Despite the good agreement between observed and simulated backscatter, it should be pointed out that the root-mean-square (rms) error, which is of the order of 2–3 dB, is large when compared with the observation error, which is about 0.2 dB. This could imply two things. On the one hand, it can be argued that the large rms error is caused by model errors. This possibility seems unlikely, however, since it is known that CMOD4, which has the bigger rms error, has a reasonable skill in retrieving wind speed and direction. On the other hand, a source of error could be provided by the analyzed wind and wave field used in the generation of the simulated backscatter with the scatterometer algorithms. Hence the observed backscatter might contain relevant information regarding, for example, the surface wind field. In view of the large difference between the rms error of Figure 1 and the measurement error the information contained in the observed backscatter might therefore be quite considerable. In order to investigate this we need to invert the VIERS algorithm so that a retrieval of wind speed and direction from the observed backscatter becomes feasible.

The program of this paper is as follows. In section 2 a review of the VIERS scatterometer algorithm is given, including a description of the energy balance of short waves. Once the wave spectrum is known, a two-scale model is used to determine the normalized backscatter. The scatterometer algorithm thus obtained gives the radar backscatter as the function of a number of geophysical parameters such as wind speed and direction, the sea state, slicks, air-sea stability, the sea surface temperature, and, of course, geometrical parameters such as the look angle and the incidence angle. For practical application discussed in this paper we restrict our attention to the usual parameters, wind speed and direction and incidence angle, and allow the sea state as an additional parameter. Effects of air-sea stability are disregarded because, except near coasts, the stability of the atmosphere over the oceans is almost always close to neutral and only an averaged effect of slicks is taken into account. Nevertheless, the VIERS algorithm is computationally very expensive and is as such not a feasible option when used in routine applications such as wind retrieval for determining the analyzed weather over the oceans.

For practical applications we therefore generated tabulated values of the normalized radar backscatter σ_0 as a function of wind speed, wind direction, incidence angle, and sea state. Once this table of σ_0 is known, we retrieve wind speed and direction from the observed ERS-1 σ_0 triplet by minimizing a measure of distance between observed and modeled backscatter under the side constraint that the retrieved wind direction does not deviate too much from the wind direction provided by the ECMWF model. This procedure allows for a unique solution of the wind vector and is discussed in section 3.

Retrieved wind vectors by means of the VIERS scatterometer algorithm are compared in section 4 with the ECMWF analyzed wind fields, and it is found that VIERS and ECMWF winds are compatible in a statistical sense. As a reference, we use results from CMOD4. A direct comparison between VIERS and CMOD4 winds reveals a good agreement between the two products, except that in the high wind speed range, VIERS has higher winds than CMOD4, while at low winds, VIERS has lower winds. In view of the problems CMOD4 has at the extreme wind speed ranges (which, as Gaffard and Roquet [1995] have pointed out, have a detrimental impact on the forecast) it is concluded that the VIERS model is performing

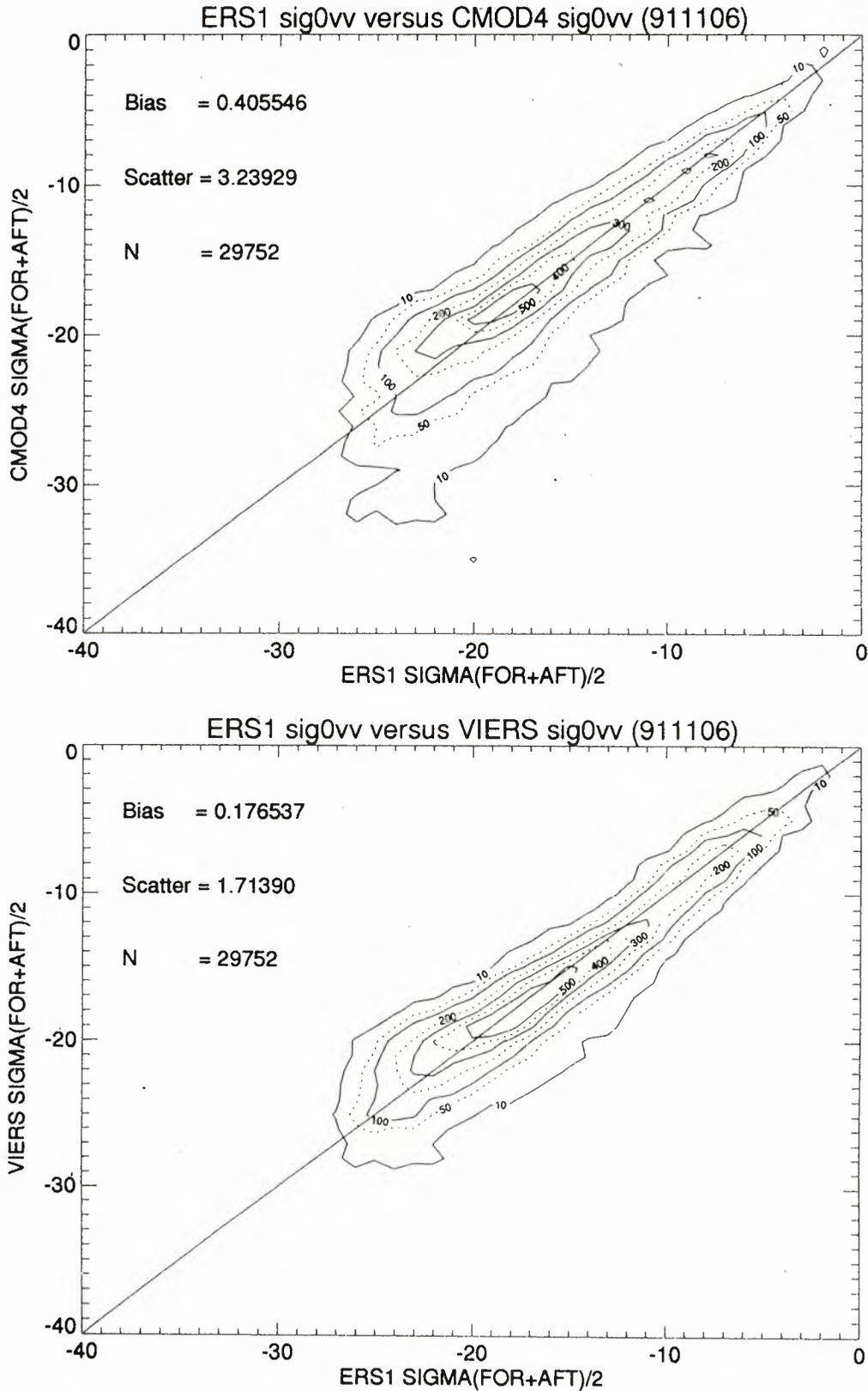


Figure 1. Simulated normalized backscatter using (top) CMOD4 and (bottom) VIERS algorithm versus observed backscatter from ERS-1 on November 6, 1991.

better. Nevertheless, it should be emphasized that the misfit in σ_0 space between modeled and observed backscatter is generally larger for VIERS than for CMOD4. This is probably caused by a too simple directional distribution of wind waves in

VIERS. After the statistical comparison we proceed with a synoptic discussion of the differences between retrieved and analyzed wind concentrating on a frontal system that occurred in the Norwegian Sea on November 6, 1991. A summary of

conclusions is presented in section 5, and suggestions for improvements are given as well.

2. Viers Model

To our knowledge to date, the most successful modeling attempt of the scattering of the radar signal at the sea surface is due to *Donelan and Pierson* [1987]. These authors obtained the spectrum of gravity-capillary waves from a simplified energy balance, consisting of wind input and dissipation through viscosity and wave breaking. Combining this short wave number with the observed directional spectra of *Donelan et al.* [1985], the complete surface wave spectrum is then known.

Scattering off a surface with a broad spectrum of waves is reasonably well modeled by means of a so-called two-scale approach [*Valenzuela*, 1978; *Plant*, 1990]. Thus the normalized radar backscatter can be found by integrating the scatter from the individual facets (which are tilted by the longer gravity waves) weighted with the probability that the water surface is tilted by a certain angle. Introducing a separation scale wave number k_c , for high wave numbers ($k > k_c$) the main scattering mechanism is assumed to be Bragg scattering, while for $k < k_c$, specular reflection is taken. Results are fairly weakly dependent on the choice of k_c [*Plant*, 1990]; *Donelan and Pierson* [1987] selected $k_c = k_b/40$, where k_b is the Bragg wave number

$$k_b = 2k_R \sin \theta_l \quad (1)$$

where θ_l is the local incidence angle, which depends on the tilt of the water surface by the long waves, and k_R is the radar wave number.

Application of the *Donelan and Pierson* [1987] scatterometer algorithm to the VIERS wave tank data set revealed a number of shortcomings in the model [*Calkoen et al.*, 1990]. First of all, considerable discrepancies between modeled and observed short wave spectral shape were found; for large winds the modeled spectrum drops off too rapidly, while at low wind speed, considerable amounts of wave energy were observed beyond the viscous cutoff wave length of *Donelan and Pierson*. In other words, the modeled *Donelan and Pierson* short wave spectrum was found to be too sensitive to the effects of water viscosity. Clear experimental evidence of this was also given by *Jähne and Riemer* [1990], who, in the framework of the VIERS project, measured the slope spectrum by means of optical techniques. Even for low wind speed, considerable contributions to the slope spectrum were found at wave number $k = 800$ rad/m, which is well beyond the viscous cutoff of *Donelan and Pierson*. *Apel* [1994] has summarized the findings of *Jähne and Riemer* [1990] and *Klinke and Jähne* [1992] in terms of a semiempirical model for the short wave spectrum.

Also, shortcomings in the electromagnetics part of the *Donelan and Pierson* [1987] scatterometer algorithm were found. In order to shield off a singularity in the Bragg scattering near nadir, a cutoff condition was applied when the local incidence angle was $< 18^\circ$. Nevertheless, for small incidence angle the contribution of Bragg scattering dominates the one of specular reflection, which is unexpected. Furthermore, the choice of the cutoff wave number is not always adequate. For example, for the wave tank data the dominant wave has a much larger wave number than at sea, sometimes even beyond k_c .

In order to alleviate the above mentioned problems it was decided to develop a new scatterometer algorithm. In partic-

ular, regarding the electromagnetic part, the introduction of a cutoff wave number on a more or less physical basis was considered. Furthermore, it was decided to remove the cutoff condition near nadir and to add the specular reflection. Finally, it was thought to include more physics in the energy balance equation, because nonlinear wave-wave interactions and effects of slicks may be relevant processes as well in determining the short wave spectrum.

Before our attempt to improve on the *Donelan and Pierson* [1987] scatterometer model is described, it is emphasized that a basic assumption of the VIERS scatterometer model is that wind wave generation is determined by the surface stress. At the air-sea interface the stress field is determined by the wind speed, the stability of the air column, and the sea state, which we characterize by the wave age of the wind sea. In order to obtain the stress τ or the friction velocity u_* the wind profile $U(z)$ is assumed to have the form

$$U(z) = \frac{u_*}{k} [\log(z/z_0) - \psi(z/L)] \quad (2)$$

Here the roughness length includes the effects of the sea state and is parameterized according to *Smith et al.* [1992]

$$z_0 = \alpha \frac{u_*^2}{g} \quad \alpha = 0.48(c_p/u_*)^{-1} \quad (3)$$

where c_p/u_* is the age of the wind sea and c_p is the phase speed of the peak of the wind sea spectrum.

For the stability function ψ we adopt the Businger-Dyer expression [*Businger et al.*, 1971; *Dyer and Hicks*, 1970]. For stable conditions ($L > 0$) we have

$$\psi = -z/L$$

while for unstable conditions ($L < 0$) we take

$$\psi = \frac{\pi}{2} + 2 \log [(1 + \Phi)/2] + \log [(1 + \Phi^2)/2] - 2 \tan \phi \quad (4)$$

where

$$\Phi = (1 - 16z/L)^{0.25} \quad (5)$$

and the Monin-Obukov length is computed according to *Stewart* [1985]

$$L = \frac{-u_* U(z)}{g\kappa} \frac{T_{\text{air}}}{T_{\text{sea}} - T_{\text{air}}} \quad (6)$$

with g as the acceleration of gravity, κ as the von karman constant, T_{air} as the air temperature, and T_{sea} as the water temperature.

For given wind speed, phase speed of the waves, air temperature, and sea temperature the friction velocity u_* is solved from (2) in an iterative manner.

2.1. Short Wave Model

The model for the short wave spectrum is based on the energy balance equation, which is solved under steady state circumstances because the short waves have a very short response timescale. Also, advection of short wave energy is disregarded, and the energy balance equation therefore reads

$$S_{\text{in}} + S_{\text{nonl}} + S_{\text{visc}} + S_{\text{br}} + S_{\text{slicks}} = 0 \quad (7)$$

where S_{in} represents the effects of wind stress, S_{nonl} describes three- and four-wave interactions, S_{visc} describes viscous dissipation, S_{br} describes dissipation due to whitecapping, and S_{slicks} describes the resonant energy transfer between surface waves and slicks (Marengoni effect). The energy balance equation (7) is solved as a boundary value problem in wave number space by providing the energy flux from the long to the short waves at a boundary $k = k_{join} \approx g/u_*^2$, which corresponds to $c/u_* \approx 1$ with c as the phase speed of gravity waves.

In order to determine the energy flux at the boundary $k = k_{join}$, knowledge of the gravity part of the wave spectrum is required. In general, the long wave spectrum consists of wind sea and swell, and the simplifying assumption is made that the energy flux at k_{join} is mainly determined by the wind sea part of the spectrum since swell usually has a small steepness. The one-dimensional wind sea wave number spectrum is assumed to be given by the Joint North Sea Wave Project (JONSWAP) shape [Hasselmann *et al.*, 1973], which is obtained from the frequency spectrum by using the linear dispersion relation for gravity waves. Hence

$$F(k) = \frac{1}{2} \alpha_p k^{-4} \exp \left[-\frac{5}{4} \left(\frac{k_p}{k} \right)^2 \right] \gamma^r \quad (8)$$

where

$$r = \exp \left[\frac{-(k^{1/2} - k_p^{1/2})^2}{2\sigma^2 k_p} \right]$$

In JONSWAP, steady state conditions were considered, and therefore the spectral parameters k_p , α_p , γ , and σ were only determined as a function of dimensionless fetch. The sea state will depend on both duration and fetch, however. In order to accommodate both circumstances, the spectral shape parameters are assumed to depend on the wave age

$$\chi = c_p/u_* \quad (9)$$

where u_* is the friction velocity and c_p is the phase speed of the peak of the wind sea spectrum, which in principle may be obtained from an ocean wave prediction model (e.g., the WAM model) [cf. *Komen et al.*, 1994]. Thus

$$\begin{aligned} k_p &= g/c_p^2 \\ \alpha_p &= A\chi^{-B} \\ \gamma &= \max [1, 1 + 3(1 - (0.038\chi)^2)] \\ \sigma &= 0.08 \end{aligned} \quad (10)$$

The parametrization of the Phillips parameter α_p was not obtained from JONSWAP, because in the JONSWAP fit for α_p , laboratory data were also used which, as is known from *Donelan et al.* [1985], belong to a different family. In a tuning exercise with the full VIERS model it was found that

$$\begin{aligned} A &= 0.24 \\ B &= 1 \end{aligned}$$

gave satisfactory results. This choice of parameters for the Phillips parameter is in fair agreement with the reanalysis of JONSWAP data performed by *Günther* [1981]. Furthermore, JONSWAP only considered young wind sea cases with a peak enhancement factor γ , which was on average 3.3. In order to assure that for old wind sea the JONSWAP spectrum asymptotes to the Pierson Moskowitz spectrum (hence $\gamma \rightarrow 1$), we have added the χ^2 factor in the expression for γ .

The JONSWAP spectrum is strictly speaking only valid for wave numbers up to 9 times the peak wave number. Recent observations of *Banner* [1990] confirm that up to a wave number of 30 rad/m the wave spectrum indeed follows a k^{-4} law, thus the region of validity of the JONSWAP spectrum may be extended to these high wave numbers. The present parameterization of the high wave number tail of the gravity wave spectrum differs, however, in one important aspect from *Banner's* fit to his observations. He chose a less sensitive dependence of the Phillips parameter on wave age (namely in his case $B = 0.5$), but the data contained only one young wind sea case. On the other hand, our parameterization is not in conflict with the data of *Jähne and Riemer* [1990], who found in the Delft wave tank a linear dependence of the gravity wave part of the spectrum on friction velocity which agrees with (10) with $B = 1$.

For wave numbers higher than k_{join} a new regime is entered because three wave interactions start to play a role in the steady state energy balance of the short waves (equation (7)). In the following we shall only develop a theory for the one-dimensional wave number spectrum, while the angular distribution of the short waves is modeled in a fairly simple fashion. The main reason for this is that we were unable to derive a reasonable parameterization of the angular dependence of the nonlinear interactions.

The one-dimensional wave number spectrum $F(k)$, which is related to the Fourier transform of the autocorrelation of the surface elevation, is normalized in such a way that $\int_0^\infty F(k)k dk = E$, where E is the wave variance. The wave energy density \mathcal{E} follows then from

$$\mathcal{E}(k) = \frac{\omega^2}{k} F(k) \quad (11)$$

where we shall only consider pure gravity-capillary waves with dispersion relation

$$\omega(k) = \sqrt{gk + Tk^3} \quad (12)$$

with g as acceleration of gravity and T as surface tension. Hence effects of current and shear in the current will be neglected.

Let us now describe some of the details of the source terms in the energy balance (7). For the input source term we adopt *Plant's* [1980] expression

$$S_{in} = \beta F \quad \beta = \delta \omega \left(\frac{u_*}{c} \right)^2 \quad (13)$$

where the dimensionless constant δ is given the value 0.03. The slowing down of wind by the short waves (the so-called quasi-linear effect) can be incorporated by renormalizing δ [cf. *Janssen et al.*, 1989; *Snoeij et al.*, 1993]. As a result, the wind input to the steep waves is reduced which has in practice a considerable impact on the backscatter which may be reduced by 5 dB.

Although *Donelan and Pierson* [1987] did not take effects of nonlinear three- and four-wave interactions into account, the work of *Valenzuela* [1978], *van Gastel* [1987], and *Janssen* [1987] suggests that three-wave interactions play an important role in the dynamical evolution of gravity-capillary waves, while *Kitaigorodskii* [1983] and *Phillips* [1985] stress the importance of four-wave interactions for short gravity waves. The exact

expressions for three- and four-wave interactions obtained by *Davidson* [1972], *Valenzuela* [1978], and *Hasselmann* [1962] will be used as a guideline to obtain an efficient parameterization of the nonlinear transfer.

Following *Kitaigorodskii* [1983], we assume that the nonlinear transfer is a local process in wave number space, and introducing the energy flux $\varepsilon(k)$, one thus has

$$S_{nl} = -\frac{1}{k} \frac{\partial}{\partial k} \varepsilon(k) \quad (14)$$

and on dimensional grounds the expression for ε reads

$$\varepsilon(k) = \frac{c^4}{v_g} [\alpha_3 B^2 + \alpha_4 B^3] \quad (15)$$

where v_g is the group velocity $\partial\omega/\partial k$, B is the angular average of the degree of saturation [*Phillips*, 1985],

$$B = k^4 F(k) \quad (16)$$

while α_3 and α_4 give the strength of the three- and four-wave interactions, respectively. The coefficients α_3 and α_4 may still depend on the ratio c/v_g . In particular, α_3 should vanish in the gravity wave regime because three-wave interactions are not possible there.

Three dissipative processes are assumed to play a role in the gravity-capillary regime, namely, viscous dissipation, wave breaking, and damping due to slicks. For viscous damping we use the exact expression [*Lamb*, 1932],

$$S_{visc} = -4\nu k^2 F \quad (17)$$

where ν is the kinematic viscosity of water.

Damping by slicks is caused by the Marangoni effect [*Alpers and Hühnerfuss*, 1989], which is the result of a resonant interaction between a sound wave in the surface film and short gravity waves. The Marangoni effect gives rise to an enhanced viscous damping

$$S_{slicks} + S_{visc} = -4\nu_{eff} k^2 F \quad (18)$$

where

$$\nu_{eff} = \nu M(k, \nu, \delta, E_s) \quad (19)$$

with M a relative damping ratio given by

$$M = \frac{1 + X(\cos \delta - \sin \delta) + XY + Y \sin \delta}{1 + 2X(\cos \delta - \sin \delta) + 2X^2} \quad (20)$$

δ is a phase angle and

$$X = \left| \frac{E_s}{\rho_w} \right| \frac{k^2}{\sqrt{2\nu\omega^3}} \quad Y = \left| \frac{E_s}{\rho_w} \right| \frac{k}{4\omega\nu} \quad (21)$$

Furthermore, E_s is the dilational modulus of the surface film, and ρ_w is the density of water. The surface film is determined by the two parameters δ and E_s . The phase angle δ is $\sim 180^\circ$, whereas E_s depends strongly on the type of slick. For a natural slick, mostly of biological origin, E_s may have the value of 0.01 N/m, whereas for chemical slicks its value may vary between 0.01 and 0.05 N/m.

Slicks may be destroyed, however, by the action of wind. We have modeled this by letting the dilational modulus vanish for strong enough winds, $E_s = 0.005[1 - \tanh(10u_* - 4.33)]$. In addition, since it is unrealistic that the ocean is covered by a single large slick, a second modification was implemented.

Since slicks come in patches, there is need for a fractional filling factor F [*Lombardini*, 1986]. With $F \leq 1$ the damping is modified according to $M_F = M/[M + F(1 - M)]$, where M is the damping when the coverage is complete. Typical F values are in the range 0.88–0.99.

It is remarked that in case of open ocean wind retrieval we have chosen to include an average effect of slicks for low wind speed. However, when comparing results from our spectral model with our laboratory data, the effect of slicks is switched off, because the experimentalists made sure that the water surface was clean so that no slicks were present. This was achieved by having an overflow at the end of the wind wave tank and by running the tank for a sufficiently long time so that after visual inspection, films had disappeared.

Individual breaking events are difficult to model because of strong nonlinearity. In a statistical description of wave evolution the whitecaps cover only a relatively small fraction of the surface, and whitecapping may therefore be regarded as a process which is weak in the mean. In work by *Komen et al.* [1994] it is then shown that the corresponding source term is quasi-linear; it consists of the spectrum at the wave number considered multiplied by a factor which is a function of the entire spectrum. Extending the *Komen et al.* [1984] expression for gravity wave dissipation into the gravity-capillary regime, we take

$$S_{br} = -\beta_d \bar{\omega} (\bar{k}^2 E)^2 (k/\bar{k}) F(k) \quad (22)$$

where β_d is a constant of the order 2, and $\bar{\omega}$ and \bar{k} are mean angular frequency and wave number, while E is the wave variance.

Combining now the explicit expressions for the source terms, the energy balance equation (7) becomes

$$\frac{\partial}{\partial k} \varepsilon(k) = \gamma \frac{\omega^2}{k^4} B \quad (23)$$

where the parameter γ is defined as

$$\gamma = \delta\omega \left(\frac{u_*}{c} \right)^2 - 4\nu M k^2 - \beta_d \bar{\omega} (\bar{k}^2 E)^2 (k/\bar{k}) \quad (24)$$

and hence gives the net effect of wind input and dissipation. The energy flux $\varepsilon(k)$ is given by

$$\varepsilon(k) = \frac{c^4}{v_g} (\alpha_3 B^2 + \alpha_4 B^3) \quad (25)$$

and we have eliminated the wave number spectrum F in favor of the degree of saturation $B = k^4 F$.

The interaction coefficient α_4 for four-wave interactions is taken as a constant, $\alpha_4 \approx 0.25$, while α_3 is allowed to depend on wave number because it is assumed that for gravity waves, three-wave interactions are not important. We take

$$\alpha_3 = \frac{3\pi}{16} \{ \tanh[\sigma_3(x-1)] + 1 \}$$

where $x = (k/k_{join})^{1/2}$ and $\sigma_3 = 2$. We remark that this choice of α_3 is, to a certain extent, arbitrary; however, a continuous transition from vanishing α_3 in the gravity range to a constant value in the gravity-capillary range is needed to avoid jumps in the spectrum.

By supplying the boundary condition at $k = k_{join}$ of continuity of flux (or spectrum) the differential equation (23) may be solved for the degree of saturation B , and the wave number

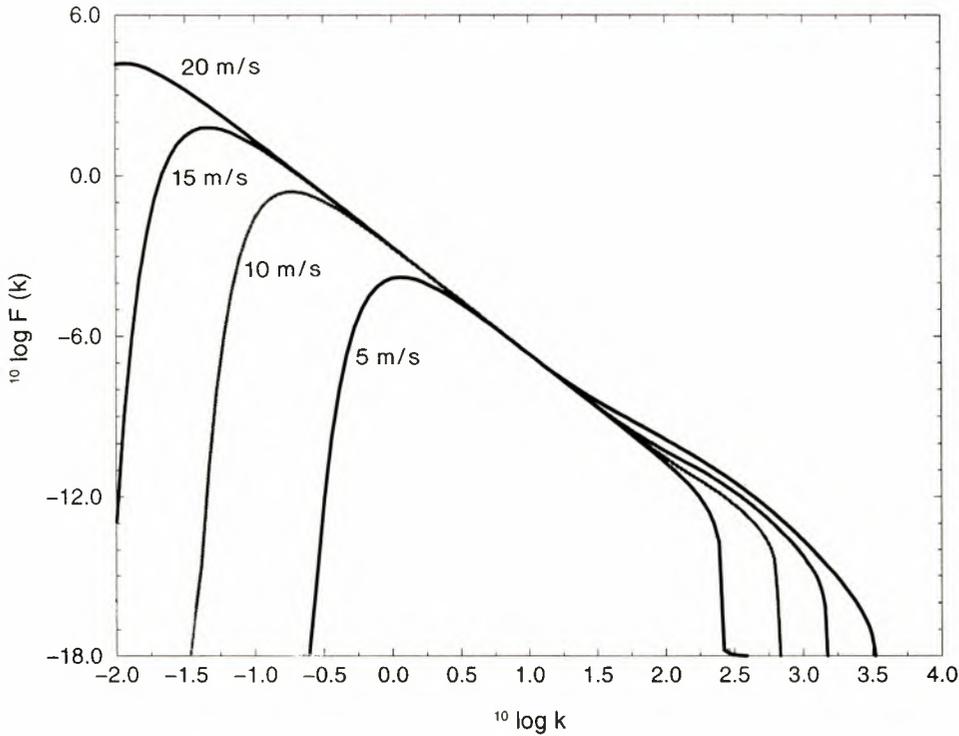


Figure 2. Wave number spectrum versus wave number for wind speeds of 5, 0, 15, and 20 m/s. The wave age parameter is 25.

spectrum $F(k)$ follows. In combination with the JONSWAP spectrum for $k < k_{\text{join}}$ the full one-dimensional wave number spectrum is obtained. Examples of the one-dimensional wave number spectrum according to the VIERS model equation (23) are shown in Figure 2 for four different friction velocities and old wind sea (wave age $\chi \approx 25$). The sensitive dependence of the high wave number part of the spectrum on friction velocity should be emphasized; this is, of course, the main reason why a scatterometer, which “observes” waves with wave numbers larger than 100, may be used as an instrument for measuring the wind field above the oceans.

In order to perform a successful wind retrieval the two-dimensional wave number spectrum is required. To that end we have taken a simple directional distribution $D(\phi)$

$$D(\phi) = \frac{1}{2\pi} [1 + 2a_2 \cos 2(\phi - \phi_w)] \quad (26)$$

where ϕ is the wave direction, ϕ_w is the wind direction, and a_2 is a parameter which measures the width of the directional distribution; a_2 is assumed to depend on friction velocity only and not on wave number. The two-dimensional wave number spectrum is then given by

$$W(k, \phi) = F(k)D(\phi) \quad (27)$$

and it will be used in a two-scale model to obtain the normalized backscatter.

We would like to discuss briefly some of the properties of the energy balance equation (23). In addition, modeled spectra are compared with observed spectra obtained in the Delft wave tank.

Since in practice the degree of saturation B is of the order 0.1 or less, it is a fair approximation to disregard four-wave interactions in the expression for the energy flux, (25). Retain-

ing therefore only three-wave interactions, the energy balance equation (23) may be solved, and the result for the degree of saturation becomes

$$B = \left(\frac{\nu_g}{\alpha_3} \right)^{1/2} c^{-2} \left\{ \varepsilon_0^{1/2} + \frac{1}{2\alpha_3^{1/2}} \int_{k_{\text{join}}}^k dk \frac{\gamma}{k^2} \sqrt{\nu_g} \right\} \quad (28)$$

where ε_0 is the value of the energy flux at $k = k_{\text{join}}$. It is of interest to discuss the respective terms in (28) separately. The first term is related to the effect of three-wave interactions. In the absence of wind input and dissipation it follows from the condition of a constant energy flux in wave number space. Using the dispersion relation for pure gravity-capillary waves (equation (12)), the degree of saturation according to the constant energy flux condition becomes

$$B_{3w} = \left(\frac{\varepsilon_0}{2\alpha_3} \right)^{1/2} c_0^{-3/2} \frac{y(1 + 3y^2)^{1/2}}{(1 + y^2)(y + y^3)^{1/4}} \quad (29)$$

where $y = k/k_0$, $k_0 = (g/T)^{1/2}$ is the wave number that separates gravity waves and capillary waves, and $c_0 = (gT)^{1/4}$. Therefore, in the gravity wave range ($k < k_0$), the degree of saturation increases with wave number like $k^{3/4}$ while, in the capillary wave range, B_{3w} decreases with wave number like $k^{-3/4}$ and B_{3w} attains its maximum value around $k = k_0$.

Effects of wind input and dissipation (γ) are represented by the second term in (28) and result in a modification of the “inertial” subrange spectrum given in (29). The degree of saturation now becomes a sensitive function of the friction velocity while, for large wave numbers, dissipation becomes important. For a large enough wave number the degree of saturation B will vanish. Let us call this particular wave number the cutoff wave number. In order to be able to compare with results from

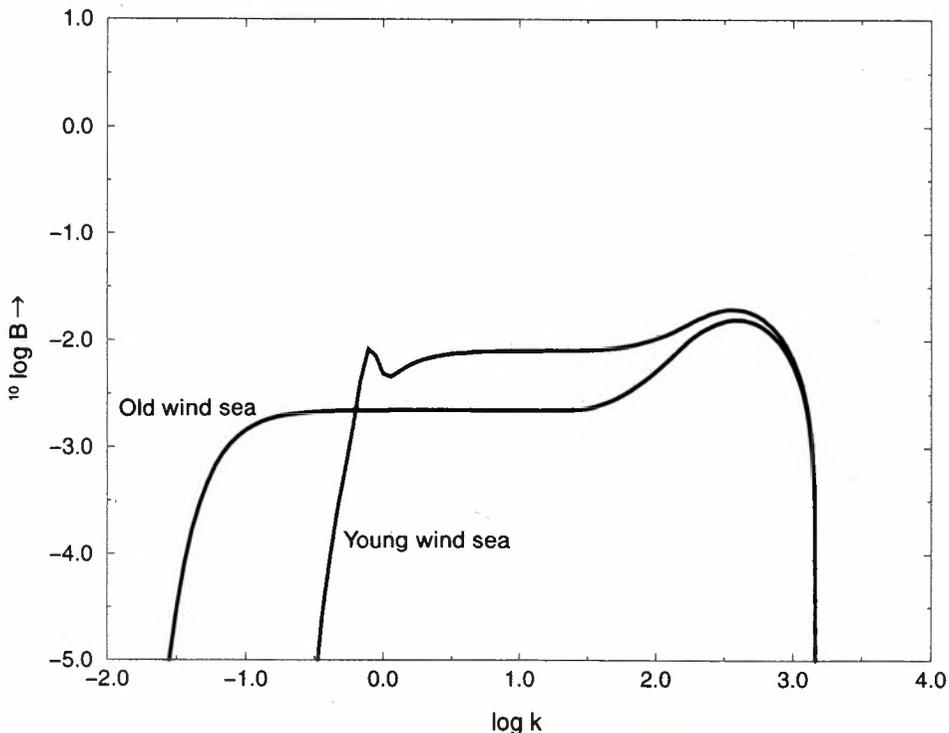


Figure 3. Comparison of degree of saturation $B(k)$ as function of wave number for young and old wind sea. The wind speed is 15 m/s.

the Donelan and Pierson [1987] model, we shall retain, in γ of (24), only the effects of wind input and viscous dissipation. Hence the cutoff wave number is determined by viscosity, and the viscous cutoff wave number in the Donelan and Pierson model follows from the condition $\gamma = 0$, or

$$\omega = \frac{1}{4} \frac{\delta u_*^2}{\nu} \quad (30)$$

In the present model, B does not depend on the local value of γ but depends on an integral in wave number space involving γ . As a consequence, the viscous cutoff wave number shifts to much larger values than given by (30). This shift in cutoff wave number is caused by the nonlinear energy transfer, which tries to maintain an inertial subrange spectrum. As a result, the present model therefore has a reduced sensitivity to changes in the water viscosity, at least in the wave number range that is relevant for scatterometry.

Furthermore, it is noted that ε_0 , which is determined by the JONSWAP spectrum (8), contains all the effects of sea state (i.e., wave age of long waves) on the short wave spectrum. As an illustration, we have compared in Figure 3 the degree of saturation B for young wind sea ($\chi \approx 7$) with old wind sea ($\chi \approx 25$), and the sea state dependence may be quite considerable, in particular in the low wave number range. From Figure 3 it is also noted that for young wind sea the increase of B in the high wave number range is less pronounced than in the case of old wind sea, which suggests that for young wind sea the short wave spectrum is controlled by nonlinear transfer because the short gravity waves are steeper.

We conclude this subsection by comparing results of the present short wave model with observed frequency spectra in the Delft wave tank. Frequency spectra were measured by means of a Lobemeier wire and a laser slope gauge (LSG) of

Jähne and Riemer [1990] for different friction velocities and fetches. Lobemeier spectra are thought to be reliable up to a frequency of 10 Hz, while LSG spectra are supposed to be valid to at least 100 Hz. If the Doppler shift due to the orbital motion of the long waves is ignored (this is a reasonable assumption in a wave tank), the modeled frequency spectrum $E(f)$ may be obtained from the wave number spectrum $F(k)$ according to

$$E(f) = \frac{2\pi k}{v_g} F(k)$$

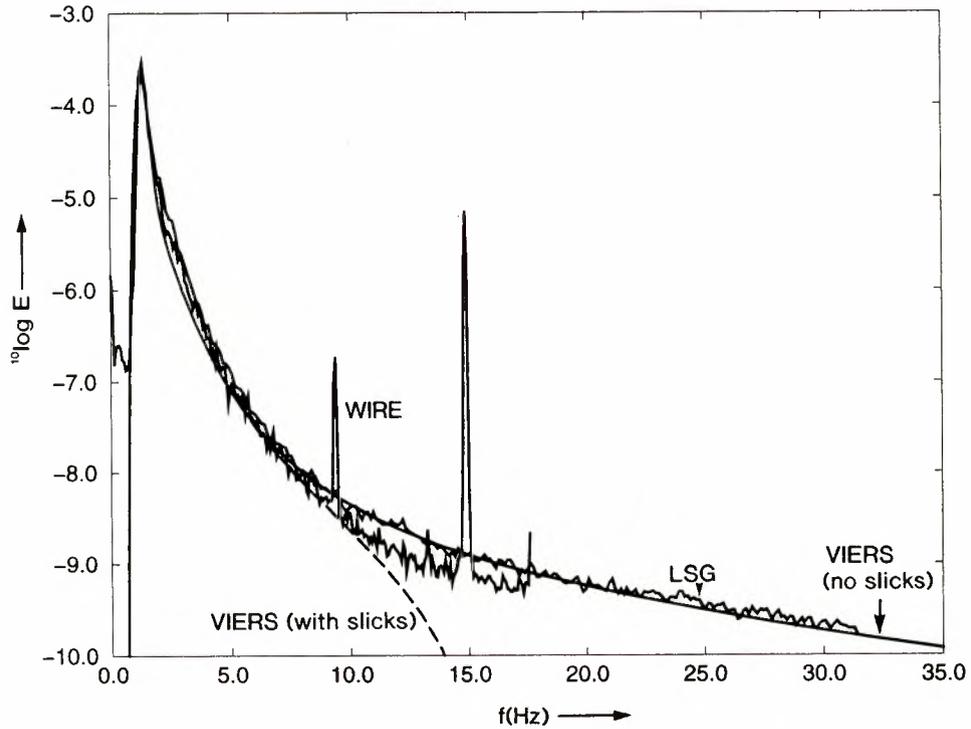
where the group velocity v_g is obtained for the dispersion relation of pure gravity-capillary waves (equation (12)). Examples of the comparison between observed and modeled frequency spectra are shown in Figure 4 for two different friction velocities and a fetch of 90 m. In view of the differences that do exist between the two types of observed spectra it may be concluded that the present short wave model shows a fair agreement with the observations. Furthermore, for comparison purposes we have plotted for the high friction velocity case the Donelan and Pierson [1987] short wave spectrum, and considerable differences with the observed wave tank spectra are found. The reason for this is that the Donelan and Pierson spectrum has an f^{-5} shape while the observed spectra in this frequency range have an f^{-4} shape. Finally, it is remarked that the water surface in the wave tank was clean, hence effects of slicks were disregarded. For low wind speed, slicks may have a dramatic impact on the spectral shape as is illustrated in the low wind speed case of Figure 4.

2.2. Radar Backscatter Model

Once the two-dimensional wave spectrum is known, the normalized backscatter may be obtained by means of a two-scale

Viers spectra compared to Observed Spectra

$u^* = 0.205$ m/s; effect of slicks is shown as well



Viers spectra compared to Observed Spectra

$u^* = 1.025$ m/s

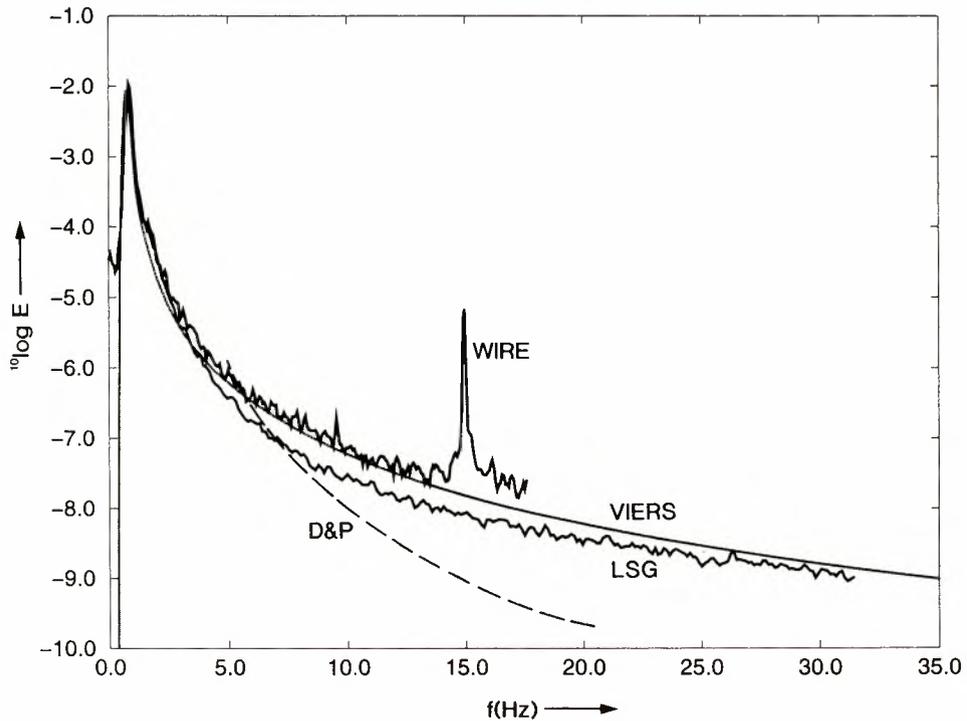


Figure 4. Comparison of simulated and laboratory frequency spectra for a friction velocity of 0.205 and 1.025 m/s, respectively, and a fetch of 90 m. For the low friction velocity case the impact of slicks on spectral shape is shown as well, while for the high friction velocity case we show the short wave spectrum according to *Donelan and Pierson* [1987] (D&P).

model. According to the wave-facet model [Plant, 1990] the normalized cross section is given by

$$\sigma_0 = \sigma_0^{\text{sp}} + \int_{-\infty}^{\infty} d(\tan \psi) \int_{-\infty}^{\infty} d(\tan \delta) P_B(\tan \psi, \tan \delta) \sigma_0^{\text{Br}}(\theta_i) \quad (31)$$

where P_B is the probability that a (Bragg) facet is oriented with tilts $\tan \psi$ and $\tan \delta$ along-wind and crosswind, respectively, while θ_i is the local incidence angle. For an anisotropic Gaussian surface one has

$$P_B = \frac{1}{2\pi s_{u,b} s_{c,b}} \exp\left(-\frac{\tan^2 \psi}{2s_{u,b}^2} - \frac{\tan^2 \delta}{2s_{c,b}^2}\right) \quad (32)$$

with $s_{u,b}^2$ and $s_{c,b}^2$ as the slope variances in upwind and crosswind. The Bragg contribution of a facet is proportional to the two-dimensional wave number spectrum at the Bragg wave number k_b (compare (1)). In fact,

$$\sigma_0^{\text{Br}} = 8\pi k_R^4 \cos^4 \theta_i |g_{\text{pol}}|^2 [W(\mathbf{k}_b) + W(-\mathbf{k}_b)] \quad (33)$$

with k_R as the radar wave number and g as a factor which depends on the polarization. The contribution due to specular reflection is given by

$$\sigma_0^{\text{sp}} = \pi |R(0)|^2 \sec^4 \theta P(\zeta_x, \zeta_y) \quad (34)$$

where P is the probability that a specular facet is oriented with tilts ζ_x and ζ_y parallel and at right angles to the radar look direction, respectively. For an anisotropic Gaussian surface one has

$$P = P(\zeta_x = \tan \theta, \zeta_y = 0) = \frac{1}{2\pi s_{u,s} s_{c,s}} \exp\left(-\frac{\tan^2 \theta}{2s_{L,s}^2}\right) \quad (35)$$

with $s_{u,s}^2$ and $s_{c,s}^2$ as the slope variances in the upwind and crosswind direction, whereas $s_{L,s}^2$ is the variance in the radar look direction. Only those waves that have a wavelength longer than the radar wavelength contribute to the slope variances as shorter waves are not seen by the radar [Stewart, 1985]. Furthermore, $|R(0)|^2$ is the reflection coefficient at normal incidence, which depends on the radar frequency via the relative dielectric constant ϵ_r ,

$$|R(0)| = |0.65(\epsilon_r - 1)/(\sqrt{\epsilon_r} + 1)|.$$

The factor 0.65 in this last equation is based on a correction of the standard reflection coefficient as specified by Valenzuela [1978]. The correction factor is needed because the remaining short wave disturbances of the water surface reduce the cross section as given by physical optics.

The above general two-scale theory has to be supplemented with a criterion to separate long waves from short ones. The wave number spectrum is separated into a low and high wave number part by means of the separation scale k_c

$$W_L(k, \phi) = \begin{cases} W(k, \phi) & k < k_c \\ 0 & k > k_c \end{cases} \quad (36)$$

$$W_H(k, \phi) = \begin{cases} 0 & k < k_c \\ W(k, \phi) & k > k_c \end{cases}$$

Hence, using the directional distribution (26), the slope variances of the tilting waves are given by

$$s_{u,b}^2 = \frac{1 + a_2}{2} \int_0^{k_c} k^3 F(k) dk \quad (37)$$

$$s_{c,b}^2 = \frac{1 - a_2}{2} \int_0^{k_c} k^3 F(k) dk$$

while the slope variance of the waves that contribute to specular reflection is given by $s_{L,s}^2 = s_{u,b}^2 + s_{c,b}^2$. Finally, the separation scale k_c is determined by the condition

$$\beta = 4k_R^2 \sigma_H^2 \quad \sigma_H^2 = \int_{k_c}^{\infty} F(k) dk \quad (38)$$

Condition (38) follows from the work of Bahar *et al.* [1983] and Brown [1978]. An optimal choice for the parameter β is then found to lie in the range 0.1–1. On the basis of a comparison with the VIERS data set [Snoeij *et al.*, 1993], $\beta = 0.13$ turns out to give optimal results for the normalized backscatter.

The present version of the two-scale model was tested against observed data obtained during the VIERS tank experiment at Delft Hydraulics. The radar operated at X band. Observed wave spectra were used as input to the backscatter algorithm. Figure 5 shows the normalized radar cross section σ (in decibels) as a function of incidence angle for vertical and horizontal polarization. The fetch was 90 m, and the friction velocity was $u_* = 0.367$ m/s. From Figure 5 it is concluded that there is a fair agreement between modeled and observed backscatter for vertical polarization but that the modeled backscatter is too low by as much as 5 dB for horizontal polarization. A similar poor performance at horizontal polarization was noted with Plant's [1990] composite surface model.

Finally, we show for C band the dependence of modeled backscatter on incidence angle for several wind speeds, where we used the parameterizations for the wind sea spectrum as appropriate for oceanic conditions. The sea state was assumed to be fully developed; for young wind seas the backscatter would increase by about 3 dB for incidence angles larger than 25°.

3. Inverse VIERS Model

The VIERS model consists of the three principal components discussed in section 2, namely, (1) a module to determine the stress for given wind, air-sea temperature difference and sea state; (2) a module to determine the short wave spectrum for given stress and sea state of the long wind waves; and (3) a module to obtain the normalized backscatter for a given two-dimensional wave spectrum. Therefore the VIERS model relates radar backscatter to wind vector, measurement geometry (e.g., incidence angle), sea state, air-sea temperature difference, and slicks. In most applications an averaged effect of slicks will be taken into account while effects of atmospheric stability on the stress will be disregarded. However, when studying some synoptic cases near the Norwegian coast, stratification may be important, and in section 4.2 it is shown that inclusion of stratification has a favorable impact on wind retrieval.

For the practical application of wind retrieval from (ERS-1) scatterometer data the model has to be inverted, however. In order to achieve this the following simple, straightforward procedure was adopted.

A table of normalized backscatter σ_0 is produced; the VIERS model is run for different incidence angles, wind vectors, and wave periods (or peak phase speeds) and the resulting σ_0 values are collected in a table. The wind parameters are wind

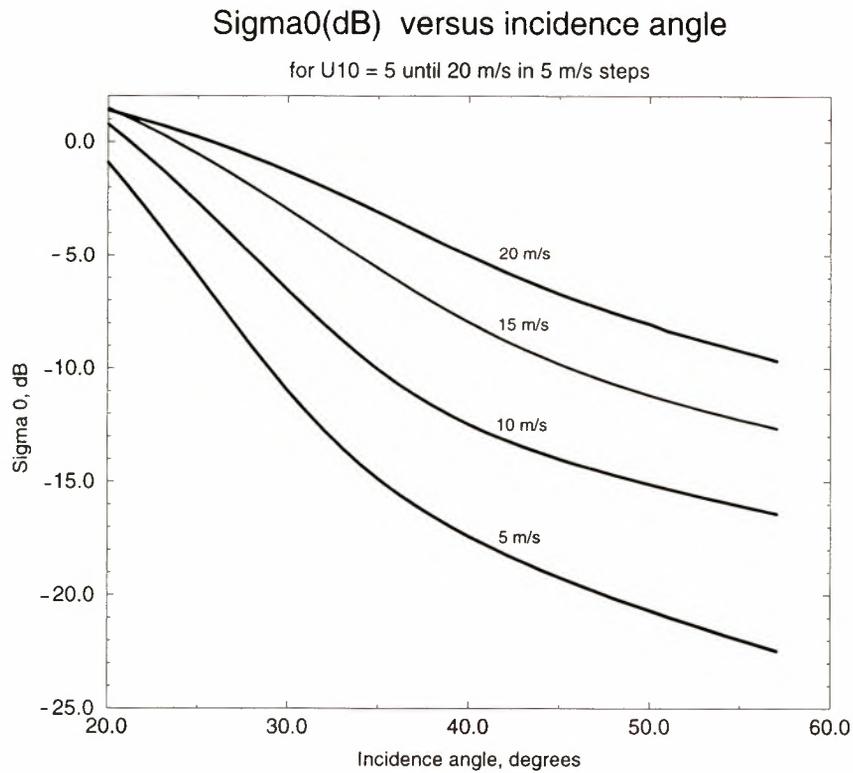
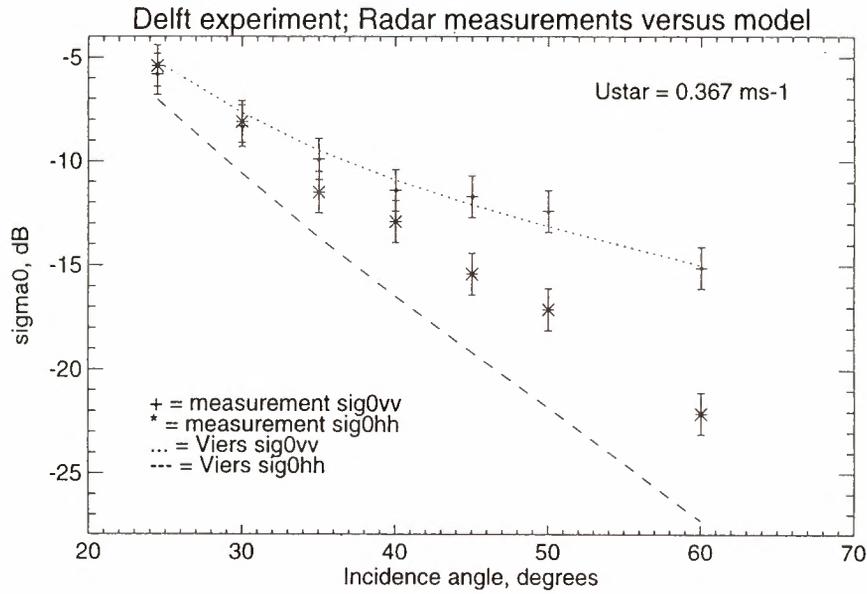


Figure 5. (top) Comparison of simulated and observed backscatter from the Delft wave tank as a function of incidence angle for a friction of 0.367 m/s, a fetch of 90 m, and two different polarizations. (bottom) Simulated backscatter as a function of incidence angle for different wind speeds for oceanic conditions.

speed U and the direction ϕ with respect to the look direction of the radar. If one accepts an accuracy of 1 m/s in retrieved wind speed and 15° in the wind direction, the incidence angle may be chosen in the range of 18° until 57° in steps of 1°, U from 1 until 30 m/s in steps of 1 m/s, and ϕ in steps of 15°, for instance.

The inversion procedure we adopted is specific for the ERS-1 configuration, where, for a certain cell i , three measurements of radar backscatter for different look angle and incidence angle were performed. The measured sigma triplet is denoted by $(\sigma_f, \sigma_m, \sigma_a)$, where the subscripts denote fore,

middle, and aft beam, respectively. The wind retrieval procedure is then as follows.

1. Determine incidence angles of fore, middle, and aft beam: $\theta_f(i), \theta_m(i), \theta_a(i)$.
2. Calculate the corresponding model triplets for all tabulated wind vectors according to

$$\begin{aligned} \sigma_f^{\text{mod}} &= \sigma^{\text{tab}}(\theta_f, U, \phi + 45, c_p) \\ \sigma_m^{\text{mod}} &= \sigma^{\text{tab}}(\theta_m, U, \phi, c_p) \\ \sigma_R^{\text{mod}} &= \sigma^{\text{tab}}(\theta_a, U, \phi - 45, c_p) \end{aligned} \quad (39)$$

Here ϕ is the wind direction with respect to the midbeam, and c_p is the wind sea phase speed obtained from a wave prediction model (e.g., the WAM model [Komen *et al.*, 1994]).

3. Determine the normalized quadratic distance between modeled and measured triplets,

$$Q_{\text{ERS-1}} = \sum_{n=1}^{n_b} Q_n \quad (40)$$

where the index n refers to the beam ($n_b = 3$) and

$$Q_n = \left(\frac{\sigma_n - \sigma_n^{\text{mod}}}{k_p \sigma_n} \right)^2 \quad (41)$$

Here k_p is the relative accuracy of the measurement (of the order of 5%), and $k_p \sigma_n$ is the measurement error in σ .

4. Determine the normalized quadratic distance between retrieved and ECMWF wind (both magnitude and direction)

$$Q_{\text{GEO}} = \left(\frac{U - U_{\text{GEO}}}{\Delta U_{\text{GEO}}} \right)^2 + \left(\frac{\phi + \chi_m - \phi_{\text{GEO}}}{\Delta \phi_{\text{GEO}}} \right)^2 \quad (42)$$

where χ_m is the look direction of the midbeam with respect to north. The errors in the ECMWF wind fields are estimated to be $\Delta U_{\text{GEO}} = 2$ m/s and $\Delta \phi_{\text{GEO}} = 20^\circ$.

5. Determine for all 30×24 tabulated wind vectors the cost function D

$$D = \sqrt{Q_{\text{ERS-1}} + Q_{\text{GEO}}} \quad (43)$$

and infer its absolute minimum.

6. The model wind that minimizes D is called the retrieved wind. The wind direction with respect to true north, ϕ_w , is given by $\phi_w = \chi_m + \phi_{\text{ret}}$, where ϕ_{ret} is the retrieved wind direction with respect to the midbeam.

The fourth step is inserted in order to remove the ambiguity in direction of 180° . This is a well-known problem in scatterometry and is evident from expression (33) of the Bragg contribution to the radar backscatter. Thus the ambiguity problem is removed by step 4, but one may introduce a spurious interdependence between retrieved and ECMWF winds.

The choice of using tabulated values of backscatter to evaluate the cost function has certain advantages. The modern approach to inversion would be to minimize the cost function using the adjoint of the forward model. However, we thought that it was far too much effort to write the adjoint of the model, while from experience it was known that the forward VIERS-1 model was too expensive to run in an operational context. The introduction of a table is far less expensive. In addition, the search for a minimum of the cost function is straightforward as the cost function may be evaluated for all possible wind speeds and directions. The added advantage is that there is no need to write an adjoint of the VIERS-1 model.

Before we discuss, in the next section, results for wind retrieval with the VIERS algorithm, it is of importance to briefly comment on the tuning procedure we followed. Several parameters in the short wave model are not fixed a priori, the most important ones being k_{join} , the directional width parameter a_2 , and the Phillips parameter α_p . However, it should be emphasized that there are empirical guidelines for the choices of a_2 and α_p . Nevertheless, the model output depends critically on the precise choice of k_{join} , a_2 as a function of u_* , and α_p as a function of wave age. These three variables were the basic tuning parameters.

Initially, we tuned the "forward" VIERS algorithm. Thus the

simulated backscatter, obtained using ECMWF winds and WAM model periods, was compared with the backscatter as observed by ERS-1. After some tuning a reasonable agreement between simulated and observed backscatter was obtained. We typically found a standard deviation of error of about 2 dB in σ_0 , which in view of the limited knowledge of the spectrum of short waves and in view of the accuracy of the analyzed wind field (which we assume to be 2 m/s) is already quite an achievement. We were therefore quite optimistic that the thus obtained algorithm would be successful in retrieving winds from the observed radar backscatter. Unfortunately, this turned out not to be the case, and some additional tuning was required to obtain reliable winds. The main reason for the additional tuning is that we needed an accuracy of the model algorithm of at least 0.5 dB, which cannot be achieved using analyzed winds with a relatively large error in magnitude (± 2 m/s) and direction ($\pm 15^\circ$). The quality of the analyzed winds obtained from the ECMWF analysis and forecasting system has been studied extensively by comparing modeled and observed wind speed from buoys over a 1-year period [Janssen *et al.*, 1997]. As a result, it is found that the ECMWF winds have on average a rms error of about 2 m/s, and this estimate of the wind speed error has been used in the cost function (42).

It was therefore decided to use the inverse of the VIERS model to do the tuning. To that end, about 30,000 σ_0 triplets, measured with the ERS-1 scatterometer (operating at C band) on November 6, 1991, together with collocated periods from the WAM model, were supplied to the inverted VIERS model. The resulting retrieved winds (magnitude and direction) were plotted against collocated analyzed winds obtained from the ECMWF atmospheric model. The tuning parameters k_{join} , a_2 , and α_p were chosen in such a way that the average bias and scatter index (between VIERS and ECMWF winds) were as low as possible and the spectra from the short wave model were of the best quality. This approach ensures that the VIERS and ECMWF winds are compatible in a statistical sense, while also reasonable choices for the tuning parameters have been obtained. For example, the wave age dependence of the Phillips parameter, as given in (10), is in fair agreement with Günther's [1981] reanalysis of the JONSWAP data.

The tuned VIERS algorithm has been used to produce the plots depicted in Figures 6 and 7. In order to visualize the density of points, contour lines of equal density (number of points per square m/s cq°) are drawn. As a reference, we have produced the same plots with the CMOD4 model, using the same inversion technique. It can be seen that the VIERS is well tuned in the sense that it produces winds that are compatible with the analyzed ECMWF winds.

4. Wind Retrieval With the VIERS Algorithm

In this section we would like to present our results for wind retrieval with the VIERS algorithm. Results are compared in detail with analyzed wind fields from ECMWF and with retrieved winds from the CMOD4 algorithm. Two approaches are followed. In section 4.1 we shall use statistical tools to compare results, while in section 4.2 we give a comparison of results with emphasis on synoptic situations. It is felt that these two approaches are to some extent complementary, and they will highlight the strong and weak points of the VIERS algorithm.

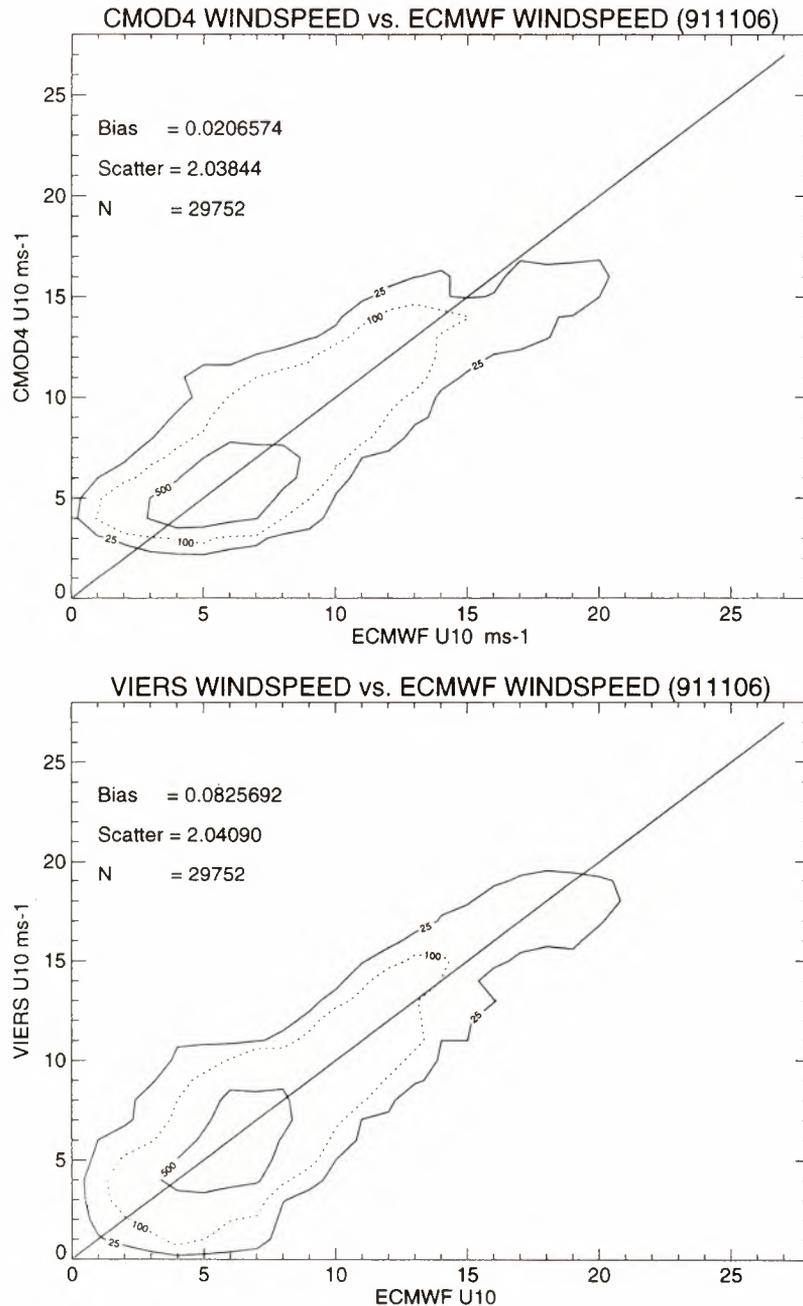


Figure 6. Retrieved wind speed using (top) CMOD4 and (bottom) VIERS algorithm versus analyzed ECMWF wind speed on November 6, 1991.

4.1. Statistical Comparison

We have applied the wind retrieval algorithms of VIERS and CMOD4 to three cases on November 6 and 7, 1991, and March 10, 1992, all on 1200 UT. To that end we collocated the σ_0 triplets, as measured by the ERS-1 scatterometer with wave periods of the WAM model and with analyzed wind fields from the ECMWF atmospheric model.

As a first result, we compare retrieved wind speed and direction from VIERS with the analyzed ECMWF winds. The comparison for the three dates is shown in Figures 6–11. As a reference, the same plots are produced with CMOD4 using the same inversion technique as VIERS. As already discussed, we have performed fine tuning of the VIERS algorithm on the November 6 case. The results of the two other dates show that

the tuning procedure was robust. Although the standard deviation of error on these last two dates has increased somewhat, it should be noted that for CMOD4 a similar remark applies.

Statistical parameters for the three dates are summarized in Table 1. Regarding the wind vector, the statistics of VIERS and CMOD4 are comparable, with VIERS having slightly better directional properties. However, as may be inferred from Figures 6, 8, and 10, CMOD4 does not allow wind speeds below 2 m/s, and this may contribute to more favorable statistics. Furthermore, CMOD4 seems to underestimate the wind speed for high winds. In order to see this point more clearly we have restricted the determination of the statistical parameters to those cases where the analyzed ECMWF wind was higher than 15 m/s. Results are given in Table 2.

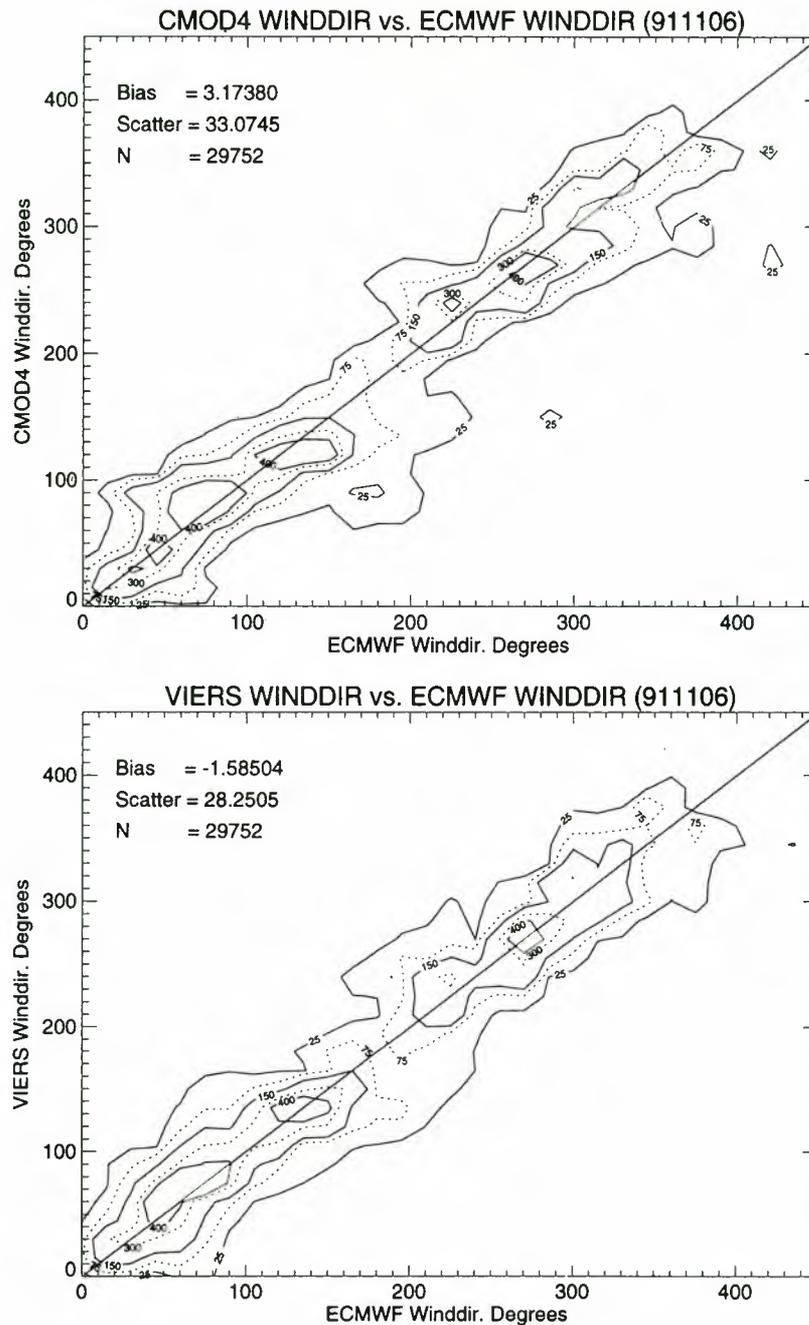


Figure 7. Same as Figure 6, but for wind direction.

The statistics in Table 2 show that both retrieval algorithms are biased low but that CMOD4 clearly underestimates the wind speed. It is emphasized that an underestimation of wind speed at high winds is an undesirable property of a retrieval algorithm. As shown by *Gaffard and Roquet [1995]*, when used in an atmospheric data assimilation scheme, the retrieved winds could result in a considerable slowing down of the major storm systems. Of course, our conclusion on the weak performance of CMOD4 at high winds depends on the quality of the analyzed ECMWF winds. However, *Gaffard and Roquet [1995]* also compared CMOD4 wind speeds with quality-controlled buoy wind measurements over a 2-year period. The data set was provided by Météo-France and consisted of buoy reports received through the Global Telecommunications System, which are closer than 100 km in space and 3 hours in time to

scatterometer measurements. CMOD4 was found to overestimate wind speeds in the low wind speed range by about 1 m/s, while in the high wind speed range, CMOD4 underestimated wind speed by as much as 2 m/s or even larger. On the basis of the comparison between CMOD4 and the buoy observations, *Gaffard and Roquet [1995]* applied a wind speed dependent bias correction to the wind speeds retrieved by CMOD4, and, in comparison with the ECMWF first-guess winds, hardly any bias was found in the wind speed range up to 20 m/s. When the corrected CMOD4 winds were used in ECMWF's analysis system, an improved agreement between radar altimeter wind speeds and analyzed wind speed was found, while also the forecast showed improvements.

It is therefore concluded that for high wind speeds the VIERS algorithm performs better than CMOD4. A similar re-

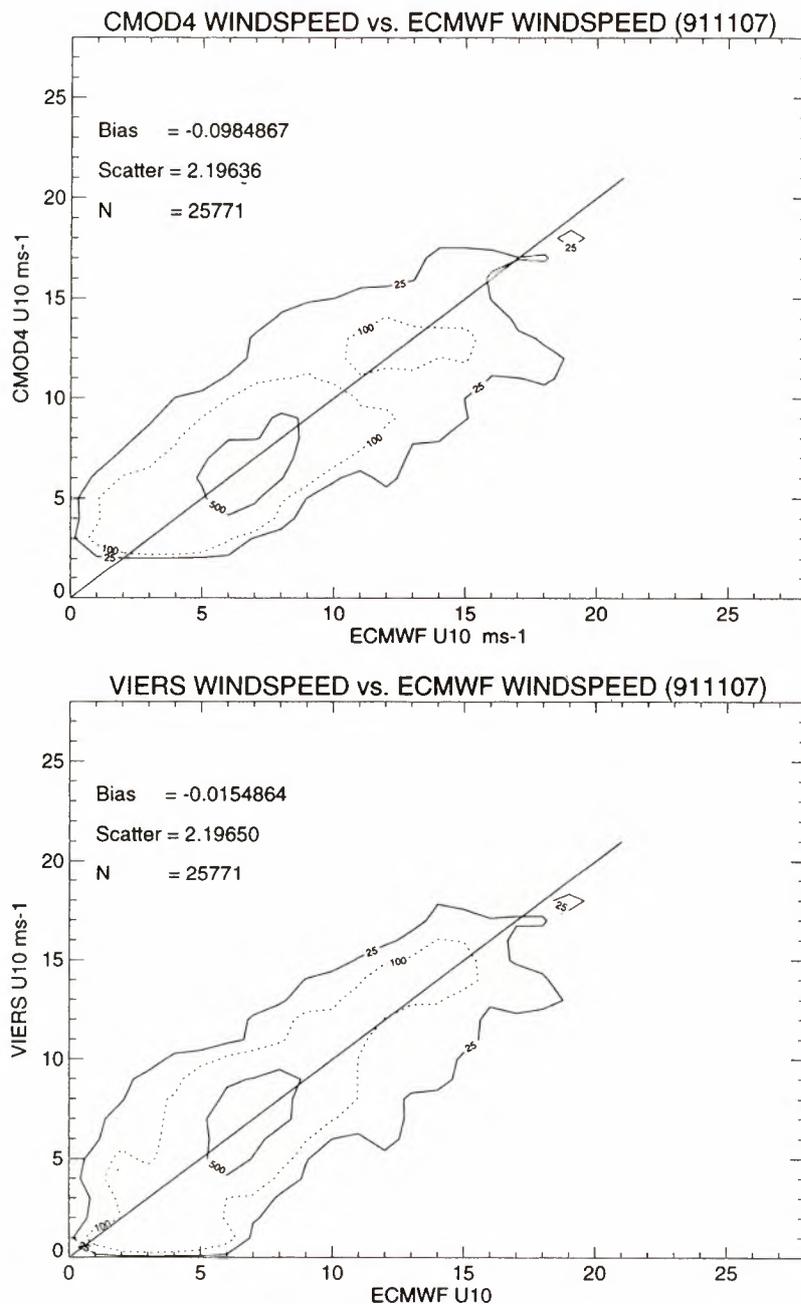


Figure 8. Same as Figure 6, but for November 7, 1991.

mark applies to the low wind speed cases. A summary of the difference in wind retrieval of VIERS and CMOD4 is given in Figure 12. The differences at low and high wind speed confirm the picture we have sketched above. Finally, H. Roquet (private communication, 1995) compared retrieved VIERS winds with the buoy data and found a good agreement, in particular at high wind speeds.

It is emphasized that high wind cases usually correspond to young wind sea because the timescale to reach equilibrium condition is proportional to wind speed. One of the reasons to develop the VIERS algorithm was that it was expected that the radar backscatter depends on the sea state. Young wind waves are usually steeper than old wind waves, and therefore for the same wind speed a larger backscatter would result (compare Figure 3). However, if one would not take the sea state de-

pendence of the radar backscatter into account (by taking, for example, a fixed wave age $c_p/u_* = 35$), then the short waves would be less steep, giving for the same wind a smaller backscatter. As a consequence, with the same observed backscatter one would expect larger winds in a sea state independent algorithm. This turns out to be the case. We reran the VIERS algorithm in sea state independent mode by fixing the wave age c_p/u_* to a constant value, $c_p/u_* = 35$. We took the period of November 6, 1991, and we restricted the wind retrieval to those cases where the ECMWF wind speed is larger than 15 m/s. When comparing the thus obtained retrieved winds with analyzed winds, we found, as expected, a positive bias of 1.64 m/s while the standard deviation of error was 2.74 m/s, which is considerably larger than obtained from the sea state dependent version of VIERS (compare Table 2). It is concluded from this

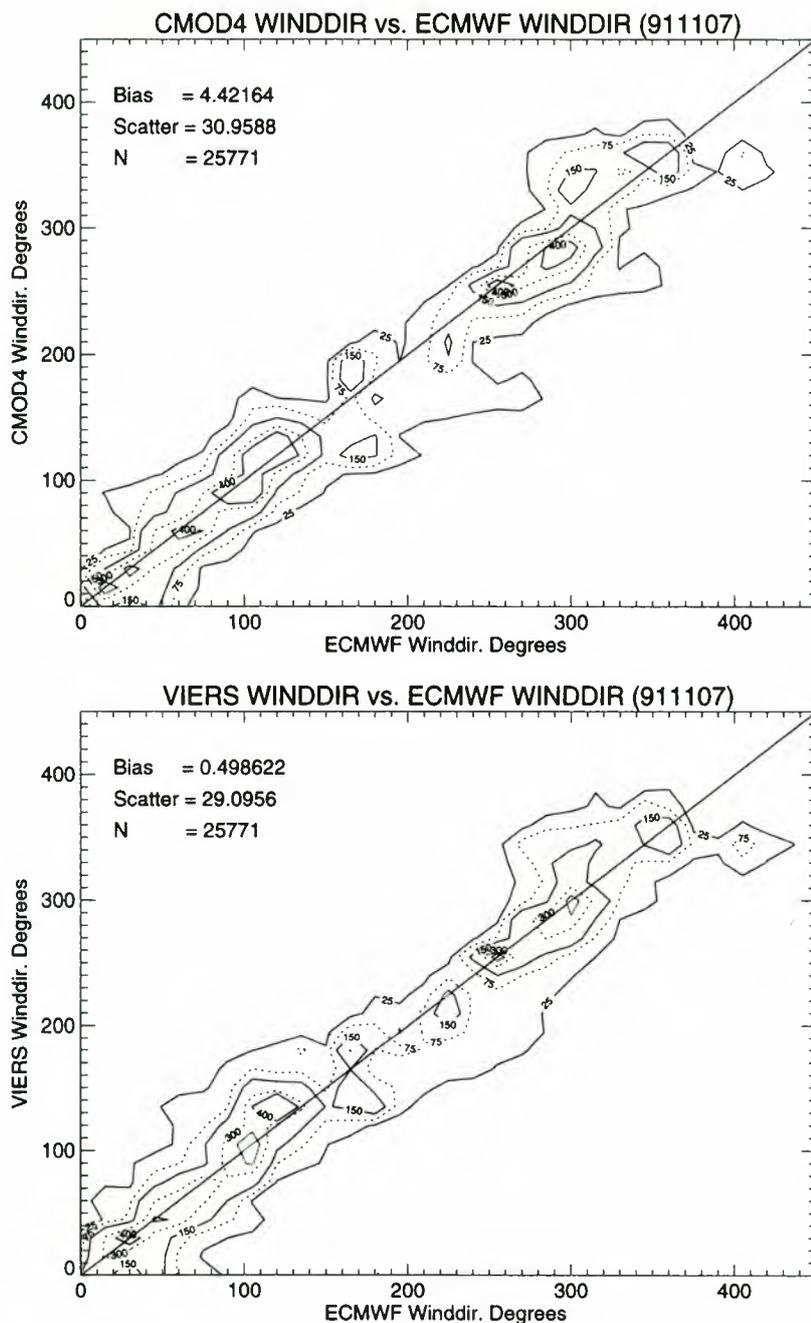


Figure 9. Same as Figure 7, but for November 7, 1991.

comparison that the sea state dependence of radar backscatter has a considerable impact on wind retrieval under a limited range of conditions. It gives rise in a change of bias of 2.5 m/s. Moreover, in view of the smaller standard deviation of error, we conclude that a sea state dependent backscatter algorithm is to be preferred.

In order to finish our discussion on the performance of the VIERS algorithm we finally concentrate on its properties in the so-called σ space. The σ space is the space spanned by the radar backscatter of fore, middle, and aft beam. Let us introduce the distance D_σ in σ space as

$$D_\sigma = \sqrt{Q_{ERS-1}} \quad (44)$$

where Q_{ERS-1} is given by (40) then, ideally, a perfect model should have a distance which is as small as possible, i.e., $D_\sigma =$

0. There are two reasons why in practice D_σ attains a finite value. The first reason is finite measurement errors. Assuming that there is no bias between model and observation, $\langle \sigma_{\text{obs}} - \sigma_{\text{mod}} \rangle = 0$, and assuming that the backscatter model is perfect, one obtains, using (41),

$$\langle Q_n \rangle = 1$$

and therefore the minimal distance in σ space becomes

$$D_\sigma = \sqrt{3}$$

Assuming, in addition, that the variable $\delta = (\sigma_{\text{obs}} - \sigma_{\text{mod}})/k_p \sigma_{\text{obs}}$ is a Gaussian variable, then for a perfect model the distribution of the distance D_σ can be calculated. Thus the statistics of D_σ are determined by three independent Gaussian

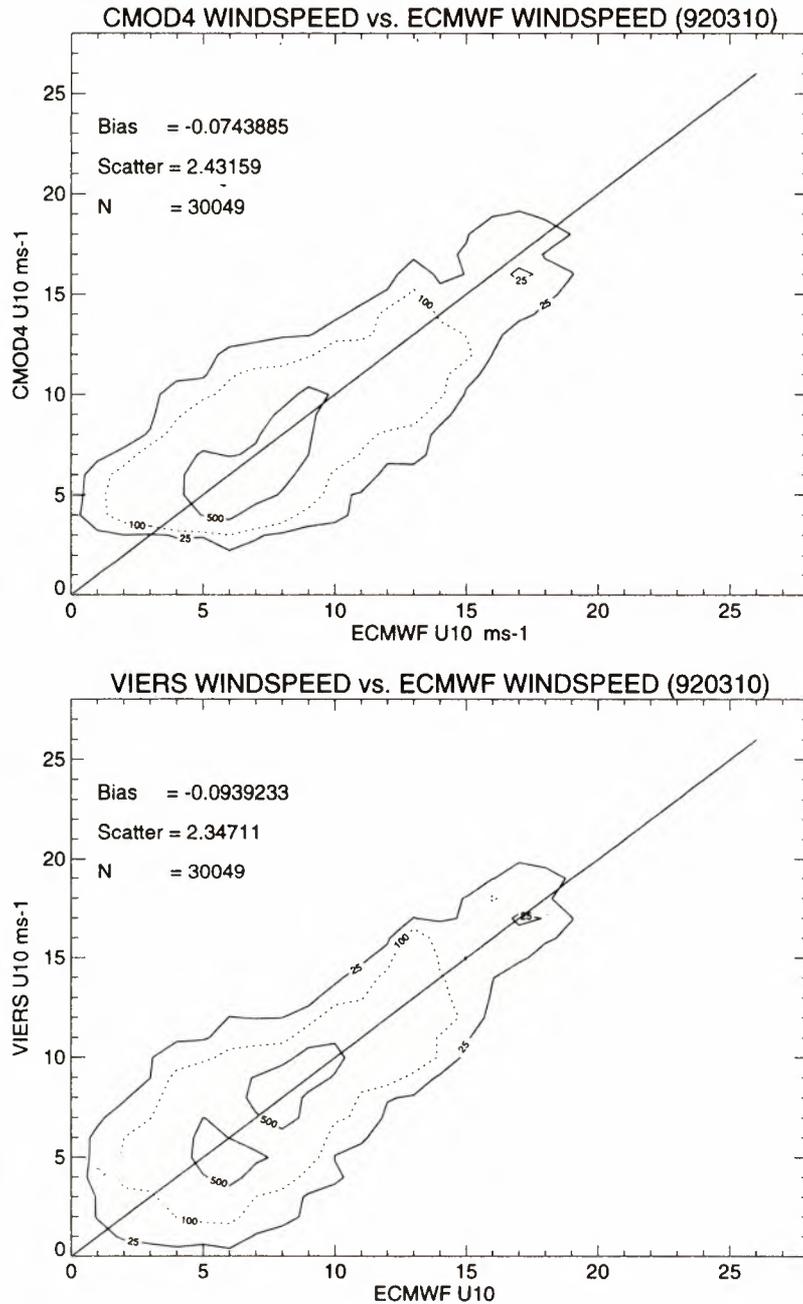


Figure 10. Same as Figure 6, but for March 10, 1992.

variables δ_f , δ_m , and δ_a with $\langle \delta \rangle = 0$ and $\langle \delta^2 \rangle = 1$. In that event the distribution of D_σ^2 is chi-square with 3 degrees of freedom. This result is valid if the assumption of independent Gaussian variables is justified and if the model at hand is perfect.

In practice, the actual distribution may deviate from the theoretical one, however. The discrepancy is caused by random model errors (assuming that all systematic errors have been eliminated) that broaden the distribution of δ . As a result, in practice, the mean distance $\langle D_\sigma \rangle$ may be larger than $\sqrt{3}$, and the distribution of D_σ^2 may be different from the chi-square distribution. In Tables 1 and 2 we show the mean values of D_σ for VIERS and CMOD4, and evidently, CMOD4 fits the observed backscatter more closely. This conclusion is supported by Figure 13, where we have plotted the distribution of D_σ for VIERS and the CMOD4 model. The period was November 6,

1991. The distribution for a perfect model is shown as well. We note from Figure 13 that CMOD4 has a more narrow distribution than VIERS, but both model distributions deviate considerably from the one of a perfect model.

Assuming that the model function describes reality in a reasonable manner, it is even possible to infer the rms error of the retrieved wind speed from the misfit in σ space. Of course, a misfit in σ space will induce an error in both wind speed and direction. It is known, however, that the mean of the backscatter from fore and aft beam

$$x = \frac{1}{2}(\sigma_a + \sigma_f)$$

is to a good approximation independent of the azimuth angle. This readily follows, assuming Bragg scattering, from the di-

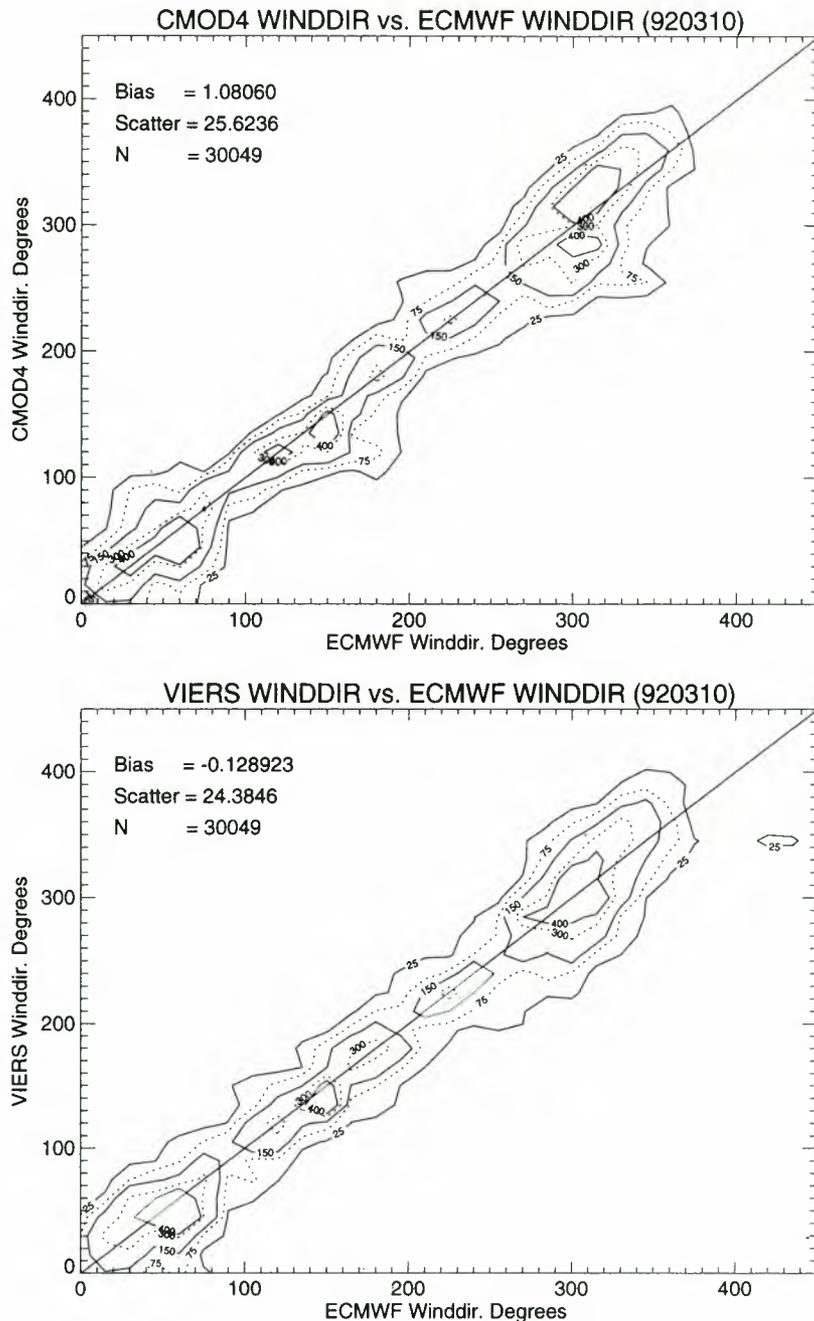


Figure 11. Same as Figure 7, but for March 10, 1992.

rectional wave spectrum, given by (27), which involves a $\cos(2\theta)$, where θ is the difference between azimuth angle and wind angle. Since the azimuth angle for the fore and aft beam is 90° apart, it follows that the sum of fore and aft beam backscatter is independent of azimuth angle. The error in wind speed then immediately follows from

$$\delta x = \frac{\partial x}{\partial U_{10}} \delta U_{10}$$

where δx is the difference between modeled and observed mean backscatter and the derivative of x with respect to U_{10} can be obtained at the minimum distance D by finite differencing. By averaging the square of the error δU_{10} over all retrievals, the overall rms error in wind speed σ_R may be determined according to

$$\sigma_R = \sqrt{\langle \delta U_{10}^2 \rangle}$$

Results of this calculation are shown in Tables 1 and 2. According to this estimate, VIERS has an rms error in wind speed of about 0.75 m/s, and CMOD4 has an rms error of 0.5 m/s, while for the high wind speed cases of Table 2 we get 1 m/s and 0.6 m/s, respectively.

All in all, it is difficult to decide which algorithm is better. On the one hand, CMOD4 has a smaller rms error in wind speed because the misfit in σ space is smaller than for VIERS. On the other hand, when compared to buoy observations and ECMWF analyses, CMOD4 underestimates wind speed considerably while VIERS has less problems in that respect. We therefore conclude that the VIERS model is an acceptable model to retrieve the wind vector from radar backscatter measurements.

Table 1. Statistical Comparison of VIERS and CMOD4 Winds Against ECMWF Analyzed Winds

	November 6, 1991		November 7, 1991		March 10, 1992	
	VIERS	CMOD4	VIERS	CMOD4	VIERS	CMOD4
Number	29752	29752	25771	25771	30049	30049
μ_u , m/s	-0.08	-0.02	0.02	0.10	0.10	0.07
σ_u , m/s	2.04	2.04	2.20	2.20	2.35	2.43
μ_ϕ , deg	1.6	-3.2	-0.5	-4.4	0	-1
σ_ϕ , deg	28	33	29	31	24	26
D_σ	5.2	3.8	5.9	3.9	5.4	3.7
σ_R , m/s	0.75	0.47	0.80	0.47	0.69	0.49

Here, μ refers to the bias and σ is the standard deviation. Also, the distance between modeled and observed backscatter is given, as well as the anticipated error in wind speed caused by the misfit in σ_0 space.

Although the VIERS model seems to perform in a reasonable manner, it is still of interest to discuss possible reasons for the larger misfit in σ space. An important factor could be the choice of the directional distribution of waves. In VIERS (compare (26)) we use a rather simple direction spectrum with a friction velocity dependent width. From observations it is known that the width also depends on the ratio of wave number to peak wave number of the spectrum [Donelan et al., 1985]. In addition, Jähne and Riemer [1990] have observational evidence for a bimodal distribution. In order to see to what extent the directional distribution of the waves plays a role in the wind retrieval, it was decided to do a retrieval experiment using only the fore and aft beam, since the mean of fore and aft beam backscatter is approximately independent of the directional distribution. A much better fit of the VIERS model to the observed backscatter was obtained in this manner. The resulting wind speed error obtained from the misfit in σ space now becomes only 0.5 m/s. Since a retrieval with three beams gives a larger rms error of 75 cm/s, this suggests that our choice of directional distribution of the waves is not optimal. After the VIERS project was finished, Janssen and Wallbrink [1997] improved the directional distribution and were able to obtain a misfit in σ space that was similar to the one of the CMOD4 algorithm, but the quality of the wind retrieval product remained the same.

4.2. Synoptic Validation Using Cal/Val Data

In this section we shall discuss in some detail results of another method of validating the retrieved wind fields obtained with the VIERS model; namely, we compare wind fields from VIERS with those of a meteorological model and CMOD4. Although this synoptic validation is only qualitative,

it has a certain number of advantages over a statistical validation.

1. One can easily verify by eye the internal consistency of the wind speed and directions of adjacent scatterometer cells; in addition, it is fairly straightforward to identify ambiguity errors and incidence angle dependent problems in the algorithm.

2. One can directly compare the structure of the wind fields of ERS-1 derived wind fields and model derived wind fields.

4.2.1. Calibration and validation campaign. When the European Space Agency (ESA) distributed an announcement of opportunity for the calibration and validation of the sensors and products of ERS-1, the VIERS group submitted a proposal for the validation of the wind scatterometer product. This proposal was granted by ESA and access was given to the calibration and validation data acquired during the Cal/Val campaign in the Norwegian part of the North Sea and the Atlantic Ocean between 5°W and 10°E and 60° and 70°N in 1991. ESA was the initiator of this large campaign in which information on the ocean and weather conditions was acquired during overpasses of the ERS-1 satellite. At that time the satellite was in a 3-day repeat period orbit, which had a scatterometer crossover point west of Norway.

The data acquired from the in situ sensors and other sensors were used together with those of the Norwegian meteorological model (hereinafter referred to as METEO) to provide the best possible estimate of the wind field over the scatterometer swath during the passage of the satellite. Besides the ERS-1 measured triplets of the radar scattering at the ocean surface the VIERS model needs to have the peak frequency of the

Table 2. Same as Table 1, but Under the Restriction of ECMWF Wind Speed Faster Than 15 m/s

	November 6, 1991		November 7, 1991		March 10, 1992	
	VIERS	CMOD4	VIERS	CMOD4	VIERS	CMOD4
Number	1324	1324	1124	1124	1237	1237
μ_u , m/s	-0.89	-2.56	-2.18	-3.04	-0.14	-1.26
σ_u , m/s	2.02	2.02	2.13	2.51	2.55	2.44
μ_ϕ , deg	-3.2	-4.8	-6.2	-6.8	4.9	4.2
σ_ϕ , deg	12.3	11.5	14.8	15.4	15.5	15.0
D_σ	5.1	3.6	4.5	3.0	4.7	3.6
σ_R , m/s	1.14	0.60	0.84	0.52	0.98	0.71

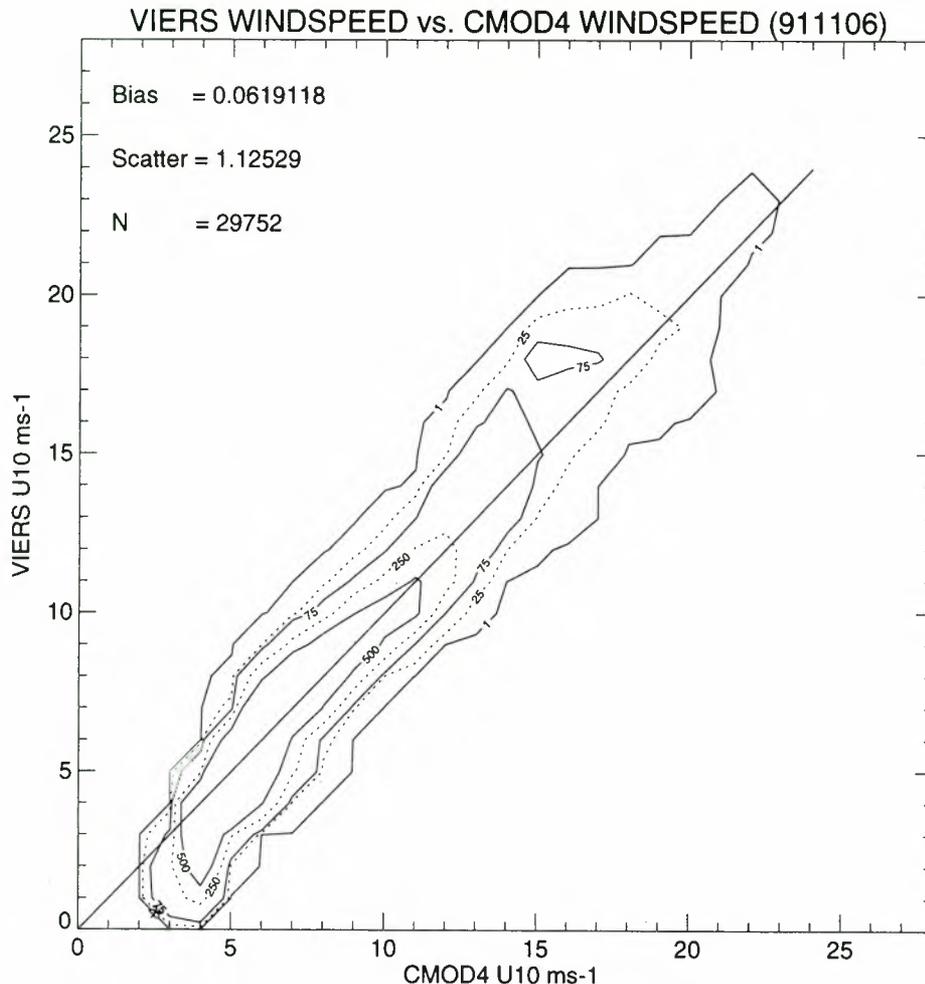


Figure 12. Comparison of VIERS and CMOD4 wind speed for November 6, 1991.

wind sea part of the dominant waves as input. This parameter was obtained from output of the operational WAM model at the Royal Netherlands Meteorological Institute (KNMI), which was kindly provided by J. Onvlee (personal communication, 1993).

Retrieved VIERS winds were then generated by running the inverse VIERS model using the collocated METEO winds as side condition. In a similar fashion, CMOD4 retrieved winds were obtained. The resulting winds were imaged on a plane tangential to the earth at 65°N and 5°E . An example is given in Plate 1. Here wind speed is coded by a color scale, where the scale ranges from 0 to 24 m/s while the arrows in the plot indicate the flow direction.

Table 3 presents an overview of the data used. It lists mean wind speeds and the differences between model and ERS-1 derived winds, as well as the differences between VIERS and CMOD4. Assuming neutral conditions, the average difference between VIERS and METEO is 0.8 m/s, and between CMOD4 and METEO the average difference is 0.3 m/s. The standard deviation of the difference between VIERS and METEO is 2.3 m/s and is 2.0 m/s for CMOD4 and METEO. If unstable conditions are assumed (we took a fixed air-sea temperature difference of -5°K), which is the usual condition for this part of the ocean in the autumn, then the mean difference between VIERS and METEO reduces to 0.1 m/s. The differences be-

tween VIERS and CMOD4 are much smaller than between each of them with the METEO winds.

4.2.2. Qualitative analysis. A qualitative analysis was performed on all data which were made available by ESA. This analysis led to a number of conclusions, which were illustrated by four case studies in *Janssen et al.* [1995]. Here we only discuss one case study, namely the detection and localization of fronts, while also the main conclusions are summarized.

We study here briefly a case in which a large front is visible in the ERS-1 data. On the southwestern part of the front the wind direction is southwesterly; on the other side of the front the wind direction is northeasterly. Plate 1 shows the VIERS retrieved wind field on November 6, 1991, in large and the corresponding METEO and CMOD4 wind field in the subimages. When comparing the images, a striking correspondence between the VIERS and CMOD4 result on the one side is seen while there is a clear discrepancy between the METEO winds and the ERS-1 derived winds. In the METEO wind field the front is not as pronounced as in the ERS-1 derived wind fields, and the position is ~ 200 km north of the front observed by ERS-1.

This example illustrates the conclusion that the ERS-1 derived wind fields show more structure than the meteorological model fields. Furthermore, the difference between VIERS wind fields and CMOD4 wind fields is generally smaller than

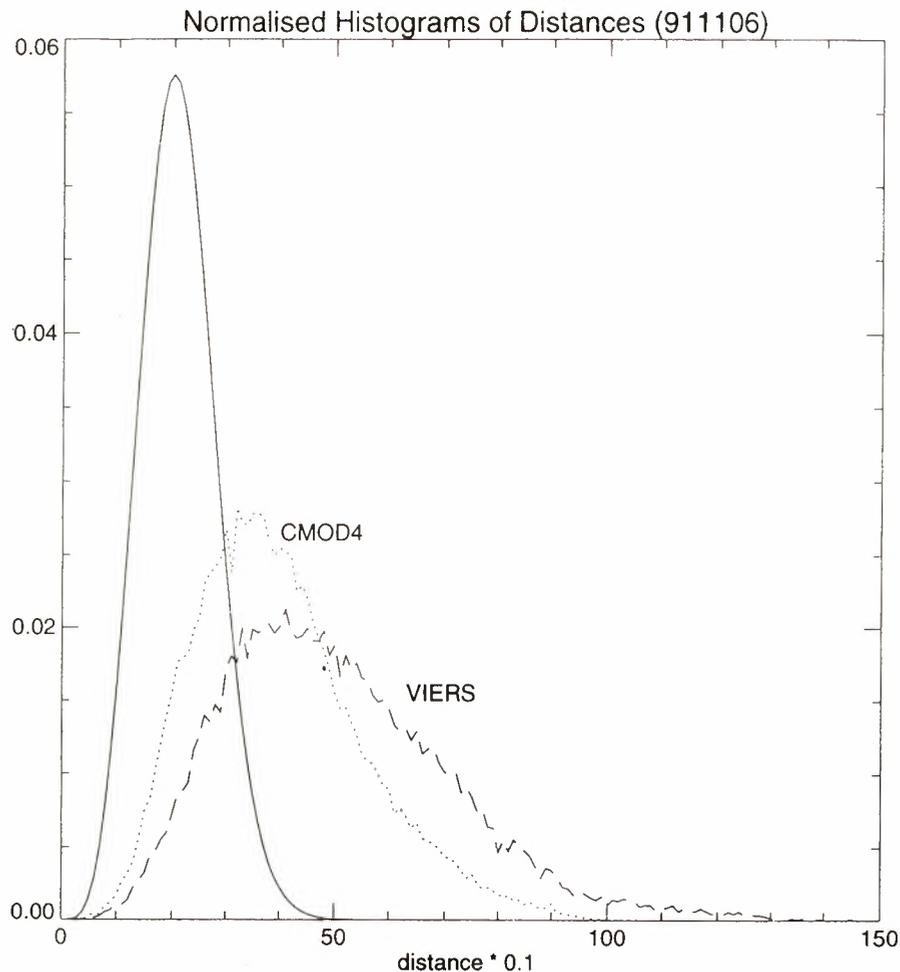


Figure 13. Normalized backscatter distance distribution for CMOD4 and VIERS. The distribution for a perfect model is shown as well.

between ERS-1 derived winds and winds from the meteorological model. It should also be pointed out that the inversion method we employ is rather successful since the ERS-1 winds are quite different from the METEO winds which are used in the minimization of the cost function D . An exception has to be made in case the METEO wind direction is orthogonal to the expected wind direction. In that event the inversion procedure is not always finding the right direction.

Additional observations we have inferred from studying the wind field maps are (1) because at small incidence angles the dependence on wind direction is weak, the retrieved wind direction in cells with these small incidence angles is less reliable; fortunately, this only occurred in the cell with the smallest incidence angle; (2) VIERS wind directions have a better internal consistency than CMOD4 directions; and (3) the VIERS model is capable of dealing with very low wind speeds.

Summarizing, we conclude that the structure of the VIERS wind fields and CMOD4 wind fields is very similar indeed, while the difference between METEO and ERS-1 derived wind fields is bigger. This is probably related to the fact that the METEO wind fields show much less structure than both fields from VIERS and CMOD4. Once more, it may be concluded that the VIERS model is an acceptable algorithm to retrieve the wind vector from radar backscatter measurements.

5. Conclusion

We have developed a scatterometer algorithm based on the present understanding of the radar backscatter process and of the relevant processes governing the short wave spectrum. The final aim was to be able to obtain wind fields from the backscatter as observed by the scatterometer on board of satellites such as ERS-1.

Using observed wave spectra and observed backscatter in the laboratory, it was readily realized that a simple two-scale model for the scattering process performed relatively well. In addition, it turned out that the short wave model was compatible with the wave measurements in the sense that spectra sufficiently close to the measured ones could be generated by tuning parameters which were not fixed a priori. As a final result, the two-scale model was combined with the wave model into the VIERS scatterometer algorithm. After a fine tuning exercise the algorithm evolved into the form described in this paper.

The present VIERS model has been shown to retrieve wind fields in a satisfactory manner; this followed both from the statistical comparison with ECMWF and CMOD4 wind fields and from the synoptic discussions. Furthermore, we have developed a method which enables us to retrieve, in a cost effective

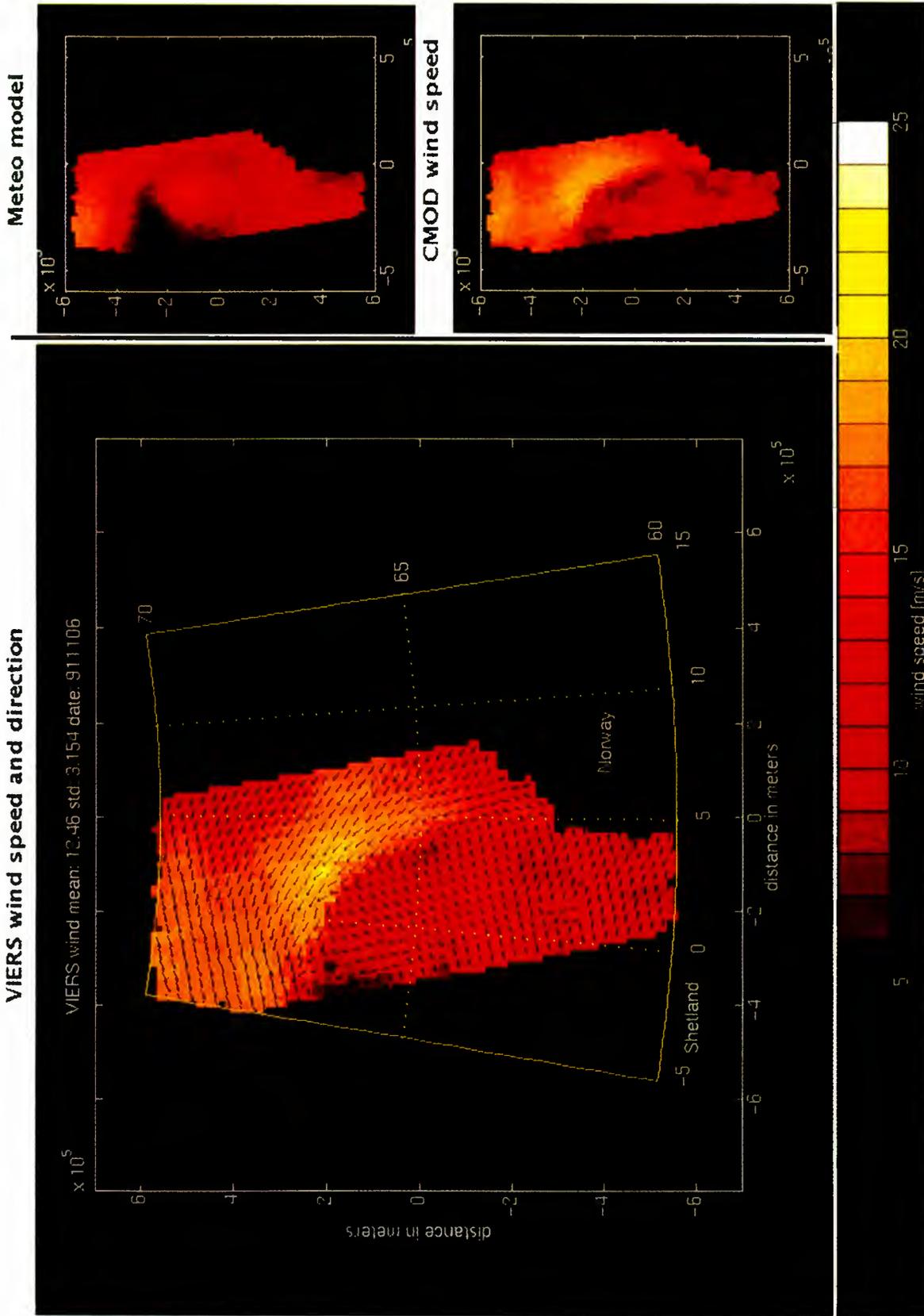


Plate 1. Retrieved wind fields from (left) VIERS, (top right) METEO, and (bottom right) CMOD4 on November 6, 1991. Area is the Norwegian Sea.

Table 3. Comparison of VIERS and CMOD4 Retrieved Winds With Analyzed METEO Winds During Cal/Val Campaign

Date	METEO	VIERS (Neutral)	CMOD4	VIERS-METEO		CMOD4-METEO		VIERS-CMOD4	
				Average	Standard Deviation	Average	Standard Deviation	Average	Standard Deviation
Sept. 19, 1991	11.6	10.1	10.1	-1.5	1.4	-1.4	1.2	-0.1	0.6
Sept. 22, 1991	14.1	9.3	8.6	-4.8	2.4	-5.5	2.5	0.8	0.7
Sept. 28, 1991	5.9	6.0	6.2	0.1	2.0	0.3	1.9	-0.2	0.6
Oct. 7, 1991	10.9	11.1	10.8	0.2	3.3	-0.0	3.1	0.3	0.6
Oct. 10, 1991	7.4	7.3	7.5	-0.2	1.6	0.0	1.5	-0.2	0.5
Oct. 19, 1991	9.9	13.6	12.9	3.7	1.5	3.1	0.9	0.6	0.8
Oct. 22, 1991	9.9	11.5	11.8	1.5	1.6	1.9	1.2	-0.4	0.7
Oct. 28, 1991	5.4	4.9	5.0	-0.5	1.4	-0.4	1.2	-0.1	0.5
Nov. 3, 1991	6.4	7.8	7.8	1.4	3.0	1.4	2.3	0.0	1.1
Nov. 6, 1991	3.1	3.4	4.1	0.3	1.8	0.9	1.7	-0.6	0.7
Nov. 12, 1991	6.5	7.8	8.1	1.3	1.9	1.6	1.8	-0.3	0.6
Nov. 15, 1991	2.0	4.3	4.7	2.3	2.1	2.7	1.7	-0.4	0.8
Nov. 21, 1991	13.2	17.1	15.6	3.8	3.1	2.4	3.0	1.4	1.1
Nov. 24, 1991	11.4	14.6	14.1	3.2	4.3	2.7	4.3	0.5	1.0
Nov. 30, 1991	14.4	14.4	13.6	0.0	2.8	-0.8	2.2	0.8	0.8
Dec. 3, 1991	11.2	13.5	13.2	2.3	1.1	2.0	1.1	0.3	0.6
Dec. 9, 1991	8.9	9.9	10.2	1.0	2.4	1.3	2.1	-0.3	0.7
Sept. 18, 1991	13.2	10.7	10.0	-2.5	3.5	-3.2	3.4	0.7	1.1
Sept. 24, 1991	10.2	10.2	9.1	-0.1	4.0	-1.1	3.8	1.1	1.0
Sept. 27, 1991	5.8	5.9	6.1	0.1	2.8	0.3	2.4	-0.2	1.0
Oct. 3, 1991	11.6	10.5	9.8	-1.1	2.2	-1.8	2.1	0.7	1.4
Oct. 6, 1991	7.7	8.2	7.9	0.5	1.6	0.2	1.1	0.3	1.2
Oct. 12, 1991	4.1	5.5	5.5	1.5	3.0	1.4	2.4	0.0	1.2
Oct. 15, 1991	2.8	2.8	3.7	0.0	1.7	0.9	1.2	-0.9	1.1
Oct. 21, 1991	6.9	9.7	9.5	2.8	1.9	2.6	1.6	0.2	0.9
Oct. 24, 1991	7.3	7.2	7.4	-0.0	1.9	0.1	1.6	-0.2	0.9
Oct. 30, 1991	12.0	6.7	6.6	-5.3	3.3	-5.4	2.7	0.1	1.1
Nov. 2, 1991	8.9	9.1	8.2	0.2	2.0	-0.7	1.5	0.8	1.0
Nov. 8, 1991	8.4	9.3	9.4	1.0	2.2	1.0	1.6	-0.0	1.3
Nov. 11, 1991	11.2	13.4	12.1	2.2	1.9	0.8	1.6	1.4	1.1
Nov. 17, 1991	4.7	7.3	7.1	2.6	2.9	2.5	2.3	0.2	1.2
Nov. 26, 1991	12.0	13.9	12.4	1.8	2.5	0.4	2.0	1.5	1.4
Dec. 2, 1991	12.3	14.9	13.2	2.6	2.0	0.9	1.6	1.7	1.2
Dec. 5, 1991	7.2	7.9	7.9	0.7	1.6	0.7	1.3	0.1	1.0
Sept. 19, 1991	12.5	16.5	14.0	4.0	1.1	1.5	1.3	2.5	1.2
Sept. 22, 1991	5.2	5.7	6.2	0.5	3.9	1.0	3.5	-0.5	1.2
Sept. 28, 1991	5.2	6.1	6.0	0.9	2.2	0.8	1.4	0.1	1.2
Oct. 1, 1991	10.9	12.4	11.4	1.5	1.9	0.5	1.6	1.0	1.2
Oct. 7, 1991	11.2	11.3	10.1	0.1	2.4	-1.1	1.7	1.1	1.3
Oct. 10, 1991	10.1	10.0	9.2	-0.0	1.5	-0.9	1.1	0.9	1.1
Oct. 16, 1991	12.6	13.6	13.3	1.0	2.3	0.7	1.9	0.3	1.1
Oct. 19, 1991	6.8	7.7	7.4	0.9	2.2	0.6	2.0	0.4	1.2
Oct. 25, 1991	8.0	8.3	7.8	0.3	1.5	-0.2	1.0	0.5	1.2
Oct. 28, 1991	8.8	8.3	7.4	-0.4	2.0	-1.4	1.1	0.9	1.3
Nov. 3, 1991	13.0	15.1	13.9	2.1	2.5	0.9	2.1	1.2	1.3
Nov. 6, 1991	10.0	12.5	12.0	2.5	3.6	2.0	3.8	0.4	1.1
Nov. 12, 1991	9.5	10.1	9.3	0.5	2.1	-0.2	1.6	0.8	1.4
Nov. 15, 1991	9.5	11.5	10.2	2.0	2.9	0.7	2.9	1.2	1.3
Nov. 21, 1991	8.5	8.8	8.2	0.3	1.9	-0.3	1.7	0.6	1.1
Nov. 24, 1991	10.4	11.2	10.0	0.8	2.8	-0.4	2.2	1.2	1.2
Nov. 30, 1991	9.0	11.5	10.3	2.5	2.1	1.2	1.9	1.3	1.3
Dec. 3, 1991	8.4	8.2	7.6	-0.2	2.2	-0.8	1.8	0.6	1.1
Statistics				0.8	2.3	0.3	2.0	0.5	1.0
Neutral				0.1	2.4	0.3	2.0	-0.2	1.1
Unstable									

tive way, wind fields using a rather complicated and expensive algorithm such as VIERS. In fact, it has been shown that in an operational environment, retrieval of VIERS winds may be done as efficiently as with the present operational CMOD4 algorithm.

Although the retrieved winds from VIERS and CMOD4 are of comparable quality in a statistical sense, we found that compared to the ECMWF wind fields, the CMOD4 winds are biased low in the high wind speed range. A similar conclusion follows from a comparison with buoy observations. The VI-

ERS bias was much less in this range. It should be once more emphasized that a reliable retrieval of winds in the high wind speed range is important. A negative bias in the wind retrieval would result in a considerably less deep analyzed depression since over the oceans the wind vector is related to a good approximation to the pressure gradient (geostrophic balance).

A weak point of the VIERS algorithm is the too simple directional distribution of the short waves. This is probably the major cause of the larger misfit in σ space (when compared to CMOD4). The strong point of the VIERS algorithm, on the

other hand, is that we have followed an approach based on physics. The framework of physical modeling as given by VIERS offers great potential for the future. Although CMOD4 at the moment shows a closer fit between modeled and observed backscatter, new insights into the directional distribution of the short waves will improve the performance of VIERS in this respect. Because of our framework this is a relatively easy step to take. In addition, in this way we were able to incorporate effects of sea state, slicks, and atmospheric stability in a natural manner. Finally, the VIERS algorithm has the added advantage that it can be applied to a fairly wide range of radar frequencies; hence without too much tuning one would expect that it should do a reasonable job of wind retrieval from backscatter from other scatterometers.

Acknowledgments. Useful discussions with David Anderson, Tony Hollingsworth, Ad Stoffelen, and Pete Woiceshyn are much appreciated. We thank J. Onvlee for providing WAM model data during the Cal/Val campaign. This work was supported by the Netherlands Remote Sensing Board (BCRS).

References

- Alpers, W., and H. Hühnerfuss, The damping of ocean waves by surface films: A new look at an old problem, *J. Geophys. Res.*, **94**, 6251–6265, 1989.
- Anderson, D., A. Hollingsworth, S. Uppala, and P. Woiceshyn, A study on the feasibility of using sea and wind information from the ERS-1 satellite, ESA report, part 1, 121 pp., Eur. Cent. for Medium-Range Weather Forecasts, Reading, England, 1987.
- Andersson, E., et al., The ECMWF implementation of three-dimensional variational assimilation (3D-Var), III, Experimental results, *Q. J. R. Meteorol. Soc.*, in press, 1998.
- Apel, J. R., An improved model of the ocean surface wave vector spectrum and its effects on radar backscatter, *J. Geophys. Res.*, **99**, 16,269–16,291, 1994.
- Bahar, E., D. E. Barrick, and M. A. Fitzwater, Computations of scattering cross sections for composite surfaces and the specification of the wave number where spectral splitting occurs, *IEEE Trans. Antennas Propag.*, **31**, 698–709, 1983.
- Banner, M. L., Equilibrium spectra of wind waves, *J. Phys. Oceanogr.*, **20**, 966–984, 1990.
- Barrick, D. E., and C. T. Swift, The Seasat microwave instruments in historical perspective, *IEEE J. Oceanic Eng.*, **OE-5**(2), 74–79, 1980.
- Brown, G. S., Backscattering from a Gaussian distributed perfectly conducting rough surface, *IEEE Trans. Antennas Propag.*, **26**, 472–482, 1978.
- Businger, J. A., J. C. Wyngaard, I. Izumi, and E. F. Bradley, Flux-profile relationships in the atmospheric surface layer, *J. Atmos. Sci.*, **28**, 181–189, 1971.
- Calkoen, C. J., et al., Viers-1 progress report, 2, *Tech. Rep. BCRS 90-27*, Neth. Remote Sens. Board, The Hague, 1990.
- Davidson, R. C., *Methods in Nonlinear Plasma Theory*, Academic, San Diego, Calif., 1972.
- Donelan, M. A., and W. J. Pierson, Radar scattering and equilibrium ranges in wind-generated waves with application to scatterometry, *J. Geophys. Res.*, **92**, 4971–5029, 1987.
- Donelan, M. A., J. Hamilton, and W. H. Hui, Directional spectra of wind-generated waves, *Philos. Trans. R. Soc. London A*, **315**, 509–562, 1985.
- Dyer, A. J., and B. B. Hicks, Flux-gradient relationships in the constant flux layer, *Q. J. R. Meteorol. Soc.*, **96**, 715–721, 1970.
- Gaffard, C., and H. Roquet, Impact of the ERS-1 scatterometer wind data on the ECMWF 3D-Var assimilation system, *ECMWF Tech. Memo. 217*, Eur. Cent. for Medium-Range Weather Forecasts, Reading, England, 1995.
- Guinard, N. W., J. T. Ransone Jr., and J. C. Daley, Variation of the NCRS of the sea with increasing roughness, *J. Geophys. Res.*, **76**, 1525–1538, 1971.
- Günther, H., A parametric surface wave model and the statistics of the prediction parameters, Ph.D. thesis, 90 pp., Univ. of Hamburg, Hamburg, Germany, 1981.
- Hasselmann, K., On the non-linear energy transfer in a gravity wave spectrum, 1, *J. Fluid Mech.*, **12**, 481–500, 1962.
- Hasselmann, K., et al., Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP), *Dtsch Hydrogr. Z.*, **A8**(12), suppl. 1–95, 1973.
- Jähne, B., and K. S. Riemer, Two-dimensional wave number spectra of small-scale water surface waves, *J. Geophys. Res.*, **95**, 11,531–11,546, 1990.
- Janssen, J. A. M., and H. Wallbrink, SATVIEW: A semi-physical scatterometer algorithm, *KNMI Sci. Rep. WR 97-03*, R. Neth. Meteorol. Inst., De Bilt, Netherlands, 1997.
- Janssen, P. A. E. M., The initial evolution of gravity-capillary waves, *J. Fluid Mech.*, **184**, 581–597, 1987.
- Janssen, P. A. E. M., P. Lionello, and L. Zambresky, On the interaction of wind and waves, *Philos. Trans. R. Soc. London A*, **329**, 289–301, 1989.
- Janssen, P. A. E. M., C. J. Calkoen, D. van Halsema, W. A. Oost, P. Snoeij, and H. Wallbrink, VIERS-1 final report phase 4, report, Neth. Remote Sens. Board, The Hague, 1995.
- Janssen, P. A. E. M., B. Hansen, and J. Bidlot, Verification of the ECMWF wave forecasting system against buoy and altimeter data, *Weather Forecasting*, **12**(4), 763–784, 1997.
- Jones, W. L., L. C. Schroeder, and J. L. Mitchell, Aircraft measurements of the microwave scattering signature of the ocean, *IEEE Trans. Antennas Propag.*, **AP-23**, 52–61, 1977.
- Jones, W. L., L. C. Schroeder, D. H. Boggs, E. M. Bracalente, R. A. Brown, G. J. Dome, W. J. Pierson, and F. J. Wentz, The Seasat-A satellite scatterometer: The geophysical evaluation of remotely sensed wind vectors over the ocean, *J. Geophys. Res.*, **87**, 3297–3317, 1982.
- Kitaigorodskii, S. A., On the theory of the equilibrium range in the spectrum of wind-generated gravity waves, *J. Phys. Oceanogr.*, **13**, 816–827, 1983.
- Klinke, J., and B. Jähne, 2D wave number spectra of short wind waves: Results from wind-wave facilities and extrapolation to the ocean, paper presented at the Conference on Optics of the Air-Sea Interface, Theory and Measurements, Int. Soc. for Opt. Eng., San Diego, Calif., 1992.
- Komen, G. J., K. Hasselmann, and S. Hasselmann, On the existence of a fully developed wind-sea spectrum, *J. Phys. Oceanogr.*, **14**, 1271–1285, 1984.
- Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen, *Dynamics and Modelling of Ocean Waves*, 532 pp., Cambridge Univ. Press, New York, 1994.
- Lamb, H., *Hydrodynamics*, 6th ed., 738 pp., Dover, Mineola, N. Y., 1932.
- Lombardini, P. P., Maximum of damping ratio in rippled water covered by monomolecular films, ONRL report on the role of surfactant films on the interface properties of the sea surface, *Rep. C-11-86*, edited by F. L. Herr and J. Williams, pp. 157–173, Off. of Nav. Res., London, 1986.
- Moore, R. J., and A. K. Fung, Radar determination of winds at sea, *Proc. IEEE*, **67**, 1504–1521, 1979.
- Moore, R. J., and W. J. Pierson, Measuring sea state and estimating surface winds from a polar orbiting satellite, in *Proceedings of the Symposium on Electromagnetic Sensing of the Earth From Satellites*, edited by R. Zirkland, pp. R1–R28, Polytechnic, New York, 1967.
- Phillips, O. M., Spectral and statistical properties of the equilibrium range in wind-generated gravity waves, *J. Fluid Mech.*, **156**, 505–531, 1985.
- Plant, W. J., On the steady state energy balance of short gravity wave systems, *J. Phys. Oceanogr.*, **10**, 1340–1352, 1980.
- Plant, W. J., Bragg scattering of electromagnetic waves from the air/sea interface, in *Surface Waves and Fluxes*, vol. II, edited by G. L. Geernaert and W. L. Plant, pp. 41–108, Kluwer Acad., Norwell, Mass., 1990.
- Schroeder, L. C., D. H. Boggs, G. Dome, I. M. Halberstam, W. L. Jones, W. J. Pierson, and F. J. Wentz, The relationship between wind vector and normalized radar cross section used to derive Seasat-A satellite scatterometer winds, *J. Geophys. Res.*, **87**, 3318–3336, 1982.
- Smith, S. D., et al., Sea surface wind stress and drag coefficients: The HEXOS results, *Boundary Layer Meteorol.*, **60**, 109–142, 1992.
- Snoeij, P., E. van Halsema, J. Vogelzang, S. Waas, S. Zecchetto, H. Janssen, W. Oost, B. Jähne, and C. Calkoen, VIERS-1 final

- report phase 3, *BCRS Rep. 92-24*, Neth. Remote Sens. Board, The Hague, 1993.
- Stewart, R. H., *Methods of Satellite Oceanography*, 360 pp., Univ. of Calif. Press, Berkeley, 1985.
- Stoffelen, A., and D. Anderson, Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4, *J. Geophys. Res.*, *102*, 5767–5780, 1997.
- Valenzuela, G. R., Theories for the interaction of electromagnetic and oceanic waves: A review, *Boundary Layer Meteorol.*, *13*, 612–685, 1978.
- Valenzuela, G. R., M. B. Laing, and J. C. Daley, Ocean spectra for the high-frequency waves as determined from airborne radar measurements, *J. Mar. Res.*, *29*, 69–84, 1971.
- van Gastel, K., Nonlinear interactions of gravity-capillary waves: Lagrangian theory and effects on the spectrum, *J. Fluid Mech.*, *182*, 499–523, 1987.
- van Halsema, D., et al., Progress report on the VIERS-1 project, *Tech. Rep. BCRS 89-24*, Neth. Remote Sens. Board, The Hague, 1989.
- Wentz, F. J., S. Peteherych, and L. A. Thomas, A model function for ocean radar cross sections at 14.6 GHz, *J. Geophys. Res.*, *89*, 3689–3704, 1984.
- Woiceshyn, P. M., M. G. Wurtele, D. H. Boggs, L. F. McGoldrick, and S. Peteherych, The necessity for a new parameterization of an empirical model for wind/ocean scatterometry, *J. Geophys. Res.*, *91*, 2273–2288, 1986.
-
- C. J. Calkoen, Argoss, Postbus 61, 8325 ZH Vollenhove, Netherlands. (e-mail: Charles.Calkoen@argoss.nl)
- P. A. E. M. Janssen, ECMWF, Shinfield Park, Reading RG2 9AX, England, U.K. (e-mail: janssen@ecmwf.int)
- W. A. Oost and H. Wallbrink, KNMI, Postbus 201, 3730 AE De Bilt, Netherlands. (e-mail: oost@knmi.nl; wallbrin@knmi.nl)
- P. Snoeij, Delft University of Technology, Faculteit Electrotechniek, Postbus 5031, 2600 GA Delft, Netherlands. (e-mail: P.Snoeij@et.tudelft.nl)
- D. van Halsema, TNO, Postbus 96864, 2509 JG Den Haag, Netherlands. (e-mail: erik.van.halsema@fel.tno.nl)

(Received December 10, 1996; revised October 3, 1997; accepted October 10, 1997.)

Analysis of ERS-1/2 synthetic aperture radar wave mode imagerettes

Vincent Kerbaol and Bertrand Chapron

Département d'Océanographie Spatiale, IFREMER, Centre de Brest, Plouzané, France

Paris W. Vachon

Canada Centre for Remote Sensing, Ottawa, Ontario, Canada

Abstract. The European Space Agency (ESA) ERS-1 and ERS-2 C band VV polarization active microwave instrument (AMI) offers the unique ability to combine interlaced wind scatterometer and high-resolution synthetic aperture radar (SAR) wave mode imagerettes. In this study, more than 2000 imagerettes were considered. Each imagerette has been statistically analyzed in comparison with normalized radar cross-section (NRCS) measurements from the scatterometer mode. During the ERS-1 AMI wave mode mission the incidence angle of the imagerette center was modified from roughly 19.9° to 23.5° . Using ERS-1 and ERS-2 NRCS, calibration has thus been completed for both configurations, which allows a better characterization of some signal data saturation effects. However, since a SAR relies on platform displacement to achieve fine resolution, surface motions reduce its nominal resolution. As the wind speed increases, scatterer motion occurring during the SAR integration time also increases, causing the characteristic large azimuth cutoff wavelength. Acknowledging the strong dependence between sea state conditions and azimuth smearing effects, our results are highlighted by the global comparison with wind estimates derived from the scatterometer. The results demonstrate the ability to define a SAR wind algorithm from a kinematic point of view. Finally, a higher-order statistical analysis shows evidence of deviation from standard Rayleigh statistics, leading to a balance between K law and lognormal distributions. This deviation is mainly due to the SAR's high-resolution properties.

1. Introduction

Since the launch of Seasat in 1978 it has been clear that synthetic aperture radar (SAR) instruments can image ocean surface wave patterns [Alpers *et al.*, 1981; Vesecky and Stewart, 1982]. Unfortunately, it is now well known that the wave-like patterns visible in a SAR image of the ocean surface may be considerably different from the actual ocean wave field. As a result, extraction of meaningful two-dimensional wave spectral properties from a SAR scene is not straightforward. The capacity of SAR to provide usable ocean wave spectra is limited by the motion of the ocean surface. Although a polar-orbiting SAR is not an ideal instrument to measure the directional spectrum of ocean surface waves, valuable information may be routinely extracted. In particular, with the ERS-1 and ERS-2 missions and the so-called SAR wave mode products [European Space Agency, 1992a], wave modelers can, for the first time, obtain continuous global information [Brüning *et al.*, 1994].

Wave mode data continuity will be available through ENVISAT.

As an attempt to extract meaningful information from ocean SAR scenes, we have focused on the unique ability of the ERS active microwave instrument (AMI) to combine interlaced scatterometer and SAR wave mode measurements. The wave mode is discussed in section 2. Our objectives are met by carrying out a systematic comparison between normalized radar cross-section (NRCS) data and wind vector estimates from the scatterometer, with various statistical and observed spectral parameters of collocated SAR imagerettes.

A primary measurable of a calibrated SAR is back-scattered power. This measurement is diffraction dominated and depends upon short sea surface wind waves. Thus an important capability of a SAR instrument should be its ability to quantitatively relate image intensity to local surface wind speed. Empirically derived wind retrieval algorithms are now routinely used to infer wind estimates from scatterometer and altimeter, but much less attention has been devoted to SAR instruments [Vachon and Dobson, 1996]. In section 3 we present a calibration procedure to use SAR data much like a higher-resolution scatterometer. In the absence of

Copyright 1998 by the American Geophysical Union.

Paper number 97JC01579.
0148-0227/98/97JC-01579\$09.00

specific information for SAR imagette radiometric analysis our method is to derive a relative calibration from direct global comparisons between collocated SAR and scatterometer central antenna measurements. The relative calibration follows a strong saturation law. Thus the relationship between scatterometer central antenna NRCS and mean values of imagette intensity must take into account saturation of the analog-to-digital converter (ADC). In particular, this study shows evidence of the change of imagette incidence angle during the ERS-1 mission. The validation of this empirical calibration process is presented in section 4 using a semiempirical wind retrieval scattering model, which has been developed at the Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER) and referred as CMOD-IFREMER, [Quilfen, 1993].

Improved information has also been obtained through global and systematic analysis of the SAR along-track resolution, i.e., the azimuthal cutoff. According to theory, this latter parameter is proportional to the root-mean-square (rms) line of sight of the wave orbital velocity field. Since the ERS-1 and ERS-2 wave mode operate at near-nadir incidence angles (19.9° and 23.5°), the radial component is supported by the surface vertical velocity which, by definition, will be dominated by shorter gravity waves than the overall energy integral [Jackson and Peng, 1985]. The portion of the spectrum supporting these waves is usually assumed to be wind dependent, and we propose that the azimuthal response is wind speed dependent. Early results based on global comparisons confirmed this assumption [Chapron et al., 1995]. In section 5, azimuth cutoff is discussed in terms of incidence angle, wind directionality, and sea state development.

From a statistical point of view, direct analysis of high-resolution radar scenes over the ocean exhibits departures from a standard Gaussian-Rayleigh model.

This deviation is demonstrated in section 6 through analysis of higher-order moments and is represented in a Pearson diagram, which is a convenient representation of distribution families. The source of this deviation is discussed in light of SAR high-resolution properties by studying SAR power image variance behavior. Indeed, the concept of large-scale inhomogeneities (long wave field) modulating a dense array of small-scale scatters can lead to non-Gaussian statistics.

2. ERS Wave Mode Data

The CERSAT (Centre for Archiving and Processing of ERS data), located at IFREMER in Brest, is tasked with archiving SAR wave mode imagettes and processing scatterometer data for the ERS-1 and ERS-2 satellites. It provides the unique opportunity to combine the analysis of both products.

The European Space Agency (ESA) ERS-1 and ERS-2 C band VV polarization AMI offers the ability to obtain interlaced scatterometer data and SAR wave mode imagettes (Figure 1). The 10×5 km imagettes are range compressed on board the satellite and are interlaced with scatterometer data every 200 km along the satellite track. These imagettes represent magnitude look detection and incoherent averaging of three independent looks. Each imagette consists of 600 pixels in range (across track) and 320 pixels in azimuth (along track), with pixel spacing of 20 m in range and 16 m in azimuth.

In this work, more than 2000 imagettes have been considered, representing a wide variety of conditions. Various statistics have been measured, including NRCS, azimuth cutoff, and higher-order statistics. Our results are highlighted by the global comparison with calibrated NRCS and wind estimates derived from the collocated scatterometer measurements.

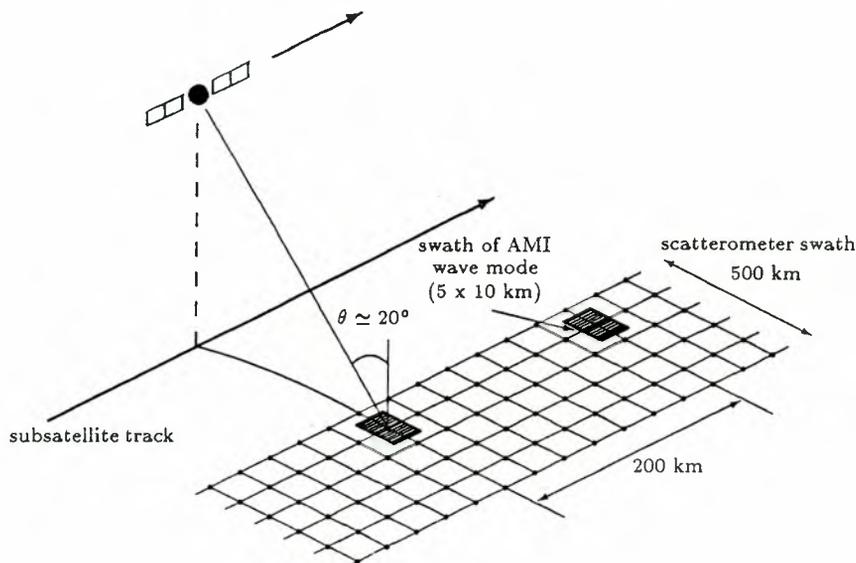


Figure 1. ERS synthetic aperture radar (SAR) wave mode and scatterometer geometry.

3. Calibration

Ocean ripples are directly related to local sea surface wind velocity and can be detected by radar techniques. In particular, SAR scattering measurements are diffraction dominated and depend upon short gravity-capillary wind waves. Thus an important capability of a SAR instrument is its ability to quantitatively relate image intensity to local surface wind speed. Empirically derived wind retrieval algorithms are now routinely used to infer wind estimates from scatterometer and altimeter data. Much less attention has been devoted to SAR instruments [Scoon *et al.*, 1996; Vachon and Dobson, 1996]. If the SAR is used like a high-resolution scatterometer to derive wind speed estimates from an empirically derived satellite wind model function such as CMOD, we need to relate pixel values in the imagettes to NRCS values.

SAR imaging of the ocean surfaces also relies on Doppler information from the relative motion between the SAR and the scene to achieve fine resolution in the along-track direction. For example, the SAR images waves by their own orbital velocities. Using the concept of facets, the effect of long waves (longer than the nominal SAR resolution cell) can be treated deterministically [Alpers and Ruffenach, 1979; Swift and Wilson, 1979] whereas short waves (shorter than the SAR resolution cell) may be treated stochastically [Tucker, 1985]. The motions due to short waves induce an azimuth smearing that acts as a low-pass filter (often modeled as Gaussian-shaped) on the azimuthal components of the image spectrum. In particular, it has been stated that this imaging process turns into a superposition of facets [Hasselmann and Hasselmann, 1991]. The study of the number of superposed facets has not been considered in detail [Kanevsky, 1993; Chapron *et al.*, 1994a]. However, this phenomenon should not influence the mean image intensity level because of the conservation of the mean between the real aperture radar (RAR) and the SAR image [Hasselmann and Hasselmann, 1991].

From the radar equation the SAR image intensity may be written

$$I_{\text{SAR}} = \frac{K \sigma^{\circ} g^2(\theta)}{R^3 \sin \theta} \text{Loss}(\text{ADC}, \dots) \quad (1)$$

where K represents a calibration constant, σ° represents the radar cross-section, θ represents the incidence angle, R represents the slant range, g represents the two-way elevation antenna pattern, and Loss represents a function accounting for any power loss processes (such as ADC saturation). As delivered, imagettes are not corrected for range spreading and antenna pattern. Yet these corrections may be considered as corrective constants due to the small incidence angle range over each imagette (0.8°). The radar equation is simplified as

$$I_{\text{dB}} = \sigma^{\circ}_{\text{dB}} + K'_{\text{dB}} + \text{Loss}_{\text{dB}} \quad (2)$$

with K'_{dB} , a modified calibration constant, defined as

$$K'_{\text{dB}} = K_{\text{dB}} + 10 \log_{10} [g^2(\theta)] - 10 \log_{10}(R^3 \sin \theta) \quad (3)$$

3.1. ADC Saturation

The quantization of I (real) and Q (imaginary) signal data channels within the ADC on a limited number of bits is known to reduce the output power. For ERS SAR Precision Images (PRI) [European Space Agency, 1992b], both channels are quantized with 5 bits [Laur *et al.*, 1993]. For SAR imagettes, quantization is applied on 4 bits [Kerbaol and Chapron, 1996]. Radar cross-section measurements are then strongly affected by ADC saturation when the backscattered power takes large values (> -6 dB), which corresponds to wind speeds greater than 6 m/s for range winds and 10 m/s for azimuthal winds. The power loss correction is given as a function of the I or Q channel standard deviation (Figure 2). The curve was obtained by simulation, assuming the I and Q channels to be normally distributed.

Fortunately, ADC power loss correction can be applied on ERS-1/2 imagettes even though raw data are not at our disposal. Indeed, the I/Q channel standard deviation is provided in the header of each imagette, and the resulting power loss can be corrected on the entire imagette since, in most cases, the homogeneity of the imagette is assured by its small size (5×10 km). Note that the ADC saturation effect is not expected to significantly affect the mean SAR intensity of ERS-2 imagettes because of a reduction in onboard gains in July 1995 [Laur *et al.*, 1996].

3.2. Change of the ERS-1 SAR Incidence Angle From 19.9° to 23.5°

The calibration of ERS-1/2 imagettes would be complete if the calibration constant K is known. Though this constant is well known for PRI images, it is not available for wave mode products, which makes the cal-

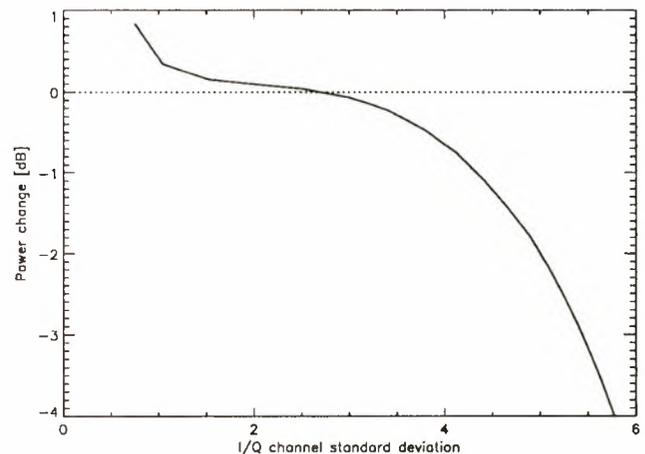


Figure 2. Power loss at analog-to-digital converter (ADC) output versus I or Q channel standard deviation for a 4 bit ADC.

ibration issue more difficult. An approach to wave mode calibration is to compare the mean intensity of ERS-1/2 SAR imagettes to the collocated scatterometer NRCS. Then, after subtracting K'_{dB} from the mean imagette intensity after correction for ADC power loss, one should obtain a 1-to-1 ratio with scatterometer σ° .

This method enabled us to refine the exact incidence angle used to acquire the ERS-1/2 SAR wave mode products. For ERS-2 imagettes the incidence angle is set to 23.5° . This was also meant to be the incidence angle for ERS-1. However, it was originally set to 19.9° . In June 1995 the ERS-1 imagette incidence angle was reset to 23.5° . Two data sets were chosen to illustrate this change of configuration. The first one represents 1092 imagettes acquired in November 1992 over the North Pacific Ocean, while the second one represents 693 imagettes acquired in June 1995 over the Indian Ocean (Figure 3).

Figure 4a shows the comparison result with collocated scatterometer NRCS at 19.9° (November 1992), while Figure 4b shows the result at 23.5° (June 1995). As expected, the change in incidence angle decreases the mean σ° value and results in a reduction of the ADC saturation effect.

After ADC saturation correction the expected 1-to-1 ratio is obtained for the 23.5° incidence angle data set, whereas it is only reached for the 19.9° data set for the lowest measured cross sections. Beyond a certain level, saturation is so large that it becomes difficult to correct for power loss. Note the spreading of data around the 1-to-1 ratio line for the smallest scatterometer NRCS values. This can be associated with the much larger scatterometer resolution cell (25×25 km) compared to

the SAR (30×30 m). Indeed, the variability at low wind speed is integrated and smoothed over the large scatterometer resolution cell while SAR imagettes only provide local measurements.

3.3. ERS-2 Onboard Gain Adjustment

For ERS-2 SAR imagettes a set of 342 imagettes, acquired over the Indian Ocean in June 1996, was selected (Figure 3). Figure 5 shows that the imagettes are almost not affected by ADC saturation. This is due to a reduction in ERS-2 onboard gains on July 13, 1995. ADC saturation correction is not essential, except for σ° values larger than -3 dB. This level is close to the -2 dB level given by Laur *et al.* [1996, p. 11] for ERS-2 PRI images. All the K' values for ERS-1 and ERS-2 are given in the appendix A.

4. Validation

Recent studies [Scoon *et al.*, 1996; Vachon and Dobson, 1996] have shown the SAR ability to work as a high-resolution scatterometer to retrieve wind speed using a scatterometer wind model (CMOD4 [Stoeffelen and Anderson, 1993] or CMOD IFREMER [Quilfen, 1993]). The method is to compare SAR range profiles to a wind retrieval model assuming a known, steady wind vector.

Here we propose to validate our empirical calibration method. For each data set we selected cases whose wind directions, as given by ERS scatterometer, are either in the range or the azimuth direction. The corresponding mean SAR intensity is then compared with CMOD IFREMER. Figure 6 shows the results for ERS-

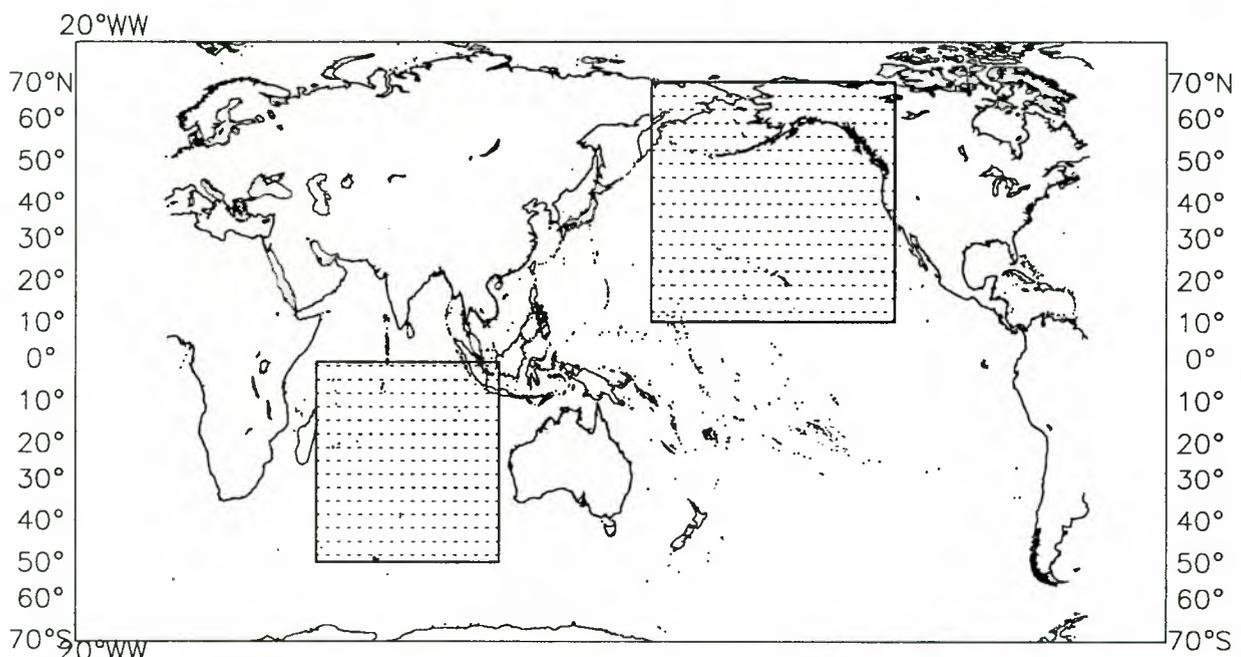
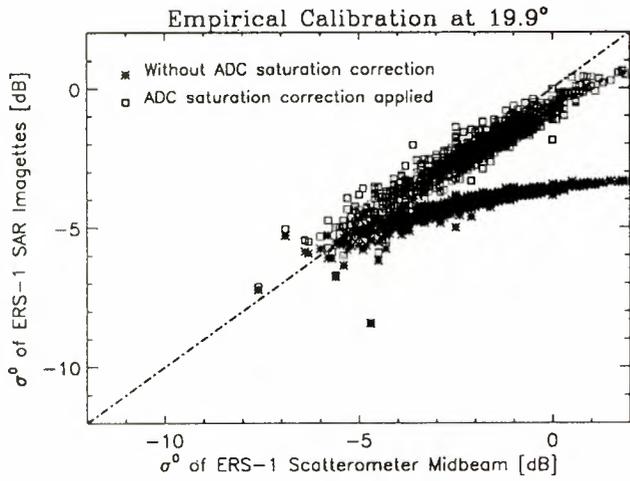
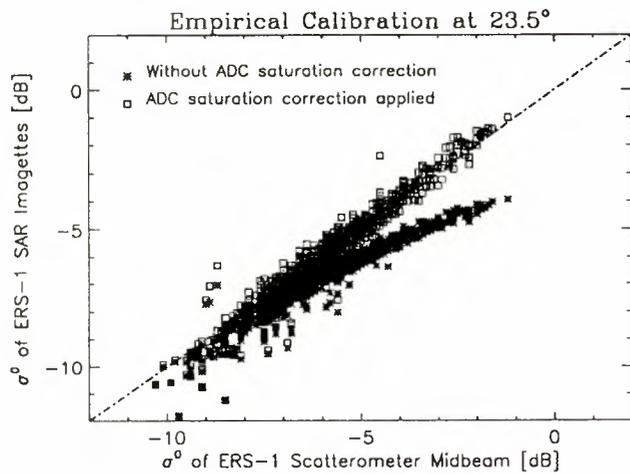


Figure 3. Areas selected for ERS-1 imagettes: North Pacific Ocean (November 1992) and Indian ocean (June 1995). Area selected for ERS-2 imagettes: Indian Ocean (June 1996).



(a)



(b)

Figure 4. ERS-1 mean SAR intensity versus collocated scatterometer σ^0 before ADC correction (asterisks) and after correction (squares). The line represents a 1-to-1 ratio. The calibration constant K' (Table A1) has been subtracted from the SAR intensity. (a) North Pacific Ocean data set ($\theta = 19.9^\circ$, November 1992). (b) Indian Ocean data set ($\theta = 23.5^\circ$, June 1995).

1 SAR in both incidence angle configurations (19.9° (Figure 6a) and 23.5° (Figure 6b)). One can see that ERS-1 SAR imagettes have been successfully calibrated for the incidence angle of 23.5° . Indeed, both wind directions (along range and azimuth) lie on the respective CMOD IFREMER curve. However, this is not the case for the data set over the North Pacific Ocean, particularly when the wind speed increases. This is due to the weakness of the ADC saturation correction at 19.9° . Incidentally, it appears that the wind retrieval model used for this study tends to underestimate high wind speeds, particularly in the range direction (up/down winds). Winds are expected to be as high as 18 m/s, rather than the maximum 15 m/s of Figure 6b. The results

for ERS-2 imagettes are presented in Figure 7. The calibration scheme also gives good results for ERS-2 SAR wave mode imagettes, although only a few examples of azimuthal wind were at our disposal. Indeed, winds generally blow from west to east in the roaring forties (e.g., mainly along the range direction).

According to these comparisons, the SAR instrument may prove its ability to work as a high-resolution scatterometer. Given a first-guess wind field direction (such as provided by a meteorological model or synoptic chart), the wind speed can be obtained from SAR images using the CMOD retrieval algorithm. Inversely, from an independent wind speed measurement a properly calibrated SAR signal might be able to provide the wind direction [e.g., *Korsbakken*, 1996].

5. SAR Azimuth Cutoff

Doppler misregistrations in azimuth are induced by gravity wave orbital motion, leading to a distortion of the imaged spectrum and a strong cutoff in the azimuthal direction. This effect is proportional to the range-to-platform velocity ratio R/V [Beal et al., 1983]. ERS-1, ERS-2, RADARSAT, and planned future SAR missions such as ENVISAT are polar-orbiting platforms and have an R/V of about 120 s or larger. Then, depending on sea state, the shortest detectable wavelength in the azimuth direction will not be less than 200 m. We calculated SAR image spectra for all ERS imagettes of our data set and plotted the spectral peak wavelength against its direction relative to the range axis (Figure 8). This plot shows that all waves shorter than 100 m can only be observed in the range direction. Note that only a few peaks arise right on the range axis. This is due to the so-called peak-splitting effect [Brüning et al., 1990]. The effect of SAR mapping is to low-pass fil-

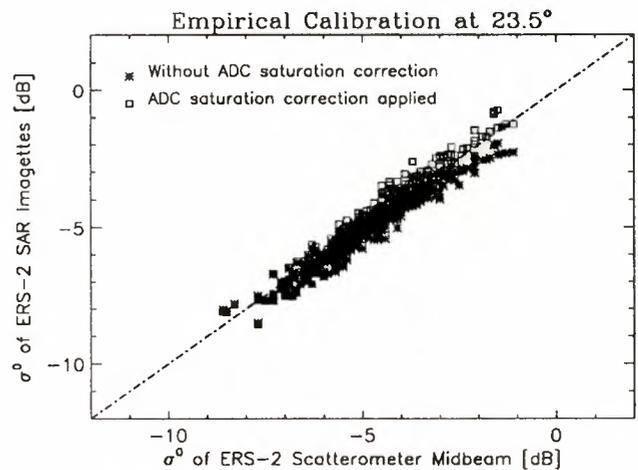
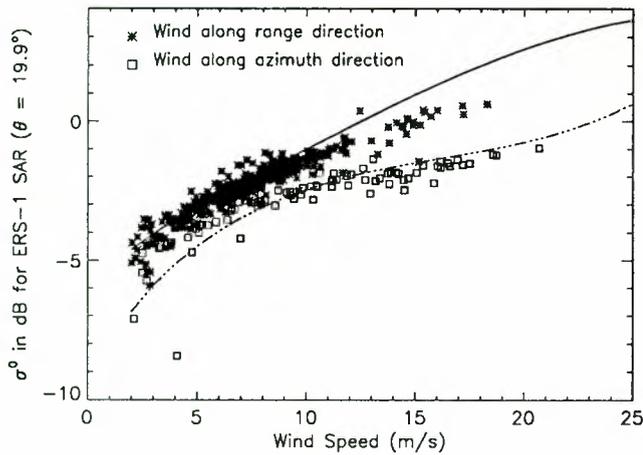
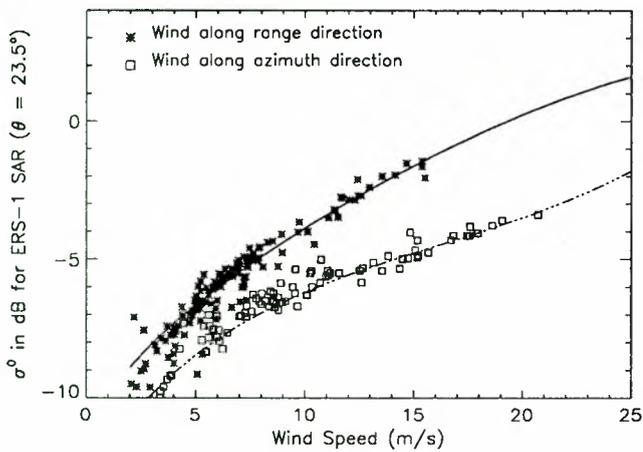


Figure 5. ERS-2 mean SAR intensity versus collocated scatterometer σ^0 before ADC correction (asterisks) and after correction (squares). The line represents a 1-to-1 ratio. The calibration constant K' (Table A1) has been subtracted from the SAR intensity. Indian Ocean data set ($\theta = 23.5^\circ$, June 1996).



(a)



(b)

Figure 6. ERS-1 SAR σ^0 versus wind speed given by ERS-1 scatterometer for wind blowing in the range direction (asterisks) and in the azimuth direction (squares). The solid curve represents the empirical C band scattering model (CMOD-IFREMER) σ^0 for a range wind direction and the dot-dashed curve for an azimuth wind direction. (a) North Pacific Ocean data set ($\theta = 19.9^\circ$, November 1992). (b) Indian Ocean data set ($\theta = 23.5^\circ$, June 1995).

ter the real aperture radar (RAR) spectral information in the azimuthal direction such that the SAR spectral domain is limited to around the range axis.

The issue of the SAR azimuth cutoff interpretation has been widely considered in the literature, and several models have been proposed as a function of wave and/or wind conditions. A first interpretation was initially proposed in terms of the significant wave height H_S [Beal et al., 1983], but the azimuth cutoff has been shown to also be wind speed dependent. Then the contribution of azimuthal displacements caused by

the long waves orbital motions was outlined [Jackson and Peng, 1985; Tucker, 1985; Hasselmann et al., 1985] and appeared to be more realistic when compared with the cutoff behavior [Vachon et al., 1994]. Referring to studies on processes moving in time [Steinberg et al., 1955; Brillinger, 1993] and taking advantage of wind vector estimates inferred from collocated ERS scatterometer data, we will show evidence of the SAR azimuth cutoff dependence with the total variance of vertical motions on the sea surface.

The power spectrum of a signal whose source is subject to random fluctuations has already been studied and has been shown to be filtered by the characteristic function of the displacement field [Brillinger, 1993]. Hence, extending this relationship to forward SAR mapping while under the simplifying assumption that the displacement field and the linear SAR modulation field (including RAR and velocity bunching modulation) are independent, the relationship between $P_{\text{lin}}(k)$, the power spectrum of the linear SAR transform of ocean spectrum, and $P_S(k)$, the complete SAR power spectrum, is

$$P_S(k) = e^{-k_{az}^2 \rho_{\xi\xi}(0)} P_{\text{lin}}(k) \quad (4)$$

where $\rho_{\xi\xi}(0)$ represents the total variance of azimuthal displacements within the SAR integration time. This exponential weighting function is identical to the standard azimuthal cutoff factor [Hasselmann and Hasselmann, 1991; Krogstad, 1992]. This relationship is often referred to as the so-called quasi-linear approximation, and this transform behaves like an azimuthal low-pass filter in the spectral domain. Noting that the azimuthal displacement ξ induced by a moving scatterer with radial velocity v is

$$\xi = \frac{R}{V} v \quad (5)$$

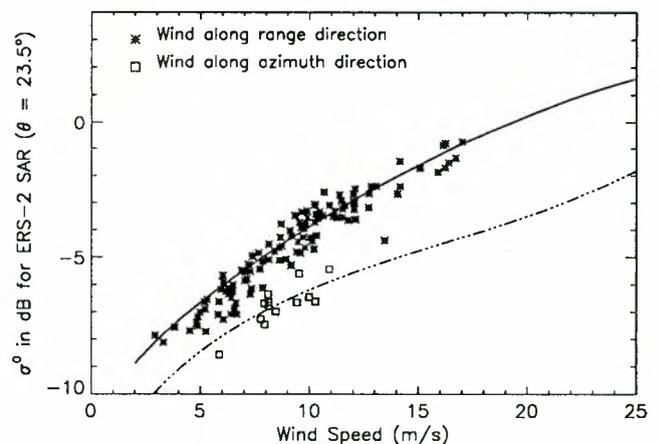


Figure 7. ERS-2 SAR σ^0 versus wind speed given by ERS-2 scatterometer for wind blowing in range direction (asterisks) and in azimuth direction (squares). The solid curve represents CMOD IFREMER σ^0 for a range wind direction and the dot-dashed curve for an azimuth wind direction. Indian Ocean data set ($\theta = 23.5^\circ$, June 1996).

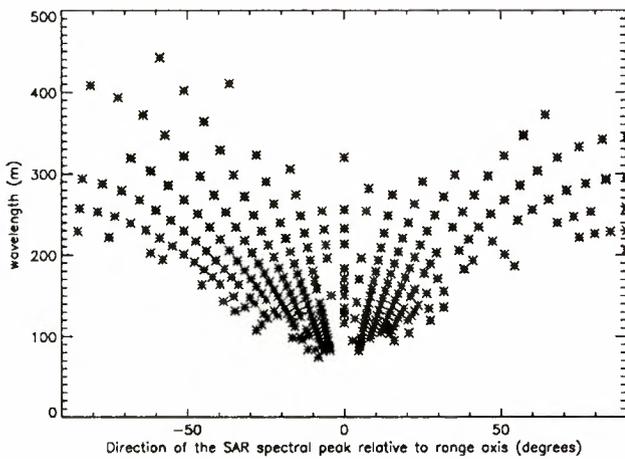


Figure 8. Wavelengths of observed SAR spectral peak versus the angle between SAR peak and range axis.

and that at an incidence angle of about 20° , v is well approximated by the vertical component of the orbital velocity (i.e., the directivity of the wave field can be neglected), $\rho_{\xi\xi}(0)$ becomes

$$\rho_{\xi\xi}(0) = \left(\frac{R}{V}\right)^2 \int_0^\infty \omega^2 S(k) dk \quad (6)$$

where ω is the wave frequency. Here $S(k)$ is integrated over the full wavenumber range in order to include the Doppler spread of the backscattered signal caused by the motion of the facets within a resolution cell. This additional azimuth smear is often expressed in terms of a coherence time.

One should note that $\rho_{\xi\xi}(0)$ does not take into account the radial acceleration term. Since the integration time of the ERS SARs is small (about 0.5 s), the acceleration contribution can be neglected [Bao et al., 1994]. However, in some cases the smearing due to acceleration may become important. This is expected for high sea state cases when specular events, such as those from breaking waves, may lead to large azimuthal streaks and an enhancement of the azimuth cutoff [Alpers and Brüning, 1986; Ouchi and Cordey, 1991].

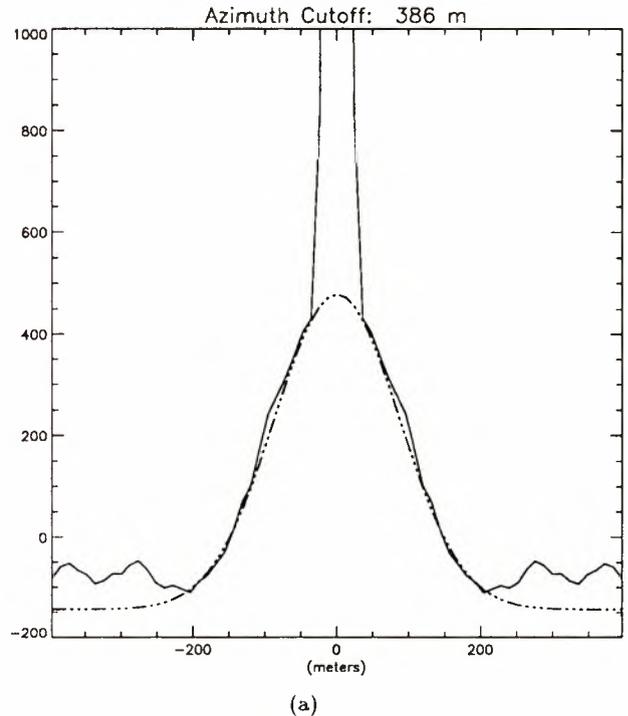
5.1. Azimuth Cutoff Estimation

An estimation method of the SAR azimuth cutoff has previously been presented [Vachon et al., 1994]. This procedure used an ensemble of model functions and various cutoff wavelengths to find the best fit between generated and observed SAR azimuth spectra. In order to gain in computing time we directly use the azimuthal autocorrelation function (ACF) by fitting a Gaussian function defined as

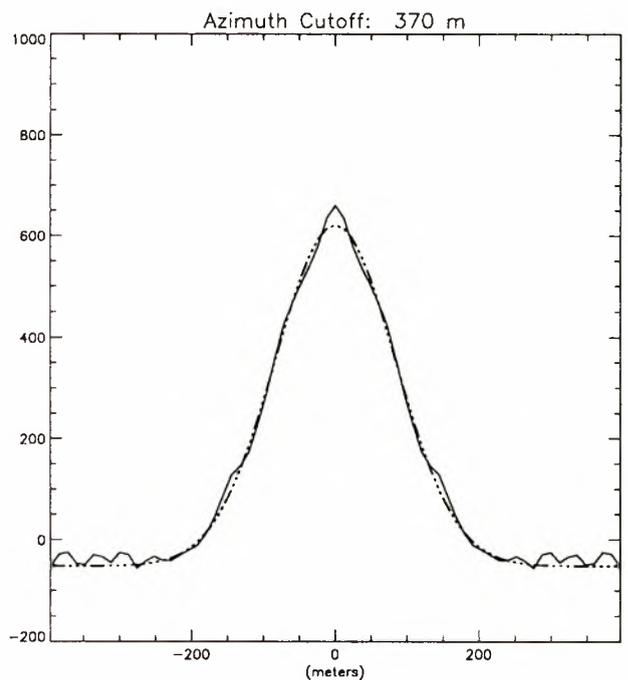
$$C(x) \sim e^{-\pi^2 \frac{x^2}{\lambda^2}} \quad (7)$$

Generally speaking, we have found that this method gives good results. However, one needs to be cautious with the interpretation of the estimated cutoff. In the

case of an imagette without any visual contrast, the ACF will be very close to the narrow point spread function associated with the speckle noise. When the sea surface motion increases, the ACF can then be considered as the sum of two spread functions, corresponding to the azimuthal resolution and the azimuth cutoff, respectively (Figure 9a).



(a)



(b)

Figure 9. (a) Estimation of the azimuth cutoff through the azimuthal correlation function. (b) The narrow central peak of the non coherent scattering has been removed by using intercorrelation between two independent looks.

In the near future the intercorrelation method between distinct looks may also be used. Indeed, when carrying the cross-spectrum of different looks, the incoherent scattering contribution will be suppressed [Engen and Johnsen, 1995]. The speckle noise contributes a large bandwidth noise floor (pedestal) in the SAR spectrum. The usual way to reduce the noise contribution is to use the multilook technique, but the azimuthal resolution is then degraded. The cross spectrum technique will lead to a much better signal-to-noise ratio by eliminating the uncorrelated noise. As a consequence, this technique is helpful when carrying out the intercorrelation between looks to estimate the falloff parameter. As noted above, when one observes the correlation function of a multilook SAR image, there is still an apparent contribution of the noise in the narrow peak of the speckle autocorrelation function. The intercorrelation function between different looks will not show this peak (Figure 9b). The estimation of the cutoff parameter will then become more consistent, especially under low wind conditions. Furthermore, the acceleration term in the azimuth smearing due to short lifetime specular events should not appear in the intercorrelation method.

5.2. SAR Wind Algorithm

Acknowledging the strong dependence between sea state conditions and azimuth-smearing effects, Jackson and Peng [1985] first suggested derivation of a wind algorithm from azimuth cutoff analysis. Global comparisons can now be done by using wind estimates derived from the ERS-1 or ERS-2 scatterometers. For each imagette the wind speed is obtained from the surrounding area (within 25 km). Linear regression gives a convincing idea of a possible empirical relationship to define a wind algorithm from ACF analysis of SAR image spectra (Figure 10). Practical application of a linear relationship has already been demonstrated [Korsbakken, 1996]. If the SAR azimuth cutoff is expressed as [Lyzenga, 1986]

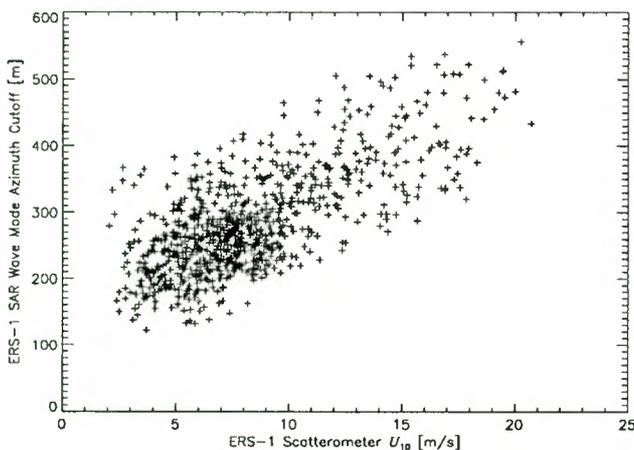


Figure 10. ERS-1 SAR azimuth cutoff versus collocated U_{10} inferred from ERS-1 scatterometer.

$$\lambda = \pi \sqrt{\rho_{\xi\xi}(0)} = \pi \frac{R}{V} \sqrt{\int_0^{\infty} \omega^2 S(k) dk} \quad (8)$$

a numerical application using a JONSWAP-like sea spectral model [Elfouhaily et al., 1996] leads to the linear relationship $\lambda \simeq 25 U_{10}$, where U_{10} is the neutral stability wind speed at 10 m above the ocean surface. This evaluation is consistent with the result of our linear regression

$$\lambda = 23.4 U_{10} + \Lambda \quad (9)$$

where Λ represents a residual cutoff involving the nominal SAR azimuthal resolution. Note that this relationship is consistent with the analysis of Vachon et al. [1994] from the Grand Banks ERS-1 SAR validation experiment but with a different notation.

5.2.1. Wave age dependency. Although the kinematic parameter λ is a robust wind speed estimator, it is necessary to consider that the degree of sea state development may also play a significant role. For example, a cutoff analysis over fetch-limited seas results in an underestimation of the wind speed [Kerbaol et al., 1996]. In this case we rewrite (8) as

$$\lambda(\Omega) = \pi \sqrt{\rho_{\xi\xi}(0)} = \pi \frac{R}{V} \sqrt{\int_0^{\infty} \omega^2 S(k, \Omega) dk} \quad (10)$$

where Ω represents the inverse wave age dependency of the wave spectral model defined as [Elfouhaily et al., 1996]

$$\Omega = 0.84 \tanh \left[\left(\frac{k_0 x}{2.2 \cdot 10^4} \right)^{0.4} \right]^{-0.75} \quad (11)$$

with x , the dimensional fetch in meters, and k_0 is g/U_{10}^2 .

As an example, and following the wind and fetch-dependent sea spectrum [Elfouhaily et al., 1996], Figure 11 shows the behavior of the theoretical λ for wind speeds ranging from 2 to 25 m/s given different fetch values. For our purpose we also consider that the degree of development is variable over different ranges of wind speed. If it is reasonable to take a fully developed model for light to moderate wind speed conditions, this will not in general be the case for high wind speeds. Following studies using altimeter data [Tournadre and Blanquet, 1994], we choose an average fetch of 200 km for the whole range of wind speed over the globe to illustrate this effect, which can be thought of as a saturation phenomenon. This corresponds to an inverse wave age parameter Ω of about 0.84 for low wind speed to about 2 for a 20 m/s wind speed.

As an example, Figure 12 gives the cutoff values estimated from a set of imagettes. These values are plotted versus the wind speed inferred from the collocated scatterometer data. The solid line represents fully developed seas while the dash-dotted line is for a fixed fetch condition (developing seas). The agreement is seen to be reasonable.

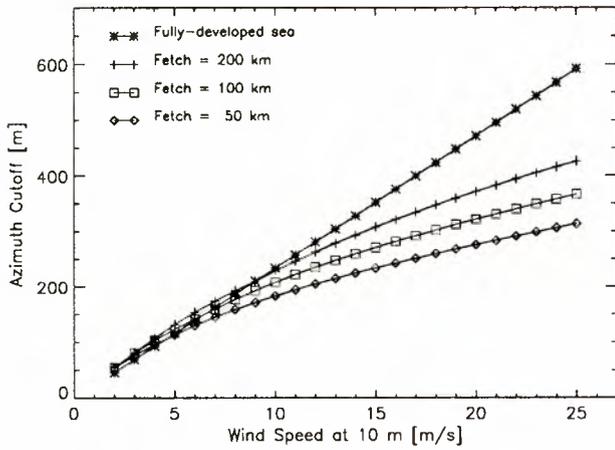


Figure 11. Estimation of the azimuth cutoff value for different wind speeds at 10 m and different fetch values (unlimited fetch, i.e., fully-developed, 200, 100, and 50 km).

5.2.2. Swell and wave relaxation biases. Apparent to these comparisons between azimuthal cutoff parameters λ and scatterometer wind speed, there are cases where the observed λ is much higher than the prediction inferred from the scatterometer wind speed. For instance, under light wind conditions, preexisting swell systems will bias the azimuthal cutoff wind speed estimate since the wind algorithm based on the empirical relationship (9) only relies on the definition of a pure wind-wave spectral model (of JONSWAP type). Thus, if we consider a sinusoidal swell system whose wavelength is l_0 and amplitude is A , the additional contribution to the total variance of azimuthal displacement is

$$\rho_{\xi\xi}(0) = \left(\frac{R}{V}\right)^2 \frac{A^2 \omega_0^2}{4} = \left(\frac{R}{V}\right)^2 \frac{\pi g A^2}{2 l_0} \quad (12)$$

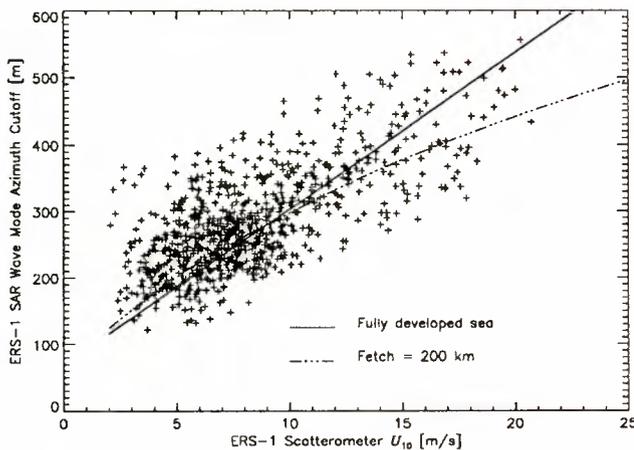


Figure 12. ERS-1 SAR wave mode imagette azimuth cutoff versus U_{10} inferred from collocated ERS-1 scatterometer data (plus). The solid line represents the theoretical relationship between SAR azimuth cutoff and U_{10} for a fully developed sea whereas the dash-dotted line represents a fetch length of 200 km.

Yet, from our comparisons, pure swell contributions to the total smear do not seem to have a strong impact except under light wind conditions (< 6 m/s). This bias is shown in Figure 13 for a swell of 200 m wavelength and 1 m amplitude. This may explain the dispersion of azimuth cutoff values for small wind speeds.

Finally, meteorological fronts can also be sources of overpredicted wind speed with the nominal relationship. Indeed, recall that spaceborne SAR images often exhibit strong gradients and striking contrasts over extended and visually well delineated spatial regions. Such patterns may be associated with a wide variety of oceanic and/or atmospheric phenomena. Under frontal conditions a conjecture is that a scatterometer or a SAR can exhibit rapid and severe attenuations due to the scarcity of capillary-gravity waves, while the azimuthal cutoff measurements which are primarily influenced by still-present longer waves react more smoothly. Such a case can be conceptualized as a low-pass filtering effect, such that some surface motion can be conserved but the Bragg-resonant waves (about 7 cm for the C band ERS SARs) almost disappear [Chapron *et al.*, 1994b]. Such a relaxation difference between scales will affect and bias the wind speed estimate from an observed λ . To resolve this problem, we will investigate the use of other statistical parameters.

5.2.3. Directionality aspects. So far, we have neglected the fact that the azimuth shifts are proportional to the surface velocity component in the direction of the radar beam. In fact, it is the rms line-of-sight orbital velocity that must be considered rather than the rms vertical velocity. One then must write

$$\lambda = \pi \sqrt{\int_0^\infty S(k) |T(k)|^2 dk} \quad (13)$$

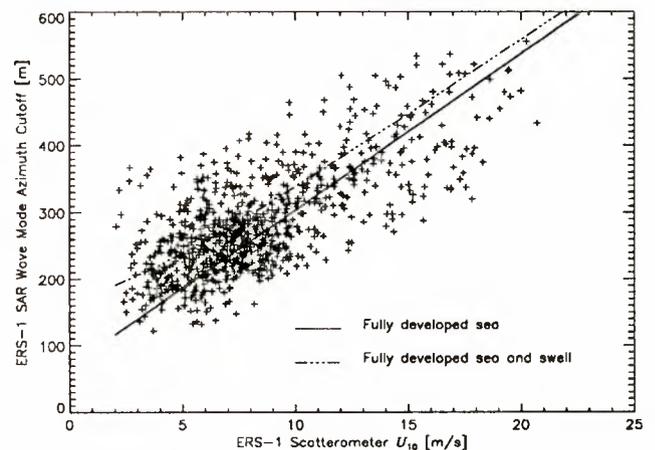


Figure 13. ERS-1 SAR wave mode imagettes azimuth cutoff versus U_{10} inferred from collocated ERS-1 scatterometer data (plus). The solid line represents the theoretical relationship between SAR azimuth cutoff and U_{10} for a fully developed sea whereas the dash-dotted line shows the contribution of an additional sinusoidal swell of wavelength 200 m and amplitude 1 m.

where $T(k)$ is the linear transfer function associated with the surface displacement. The transfer function is defined as [Krogstad and Vachon, 1994]

$$T(k) = \frac{R}{V} |\omega| \left[j \cos \theta + \frac{k_r}{k} \sin \theta \right] \quad (14)$$

where θ is the local incidence angle and k_r is the range wavenumber. If we further suppose that the spreading function of the wave spectral model is represented by a second-order cosine Fourier coefficient, the unidirectional sea spectrum is

$$S(k, \phi) = \frac{1}{k} S(k) \frac{1}{2\pi} \{1 + \Delta(k) \cos [2(\phi - \phi_0)]\} \quad (15)$$

with ϕ_0 the wind mean direction relative to the range axis. We obtain

$$\lambda = \pi \frac{R}{V} \sqrt{\int_0^\infty \omega^2 S(k) F(k, \theta, \phi_0) dk} \quad (16)$$

where,

$$F(k, \theta, \phi_0) = \cos^2 \theta + \left(\frac{\sin^2 \theta}{2} + \Delta(k) \frac{\sin^2 \theta}{4} \cos 2\phi_0 \right) \quad (17)$$

For the ERS SARs the incidence angle is 19.9° or 23.5° , so $\cos^2 \theta + \frac{1}{2} \sin^2 \theta > 0.92$. Since the absolute value of $\Delta(k)$ is bound by unity, the ratio between range and azimuth cases is at most given by

$$r = \sqrt{\left(\cos^2 \theta + \frac{3}{4} \sin^2 \theta \right) / \left(\cos^2 \theta + \frac{1}{4} \sin^2 \theta \right)} \quad (18)$$

For $\theta = 23.5^\circ$ the differences between λ_{range} and λ_{azimuth} are then negligible ($< 5\%$). The directional dependence will be weak, and only considering the rms vertical velocity is, at first, a good approximation.

6. Statistical Analysis

Besides calibration and azimuth cutoff analysis of ERS-1/2 imagettes, higher-order statistical analysis, up to fourth order, was also carried out. Indeed, the understanding and determination of SAR image distributions remain problems. The results of this study will be discussed in this section as well as the interpretation of the statistical information.

6.1. ERS-1/2 SAR Imagette Distributions

The problem of determining the distribution followed by the cross-section or the amplitude of SAR images has been widely studied and several models have been proposed: Rayleigh, K [Jakeman and Pusey, 1976], lognormal [Trunk, 1972], Weibull [Fay et al., 1977], etc. A useful method is to compare their higher-order

moments, particularly the couple (β_1, β_2) , the skewness-squared, and the kurtosis.

When the radar illuminates a large area of the sea, it is usually found that the amplitude of the signal is Rayleigh distributed. This is in accordance with the central limit theorem (CLT) since the signal can be thought of as being the sum of a large number of randomly phased, independent scatterers. In the case of a summation process of N_l independent looks in amplitude, the distribution will follow the convolution of N_l Rayleigh probability density functions (pdf). The couple (β_1, β_2) is given by

$$\beta_1 = \frac{m_{3c}^2}{m_{2c}^3} = \frac{4\pi(\pi-3)^2}{N_l(4-\pi)^3} \quad (19a)$$

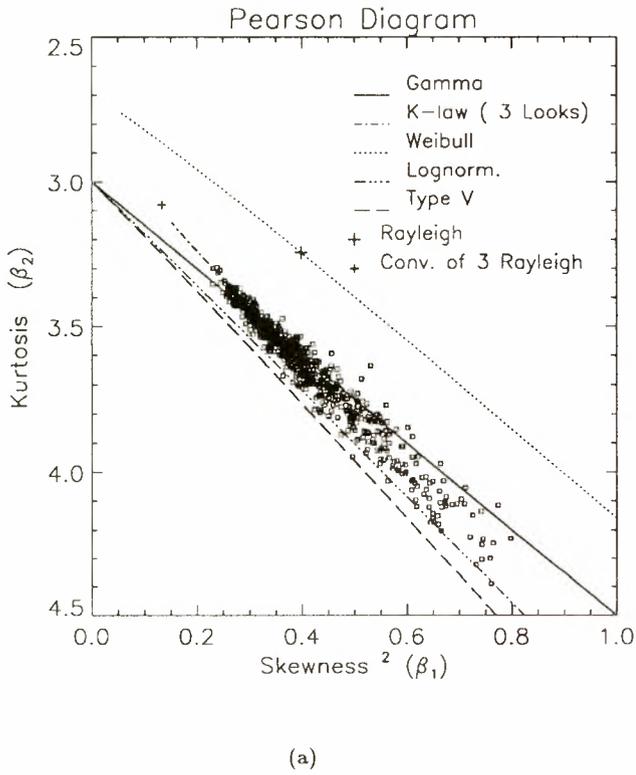
$$\beta_2 = \frac{m_{4c}}{m_{2c}^2} = 3 + \frac{24\pi - 6\pi^2 - 16}{N_l(4-\pi)^2} \quad (19b)$$

where m_{ic} represents the centered moment of i th order. For pure speckle ERS imagettes (i.e., with no texture in the scene), (β_1, β_2) is (0.13, 3.08). Examples of pure speckle have been found for SAR scenes over lakes or large glacier areas.

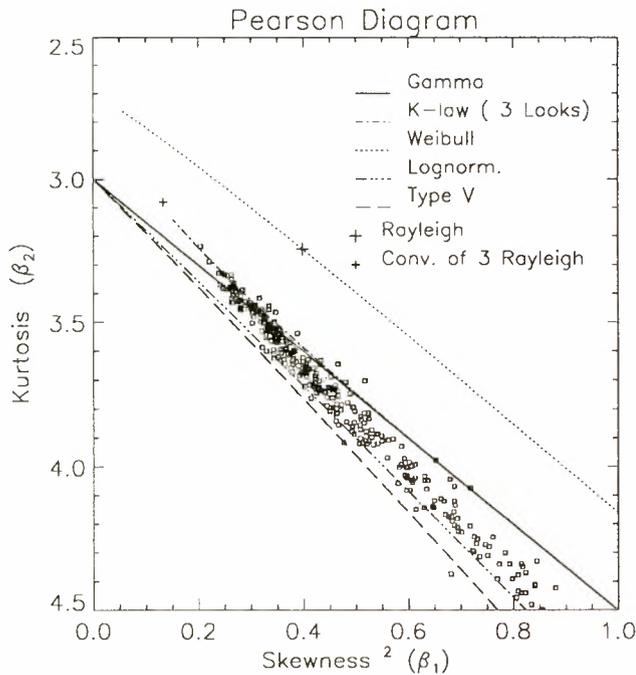
Very few imagettes have Rayleigh statistics over the ocean. Indeed, the return signal generally has a more spike-like structure, which leads to an extended tail of the SAR amplitude image pdf. The couple (β_1, β_2) for all ERS-1 and ERS-2 imagettes considered has been plotted in a Pearson diagram (Figure 14) which also shows the different distribution families [Johnson and Kotz, 1969]. The data lie on the K law three-looks in amplitude for small values of β_1 . For large values of β_1 the data lie between K and lognormal. Moreover, β_1 and β_2 take lower values for ERS-1 imagettes because of ADC saturation effect. This suggests that ADC saturation does not modify the pdf family but rather acts on the range of β_1 and β_2 . Note that only a few points are close to the Rayleigh three-looks couple. The evident departure from standard Rayleigh statistics demonstrates the violation of the CLT for most imagettes. Various explanations for such statistics have been proposed, including either non-Bragg scattering contributions [Jakeman and Pusey, 1976] or fluctuations characterized by two spatial coherence lengths due to SAR high-resolution properties [Ward, 1981; Jakeman and Tough, 1988].

6.2. Contribution of Non-Bragg Scattering

The extended tail of the SAR pdf has often been interpreted as the result of non-Bragg scattering. Thus, in the case of rough sea states, the cross-section is dominated by sea spike density due to non coherent specular scattering (e.g., wave-breaking events). Tilley and Sarma [1993] suggested that when the cross-section is dominated by sea spike density clipped in amplitude, the cross-section measurement can be approximated by the number of point-scattering processes occurring



(a)



(b)

Figure 14. Imagettes kurtosis versus skewness-squared in a Pearson diagram (squares). (a) ERS-1. (b) ERS-2.

within the integration time. This theory implies that SAR distribution degenerates to a Poisson law. *Jakeman and Pusey* [1976] suggested that the considerable enhancement of intensity fluctuation can be interpreted in terms of facet-like scattering. Indeed, large fluctua-

tions of individual form factors will lead to a decrease in the apparent number of scatterers contributing to the scattering intensity. So far, there has been no attempt to relate SAR statistics to sea spike density. It has been shown that a SAR Doppler spectrum analysis should provide useful information on spike density and time coherence [*Kerbaol et al.*, 1996].

6.3. Modulation Contribution

The departure from standard Rayleigh statistics can be interpreted from the SAR high-resolution property. The presence of only a small number of scatterers in the illuminated area might explain the violation of the CLT. However, the existence of pure speckle imagettes, that are well modeled as a Rayleigh three-look distribution, makes this rather unlikely. The average number of scatterers within a resolution cell can then be considered asymptotically large. Nevertheless, long wavelength gravity waves and swell contribute to the distortion of the overall distribution. This concept of large-scale inhomogeneities modulating a dense array of small-scale scatterers has been introduced in various ways in the literature, including the effect of fluctuations of the number of scatterers within a cell [*Jakeman and Tough*, 1988] or the representation of the two-scale ocean surface by a compound process [*Ward*, 1981]. Both models lead to the K law.

Let us consider that the observed SAR power image as being equal to $Y(x) = X(x)S(x)$ where S is the speckle governed by the Rayleigh three-look distribution and X is the unspeckled image containing the ocean wave modulation. If we assume that S and X are independent, the spectrum of Y is given by [*Goldfinger*, 1982]

$$\Phi_Y(k) = E[S]^2 \Phi_X(k) + [\Phi_X(k) + E[X]^2 \delta(k)] \otimes \Phi_S(k) \tag{20}$$

where \otimes is the convolution operator. Acknowledging that the characteristic correlation distance for the speckle is shorter than for the modulation in X , it was suggested that $\Phi_Y(k)$ could be approximated as [*Krogstad*, 1994]

$$\Phi_Y(k) = E[S]^2 \Phi_X(k) + [\text{var}(X) + E[X]^2] \Phi_S(k) \tag{21}$$

Hence, considering the relative wave modulation m and χ the normalized speckle spectrum such that $\Phi_S(k) = \text{var}(S) \chi(k)$, $\Phi_Y(k)$ becomes:

$$\Phi_Y(k) = \alpha E[Y]^2 \chi(k) + E[Y]^2 [\Phi_m(k) + \alpha \text{var}(m) \chi(k)] \tag{22}$$

where $\alpha = \sigma_S^2/E[S]^2 = .253$ for ERS-1/2 three-look power imagettes (α is 1/3 for three-look power summation). The first term in (22) represents the underlying speckle spectrum while the second represents the modu-

lation contribution. SAR power image variance to mean squared ratio is derived from (22) such that

$$\frac{\text{var}(I_{\text{SAR}})}{E[I_{\text{SAR}}]^2} = \alpha + [\text{var}(m)(1 + \alpha)] \quad (23)$$

This last equation illustrates the departure from a Rayleigh three-look distribution, providing a wave system is imaged by the instrument. SAR normalized variances (corrected for ADC saturation) have been plotted as a function of scatterometer wind speed for ERS-1 (Figure 15). Azimuth and range wind direction were selected. One can see that the wind direction can be discriminated when analyzing SAR variance to mean squared ratio. An attempt was made to simulate this parameter using a JONSWAP-like two-dimensional sea spectrum and the modulation transfer functions (MTF) as described by *Hasselmann and Hasselmann* [1991]. In addition, an average swell system was added to the sea spectrum to account for the fact that modulation was always detected, even at lowest wind speeds. There are still discrepancies between theoretical predictions and observed data. However, most of the salient features are captured by the simulation. Since the RAR modulation transfer function is based on the Bragg-scattering model, discrepancies may be related to the specular contribution. Indeed, at such an incidence angle (about 20°), quasi-specular scattering mechanisms cannot be neglected. Thus the analysis of the SAR variance to mean squared ratio parameter should help to improve the RAR MTF definition.

7. Summary

The ERS-1 and ERS-2 wave mode imagettes can offer new insights to wind/wave modellers. The interpreta-

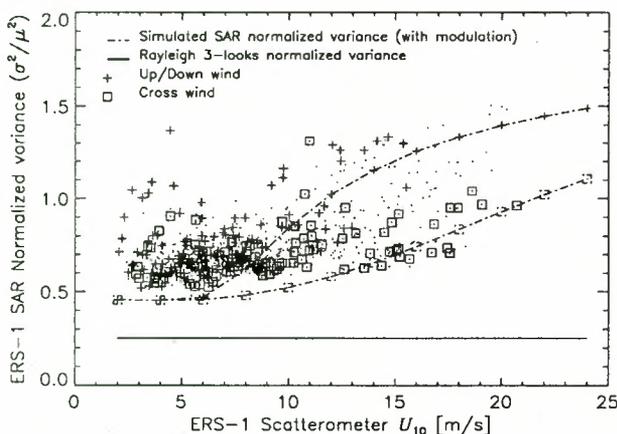


Figure 15. ERS-1 SAR wave mode imagettes power normalized variance versus U_{10} inferred from collocated ERS-1 scatterometer data for wind blowing in range direction (asterisks) and in azimuth direction (squares). The solid line represents the theoretical normalized variance value ($\alpha = 0.253$) corresponding to the square of the convolution of 3 Rayleigh laws and the dash-dotted line represents the simulated SAR normalized variance.

tion of ERS-1 SAR wave mode data in terms of normalized radar cross section (NRCS) through intensive comparisons with scatterometer estimates, clearly show saturation effects for ocean scenes. Such a problem seems to be adequately compensated for ERS-2 wave mode products after reduction of onboard gains and for ERS-1 after ADC corrections. Our results show that SAR NRCS products can be successfully interpreted using a standard scatterometer wind speed retrieval algorithm. This affirms the use of the SAR instrument as a high-resolution scatterometer. The difference in resolution between the SAR and the scatterometer does not appear to present any obstacle to the use of an operational wind speed algorithm.

The along track resolution has been studied using a time-domain method to obtain cutoff wavelength estimates. Observed azimuthal falloff confirmed the predicted limitation of high-altitude synthetic aperture radars. Convincing comparisons highlight the development of a SAR wind algorithm, from a kinematic point of view, that accounts for vertical motions of sea surface (i.e., the directivity can be neglected). This analysis suggests the use of an alternative method to obtain an independent wind speed measurement. However, one needs to be cautious with the estimation bias caused by the inertia of gravity waves in the case of fetch limitation or relaxation effects. Further investigation is required to extend this algorithm to other sensors such as RADARSAT or ENVISAT. In particular, sea surface motions can no longer be reduced to vertical displacements at larger incidence angle, and the directivity cannot be neglected. Moreover, acknowledging that HH polarization SAR instruments are more sensitive to short lifetime, fast-moving scatterers [*Jessup et al.*, 1990; *Apel*, 1994; *Smith et al.*, 1996], collocation between ERS-2 and RADARSAT SAR data may be used to address this issue.

It was also shown that the statistical distribution of the SAR wave mode products is well characterized by an extended-tail distribution lying between K law and lognormal law in the Pearson diagram. Furthermore, an attempt was made to reproduce SAR power image variance by simulating the SAR transform of a JONSWAP-like two-dimensional sea spectrum using standard modulation transfer functions as described by *Hasselmann and Hasselmann* [1991]. This leads to the conclusion that the departure from the standard Rayleigh three-look pdf was mainly due to the SAR high resolution. Further investigations may be envisaged to globally define a standard RAR modulation transfer function to improve SAR inversion algorithms.

To conclude, following the methodology presented here, the analysis of SAR data in the spatial domain rather than exclusively in the spectral domain also offers useful information for operational purposes. An independent wind speed product, based on azimuth cutoff analysis, is proposed. This would complement the scatterometry-type analysis.

Table A1. Summary of K' Values.

Sensor	Number of Imagettes	K'_{dB}	Incidence Angle	Data Set
ERS-1	1092	76.21	$\theta = 19.9^\circ$ (prior to June 1995)	North Pacific Ocean
ERS-1	693	77.55	$\theta = 23.5^\circ$ (after June 1995)	Indian Ocean
ERS-2	342	73.78	$\theta = 23.5^\circ$ (after April 1995)	Indian Ocean

Appendix: Calibration Constant Retrieval

Following the assumption of a 1-to-1 ratio between $I_{SAR} - Loss_{dB} - K'_{dB}$ and σ°_{SCAT} , the mean square estimator for K' is given by

$$K' = \frac{1}{n} \sum_{i=1}^n (I_{SAR} - Loss_{dB} - \sigma^{\circ}_{SCAT}) \quad (A1)$$

This estimator is expected to be consistent for both data sets over the Indian Ocean (ERS-1 and ERS-2 at 23.5°). In the case of the North Pacific Ocean (ERS-1 at 19.9°) the estimator should only be applied to σ° values of scatterometer that are below -2 dB (above this level, ADC correction is not expected to give satisfactory results). Table A1 presents the results of this estimator for both incidence angle configurations of ERS-1 and for ERS-2.

Acknowledgment. This work was partly supported by a contract from the European Union, program MAST, entitled European Community Atmosphere Wave Ocean Model (ECAWOM), MAS2-CT94-0091.

References

- Alpers, W.R., and C. Brüning, On the relative importance of motion-related contributions to the SAR imaging mechanism of ocean surface waves, *IEEE Trans. Geosci. and Remote Sens.*, GE-24(6), 873-885, 1986.
- Alpers, W.R., and C.L. Ruffenach, The effect of orbital motions on synthetic aperture radar imagery of ocean waves, *IEEE Trans. Antennas Propag.*, AP-27(5), 685-690, 1979.
- Alpers, W.B., D.B. Ross, and C.L. Ruffenach, On the detectability of ocean surface waves by real and synthetic aperture radar, *J. Geophys. Res.*, 86(C7), 6481-6498, 1981.
- Apel, J.R., An improved model of the ocean surface wave vector spectrum and its effect on radar backscatter, *J. Geophys. Res.*, 99(C8), 16269-16291, 1994.
- Bao, M., C. Brüning, and W. Alpers, A generalized nonlinear ocean wave SAR spectral integral transform and its application to ERS-1 SAR ocean wave imaging., in *Proceedings Second ERS-1 Symposium*, 219-224, ESA Publications Division, ESTEC, Noordwijk, The Netherlands, 1994.
- Beal, R.C., D.G. Tilley, and F.M. Monaldo, Large and small-scale spatial evolution of digitally processed ocean wave spectra from Seasat synthetic aperture radar, *J. Geophys. Res.*, 88(C3), 1761-1778, 1983.
- Brillinger, D.R., Distributions of particle displacements via higher-order moment functions, *Proc. IEEE, Part. F*, 140(6), 390-394, 1993.
- Brüning, C., W. Alpers, and K. Hasselmann, Monte-Carlo simulations studies of the non-linear imaging of a two-dimensional surface wave field by a synthetic aperture radar, *Int. J. Remote Sens.*, 11(10), 1695-1727, 1990.
- Brüning, C., S. Hasselmann, and K. Hasselmann, First evaluation of ERS-1 synthetic aperture radar wave mode data, *Global Atmos. Ocean Syst.*, 2, 61-98, 1994.
- Chapron, B., R. Garello, V. Kerbaol, and J.M. Lefevre, Non-linear theory of ocean-SAR transformation and statistical analysis of ERS-1 SAR wave mode imagettes, in *Proceedings Second ERS-1 Symposium*, 247-250, ESA Publications Division, ESTEC, Noordwijk, The Netherlands, 1994a.
- Chapron B., D. Vandemark, and F.C. Jackson, Airborne measurements of the ocean's Ku-band radar cross-section at low incidence angle, *Atmos. Ocean*, 32(1), 179-193, 1994b.
- Chapron, B., T. ElFouhaily, and V. Kerbaol, Calibration and validation of ERS wave mode products, *Tech. Rep. DRO/OS/95-02*, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, 1995.
- ElFouhaily, T., B. Chapron, K. Katsaros, and D. Vandemark, A unified directional spectrum for long and short wind-driven waves, *J. Geophys. Res.*, 102(C7), 15781-15796, 1997.
- Engen, G., and H. Johnsen, Analysis and inversion of ERS-1 image cross-spectra, in *IGARSS'95*, 1863-1865, IEEE Press, Piscataway, N.J., 1995.
- European Space Agency, ERS-1 system, Eur. Space Agency Spec. Publ., ESA SP-1146, 26-27, 1992a.
- European Space Agency, ESA ERS-1 product specification, Eur. Space Agency Spec. Publ., ESA SP-1149, 1992b.
- Fay, F.S., J. Clarke, and R.S. Peters, Weibull distribution applied to sea clutter, *IEE Conf. Publ.*, 105, 101-104, 1977.
- Goldfinger, A.D., Estimation of spectra from speckled images, *IEEE Trans. Aerosp. Electron. Syst.*, AES-18(5), 675-681, 1982.
- Hasselmann, K., and S. Hasselmann, On the nonlinear mapping of an ocean wave spectrum into a synthetic aperture radar image spectrum and its inversion, *J. Geophys. Res.*, 96(C6), 10713-10729, 1991.
- Hasselmann, K., R.K. Raney R.K., W.J. Plant, W. Alpers, R.A. Shuchman, D.R. Lyzenga, C.L. Ruffenach, and M.J. Tucker, Theory of synthetic aperture radar ocean imaging: A MARSEN view, *J. Geophys. Res.*, 90(C3), 4659-4686, 1985.
- Jackson, F.C. and C.Y. Peng, Comment on "Imaging radar observations of directional properties of ocean waves" by W. McLeish and D.B. Ross, *J. Geophys. Res.*, 90(C4), 7367-7370, 1985.
- Jakeman, E., and P.N. Pusey, A model for non-Rayleigh sea echo, *IEEE Trans. Antennas Propag.*, AP-24(6), 806-814, 1976.
- Jakeman, E., and R.J.A. Tough, Non-gaussian models for the statistics of scattered waves, *Adv. Phys.*, 37(5), 471-529, 1988.
- Jessup, A.T., W.C. Keller, and W.K. Melville, Measurements of sea spikes in microwave backscatter at moderate incidence, *J. Geophys. Res.*, 95(C6), 9679-9688, 1990.

- Johnson, N.L., and S. Kotz, *Distribution in Statistics: Continuous Univariate Distributions*, vol. I, John Wiley, New York, 1969.
- Kanevsky, M.B., On the theory of SAR ocean wave imaging, *IEEE Trans. Geosci. and Remote Sens.*, 31(5), 1031-1035, 1993.
- Kerbaol, V., and B. Chapron, Calibration and validation of ERS-1/2 wave mode products, *Eur. Space Agency Rep. PO NO 160709*, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, October 1996.
- Kerbaol, V., B. Chapron, T. ElFouhaily, and R. Garello, Fetch and wind dependence of SAR azimuth cutoff and higher order statistics in a mistral wind case, in *IGARSS'96*, 621-624, IEEE Press, Piscataway, N.J., 1996.
- Korsbakken, E., Quantitative wind field retrievals from ERS SAR images, *Tech. Rep.*, Eur. Space Agency (ESTEC), 1996.
- Krogstad, H.E., A simple derivation of hasselmann's nonlinear ocean-synthetic aperture radar transform, *J. Geophys. Res.*, 97(C2), 2421-2425, 1992.
- Krogstad, H.E., Some comments about the processing of ERS-1/2 imagettes and the wave mode product, *Tech. Rep. STF10 A94008*, SINTEF Industrial Mathematics, Trondheim, Norway, 1994.
- Krogstad, H.E., and P.W. Vachon, Generalizations of the non-linear ocean-SAR transform and a simplified SAR inversion algorithm, *Atmos. Ocean*, 32(1), 61-82, 1994.
- Laur, H., P.J. Meadows, J.I. Sanchez, and E. Dwyer, ERS-1 SAR radiometric calibration, in *Proceedings of the CEOS SAR Calibration Workshop*, 257-281, ESTEC, Noordwijk, The Netherlands, ESA WPP-048, 1993.
- Laur, H., P. Bally, P. Meadows, J. Sanchez J., B. Schaeffler, and E. Lopinto, ERS SAR calibration - Derivation of backscattering coefficient in ESA ERS SAR PRI products, *Tech. Note ES-TN-RS-PM-HL09*, Eur. Space Agency - Eur. Space Res. Inst., Frascati, Italy, 1996.
- Lyzenga, D.R., Numerical simulation of synthetic aperture radar image spectra for ocean waves, *IEEE Trans. Geosci. and Remote Sens.*, GE-24(6), 863-872, 1986.
- Ouchi, K., and R.A. Cordey, Statistical analysis of azimuth streaks observed in digitally processed CASSIE imagery of sea surface, *IEEE Trans. Geosci. and Remote Sens.*, 29(5), 727-735, 1991.
- Quilfen, Y., ERS-1 off-line wind scatterometer products, *Tech. Rep. IFREMER-SCAT/IOA/DOS-01*, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, 1993.
- Scoon, A., I.S. Robinson, and P.J. Meadows, Demonstration of an improved calibration scheme for ERS-1 SAR imagery using a scatterometer wind model, *Int. J. Remote Sens.*, 17(2), 413-418, 1996.
- Smith, M.J., E.M. Poulter, and J.A. McGregor, Doppler radar measurements of wave group and breaking waves, *J. Geophys. Res.*, 101(C6), 14269-14282, 1996.
- Steinberg, H., P.M. Schulteiss, C.A. Wogrin, and F. Zweig, Short-time frequency measurements of narrow-band random signals by means of a zero-counting process, *J. Appl. Phys.*, 26(2), 195-201, 1955.
- Stoeffelen, A., and D.L.T. Anderson, ERS-1 scatterometer data characteristics and wind retrieval skill, *Adv. Space Res.*, 13, 553-560, 1993.
- Swift, C.T., and L.R. Wilson, Synthetic aperture radar imaging of moving ocean waves, *IEEE Trans. Antennas Propag.*, AP-27(6), 725-729, 1979.
- Tilley, D.G., and Y.V. Sarma, A comparison of synthetic aperture radars applied for satellite remote sensing of the ocean surface, *Trends Geophys. Res.*, 2, 467-486, 1993.
- Tournadre, J., and S. Blanquet S., Wind and wave mesoscale variability from in-situ and altimeter data, *Global Atmos. Ocean Syst.*, 2, 221-247, 1994.
- Trunk, G.V., Radar properties of non-Rayleigh sea clutter, *IEEE Trans. Aerosp. Electron. Syst.*, AES-8(2), 196-204, 1972.
- Tucker, M.J., The imaging of waves by satelliteborne synthetic aperture radar: the effect of sea-surface motion, *Int. J. Remote Sens.*, 6(7), 1059-1074, 1985.
- Vachon, P.W., and F.W. Dobson, Validation of wind vector retrieval from ERS-1 SAR images over the ocean, *Global Atmos. Ocean Syst.*, 5, 177-187, 1996.
- Vachon, P.W., H.E. Krogstad, and J. Scott Paterson, Airborne and spaceborne synthetic aperture radar observations of ocean waves, *Atmos. Ocean*, 32(1), 83-112, 1994.
- Vesceky, J.F., and R.H. Stewart, The observation of ocean surface phenomena using imagery from the Seasat synthetic aperture radar, *J. Geophys. Res.*, 87(C5), 3397-3430, 1982.
- Ward, K.D., Compound representation of high resolution sea clutter, *Electron. Lett.*, 17, 561-563, 1981.

B. Chapron and V. Kerbaol, Département d'Océanographie Spatiale, IFREMER, centre de Brest, BP 70, 29280 Plouzané, France. (e-mail: Bertrand.Chapron@ifremer.fr; Vincent.Kerbaol@ifremer.fr)

P. W. Vachon, Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Ont. K1A 0Y7, Canada. (e-mail: vachon@ccrs.emr.ca)

(Received December 9, 1996; revised May 16, 1997; accepted May 29, 1997.)

Mesoscale wind measurements using recalibrated ERS SAR images

S. Lehner

Deutsche Forschungsanstalt für Luft- und Raumfahrt, Deutsches Fernerkundungsdatenzentrum
Weßling, Germany

J. Horstmann, W. Koch, and W. Rosenthal

GKSS Research Centre, Geesthacht, Germany

Abstract. The precision images (PRI) of the synthetic aperture radars (SAR) on board the European Remote Sensing Satellites ERS-1 and ERS-2 are used to derive mesoscale wind fields over the ocean. For calculation of the wind speed the C-band model (CMOD4) is used, which was originally developed by *Stoffelen and Anderson* [1993] for the European Space Agency (ESA) to derive wind fields from measurements of the wind scatterometer (SCAT). In the case of the ERS-1/2 SAR the CMOD4 is used to compute the wind speed from the normalized radar backscatter cross section (NRCS) and the incidence angle of the radar beam, both computed from the SAR.PRI data. The third input variable is the wind direction, which is estimated from the wind streaks in the images or from ground truth measurements. The SAR data are affected by a power loss, caused by saturation of the analog to digital converter (ADC) of the SAR. Therefore the images have to be recalibrated. Errors in the derived wind speed are mainly due to ADC saturation and uncertainties of the input wind direction. These errors are estimated for various wind conditions. Mesoscale wind fields computed from ERS-1/2 SAR.PRI images taken between the Shetland Islands and the west coast of Norway are compared to ground truth measurements and modeled wind fields from the German weather service (DWD). Wind fields of the nonhydrostatic mesoscale model Geesthacht simulation model of the atmosphere (GESIMA) are compared to the derived wind field of the ERS-1 SAR.PRI image at the island Rügen in the Baltic Sea.

1. Introduction

The grid cell of state of the art meteorological models is of the order of 10 km. Finer spatial resolution is needed to model a variety of coastal processes, e.g., currents, waves, wind, and related transport processes. Validation with conventional ground truth measurements usually requires enormous efforts in large campaigns.

The spatial resolution of 25 m together with the coverage of 100×100 km makes the synthetic aperture radar (SAR) on board of the European Remote Sensing Satellite (ERS-1/2) especially valuable for measuring spatially inhomogeneous wind and wave fields. Therefore radiometrically calibrated SAR images offer a unique opportunity to make synoptic mesoscale wind measurements that were never possible before. Figure 1 shows an ERS-1 SAR image at the south coast of

Norway with superimposed wind speeds derived with the C-band model (CMOD4) out of the SAR precision image (PRI) data.

The object of the investigation is to derive mesoscale wind fields over the ocean surface from the normalized radar backscatter cross section (NRCS) of ERS SAR data and compare these results to model and ground truth data. Wind speeds from ERS SAR images of the ocean surface have been derived, e.g., by *Alpers and Brümmer* [1994], *Chapron et al.* [1994], *Rosenthal et al.* [1995], *Johannessen et al.* [1996], and *Scoon et al.* [1996]. In these articles either recalibration is not considered or extensive comparisons to models and ground truth are not given.

The backscatter from the rough ocean surface for moderate incidence angles (20° – 70°) is dominated by resonant Bragg scattering [*Valenzuela*, 1978]. Additionally, Raleigh scattering and specular reflection may contribute to the backscatter. Following *Wright* [1966], the backscatter signal is caused by the water wave component which is in resonance with the incidence radiation. The prime resonant water wave number k_w is related to the electromagnetic wave number k_{el} of the radar

Copyright 1998 by the American Geophysical Union

Paper number 97JC02726.
0148-0227/98/97JC-02726\$09.00

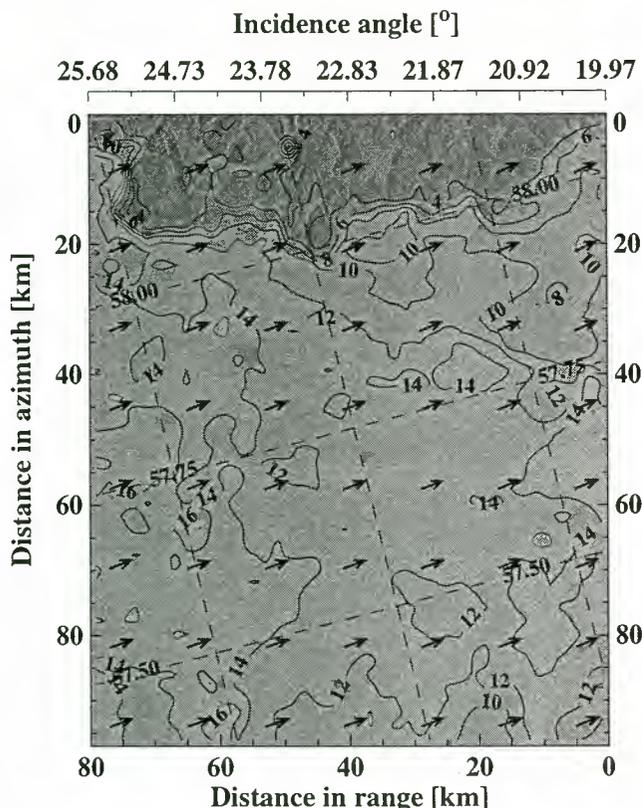


Figure 1. ERS-1 synthetic aperture radar (SAR) image from September 26, 1995, at 1036 UTC. The isotachs are given in m s^{-1} and were derived using the C band model (CMOD4) with a constant wind direction of 270° .

according to

$$k_w = 2k_{el} \sin \alpha, \quad (1)$$

where α is the incidence angle of the radar beam. In case of the ERS SAR, operating at C-band with incidence angles between 20° and 26° , the range of scattering wavelengths extends from 8.2 to 6.5 cm. Therefore the NRCS can be used to evaluate parameters which influence the small-scale roughness, like the wind speed.

For the determination of the wind speed and direction over the ocean surface from the three antennae of the ERS scatterometer (SCAT), an empiric C-band model CMOD4 [Stoffelen and Anderson, 1993] was developed, a revision of which is described by Stoffelen and Anderson [1997]. The SCAT and SAR on board ERS-1/2 operate at the same frequency. Thus the CMOD4 can be applied to the SAR. In contrast to the SCAT, the SAR collects data only from one antenna. Therefore the wind direction is needed as further input to derive the wind speed by the CMOD4. Usually, SAR images show distinct features like wind streaks or shadowing behind coasts from which the large-scale wind direction can be derived.

It is obvious that for this procedure, accurately calibrated SAR images are needed. Unfortunately, while studying the characteristic range dependence of backscatter [Meadows and Wright, 1994; Scoon et al. 1996]

on the ocean surface, it became evident that ERS-1 SAR data are not properly calibrated; the measured drop-off in backscatter between near range to far range of the SAR image can be up to 5 dB lower than the expected value of about 6 dB for an ocean surface.

This phenomenon can be assigned to the receiver gain setting causing a too high input power into the analog to digital converter (ADC). It gets saturated when large homogenous areas of relatively high backscatter are imaged [Meadows and Wright, 1994]. Because of ADC saturation the SAR data are affected by a power loss. This is especially the case for ocean surfaces in medium to strong wind conditions and on inland ice, but inhomogeneous areas of large backscatter, like cities, show the same effect. In the case of the ocean surface the near-range part of the image, where the backscatter is higher, is most strongly affected. The European Space Agency (ESA) tried to avoid this problem for the second satellite ERS-2 by a reduction of the receiver gain setting of the SAR system (about 4 dB lower). Therefore the problem of ADC saturation is less severe for ERS-2 SAR images, but it is still present. In consequence, ERS SAR images used for determination of wind fields over the ocean have to be recalibrated.

2. Estimation of Wind Direction

Figure 2 depicts an ERS-1 SAR image acquired on the August 12, 1991, at 2107 UTC, showing the island of Rügen, situated at the German coast in the Baltic Sea. Typical for this area is brackish water and a structured coastline with several lagoons called "Bodden." For this area, mesoscale models of wind speed, sea state, currents, and transport-related processes were developed that are compared to respective SAR images. An application can be found in work by Wolf and Gayer [1995] and in work by Horstmann and Wolf [1997].

On the sea surface the typical features of a mesoscale wind field, like shadowing behind the cliffs at the eastern part of the island, wind streaks, and varying wind speed in the Bodden can be observed. Where wind streaks are visible, they can be used to estimate the wind direction up to 180° ambiguity, finding the dominant direction by fast Fourier transform (FFT) methods.

The solid black lines in Figure 2 give the orientation perpendicular to the main spectral power of wind streaks as derived by FFT. For the thin lines an area of 5×5 km was used, and for the thick lines an area of 10×10 km was used. For the estimation a regression weighted with the energy densities for wavelengths between 500 and 1500 m was computed. As an example, Figure 3 shows the image power spectrum of a 10×10 -km subset taken north of Rügen. The main spectral energy is located perpendicular to the orientation of the streaks, which give the wind direction with a 180° ambiguity. Because of the shadowing the 180° ambiguity could be removed, resulting in a wind direction of 296° .

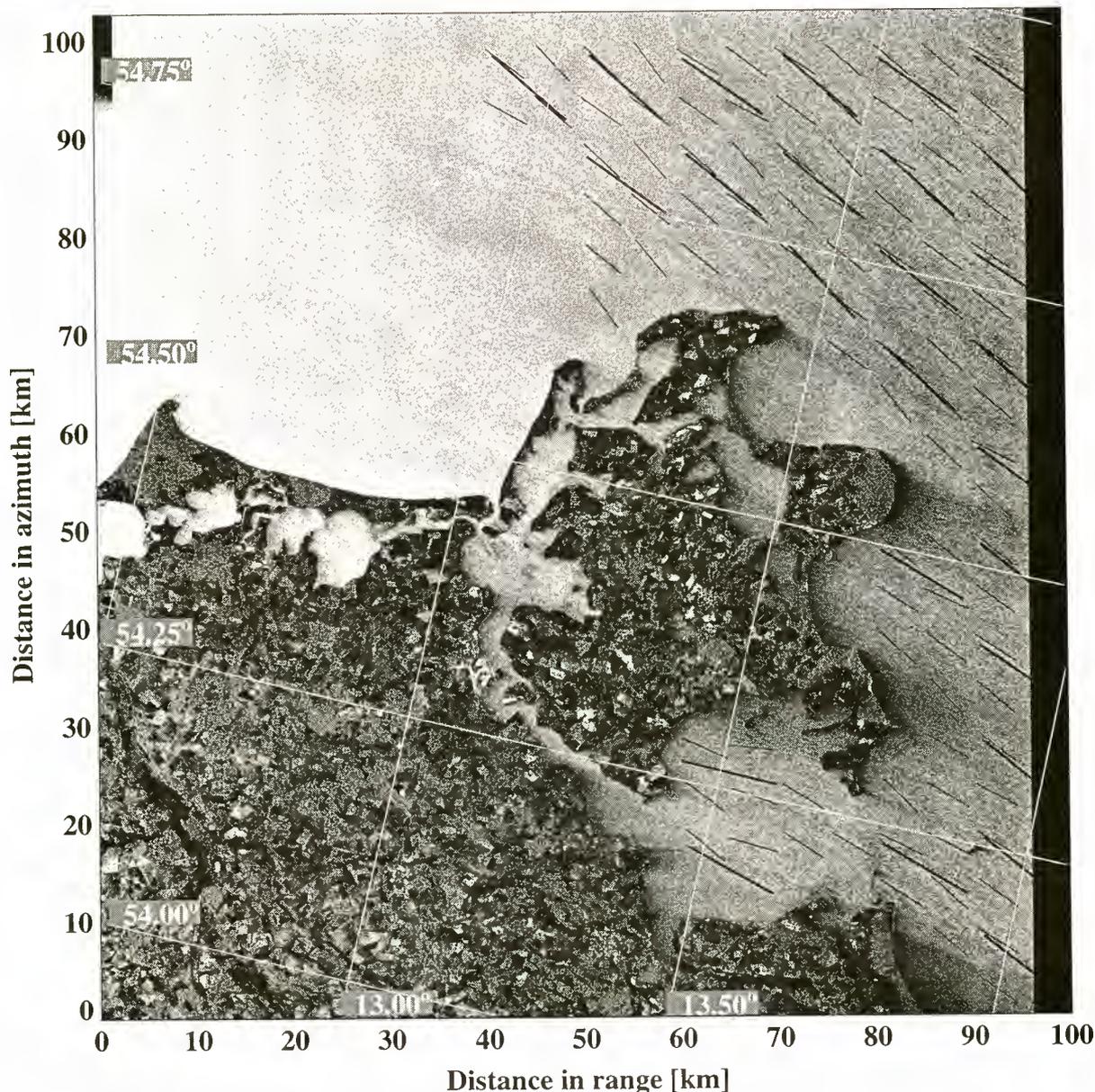


Figure 2. ERS-1 SAR image of the island of Rügen at the German coast of the Baltic Sea, taken on August 12, 1991, at 2107 UTC. The solid black lines show the orientation of streaks in the image as computed by fast Fourier transform. The thick lines were computed for a 10- × 10-km area and the thin lines for a 5- × 5-km subarea.

This agrees well with the measured ground truth wind direction of 282° at Arkona at the north tip of Rügen.

In the left part of Figure 2, no wind streaks could be detected, and the spectra yielded no reasonable directional information. For the calculation of wind speed for Figure 2 the wind direction was taken to be a constant 290° over the whole image.

Using this procedure, wind direction and wind speed are derived at different scales attributing all the variability to wind speed, while the direction is held constant. In the upper right corner of Figure 2 it can be seen that the error due to this procedure between the 5- and 10-km scale is rather small at $\sim 10^\circ$. In the Bodden area a higher variability is to be expected.

This method fails when no wind streaks or wind shadows are detectable, but in 70% of the SAR images considered, wind streaks could be detected and were used as input for the wind direction into the CMOD4. The error in wind speed due to error in wind direction is discussed in section 6.

3. Effect of Power Loss on Wind Measurements

A first-calibration procedure converting intensities, I (amplitudes squared), of the SAR.PRI data to the NRCS is given by Laur [1992]

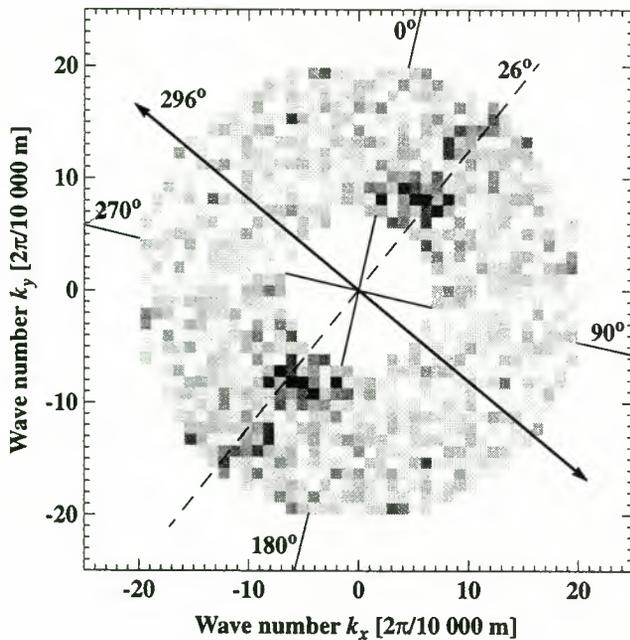


Figure 3. Image power spectrum of an area in the open water north of Rügen. Wave numbers taken for estimation of wind direction correspond to wavelengths from 500 to 1500 m. The main energy is along an orientation of 26° versus north; the wind streaks are orientated perpendicular.

$$\text{NRCS} = \frac{I}{K} \frac{\sin \alpha}{\sin \alpha_{\text{ref}}}, \quad (2)$$

where K is a calibration constant at the reference angle α_{ref} (23°) of the SAR.PRI images, which unfortunately depends on the individual processing and archiving facility processor and depends on the processing date. A list of K values is available at ESA; for the German processing and archiving facility (D-PAF), K is 666, 110. for ERS-1 and 944, 064. for ERS-2. The factor $\sin \alpha \times \sin^{-1} \alpha_{\text{ref}}$ takes the variation of K with the incidence angle α between 19° and 26° into account. Depending on the PAF and the processing date, a cor-

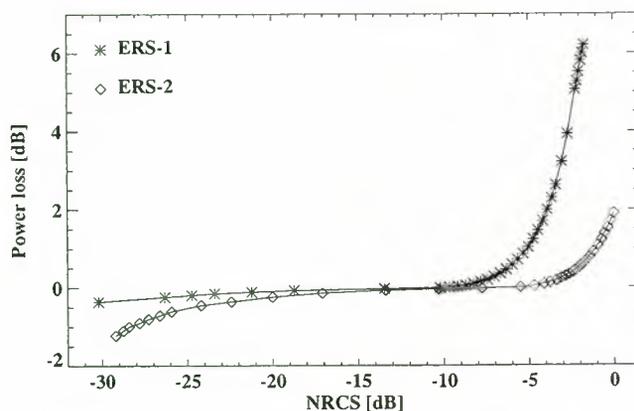


Figure 4. Power loss correction for ERS-1 and ERS-2 SAR images.

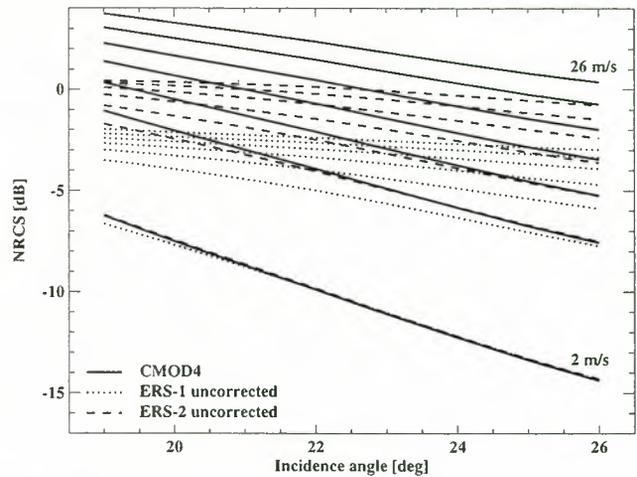


Figure 5. Estimate of the effect of power loss on wind measurements derived with the CMOD4 from ERS-1 and ERS-2 SAR images in the case of upwind for wind speeds from 2 to 26 m s⁻¹ in steps of 4 m s⁻¹.

rection for the antenna pattern has to be applied. A detailed list is available at the ESA help desk. An initial contact is available at <http://earthnet.esrin.esa.it/>.

In addition to the effects mentioned in (2), the new calibration procedure takes the power loss due to the ADC saturation and the changes in replica power into account. The replica power is part of the internal SAR calibration. Details are given in the paper by Laur *et al.* [1996]. The resulting complete calibration formula is of the form

$$\text{NRCS} = \frac{I}{K} \frac{\sin \alpha}{\sin \alpha_{\text{ref}}} \frac{\text{IRP}}{\text{RRP}} \text{power loss}, \quad (3)$$

where IRP denotes the replica power of the image and RRP denotes the replica power of the reference image. While all the other corrections are in the range of 1 dB and may be important for an accurate calibration, the power loss correction can be higher than 6 dB and is thus crucial.

Figure 4 shows the power loss correction for ERS-1 and ERS-2 SAR images. The NRCS is averaged over the area which contributes to the signal at the input to the ADC. It can be seen that for ERS-1 a power loss correction larger than 1 dB needs to be applied only for scenes with a NRCS higher than -6 dB. This is usually only the case for rough ocean surfaces, inland ice, and areas containing a lot of targets with high backscatter values such as towns. Because of the different gain settings of the ADC on the ERS-2 SAR, power loss corrections above 1 dB are only necessary for a NRCS larger than -1 dB. For ERS-2, there is a power gain for NRCS values below -20 dB, which is due to quantization effects for low return values. As NRCS values below -17 dB would correspond to wind speeds below 2 m s⁻¹, this effect is of no relevance to wind speed measurements.

In Figure 5 we give a theoretical estimate of the effect of power loss on wind measurements from ERS-1

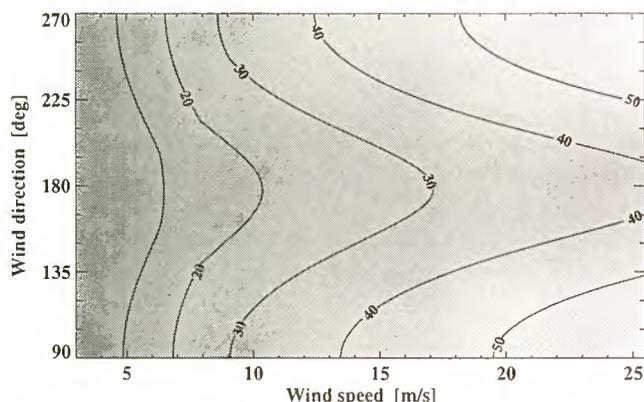


Figure 6. Estimated relative reduction in percentage of wind speed derived by CMOD4 due to the power loss of ERS-1 SAR for an incidence angle of 23° . The wind direction is given clockwise from flight direction; for example, 90° is corresponding to upwind.

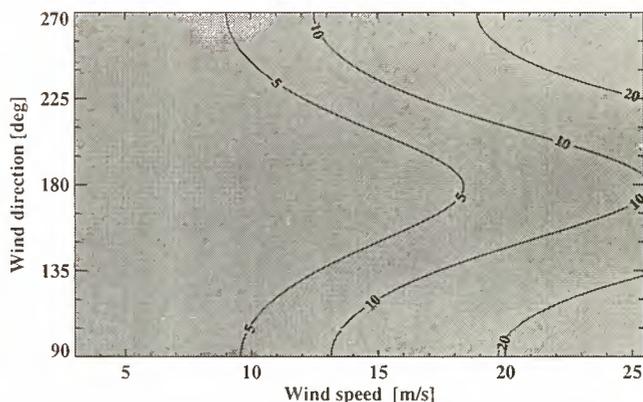


Figure 7. Estimated relative reduction in percentage of wind speed derived by CMOD4 due to the power loss of ERS-2 SAR for an incidence angle of 23° . The wind direction is given clockwise from flight direction; for example, 90° is corresponding to upwind.

and ERS-2 SAR images. The solid lines show the expected range dependence of NRCS values for ERS-1/2 calculated by CMOD4 at wind speeds between 2 and 26 m s^{-1} when the wind is blowing toward the antenna. Using the table of the power loss correction algorithm, the dotted lines show the expected NRCS for ERS-1 SAR, and the dashed ones show the expected NRCS for ERS-2 SAR at the respective wind speeds. The NRCS for upwind and downwind is higher than for cross wind. In the former case the power loss has the largest influence. For ERS-1 SAR images in near range (α about 19.6°), already for wind speeds above 6 m s^{-1} , the NRCS exceeds the critical value of -2 dB , where power loss above 1 dB occurs. In far range, significant power loss occurs for wind speeds above 10 m s^{-1} . For ERS-2 the situation is less severe; power loss occurs in near range above wind speeds of about 10 m s^{-1} and at far range of about 18 m s^{-1} .

In Figures 6 and 7 the relative reduction in percentage of wind speed derived by the CMOD4 due to the power loss is plotted as a function of wind direction and wind speed for the ERS-1 and the less affected ERS-2 SAR images for an incidence angle of 23° . The conclusion is

that at least for ERS-1 images, recalibration is always necessary.

4. Comparisons to Ground Truth

CMOD4-derived wind speeds from 20 recalibrated ERS-1/2 SAR images at different geographical locations are compared to the analyzed wind fields of the German weather service (DWD) model, with a grid size of $0.5^\circ \times 0.5^\circ$, 10-min-mean wind speed measurements of the German research vessel (R/V) *Gauss* operated by the Bundesamt für Seeschifffahrt und Hydrographie (BSH) and from measurements of the Royal Netherlands Meteorological Institute (KNMI) [Koek, 1995a, b, c] at the oil platform Cormorant and the lighthouse Lista Fur. Table 1 lists the SAR frames considered.

Figure 8 shows an ERS-1 SAR image taken on December 1, 1995, at 1101 UTC near the Gullfaks oil field. The oil platforms can be seen as bright spots, and one of them is pointed at by the white arrow. The position of the R/V *Gauss* is denoted by a star.

In this case, wind direction can be taken from the wind streaks which give the wind direction with a 180°

Table 1. List of the ERS-1/2 SAR Images Considered for Comparison

Orbit	Frame	Date	Time, UTC	Orbit	Frame	Date	Time, UTC
21539	2439	Aug. 28, 1995	1047	1866	2439	Aug. 29, 1995	1047
21682	2457	Sep. 07, 1995	1033	2009	2457	Sep. 08, 1995	1033
21789	1161	Sep. 14, 1995	2140	2095	2439	Sep. 14, 1995	1045
21811	2421	Sep. 16, 1995	1050	2138	2421	Sep. 17, 1995	1050
21911	2439	Sep. 23, 1995	1030	2216	1161	Sep. 22, 1995	2120
21954	2439	Sep. 26, 1995	1036	2259	1161	Sep. 25, 1995	2125
21975	1161	Sep. 27, 1995	2131	2281	2439	Sep. 27, 1995	1036
21997	2439	Sep. 29, 1995	1042	2302	1161	Sep. 28, 1995	2131
22018	1161	Sep. 30, 1995	2136	2324	2439	Sep. 30, 1995	1042
22899	2367	Dec. 01, 1995	1101	3226	2367	Dec. 02, 1995	1101

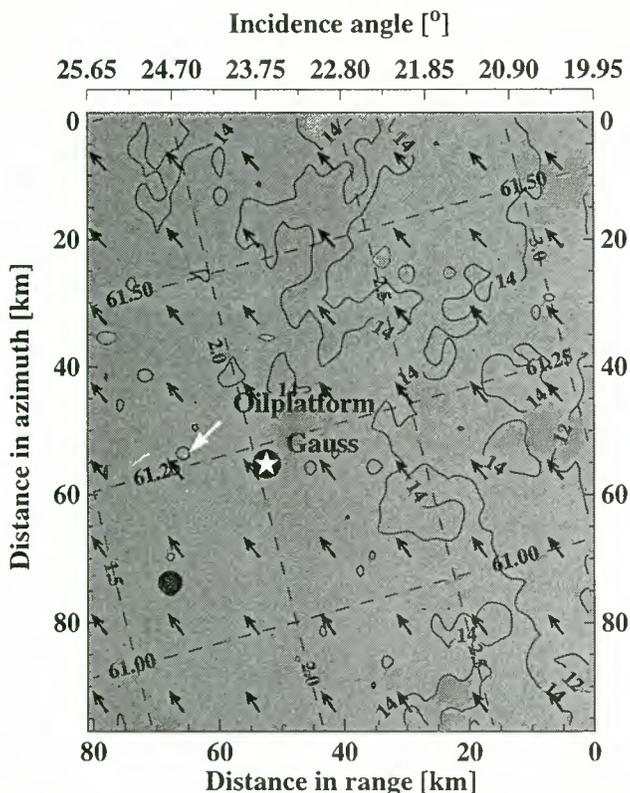


Figure 8. ERS-1 SAR image 22899-2367 taken on December 1, 1995, at 1101 UTC in the area of the Gullfaks oil field west of the Norwegian coast. The isotachs are given in m s^{-1} . The star denotes the position of the research vessel (R/V) *Gauss*. The arrow points to a platform in the Statfjord field.

ambiguity. By looking at the wind shadowing behind the oil platforms the ambiguity can be removed, and the wind direction is estimated to 170° (wind streaks and shadows are present in the SAR image but are not reproduced in Figure 8). Using the CMOD4 and this wind direction, wind speeds are calculated for the whole SAR image. Figure 9 shows the mean NRCS from the recalibrated SAR image in comparison to the expected NRCS from CMOD4 for constant wind speeds of 11, 13, and 15 m s^{-1} . The mean wind speed over the whole image is 14.3 m s^{-1} ; only in the immediate near range it drops down to $\sim 12.5 \text{ m s}^{-1}$. Table 2 gives a comparison of the computed wind speeds to the different ground truth measurements.

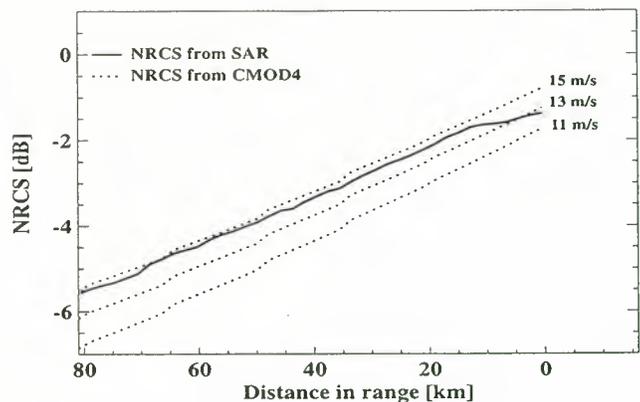


Figure 9. Mean normalized radar backscatter cross section (NRCS) in range of ERS-1 SAR image versus incidence angle in comparison to the NRCS for 11, 13, and 15 m s^{-1} computed by CMOD4.

The KNMI received the values from offshore platforms. It was not possible to get detailed information on measurement and calibration of the data. The sensor height at those platforms is normally higher than 50 m, and wind speed is reduced by a prescribed wind profile. The analyzed wind fields of the DWD are computed from an atmospheric model on a $0.5^\circ \times 0.5^\circ$ grid and improved by inserting available ground truth measurements. The data from R/V *Gauss* are taken by a sensor in 19.5-m height, which is regularly maintained.

The SAR measurements of wind direction and wind speed agree well with the R/V *Gauss* measurements. The wind speed from KNMI is 4.2 m s^{-1} larger than from SAR. The direction is turned 20° counterclockwise. The DWD and KNMI wind speeds were taken 1 hour later, differing from each other by about 6 m s^{-1} . The SAR measurement lies between these two values. From the weather charts we conclude a homogeneous wind situation for the area covered by the SAR image. This is supported by the time series measured from the R/V *Gauss*.

Figures 10 and 11 show scatterplots of ground truth measurements versus the values from the SAR for all scenes considered. Out of the 20 SAR images, 14 showed distinct wind streaks. In Figure 10 a comparison of wind direction from SAR images to KNMI ground truth measurements is given. The measurements compare well with a bias of 1.53° and a correlation of 0.99.

Table 2. Ground Truth and ERS SAR Data

	Wind Speed, m s^{-1}	Wind Direction, deg	Time, UTC
R/V <i>Gauss</i>	15.3	167	1100
KNMI	18.5	150	1200
DWD	12.3	156	1200
ERS SAR	14.3	170	1100

Wind speeds are u_{10} , and the time is on December 1, 1995. KNMI is the Royal Netherlands Meteorological Institute, and DWD is the German weather service.

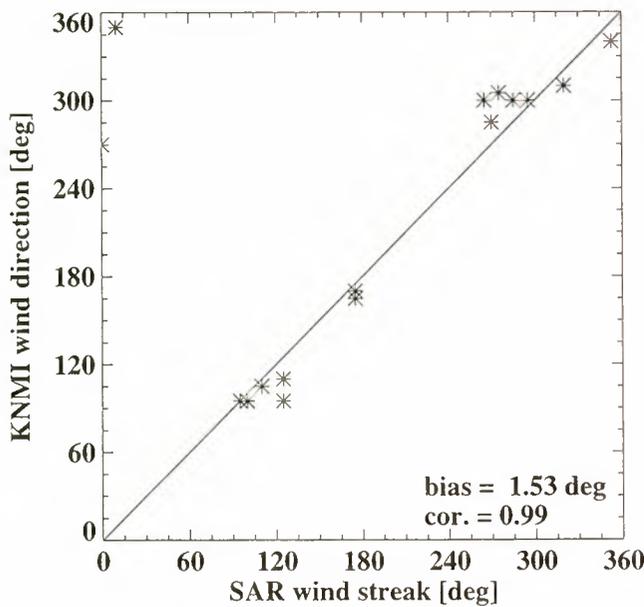


Figure 10. Scatterplot of the wind directions of the SAR versus the Royal Netherlands Meteorological Institute (KNMI) data set.

Comparisons to wind directions of DWD showed a much higher bias of 24° with wind directions of DWD being generally turned to the right. Figure 11 shows the scatterplot of wind speed with a bias of 0.6 m s⁻¹ and a correlation of 0.78.

From the close agreement of SAR data with the R/V *Gauss* and KNMI data we conclude that the SAR is capable to derive valid wind fields on a large scale. This is a strong motivation to investigate the SAR wind fields on a smaller scale.

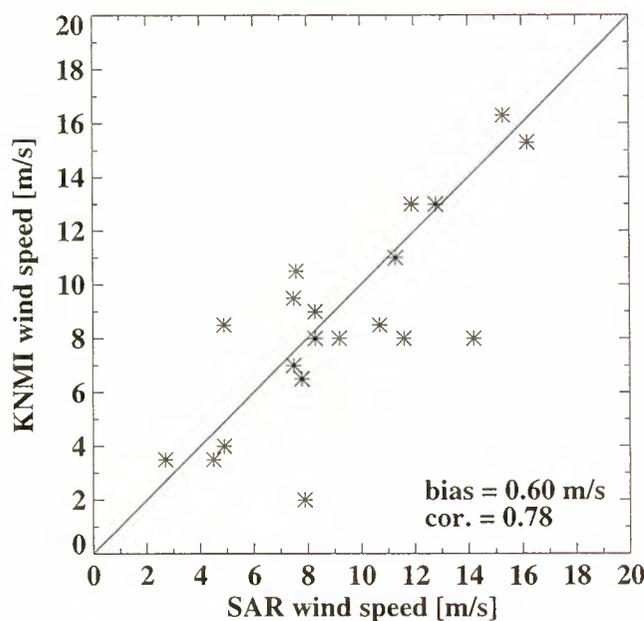


Figure 11. Scatterplot of the wind speeds of the SAR versus the KNMI data set.

5. Comparisons to the Mesoscale Wind Model GESIMA

High-resolution wave and current models require detailed wind information to calculate hydrographic parameters. It was important for us to test the capability of our algorithm to reproduce spatial variations of the wind field. Since no area-covering measuring methods exist, we used the mesoscale atmospheric model Geesthacht simulation model (GESIMA), developed at GKSS Research Centre to estimate the wind field around the island Rügen situated before the German coast in the Baltic Sea. It is a three-dimensional nonhydrostatic model of the atmospheric circulation with terrain-following coordinates. A detailed description of the model is given by Kapitza and Eppel [1992]. An overview of sample applications is given by Eppel et al. [1995]. The ERS-1 SAR image from August 12, 1991, at 2107 UTC was used to compute the wind field with a constant wind direction of 290°. For our comparisons, GESIMA was run for neutral atmosphere with a horizontal grid cell size of 1 km and variable grid size in the vertical. The lower boundary condition for the friction velocity over the sea is given by Charnock's relation. Over land the variable bottom stress is taken into account from land use charts. For the upper boundary condition at 1500 m, height stationary wind speed of 16 m s⁻¹ coming from 306° was assumed.

Figure 12 shows isolines of wind speed as computed by GESIMA, and Figure 13 shows isolines as computed by CMOD4. Wind direction and relative changes in the wind speed due to shadowing effects, particularly behind the white cliffs of Rügen, are clearly visible.

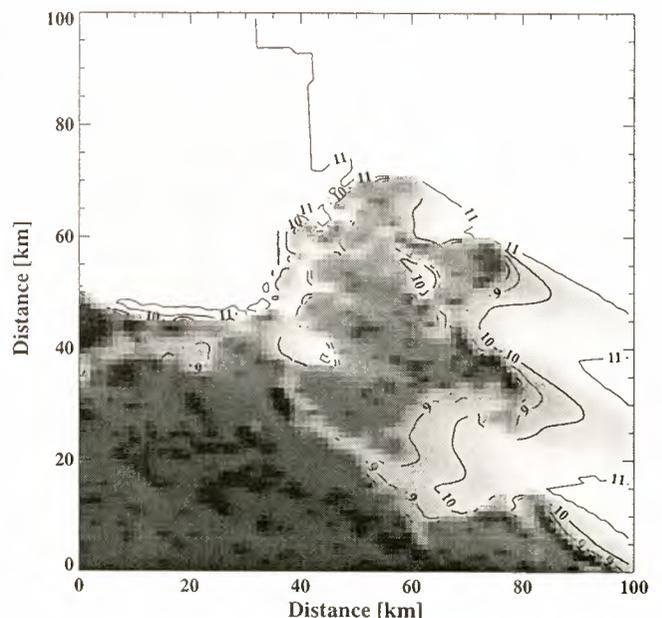


Figure 12. Wind speeds computed by the Geesthacht simulation model of the atmosphere (GESIMA). The solid lines represent isotachs in m s⁻¹. The image center is at 54.5°N and 13.25°E.

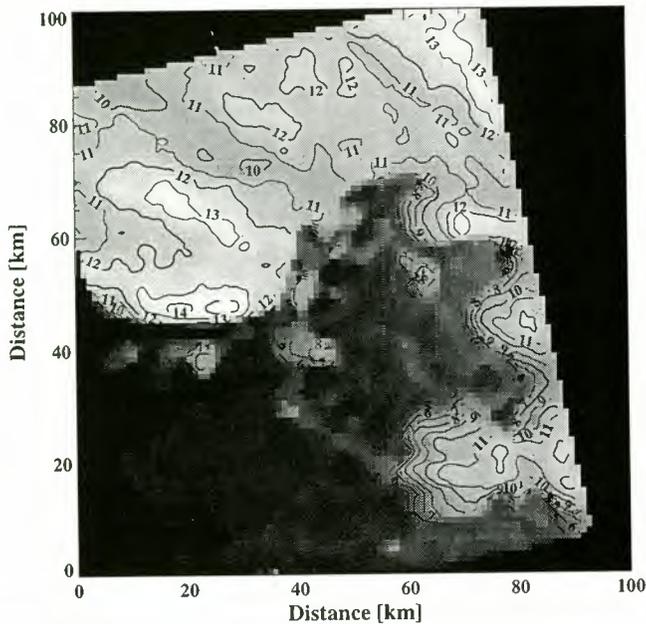


Figure 13. Wind speeds computed by the CMOD4 out of the ERS-1 SAR data. The solid lines represent isotachs in m s^{-1} . The image center is at 54.5°N and 13.25°E .

The SAR image shows much finer detail in wind structure and a higher variability in wind direction. This is due to the difference between the snapshot of ERS-1 SAR of a highly turbulent wind field and the mesoscale model simulation assuming a stationary situation. On the ocean surface north of Rügen the SAR image shows wind rolls from which the wind direction is derived. These features cannot be reproduced by the GESIMA model. So the SAR data give a lot of additional information on the turbulent structure of the atmosphere. East of the island of Rügen, the wind shadowing shows the same order of magnitude. The wind speed drops down to about 6 m s^{-1} behind the island before picking up again toward the open ocean. The geometrical location of the wind shadowing in the SAR and GESIMA data is slightly different, showing again that the results of GESIMA yield a much smoother solution than the distribution of wind speed in real nature.

Since the model has been run for stationary conditions, a full coincidence with the SAR wind field cannot be expected. We consider the comparison as a proof that the developed SAR algorithm for the wind field extraction can indeed resolve spatial inhomogeneities in the wind field, which in this case come from the variable topography and the variable bottom stress for the atmosphere.

6. Error Analysis of CMOD4-Derived Wind Speed

In the following we discuss three major sources of error when deriving wind speed from SAR images by CMOD4: the effect of the grid cell size, the effect of the

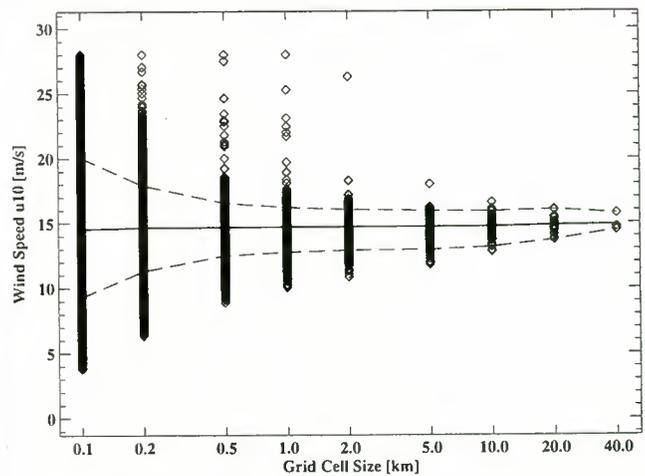


Figure 14. Wind speeds computed with a constant wind direction of 270° from the ERS-1 SAR image on December 1, 1995, at 1101 UTC. The solid line indicates the mean; 90% of the values lie between the dashed lines.

uncertainty in wind direction, and the computation of power loss. Figure 14 shows the wind speed values computed with the CMOD4 for different grid cell sizes of the ERS-1 SAR image from December 1, 1995, at 1101 UTC, shown in Figure 8. Four 40×40 -km areas were selected from the image, and the power-loss-corrected NRCS expressed in linear units was averaged over each area. As SAR incidence angle, the value in the middle of the grid cell was taken. Then the wind speed was computed with a fixed direction of 355° that was deduced from wind streaks. The 180° ambiguity was resolved by looking at wind shadowing behind the oil platforms. The procedure was repeated using smaller areas down to a size of 0.1×0.1 km. In Figure 14 the resulting wind speed values are plotted versus the corresponding grid cell size. The mean values for each area, shown as a solid line, are nearly constant at 15 m s^{-1} . The 90% confidence intervals, shown as dashed lines, are narrow and between 1- and 10-km grid cell size only, slightly de-

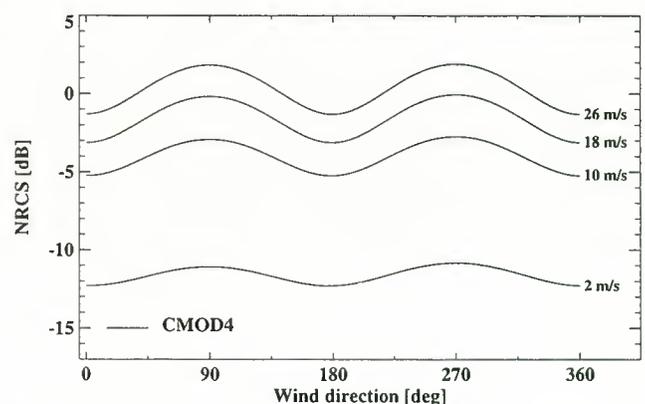


Figure 15. The dependency of NRCS on wind direction for different wind speeds. The wind direction is given clockwise to flight direction.

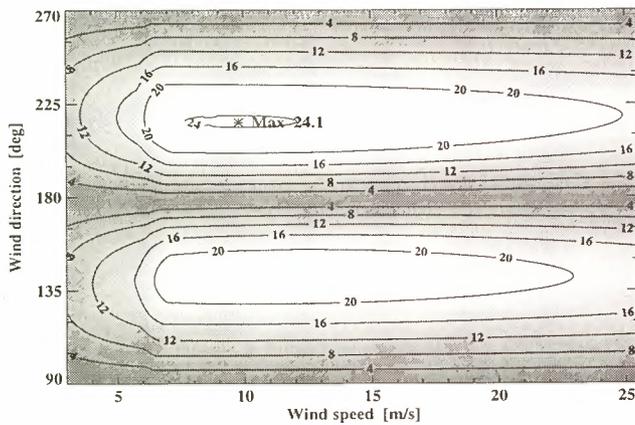


Figure 16. Estimated relative error in percentage of wind speed due to an uncertainty in wind direction of $\pm 10^\circ$ for an incidence angle of 23° . The wind direction is given clockwise from flight direction; for example, 90° is corresponding to downwind.

pending on area size. It is obvious that at grid cell sizes of < 0.5 km the estimated wind speeds are influenced by speckle effects [Bally and Fellah, 1995] and error in the mean wind direction. For small grid cell sizes up to 5 km, point targets like ships or oil platforms influence strongly the mean NRCS resulting in too large wind speeds.

The next source of error when deriving wind speed by ERS SAR is the necessity to have information on the wind direction as input into CMOD4. Usually, at wind speeds above 8 m s^{-1} the SAR images show patterns from which the wind direction can be derived. This direction compares very well to ground truth; see Table 2. However, Figure 15 shows the dependency of backscatter of the ocean surface at different wind speeds versus wind direction (clockwise to flight direction) at a radar incidence angle of 23° . The backscatter is obviously highest when the wind is blowing downwind and upwind at 90° and 270° , respectively, and lowest when blowing sideways to the antenna.

From CMOD4 the relative error in wind speed due to a 10° error in wind direction is derived. Figure 16 shows a contour plot of this error for wind direction versus wind speed. The error is minimal when the wind is blowing sideways to the antenna. The maximum error of 24.1% occurs at a wind speed of about 10 m s^{-1} , and the wind direction is 220° relative to flight direction.

Another source of error is the approximation used when recalibrating the ERS SAR images. To avoid this, the SAR preprocessing data (RAW) would have to be recalibrated and processed at high costs. Using a large sliding window to derive the power loss over the already processed PRI image is only an approximation. Generally, the algorithm works well for large, fairly homogeneous surfaces like the open ocean, but it fails at sharp boundaries such as coast lines in near range at high wind conditions. In this case the land is relatively

dark in comparison to the water surface, and the sharp boundary leads after recalibration to a reduced NRCS on the water surface because of the fact that the sliding window used for estimating the power loss is partly over land of low backscatter. Especially for ERS-1, this leads to unrealistically low values of wind speeds near the coast.

7. Conclusions and Outlook

It is shown that application of the ESA CMOD4 algorithm to ERS-1/2 SAR images can give valuable information on mesoscale wind fields over the ocean. If wind streaks show up in the SAR images, wind direction can be measured to an accuracy of 5° . Comparisons to available ground truth wind speed measurements show that the wind extraction algorithm is reliable for the prediction of coastal wind fields for distances > 3 km from shore. Recalibration of the SAR images to calibrated backscatter values remains a problem, especially for ERS-1 images near coastlines. In principal, the use of ERS-2 images for wind measurements is recommended. The method will be used in the future to calibrate mesoscale atmospheric models like the described GESIMA model. The direct use of the measured wind data, without employing a model calibration as an intermediate step, will only be possible if SAR imagery for individual regions becomes available more frequently. Therefore the increased availability of SAR images of coastal areas is strongly needed for the worldwide demanded, improved coastal zone management.

Acknowledgments. We are grateful to the ESA, who supplied us with the ERS-1 and ERS-2 SAR data under the AO project AO2.D113, and to the German weather service DWD for delivering the wind fields. This work was sponsored by the German Ministry of Research and Technology BMBF under contract 03F0165C.

References

- Alpers, W.R., and B. Brümmer, Atmospheric boundary layer rolls observed by the synthetic aperture radar aboard the ERS-1 satellite, *J. Geophys. Res.*, **99**, 12,613–12,621, 1994.
- Bally, P., and K. Fellah, Evaluation of the accuracy of the backscattering coefficient measurement in SAR data products, technical note, Eur. Space Res. and Technol. Cent., Noordwijk, Netherlands, July 1995.
- Chapron, B., T. Elfouhaily, and V. Kerbaol, A SAR speckle wind algorithm, in *Proceedings of the Second ERS-1 Workshop*, pp. 5–40, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, 1994.
- Eppel, D.P., H. Kapitza, M. Claussen, D. Jacob, W. Koch, L. Levkov, H.-T. Mengelkamp, and N. Werrmann, The non-hydrostatic mesoscale model GESIMA, II, Parameterizations and applications, *Contrib. Atmos. Phys.*, **68**(1), 15–41, February 1995.
- Horstmann, J., and T. Wolf, Comparison of ERS-SAR with numerical wind wave and current model results in a coastal area, in *Proceedings of CEOS Wind and Wave Validation*

- Workshop*, Eur. Space Agency Publ. Div., Noordwijk, Netherlands, in press, 1997.
- Johannessen, O.M., et al., COAST WATCH-95: ERS-1/2 SAR applications of mesoscale upper ocean and atmospheric boundary layer processes off the coast of Norway, in *IGARSS '96 / IEEE Proceedings*, vol. 2, pp. 1158-1161, Inst. of Elec. and Electron. Eng., New York, 1996.
- Kapitza, H. and D.P. Eppel, The non-hydrostatic mesoscale model GESIMA., I, Dynamical equations and tests, *Contrib. Atmos. Phys.*, 65(2), 129-146, May 1992.
- Koek, F.B. (Ed.), Time series, *Mon. Bull. N. Sea*, 12, 13-26, Aug. 1995a.
- Koek, F.B. (Ed.), Time series, *Mon. Bull. N. Sea*, 12, 13-26, Sept. 1995b.
- Koek, F.B. (Ed.), Time series, *Mon. Bull. N. Sea*, 12, 13-26, Dec. 1995c.
- Laur, H., Derivation of backscattering coefficient σ_0 in ERS-1 SAR PRI products, technical note, issue 1, rev. 0, 17 pp., Eur. Space Agency, Frascati, Italy, October 1992.
- Laur, H., P. Bally, P. Meadows, J. Sanchez, B. Schaettler, and E. Lopinto, Derivation of the backscattering coefficient σ_0 in ESA ERS-1/2 SAR PRI data products, *Tech. Note ES-TN-RS-PM-HL09*, issue 2, rev. 2, 41 pp., Eur. Space Agency, Frascati, Italy, June 1996.
- Meadows, P.J., and P.A. Wright, ERS-1 SAR analogue to digital convertor saturation, in *Proceedings of the CEOS SAR Calibration Workshop*, edited by K. Sarabandi, pp. 24-37, Univ. of Mich. Press, Ann Arbor, 1994.
- Rosenthal, W., S. Lehner, J. Horstmann, and W. Koch, Wind measurements using ERS-1 SAR, in *Proceedings of the Second ERS Applications Workshop*, Eur. Space Agency Spec. Publ., ESA SP-383, 355-358, December 1995.
- Scoon, A., I.S. Robinson, and P.J. Meadows, Demonstration of an improved calibration scheme for ERS-1 SAR imagery using a scatterometer wind model, *Int. J. Remote Sens.*, 17(2), 413-418, 1996.
- Stoffelen, A., and D. Anderson, Characterisation of ERS-1 scatterometer measurements and wind retrieval, in *Proceedings of the Second ERS-1 Symposium - Space at the Service of Our Environment*, Eur. Space Agency Spec. Publ., ESA SP-361, 997-1001, October 1993.
- Stoffelen, A., and D. Anderson, Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4, *J. Geophys. Res.*, 102, 5767-5780, 1997.
- Valenzuela, G.R., Theories for interaction of electromagnetic and oceanic waves - A review, *Boundary Layer Meteorol.*, 13, 61-85, 1978.
- Wolf, T., and G. Gayer, Coupled modelling of hydrodynamical parameters in the Greifswalder Bodden, in *Oceans '95 MTS / IEEE Proceedings*, vol. 2, pp. 769-774, Inst. of Elec. and Electron. Eng., New York, 1995.
- Wright, J.W., Backscattering from capillary waves with application to sea clutter, *IEEE Trans. Antennas Propag.*, AP-14(6), 749-754, November 1966.

J. Horstmann, W. Koch, and W. Rosenthal, GKSS Research Centre, 21502 Geesthacht, Germany. (e-mail: Horstmann@gkss.de; frkochw@gkss.de; Rosenthal@gkss.de)
 S. Lehner, DLR/DFD, 82230 Weßling, Germany. (e-mail: Susanne.Lehner@dlr.de)

(Received December 23, 1996; revised August 21, 1997; accepted September 26, 1997.)

Coastal wind field retrievals from ERS synthetic aperture radar images

Erik Korsbakken

Nansen Environmental and Remote Sensing Center, Solheimsviken, Bergen Norway

J. A. Johannessen

European Space Agency, European Space Research and Technology Centre, Noordwijk, The Netherlands

O. M. Johannessen

Nansen Environmental and Remote Sensing Center /Geophysical Institute, University of Bergen, Bergen, Norway

Abstract. A unique series of European Remote Sensing Satellite (ERS) 1 and 2 C band synthetic aperture radar (SAR) images was obtained off the southern coast of Norway during the Coast Watch'95 experiment in September 1995. In this paper we carry out a systematic analysis of the mesoscale coastal wind field conditions expressed in the SAR images. Four different categories of phenomena including windrows, fetch-limited seas, wind fronts, and oceanographic fronts are examined and discussed. The quantitative retrievals of the wind field are based on examination of both the SAR image backscatter characteristics and the spectral properties. Results are compared and validated against coincident ship and buoy data, providing complementary and independent observations of the oceanographic and meteorological conditions.

1. Introduction

Wind scatterometer observations over the ocean provide direct estimates of the global wind vector field at a spatial resolution of 50 km with an accuracy of ± 2 m/s in speed, $\pm 20^\circ$ in direction and a directional ambiguity of 180° [Stoffelen and Anderson, 1993]. However, for some applications, such as in semi-enclosed seas, straits, coastal regions, estuaries, sea ice polynyas, and along the marginal ice zones, this resolution is too coarse. For monitoring and forecasting in these regions, wind field estimates retrieved from high-resolution synthetic aperture radar (SAR) images can therefore be very useful.

Quantitative estimates of the wind field from SAR images have not been systematically examined until recently. Chapron *et al.* [1995] have proposed a method for extracting the wind speed from the spectral properties of a SAR image. Vachon and Dobson [1996] have systematically applied the C band model (CMOD4) [Stoffelen and Anderson 1993] to ERS SAR data and compared with in situ measurements to provide qualitative performance assessments. Vachon and Dobson [1996] obtained an accuracy better than 1.5 m/s (compared to in situ measurements) if the wind direction was known and corrections for the ERS-1 SAR analog to digital conversion (ADC) saturation were applied. Scoon *et al.* [1996] have compared wind speed obtained using CMOD4 on SAR images from the English Channel to those from synoptic weather maps and conclude that a proper ADC correction is necessary to obtain reliable wind speed estimates.

In order to better quantify SAR imaging of upper ocean and atmospheric boundary layer processes, an experiment, Coast

Watch'95 [Johannessen *et al.*, 1996], was carried out off the southwest coast of Norway during September 1995 as a ERS-1 and ERS-2 tandem announcement of opportunity project. The tandem operation provided a unique SAR coverage of the experiment area consisting of 148 SAR images (frames). Along-track scanning radiometer (ATSR) and National Oceanic and Atmospheric Administration (NOAA) advanced very high resolution radiometer (AVHRR) data were also obtained during the experiment. In situ meteorological and oceanographic observations were provided from the research vessel (R/V) *Håkon Mosby* of the University of Bergen, and an advanced metocean buoy from the Naval Postgraduate School (NPS), Monterey. Digital transmission of near real-time analysis of the SAR data to the vessel was continuously used in the planning and execution of the field study.

The main study objective reported in this paper is to optimize the wind retrieval from SAR by combinations of the two methods suggested above. We emphasize that our intention is not to improve the methods themselves, but to examine their responses under different conditions and taking advantage of combining the methods to derive the wind direction. We also want to compare the different algorithms against Coast Watch'95 in situ measurements and a meteorological hindcast model.

Because of the spatial resolution of the hindcast data as provided from the Norwegian Meteorological Institute, a direct validation based on these data becomes difficult, but the data are still used as an indication of the wind direction. The model is run on a 75 km resolution and based on a large number of in situ observations. Results are stored for off-line use with an interval of 6 hours.

In our analysis we use the procedure outlined by Korsbakken [1996] and Korsbakken and Johannessen [1996], and section 2 contains a brief review of the two retrieval methods. Application of the methods to the Coast

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02580.
0148-0227/98/97JC-02580\$09.00

Watch'95 data is reported in section 3 including a systematic comparison of the separate wind speed estimates from the spectral and radar backscatter properties (CMOD4) of the SAR images as well as a combination of the two methods. Three different categories of wind regimes are used, and in section 4 the results are discussed in accordance with the characteristics of these different regimes.

In particular, we present a comparison of the wind speed results, in areas where we do not have supporting ground truth, motivated by the following.

1. We believe that there are basically two different mechanisms responsible for the wind speed response in the two methods, namely, the short centimeter-scale waves for the CMOD4 and the longer wind-generated waves for the azimuth cutoff method. This makes it interesting to compare the two different algorithms under different conditions.

2. Consistency in the derived wind speed results is important for the method to obtain the wind direction.

3. Comparisons to local in situ measurements are not satisfactory to validate these algorithms to obtain a high spatial resolution wind field for a larger area (of course, limited by the SAR coverage).

4. Different responses in the methods may provide new oceanic information.

5. We want to demonstrate local variability, especially in coastal regions.

Finally, in section 5 the main results including the feasibility of the methods are summarized.

2. Wind Retrieval Algorithms

The two SAR wind retrieval algorithms are based on the extraction of the wind field from different parts of the ocean surface wave spectrum, in particular the medium wind-wave regime and the small centimeter-scale regime. Figure 1 shows a conceptual overview of the wind field estimation described below.

2.1. SAR Wind Algorithm

The SAR wind algorithm (SWA; the notation is used for the first time in this paper), proposed by *Vachon et al.* [1994] and further examined by *Chapron et al.* [1995], is based on a relation between the smearing effects [*Hasselmann and Shemdin* 1982] in the SAR image and the wind field. Smearing effects tend to increase the coherence (correlation) length of the radar returns in the spatial image domain and to influence the spectral properties of the SAR image.

The SWA is derived from a comparison of ERS SAR wave-mode data and the ERS wind scatterometer data which can be simultaneously acquired with the same satellite. In the case of a fully developed sea (no fetch limitation), the empirical SWA relation, based on an evaluation of 1200 SAR wave-mode images with a central incidence angle of 20.2°, is given by *Chapron et al.* [1995] as

$$U_{10} = 4.75 \left(\frac{\lambda_c - 30}{110} \right) \text{ m s}^{-1} \quad (1)$$

where U_{10} is the wind speed at 10 m above the surface and λ_c is the azimuth cutoff wavelength.

Expression (1) is also consistent with the expression $\lambda_c = 23.4U_{10} + \Lambda$ as given by *Kerboal et al.* [this issue], who

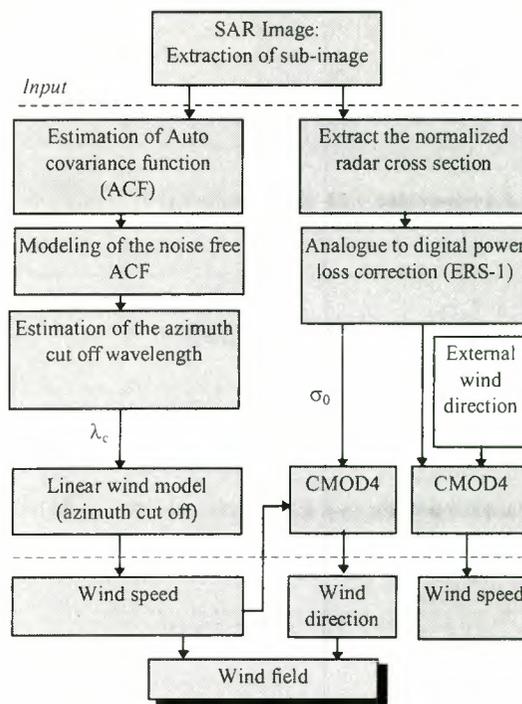


Figure 1. Conceptual overview of wind retrieval models. The external wind direction may be obtained from windrows in the SAR image as proposed in the text.

discuss the method and its capabilities in greater detail. In this expression Λ represents a residual cutoff involving the nominal SAR azimuth resolution. In case of the SAR precision image (SAR.PRI) as used in this study, the nominal azimuth resolution is about 30 m ($\Lambda = 30$ m).

Assuming a Gaussian shaped low-pass filter for the azimuth cutoff, the cutoff wavelength can be estimated from the autocovariance function (ACF) derived from an inverse Fourier transform of the SAR image power spectrum (the Wiener Khinchin theorem). The effect of strong azimuth filtering effects due to the smearing and, finally, the azimuth cutoff are clearly seen in the presentation of the case studies in this paper. As proposed by *Chapron et al.* [1995], the ACF can be regarded as a sum of the narrow peak due to the system response of the broadband noise and a broadened "shoulder" corresponding to the Gaussian-shaped filter (in the spectral domain) defined as $c(x) = \exp(-\pi x / \lambda_c)$, where x is the lag in the ACF. A Gaussian function $f(x) = \exp(-ax^2)$ is then fitted to the ACF in the azimuth direction over the broadened "shoulder" part of the function to determine a . In return, λ_c is then obtained from the relation $f(x) = c(x)$.

The spatial resolution of the SWA method is limited by the needed full-resolution subimage to obtain statistical confidence in the estimation of the SAR image power spectrum.

2.2. The CMOD4 Model

The CMOD4 wind retrieval model [*Stoffelen and Anderson* 1993] was developed for the ERS-1 C band scatterometer, but it has also been shown to give good estimates of wind speed when applied to ERS-1 SAR images [e.g. *Johannessen et al.*, 1995; *Vachon et al.*, 1995; *Vachon and Dobson* 1996; *Wackerman et al.*, 1996]. The CMOD4 model provides σ_0 values as a function of relative wind

direction ($\phi=0$ for a wind blowing toward the radar), wind speed, and incidence angle, expressed as

$$\sigma_0 = B_0[1 + B_1 \cos(\phi) + B_2 \cos(2\phi)] \quad (2)$$

The coefficients B_0 , B_1 , and B_2 depend on the local incidence angle of the radar beam and wind speed. The accuracy in the model is $\pm 20^\circ$ in relative wind direction and ± 2 m/s in wind speed when applied to scatterometer data.

The CMOD4 model is derived for a neutral stratification. In order to compute the wind speed (U_{10}) from the radar backscatter accounting for the stratification (ΔT) in the atmospheric boundary layer (ABL) the CMOD4-derived wind speed must be modified following the method suggested by Wu [1993] or Smith [1988].

The saturation effect of the ADC in the satellite is described by Meadows and Willis [1995], Laur *et al.* [1993], and Scoon *et al.* [1996]. The effect is strongest in the near range and increases with radar backscatter intensity (e.g. at medium to high winds over the ocean surface) and leads to an underestimation of σ_0 , which has to be compensated in order to properly estimate the absolute value of the radar backscatter. In turn, underestimation of σ_0 leads to an underestimation of the wind speed.

In this work the absolute calibrated σ_0 from ERS SAR is derived in accordance with a comprehensive calibration scheme provided by European Space Agency (ESA) [Laur *et al.*, 1996] except for correction for variance in the replica pulse power. The latter is shown to be negligible using the CMOD4 algorithm for wind speed retrieval in SAR data [Scoon *et al.*, 1996].

In this paper we also emphasize that the wind direction can be estimated from the CMOD4 model for different incidence angles, provided the wind speed, derived from the SWA method, can be associated with the corresponding measured radar backscatter (σ_0). In such cases we will show that four solutions, i.e., two pairs, each with a 180° ambiguity can be found, except in the cases when the direction is close to upwind (the wind blowing toward the radar) or downwind, for which only one pair is found. (Note that for the three beam scatterometers on ERS-1 and ERS-2 the number of solutions is reduced to a single pair with a 180° ambiguity.)

It has also been demonstrated [Johannessen *et al.*, 1995] that windrows manifested in the SAR images can be used to indicate the near-surface wind direction during the SAR integration time. In such cases, the number of wind direction solution pairs is also reduced to one (180° ambiguity). In some cases, e.g., off-ice and offshore winds the ambiguity can be completely resolved. Estimating the wind direction has also been proposed derived from the long-wave part (1000 m or more) of the SAR image power spectrum [Vachon and Dobson 1996; Wackerman *et al.*, 1996]. This requires that the spectra are estimated from subimages large enough to resolve the low-frequency part of the spectra and that the wind streak pattern appear as a stationary wave field within the subimage.

3. Coast Watch'95 Analysis

The analyzed SAR data is three-looks ground range data in ERS PRI format, which is corrected for antenna elevation gain and range spreading loss. The ground resolution in range

(normal to the satellite track) direction and azimuth is about 30 m, and the pixel size is 12.5 m in each direction. All satellite passes are descending tracks, and here the northernmost SAR image is always referred to as the first. Each SAR image (100×100 km) has been regularly divided into 9×9 subimages covering about 10×10 km at the ground.

The SWA wind speed and SAR image power spectrum (SIPS) are estimated from each subimage following the method suggested by Monaldo [1991]. For the estimation of the CMOD4 wind speed and direction, each subimage is averaged to a pixel size of 100×100 m before the ERS-1 data are corrected for the ADC power loss according to Laur *et al.* [1996]. The CMOD4 wind speed is then calculated and compared to the SWA-derived wind speed assuming the in situ wind direction from the R/V *Håkon Mosby* to be valid in all subimages. The wind direction is also independently estimated from the CMOD4 model function by a combination of the SWA wind speed and the calibrated radar backscatter. The ambiguity in wind direction is then resolved by comparing the four results with the in situ measured wind direction and also the windrows seen in the SAR image.

Three case studies from the Coast Watch'95 database are presented, including case 1, windrows under fetch-limited conditions; case 2, wind front; and case 3, oceanic front (jet) together with a local wind front.

A composite overview of these three cases including the SAR image expressions together with the retrieved wind vector maps is shown in Figures 2-5. They can be briefly characterized as follows.

Case 1 includes the ERS-1 SAR images from September 16, 1995 (Figure 2, left), which show a very characteristic pattern of windrows aligned in the wind direction of about 120° . There is also a gradual increase in the backscatter toward the south. The corresponding wind vector map is shown in Figure 2 (right). The wind speed in the vicinity of the ship is about 13 m/s and reveals a clear northwest gradient, as expected.

The SAR images from September 17, 1995 (Figure 3, left), also show a pattern of windrows with an orientation of about 110° as well as higher backscatter in the southern parts. The corresponding wind vector map (Figure 3, right) confirms this. Overall, the wind speed has dropped by about 4-5 m/s in comparison to the previous image.

Case 2 includes the SAR images from September 23, 1995 (Figure 4, left), which show a characteristic local wind front appearing as a bright-dark boundary. Windrows are also present. The corresponding wind vector map shows a northeasterly wind speed gradient and a wind direction of 260 - 300° .

Case 3 the final case study, is based on the SAR images from September 27, 1995 (Figure 5, left), which shows a westward flowing coastal jet bounded by two distinct, and in some places parallel fronts. A local, 5-10 km wide wind feature is also running diagonally, in the second image. The corresponding wind vector map (Figure 5, right) shows a wind speed of about 10 m/s from the southwest at about 215° .

A more detailed analysis of the three cases is presented systematically below. Since many of the SAR images express distinct areas associated with the different surface conditions, they are divided into subareas to ensure near-homogeneous backscatter within the analyzed subimages. A selection of the SIPS as obtained within the different regions of the image is

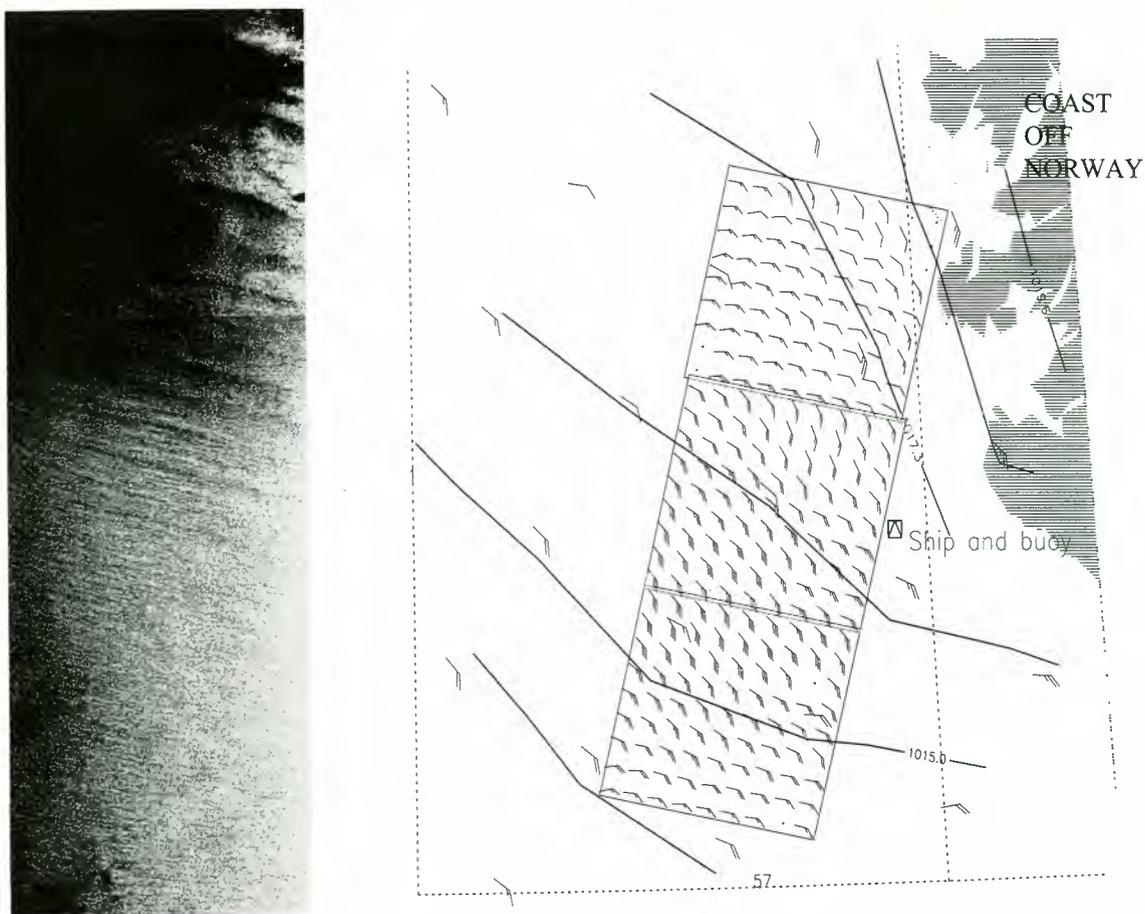


Figure 2. Case 1: SAR images from September 16 (left) and corresponding wind vector map derived from inverting the CMOD4 model function using the SWA wind speed and the calibrated backscatter (right). Isobars and corresponding surface wind vectors from diagnostics are superimposed. The position of the R/V *Håkon Mosby* and the “metocean” buoy at the satellite overpass is indicated. The observed discontinuity in wind direction between the northernmost and the center image is due to the processing of the SAR images as discussed in the text.

also plotted to demonstrate the effects of variability in the wind speed from one subarea to another for case 2 and 3. The SIPS do not provide direct quantitative information about the real sea state, but they are still valid in assessing the SWA performance, since the wind information is directly derived from the azimuth response in the spectra. Moreover, if we assume a linear ocean to SAR mapping of the ocean waves, the SIPS will still contain valuable information of the main wave modes and their propagation direction.

3.1. Case 1: Windrows (Fetch)

Four distinct subareas are identified in the composite of the three SAR images from September 16 as illustrated in Figure 6 (left). They include nearly homogeneous backscatter with no structure (area 1), low frequent streaks (area 2), homogeneous (periodic) wind streaks (area 3), and windrows with dark patches in between (area 4). Superimposed on the figure is the ship track.

In comparing the SWA and CMOD4 derived wind speed, about 80% of a total of 238 subimages have a wind speed difference less than 2 m/s using the in situ measured wind direction to present all subimages in the CMOD4 model. In the scatter plots in Figure 7 the wind speed ranges from 3 to 15 m/s. The correlation is good for areas 1 and 2, and the

SWA wind speed retrievals tend to be higher than the CMOD4 wind speed for area 3, although some impression of saturation is seen in CMOD4 at winds around 9 m/s. For the low number of subimages in area 4, the correlation seems to be relatively good. Deviations in the obtained wind speeds will be further discussed in terms of surface conditions and spatial variability in section 4.

The second and third SAR image were originally processed using a nominal chirp function in the processing of the SAR image (information provided by ESA). Compared to standard processing, where the chirp function is extracted from the received raw SAR data, this leads to a reduction in the average σ_0 value. To compensate for this, ESA provided a correction constant of 4.1 dB to be added. This correction may not preserve the full dynamical range of backscatter and may cause the saturation-like effect in area 3. Similarly, a discontinuity in the derived wind direction is seen in Figure 2; on the other hand, this is not supposed to impact on the SWA wind speed. Nonetheless, the wind direction retrieved from the combination of SWA and CMOD4, as shown in Figure 2, is in good agreement with the observed windrows in the SAR images as well as the in situ measurements.

Table 1 summarizes the observations in areas 1 to 4, listing the main properties of the subareas and the corresponding

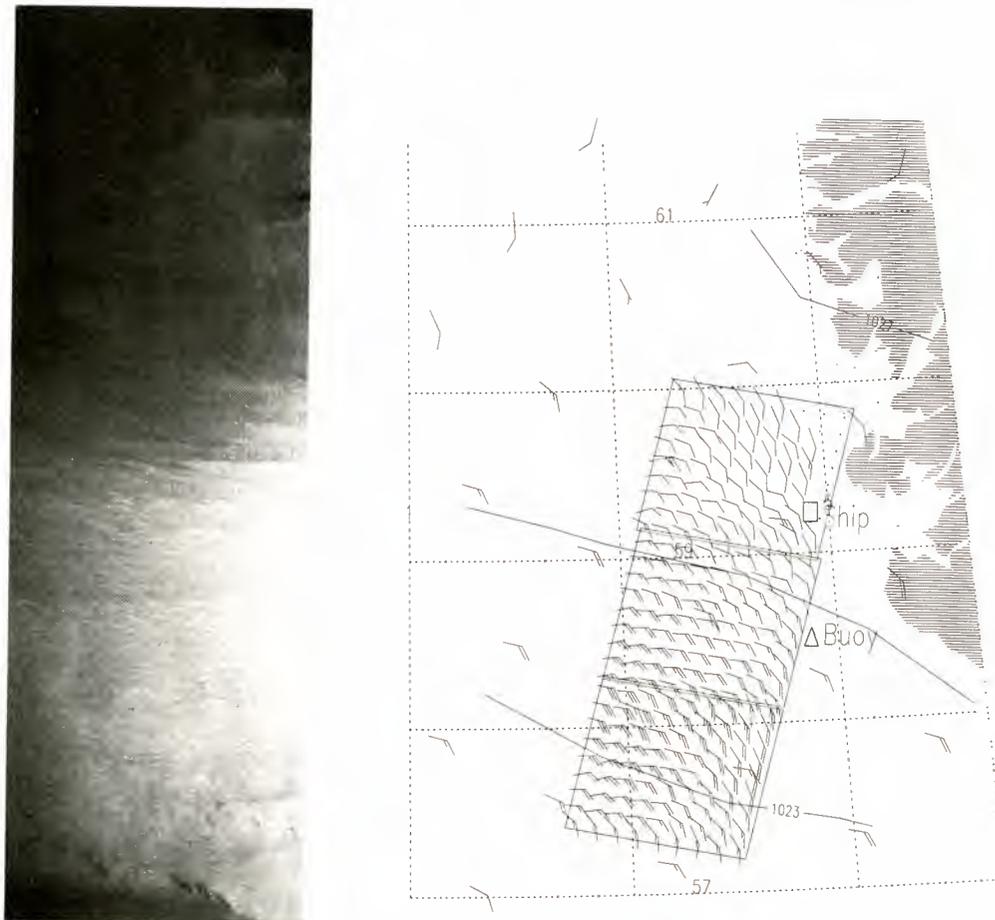


Figure 3. Case 1: SAR images from September 17 (left) and corresponding wind vector map derived from inverting the CMOD4 model function using the SWA wind speed and the calibrated backscatter (right). Isobars and corresponding surface wind vectors from diagnostics are superimposed. The positions of the R/V *Håkon Mosby* and the “metocean” buoy at the satellite overpass are indicated.

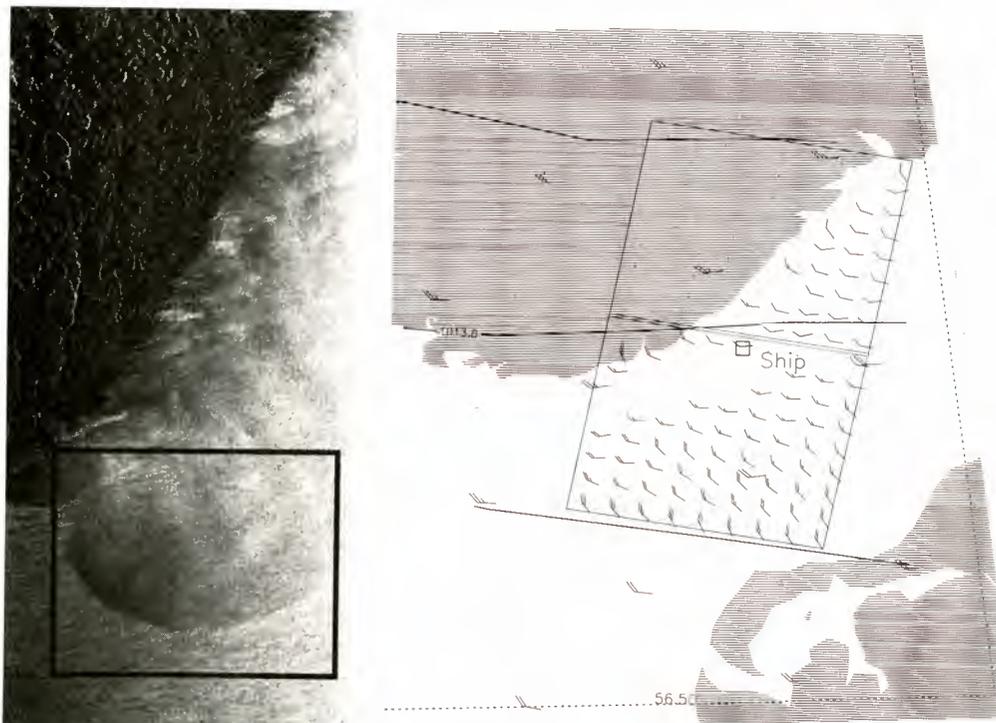


Figure 4. Case 2: SAR images from September 23 (left) and corresponding wind vector map derived from inverting the CMOD4 model function using the SWA wind speed and the calibrated backscatter (right). Isobars and corresponding surface wind vectors from diagnostics are superimposed. The positions of the R/V *Håkon Mosby* at the satellite overpass is indicated. Boxed area indicated in the SAR image is a subimage used in Figure 12.

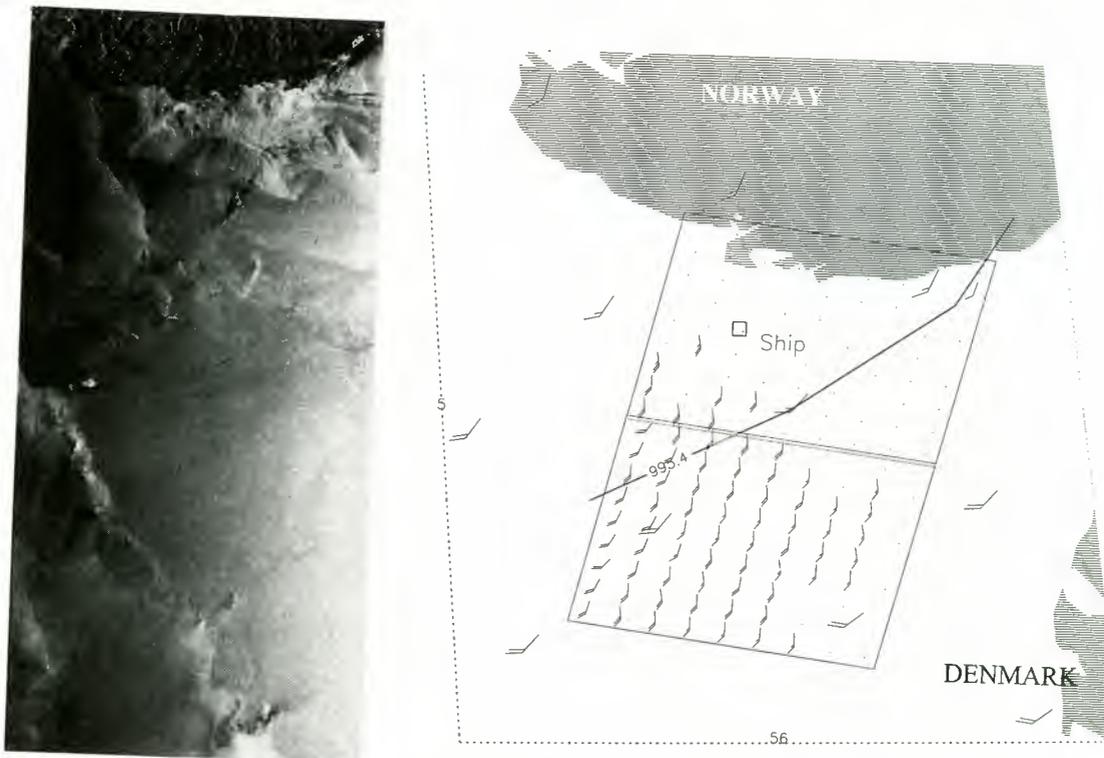


Figure 5. Case 3: SAR images from September 27 (left) and corresponding wind vector map derived from inverting the CMOD4 model function using the SWA wind speed and the calibrated backscatter (right). Isobars and corresponding surface wind vectors from diagnostics are superimposed. The position of R/V Håkon Mosby at satellite overpass is indicated.

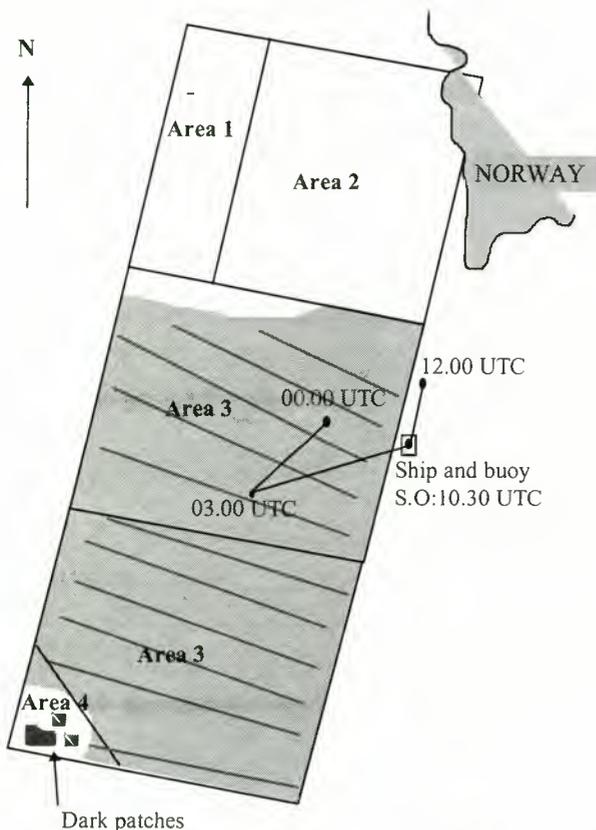


Figure 6. Identification of subareas of the SAR images on September 16, 1995 (left) together with the ship track prior to and during satellite overpass. Shaded area marks the presence of windrows and the lines roughly indicate the orientation but not the spacing of the windrows. The ship and buoy were in the same position at satellite overpass as marked by S.O.

SAR image power spectra as extracted from the subareas, and finally, the difference in the derived SWA and CMOD4 wind speeds and wind directions from the combination of SWA and CMOD4 provided collocated with wind direction from the hindcast model and in situ observation if available. Because of the spatial resolution of the hindcast model of 75 km the number of collocated data sets become limited.

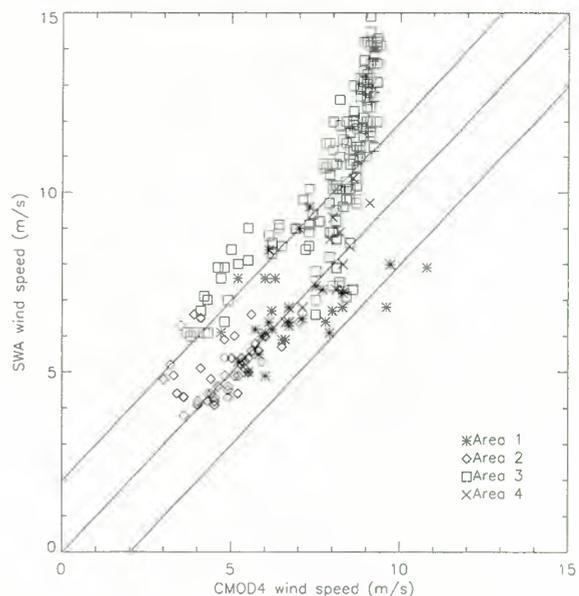


Figure 7. CMOD4 wind speed versus SWA wind speed derived from the SAR images on September 16, 1995. The interval of ± 2 m/s is indicated.

Table 1. Summary of Analyses of September 16 SAR Images

Area Location (SAR Images 1,2, and 3, From Top to Bottom in Figures 2 and 6)	Features Observed in the Different Areas	Waves Obtained From SIPS	SWA-CMOD4 Wind Speed Comparison (Magnitude of Difference in the Area)	Wind Direction		
				SWA Combined With CMOD4	Hindcast	In Situ
Area 1, westernmost part of the first image	Nearly homogeneous, no presence of streaks	Mode of about 300 m propagating 45° to range direction	CMOD4>SWA (7 - 3 m/s)	84°	114°	
Area 2, main part of the first image	Streaks which are more low frequent than the characteristic windrows	Mode about 500 m propagating about 20° to range direction	CMOD4~SWA (1 - 2 m/s)	105°	152°	
Area 3, the second and third image	Characteristic homogeneous windrows	A strong mode of about 100 m in near range direction and a long mode (>300 m) in near azimuth direction	CMOD4<SWA (0 - 2 m/s)	125°	124°	130°
				136°	114°	
Area 4, lower south-east parts of the third image	Characteristic windrows and some dark spots	A strong mode of about 100 m near 45° to range direction and a long mode (>300 m) in near azimuth direction	CMOD4>SWA (2 - 4 m/s)	96°	123°	

In order to validate the wind field estimates further, these results are compared to meteorological records (obtained by the R/V *Håkon Mosby* along the ship track) prior to and at the satellite overpass (see Table 5). The wind speed shows fluctuations between 10 and 13 m/s before the satellite overpass, while the wind direction was stable at 120° prior to the satellite overpass. Very good agreement between the wind direction along the ship track and the observed windrows in the SAR image is found, suggesting that the use of the in situ direction in all CMOD4 estimates is valid. The air-sea temperature difference shows only a weak unstable ($\Delta T = -1^\circ\text{C}$) atmospheric boundary layer stratification at the satellite overpass, allowing us to neglect this in the CMOD4-SWA comparison.

Another SAR image with clear expressions of windrows is obtained for exactly the same area 1 day later on September 17 (Figure 3). Three distinct areas are identified in this SAR image, as shown in Figure 8 (left). Again, the northernmost area, west of the coast, lacks expression features, while in the southern areas are windrows with a west-northwestward orientation of 110°.

Wind retrievals from a total of 238 subimages were compared for the September 17 images. The CMOD4 - SWA wind speed difference was less than 2 m/s for about 90% of the subimages. The scatter plot in Figure 9 reveals good agreement between the SWA and the CMOD4 derived wind speeds in the range of 3 to 15 m/s for all three images. In contrast to the previous results, no evidence of saturation is found for the high CMOD4 winds in this case. The outliers in area 1 (represented in Figure 9 as squares) arise from a very dark patch (low σ_0) close to the coastline where the SWA wind retrieval method fails because of fetch-limited seas with

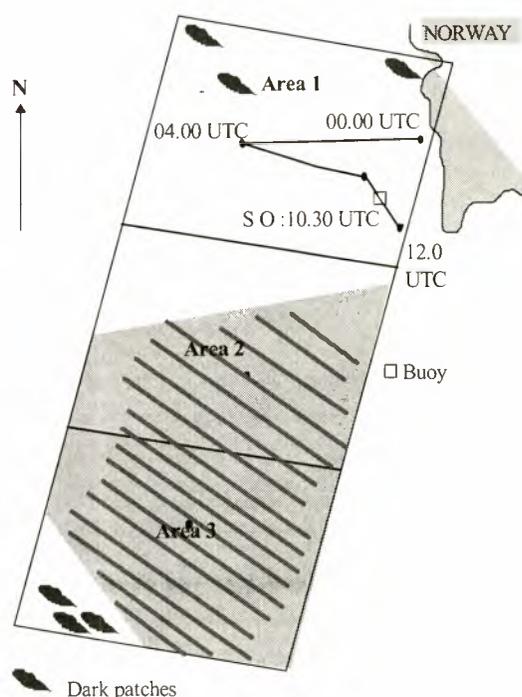


Figure 8. Identification of subareas of the SAR images on September 17, 1995 (left) together with the ship track prior to and during satellite overpass. Shaded marks presence of windrows and the lines roughly indicate the orientation but not the spacing of the windrows. The ship position at satellite overpass is marked by S.O.

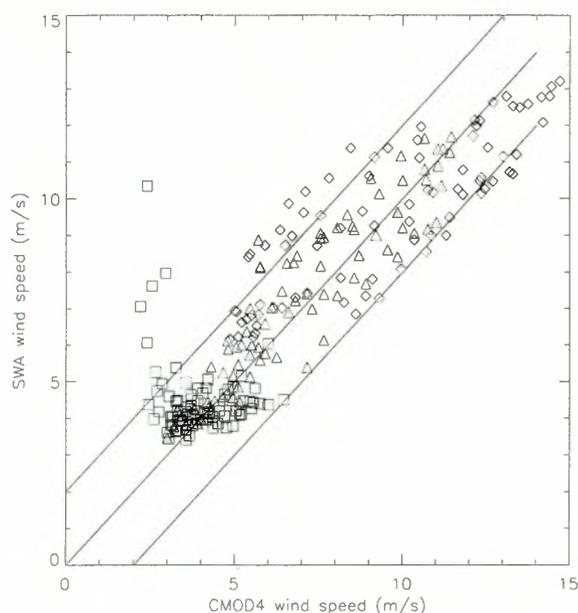


Figure 9. CMOD4 wind speed versus SWA wind speed, derived from the SAR images on September 17, 1995. The interval of ± 2 m/s is indicated. The subimages containing land are rejected. Squares are used for the northernmost SAR image, triangles for the center image, and diamonds for the southernmost SAR image.

subsequent lack of fully developed waves and corresponding wave modulation. In such regions, no particular azimuth cutoff is present, and the azimuthal resolution is only dependent on the response function of the system (speckle noise response), and the SWA method produces high wind speed values. Therefore no wind vectors could be produced in

this area (see Figure 3), since there were no intersecting points between observed σ_0 and modeled σ_0 based on the SWA wind speeds.

There is generally very good agreement between the wind vector map and the isobars superimposed on the map (Figure 3) both in regard to direction and to changes in magnitude from north to south.

Table 2 summarizes the observations for the three areas including a listing of the main properties of the subareas and the corresponding SAR image power spectra as extracted from the sub-areas and, finally, the response in the derived wind speeds and wind directions.

Although the range of winds in the September 16 to 17 SAR images is about the same, there is a gradual decrease in wind speed at the buoy and ship locations. This explains why the wave field is undergoing moderate changes.

The meteorological records from the R/V *Håkon Mosby* in the time period before, during, and after the satellite overpass, show that the wind speed is gradually decaying in the area before the satellite overpass reaching about 7 m/s as mentioned above, while the wind direction was relatively stable at 100° prior to and at satellite overpass. The air-sea temperature difference is weak and slowly changing to neutral. A southward increase in the wind speed of about 4 m/s over the distance from the ship (6-7 m/s) to the buoy (10-11 m/s) is also encountered, in agreement with the SAR-derived wind vector map (Figure 3).

3.2. Case 2: Wind Front

The SAR images from September 23 display the existence of a clear wind front (Figures 4 and 10). The images have moreover been arranged into three areas (Figure 10, center), in which area 1 is in the near-shore zone, area 2 is immediately north of the wind front, and area 3 is out of the front. Weak

Table 2. Summary of Analyses of September 17 SAR Images

Area Location (SAR Images 1, 2, and 3, From Top to Bottom in Figures 3 and 8)	Features Observed in the Different Areas	Waves Obtained From SIPS	SWA-CMOD4 Wind Speed Comparison (Magnitude of Difference in the Area)	Wind Direction		
				SWA Combined With CMOD4	Hindcast	In Situ
Area 1 (Image 1)	Dark patches	Weak range traveling mode at about 100 m	CMOD4 > SWA (0 to 1 m/s)	98°	124°	
				108°	147°	
Area 2 (Image 2)	Windrows	Strong short mode of about 100 m 20° to range and a near azimuth traveling mode at about 500 m	SWA > CMOD4 (0 to 1 m/s)	101°	124°	100°
Area 3 (Image 3)	Windrows and some dark patches in low left corner	Short mode of about 100 m 20° to range and a near azimuth traveling mode at about 500 m	SWA > CMOD4 (0 to 1 m/s)	100°	107°	
				87°	96°	
				140°	101°	

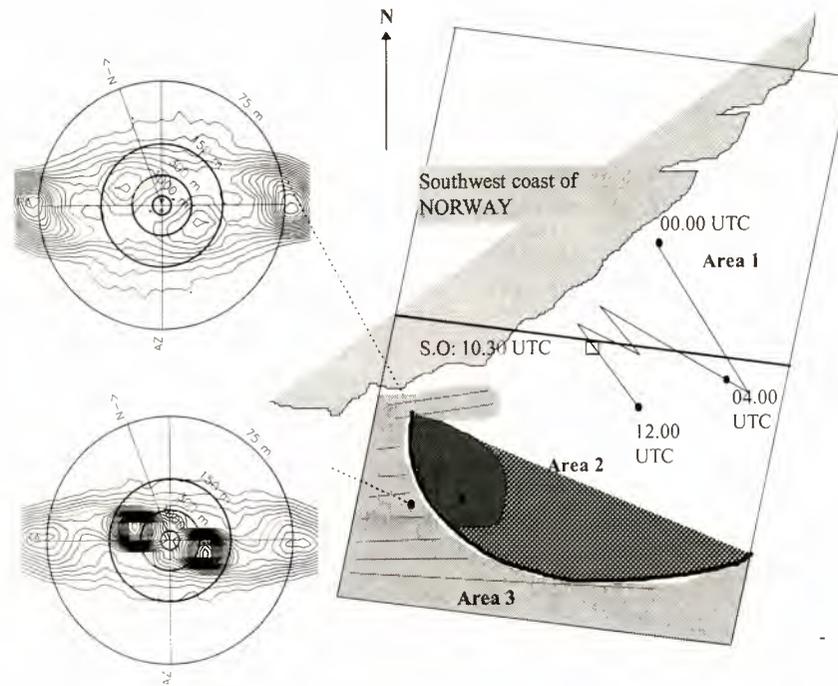


Figure 10. Identification of subareas of the SAR images on September 23, 1995 (center) together with the ship track prior to and during satellite overpass. Shaded area marks the presence of windrows, and the lines roughly indicate the orientation but not the spacing of the windrows. The ship position at satellite overpass is marked S.O. A selection of SAR image power spectra in the different areas (left). The concentric circles indicate wavelengths of 1000 m (inner ring), 300, 150, and 75 m, RA is range direction, AZ is azimuth direction, and N is north.

evidence of windrows is found in areas 2 and 3. Superimposed on the figure is the ship track, while the locations and extracts of the SAR image power spectra are shown to the left.

The wind front curves through the image with a width or transition zone of about 15 to 20 km. The largest gradient in σ_0 appears to be oriented approximately perpendicular to the wind direction (darkest area in the front as illustrated in Figure 10), while the σ_0 gradient becomes weaker as the frontal orientation becomes more closely aligned with the wind direction. A pattern of northeast oriented windrows is seen in the upper left part of the second image between the front and the coastline. Some windrows are also manifested to the south of the front. The in situ wind direction indicates that the fetch distance should be more than 200 km in the second image, and fetch-limited seas are expected in the northwest sector, in area 1, of the first image.

A total of 116 subimages were examined, of which 60% have a SWA-CMOD4 wind speed difference of less than 2 m/s in the range from 3 to 15 m/s as shown in Figure 11. The SWA wind speed tends to be somewhat less than the CMOD4 wind speed in area 1. In area 2 the majority of the estimates are still within ± 2 m/s, but we also observe that some of the CMOD4 wind speed estimates exceed the SWA wind by more than 2 m/s. However, in area 3 the agreement is generally poor, with the SWA wind speed generally larger (more than 2 m/s) than the CMOD4 wind speed. Possible reasons for the lack of agreement in this area may be due to incoming swell as derived from the SIPS analysis (Figure 10), which in turn, increases the SWA wind speed.

The change in the CMOD4 wind speed across the front from area 2 to 3 is about 2-3 m/s, as indicated in the contour

plot (Figure 12) of the derived CMOD4 wind speed obtained using the in situ northeasterly wind direction of 220° .

The wind field derived from inverting the CMOD4 model function combined with the SWA wind speed and the calibrated σ_0 is shown in Figure 4, together with the hindcast isobaric map and wind vectors from the hindcast model 90

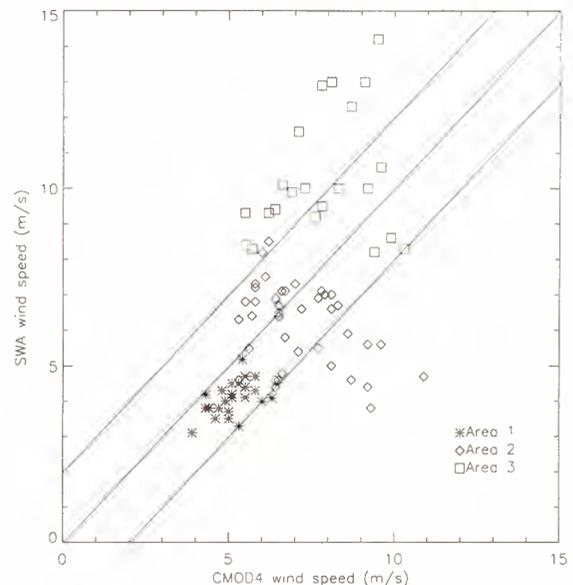


Figure 11. CMOD4 wind speed versus SWA wind speed derived from the SAR images on September 23, 1995, and classified according to the three areas. The interval of ± 2 m/s is indicated.

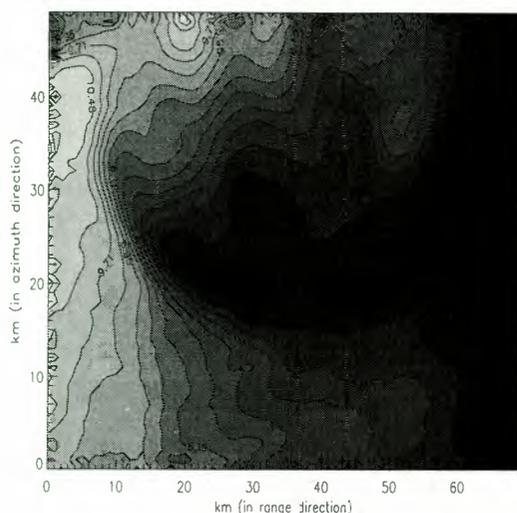


Figure 12. Contour plot of the CMOD4 wind. The structure reflects the change in wind speed across the front seen in the image. Also the local increase in wind close to the coast is clearly seen in the upper left part of the plot. The subimage frame is shown in Figure 4.

min after the satellite overpass. In area 2 a patch of windrows aligned with the in situ wind direction stretching along the coast and eastward in the top of the second image has no wind vectors which can be associated with diverging SWA and CMOD4 wind speeds. The presence of these windrows suggests that there may be a local increase in the wind speed in this area. This is also reflected in the CMOD4-derived wind speed (in area 2), which is about 2 m/s higher than the CMOD4 wind speed in the surrounding areas, whereas the SWA wind speed does not respond similarly. The increased CMOD4 wind speed is also seen in Figure 12 as a peak in the upper left part of the plot.

During the satellite overpass, the ship position was about 75 km northeast of the wind front. Prior to the overpass, the winds varied from 8 to 15 m/s, and at overpass the wind speed in the area of the R/V *Håkon Mosby* was about 12 m/s, while the wind direction remained constant around 250° during the observation time.

Unstable stratification of about 2°C slowly changing toward neutral is encountered along the ship track during the observation interval. This may help to explain why CMOD4 seems to overestimate the wind speed in area 3 (Figure 11).

Table 3 summarizes the observations for the three areas analyzed in case 2 including a brief characterization of areas, features, SIPS results and the derived SWA versus CMOD4 wind speed differences, as well as comparison of the wind direction retrievals.

3.3. Case 3: SST Oceanographic Front and Wind Front

As shown in the beginning of this section, the SAR images from September 27, 1995 (Figure 5, left) express a westward flowing coastal jet bounded by two distinct, and in some places parallel fronts as well as a narrow-banded diagonal wind feature with some backscatter patches possibly originating from rain showers and downdraft. In Figure 13 the SAR is shown to be classified into five distinct zones covering the vicinity of as well as inside these frontal features. Superimposed on the figure is the ship track and the locations and extracts of four SAR image power spectra.

A total of 144 subimages were analyzed for the September 27 image (Figures 5 and 13), of which only 8% have a wind speed difference of less than 2 m/s in the range from 3 to 15 m/s. Figure 14 shows that the SWA wind speed is up to 10 m/s higher than the CMOD4, in particular, for areas 2 and 3. A shift of about 1.5 m/s to 2 m/s in the CMOD4 wind speed is observed across the jet and across the wind front between areas 4 and 5, as seen in the left plot in the Figure 14. In contrast to CMOD4 wind speeds, there is a large spread in the

Table 3. Summary of Analyses of September 23 SAR Images

Area Location (SAR Images 1,2, and 3, From Top to Bottom in Figures 4 and 10)	Features Observed in the Different Areas	Waves Obtained From SIPS	SWA-CMOD4 Wind Speed Comparison (Magnitude of Difference in the Area)	Wind Direction		
				SWA Combined With CMOD4	Hindcast	In Situ
Area 1, all open ocean sub-images in image 1	Some undefined brighter areas	Near azimuth propagating mode (about 250 m)	SWA < CMOD4 (0 to 1 m/s)	260°	250°	
Area 2, north-east of front	Some windrows (near the coast) that disappear in the area bounded by the front	Short mode of 75 m in near range direction and a weak strong mode of 200 m about 45° to range direction	SWA < CMOD4 (0 - 4 m/s) closer to front: SWA > CMOD (1 - 2 m/s)	305° 277°	262° 266°	260°
Area 3, south of front	Windrows	Short mode of 75 m in near range direction and a strong mode of 200 m about 45° to range direction	SWA > CMOD4 (2 - 6 m/s)	265°	261°	

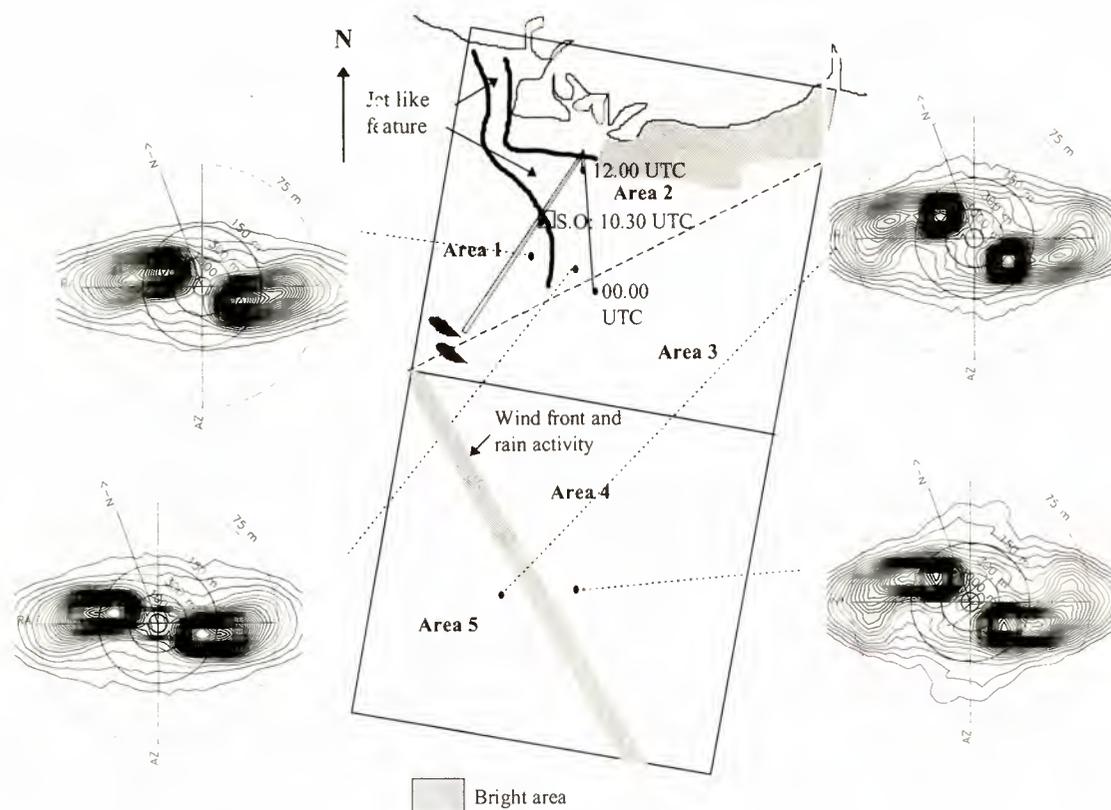


Figure 13. Identification of subareas of the SAR images on September 27, 1995 (center) together with the ship track prior to and during satellite overpass. The gray area marks the presence of windrows, and the lines roughly indicate the orientation but not the spacing of the windrows. The ship position at satellite overpass marked S.O. A selection of SAR image power spectra in the different areas is shown at left and right. The concentric circles indicate wavelengths of 1000 m (inner ring), 300, 150, and 75 m, RA is range direction, AZ is azimuth direction, and N is north.

SWA-derived wind speed for the areas. However, no evidence of wind shifts is found across the fronts.

From the SIPS analysis, a strong peak in the energy occurs between 150 and 300 m wavelength, 30° to 50° to the range axis. This peak in the SIPS is most likely caused by a

corresponding peak in the real ocean wave spectrum, suggesting the presence of incoming swell. As proposed previously, the swell seems to strongly affect the wind field results, in particular, the SWA wind speed and the directional estimates.

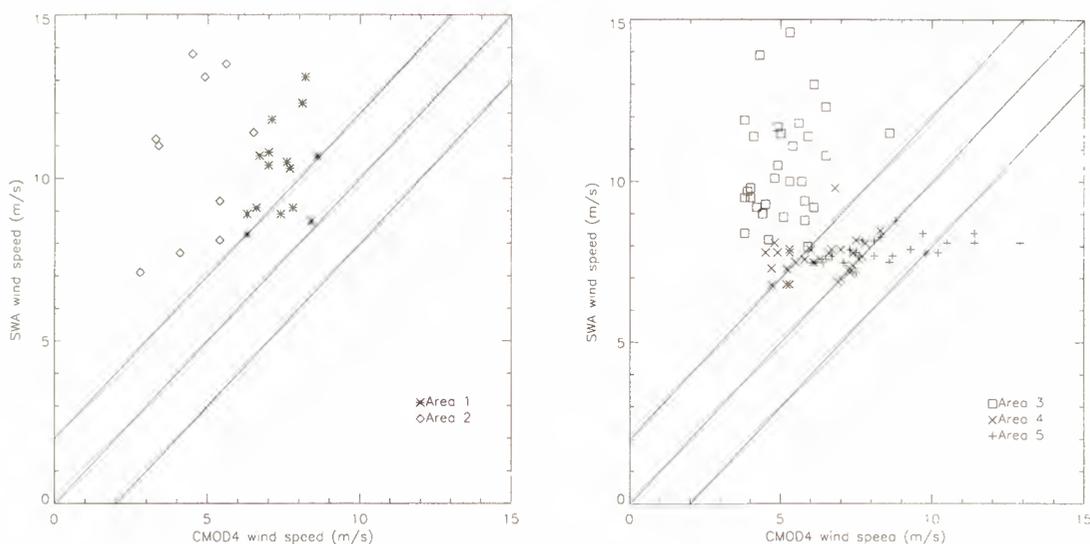


Figure 14. CMOD4 wind speed versus SWA wind speed derived from the SAR images on September 23, 1995. The interval of ± 2 m/s is indicated.

Table 4. Summary of Analyses of September 27 SAR Images

Area Location (SAR Images 1,2, and 3, From Top to Bottom in Figures 2 and 6)	Features Observed in the Different Areas	Waves Obtained From SIPS	SWA-CMOD4 Wind Speed Comparison (Magnitude of Difference in the Area)	Wind Direction		
				SWA Combined With CMOD4	Hindcast	In Situ
Area 1, west of the jet feature in the first image	Uniform area some dark patches	200 m mode traveling about 30° relative to range direction	SWA>CMOD4 (4 to 7 m/s)			220°
Area 2, inside the jet feature in the first image	Bright patches	150 to 200 m mode traveling about 15° to 30° relative to range direction	SWA>CMOD4 (3 to 11 m/s)			260°
Area 3, south-east of diagonal in the first image	Uniform area	about 150 m mode traveling about 15° relative to range direction	SWA>CMOD4 (4 to 12 m/s)		205°	
Area 4, north-east of diagonal in the second image	Uniform area	150 to 200° traveling about 15° to 30° relative to range direction	SWA>CMOD4 (4 to 8 m/s)		220°	
Area 5, south-west of diagonal in the second image	Some bright spots (rain cells)	Strong 200 m mode traveling 45° relative to range axis	SWA>CMOD4 (4 to 8 m/s)	206°	219°	

Because of the large deviation in the SWA and CMOD4 wind speed, the retrievals of wind directions and in turn, wind vectors were possible only from area 5.

Table 4 summarizes the observations for the five areas analyzed in case 3 including a brief characterization of areas, features, SIPS results and the derived SWA versus CMOD4 wind speed differences, as well as comparisons of the wind direction retrievals. In comparison to the previous results, in particular from case 1, the results are poor. The SWA is consistently larger than CMOD4 by at least 4 m/s and completely without any correlation for the conditions encountered in case 3.

The boundary layer conditions prior to and during SAR acquisition observed from the R/V *Håkon Mosby* reveal a complicated and variable pattern in both time and space. During the ERS-1 overpass, the ship is located at the edge of the SST front. The measured wind speed varies from 3 to 10 m/s and is about 9 m/s at the time of satellite overpass. There is also a rapid variation in wind direction from 150° to 300° with a direction of approximately 220° at the satellite overpass. The observed air-sea temperature difference of about -1.5°C gave unstable ABL stratification west of the jet, while the air-sea temperature difference appears to be small (neutral ABL stratification) inside the jet. Under moderate wind speed, these gentle changes in stratification can cause a shift in the surface roughness, which is expressed by the SAR.

4. Discussion of Results

Table 5 provides the available in situ measurements as taken from the R/V *Håkon Mosby* and the buoy (for case 1) and corresponding SAR measurements taken from a SAR 10

× 10 km² subimage centered at the position of the R/V *Håkon Mosby* using the geographical corner and center coordinates provided with the SAR image and global positioning system (GPS) records from the ship. The database has also been enlarged, exceeding the number of SAR images in the case studies to include 10 SAR images where corresponding in situ measurements were available. In evaluating the result of this comparison, it should be noted that the two SAR wind speed retrieval methods are based on spatial averaging 10 × 10 km², while the in situ measurements represent 15 min temporal averages prior to SAR acquisition (corresponding to a 9 km travel distance at a speed of 10 m/s). In all cases, it is seen that the estimated wind speeds from the SWA and CMOD4 agree to within 2 m/s except on September 27. On the other hand, all cases show that the SAR-derived wind speeds are generally lower than the directly measured wind speed.

Table 6 summarizes the obtained wind direction as given in Tables 1-4 compared to the sparse number of in situ measurements and the available hindcast wind directions within the SAR frames. Table 6 reveals good agreement between the estimated wind directions and the in situ and modeled results. A deviance of less than 30° is obtained.

On September 16, when the ship was in the same position as the buoy at the satellite overpass, the wind speeds obtained from the SAR images (case 1) are 4 to 5 m/s lower than the in situ buoy- and ship-measured winds. Almost the same results are obtained when the SAR-derived and ship-measured wind speeds are compared on September 17. In contrast, the SAR-derived wind speeds from the area close to the buoy are in very good agreement with the in situ measurements. The SAR images for these 2 days had in common clear expressions of large areas of homogeneous windrows.

Table 5. Wind Speeds Obtained From SAR Images and In Situ Measurements From the R/V *Håkon Mosby* (H.M.) and the “Metocean” Buoy

Location of In Situ Measurement	Date in 1995	Estimated Wind Speed From SAR Image m/s		In Situ Wind Speed at Overpass (± 2 m/s)	Deviations (Absolute Value) m/s		
		SWA	CMOD4		SWA - CMOD4	In Situ - SWA	In Situ - CMOD4
H.M.	Sept. 16	8	8	13	0	5	5
Buoy	Sept. 16	8	8	12	0	4	4
H.M.	Sept. 17	4	3	5	1	1	2
Buoy	Sept. 17	9	10	10	1	1	0
H. M.	Sept. 23	5	6	12	1	7	6
H. M.	Sept. 22	6	8	12	2	6	4
H. M.	Sept. 14	5	4	6	1	1	2
H. M.	Sept. 26	9	7	14	2	5	7
H. M., (ERS-1)	Sept. 27	9	5	6	4	3	1
H. M., (ERS-2)	Sept. 27	10	4	10	6	0	6
Deviation (Mean)					1.8	3.3	3.7

In case 2 on September 23, the deviations in the SWA and CMOD4 retrievals and the in situ measured wind speed have increased to 6–7 m/s, while the SWA wind speed compares well with the in situ ship measurement in case 3 on September 27. In this latter case the CMOD4 wind speed is far too low. Overall, the magnitude of the in situ measured wind speed in the region remains from 10 to 13 m/s (except for September 17) for these three cases and cannot therefore explain the wide span that results from comparison with the SAR-derived winds. On the other hand, the SAR images from the two days in cases 2 and 3 have in common an increase in spatial variability in backscatter at the expense of clear expressions of homogeneous windrow areas. The effect of this on the wind and wave field retrievals from the SAR images may cause the larger discrepancies.

On average, Table 5 shows that the mean deviation of the two SAR wind speed retrieval methods is less than 1.8 m/s, and grows to 3.3 m/s and 3.7 m/s for the in situ SWA and the in situ CMOD4 wind speeds, respectively. Regarding the wind direction retrievals, on the other hand, there is a good overall agreement, except for the case 3 on September 27 and the additional SAR scene from the same date. Again, we emphasize that this is locally in the position of the in situ measurements and does not reveal the spatial variability.

Although the two SAR retrieval methods agree quite well, as demonstrated in the previous section and summarized in Table 5, several possible effects and error sources in the

CMOD4 and SWA wind field retrievals may explain the wide range of results when locally compared to in situ measurements as suggested in Table 7. This may also explain some of the spatial variability we obtain in comparing the SWA and the CMOD4 results in cases 1 to 3. In the following, some of these are considered in more detail, without reflecting on possible uncertainties and errors from the in situ measurements.

As mentioned above, fetch-limited seas can lead to underestimation of the local SWA wind retrievals since the sea state has not reached equilibrium with the local wind speed [Kerbaol *et al.*, 1996; Korsbakken and Johannessen, 1996]. This effect can be compensated for, provided the wind direction and fetch distance can be determined. This effect may explain the deviation between the SWA and in situ measurements for case 1, September 16 and 17 SAR images, where we expect fetch-limited seas at least in the northern parts of the images. The fetch dependency is theoretically estimated by Kerbaol [this issue] for fetch of 50, 100, and 200 km. The azimuth cutoff wavelength decays as a function of the wind speed at different fetch compared to a fully developed sea. The corresponding underestimation of wind speed in SWA becomes significant at 10 m/s wind speed if the seas are not fully developed. An approximate correction for our results in case 1 is estimated if the fetch is assumed to have the following characteristics: (1) 50 to 100 km in the position of in situ measurements (Figure 2 and Table 5) on September 16,

Table 6. Wind Directions Obtained From SAR Images and In Situ Measurements From the R/V *Håkon Mosby* (H.M.), "Metocean" Buoy and Hindcast Data

Date in 1995	Estimated Wind Direction From SAR Image.	In Situ Wind Speed Direction at Overpass H. M. or Buoy	Nearest Available Hindcast Data	Deviations (Absolute Value) Wind From SAR	
				In Situ	Hindcast
Sept. 16	84	-	114		30
	105	-	152		47
	125	-	124		1
	136	130	114	6	22
	96	-	123		27
Sept. 17	98	-	124		26
	108	-	147		39
	101	100	124	24	23
	100	-	107		7
	87	-	96		9
	140	-	101		39
Sept. 23	260	-	250		10
	305	-	262		43
	277	260	266	6	11
Sept. 27	206		219		13
Deviation (Mean)				12	23

All numbers except dates in degrees.

and the correction in wind speed 1 to 3 m/s; (2) 50 km in the ship position (Figure 3) on September 17, yielding a correction of 1 m/s; and (3) 100 km in the buoy position (Figure 3) on September 17, yielding a correction of 1 m/s. Adding these corrections to the SWA wind speed will reduce the difference between the SWA wind speeds and in situ measurements.

In contrast, incoming swell generated by surface winds outside the region will, of course, introduce a surface wave field that is not, again, directly in equilibrium with the local wind field. Hence, the SWA method might overestimate the local wind speed because of increased smearing.

The CMOD4 wind speed estimates appear not to be affected by fetch-limited seas. Instead, this method is affected by surface boundary layer conditions such as the stratification and presence of surface films. While the former effect can be corrected for via a boundary layer model for stable (intensify wind speed at 10 m) or unstable (reduce wind speed at 10 m) stratification [e.g. Wu 1993], the effect of the latter will always dampen the surface roughness and hence reduce the wind speed. Assuming the weakly unstable (-1.0 to -2.0°C)

boundary layer stratification reported from the ship to be valid across the entire images, the CMOD4 wind speed at 10 m above the surface will be reduced. At the same time an increased presence of film due to accumulation of breaking air bubbles returning to the surface after the wind maximum passed early on September 16 will further magnify this reduction. Quantifying here is difficult because the attenuation of σ_0 depends on the unknown thickness of the surface films.

An uncertainty factor in applying the CMOD4 model to SAR data is local variations in the surface wind direction and, in turn, our assumption of a local in situ measurement to be valid in all subimages in the imaged area. The only confidence in this assumption applies to case 1, where windrows confirm the homogeneity of the wind direction. The variability in wind speed due to variations in the wind direction is illustrated in Figure 15 for examples of 5 m/s and 15 m/s CMOD4 wind speed. Deviance in CMOD4 wind speed due to error in the wind direction used, as taken from in situ measurements or windrows in SAR images, is plotted for variations around 0°, 45°, and 90° wind direction (relative to the satellite look direction). Because of the wind-directional asymmetry in the

Table 7. Overview of Observed Problems and Possible Explanations for the Case Studies

Date in 1995	Observed Problem	Possible Explanation(s)
16 September	SWA wind speeds lower than in situ measurements	Fetch limited seas
	CMOD4 wind speed lower than in situ measurements	Local variations in wind direction; surface films
17 September	SWA wind speeds lower than in situ measurements	Fetch limited seas
	CMOD4 wind speeds lower than in situ measurements	Variations in wind direction
23 September	SWA wind speed generally higher than CMOD4 wind speed	Incoming swell (increase in SWA); all effects reducing CMOD4 wind speed
27 September	SWA wind speed generally higher than CMOD4 wind speed	Incoming swell
	CMOD4 wind speed less than in situ wind speed	Stratification; variations in wind direction

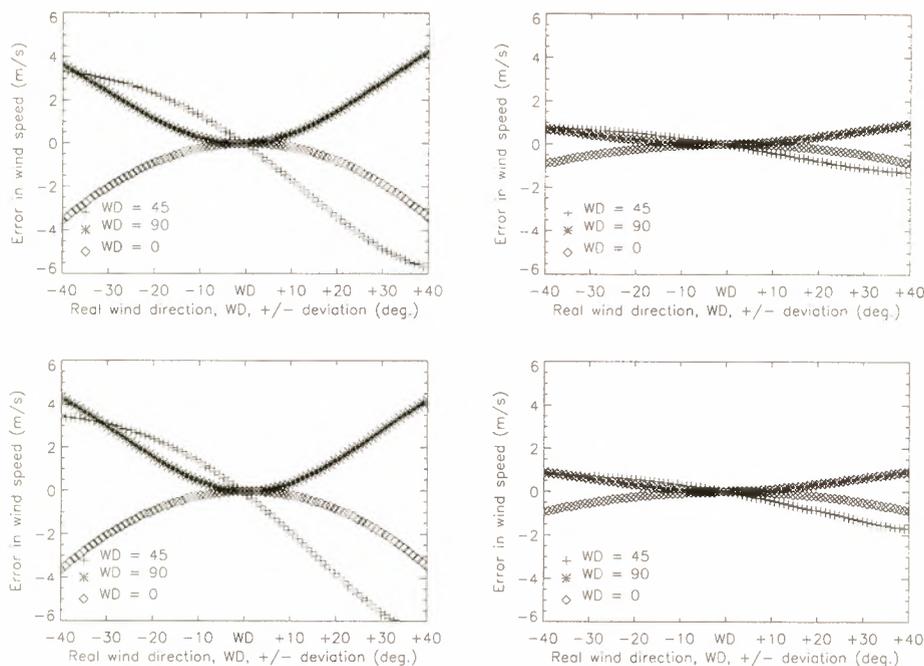


Figure 15. The error in the estimated CMOD4 wind speed due to deviations in the wind direction. Upper left plot is for a wind speed of 15 m/s for 0°, 45°, and 90° wind direction (WD) and WD ± 0° to 40° relative to the satellite look direction; the incidence angle is 20°. Upper right is a similar plot for a wind speed of 5 m/s at 20° incidence angle. At WD the error in wind speed is 0 m/s. Lower left and right plots follow the same pattern as the upper left and right plots except for the incidence angle, which is changed to 26°.

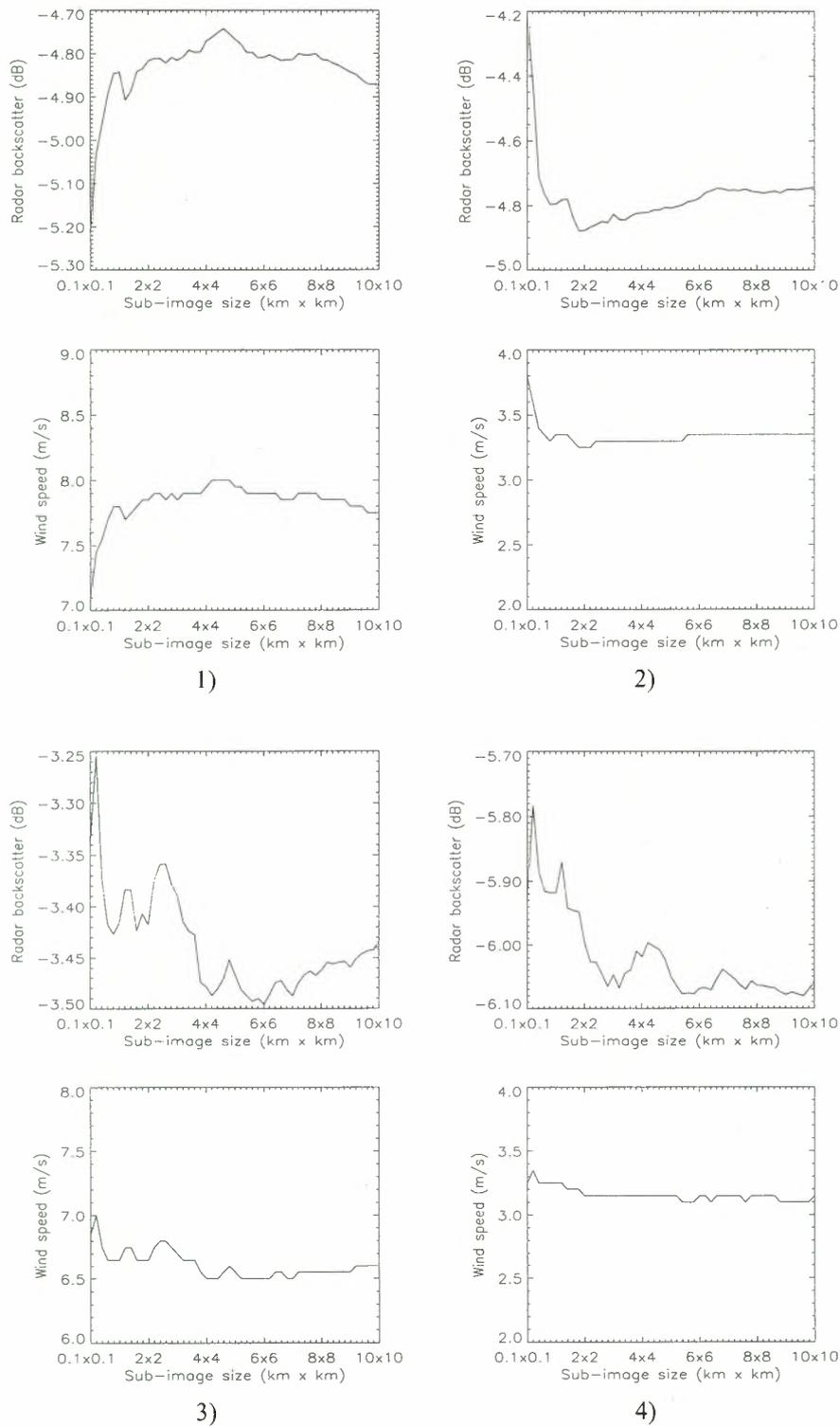


Figure 16. Plots of the σ_0 and corresponding CMOD4 wind speed measurement as a function of subimage size. Plots 1 and 2 are taken from the same position as the in situ measurements from the R/V *Håkon Mosby* in case 1 on September 16 and 17. Plots 3 and 4 corresponds to in-situ measurements in case 2 and 3 on September 23 and 27, respectively.

CMOD4 model, the largest errors occur, if the used wind direction has positive deviations from a real wind direction of 45° . Figure 15 also shows that the error estimate does not vary in the range of ERS SAR incidence angles of about 19° to 26° . For the relatively high wind speeds observed in most of the cases in this study, the wind direction used to derive the

CMOD4 wind speed becomes a significant error source and may cause errors up to 5 m/s in the derived wind speed.

In comparing the CMOD4 and in situ measurements, we do not obtain as good quantitative agreement as *Vachon and Dobson* [1996] in their comparison to in situ measurements. The discrepancy between our CMOD4 results and the results

obtained by *Vachon and Dobson* [1996] seems to be difficult to explain and has to be further investigated. In this work the spatial resolution of the CMOD4 estimates was primarily chosen to be the same as the spatial resolution for the SWA method, which in turn, may explain the better agreement when comparing SWA and CMOD4 wind speed retrievals than when comparing CMOD4 and local in situ measurement. *Vachon and Dobson* [1996] used a subimage size of 4×4 km. However, choosing this same size of the subimage did not significantly improve the agreement with in situ measurements in our case, as illustrated in Figure 16, which shows the variability in σ_0 and the corresponding CMOD4 wind speed as a function of increasing the subimage from 100×100 m to 10×10 km. The enlargement of the subimages is done symmetrically around the location of the in situ measurements from the R/V *Håkon Mosby*. In general, the plots in Figure 16 show small variations in σ_0 as a function of the subimage size and the corresponding CMOD4 winds do not change significantly in any of the cases.

Moreover, the CMOD4 is capable of quantifying the change in wind speed and direction across wind fronts. The radar backscatter will have a maximum change, if for a given wind speed change the wind direction at the same time changes from the near range in the high wind zone to the near azimuth on the low wind side of the front [*Johannessen et al.*, 1991].

In summary, the performances of the two SAR wind retrieval methods are best for case 1 on September 16 and 17. Images for these days are similar with respect to the meteorological situation, the windrows in the SAR images and the SIPS. Gradual spatial variations in both the SWA and CMOD4 wind speeds over several subimages (considered as independent samples) therefore give confidence in the wind speed estimates, assuming that the wind and wave field undergoes small changes at a scale similar to the subimage.

5. Summary

In this paper we have shown that the radar backscatter and spectral signatures of the ocean surface obtained from ERS-1 and 2 SAR images can provide valuable and quantitative information on near-surface wind speed and wind direction. Ten ERS SAR images, from September 16, 17, 23, and 27, 1995 have been examined by studying their σ_0 and spectral properties in 736 subimages. Two different wind retrieval models, SWA and CMOD4, have been applied to the data. We have found by comparing the SWA and CMOD4 wind speed retrievals that 65% of the sub images gives a wind speed difference less than 2 m/s, but the comparison also reveals relatively large spatial variations in the agreement due to different scattering clusters. In general, we also obtain an underestimation of the wind speed from SAR in comparison with the available in situ observations. The obtained wind directions are also in good agreement with the hindcast results and in situ measurements and are consistent with the good agreement in wind speed. However, we also observe diverging results in specific areas.

It is demonstrated that the surface conditions impact on the performances of the different wind field retrieval methods and their corresponding results. The presence of homogeneous windrows is clearly favored. For fetch-limited seas and in vicinity of wind fronts, on the other hand, the SWA method underestimates the wind speed. For fetch-limited seas the

waves are not in equilibrium with the near-surface wind speed. Hence the distribution of the velocity field of surface scatters, as introduced by the orbital velocity of the waves, will be narrower than for fully developed seas, leading to underestimation. Moreover, the relatively large relaxation rate for the longer wind waves (also the medium wavelengths suppressed by the general resolution of the SAR) allows these waves to propagate across a wind front. Hence they will maintain their original equilibrium state over some distance away from the front. In turn, the wind front will not be resolved properly by the SWA method. For the same reason, the presence of swell may cause an increase in the derived wind speed. In contrast, the CMOD4 method is not limited by these conditions.

While absolute image calibration is necessary for the CMOD4 method, it is not required for the SWA method, since the former method uses radar backscatter values while the latter uses spectral characteristics. However, as mentioned in the analysis of the September 17 image, the SWA method is limited to SAR images containing clear wave modulation from which the azimuth cutoff can be derived. In the case of very low wind conditions (lower than approximately 3 m/s) or in the presence of slicks, the SWA method will therefore break down, but these conditions will also affect the CMOD4, since the threshold wind speed for C band waves is around 3 m/s [*Johannessen et al.*, 1996].

Further investigation of these methods is necessary in order to fully understand their limitations and strengths, particularly in regard to the surface conditions and synoptic weather situation. However, as reported in this and other papers on wind retrievals from SAR (included in this special section), the possibility looks promising in regard to the continuation of regular spaceborne SAR observations.

Acknowledgments. This work was funded by the Norwegian Research Council as a part of a strategic SAR program at the Nansen Environmental and Remote Sensing Center. The SAR data were provided under the ESA AO2/N108 (announcement of opportunity) program. The ship time for the Coast Watch'95 experiment was provided and funded by the University of Bergen. The following contributions are highly appreciated: Ken Davidson and Paul Fredrickson at the Naval Postgraduate School for providing the buoy measurements and Magnar Reistad at the Norwegian Meteorological Institute for providing the weather diagnostics.

References

- Chapron, B., T. E. Fouhaily and V. Kerbaol, Calibration and validation of ERS wave mode products, *Doc. DRO/OS/95-02*, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, 1995.
- Hasselmann, K. and O.H. Shemdin, Remote Sensing Experiment Marsen. *Int. J. Remote Sens.*, 3, 139 - 361, 1982.
- Johannessen, J.A., R. Shuchman, O. M. Johannessen, K. Davidson and D. R. Lyzenga, Synthetic aperture radar imaging of ocean circulation features and wind fronts, *J. Geophys. Res.*, 96, 10,411-10,422, 1991.
- Johannessen, J.A., P.W. Vachon and O.M. Johannessen, ERS-1 SAR imaging of marine boundary layer processes, *Earth Observation Quarterly*, European Space Agency, Paris, France, 1995.
- Johannessen, O. M. et al., Coast Watch'95 ERS-1,2 SAR applications of mesoscale upper ocean and atmospheric boundary layer processes off the coast of Norway, paper presented at IGARSS'96, Int. Geosci. and Remote Sens. Symp., Lincoln, Nebr., 1996.
- Kerbaol, V., B. Chapron, T. El Fouhaily, and R. Gareilo, Fetch and wind dependence of SAR azimuth cutoff and higher order statistics in a mistral case, paper presented at IGARSS'96, Int. Geosci. and Remote Sens. Symp., Lincoln, Nebr., 1996.

- Kerbaol, V., B. Chapron, and P. W. Vachon, Analysis of ERS-1/2 synthetic aperture radar wave mode images, *J. Geophys. Res.*, this issue
- Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, K. Davidson and P.A.E.M Janssen, *Dynamics and Modelling of Ocean Waves*, Cambridge Univ. Press, Cambridge, 1994.
- Korsbakken, E., Quantitative wind field retrievals from ERS SAR images, Young graduate trainee report, Ocean and Sea Ice Unit, Earth Sci. Div., Eur. Space Res. and Technol. Cent., Eur. Space Agency, Noordwijk, The Netherlands, 1996.
- Korsbakken, E and J. A. Johannessen, Quantitative wind field retrievals from ERS SAR images, paper presented at the Third ERS Workshop, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, June 18-20, 1996.
- Laur, H., P. Meadows, J. I. Sanchez and E. Dwyer, ERS-1 radiometric calibration, technical report, Eur. Space Res. Inst., Frascati, Italy, 1993.
- Laur, H., P. Bally, P. Meadows, P. J. Sanchez, B. Schaettler, E. Lopinto, Derivation of the backscattering coefficient σ_0 in ESA ERS SAR PRI products, Doc. *ES-TN-RS-PM-HL09*, issue 2, rev. 2, Eur. Space Res. Inst., Frascati, Italy, June 28, 1996.
- Meadows, P. J. and C. J. Willis, Derivation of radar cross section coefficient in UK-PAF ERS-1 SAR PRI products, GEC-MARCONI Res. Cent, Chelmsford, Essex, 1995.
- Monaldo, F. M., The consequences of sampling variability on the estimation of wave number and propagation direction from spaceborne SAR image spectra, *IEEE Trans. Geosci. Remote Sens.*, 29 (1), 113-119, 1991.
- Scoon, A., T. S. Robinson, and P. J. Meadows, Demonstration of an improved calibration scheme for ERS-1 SAR imagery using a scatterometer wind model, *Int. J. Remote Sens.*, 17 (2), 413-418, 1996.
- Smith, D. S., Coefficients for sea surface wind stress, heat flux, and wind profiles as a function of wind speed and temperature, *J. Geophys. Res.*, 93, 15,467-15,472, 1988.
- Stoffelen, A. and D.L.T. Anderson, ERS-1 scatterometer data and Characteristics and wind retrieval skills, Proceeding of First ERS-1 Symposium, *Eur. Space Agency Spec. Publ.*, ESA SP-359, 1993.
- Vachon, P.W, and F.W Dobson, Validation of wind vector retrieval from ERS-1 SAR images over the ocean, *The Global Atm. and Ocean Sys.*, 5, 177-187, 1996.
- Vachon, P.W., H. E. Krogstad, and J. S. Paterson, Airborne and spaceborne synthetic aperture radar observations of ocean waves, *Atmos.-Ocean* 32, 83-112, 1994.
- Vachon, P.W., J.A Johannessen, and D. P Browne, ERS-1 SAR images of atmospheric gravity waves, *IEEE Trans. Geosci. Remote Sens.*, 33 (4), 1014-1025, 1995.
- Wakeman, C.C, C. L. Rufenach, R. Shuchman, J. A. Johannessen, and K. Davidson, Wind vector retrieval using ERS-1 synthetic aperture radar imagery, *IEEE Trans. Geosci. Remote Sens.*, 34, 1343-1352, 1996.
- Wu, J., Ripples and oceanic remote sensing, in *Proceedings Environment '93*, The Hong Kong Univ. of Sci. and Technol., The Commercial Press Ltd., Hong Kong, 1993.

J. A. Johannessen, ESA-ESTEC, P.O. Box 299, 2200 AG Noordwijk, The Netherlands. (email: jjohnne@estec.esa.nl)

O. M. Johannessen and E. Korsbakken, Nansen Environmental and Remote Sensing Center, Edv. Griegs Vei 3A, N-5037 Solheimsviken, Bergen, Norway. (email: ola.johannessen@nrsc.no; erik.korsbakken@nrsc.no)

(Received March 18, 1997; revised August 25, 1997; accepted September 8, 1997.)

Katabatic wind fields in coastal areas studied by ERS-1 synthetic aperture radar imagery and numerical modeling

Werner Alpers

Institut für Meereskunde, Universität Hamburg, Hamburg, Germany

Ulrike Pahl and Günther Gross

Institut für Meteorologie und Klimatologie, Universität Hannover, Hannover, Germany

Abstract. Synthetic aperture radar (SAR) images acquired over Mediterranean coastal waters from the first European Remote Sensing Satellite (ERS-1) showing sea surface manifestations of katabatic wind fields are presented. In particular, sea surface roughness patterns generated by katabatic winds blowing from 1800 m high mountains through a broad valley at the western coast of Calabria (southern Italy) into the Gulf of Gioia is investigated. The ERS-1 SAR images show that their areal extent and shape vary strongly, depending on the meteorological conditions. The roughness pattern sometimes has the form of a mushroom, an elongated tongue, a broad blob, or a narrow truncated band. For one event (September 8, 1992) we have simulated the wind field at the sea surface by using a nonhydrostatic mesoscale atmospheric model and then compared it with the wind field derived from the ERS-1 SAR image by using the C band wind scatterometer model CMOD4. The comparison shows that the atmospheric model reproduces quite well the mushroom-like form of the wind field pattern, while the wind speed is obtained somewhat lower than the one inferred from the SAR image. This study demonstrates that SAR images acquired over coastal waters are well suited to study local wind fields in coastal areas and to test numerical models that describe local wind fields which extend from the coast onto the sea.

1. Introduction

It is well known that on synthetic aperture radar (SAR) images of the sea surface, often "imprints" of atmospheric phenomena occurring in the marine boundary layer can be delineated. They include atmospheric boundary layer rolls [Thompson *et al.*, 1983; Alpers and Brümmner, 1994], atmospheric convective cells [Alpers, 1995], atmospheric internal waves [Vachon *et al.*, 1994; Alpers and Stilke, 1996]), tropical rain cells [Atlas, 1994; Atlas *et al.*, 1995; C. Melsheimer *et al.*, Investigation of multifrequency / multipolarization radar signatures of rain cells over the ocean using SIR-C/X-SAR data, submitted to the *Journal of Geophysical Research*, 1997], and meteorological fronts [Johannessen *et al.*, 1991; Rufenach *et al.*, 1991]. These atmospheric phenomena become visible on radar images of the sea surface because they are associated with a spatially varying wind stress at the sea surface which modulates the sea surface roughness and thus the backscattered radar power,

which usually is expressed in terms of the normalized radar cross section (NRCS). By using a wind scatterometer model one can convert the NRCS values of the SAR image into wind speed at the sea surface, provided the wind direction is known [see, e.g., Vachon *et al.*, 1994; Alpers and Stilke, 1996]. On SAR images of the sea surface the wind direction can often be inferred from the orientation of "wind streaks," from the position of wind shadows behind isolated islands, or, as in the case of katabatic winds, from the direction of the wind "tongues" (see section 2) emanating from coastal valleys.

In this paper we present several SAR images that were acquired by the first European Remote Sensing Satellite (ERS-1) over the Mediterranean Sea off the coast of Calabria (southern Italy), which show different forms of roughness patterns generated by katabatic winds blowing from the coastal mountains onto the sea. Katabatic winds are caused by the gravitational flow of cold air off high terrain. The cold air that is formed at night near the ground in sloping terrain flows downslope because of the differential cooling between the ground surface and the free atmosphere at the same height level some distance away. The direction of the cold air flow is

Copyright 1998 by the American Geophysical Union.

Paper number 97JC01774.
0148-0227/98/97JC-01774\$09.00

controlled almost entirely by orographic features. Katabatic winds occur late in the evening and at night after the air in the mountains has cooled down. Superimposed on the katabatic wind fields are often other local wind fields, like sea or land breezes and valley winds [Mahrt and Larsen, 1982]. When the airflow associated with these local coastal winds hits the sea surface, it increases its short-scale roughness. According to Bragg scattering theory [Valenzuela, 1978] the short-scale sea surface roughness determines the backscattered radar power or the NRCS. The higher the wind speed, the larger the NRCS and thus the brighter the SAR image intensity. The average NRCS is determined mainly by the local wind speed, but also, other oceanic phenomena, like the long ocean surface wave field, have an effect on the average NRCS. However, these effects are of second order and can often be neglected. The sea surface roughness patterns generated by katabatic winds often have shapes that resemble "tongues," "mushrooms," or broad "blobs" emanating from coastal valleys. The areal extent and shape of the sea surface roughness pattern depends, among others, on the orography of the valley, the synoptic weather situation, the thermal state of the air in the valley, and the sea-land temperature difference.

In this investigation we concentrate on studying the sea surface roughness pattern generated by katabatic winds blowing through the broad valley of Gioia in southern Italy into the Gulf of Gioia in the Tyrrhenian Sea (Mediterranean Sea). This roughness pattern is usually very large because this valley has a favorable orography for channeling the wind and is surrounded by 1800 m high mountains, where large volumes of air are cooled down after sunset. For this study we had available in total 36 ERS-1 SAR images of the Gulf of Gioia, which were acquired always at 2213 LT \pm 1 min in the years 1992-1995. Out of these 36 images, 20 show sea surface manifestations of katabatic winds in the Gulf of Gioia. Five typical examples, which encompass the full range of the observed patterns, are depicted in Figures 3-7.

The SAR images show the actual state of the sea surface roughness which can be caused by oceanic as well as by atmospheric phenomena [Alpers, 1995]. However, from the form and the location of the roughness pattern one can in most cases, but not always, determine unambiguously whether they are of oceanic or atmospheric origin. The best way to extract maximum information from SAR images of the sea surface showing sea surface manifestations of mesoscale atmospheric phenomena is to combine them with surface observations and/or with mesoscale numerical models [McNider and Pielke, 1984; Gudiksen et al., 1992]. The great advantage of SAR images of the ocean surface is that they provide very detailed information on the areal extent of mesoscale atmospheric phenomena in the marine boundary layer.

For the ERS-1 SAR image acquired on September 8, 1992, which shows a roughness pattern resembling a mushroom, we have carried out numerical simulations by using a nonhydrostatic mesoscale atmospheric

model. This model yields as output, among others, the wind field at the sea surface, which can be compared with the sea surface wind field inferred from the SAR image intensity. Furthermore, we discuss qualitatively how different meteorological parameters affect the areal extent and shape of the sea surface roughness pattern generated by the katabatic wind.

The present analysis shows that SAR images acquired over coastal waters are a valuable means to study katabatic wind fields in mountainous coastal regions. Furthermore, they can also be used to validate numerical models describing coastal wind fields. These local winds strongly affect the microclimate in coastal regions. Furthermore, they determine to a large extent the advection and dispersion of pollutants in the atmosphere, which are released by industrial plants located in coastal areas. Since these local winds also act on the sea surface and generate there local surface currents, they also have a strong effect on the advection and dispersion of pollutants in coastal waters. Since most of the world's population lives in coastal areas and since most pollutants are released into the environment near coasts, the study of these local winds is also of great relevance for environmental protection purposes.

2. ERS-1 SAR Images

Figures 2-7 show synthetic aperture radar (SAR) images that were acquired from the first European Remote Sensing Satellite ERS-1 (launched on July 17, 1991) over the Mediterranean Sea at the western Calabrian coast (southern Italy). This satellite carries a SAR that operates at a radar frequency of 5.3 GHz (C band), transmits and receives the radar signals at vertical polarization, and has a right-looking antenna which illuminates the Earth's surface at a mean incidence angle of 23°. The ERS-1 SAR images presented in this paper are precision-processed SAR images provided by the European Space Agency (ESA).

The locations of the imaged scenes, which all cover areas of 100 \times 100 km², are marked in the map shown in Figure 1. One of the six ERS-1 SAR images analyzed in this study is from area 1, and the five others are from area 2. On all of these images, roughness patterns generated by katabatic winds blowing from the mountains of the Calabrian Apennine through coastal valleys onto the Mediterranean Sea can be delineated. The SAR image of area 1 depicted in Figure 2 was acquired on February 20, 1995, at 2113 UTC (orbit 18839, frame 783), and the five SAR images of area 2 shown in Figures 3-7 were acquired on September 8, 1992, at 2114 UTC (orbit 6014, frame 765); February 20, 1995, at 2112 UTC (orbit 18839, frame 765); June 3, 1994, at 2113 UTC, (orbit 15079, frame 765); August 24, 1993, at 2114 UTC (orbit 11024, frame 765); and May 4, 1995, at 2113 UTC (orbit 19885, frame 765), respectively. Note that the SAR scene shown in Figure 2 is the northern continuation of the SAR scene shown in Figure 4.

Meteorological data at times close to the time of the SAR data acquisitions, which took place always at 2113

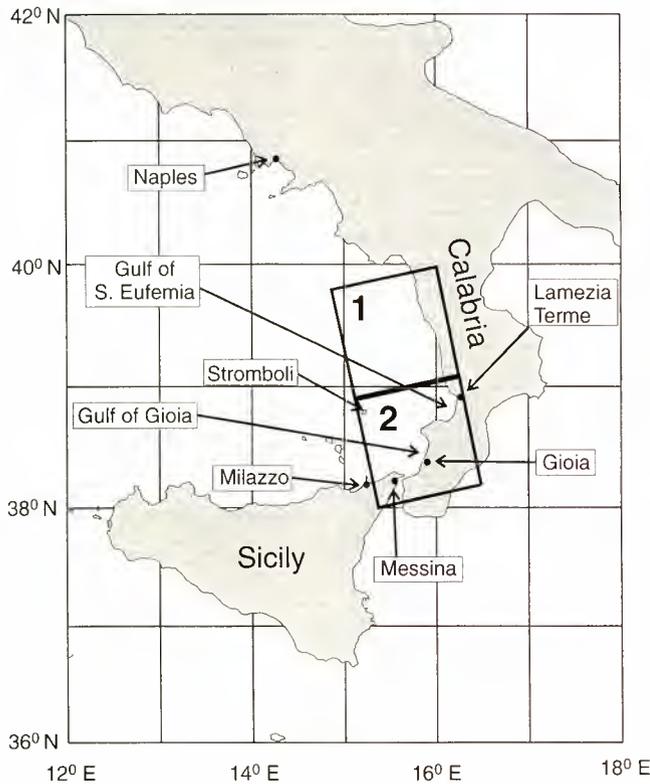


Figure 1. Map of a section of the Tyrrhenian Sea (Mediterranean Sea) with the position of the ERS-1 synthetic aperture radar (SAR) scenes shown in Figures 2-7 inserted. The SAR scene of Figure 2 is from area 1, while the SAR scenes of Figures 3-7 are from area 2.

UTC \pm 1 min or 2213 LT \pm 1 min are available from the Italian stations, Lamezia Terme (38°53'N, 16°15'E) and Messina (38°13'N, 15°33'E) and from weather maps. The locations of these weather stations are marked in Figure 1. They are located in mountainous coastal terrain, which implies that the measured meteorological data are influenced by local atmospheric effects and thus are not necessarily representative for the synoptic weather situation. Wind speed, wind direction, air temperature, and cloud coverage reported by these stations at 2100 UTC for those 5 days on which the SAR scenes shown in Figures 2-7 were acquired are listed in Table 1. Note that according to the meteorological convention, wind direction denotes here the direction where the wind is blowing from, and the angle of the wind direction is counted positively in clockwise direction from north (e.g., a wind direction of 90° means that the wind is blowing from the east). The water temperatures at the sea surface given in the last column of Table 1 are only estimates obtained from in situ temperature measurements carried out from ships transiting the area and from climatological data.

The SAR image depicted in Figure 2 shows a roughness pattern that stretches like a band along the western Calabrian coast and that has a width of up to 28 km. Close to the shoreline, the pattern is strongly influenced by orographic features. Coastal valleys can be localized

by the associated tongue-like bright patterns on the sea surface. Note also the mottled pattern in the left-hand section of this image, which we interpret as sea surface manifestations of atmospheric convective cells. They are generated in calm weather conditions when the air is cooler than the underlying water such that the air is heated from below [Agee, 1989; Mitnik and Viktorov, 1990; Mitnik, 1992]. From Table 1 we see that the water was almost 5°C warmer than the air, and thus a condition for the formation of atmospheric convective cells was fulfilled. The convective cells are associated with a variable horizontal wind field at the sea surface, which modulates the sea surface roughness and thus the SAR image intensity.

The SAR image depicted in Figure 3 shows, on the left-hand side at the bottom, the northeastern part of Sicily and, on the right-hand side, a part of the southern Italian province Calabria, which are separated by the Strait of Messina. This image was acquired on September 8, 1992, \sim 4 hours after sunset. The most noticeable feature visible on this SAR image is, in the central right-hand section, the "mushroom-like" sea surface roughness pattern that seems to emanate from the valley of Gioia. This valley opens to the Gulf of Gioia and is bordered on the eastern side by the mountains of the Calabrian Apennine, which have heights up to 1800 m. By carrying out numerical simulations with a mesoscale atmospheric model we will show in section 5 that this mushroom-shaped pattern is indeed caused

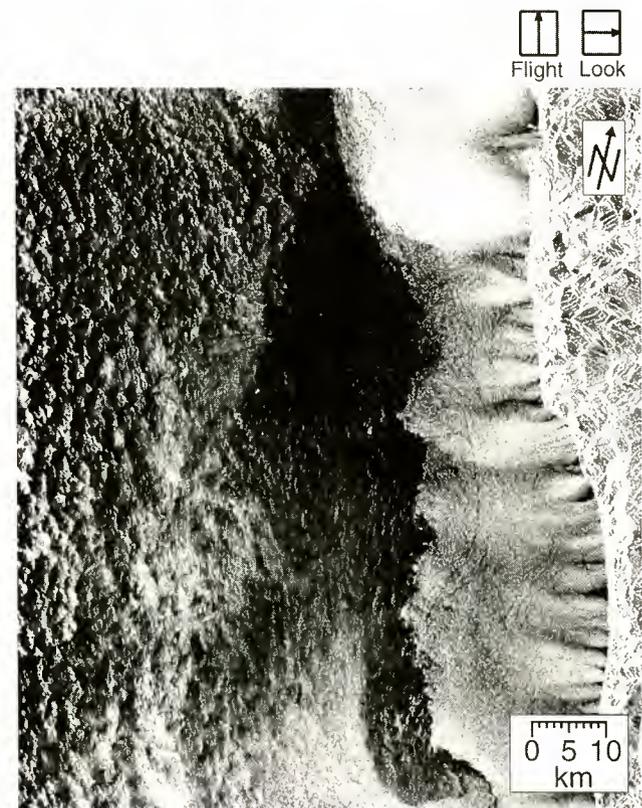


Figure 2. ERS-1 SAR image acquired on February 20, 1995, at 2113 UTC showing a sea surface roughness band stretching along the western Calabrian coast.

Table 1. Meteorological Data Reported by Weather Stations at 2100 UTC for Dates When the SAR Scenes Shown in Figures 2-7 Were Acquired

Figure	Date	Lamezia Terme				Messina				
		Wind Speed, $m\ s^{-1}$	Wind Direction, deg	Air Temperature, $^{\circ}C$	Cloud Cover	Wind Speed, $m\ s^{-1}$	Wind Direction, deg	Air Temperature, $^{\circ}C$	Cloud Cover	Water Temperature, $^{\circ}C$
2,4	February 20, 1995	2.1	270	9.8	2/8	2.6	10	12.4	2/8	14.5 ^a
3	September 8, 1992	1.5	70	16.1	0	0	0	23.6	0	25
5	June 3, 1994	1.0	260	21.3	0	2.1	20	24.2	0	22
6	August 24, 1993	2.1	110	18.3	0	0	0	27.0	0	28 ^b
7	May 4, 1995	1.5	70	9.8	1/8	3.1	310	14.6	2/8	14 ^c

Water temperature was obtained from ship measurements and climatological data. The time of the satellite overflights was 2113 UTC \pm 1 min on all days.

^a Measured at 38.3°N 15.2°E on February 21, 1995, at 0000 UTC.

^b Measured at 38.5°N 16.8°E on August 24, 1993, at 1200 UTC.

^c Measured at 38.5°N 15.5°E on May 6, 1995, at 0600 UTC.

by katabatic winds. We have drawn a line through this pattern, along which we will determine the variation of the SAR image intensity. As will be shown in the next section, we are able to convert the SAR image intensity along this line into wind speed at the sea surface. Also visible in Figure 3, in the upper left-hand sections, are sea surface manifestations of atmospheric convective rolls and, in the lower left-hand section, small "tongue-like" sea surface roughness patterns caused by katabatic winds blowing through narrow valleys at the northern Sicilian coast onto the sea. Another noticeable feature is the broad darkish band stretching along the northern Sicilian coast. We have noticed such darkish bands on many ERS SAR images of the northeastern Sicilian coast (see, e.g., the SAR image shown in Figure 6). We suspect that they result from natural surface films which are accumulated in this coastal region due to blocking and shadowing by the narrow Milazzo Peninsula, which sticks out into the Tyrrhenian Sea like a thorn (see Figure 1). Natural surface films are produced by marine plants and animals and are ubiquitous in coastal waters with high-biological activity like the southern Tyrrhenian Sea, especially during spring and autumn.

The SAR image depicted in Figure 4 shows almost the same area as Figure 3, but it is shifted slightly to the east such that north of the Gulf of Gioia the Gulf of Santa Eufemia becomes visible in the upper right-hand section of the image. Figure 4 was acquired in winter (February 20, 1995) \sim 5 hours after sunset. Remember that this SAR scene is the southern continuation of the SAR scene shown in Figure 2. From the weather map at 2400 UTC we infer that at the time of the SAR data acquisition a light ambient wind was blowing from NNW. On this day the roughness pattern, which is associated with the katabatic wind in the Gulf of Gioia, is broader than on September 8, 1992 (Figure 3) and extends quite far south. As will be discussed in section 6, this can be explained by the NNW ambient wind that pushes the pattern into the SSE direction. Note that a part of the front of the roughness pattern is fringed. We have noticed this phenomenon also in many other SAR images showing sea surface manifestations of katabatic wind fields (e.g., in the ERS-1 SAR image of the Gulf of Gioia acquired on October 12, 1994, at 2113 UTC (orbit 16959, frame 765)). Another feature faintly visible on the SAR images depicted in Figures 4 and 2 is the alignment of the sea surface manifestations of the atmospheric cells in the NNW-SSE direction, i.e., in the direction into which the ambient wind is blowing. It is a well-known fact that atmospheric convective cells tend to align themselves in wind direction [Etling and Brown, 1993; Atkinson and Zhang, 1996]. Therefore such alignments can be used to extract the direction of the ambient wind from SAR images. Note also on the SAR image depicted in Figure 4 the funnel-like roughness pattern in the Gulf of Santa Eufemia in which a small "blob-like" shaped roughness pattern adjacent to the coast is embedded. Very likely the funnel-like pattern is generated by a local wind blowing from the east



Figure 3. ERS-1 SAR image acquired on September 8, 1992, at 2114 UTC showing a “mushroom-like” roughness pattern in the Gulf of Gioia.

coast of Calabria over the mountain pass to the west coast. The blob-like pattern seems to be associated with a katabatic wind blowing from the mountains at the southern side of the valley into the Gulf of Santa Eufemia.

The SAR image depicted in Figure 5 was acquired on June 3, 1994, ~3 hours after sunset. The roughness pattern in the Gulf of Gioia visible on this image has the form of an elongated tongue. On this day the Gulf of Gioia was lying almost in the center of a weak high-pressure system. From the weather map and from Table 1 we may conclude that at the time of the SAR data acquisition a $2\text{--}3\text{ m s}^{-1}$ westerly to southwesterly ambient wind was blowing. On this day the air temperature measured at Lamezia Terme at 2100 UTC was 21.3°C , which was exceptionally high (see Table 1) and almost equal to the water temperature (22°C). West of this “wind tongue,” the sea surface is imaged almost completely dark with isolated bright patches in it. We can think of only one reason for this to happen: the whole area was covered with natural surface films, but it had isolated film-free patches in it where the wind roughened the sea surface. This interpretation is supported by the fact that in early June the spring plankton bloom is still going on and that at this time much surface active material is produced that rises to the sea surface. Another support for this interpretation comes from the sea surface manifestation of an oceanic eddy which can be delineated at the top of Figure 5. The

neighboring ERS-1 SAR scene (not shown here) reveals that this eddy has a diameter of $\sim 40\text{ km}$. It becomes visible on the SAR image because surface films are entrained by the surface current associated with this eddy. Another noteworthy feature visible in the upper right-hand section of Figure 5 is the spherical wave pattern that seems to have its origin in the broad valley that opens to the Gulf of Santa Eufemia. Very likely this is a sea surface manifestation of an atmospheric internal wave generated by the funneled air flow over the above mentioned mountain pass.

The SAR image depicted in Figure 6 was acquired on August 24, 1993, ~4 hours after sunset. At this time an ambient wind was blowing from 110° as evidenced by Table 1 and by the position of the wind shadow behind the island of Stromboli visible in the upper left-hand section of Figure 6. Knowing this wind direction, we can determine the wind speed from the SAR image intensity (see next section). In the area southeast of Stromboli, which is marked in Figure 6 by a square, we measured this wind speed to be 2.7 m s^{-1} , which is consistent with the values given in Table 1. The roughness pattern generated by the katabatic wind blowing through the valley of Gioia merges at the northwestern side with the pattern generated by the ambient wind. Thus, in this section of the image, the front line of the wind tongue cannot be determined.

The SAR image depicted in Figure 7 was acquired on May 4, 1995, ~3.5 hours after sunset. On this day



Figure 4. ERS-1 SAR image acquired on February 20, 1995, at 2112 UTC showing a broad roughness pattern in the Gulf of Gioia.



Figure 5. ERS-1 SAR image acquired on June 3, 1994, at 2113 UTC showing a “tongue-like” roughness pattern in the Gulf of Gioia.

the Gulf of Gioia was under the influence of a high-pressure system with weak pressure gradients, which was favorable for the development of local wind systems. Therefore we attribute the easterly wind of 2.5 m s^{-1} measured at Lamezia Terme at 2100 UTC to a local wind. From the position of the wake behind the island of Stromboli we infer that the ambient wind was blowing over the Gulf of Gioia from WNW, which implies that the coast near Gioia was attacked by a head wind. We have determined the sea surface velocity from the SAR image intensity in the area marked by a square in the SAR image and obtained 3.0 m s^{-1} . Because of this head wind the local winds from the shore onto the sea are significantly attenuated and remain confined to small regions close to the shore. This SAR image shows that the katabatic wind blowing through the valley of Gioia onto the sea is almost completely suppressed while the slope winds at the southwestern slope of Monte Poro are weakened and deflected in direction.

3. SAR-Derived Wind Speed

As already stated in the introduction, we can retrieve the wind speed at the sea surface from the SAR image intensity when the wind direction is known. First, we have to convert the image intensity into values of the normalized radar cross section (NRCS), which we have done by using the recently revised conversion formula supplied by ESA [Laur *et al.*, 1996]. Then the NRCS

values are related to sea surface wind speed by applying a so-called wind scatterometer model. In this investigation we use the C band wind scatterometer model CMOD4, which is presently used by ESA for retrieving sea surface winds from ERS scatterometer data [Stoffelen and Anderson, 1997]. This model converts NRCS values into wind velocities at a reference height of 10 m above the sea surface for a neutrally stable boundary layer. However, at the times the SAR images shown in Figures 3-7 were acquired, the air-sea boundary layer was hardly ever neutrally stable but most of the time unstable. But it is permissible to use the CMOD4 wind scatterometer model also in this case because previous C band scatterometer measurements carried out from a sea-based platform in the North Sea have shown that the wind speed dependence of the NRCS does not change significantly when the air-sea interface changes from neutrally stable to unstable [Keller *et al.*, 1989].

An image intensity scan was performed through the roughness pattern in the Gulf of Gioia along the line inserted in the SAR image depicted in Figure 3. The scan direction is from the shore to the sea. The NRCS values were calculated by averaging in cross-scan line direction over 100 m (8 pixels) and by applying a moving average filter in scan line direction with a length of 50 m (4 pixels). In Figure 8 the NRCS values along this scan line are plotted (upper curve). On the left-hand vertical axis of the plot the NRCS is shown, and on the right-hand vertical axis the wind speed calculated from



Figure 6. ERS-1 SAR image acquired on August 24, 1993, at 2114 UTC showing a partly “concealed” roughness pattern in the Gulf of Gioia.



Figure 7. ERS-1 SAR image acquired on May 4, 1995, at 2113 UTC showing a weak roughness pattern in the Gulf of Gioia.

the NRCS by using the CMOD4 wind scatterometer model is shown. Note that the NRCS increases slightly with distance from shore until it reaches a maximum at a distance 15–20 km offshore. Then it decreases again and drops sharply at the front end of the pattern. The two vertical dash dot lines inserted in Figure 8 mark the shoreline (left line) and the front end of the roughness pattern (right line). We have defined the front end as the point where the wind speed drops below 2.5 m s^{-1} . In this case the roughness pattern has a length of 30.5 km.

We have also made scans through the other roughness patterns associated with katabatic winds which are visible on the SAR images shown in Figures 2, 4, 5, and 6. In all cases we found that the maximum wind speed inferred from the NRCS values by using the CMOD4 model lies between 3 and 4 m s^{-1} .

4. The Mesoscale Atmospheric Model

Katabatic wind fields that generate the roughness patterns visible on the SAR images have been simulated by using the three-dimensional version of the non-hydrostatic mesoscale model “Flow Over Irregular Terrain With Natural and Anthropogenic Heat Sources” (FITNAH) of the University of Hannover. Since there exist several publications in which the FITNAH model is described in detail [Gross, 1989, 1990, 1991], we give here only a brief description. The model is based on

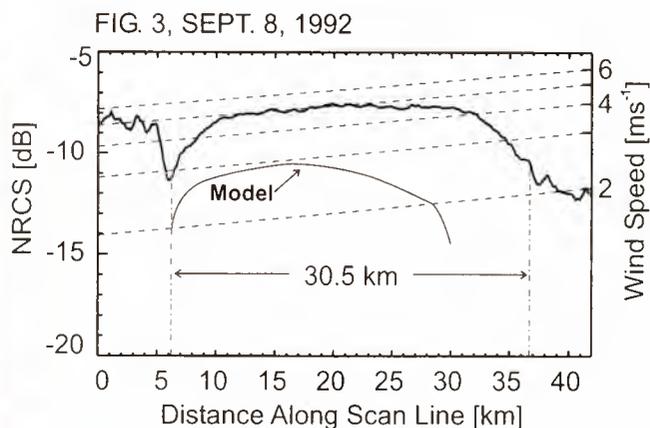


Figure 8. Variation of the surface wind speed along the scan line inserted into the ERS-1 SAR image of September 8, 1992 (Figure 3). The upper curve represents the wind speed inferred from the SAR image, and the lower curve represents the wind speed calculated from the atmospheric model described in section 5. On the left-hand vertical axis of the plot the normalized radar cross section (NRCS) is shown, and on the right-hand vertical axis the wind speed calculated from the NRCS by using the CMOD4 wind scatterometer model is shown. The left-hand vertical dash dot line marks the coastline, and the right-hand dash dot line marks the front of the roughness pattern generated by the katabatic wind.

the equation of motion using the Boussinesq approximation, the first law of thermodynamics, the conservation equation for water vapor, and the continuity equation in the anelastic form. The equations are Reynolds averaged. The resulting turbulent fluxes are parameterized by using the flux gradient relationships where the eddy exchange coefficients are calculated by using the turbulent kinetic energy determined by an additional prognostic equation. The divergence of the flux of the long wave radiation is parameterized by taking into account only the emission of water vapor and CO_2 . A terrain-following coordinate system with the new vertical coordinate $\eta = (z - h)/(H - h)$, where h is the ground elevation and H is the top of the model domain, is used to simplify the application of the lower boundary conditions. At the top of the model domain, which in our case is at 10,200 m above ground, the variables are unaffected by disturbances generated by the topography. At the surface a nonslip boundary condition for the velocity components is applied. In the model the surface humidity depends on the soil water content and on the degree of water vapor saturation in the lowest level of the atmosphere. The surface temperature is obtained by a surface energy budget in which the sensible, the latent, and the soil heat flux enters as well as the flux of short and long wave radiation. Over the sea the surface temperature is taken as constant over the whole integration time. The model equations are discretized using centered difference, a forward scheme is adopted for the time integration, and the advection is approximated by a selective vector upstream algorithm. The

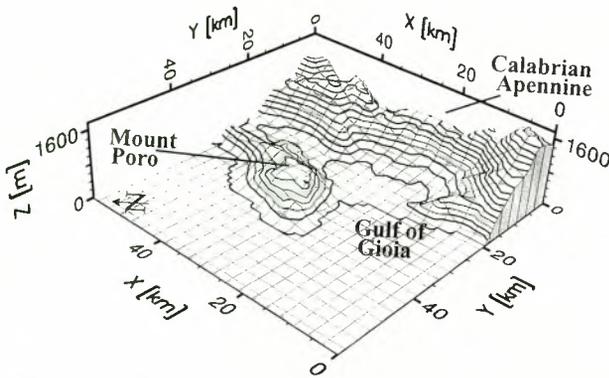


Figure 9. Three-dimensional perspective view of the model topography of the valley of Gioia used in the simulations. The spacing of the contour lines is 100 m. Note that in this plot the vertical scale is exaggerated.

equations are solved on a staggered grid, which is extended into the vertical direction in order to achieve a high spatial resolution near the surface. Horizontally, a constant mesh width of 1500 m is used. Below the ground, six grid levels are used for calculating the heat flow within the soil.

5. Simulations for the Event of September 8, 1992

In this section we present simulation results obtained with the FITNAH model for the SAR scene shown in Figure 3 (September 8, 1992). In the simulations we have used a digital elevation model of the valley of Gioia provided by the German Aerospace Research Establishment (DLR). A three-dimensional perspective view of the model topography of the valley of Gioia is shown in Figure 9. In this plot the contour lines have a spacing of 100 m and the horizontal grid lines a spacing of 3 km, which is twice the spacing used in the simulations. In the back the mountains of the Calabrian Apennine are visible, and in the front, Mount Poro is visible with a model height of 700 m. In Figure 9 the coordinate system is chosen such that a good perspective view of the topography is obtained. However, in the simulations we use another coordinate system: the x axis is parallel to the mountain ridge and the y axis is directed toward the sea (see Figure 10).

The horizontal grid spacing is 1.5 km, and 48 vertical levels are used in the atmosphere with a dense arrangement in the lower part in order to resolve the strong vertical variation of the meteorological variables close to the ground. With 40 grid points in the horizontal the model domain is 58.5 km in both horizontal directions. In order to suppress reflections we use open lateral boundary conditions and an absorbing layer in the upper part of the domain. For the synoptic situation chosen in this simulation the effects of lateral boundaries on the mountain wind are negligible. The cold air production area for the mountain wind is separated by the mountain peaks from the boundaries. From the

summit of the Calabrian Apennine the cold air moves into the valley of Gioia and also in the direction of the boundaries. Therefore no erroneous informations are advected from the boundaries into the model domain.

To run a mesoscale model, a broad spectrum of input parameters must be defined. However, only the numerical parameters can be specified exactly. The synoptic situation as well as the site parameters are usually unknown with the spatial density needed for numerical simulations. Therefore sparse observations and plausible assumptions must fill this gap. For the simulation presented here the synoptic situation is fixed with a vanishing geostrophic wind and a vertical profile for water vapor with 70% relative humidity up to a height of 1800 m and a linear decrease to zero up to 10,000 m. For the temperature a neutral, stratified atmosphere up to 1800 m is adopted with an increase of potential temperature according to the standard atmosphere above. The various site parameters have to be specified from grid point to grid point. Since this information for September 8, 1992, is not available, a standard input data set for grass coverage of the terrain is used (roughness length, 0.2 m; heat conductivity, $1.0 \text{ W m}^{-1} \text{ K}^{-1}$; field capacity, 0.05 m; thermal diffusivity, $4 \times 10^{-7} \text{ m}^2 \text{ s}^{-1}$). The roughness length over water is assumed to be 10^{-3} m .

Although the accuracy of the synoptic and the site input parameters is poor, the combination is sufficient to reproduce the observed temperature characteristics of the weather station at Lamezia Terme. The sea surface temperature is kept constant during the integration period with a value of 25°C . We started the simulations during a cloudfree and calm weather situation after sunset when a stable boundary layer had already developed over land. Above the boundary layer, there exists a residual layer which had developed during the previous day. At a height of 1800 m the stratification is stable again. It is assumed that the temperature difference between the water and the land was initially 2°C .

One output of the simulation is the wind field at a height of 15 m above ground 4 hours after sunset, which is approximately the SAR data acquisition time. At this time the daytime sea breeze should have vanished, and the nocturnal land breeze is not yet well established. Therefore the sea-land breeze only slightly affects the wind field investigated here. Note that the complex interaction of drainage flows and mountain winds is included in the model. In Figure 10a the resulting wind vectors are plotted together with the topography of the valley (spacing of the contour lines is 100 m), and in Figure 10b the corresponding isotachs are plotted together with the topography represented by grey levels.

The orography of the mountains and the valley as well as the shape of the coastline strongly affect the wind field. Mount Poro diverts the cold air flow in all directions, with maximum wind speed at the seaside and a convergence of the drainage winds in the small valley located to the east. The wind blowing through the basin-shaped valley of Gioia merges with the downslope winds from the mountains of the Calabrian Apennine

heading toward the eastern valley. When the cold air current reaches the coast, it spreads out over the Gulf of Gioia. Here the wind field is modified by the katabatic winds from Mount Poro and by the shape of the coastline. In the simulations a zone of convergence with wind speeds up to 2.5 m s^{-1} can be delineated in the Gulf of Gioia. When the air flows out of the valley onto the open sea, the flow diverges in direction since the channeling effect of Mont Poro vanishes, whereas the continuous mass flow from land remains. Therefore the structure of the horizontal wind field looks like a tongue or mushroom emanating from the valley of Gioia (Figure 3). At 4 hours after sunset the flow of cold air extends 30 km onto the sea. At the front end the wind speed drops down to zero quite sharply.

When the channeled flow reaches the sea, the cold air close to the warm water surface is heated. The cold air mass is modified with time because its temperature equilibrates with that of the sea surface. Because of the enhanced production of turbulence in the unstable surface layer, heat and momentum are mixed by enhanced turbulent fluxes in the lowest part of the atmosphere. This mixed layer grows with distance from shore. At the front end of the cold air current, there exists a mixed layer up to the height of the flow. The difference in density between the cold and surrounding air decreases faster over the sea than over land. Therefore the flow of cold air terminates faster over sea than over land.

We now want to compare the simulated horizontal sea surface wind field with the one inferred from the ERS-1 SAR image shown in Figure 3. Comparing the 1.5 m s^{-1} isoline in Figure 10b with the front line of the roughness pattern visible on the ERS-1 SAR image depicted in Figure 3, we see that the observed mushroom-like pattern is approximately reproduced if we take the 1.5 m s^{-1} isotach as the front line of the pattern. A wind speed of 1.5 m s^{-1} can be considered as a threshold wind speed for generating short-scale sea surface roughness. In order to compare wind speed calculated by the model with the wind speed inferred from the SAR image we have calculated the model wind speed along the same scan line as shown in Figure 3 and have inserted these values in the plot of Figure 8 (lower curve). Comparing the two wind speed curves plotted in Figure 8, we see that the length of the katabatic wind field is slightly underestimated by the atmospheric model. Furthermore, the wind speed calculated by this model is lower than the one inferred from the SAR image. This may be due to deficiencies of the CMOD4 model by which the NRCS values are converted into wind speeds. Indeed, there are indications that the CMOD4 model overestimates the wind speed when the wind speed is below 3 m s^{-1} [Rufenach, 1995]. On the other hand, we can also not expect that the FITNAH model reproduces exactly the SAR-derived wind speed because the input parameters for the simulations are not well known. The simulations should have given more accurate results if we had available data from simultaneous radiosondes ascents. On the whole we consider the agreement between the

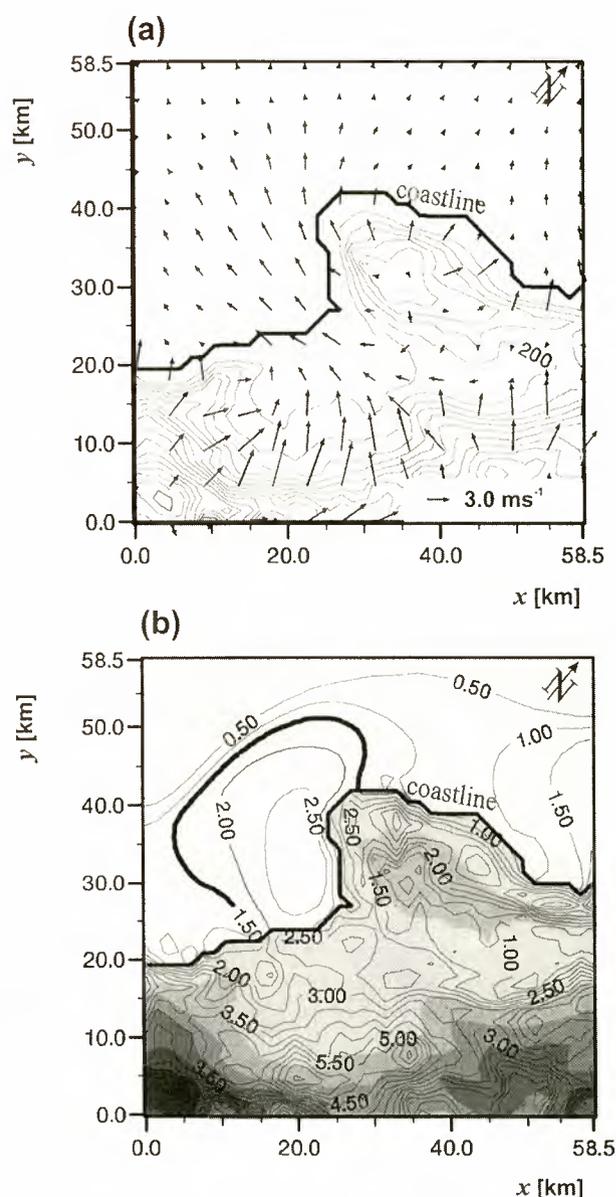


Figure 10. Wind field at a height of 15 m above the surface calculated by the atmospheric model for September 8, 1992, 4 hours after sunset. (a) The wind vectors are shown together with the topography of the valley, and (b) the isotachs are shown together with the topography (represented in grey shades). The numbers attached to the isotachs denote the wind speed in m s^{-1} . Note that in these plots the coordinate system is rotated by -130° relative to the one used in Figure 9.

measured and simulated wind fields to be fairly good, considering all the uncertainties involved.

6. Discussion

The mushroom-like sea surface roughness pattern observed on September 8, 1992, results only from one possible realization of the nocturnal cold air outflow through the valley of Gioia. (A very similar roughness pattern is visible on an ERS-1 SAR image acquired on

June 7, 1994, at 2112 UTC (orbit 15897, frame 765) over the same area.) As evidenced by the other four ERS-1 SAR images shown in Figures 4-7 and by many more ERS-1 SAR images of the same area not shown here, other roughness patterns can be formed, which have shapes and areal extents ranging from very elongated tongues with lengths of several tens of kilometers to narrow strips with widths of only few kilometers.

These very different roughness patterns generated by the sea surface winds are caused by differences in meteorological and environmental conditions. It is beyond the scope of the present paper to investigate systematically the dependence of the areal extent and shape of this roughness pattern on the various parameters. Such parameters are, among others, the time after sunset, the temperature difference between sea and land, the cloud coverage, the air humidity, the soil moisture, the vegetation coverage of the terrain, the height of the boundary layer, and the ambient wind field [McNider and Pielke, 1984; Gudicksen et al., 1992]. However, from the physics that enters into the model and from simulations that we have carried out with different meteorological input parameters [Pahl, 1996] we are able to discuss qualitatively how the meteorological conditions affect the cold air outflow and thus the roughness pattern on the sea surface.

The main factors that determine the areal extent and the shape of the pattern associated with the katabatic wind are speed and direction of the superimposed ambient wind field, the cloud coverage, and the temperature difference between sea and land. The nocturnal mountain wind from the valley will be strong in calm synoptic weather situations when there are no or only a few clouds present and when the difference in temperature between sea and land is small. A low cloud coverage accelerates the nighttime cooling of the air close to the ground and gives rise to a strong downslope flow of cold air. The stronger the surface cooling, the larger is the wind speed at the surface and the farther the cool air can advance onto the sea. The acceleration force of the outflow is proportional to the horizontal temperature gradient near the head of the current.

When the temperature difference between the sea surface and the land is large, then the cold air moving over the warm water surface gives rise to an unstable atmosphere near the sea surface in which strong turbulent mixing takes place that destroys the low-level wind maximum. Therefore, in the case of a large sea-land temperature difference, the horizontal extent of the cold air flow over the sea current will be smaller than in the case of a small sea-land temperature difference.

Local wind systems in mountainous regions, like katabatic winds, are thermally induced air circulations having small to moderate wind speeds. If a strong ambient wind is present, then the local wind systems can be completely concealed by the ambient wind and sea surface manifestations of the katabatic field cannot be delineated on the SAR image. In the case of a weak to moderate ambient wind the cold air outflow will be modified depending on its speed and direction. If the

ambient wind blows in the same direction as the katabatic wind, then the cold air current, driven by the horizontal pressure gradient at the head and intensified by a tail wind, will move far out onto the sea. However, if the ambient wind blows in the opposite direction, the katabatic wind field will be pushed back to the shore and can even be completely suppressed by the head wind.

Having the above mentioned mechanisms in mind, we are able to explain qualitatively the shapes of the roughness patterns visible on the SAR images shown in Figures 3-7.:

6.1. Event of September 8, 1992 (Figure 3)

This mushroom-like pattern develops in a cloudless and calm weather situation. The cool air shooting down the valley experiences an enhanced friction due to turbulence generated by the large temperature difference between the water and the air ($\sim 9^\circ\text{C}$). The front of the cold air jet comes to a stop ~ 10 km seaward behind the point where the northern coastline of the Gulf of Gioia abruptly bends northward. Thus, behind this point, the decelerated airflow can divert in all directions and thus form a mushroom-like pattern.

6.2. Event of February 20, 1995 (Figure 4)

The broad roughness pattern extending almost over the whole area of the Gulf of Gioia and reaching far south almost to the Strait of Messina is caused by the northwesterly wind deflecting the katabatic wind field into a southeasterly direction. The large area of the pattern is favored by the moderate temperature difference between the water and the air ($\sim 5^\circ\text{C}$). The fringed structure of the front of the roughness pattern is very likely generated by advancing lobes in the head of the gravity current, which are separated by deep clefts. Near the surface the gravity current is retarded by surface friction resulting in an overhanging nose. This leads to a highly unstable situation which gives rise to various instabilities in the head zone [Simpson, 1994].

6.3. Event of June 3, 1994 (Figure 5)

On this day the air temperature measured at Lamezia Terme at 2100 UTC was quite high (21.3°C), which implies that the air near the coast had not cooled off very much in the evening. Therefore we expect that on this day, less cold air was generated than on September 8, 1992. But a small amount of cold air was shooting down from the high mountains through the central valley onto the sea. Since the water-air temperature difference was quite small ($\sim 1^\circ\text{C}$), only a little turbulent mixing took place at the sea surface, and thus the cold air could move quite far out onto the sea. Therefore we attribute the generation of the long tongue-like pattern to the combined effect of a small volume of cold air and a reduced friction at the sea surface.

6.4. Event of August 24, 1993 (Figure 6)

The roughness pattern generated by the katabatic wind merges on the northwest side with the roughness

pattern generated by the ambient wind. Thus, on this side, the front of the katabatic wind field becomes (almost) undetectable on the SAR image. This may be the explanation why on several other SAR images of this area acquired in the evening (16 out of 36), no sea surface manifestations of this katabatic wind field could be delineated.

6.5. Event of May 4, 1995 (Figure 7)

On this day the roughness pattern is very weak and is confined to a small area adjacent to the northern coastline of the Gulf of Gioia. The main reason for this weak pattern is that the katabatic wind is attacked by a northwesterly head wind which attenuates and deflects the katabatic wind field. Note also the dark band at the front line of the katabatic wind field pattern, which we often have observed also on other SAR images showing sea surface manifestations of katabatic wind fields. At this border line between the two wind systems the horizontal wind speed is zero, and thus here the wind cannot roughen the sea surface. Furthermore, this is a line of maximum horizontal convergence, where sea surface slicks can accumulate, which reduce the sea surface roughness.

7. Conclusions

The present investigation has shown that SAR imagery of coastal waters can yield valuable information on katabatic winds in coastal mountainous regions provided that some information on the synoptic weather situation is available. From such SAR images the areal extent and the shape of the katabatic wind field extending from land over the sea can be obtained. Furthermore, the wind speed at the sea surface can be determined from the SAR image intensity when the wind direction is known. Such detailed information on the structure of the wind field over coastal waters is difficult to obtain by other means. SAR images acquired over coastal waters can also be used to validate models describing katabatic wind fields or, more generally, local wind fields that extend from the coast onto the sea.

The knowledge of local wind fields in coastal regions is of great importance for predicting the advection and dispersion of pollutants discharged into the atmosphere and the sea at the coast.

Acknowledgments. We thank A. Michelsen for preparing the figures, the European Space Agency for supplying the ERS-1 SAR images, the Deutscher Wetterdienst (Seewetteramt Hamburg) for supplying the meteorological data, and the DLR for supplying the digital terrain data. This investigation was supported by the German Space Agency (Dara) under contract 50 EE 9413.

References

- Agee, E.M., Mesoscale cellular convection over the oceans, *Dyn. Atmos. Oceans*, 10, 317-341, 1987.
- Alpers, W., Measurement of mesoscale oceanic and atmospheric phenomena by ERS-1 SAR, *Radio Sci. Bull.* 275, pp.14-22, Int. Union of Radio Sci., c/o University of Gent (INTEC), Gent, Belgium, 1995.
- Alpers, W., and B. Brümmer, Atmospheric boundary layer rolls observed by the synthetic aperture radar aboard the ERS-1 satellite, *J. Geophys. Res.*, 99, 12613-12621, 1994.
- Alpers, W., and G. Stilke, Observation of a nonlinear wave disturbance in the marine atmosphere by the synthetic aperture radar aboard the ERS-1 satellite, *J. Geophys. Res.*, 101, 6513-6525, 1996.
- Atkinson, B.W., and J.W. Zhang, Mesoscale shallow convection in the atmosphere, *Rev. Geophys.*, 34, 403-431, 1996.
- Atlas, D., Footprints of storms on the sea: A view from spaceborne synthetic aperture radar, *J. Geophys. Res.*, 99, 7961-7969, 1994.
- Atlas, D., T. Iguchi, and H.F. Pierce, Storm induced wind patterns on the sea from spaceborne synthetic aperture radar, *Bull. Am. Meteorol. Soc.*, 76, 1585-1592, 1995.
- Etlng, D., and R.A. Brown, Roll vortices in the planetary boundary layer: A review, *Boundary Layer Meteorol.*, 65, 215-248, 1993.
- Gross, G., Numerical simulation of the nocturnal flow systems in the Freiburg area for different topographies, *Contrib. Atmos. Phys.*, 62, 57-72, 1989.
- Gross, G., On the wind field in the Loisach Valley - Numerical simulation and comparison with the LOWEX III data, *Meteorol. Atmos. Phys.*, 42, 231-247, 1990.
- Gross, G., Anwendungsmöglichkeiten mesoskaliger Simulationsmodelle dargestellt am Beispiel Darmstadt, I, Wind- und Temperaturfelder, *Meteorol. Rundsch.*, 43, 97-112, 1991.
- Gudiksen, P.H., J.M. Leone, C.W. King, D. Ruffieux, and W.D. Neff, Measurements and modeling of the effects of ambient meteorology on nocturnal drainage flows, *J. Appl. Meteorol.*, 31, 1023-1032, 1992.
- Johannessen, J.A., R.A. Shuchman, O.M. Johannessen, K.L. Davidson, and D.R. Lyzenga, Synthetic aperture radar imaging of upper ocean circulation features and wind fronts, *J. Geophys. Res.*, 96, 10411-10422, 1991.
- Keller, W.C., V. Wisman, and W. Alpers, Tower-based measurements of the ocean C band radar backscattering cross section, *J. Geophys. Res.*, 94, 924-930, 1989.
- Laur, H., P. Bally, P. Meadows, J. Sanchez, B. Schaeffler, and E. Lopinto, ERS SAR calibration: Derivation of the backscatter coefficient σ^0 in ESA ERS SAR PRI products, *Document No: ES-TN-RS-PM-HL09*, Eur. Space Agency, Eur. Space Res. Inst. (ESA/ESRIN), Frascati, Italy, June 1996.
- Mahrt, L., and S. Larsen, Small scale drainage front, *Tellus*, 34, 579-587, 1982.
- McNider, R.T., and R.A. Pielke, Numerical simulation of slope and mountain flows, *J. Appl. Meteorol.*, 23, 1441-1453, 1984.
- Mitnik, L.M., Mesoscale coherent structures in the surface wind field during cold air outbreaks over the Far Eastern seas from satellite side looking radar, *La Mer*, 30, 287-296, 1992.
- Mitnik, L.M., and S.V. Viktorov (Eds.), (in Russian), *Radar Sensing of the Earth's Surface From Space*, 200 pp., Gidrometeoizdat, St. Petersburg, Russia, 1990.
- Pahl, U., Numerische Simulationen nächtlicher Kaltluftabflüsse zur Interpretation von SAR Aufnahmen, diploma thesis, Inst. of Meteorol. and Climatol., Univ. of Hannover, Germany, 1996.
- Rufenach, C.L., A new relationship between radar cross-section and ocean surface windspeed using ERS-1 scatterometer and buoy measurements, *Int. J. Remote Sens.*, 18, 3629-3647, 1995.

- Rufenach, C.L., R.A. Shuchman, N.P. Malinas, and J.A. Johannessen, Ocean wave spectral distortion in airborne synthetic aperture radar imagery during the Norwegian Continental Shelf Experiment of 1988, *J. Geophys. Res.*, *96*, 10453-10466, 1991.
- Simpson, J. E., *Sea Breeze and Local Winds*, 234 pp., Cambridge Univ. Press, New York, 1994.
- Stoffelen, A., and D. Anderson, Scatterometer data interpretation: Estimation and validation of the transfer function CMOD4, *J. Geophys. Res.*, *102*, 5767-5780, 1997.
- Thompson, T.W., W.T. Liu, and D.E. Weissman, Synthetic aperture radar observation of ocean roughness from rolls in an unstable marine boundary layer, *Geophys. Res. Lett.*, *10*, 1172-1175, 1983.
- Vachon, T.W., O.M. Johannessen, and J.A. Johannessen, An ERS-1 synthetic aperture radar image of atmospheric lee waves, *J. Geophys. Res.*, *99*, 22483-22490, 1994.
- Valenzuela, G.R., Theories for the interaction of electromagnetic and oceanic waves - A review, *Boundary Layer Meteorol.*, *13*, 61-85, 1978.
-
- W. Alpers, Institut für Meereskunde, Universität Hamburg, Troplowitzstr. 7, D-22529 Hamburg, Germany. (e-mail: alpers@ifm.uni-hamburg.de)
- G. Gross and U. Pahl, Institut für Meteorologie und Klimatologie, Universität Hannover, Herrenhäuserstr. 12, D-30419 Hannover, Germany. (e-mail: gross@muk.uni-hannover.de; pahl@muk.uni-hannover.de)

(Received February 4, 1997; revised May 29, 1997; accepted June 16, 1997.)

Assimilation of ERS SAR wave spectra in an operational wave model

Lars-Anders Breivik, Magnar Reistad, Harald Schyberg, and Jens Sunde

Norwegian Meteorological Institute, Oslo, Norway

Harald E. Krogstad

Department of Mathematical Sciences, Norwegian University for Science and Technology, Trondheim, Norway

Harald Johnsen

NORUT Information Technology, Tromsø, Norway

Abstract. Inverted wave spectra from ERS-1 synthetic aperture radar (SAR) image and wave mode products have been assimilated in the operational wave model for the ocean wave forecast service at the Norwegian Meteorological Institute. Elements of the operational system are explained briefly, and the impact of including the SAR wave data in the operational wave model runs is shown both for individual cases and as overall statistics. Although individual cases clearly show that the satellite observations are able to influence the forecast in a generally positive way, the average improvement is minor for the areas covered by the wave model. Reasons for this are the intermittency of the data, on the average small differences between inverted SAR and model first-guess wave spectra, and to some extent, limitations in the analysis method.

1. Introduction

This study examines the value of inverted synthetic aperture radar (SAR) wave spectra in the operational wave forecasting service for the North Sea and Norwegian Sea carried out by the Norwegian Meteorological Institute (DNMI). In order to use the data operationally, the information needs to be obtained in near-real time and interpolated in space and time from a very limited number of observation points to the larger area of interest. This is achieved by assimilating the data in the operational wave forecasting model. The model wave field is thus updated to produce a new and improved wave field as the initial condition for the next model forecast.

Traditionally, operational wave forecast models are driven by wind input from weather forecast models without taking any wave observations into account. The wind input is continuously updated in the weather forecast model analysis/assimilation cycle, and the waves develop according to the appropriate wind-wave relations. However, in later years an increasing amount of wave observations have become available from satellites, and work has been carried out to utilize this information for improving the wave model results. There have been several attempts to assimilate altimeter wave heights [e.g., Thomas, 1988; Esteva, 1988; Janssen *et al.*, 1989; Lionello *et al.*, 1992; Gunther *et al.*, 1993; Breivik and Reistad, 1994]. These are sequential methods well suited for operational use, and presently, altimeter wave heights from the European Remote Sensing Satellite (ERS) are operationally assimilated at several meteorological centers, including DNMI. Advanced methods, including the wave model dynamics in the assimilation scheme for a consistent analysis of wave field and wind input, have also

been developed [e.g., de Valk and Calhoun, 1989; de las Heras *et al.*, 1994; Bauer *et al.*, 1996]. These more sophisticated techniques are computationally more expensive and so far not commonly in operational use at the wave forecast centers. A further development in the assimilation of satellite wave data is to utilize the full spectral information from the synthetic aperture radar (SAR). A method for assimilation of two-dimensional wave spectra by dividing the full spectrum into separate partitions suitable for optimal interpolation has been developed by Hasselman *et al.* [1994, 1996] and demonstrated in an operational framework by Voorrips *et al.* [1997]. In the present study we apply a simpler sequential method for assimilation of full two-dimensional SAR wave spectra.

Unfortunately, the currently inverted SAR wave spectra are not completely independent from the wave model spectra since the SAR inversion algorithm depends on a priori information from the wave model itself. Previous studies [Breivik *et al.*, 1995] have shown that there is nearly full correlation between a priori and inverted SAR wave spectrum parameters in the wind sea part. This is actually to be expected since the SAR information is confined to frequencies below ~ 0.15 Hz. However, the SAR spectra usually cover the swell part and therefore have a potential for improving the model results by data assimilation. Very high average impact cannot be expected in the present case since the model is known to produce reliable results. However, in special situations, e.g., with poor wind history and correspondingly poorly defined swell, more significant improvement might be expected.

In order to assess the impact of the SAR data an additional operational wave model routine including assimilation of the inverted SAR spectra was run parallel to the regular operational wave model forecasts for 4 months during the fall and winter of 1995–1996. The results from these runs have been compared to independent measurements from North Sea platforms and the ERS-1/2 altimeter. The evaluation of the fore-

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02728.
0148-0227/98/97JC-02728\$09.00

casts demonstrates only a small positive impact from the SAR data. In broad terms, the observations are, so far, too spread out and intermittent.

Section 2.1 contains a brief description of the SAR data and the inversion procedure. The operational wave model Wave Incidents on Norwegian Coast: Hindcasting (WINCH) is described in section 2.2, and comparisons between inverted and a priori spectral are described in section 2.3. The assimilation method is described in section 3, and three specific case studies of how the impact of the assimilation affects the wave development are shown in section 4. The results from the parallel runs are discussed in section 5, and finally, the conclusions are found in section 6.

2. Data and Algorithms

2.1. SAR Observations

DNMI receives the fast delivery active microwave instrument (AMI) wave mode product, denoted UWA by the European Space Agency (ESA), in near-real time on the Global Telecommunication System (GTS). The product is briefly described by *Vass and Battrick* [1992] and also by *Krogstad et al.* [1994] and *Krogstad* [1994]. A second wave product has been received on request in near-real time from Tromsø Satellite Station (TSS) and consists of spectra of the high-resolution SAR images processed at TSS.

2.1.1. ERS AMI wave mode product. The wave mode product is a SAR image spectrum obtained from a three-look, amplitude-averaged, 5×8 km imagette recorded every 200 km along the track. The resolution of the imagette is similar to the regular full-resolution images (20×16 m). Only every second recording is distributed daily on GTS. The spectrum consists of spectral values given in 12 directional and 12 wavenumber bins ranging approximately from 100 to 1000 m in wavelength and covering a 180° directional sector. The ERS wave mode spectra used for the present study suffer from a nonoptimal processing which complicates its use. The ocean-SAR inversion theory [see, e.g., *Hasselmann and Hasselmann*, 1991] has been developed for the spectrum of the relative wave modulation in the SAR image. However, the received spectrum is essentially a linear combination of the wave modulation spectrum and the speckle or clutter spectrum with unknown scaling and relative weight between the two. Moreover, the imagettes are look summed in amplitude instead of in intensity as required by the theory. The speckle spectrum has been determined by analysis of a separate set of imagettes, and the scaling procedure is described by *Krogstad* [1994].

The operational system includes an automatic data control for the UWA spectra which is run prior to the inversions in order to remove spectra with excessive low wavenumber noise and no definite wave features (e.g., scenes from near shore, from ice-infected areas, etc.). A limited intercomparison between UWA spectra and spectra obtained from the same imagette indicates that the low wavenumber noise may basically be due to the UWA processing. ESA is now implementing new algorithms which will not suffer from these shortcomings [Brooker, 1996]. As a fully automated control, the rejection criteria have been set rather tight, and this leads to $\sim 50\%$ of the spectra being rejected. In September, October, and November 1995 a total of 217 ERS-2 wave spectra and 583 ERS-1 wave spectra passed the data control. This gives ~ 9 observations per day or 4.5 observations per assimilation cycle. A typical coverage for a 2 week period is illustrated in Figure 1.

2.1.2. TSS ERS image mode spectra. In image mode the ERS SAR and the associated ground processing provide 100×100 km images with the same resolution as the imagettes. The TSS production is in accordance with the ESA UI16 product [*Vass and Battrick*, 1992]. For the present application, four spectra are computed from subimages within one scene. Each subimage covers an area of $\sim 18 \times 18$ km and is transferred to DNMI over the Internet. The spectra are computed from the relative modulation intensity using squared pixel values and the averaged periodogram method after a parabolic detrending. This yields spectral values with about 5% uncertainty. The spectra are encoded using 24 wavelength bins between 37 and 845 m and 18 directional bins between 0° and 180° .

In the period from September 15 to December 31, 1995, DNMI received over 1000 wave spectra from TSS, computed from a total of 315 images. Although representing a considerable amount of SAR data, this results only in about three independent SAR wave spectra per day.

2.1.3. Ocean-SAR inversion procedure. The procedure is based on a quasilinear (QL) version of the Hasselmann forward ocean-SAR spectral transform expressing the SAR image modulation spectrum $S_{\text{ql}}(\underline{k})$ as

$$S_{\text{ql}}(\underline{k}) = \frac{1}{2}A(k)[|T_{\text{lin}}(\underline{k})|^2\Psi(k) + |T_{\text{lin}}(-\underline{k})|^2\Psi(-k)] \quad (1)$$

Here Ψ is the wave spectrum, \underline{k} is the wavenumber, T_{lin} is the linear ocean-SAR transfer function, and A is an optimized azimuth cut-off function dependent on the spectrum (see *Krogstad et al.* [1994] for a derivation of A based on the series expansion of the fully nonlinear transform). The QL expression is reasonable around the most energetic part of the SAR spectrum but fails to spread sufficient spectral energy to higher wavenumbers. Since the inversion system was originally designed for UWA spectra, the QL transform was considered sufficient because of the UWA product's 100 m cutoff and relatively crude resolution. Standard forms of the tilt and hydrodynamic transfer functions have been applied, although it is realized that this may well be the main error source, as it would still be for the full nonlinear transform [see, e.g., *Schmidt et al.*, 1995; *Hara and Plant*, 1994].

The ocean-SAR inversion amounts to find an ocean wave spectrum such that when the forward ocean-SAR transform is applied, the result conforms to the observed SAR spectrum. Following *Hasselmann and Hasselmann* [1991], we minimize the penalty functional

$$J(\psi) = \int |\psi(\underline{k}) - \psi_o(\underline{k})|^2 W_\psi(\underline{k}) d\mathbf{k}^2 + \int |T[\psi(\underline{k})] - S_{\text{mes}}(\underline{k})|^2 W_s(\underline{k}) d\mathbf{k}^2 \quad (2)$$

subject to the nonnegativity constraint $\psi \geq 0$, where ψ_o is a best guess a priori spectrum, $T[\psi(\underline{k})]$ is the forward mapping, S_{mes} is the measured SAR spectrum, and W is the nonnegative weight functions. We refer to *Krogstad et al.* [1994] for details as to how the weight functions are selected and the iterative solution is carried out.

The derived wave spectra are to some extent dependent on the inversion strategy. In the present case the a priori wave information is a wave spectrum taken from a neighboring grid point of the wave model. The 180° directional ambiguity in the

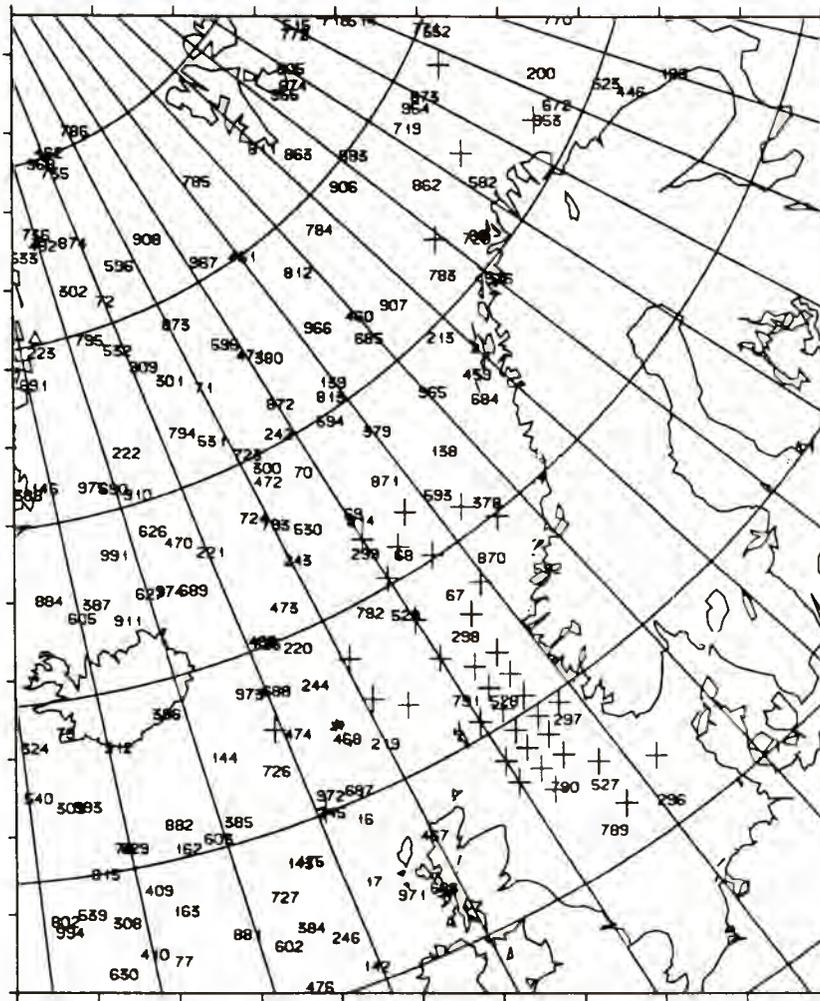


Figure 1. Locations of ERS-1 wave mode spectra over the Norwegian Sea, November 8–22, 1993, shown with numbers. Crosses are grid points of the wave model where a priori spectra are available for the inversion.

SAR image spectrum is removed by reading probable directions from the a priori spectrum and selecting the weight functions accordingly. Outside the azimuth pass band and for high-range wavenumbers in the SAR spectrum, there is no wave-related information, and it is therefore reasonable to keep the a priori wave spectrum unchanged. Inside this domain it is possible to give the a priori spectrum some weight away from the peaks in the SAR spectrum, but this will sometimes lead to dubious double-peaked spectra having maxima corresponding to the peaks both in the a priori wave spectrum and the SAR spectrum. It is therefore better to let the SAR spectrum determine the wave spectrum completely within the domain. The inversion strategy is denoted “blind inversion” and is the one used operationally here. In blind inversion the spectral level in the inverted spectra is completely determined by the SAR image spectrum for domains of the wavenumber plane where there is SAR spectral energy. Further examples are shown by Sunde *et al.* [1994].

One weakness in the inversion algorithm is that the first-guess model spectra are only modified within a 180° sector around a main direction in the a priori spectrum. This means, for example, that unrealistic swell outside this sector in the model spectrum will not be removed in the inversion, and examples of this happening have actually been observed [Breivik *et al.*, 1995]. Improved procedures are given by Hasselmann

et al. [1996], but the best way to avoid the directional ambiguity is to apply complex cross spectra in the inversion [Engen and Johnsen, 1995].

The wave model spectra are defined for the range from 0.04–0.3 Hz, whereas the inverted AMI wave mode spectra only cover 0.04–0.125 Hz (in deep water). A final merging of the inverted and a priori spectra has therefore to be carried out. The inverted spectrum is first converted to the wave model format and extended above 0.125 Hz using a f^{-4} frequency spectrum and a constant directional distribution equal to the distribution at 0.125 Hz. In the final spectrum the inverted spectrum is used below 0.125 Hz, the a priori wave model spectrum is used above 0.153 Hz, and a gradual transition is used in between. The TSS spectra cover a larger frequency range (0.04–0.2 Hz), and the upper limit for the crossover between the inverted and the input wave model spectrum is at 0.2 Hz. As discussed by Hasselmann *et al.* [1996], this is also a rather simplistic approach that could be improved.

The operational SAR to wave spectrum inversion is run prior to the assimilation. All wave mode and TSS spectra where the distance in space and time to the nearest wave model spectrum is less than specified limits are used, and a previous wave model forecast (2–12 hours) is used as a priori information.

Table 1. Statistics for ERS-1 Wave Mode, September 1 to November 30, 1995

ERS-1	Total Sea	Wind Sea	Swell
Number of entries	583	583	583
SAR, ^a m	3.28	1.19	2.89
Bias (SAR – model), ^b m	0.00	–0.07	0.07
Standard deviation, ^b m	0.49	0.22	0.48
Correlation coefficient, ^b	0.94	0.99	0.91
SAR, ^c s	9.98	4.25	10.29
Bias (SAR – model), ^d s	0.37	–0.10	0.50
Standard deviation, ^d s	1.48	0.34	1.51
Correlation coefficient, ^d	0.72	1.00	0.66

^aMean significant wave height (SWH).^bSWH.^cMean peak period (PP).^dPP.

2.2. Wave Model

In the present study the wave model is used in different contexts. First, wave model spectra are used as a priori information for the SAR inversion. Then the wave model spectra are compared to inverted spectra to decide whether the SAR has picked up new wave information. Finally, the inverted SAR wave spectra are assimilated into the wave model.

The WINCH wave model is a second generation, deep water, spectrally discrete model developed by Oceanweather Inc. It is a modified version of the Sea-Air Interaction Laboratory (SAIL) model, which was one of the models used in the Sea Wave Modeling Project (SWAMP) study [Greenwood *et al.*, 1985]. WINCH runs on a rectangular grid with a fine mesh (~75 km) covering the North Sea, the Norwegian Sea, and the Barents Sea and a coarse mesh (~150 km) covering the western parts of the North Atlantic. The time steps in the model are 2 hours for the propagation and 1 hour for the growth term. The wave spectrum is divided into 24 directional bands with 15° bandwidth and 15 frequency bands ranging from 0.04 to 0.30 Hz. Each spectral bin is propagated individually with the energy input proportional to the existing energy multiplied by a function of the friction velocity and the angle between the wind and wave directions. The wind input is obtained from the operational weather forecast High-Resolution Limited Area Model (HIRLAM) [Gustafsson, 1993]. Nonlinear interactions and dissipation are achieved by using a reference spectrum and directional relaxation. Spectral energy below the wind-sustained frequencies is treated as swell, and these bins are subject to advection and geometric damping only. The WINCH model has been run operationally at DNMI since 1985 [Breivik and Reistad, 1994]. DNMI has also implemented and tested a version of the third-generation wave model (WAM) [WAM Development and Implementation (WAMDI) Group, 1988], but it has turned out to provide little evidence of improved performance for the operational wave forecasting. Thus, for the present study, we believe that the choice of wave model is secondary for the impact of the SAR assimilation.

2.3. Comparisons Between Inverted and a Priori Spectra

For the period from September 1 to November 30, 1995, inverted wave spectra based on AMI wave mode both from ERS-1 and ERS-2 have been compared with collocated wave spectra from WINCH. A similar exercise was carried out using TSS-processed ERS-1 and ERS-2 AMI image mode SAR spectra from September 15 to November 30. Because of hard-

ware constraints, full-resolution SAR images and imagerettes cannot be recorded simultaneously by the satellite [Battrick, 1993].

The statistics are presented separately for total sea, wind sea, and swell. The wave parameters presented in Tables 1 and 2 are significant wave height (SWH, in meters) and peak period (PP, in seconds). The scatterplot for the swell part shown in Figure 2 also includes the peak direction (PD). We notice that the results for ERS-2 wave mode are very similar to ERS-1. Since the SAR inversion only affects frequencies below ~0.13 Hz, the correlation between inverted and a priori spectra is quite high in the wind sea part. This is the case both for SWH and peak period. Even the correlation for swell SWH is above 0.9. There is a small negative mean difference both in wind sea and swell for ERS-2, while the mean difference is 0.07 m for ERS-1 swell. The peak period is longer for ERS swell compared to WINCH, which is expected, since operational experience with the WINCH model indicates that the model tends to underestimate the swell period. A scatter diagram for the peak direction, which is the direction giving maximum energy integrated over frequency, is also shown in Figure 2. Somewhat surprisingly, the scatter diagram shows little change, and there seems to be little additional directional information in the SAR data not already available in the a priori spectrum.

A similar analysis for spectra from TSS is shown in Table 3 and Figure 3 (for the swell part). Again, there seems to be little new directional information gained from the SAR spectra. As for ERS wave mode, nearly all change is in swell SWH and period. From Table 3 it is seen that inverted TSS SAR wave data, similar to wave mode, gives higher values for the swell period. The TSS SAR wave heights are also high compared to WINCH. This difference of behavior between the inverted TSS and wave mode spectra is difficult to explain since the actual inversion system is the same, but different spectral resolution and somewhat different spectral estimations may be part of the reason.

We might have expected some more wind sea information in the TSS data compared to wave mode, but on average the wind sea contribution seems to be small, and the corresponding difference between the wave mode and TSS is of relatively little importance. In addition, the two data sets are sampled at different times and locations, which complicates the interpretation.

The bias in the wave height can most probably be explained

Table 2. Statistics for ERS-2 Wave Mode, September 1 to November 30, 1995

ERS-2	Total Sea	Wind Sea	Swell
Number of entries	217	217	217
SAR, ^a m	3.42	1.07	3.09
Bias (SAR – model), ^b m	–0.06	–0.08	–0.01
Standard deviation, ^b m	0.53	0.22	0.52
Correlation coefficient, ^b	0.93	0.99	0.92
SAR, ^c s	10.14	3.99	10.40
Bias (SAR – model), ^d s	0.34	–0.05	0.47
Standard deviation, ^d s	1.47	0.28	1.55
Correlation coefficient, ^d	0.71	1.00	0.67

^aMean SWH.^bSWH.^cMean PP.^dPP.

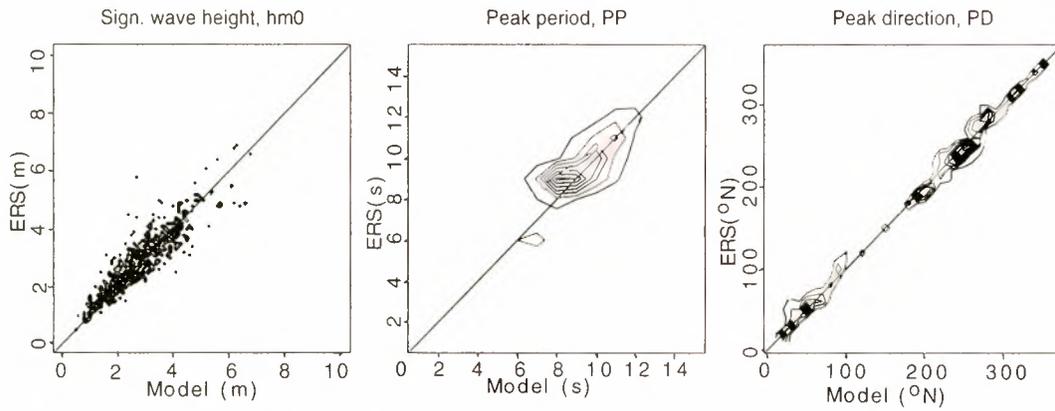


Figure 2. Scatterplots of inverted wave mode spectra; swell part against Wave Incidents on Norwegian Coast: Hindcasting (WINCH) model data.

by insufficient knowledge of the modulation transfer function (MTF). The hydrodynamic part of the real aperture radar (RAR) MTF shows a significant dependency of the wind speed for the wavelength range, radar frequency, and polarization [Schmidt *et al.*, 1995; Hara and Plant, 1994]. For example, the modulus of the RAR MTF, after subtraction of the range and tilt contribution, decreases with a factor of 2 over a wind speed region from 4 to 16 m/s for a C band VV polarization radar [Schmidt *et al.*, 1995]. However, although showing the correct trends, the available theoretical expressions of the RAR MTF cannot so far fully explain what is observed.

3. Analysis

The assimilation scheme is based on a modification of the successive correction method proposed by Bratseth [1986]. The scheme was developed for three-dimensional multivariate analysis of the atmosphere. The basic idea is similar to statistical interpolation [Hollingsworth, 1987], which is commonly used for operational data analysis in atmospheric models at a number of meteorological centers, including DNMI. A simplified version of the scheme was developed for the assimilation of altimeter wave heights [Breivik and Reistad, 1994]. The wave data assimilation requires the complete wave model spectrum to be updated. Since wave height depends only on the total wave energy, some preassumptions about the shape of the wave spectrum have to be made, and the wave height assimilation

therefore works as a scaling of the wave energy while preserving the ratio between wind sea and swell. With the inverted SAR spectra, full-directional wave spectra are available, and the assimilation may be carried out without pre-defined assumptions about the spectral shape.

The analysis starts with a first-guess field of wave spectra on the model grid derived from a wave model short forecast. This field is then corrected by the available observations. The method is based on the modified successive correction method [Bratseth, 1986], which uses two iterative equations where wave energy values for iteration number $v + 1$ are derived from iteration number v according to

$$E_x^A(v+1) = E_x^A(v) + \sum_{j=1}^N a_{xj} [E_j^O - E_j^A(v)] \quad (3)$$

$$E_i^A(v+1) = E_i^A(v) + \sum_{j=1}^N a_{ij} [E_j^O - E_j^A(v)] \quad (4)$$

Here the successive corrections are performed both for model grid points (with subscript x) and observation points with subscripts i and j . The wave energy $E = E(f, \theta)$ is as a function of frequency and direction. Superscripts O and A refer to observed and analyzed values, and N is the number of observations. The iteration is initiated by a spatial interpolation of the first-guess value to the observation points; that is, $E_j^A(v=0)$ is E_x (first guess) interpolated.

The analysis weights, a_{ij} and a_{xj} , are functions of model and observation error covariances and are specified for each frequency, f ,

$$a(f)_{ij} = [m(f)_{ij} + d(f)_{ij}] / M(f)_j \quad (5)$$

$$a(f)_{xj} = m(f)_{xj} / M(f)_j \quad (6)$$

where $m(f)_{ij}$ is the element of the first-guess error covariance matrix and $d(f)_{ij}$ is of the observation error covariance matrix. $M(f)_j$ is calculated as a function of $m(f)_{ij}$ and $d(f)_{ij}$ in a way that makes (4) converge. In (3) and (4) it is important to note the inclusion of the iteration variable E_j^A defined at the observation points, which distinguishes this modified successive correction scheme from earlier successive correction schemes. This variable converges to the observed values for large v , and in this modified scheme the analysis E_x^A converges

Table 3. Statistics for TSS ERS-1/2 Wave Spectra, September 11 to November 27, 1995

TSS ERS-1/2	Total Sea	Wind Sea	Swell
Number of entries	715	707	715
SAR, ^a m	3.96	1.59	3.40
Bias (SAR - model), ^b m	0.37	-0.04	0.48
Standard deviation, ^b m	0.53	0.12	0.57
Correlation coefficient, ^b	0.93	1.00	0.88
SAR, ^c s	10.27	5.08	10.74
Bias (SAR - model), ^d s	0.59	-0.05	0.78
Standard deviation, ^d s	1.62	0.21	1.65
Correlation coefficient, ^d	0.58	1.00	0.58

TSS is Tromsø Satellite Station.

^aMean SWH.

^bSWH.

^cMean PP.

^dPP.

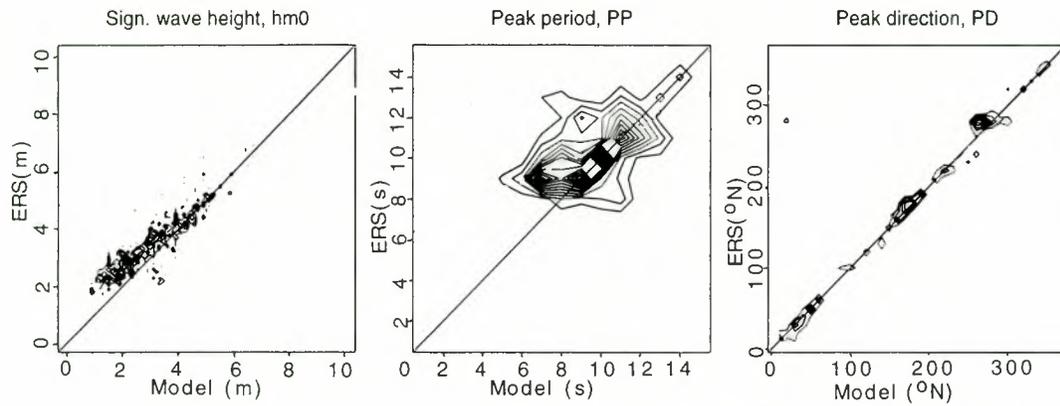


Figure 3. Scatterplots of inverted Tromsø Satellite Station (TSS) spectra, swell part against WINCH model data.

(for large v) toward the result of statistical interpolation [Bratseth, 1986].

The iteration ((3) and (4)) is performed on each (f, θ) -bin independently, which is equivalent to assuming that the first-guess errors in different (f, θ) -bins are uncorrelated. This is not correct since energy in different bins is linked by nonlinear interaction within the wave systems and by the wind forcing. However, a proper description of the error correlations is complicated. A statistical model of the first-guess error correlations could be obtained by analysis of a large number of wave models and accurate observation spectra. An ideal model for the error correlations would be situation dependent since energy in different parts of the spectra would originate from different wave systems.

An alternative way of performing the analysis would be to divide the spectrum into a small number of distinct partitions characterized by bulk wave parameters like the total energy, mean frequency, and direction and then perform the analysis on each of the partitions. This is equivalent to assuming that the model errors of different partitions are uncorrelated, which would be justified if one could assume that the partitions represented separate wave systems originating from independent meteorological systems. The method is described by Hasselmann *et al.* [1994, 1997] and implemented and tested on several cases with spectral buoy observations by Voorrips *et al.* [1997]. Another aim of the method is to drastically reduce the number of variables to be analyzed. This is important when statistical interpolation, which is computer intensive, is used for the analyses but less important for the successive correction method used here. A main problem with this way of modeling the spectral correlations is how to assign each partition of the observed spectrum to the correct partition of the first-guess spectrum. Some ad hoc method, which cannot be strictly justified, has to be chosen to define this assignment. Such assignment methods are prone to errors when there are large differences between observations and the first guess.

In the absence of any well-justified general model for spectral component error correlations, we chose to keep the simplicity of independent analysis of each bin, although this is not an optimal choice. Evaluating the analyzed spectra in numerous cases has shown that this works well in terms of reproducing realistic spectral shapes. Two examples are given in Figure 6 showing model first guess, SAR image mode spectra, and the analyzed result for an observation point.

The analysis method with individual correction of energy in

each bin within a certain influence radius might in some cases lead to negative energy. In these cases the energy in the actual bin is set to zero, and the iteration continues.

Since there is no spatial overlap between the SAR observations, we assume that the errors in different observation points are uncorrelated, i.e., that $d(f)_{ij} = 0$ for $i \neq j$. In the observation error covariance matrix we then only need the diagonal elements, the observation variance, and $d(f)_{ii} = [\sigma(f)^o]^2$, where $\sigma(f)^o$ is the standard deviation of the observation errors, assumed constant for all SAR observations.

We assume a Gaussian form of the first-guess error spatial correlation

$$m(f)_{ij} = \sigma(f)^{o^2} \exp \left[-\frac{r_{ij}^2}{2b(f)^2} \right] \quad (7)$$

where $\sigma(f)^o$ is the standard deviation of the model errors; r_{ij} is the distance between observation points i and j ; and $b(f)$ is the radius of influence, a typical distance for which the model errors become decorrelated. In order to determine the size of the influence radius $b(f)$, experience from the assimilation of altimeter wave heights has been used [Breivik and Reistad, 1994]. From collocated WINCH model first guess and altimeter significant wave heights from January to April 1994 the first-guess error covariance function for wave heights was derived empirically. The influence radius in a Gaussian covariance function similar to (7) was found to be around 200–250 km. Significant wave height is related to the wave energy integrated over frequencies and direction, and it is reasonable to assume that the spatial scale of the errors in the first-guess energy spectra is of similar size. In the present case the influence radius has been set to depend on frequency so that long wave information from the observations influence a larger area of the analyzed field than shorter waves. The influence radius is thus ranging from 300 km for $f = 0.04$ Hz to 180 km at 0.11 Hz and 60 km at 0.24 Hz.

The method assumes that the first-guess and observation errors are uncorrelated. This is not fully correct since the first-guess background has been used in the inverted spectra for high frequencies as described in the previous section. The effect on the analysis will be that for this part of the spectrum the relative confidence in the first guess will increase. If SAR spectra were to be used in the analyses together with other independent wave observations, the weight of the first-guess information would increase compared to the observations. To avoid this, the weights in (5) are reduced by a certain multi-

plication factor R , equal to the one used to merge first-guess and observation spectrum given in section 2.1. The effect is that no information from the wind sea part of inverted SAR wave spectra is used, consistent with the results shown in section 2.3 that this part of the spectra contains little new information. Consequently, there is no need for updating the wind input at the analysis time, e.g., similar to what is done when analyzing altimeter wave heights measurements [Breivik and Reistad, 1994].

The relative magnitude of the standard deviation of the errors $\sigma(f)^g$ and $\sigma(f)^o$ determines the relative influence of first guess and observation at observation point. For available analyses where a larger number of data has been used to derive statistics for the quality of SAR wave data (e.g., section 2.3), the observations have not been compared to a "true" value but to numerical wave models. Such comparisons are influenced by both errors in the model and errors in the observations and are insufficient for estimating the observation error contributions. The standard deviations of differences of significant wave heights from the model are less for the SAR (Table 1) than for the Gullfaks and Ekofisk measurements (Tables 4 and 5) and the altimeter measurements (Table 7), which indicates that the accuracy of the SAR data compared to a "model truth" is better than for the other instruments. Because of this and the strict quality control described above we choose to have high confidence in the observations when assimilating them in the long wave part of the spectrum, and we choose $\sigma(f)^g = 3\sigma(f)^o$. More studies are needed to decide how near this is an optimal choice.

4. Case Studies

Below, three cases are studied in more detail in order to investigate the impact of assimilating inverted SAR wave spectra. The results are not evaluated against independent measurements. The intention is to demonstrate how the model/assimilation system absorbs the information in the presence of wind forcing and to demonstrate that it can be utilized in the proceeding model forecast. In the first two cases, spectra derived from TSS image mode spectra are used. In these cases the wave model is run with analyzed winds from HIRLAM. The third example is from the parallel runs with assimilation of wave spectra derived from wave mode data received on GTS. These runs are carried out in near-real time, and the wave model is again advancing forward in time using wind prognoses from HIRLAM. SAR data are assimilated in the wave model every second hour when available.

Table 4. Assimilation Run and Operational Run Compared to Gullfaks Wave Heights (MIROS Radar) From November 1, 1995, to January 29, 1996

Forecast Length, hours	Mean Observations SWH, m	Mean SWH, m		RMS SWH, m		Correlation Coefficient	
		noasm	asm	noasm	asm	noasm	asm
0 (analysis)	3.11	3.55	3.55	0.80	0.78	0.89	0.90
3–12	3.09	3.44	3.44	0.80	0.79	0.87	0.88
15–24	3.08	3.31	3.31	0.81	0.80	0.85	0.86
27–36	3.06	3.24	3.23	0.81	0.80	0.84	0.85
39–48	3.03	3.20	3.20	0.82	0.82	0.84	0.84

Assimilation run is asm, operational run is noasm, and root-mean-square is RMS. Number of entries is 671.

Table 5. Assimilation Run and Operational Run Compared to Ekofisk Wave Heights (Wave Buoy) From November 1, 1995, to January 29, 1996

Forecast Length, hours	Mean Observations SWH, m	Mean SWH, m		RMS SWH, m		Correlation Coefficient	
		noasm	asm	noasm	asm	noasm	asm
0 (analysis)	2.50	2.96	2.95	0.70	0.69	0.91	0.91
3–12	2.48	2.81	2.80	0.63	0.62	0.89	0.90
15–24	2.46	2.70	2.69	0.61	0.60	0.88	0.89
27–36	2.46	2.68	2.67	0.63	0.62	0.87	0.87
39–48	2.47	2.65	2.65	0.66	0.66	0.85	0.85

Number of entries is 671.

The first case is from October 19, 1995, with SAR image mode spectra northwest of Scotland around 1150 UTC. The inverted SAR wave spectra have less wave energy than the first-guess spectra from the WINCH model giving a reduction in significant wave height up to 1 m. The peak wave direction is shifted 12° toward the north. To give an impression of the wind input driving the wave model, the mean sea level pressure map at October 19, 1995, 0000 UTC is shown in Figure 4.

Figure 5 shows isolines of the difference of significant wave height between the run with assimilation and the run without assimilation 2, 6, 12, and 18 hours after the assimilation. An

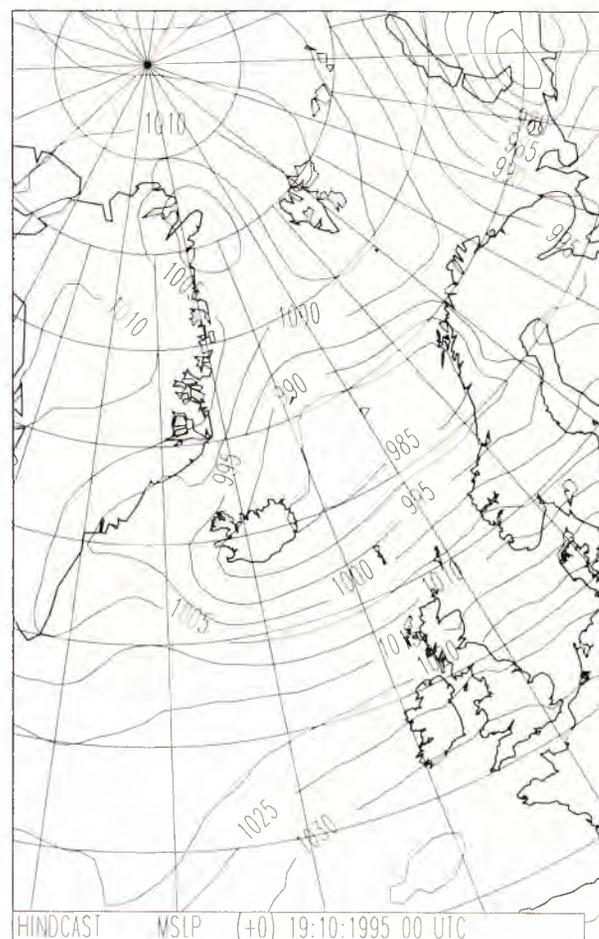


Figure 4. Mean sea level pressure map at October 19, 1995, 0000 UTC.

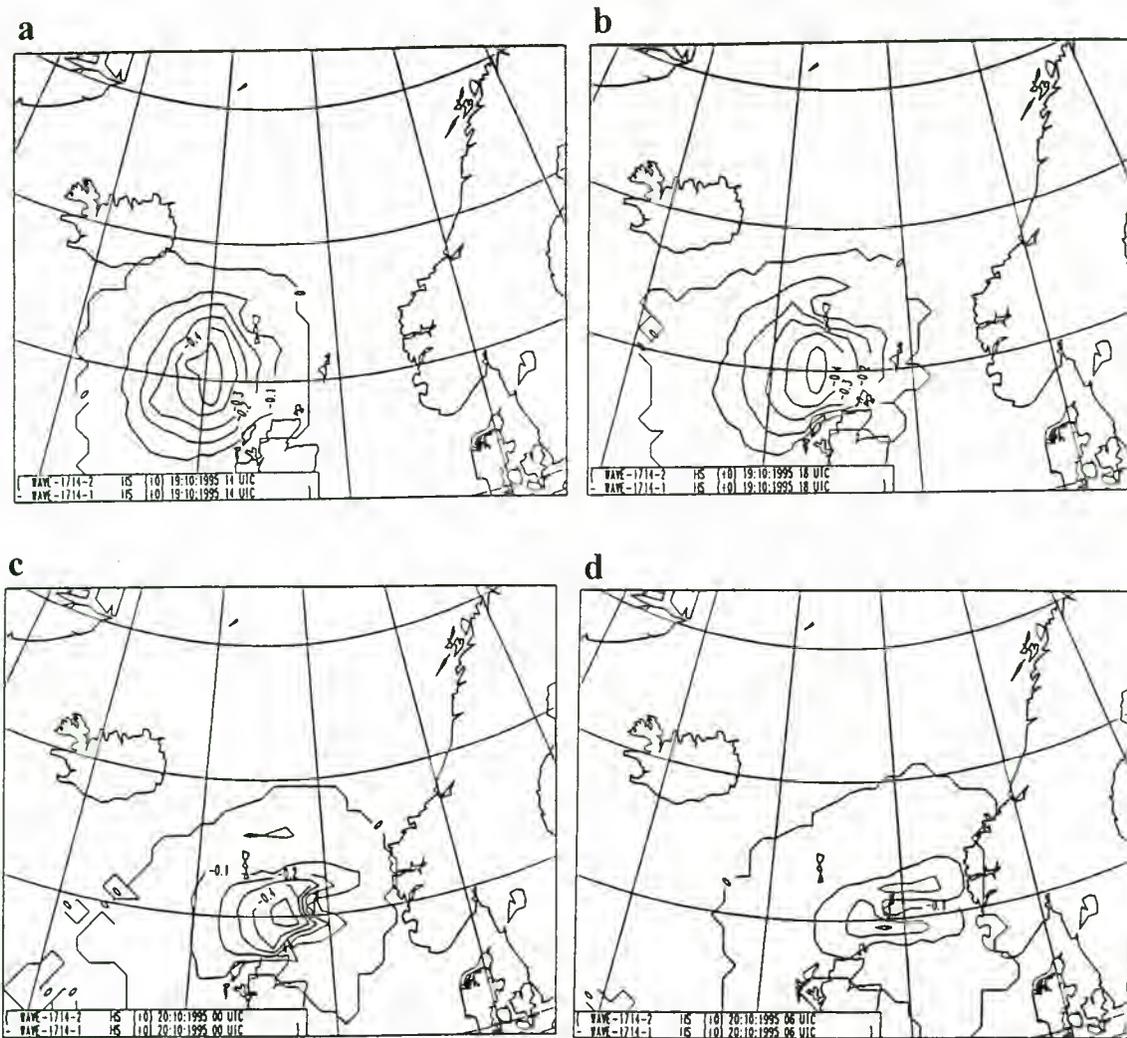


Figure 5. Differences in significant wave height between model runs with and without assimilation (a) 2, (b) 6, (c) 12, and (d) 18 hours after the assimilation.

eastward moving area of differences larger than 0.5 m is present the first 12 hours. After 18 hours the difference is reduced to about 0.3 m, partly because the area of maximum difference reaches Shetland, where there is a land point in the wave model, and partly because the waves are influenced by the wind input from HIRLAM, which is identical for the two runs. After 24 hours most of the influence has disappeared.

In the second example from November 4, 1995, there are SAR image mode spectra northeast of Iceland and northwest of Scotland. In this case the SAR wave spectra have more energy than the first guess model spectra. Examples of spectra are seen in Figure 6. Model first guess, SAR wave spectra, and assimilated spectra are shown for two positions. The wave height increases from 2.5 to 3.2 m in the first point and from 3.7 to 4.2 m in the second point. Figure 7 shows differences between model runs with and without assimilation 2, 6, 12, and 18 hours after the analysis. From Figure 7 it is seen that the largest differences are about 0.8 m and occur northeast of Iceland. Although the differences shortly after the assimilation time are larger than in the first example, the differences decrease more rapidly, and 12 hours after the assimilation the maximum difference is only 0.3 m. This indicates that the parts of the wave spectra that are changed by the assimilation are

more influenced by the wind during the first hours after assimilation than in the first example.

In the last example, with assimilation of wave mode data from November 22, 1995, there are two passages of assimilated ERS observations. In an area off the northern Norwegian coast the wave energy is significantly reduced by the assimilation. Figure 8 shows the first-guess and wave mode spectra at 2000 UTC, and Figure 9 shows the impact on the forecasts. Four hours after the analysis time (0000 UTC, November 23) the maximum reduction of significant wave height is about 1.3 m in the area outside northern Norway. The other observations are between Scotland and Iceland, and here the model run with assimilation has more than 0.6 m lower wave heights at 0000 UTC. In the north the area influenced by the assimilation propagated toward ice-covered areas southeast of Svalbard. Even 16 hours after the assimilation (1200 UTC, November 23) there is more than a 0.7 m difference in wave height. Later, the area of maximum difference reaches the ice-covered ocean, which is treated as land in the model. The differences in wave heights between Iceland and Scotland are moving east, and 22 hours (1800 UTC) after the assimilation, the maximum difference was about 0.3 m.

These three cases, which are picked because of significant

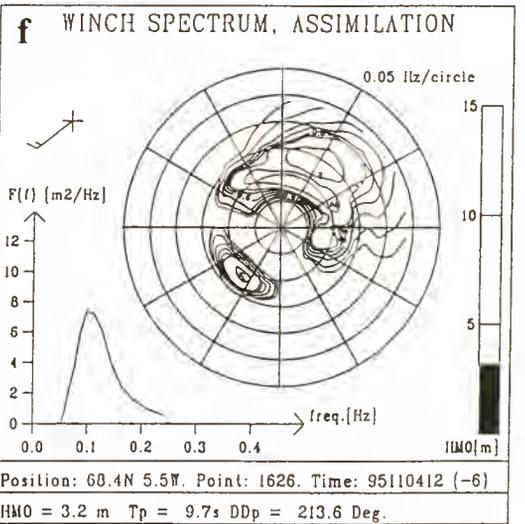
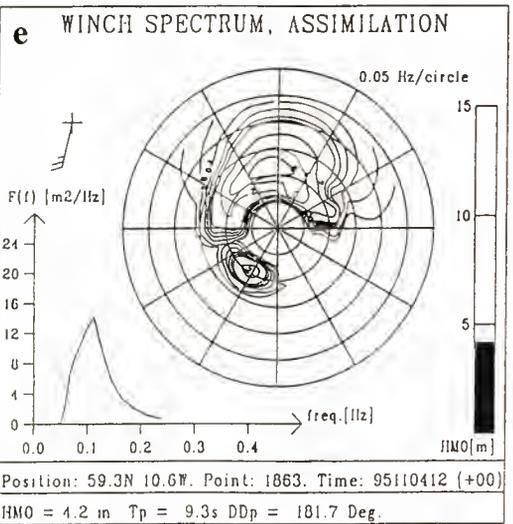
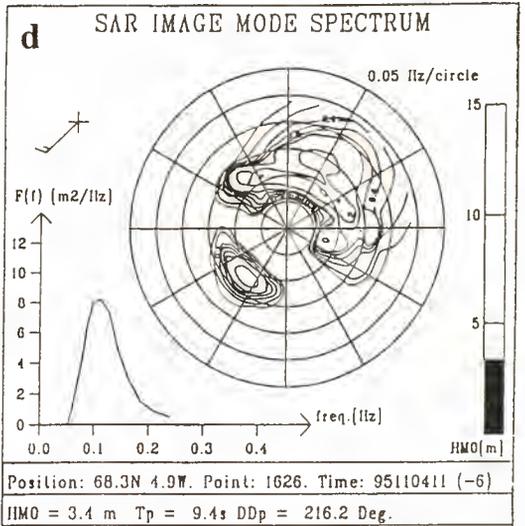
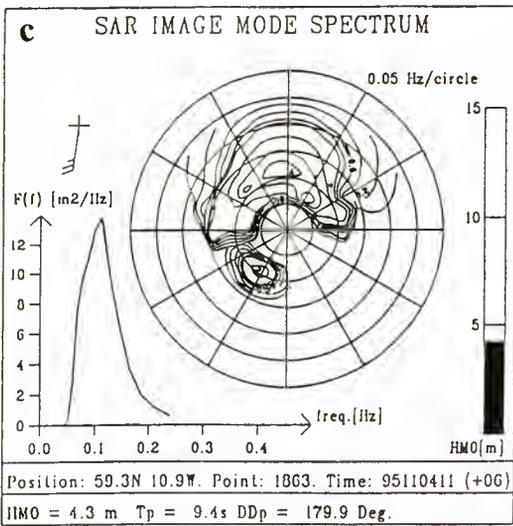
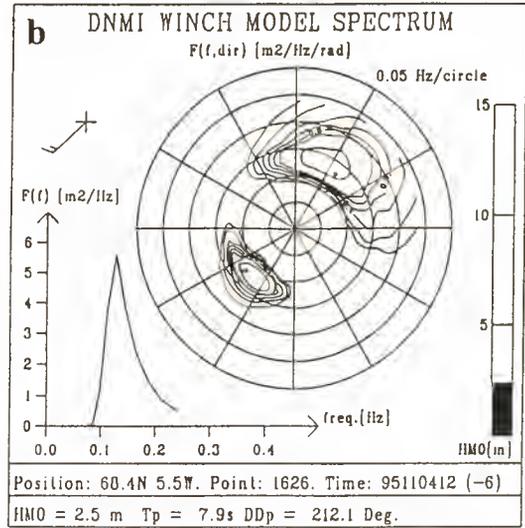
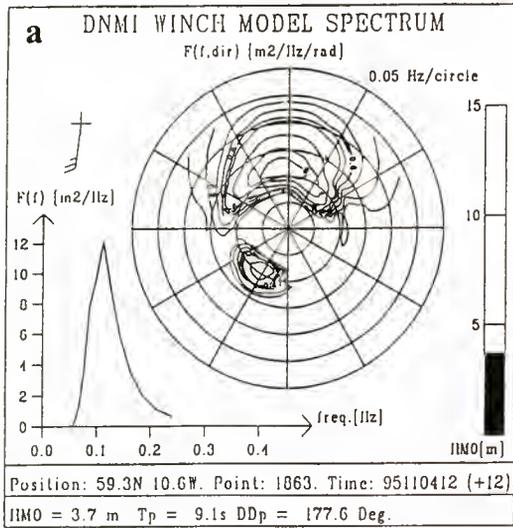


Figure 6. Wave spectra from the (a, b) WINCH model, (c, d) SAR image mode, and (e, f) WINCH model after assimilation from two different positions. Figures 6a, 6c, and 6e, 59.3°N, 10.6°W; Figures 6b, 6d, and 6f, 68.4°N, 5.5°W.

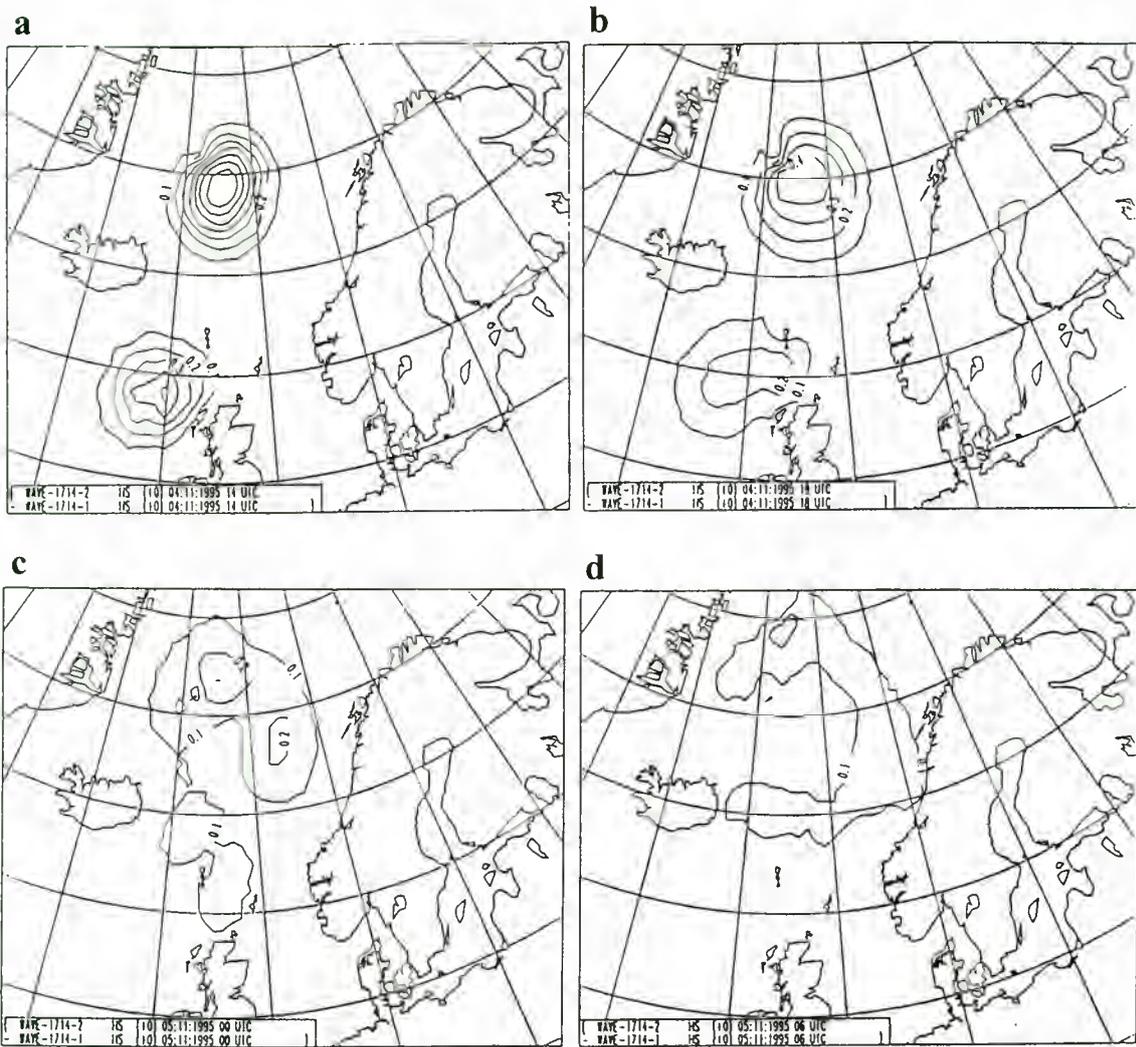


Figure 7. Differences in significant wave height between model runs with and without assimilation (a) 2, (b) 6, (c) 12, and (d) 18 hours after the assimilation.

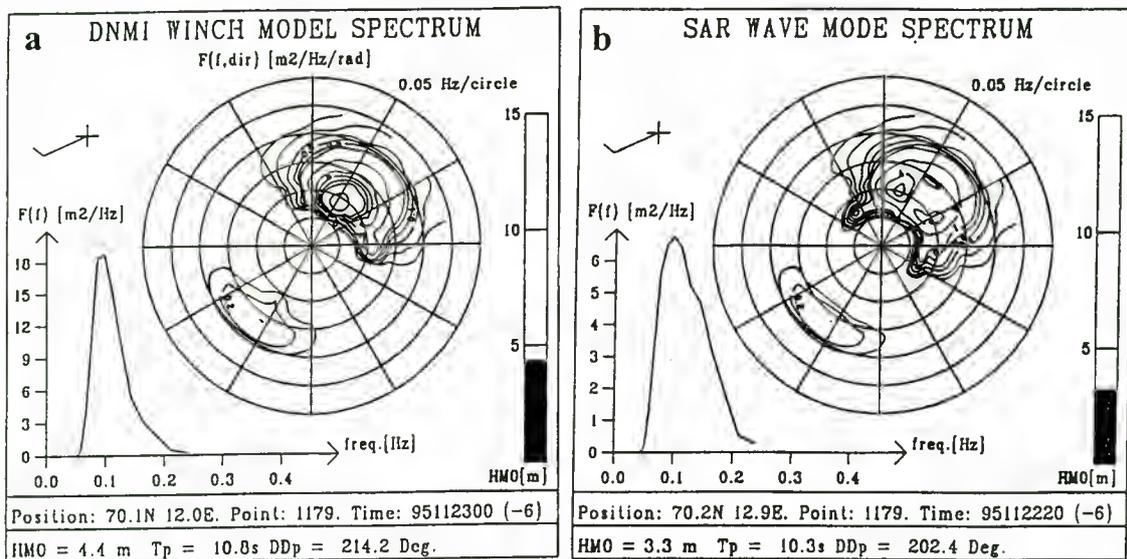


Figure 8. Wave spectrum from the (a) WINCH model and (b) inverted SAR wave mode.

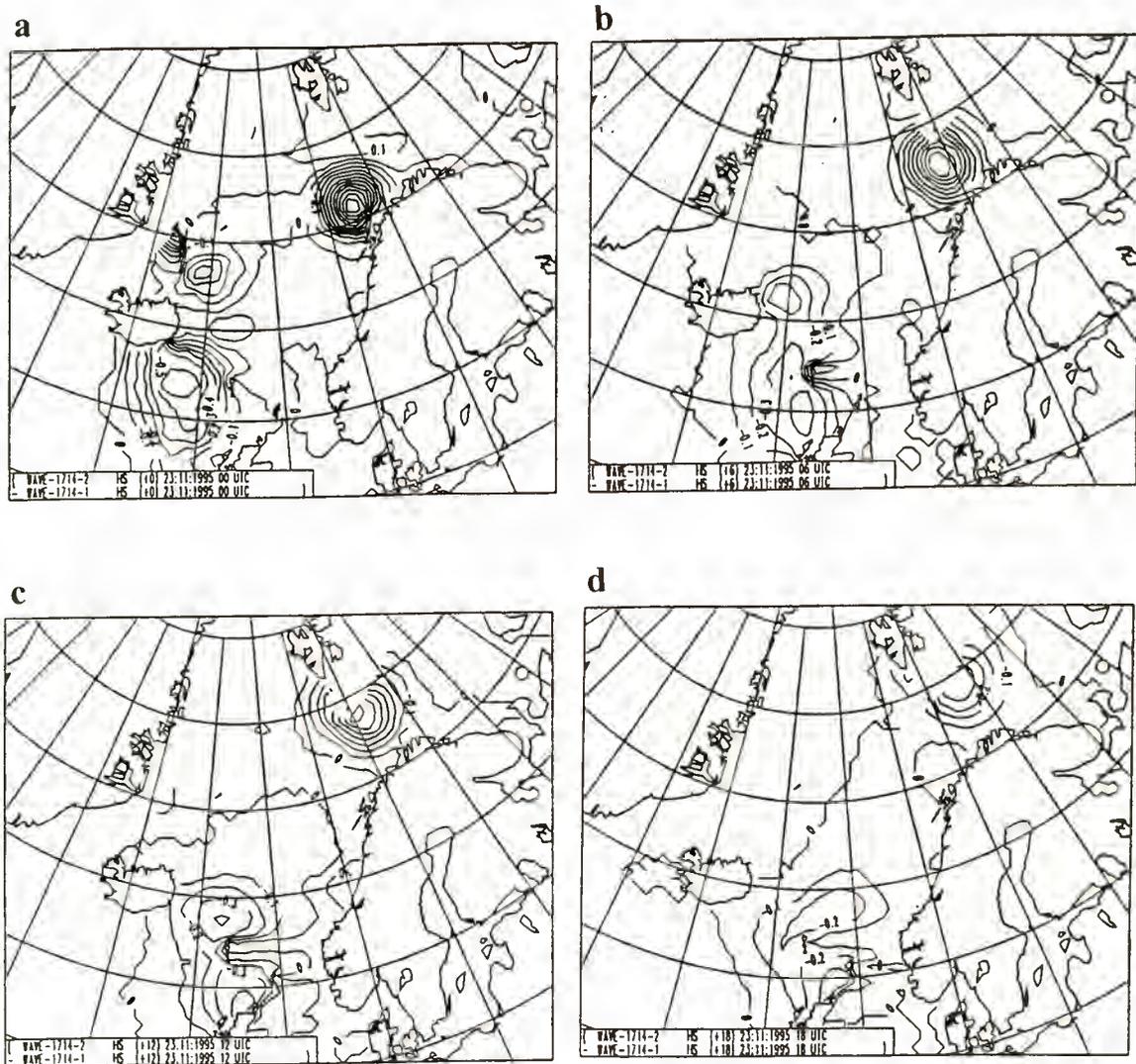


Figure 9. Differences in significant wave height between model run with and without assimilation (a) 4, (b) 10, (c) 16, and (d) 22 hours after the assimilation time.

differences between model first-guess and observations, show that the information is well assimilated with an area of impact characterized by the influence radius. The information is further propagated by the model and seen up to 24 hours after assimilation. The impact gradually decreases as new wind input enters the model or the wave systems reach the coast or areas covered by sea ice.

5. Experience From 4 Months of Assimilation of ERS Wave Mode

A wave model routine with assimilation of wave mode wave spectra from ERS-1 and ERS-2 was run in parallel to the operational model run from October 1995 to January 1996. The results from the assimilation run (asm) and the operational run (noasm) was compared to in situ measurements of significant wave heights (SWH) from two platforms in the North Sea (Gullfaks and Ekofisk) and with the ERS-2 altimeter SWH measurements. Data from analysis and forecasts up to 24 hours were compared, and SWH is the only evaluated parameter. Since the forecast SWH field depends not only on the initial SWHs but also on wave propagation and hence the

complete spectra, the result for SWH is actually a measure of the quality of assimilating all spectral components.

The in situ SWH are 3 hourly with 0.5 m resolution. At Gullfaks the waves are measured by a MIROS radar and at Ekofisk by a wave buoy. Table 4 and Table 5 show the mean SWH from the measurements and the model data, the root-mean-square (RMS) difference, and the correlation coefficient for the validation period November 1, 1995, to January 29, 1996. The statistical parameters are given for model data at analysis time and for four different forecast intervals (3–12, 15–24, 27–36, and 39–48 hours).

There is only a very small effect of the assimilation of SAR data. The limited effect is partly due to the relatively small number of SAR observations, approximately nine observations per day in the whole integration area for the model. Moreover, as seen in the previous chapter, the influence of the observations is roughly limited in space by the size of the influence radius. To get large increments caused by the SAR observations on specific locations such as the North Sea platforms, one actually needs nearby SAR observations, which are indeed very few.

Because of the on average small increments between asm

Table 6. Assimilation Run and Operational Run Compared to Ekofisk Wave Heights From November 1, 1995, to March 6, 1996

Forecast, hours	Number of Entries	Mean Observations SWH, m	Mean SWH, m		RMS SWH, m		Correlation Coefficient	
			noasm	asm	noasm	asm	noasm	asm
0	56	2.74	3.38	3.37	0.90	0.87	0.83	0.88
3–12	32	2.41	3.00	3.12	0.94	0.88	0.84	0.87
15–24	16	2.75	3.30	3.42	0.92	0.94	0.67	0.68

Comparisons are for situations where the difference in SWH between asm and noasm exceeds 0.3 m.

and noasm it is difficult to assess the statistical significance of the improved correlation coefficients from the data given in Tables 4 and 5, where only small changes occur in the last presented digit. It is more interesting to assess this on data sets with larger changes caused by the observations. To assess the possibility that the improvements caused by the SAR assimilation occurred by chance, we define for each in situ observation two quantities $x = \text{asm} - \text{noasm}$ and $y = \text{obs} - \text{noasm}$, which are the changes caused by the SAR assimilation and the deviation of the reference in situ observation from the reference run, respectively. The null hypothesis is that the SAR assimilation changes are purely random and hence that x and y are uncorrelated. The computed correlation coefficient (Pearson's r) between x and y will, because of the sampling variability, in general differ from the real correlation (equal to zero if the null hypothesis holds). The sampling distribution for r may be expressed in terms of the related variable $z' \equiv 1/2 \ln [(1+r)/(1-r)]$, which has a normal distribution with standard deviation $\sigma_{z'} = 1/\sqrt{n-3}$, where n is the number of observations.

In Table 6, with data from Ekofisk, only situations where the difference between asm and noasm exceeds 0.3 m are included in the statistics. There are relatively few cases. For the short forecasts, there is an improvement both in RMS and correlation. However, for the longer forecasts the RMS is slightly increased in asm compared to noasm. Computing the correlations between x and y defined above, we find correlation coefficients of 0.8, 0.4, and 0.1 for 0, 3–12, and 15–24 hours, respectively. Computing the corresponding z' values and $\sigma_{z'}$ based on the number of observations, we find that for reasonable levels of significance the improvements for the 0 and the 3–12 hour forecasts are statistically significant, while the small improvement found for 15–24 hours is not. Because of this we cannot put any confidence in the small degrading in the RMS found with SAR assimilation for these long prognoses.

The two runs have also been compared to SWH measurements from the ERS-2 altimeter. Results from October and November 1995 are presented in Table 7. Altimeter data are received in near-real time via GTS. Altimeter SWH observations have been operationally assimilated in WINCH since 1993. However, for the test period when the observations have been used to evaluate the SAR impact, assimilation of altimeter SWH has been switched off in order to ensure that the altimeter data is an independent data source for evaluating the model runs.

The distance between each altimeter observation on the sea surface is ~ 7 km. For the present evaluation, along-track average values over ~ 50 km have been used. Two forecast intervals, a short prognosis (2–12 hours) and a long prognosis (14–24 hours), are evaluated. As seen from Table 7, there are

virtually no increments from assimilating ERS wave mode spectra on the average forecast results compared to altimeter wave heights. When interpreting these numbers, one must keep in mind that there are two effects which will make the impact somewhat less than if the observations were sampled randomly. First, the impact of ERS SAR assimilation is rather local, and the ERS altimeter data are necessarily located to the side of the wave mode observations. ERS-2 altimeter observations are 250 km to the left (relative to the satellite track direction) of the ERS-2 wave mode observations and even farther away from the ERS-1 wave mode data. Second, within the 2–12 hour interval most altimeter observations in our area turn out to appear near the end, 10–12 hours after the assimilation.

In Table 8 and Figure 10, only situations where the absolute difference between asm and noasm exceeds 0.5 m are included in the statistics. Data are taken during the period from October 1, 1995, to January 9, 1996. This reduces the number of cases dramatically. As seen from the table, the assimilation in these cases gives a small, but positive impact. For the short prognosis the standard deviation of errors is reduced by 0.13 m, and the correlation coefficient is increased by 0.08. For the long prognosis the bias in these cases is reduced by 0.43 m, which is a large improvement.

To see if the improvements are statistically significant, we again compute the correlation between the x and y defined above. We find a correlation coefficient of 0.6 for both the short and long prognosis. Using the method described above involving z' , we find that a positive effect is statistically significant in both cases, even if the actual numbers are subject to sampling variability. This shows that for the few cases where SAR data are available, they have a potential for improving the wave forecast.

Table 7. Assimilation Run and Operational Run Compared to ERS-2 Altimeter Wave Heights From October and November 1995

	Short Prognosis		Long Prognosis	
	noasm	asm	noasm	asm
Number of entries	15,143	15,143	15,143	15,143
Mean altimeter, ^a m	2.42	2.42	2.42	2.42
Mean model, ^a m	2.76	2.76	2.70	2.69
Bias (altimeter – model), ^a m	–0.34	–0.34	–0.28	–0.27
Standard deviation, ^a m	0.64	0.64	0.63	0.63
Correlation coefficient, ^a	0.87	0.87	0.86	0.86

The comparison is carried out for two forecast intervals, short (2–12 hours) and long (14–24 hours).

^aSWH.

Table 8. Assimilation Run and Operational Run Compared to ERS-2 Altimeter Wave Heights From October 1, 1995, to January 9, 1996

	Short Prognosis		Long Prognosis	
	noasm	asm	noasm	asm
Number of entries	190	190	95	95
Mean altimeter, ^a m	2.38	2.38	2.28	2.28
Mean model, ^a m	2.72	2.67	2.86	2.43
Bias (altimeter – model), ^a m	–0.34	–0.30	–0.59	–0.16
Standard deviation, ^a m	0.79	0.66	0.53	0.53
Correlation coefficient, ^a	0.75	0.83	0.80	0.86

The comparison is carried out for two forecast intervals, short (2–12 hours) and long (14–24 hours), for situations where the difference in SWH between asm and noasm exceeds 0.5 m.

^aSWH.

6. Conclusions and Recommendations

In this paper we have considered the impact of including inverted ERS-1 and ERS-2 SAR spectra in the operational wave model at the Norwegian Meteorological Institute. Operational use of inverted SAR spectra needs a careful data control. The fairly simple automatic data control implemented in the present system works satisfactory and eliminates spectra that have no wave features or contain excessive noise. About 50% of the data are rejected by the data control.

There appears to be useful information about wave height and wave period in the swell part of the inverted spectra but, somewhat surprisingly, relatively little new directional wave information compared to what the wave model already gives. For wave spectra derived from image mode ERS data delivered through TSS, evaluation of wave heights and periods against WINCH shows that the main contribution of the observations is increased swell energy in cases with low significant wave height.

The relatively simple assimilation method, where the energy in every frequency and directional bin is analyzed separately, works technically well in an operational environment. The new information gained from the SAR spectra are kept in the model for more than 24 hours in some cases as shown in

section 4. However, the situations studied in section 4 are picked because of significant differences from the reference background run without SAR data. Usually, the differences from the background analyses without SAR are smaller.

Evaluated against independent wave measurements in the North Sea and ERS-2 altimeter wave heights, the assimilation is only to a limited extent capable of improving the average results at a given location. There are several reasons for the low average impact. First, there are very few observations, typically two-three observations in each assimilation cycle. As mentioned in section 2.1.1, only every second AMI wave mode product is distributed on GTS. In addition, the automatic data control procedure, which is run operationally prior to the inversions in order to remove excessive low wavenumber noise and spectra with no definite wave features, removes ~50% of the received data. Second, some information from the wave model is used in the derivation of SAR wave spectra, and the amount of new information included in the wave model spectra by assimilation of a SAR spectrum is limited. This might be improved by new methods of deriving wave spectra from SAR without any prior wave model information, like in the cross-spectrum method [Engen and Johnsen, 1995]. The current algorithm for the ERS AMI wave mode excludes utilization of this method, but this will be improved for the wave mode products from ENVISAT scheduled for 1998. A third important reason for limited impact is the fact that the model (and real-world) wave fields are strongly controlled by the input wind field. A short time after the start of the wave forecast the input wind field has much more influence on the sea state than the initial state itself. The effect of correcting the initial state therefore gradually fades away. It is only for swell, propagating undisturbed by the wind, that new information introduced by the assimilation of observed spectra will be kept in the model for a long time. Since the inverted wave spectra only contains long wave information, there is little to obtain by correcting the wind at the analysis time. However, improved results can be expected when assimilation methods taking the time development of the total wind and wave field into account are operationally available [see, e.g., de las Heras *et al.*, 1994; Bauer *et al.*, 1996].

The somewhat negative conclusion is therefore that assimilation

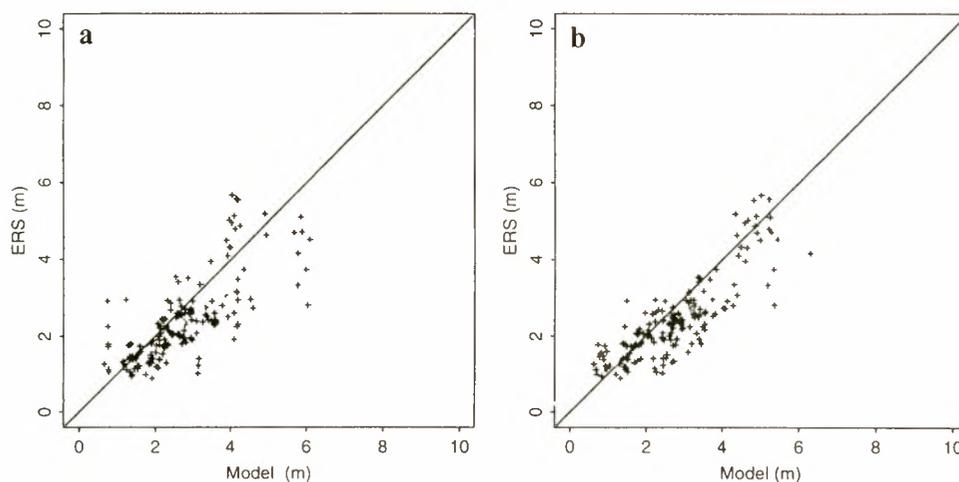


Figure 10. Scatterplots comparing ERS-2 altimeter SWH with model forecasts, 2–12 hours, from the operational run (noasm) and the assimilation run (asm) for situations where the absolute difference in SWH between asm and noasm exceeds 0.5 m.

lation of SAR sea state data in the operational Norwegian wave model, with the present analysis system and the present ERS coverage and data amount received on GTS, is not sufficiently useful to warrant an operational setup. It is, however, quite likely that new SAR processing methods (e.g., the cross-spectrum method), together with increased data coverage from additional satellites, may enable an assimilation system for waves with a significant positive impact. In fact, the positive impact from assimilation of radar altimeter wave heights [Breivik and Reistad, 1994], for which the data coverage is much higher, does indeed show the potential for satellite observations to improve wave analysis and forecasts.

Acknowledgments. This work has been funded by the Norwegian Space Centre under contract JOP 8.3.3.15.93.2 and DNMI. The activities are a part of the ESA program AO2.N104, SAR and Altimeter Data for Ocean Wave Monitoring and Climatology.

References

- Batrick, B. (Ed.), ERS user handbook, *Eur. Space Agency Spec. Publ., ESA SP-1148*, 129 pp., 1993.
- Bauer, E., K. Hasselmann, I. R. Young, and S. Hasselmann, Assimilation of wave data into the wave model WAM using an impulse response function method, *J. Geophys. Res.*, **101**, 3801–3816, 1996.
- Bratseth, A. M., Statistical interpolation by means of successive corrections, *Tellus. Ser. A*, **38**, 439–447, 1986.
- Breivik, L. A., and M. Reistad, Assimilation of ERS-1 altimeter wave heights in an operational numerical wave model, *Weather Forecasting*, **9**(3), 440–451, 1994.
- Breivik, L. A., B. Haugse, H. Krogstad, H. Johnsen, and M. Reistad, Evaluation of SAR wave products, *DNMI Res. Rep.*, **2**, Norw. Meteorol. Inst., Oslo, Norway, 1995.
- Brooker, G., UWA processing algorithm specification, version 7, *ESA Rep. ER-TN-ESA-GS-0342*, Eur. Space Agency, Paris, 1996.
- de las Heras, M. M. G. Bruggers, and P. A. E. M. Janssen, Variational wave data assimilation in a third-generation wave model, *J. Atmos. Oceanic Technol.*, **11**(5), 1350–1369, 1994.
- de Valk, C. F., and C. J. Calkoen, Wave data assimilation in a third generation wave prediction model for the North Sea: An optimal control approach, *Rep. X38*, 123 pp., Delft Hydraul. Lab., Delft, Netherlands, 1989.
- Engen G., and H. Johnsen, SAR-ocean wave inversion using image cross spectra, *IEEE Trans. Geosci. Remote Sens.*, **33**(4), 1047–1056, 1995.
- Esteve, D. C., Evaluation of preliminary experiments assimilating Seasat significant wave heights into a spectral wave model, *J. Geophys. Res.*, **93**, 14,099–14,106, 1988.
- Greenwood, J. A., V. J. Cardone, and L. M. Lawson, Intercomparison test version of the SAIL wave model, in *Ocean Wave Modelling*, pp. 221–233, Plenum, New York, 1985.
- Gunther, H., P. Lionello, and B. Hansen, The impact of ERS-1 altimeter on the wave analysis and forecast, *GKSS Rep. GKSS 93/E/44*, 56 pp., GKSS Forschungszentrum, Geesthacht, Germany, 1993.
- Gustafsson, N., HIRLAM 2 Final Report, *HIRLAM Tech. Rep. 9*, Norw. Meteorol. Inst., Oslo, Norway, 1993.
- Hara, T., and W. J. Plant, Hydrodynamic modulation of short wind-wave spectra by long waves and its measurements using microwave backscatter, *J. Geophys. Res.*, **99**, 9767–9784, 1994.
- Hasselmann, K., and S. Hasselmann, On the nonlinear mapping of an ocean wave spectrum into a synthetic aperture radar image spectrum and its inversion, *J. Geophys. Res.*, **96**, 10,713–10,729, 1991.
- Hasselmann, S., C. Brüning, and P. Lionello, Towards a generalized optimal interpolation method for the assimilation of ERS-1 SAR retrieved wave spectra in a wave model, in *Proceedings of the Second ERS-1 Symposium, October 11–14, 1993, Hamburg, Germany*, *Eur. Space Agency Spec. Publ. ESA SP-361*, 21–25, 1994.
- Hasselmann, S., C. Brüning, K. Hasselmann, and P. Heimbach, An improved algorithm for retrieval of ocean wave spectra from synthetic aperture radar image spectra, *J. Geophys. Res.*, **101**, 16,615–16,629, 1996.
- Hasselmann, S., P. Lionello, and K. Hasselmann, An optimal interpolation scheme for the assimilation of spectral wave data, *J. Geophys. Res.*, **102**, 15,823–15,836, 1997.
- Hollingsworth, A., Objective analysis for numerical weather prediction, *J. Meteorol. Soc. Jpn.*, **65**, 11–60, 1987.
- Janssen, P. A. E. M., P. Lionello, M. Reistad, and A. Hollingsworth, Hindcasts and data assimilation studies with the WAM model during the Seasat period, *J. Geophys. Res.*, **94**, 973–993, 1989.
- Krogstad, H. E., Ocean-SAR Inversion with ERS-1 Wave Mode Spectra, in *Proceedings of the 2nd ERS-1 Workshop*, edited by K. B. Katsaros, pp. 54–59, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, 1994.
- Krogstad, H. E., O. Samset, and P. W. Vachon, Generalization of the non-linear ocean-SAR transform and a simplified SAR inversion algorithm, *Atmos. Ocean*, **32**, 61–82, 1994.
- Lionello, P., H. Gunther, and P. A. E. M. Janssen, Assimilation of altimeter data in a global third generation wave model, *J. Geophys. Res.*, **97**, 14,453–14,474, 1992.
- Schmidt, A., C. Wismann, R. Romeiser, and W. Alpers, Simultaneous measurements of the ocean wave-radar modulation transfer function at L, C, and X bands from the research platform Nordsee, *J. Geophys. Res.*, **100**, 8815–8827, 1995.
- Sunde, J., L. A. Breivik, B. Haugse, H. Johnsen, H. E. Krogstad, and M. Reistad, Processing and application of wind scatterometer data and SAR wave spectra in numerical wave and weather prediction, in *Proceedings of the ERS-1 Pilot Project Workshop*, *Eur. Space Agency Spec. Publ., ESA SP-365*, pp. 29–35, 1994.
- Thomas, J., Retrieval of energy spectra from measured data for assimilation into a wave mode, *Q. J. R. Meteorol. Soc.*, **114**, 781–800, 1988.
- Voss, P., and B. Batrick (Eds.), ESA ERS-1 product specification: General information, *Eur. Space Agency Spec. Publ., ESA SP-1149*, pp. 2–4, 1992.
- Voorrips, A. C., V. K. Makin, and S. Hasselmann, Assimilation of wave spectra from pitch-and-roll buoys in a North Sea wave model, *J. Geophys. Res.*, **102**, 5829–5849, 1997.
- WAMDI Group, The WAM model: A third generation wave prediction model, *J. Phys. Oceanogr.*, **18**, 1775–1810, 1988.

L.-A. Breivik, M. Reistad, H. Schyberg, and J. Sunde, Norwegian Meteorological Institute, P. O. Box 43, Blindern, 0313 Oslo, Norway.
H. Johnsen, NORUT Information Technology, N-9005 Tromsø, Norway.
H. E. Krogstad, Department of Mathematical Sciences, Norwegian University for Science and Technology, N-7034 Trondheim, Norway.
(e-mail: harald.krogstad@math.ntnu.no)

(Received December 30, 1996; revised August 26, 1997; accepted September 24, 1997.)

The effect of assimilating ERS-1 fast delivery wave data into the North Atlantic WAM model

E. M. Dunlap,¹ R. B. Olsen,² L. Wilson,³ S. De Margerie,¹ and R. Lalbeharry³

Abstract. The launch of the European ERS satellites has provided a new source of wave information that is particularly suitable for use in improving wave forecasts in the open ocean. We have implemented and tested a simple system for assimilating corrections to model wave fields produced by the WAM model, where the corrections are derived from inverted synthetic aperture radar (SAR) image spectra from ERS-1. Corrections are applied to significant wave height, mean period and direction for wave modes that are detectable in both the model and the SAR data. The system has been tested in a storm situation and in moderate conditions using buoy data and altimeter data, as well as SAR observations for verification. Overall, it is demonstrated that the net effect of assimilating SAR data is beneficial but very small. The small impact is due at least partly to relatively small spatial and temporal coverage of the SAR wave mode data. Locally larger impacts were found in the storm situation in individual cases where SAR observations were collocated with independent buoy observations.

1. Introduction

Up until a few years ago, ocean wave models were run without the use of any wave observations. Model simulations of wave growth, propagation, and decay were obtained using marine winds as input, either as a series of analyses (“hindcast mode”) or as forecast winds from an atmospheric model (“forecast mode”). Wave observations were not used for two reasons. First, wave models have been shown to simulate the wave field quite well if they are driven by a consistent high-quality wind field [Graber *et al.*, 1994]. Second, so few wave observations were available that they could not be expected to have a significant impact on regional or ocean basin scale wave simulations. With the launch of ERS-1 in July 1991 the spatial and temporal coverage of wave observations increased dramatically, making their use to initialize wave models operationally practical. ERS-1 (and recently ERS-2) wave observations are available in two forms, wave heights from the radar altimeter, and estimates of the two-dimensional (2-D) wave spectrum from the active microwave instrument (AMI) operating in synthetic aperture radar (SAR) mode. The latter offers the opportunity of obtaining real-time wave observations that are the most complete and consistent with the output of a wave model, an estimate of the two-dimensional spectral wave energy.

The general purpose of data assimilation is to change a model’s estimate of the state of its geophysical variables toward their true state, using information obtained from observations. Both model and data are assumed to contain errors, which must be accounted for in the assimilation procedure. With remotely sensed observations it is often the case that the geophysical variable is not directly observed; its values must be inferred or estimated. The raw SAR mode data from ERS-1

are in the form of radar backscatter measurements. The 2-D SAR image spectra estimated from analysis of spatial patterns in the SAR image intensities are known to be a representation of the corresponding ocean wave spectra, subject to certain limitations. Hasselmann and Hasselmann [1991] (hereinafter referred to as HH) were the first to propose a practical and robust scheme for translating SAR spectra into ocean wave spectra. The HH algorithm accounts for the limitations of the SAR spectrum and produces a SAR enhanced estimate of the ocean spectrum by systematically combining the SAR information with estimates of the ocean spectrum from a wave model. The wave model can be expected to provide a full spectrum that is spatially consistent and consistent with the physics of wave growth, propagation, and decay, while the SAR spectrum can be used to correct for errors in the model simulation due, for example, to inaccuracies in the input wind field. The HH algorithm has been run through many thousands of ERS-1 SAR images and has been shown to function reliably [Brüning *et al.*, 1994a].

The HH algorithm provides an estimate of the ocean spectrum from SAR data at the SAR observation location. The process of data assimilation also must include a systematic means of correcting the model wave spectra at all locations within a reasonable range of influence of the observation site. Two-dimensional interpolation methods have been in use in meteorological applications for many years. For example, Cressman-type methods [Cressman, 1959] are the simplest and involve simple interpolation of the differences between model estimates (“first guess”) and the observations to nearby grid points using a weighting function that is inversely related to the distance of the observation site from the grid point. If the weights are determined using the error statistics of the model and their spatial correlation, the method is known as optimum interpolation [Gandin, 1963]. Collectively, these methods are referred to as sequential insertion methods, since the strategy is to run the model forward in time, stopping at regular intervals to assimilate data that are available and all valid at about the same time, then continuing the model run with the corrected model state.

¹ASA Consulting Ltd., Halifax, Nova Scotia, Canada.

²Satlantic, Inc., Halifax, Nova Scotia, Canada.

³Atmospheric Environment Service, Environment Canada, Dorval, Quebec.

Data assimilation methods used so far in wave modeling are of the sequential insertion type and have been applied mostly to altimeter data. Examples are the assimilation of Seasat altimeter wave heights into the U.K. Meteorological Office global wave model [Francis and Stratton, 1990] and the assimilation of altimeter wave heights into the wave model WAM [Lionello et al., 1992]. Ideally, one would like to combine the algorithm that matches the first guess and observation at the observation point with the assimilation step, which spreads the corrections spatially. In meteorological applications this is now being accomplished by 3-D variational methods [e.g., Talagrand and Courtier, 1987]. The variational approach is being pursued actively in operational meteorological data assimilation applications in many national weather centers, but its application in wave data assimilation is so far limited to tests with idealized data (see for example, de las Heras and Janssen [1992]). Variational assimilation will probably be the optimal methodology in wave assimilation as well, but it is more computationally demanding than the older methods. We have thus opted for this first wave data assimilation system to adopt a step-by-step approach, to use methods which had already been tested in wave applications before embarking on experiments with new methods applied to a new data source.

Our goal was to develop a data assimilation system that would be sophisticated enough to give an initial assessment of the impact of SAR wave data on analyses and forecasts from the WAM model. To accomplish this, we put together a full assimilation system by adapting simple assimilation methods and tested it on a variety of North Atlantic cases, including both specific extreme storms and nonstorm situations. The results of these tests are described in this paper. First, the design and construction of the assimilation procedure is described; then results are shown for one storm case, the "Storm of the Century," and for a 1-month period of regular wave forecasts.

2. Assimilation System

Our assimilation system consists of three main components: the wave model WAM, the SAR data preprocessor, and the assimilation module. The assimilation module consists of two main subcomponents: SAR inversion at observation points and interpolation of SAR-induced corrections within the model grid domain. The following sections describe all these components.

2.1. WAM

In this study we have used the wave model WAM, described by the WAMDI Group [1988] with a coupling to the atmospheric boundary layer following Janssen [1991]. This version of WAM, referred to as Cycle-4, was implemented on a $1.0^\circ \times 1.0^\circ$ latitude-longitude grid that covers the northwest Atlantic extending from 25°N to 70°N and from 80°W to 15°W and includes 2318 water points. The model simulates the 2-D spectrum of wave energy discretised into 24 directional bands, 15° wide, and 25 frequency bands logarithmically spaced from 0.042 Hz to 0.41 Hz with an increment-to-frequency ratio equal to 0.1.

2.2. SAR Data Processing

The ERS-1 and 2 satellites carry the active microwave instrument package as their main payload. In the SAR wave mode the AMI acquires signal data to produce a SAR image

(imagette) every 200 km along track. Each imagette is nominally $5 \text{ km} \times 5 \text{ km}$. The across-track position of the imagette is programmed by the European Space Agency (ESA) within the 100-km SAR image mode swath. The imagettes are subsequently transformed into image spectra using a fast Fourier transform (FFT)-based approach. The spectral calculations are performed on intensity images (amplitude squared) where the mean value has been subtracted. For a more detailed description, see Brooker [1994]. A Hamming window is applied to the data prior to the FFT, and the data are subsequently zero-padded to a sample size of 512×512 values before applying the Fourier transform. The resulting spectrum in Cartesian coordinates is converted to polar coordinates on a scale that is logarithmic in wave length, sampled at 12 wavelengths, and specified in 12 directions between 0° and 180° , rescaled to eight bits on a linear scale. The sampling is designed to match the typical spectral representation used in global wave forecast models such as WAM. The spectra are also filtered to wavelength limits of 100 m to 1000 m. The data are subsequently transmitted to ESA/European Space Research Institute (ESRIN) for compilation of global data sets and redistributed to national users. In Canada the data are received at the Atmospheric Environment Service (AES), where they are decoded and converted into a format suitable for ingestion into the Assimilating WAM model, AWAM. During the conversion process, the data are sorted and divided into files, each containing 3 hours worth of data, centered on each WAM model time step. This procedure also filters out data from outside the model domain.

2.3. Inversion

In order to extract ocean wave information from SAR wave data, the mechanisms by which SAR images ocean surface waves must be known. These mechanisms have been discussed extensively in the literature [e.g., Jain, 1981; Alpers, 1983; Hasselmann et al., 1985; Alpers and Brüning, 1986; Hasselmann and Alpers, 1986; Brüning et al., 1990, 1994b; Jacobsen and Høgda, 1994]. The predominant backscattering mechanism at incidence angles encountered with a spaceborne SAR is attributed to Bragg resonant scattering from short surface ripples [Hasselmann et al., 1985]. This backscattering is modulated by three major processes, which contribute to long wave imaging: tilt modulation, hydrodynamic modulation and velocity bunching. The first two are related to the occurrence and local imaging geometry of scattering elements on the ocean surface which vary across the larger ocean waves. Wave orbital motion results in a shift of the backscattering elements, "bunching" them in the SAR image to form wave patterns. This is an effect that is specific to coherent imaging systems such as SAR. The relative importance of the modulation mechanisms is dependent on the propagation direction of the ocean surface gravity waves relative to the radar look direction, the sea state, the viewing geometry, and the satellite height and velocity. The velocity-bunching effect is proportional to the gradient of the radial orbital wave velocity in the azimuth direction.

One of the main limitations of SAR wave data results from the fact that the velocity-bunching effect is nonlinear [Hasselmann et al., 1985], which causes rotation of the spectra toward the range direction and stretching [Brüning et al., 1990]. Another limitation is the incompleteness of the ocean wave information contained in SAR image data. Loss of scene coherence due to orbital motion and the limited lifetime of the scattering elements leads to a loss of resolution in the azimuth

direction [Vachon *et al.*, 1989, 1993], resulting in a scene-dependent loss of SAR imaging capability in the higher-azimuth wave numbers. SAR spectra therefore typically exhibit a cutoff in energy along the azimuth wave number axis that is dependent on sea state and wind speed. Furthermore, the ERS-1 SAR wave mode spectra are generated from single frames of imagery, producing a 180° ambiguity in the wave direction.

In the present approach we use the pure nondispersive velocity-bunching theory. The extension to a dispersive case is rather straightforward [Krogstad *et al.*, 1994]; however, it would introduce an additional azimuth spectral cutoff factor and is neglected for simplicity of calculation. The scanning distortion due to the motion of the surface during the scanning, which may be significant for airborne SAR [Krogstad *et al.*, 1994; Krogstad, 1992; Raney and Lowry, 1978] is negligible for satellite-borne SAR and is therefore neglected.

The HH SAR spectral inversion technique, which takes into account the limitations inherent in SAR wave data, is now a well-established procedure [HH; Krogstad, 1992; Bao *et al.*, 1994; Brüning and Hasselmann, 1993; Brüning *et al.*, 1994b; Engen *et al.*, 1994; Hasselmann *et al.*, 1996] and will not be described in detail here. In summary, the inverted SAR spectrum is obtained by a process of optimization based on the nonlinear transform from the ocean wave spectrum to the SAR image spectrum. This transform expresses the nonlinear effects of velocity bunching as a closed form integral. The a priori estimate of the wave spectrum (which is necessary to provide the information missing in SAR data) is supplied by a wave model, in our case WAM.

The nonlinear relationship between the image spectrum S_{SAR} and the ocean surface wave spectrum S_{wave} may be expressed generically as

$$S_{\text{SAR}} = M_{\text{nl}}(S_{\text{wave}}) \quad (1)$$

Because of the limitations described above, this is not a one-to-one mapping.

The inversion scheme is based on the cost function and weighting functions defined as follows:

$$J = \sum_{\alpha=\text{SAR, wave}} \int_{\mathbf{k}} (S_{\alpha}(\mathbf{k}) - S_{\alpha}^0)^2 W_{\alpha}(\mathbf{k}) d\mathbf{k} \quad (2)$$

with the relative weights defined following HH. The shape of the data weight function $W_{\text{SAR}}(\mathbf{k}) = S_{\text{SAR}}^0(\mathbf{k})$ was selected to suppress the noise part of the observed SAR spectrum S_{SAR}^0 , while the main role of the weight $W_{\text{wave}}(\mathbf{k}) = \mu(B + \min(S_{\text{wave}}(\mathbf{k}), S_{\text{wave}}^0(\mathbf{k})))^{-1}$ is to remove the 180° ambiguity in the SAR image spectrum. The term in the cost function measuring the distance between the simulated and observed SAR spectrum gives identical values for the spectral components at the locations \mathbf{k} and $-\mathbf{k}$. However, the term measuring the distance between the fitted wave spectrum and first-guess spectrum will give very large values to low energy spectral components and will favor the wave component at the peak location. The role of a small constant B is to avoid numerical infinity when the spectrum vanishes. The constant $\mu = 10^{-3} \max((S_{\text{SAR}}^0)^3)$ gives a low relative weight to the first-guess term in the cost function J , and also assures the dimensionless form of J . Other choices of weight functions used to control relative weighting of terms in the cost function are also possible [Engen *et al.*, 1994; Krogstad *et al.*, 1994; Lasnier *et al.*, 1994].

The SAR inversion algorithm described above results in a spectrum that combines information from the SAR, within the wave number band it is capable of sensing, and information from the first-guess wave spectrum. The end result of the inversion procedure in some cases leads to nonphysical discontinuities in the transition zone between the SAR capable wave number band and high wave numbers. The proposed remedy for this problem [Hasselmann *et al.*, 1996] implies additional adjustments of the first-guess wave systems at each inversion step, which requires additional computer time and was not implemented here.

2.4. Assimilation

The assimilation is carried out using a simple two-step scheme [Hasselmann *et al.*, 1994; Komen *et al.*, 1994] in which the corrections introduced by SAR wave data are first computed at the observation location and then spread over the model grid domain within a prescribed range of influence for the data. All SAR wave data available in a time window of 3 hours are assimilated simultaneously.

In order to reduce the dimensionality of the assimilation problem, and to make it practically feasible in an operational system, the corrections of only a small number of spectral parameters characterizing the main wave systems in each spectrum were assimilated. To accomplish this, spectral modes were identified in the SAR-derived wave spectra, and in WAM spectra at the grid points within the range of influence of each SAR spectrum location. For each mode the mean energy $\langle E \rangle$, mean frequency $\langle f \rangle$, and mean direction $\langle \theta \rangle$ were calculated. The mode separation algorithm used here was first proposed by Gerling [1992] and was described in detail by Bauer *et al.* [1995], Hasselmann *et al.* [1994], Komen *et al.* [1994, chapter V.4.3], and Brüning *et al.* [1994a].

Secondary partitions resulting from less significant peaks in the spectrum, characterized for example by relatively close peak locations or insufficiently deep valleys between partitions, were merged together. Also, all wind-sea partitions were combined into a single wind-sea mode, associated with the local wind forcing. Corresponding wave systems were subsequently identified for each SAR-extracted spectrum and the WAM spectra in its range of influence, using the approach of Hasselmann *et al.* [1994, 1996]. Modeled wave systems that could be correlated with SAR-derived wave systems were corrected, while SAR-derived wave systems that did not correspond to any of the modeled wave systems were simply added to the first-guess spectrum. Once the three spectral parameters were computed and matched for each distinct wave system, differences between the SAR-based estimates and model estimates were spread to neighboring grid points of the model.

The observed and corresponding modeled parameters were spread over a region of neighboring model grid points using a weighting function and a single-pass version of Cressman's approach [see Cressman, 1959; Francis and Stratton, 1990]. We define the spreading function as

$$P^{\text{new}}(\mathbf{r}_j) = P^{\text{WAM}}(\mathbf{r}_j) + \frac{\sum_{k=1}^{N_{\text{obs}}} \varepsilon^2 w(\mathbf{r}_j - \mathbf{r}_k) (P^{\text{obs}}(\mathbf{r}_k) - P^{\text{WAM}}(\mathbf{r}_k))}{\sum_{k=1}^{N_{\text{obs}}} \varepsilon^2 w(\mathbf{r}_j - \mathbf{r}_k) + 1} \quad (3)$$

Table 1. Dates for Case Studies

Case Study	Data Set	Period
1	Storm of the Century	March 11–21, 1993
2	Near real time	March 16 to April 16, 1996

where the weights $w(|\mathbf{r}_j - \mathbf{r}_k|)$ are a function only of the relative distance between the model and observation locations and where $\varepsilon^2 = \langle \varepsilon_{\text{WAM}}^2 \rangle / \langle \varepsilon_{\text{Obs}}^2 \rangle$ is the expected model error variance, normalized by the expected observation error variance. The observation errors are assumed to be uncorrelated. At present, we assume $\varepsilon^2 = 1$ and use a simple functional form for the weights $w(|\mathbf{r}_j - \mathbf{r}_k|)$:

$$w(|\mathbf{r}_j - \mathbf{r}_k|) = \exp(-R_{jk}) \approx 1 - R_{jk} \quad R \leq 1 \quad (4)$$

$$w(|\mathbf{r}_j - \mathbf{r}_k|) = \exp(-R_{jk}) = 0 \quad R > 1$$

where R is a dimensionless distance, computed in a spherical coordinate system as follows:

$$R = \sqrt{\frac{1}{2} \left(\frac{\Phi_{jk}^2}{L_{\text{lat}}^2} + \frac{\Theta_{jk}^2}{L_{\text{long}}^2} \right)} \quad (5)$$

The distances in the latitude and longitude direction (Φ_{jk} and Θ_{jk} , respectively) between observation and model grid points are normalized by the corresponding scales of the region of influence, L_{lat} and L_{long} . A linear weighting function is used here mainly for efficiency. It may be argued that a single correction pass, as compared with the optimal interpolation scheme [Lionello *et al.*, 1995], results in cases where areas with a high density of observations are given too much weight relative to observations in areas of low data density. This is not likely to be important in our case, however, as the SAR spectra are spaced evenly along the track and are spaced quite sparsely across the satellite tracks. We also note that this interpolation method is not strictly optimal, as we do not have available statistics for model and observation errors. The resulting estimate of the ocean surface wave field is therefore not optimized.

In principle, at least two distinct correlation length scales for wind sea and swell should be used. The wind sea correlation length scales correspond roughly to the storm generation area where the winds are strong, and would be of the order of 200 km. Correlation length scales of the order of 1000 km were used in assimilating altimeter data [Lionello *et al.*, 1992, 1995; Bauer *et al.*, 1992] in order to spread corrections to all grid points lying between satellite orbits. Here a moderate correlation length scale corresponding to five model grid points (about 500 km) was used. This represents not only a compromise of the two length scales above, but also a typical propa-

Table 2. Summary of Experimental Data Sets

Case Study	Model Input Data		Verification Data		
	CMC Forecast Winds	ERS-1 SAR Wave Mode	Buoys	ERS-1 Altimeter	ERS-1 SAR Wave Mode
1	yes	FDC	yes	FDC	FDC
2	yes	FDP	no	FDP	FDP

CMC, Canadian Meteorological Center; FDP, fast delivery product; FDC, fast delivery copy.

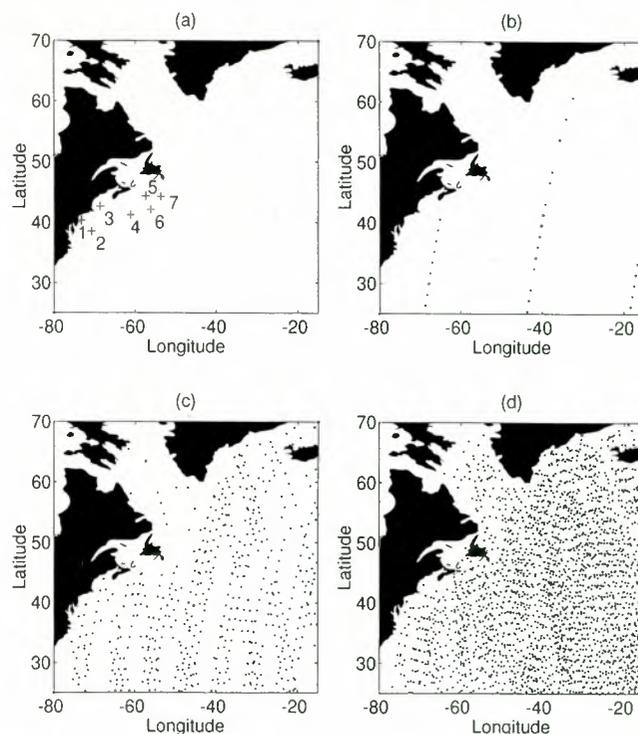


Figure 1. (a) Buoy locations used in verification, and coverage of the input ERS-1 SAR wave data: (b) during one synoptic period of 12 hours, (c) for case 1 (March 11–20, 1993), and (d) for case 2 (March 16 to April 16, 1996).

gation distance for wave energy in a 3-hour time step at wavelengths dominated by swell. Analyzed wave spectra were constructed by topologically adjusting (scaling, stretching and rotating) each of the partitioned WAM wave systems to match its spectral parameters with the interpolated ones.

Unlike other SAR data assimilation systems under development [Hasselmann *et al.*, 1996], we make no attempt to correct the wind fields locally near the SAR observation points. This is mainly because the temporal and spatial distribution of ERS wave mode data is too sparse to maintain wind corrections in the region associated with the storm track, and the impact of locally correcting the wind is expected to be minimal. One might expect that this will limit the lifetime of the corrections to the wave field, since the (possibly) inconsistent wind field would damp them out. However, even though SAR can detect some wind-field-driven wave components, particularly if they are range traveling, it primarily picks up low-frequency swell components, which are decoupled from the wind field. Eventually, it should be possible to use the swell corrections to adjust the wind at an earlier time in the simulation, but it will be necessary to use continuous (time-dependent) insertion methods to accomplish this. Owing to the scarcity of SAR wave data the corrections to the wind field will be most effective after being assimilated into an atmospheric model. We therefore believe that local wind corrections would have little or no effect on the demonstration of the utility of SAR data.

3. Data and Test Method

The system has been tested on two cases, representative of both a storm situation and a case of moderate sea state and wind conditions. The data available for the study consist of

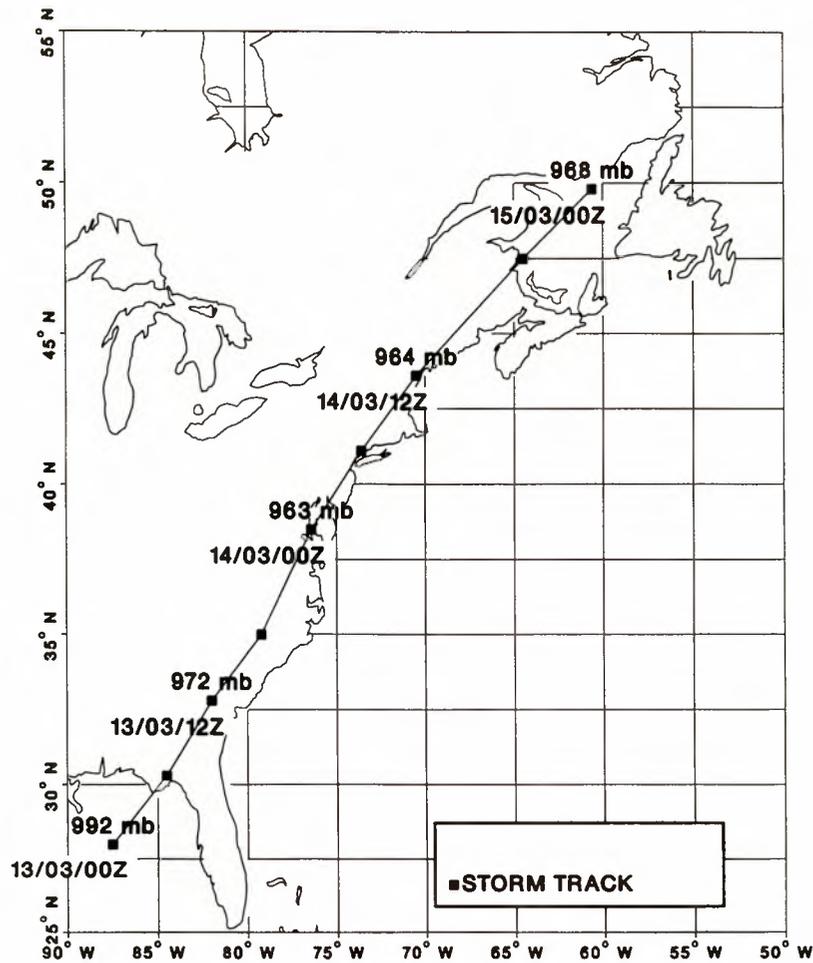


Figure 2. Track of the Storm of the Century, March 11–20, 1993.

model input data (wind fields for driving the model, and ERS SAR wave mode data for assimilation into the model) and verification data (ERS altimeter data (significant wave height), nondirectional wave buoy data, and inverted ERS SAR data not used in the assimilation).

The cases we have considered are listed in Table 1, details of the different data sets are shown in Table 2, and buoy locations and coverage of the satellite data for a typical 12-hour synoptic period are shown in Figure 1 along with the total coverage for the two cases. Case 1 is a particularly intense storm, referred to as the “Storm of the Century,” which produced measured significant wave heights of 16.3 m south of Nova Scotia. The track of the storm is shown in Figure 2. Although the center of the storm remained inland, it was sufficiently close to the coast that very intense winds were maintained in the eastern half of the storm over the Atlantic. By 1800 UTC on March 13 a southeasterly flow of 45 knots (22.5 m s^{-1}) had developed over a large area south of 37°N . Six hours later, winds east of the storm center were reported as high as 60 knots (30 m s^{-1}), and 50-knot (25 m s^{-1}) southwesterly winds were reported behind the cold front south of the storm. As the storm moved northeastward, an intense southerly low level jet (60 knots reported) developed south of Nova Scotia and persisted for at least 12 hours. The wind sea developed in response to this jet later was supported by seas developed by an equally strong west-southwesterly flow, which followed the cold front as it swept eastward across the ocean south of Nova Scotia. It is the

interaction of these wave systems that likely led to the reported maximum significant wave height at 0000 UTC on March 15. In general, the maximum significant wave heights were in the cold air closer to the coast, located generally well south of the storm center but following its track northeastward. The maxima were in excess of 14 m from 0000 UTC on March 14 until 1200 UTC on March 15. By March 16 the low had reached Iceland, producing an elongated area of westerly (eastward moving) high seas south of the storm from east of Newfoundland to south of Iceland. Seas had subsided over the western Atlantic by this time as a high-pressure area moved offshore.

In order to compare and contrast our results from a storm situation, we chose a contiguous month of data to use as a second case for assessing the assimilation method. This period, a late winter–early spring case, does not contain any major storms but represents sea state conditions closer to normal for this time of year. The highest significant wave heights in this period were in the 8-m range.

3.1. Independent (Control) Data Sets

In this study, we have used altimeter, buoy, and SAR data for verification. The altimeter data, while offering only a bulk measurement of the sea state in the form of a significant wave height, offer the advantage of large spatial coverage and provide enough data to perform comparisons with model results in a significant sense. The altimeter data used in this study are summarized in Table 2. For case 1 we obtained 1106 observa-

Table 3. Buoy Station Information

Station	Name	Depth, m	Latitude, °N	Longitude, °E	
1	44005	Gulf of Maine	202	42.60	-68.60
2	44025	Long Island	...	40.30	-73.20
3	44004	Hotel	3231	38.50	-70.70
4	44141	Laurentian Fan	4500	42.07	-56.15
5	44139	Banquereau	1100	44.32	-57.35
6	44138	SW Grand Banks	1500	44.23	-53.35
7	44137	Scotian Shelf	4500	41.20	-61.13

Numbers in left column correspond to labels in Figure 1a.

tions that match a model counterpart value, and for case 2 we obtained 1289 observations. The altimeter data were processed as follows.

3.1.1. Averaging. The original altimeter data are sampled at approximately 0.06° , while the modeled data are represented on a $1^\circ \times 1^\circ$ grid. For analyses involving comparisons of wave heights from the altimeter and the model, we computed an average wave height for every 1° of the orbit latitude along the track. Groups with fewer than 7 points were rejected before averaging, which removed most outliers.

3.1.2. Ice point removal. The altimeter data were provided with standard deviation estimates for wave height. Data with a standard deviation larger than 2 m were rejected before further use, as this is commonly used as an indicator for presence of ice [Breivik and Reistad, 1992].

Significant wave height data collected by seven nondirectional, 6-m NOMAD buoys were available only for case 1 (March 1993). The buoy locations are shown in Figure 1a, with details in Table 3.

3.2. Model Runs

In the present study, the WAM was driven in both hindcast and forecast modes by 10-m level winds obtained from the

regional finite element (RFE) weather prediction model of the Canadian Meteorological Center (CMC) [Mailhot et al., 1995]. The RFE model grid is a variable resolution grid with a central window of uniform resolution covering the region including the continental United States, Canada, and the Canadian Atlantic. The RFE model was run twice daily at 0000 UTC and 1200 UTC and generated forecast winds at 3-hourly intervals valid up to 48 hours, which were then interpolated onto the WAM grid.

In the hindcast mode, WAM was run using the first 12 hours (fields at 0, 3, 6, and 9 hours, referred to as 0h, 3h, 6h and 9h) of winds from each wind file. The 0h fields are analysis winds, while the 3h to 9h fields are forecasts. Two types of runs were done: baseline runs, with no assimilation, and assimilation runs, where for each 3-hour time step, SAR data were ingested, inverted, and assimilated. In order to perform comparisons between the model and the observations and to analyze the results for determining the impact of assimilating the SAR data, the model fields of wave height, period, and direction were stored for each time step. The inverted SAR spectra were also stored, along with the corresponding wave parameters calculated from these spectra.

In forecast mode the model was set up to run in a similar fashion. However, after each 12-hour assimilation period, forecast winds for the following 48 hours, again at 3-hour intervals, were used to produce a wave forecast. Thus forecast runs use forecast wind time series to drive the wave model. A new series was initiated every 12 hours through the period of each case. Again, these runs were conducted as baseline and assimilation runs. For the baseline runs there was no assimilation at all. For the assimilation runs, assimilation was done up to the 0h time of the forecast, as would be the case in operational forecasting.

For the detailed comparison of model results and observations from the satellite and wave buoys, the model data from grid points surrounding each observation location were extracted, and a bilinear interpolation was used to obtain a model

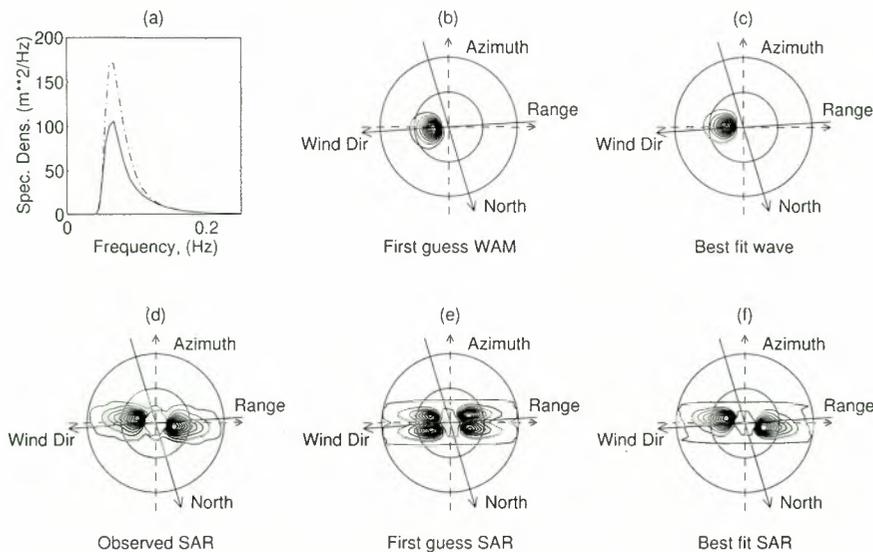


Figure 3. Example of inversion results for a single wave system. First guess and best fit frequency spectra (Figure 3a) are marked with the continuous and dot-dashed lines, respectively. Significant wave heights are 8.82 m for the modeled spectrum (Figure 3b) and 10.50 m for the inverted spectrum (Figure 3c). SAR image spectra are shown for the observed (Figure 3d), first-guess (Figure 3e), and best fit (Figure 3f) spectra. The radii of the outer and inner circles represent wave numbers corresponding to 100 m and 200 m wavelength, respectively.

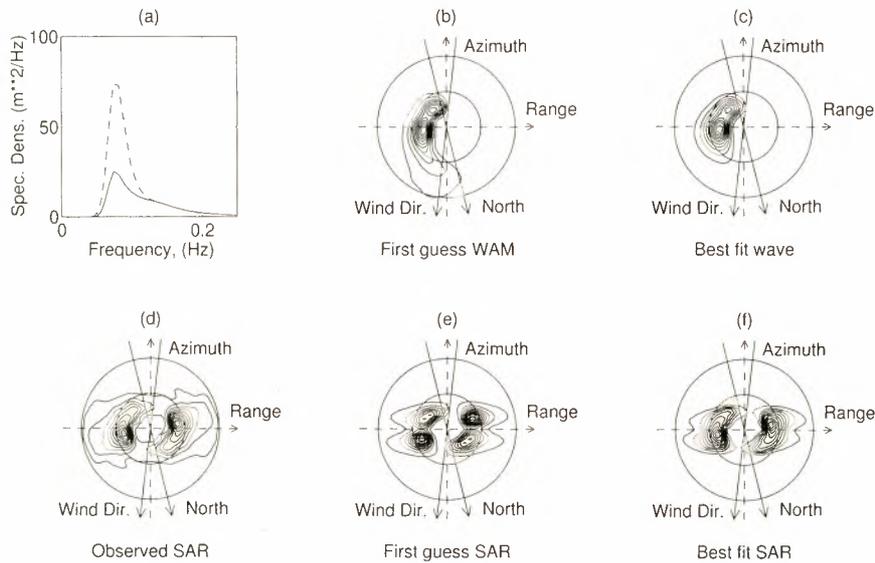


Figure 4. Same as Figure 3, but for two wave systems. Significant wave heights are 4.97 m for the modeled spectrum and 7.05 m for the inverted spectrum.

counterpart for each observation. This was done for the model time steps immediately before and after each observation. The model counterparts were then interpolated in time to match the time of the observation.

4. Results

The data sets and test methods described above were designed to allow us to determine the impact of the SAR data on wave analyses and forecasts from WAM. To assess the impact, we have compared the summary wave parameters from the different runs in different combinations. The statistics used in the comparison are defined in the appendix. The results of the comparisons are presented and discussed in this section, beginning with a discussion of the performance of the inversion algorithm and followed by an assessment of impacts on both hindcast and forecast runs compared to independent data. The verification was limited to comparisons of summary parameters because no directional spectral observations were available and because it is the summary wave parameters that were assimilated. Throughout this section, the run without assimilation is referred to as the “baseline run” and the runs with assimilation are called “assimilation runs.”

4.1. Inversion

Examples of the inversion, for one and two wave systems are shown in Figures 3 and 4, respectively. Spectra are given in the azimuth and range coordinate system. True north and the corresponding wind velocity are also indicated. Wind direction is shown using the oceanographic convention, (i.e., “going to”). The first-guess and best fit SAR spectra are created by mapping the first-guess WAM and best fit wave spectra, respectively. Figure 3 shows that the inversion procedure leads to a spectral estimate with a slight directional change. The inversion also results in a larger significant wave height and a slight increase in the peak period. The result is seen to be closer to the corresponding wind direction. This case represents an almost range-traveling wave system. Thus nonlinear effects are small. However, the relatively broad wave spectrum with a

single peak traveling in the range direction is often mapped into a double-peak SAR image spectrum on account of too small values of the RAR transfer function relative to the velocity bunching component [e.g., Brüning *et al.*, 1990]. The velocity-bunching mechanism vanishes in the range directions and rapidly resumes its large value on both sides of this direction. This leads to a deep, nonphysical trough in the spectrum. The calculated SAR spectrum was significantly modified in this case, compared with the first-guess SAR spectrum, changing from a split peak spectrum to a spectrum in which the peaks around the range wave number axis are joined together. After the inversion the simulated and observed SAR spectra are in much better agreement, with a pattern correlation coefficient increasing from 82.4% for the first guess to 99.4% for the best fit.

Figure 4 exhibits a case with two dominant wave modes. The inverted SAR spectrum shows a very large difference in wave energy (it actually doubles the wave height). Also, we observe that the relative peak strength changes. Here the pattern correlation coefficient increased from 89.6% to 95.7%.

Table 4 documents the overall inversion success rate, which is about 10% higher for the storm case than for case 2. This result is expected in the sense that the average wave height, and therefore the average signal strength in the SAR data is higher for the storm case. This in turn means that fewer spectra are likely to fail to invert because of low signal-to-noise ratio.

In general, we have found that the mean values of the significant wave height extracted from ERS-1 SAR wave mode spectra tend to exceed wave heights calculated by WAM by

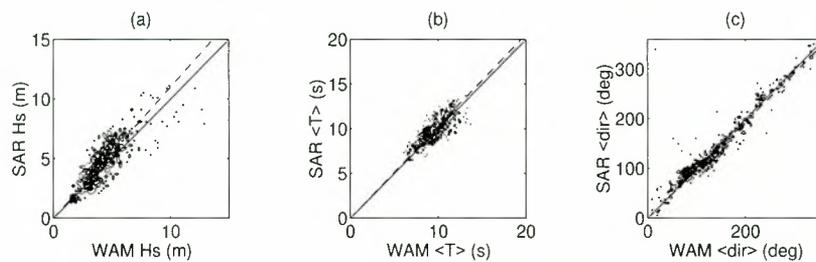
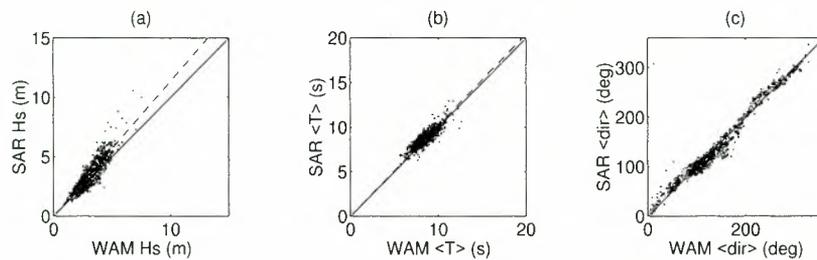
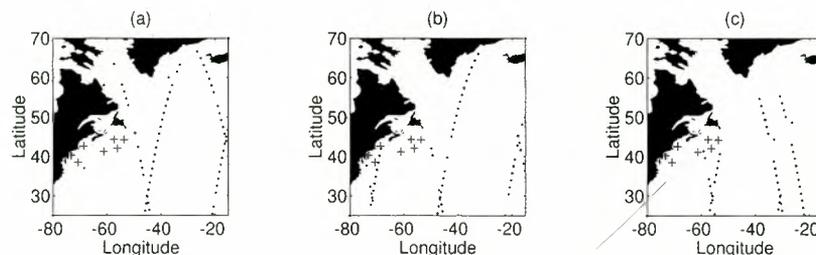
Table 4. Number and Percent of Inverted SAR Image Spectra

Case	Total	Number Inverted	Percent Inverted	Percent Rejected
1	714	548	76.8	23.2
2	3504	2423	69.15	30.85

Table 5. Comparison of WAM Model Versus SAR-Based Estimates of Wave Height, Period and Direction for Case 1 (548 Samples) and Case 2 (2423 Samples)

	H_S , m		$\langle T \rangle$, s		$\langle k \rangle$, m^{-1}		$\langle \text{dir} \rangle$, deg	
	Case 1	Case 2	Case 1	Case 2	Case 1	Case 2	Case 1	Case 2
Mean (WAM)	4.09	2.88	9.24	8.27	0.05	0.06	140.86	152.13
Bias	0.43	0.38	0.28	0.31	-0.004	-0.005	-2.38	-4.59
Std	1.00	0.73	0.87	0.77	0.01	0.01	46.23	47.82
SI_D , %	25.36	6.70	9.27	9.84	18.59	19.67	33.12	31.24
Slope	1.10	1.13	1.03	1.02	0.93	0.92	0.98	1.001
Correlation, %	97.66	98.15	99.58	99.61	98.49	98.56	95.86	96.25

Notation is defined in the appendix. Cor, correlation.

**Figure 5.** Scatter diagrams of the ERS-1 SAR based wave parameter estimates versus WAM based wave parameter estimates for case 1: (a) H_S , (b) average period, and (c) vector mean direction.**Figure 6.** Same as Figure 5 but for case 2: (a) H_S , (b) average period, and (c) vector mean direction.**Figure 7.** Coverage of assimilated SAR data in relation to buoys for case 1: (a) March 11, (b) March 14, and (c) March 16. Each panel shows 24-hour coverage.

approximately 10–12% (Table 5, Figure 5, and Figure 6). The bias is positive and larger for higher wave energies. This may be caused by a deficiency in the wind field driving WAM or by too small a value of the real aperture radar transfer function. A similar but negative bias, about 8%, was observed for the mean wave number. These results agree well with monthly statistics published by *Brüning and Hasselmann* [1993]. Figures 5 and 6 also show a comparison of wave directions. We observe that the predominant wave direction is within 50°–100°.

4.2. Hindcast Comparison With Buoy and Altimeter Data

In order to assess the impact of SAR data assimilation throughout the Storm of the Century, we compared model output with buoy observations. The results, however, are dependent on the buoy locations. The available buoy data locations are along the western fringe of the North Atlantic and are windward of most of the satellite observations. Therefore most of the wave systems and associated corrections arising from data assimilation propagate away from the buoy locations. The measurable impact is directly related to the region of influence and the spreading function applied. It is also very closely related to the spatial and temporal coverage of the SAR observations. In Figure 7 we show three examples of the 24-hour satellite coverage for March 11, 14, and 16, 1993. In Figures 8 and 9 we show time series comparing wave parameters from the model with buoy observations in the Gulf of Maine and off Long Island. The results are typical of all the buoy results and show that there is little difference in the two model runs, baseline and assimilation, early in the period but greater impact later on, coinciding with satellite passes closer to the buoy locations. We also see that the assimilation of SAR data results in a reduced overall difference between the model and the buoy observations. The improvement is most pronounced during periods of decaying waves. Table 6 shows a summary of the impact of the assimilation for each of the buoys, in terms of the changes in scatter index and correlation between buoy significant wave height and period and model significant wave height

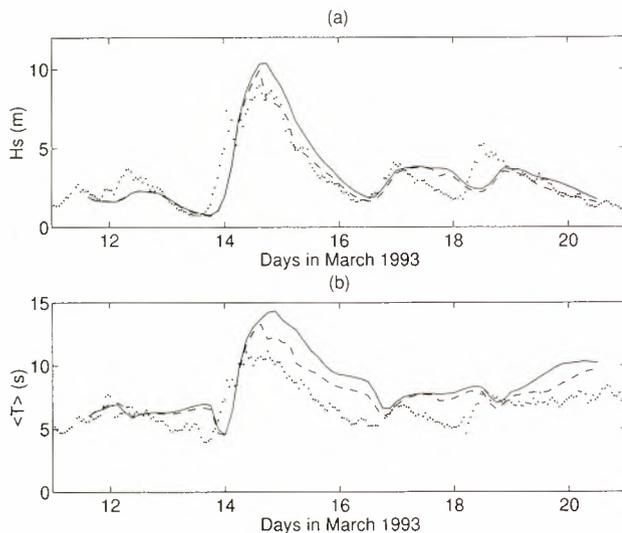


Figure 8. Comparison of hindcast series of (a) wave height and (b) mean period for buoy 44005 in the Gulf of Maine (case 1). Modeled wave parameters with and without assimilation are indicated by dashed and solid lines, respectively. Dots represent buoy observations.

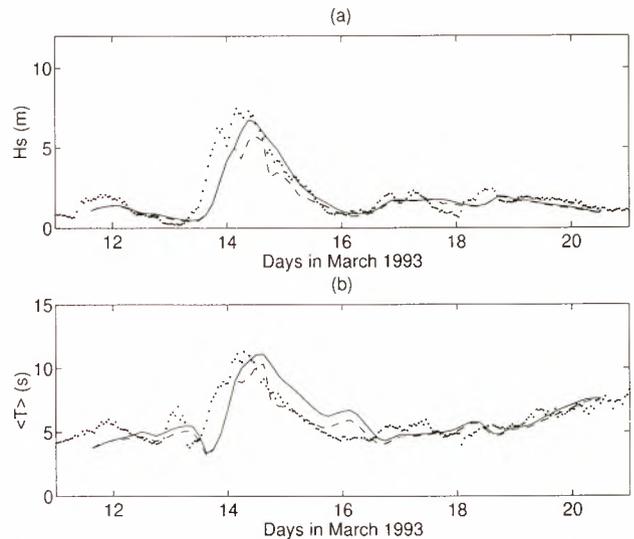


Figure 9. Same as Figure 8, but for buoy 44025 (Long Island).

and period. Results are averaged over the 9-day period of case 1 and are based on about 183 observations for each buoy. Both the sample sizes and the observed changes are small. The overall scatter index is reduced by 0.3% and 1.55%, and the correlation is increased by 0.2% and 0.26% for the significant wave height and for the mean period, respectively. The impact on wave period is slightly greater than the impact on H_s . A typical example comparing wave height and period for one of the buoys on the Scotian Shelf is shown in Figure 10. The small effect of the assimilation is consistent with the fact that the buoys were very seldom within the influence region of ERS-1 data.

In order to further understand the effect of assimilating SAR data in the case of the Storm of the Century, we measured the impact against observations from the ERS altimeter. The comparisons were made both in terms of summary statistics for all the data and in terms of individual satellite passes. In Figure 11a we show a selected satellite pass, which coincides with the storm area. The effect of assimilating the SAR observations for this pass, presented in Figure 11b (dashed line),

Table 6. Change in Scatter Index and Correlation Coefficient Between Buoy Significant Wave Height and Period and WAM/AWAM Model Significant Wave Height and Period, Due to SAR Data Assimilation

Station	N	H_s , m		$\langle T \rangle$, s	
		ΔSI_D , %	ΔCor_D , %	SI_D , %	ΔCor_D , %
44005	187	-2.46	0.23	-3.13	0.28
44025	185	3.53	0.72	-2.83	0.52
44004	182	-2.18	1.08	-1.91	0.64
44141	180	0.82	-0.24	-0.29	-0.05
44139	185	1.59	-0.22	-3.36	0.34
44138	179	-4.00	2.21	-3.02	0.83
44137	184	2.95	-0.18	-1.40	0.03
All buoys	1282	-0.30	0.21	-1.55	0.26

N is number of observations.

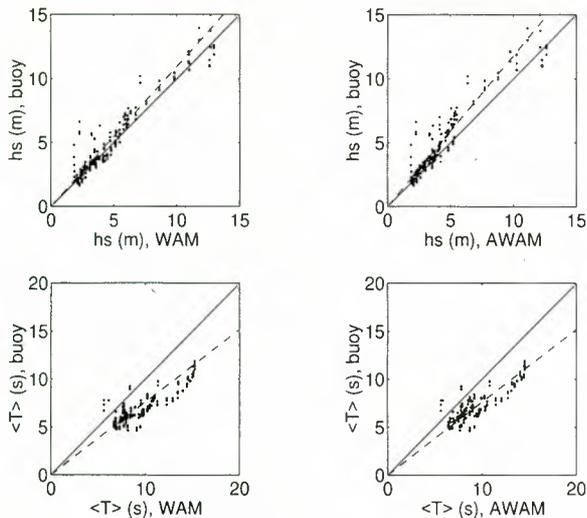


Figure 10. Comparison of buoy and model hindcast (top) wave heights and (bottom) mean periods for Scotian Shelf (buoy 44137), including (left) baseline run results and (right) assimilation run results. The dashed line is a least squares regression fit to the data points.

shows an improvement over the baseline case (solid line). In this case, the scatter index is decreased from 21.33% for the baseline run to 16.03% for the assimilation run, while the correlation coefficient is increased from 98.42% (baseline) to 99.02% (assimilation). The altimeter data show a dip in wave height between 38° and 40° along the track, which apparently was not resolved in the model. This feature is actually seen on several passes in the vicinity and is therefore not simply an artifact in the data. A glance at the analysis maps for the time of these altimeter measurements suggests that the dip in the significant wave height is related to a trough and associated wind shift from west to NW. The gradient in this area was also slackening with time. The feature was apparently not resolved in the modeled wave field. Even though the dip at the northern

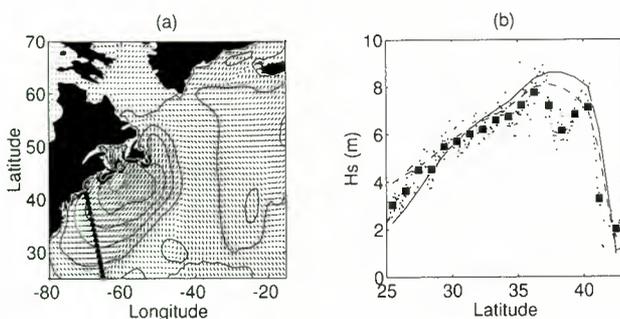


Figure 11. Comparison of altimeter and WAM significant wave heights along a specific satellite track for March 15, 0300 UTC. (a) Map showing contours of significant wave height and vector wave direction. Dots indicate location of averaged altimeter observations. (b) Significant wave height as a function of the latitude along the satellite track, for baseline run (solid line) and assimilation run (dashed line). Altimeter data averaged over a distance of 1° are indicated by squares. Maximum model $H_s = 12.7$ m.

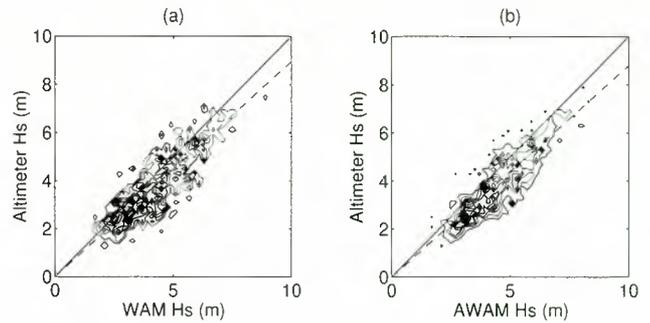


Figure 12. Scatter plots of WAM H_s versus ERS-1 altimeter wave heights for (a) the baseline run and (b) the assimilation run, for case 1. All hindcast series data are included. The dashed line is a least squares regression fit to the plotted data.

end of the track is not reproduced, the overall fit is better than that for the run without assimilation.

Our summary comparison of model results and altimeter data is shown in Figures 12 and 13 for the hindcast runs. The accompanying collocation statistics are given in Table 7. In both case 1 and case 2 we observe a clear tendency for the model to produce higher wave heights than the ERS-1 altimeter observes. In the storm case (see Table 7, case 1) the scatter index and correlation are slightly improved as a result of assimilation. An opposite tendency is observed for case 2. However, both cases have slopes that are farther from the ideal slope of 1 than the baseline run WAM result. Thus the overall trend of the assimilation has been to decrease the level of agreement between the altimeter and the model. This is in agreement with recently published results for the ERS-1 altimeter [e.g., Breivik and Reistad, 1992; P. A. E. M. Janssen, personal communication, 1996] which show that the ERS-1 altimeter systematically underestimates significant wave heights as compared with wave model results. Assimilation of SAR data augments this effect, as the SAR-enhanced wave spectra tend to overestimate the wave energy as compared with the model [Brüning and Hasselmann, 1993].

4.3. Forecast Comparisons

In the forecast assimilation runs, data were assimilated for a 12-hour period, and then a forecast was run for 48 hours. This was repeated for each 12-hour period for case 1 and case 2. In Figure 14 we show the root-mean-square of the difference and the bias between the baseline and assimilation runs for wave height, period, and direction. As we expected, the impact of

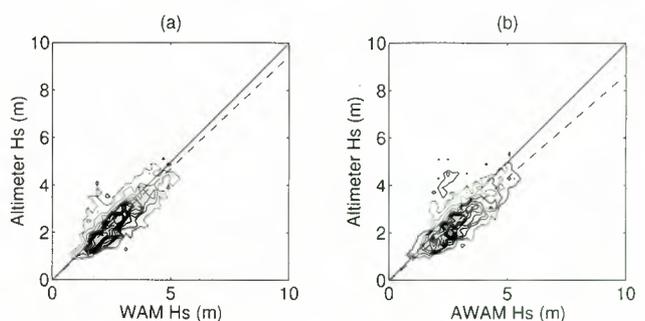


Figure 13. Same as Figure 12, but for case 2.

assimilating the SAR based corrections decays over time. This decay is more rapid for the storm situation in case 1 than in the more moderate conditions of case 2. In both cases the impact lasts for more than 36 hours into the forecast.

4.3.1. Comparison with altimeter. The results of inter-comparison of significant wave heights measured by the ERS-1 altimeter and predicted by the wave model with and without SAR wave data assimilation are shown in Figure 15 for case 1 (1103 samples) and case 2 (3313 samples). The results include the scatter index and the correlation for both the analysis and the forecast. The results illustrate a slight but persistent improvement of the forecast result. The effect is stronger for the storm case, where the scatter is reduced by 1.6% and the correlation is increased by 0.2%, for the first 12 hours of forecast.

4.3.2. Comparison with SAR. To strengthen our verification of the assimilation results, we have also compared forecast model output with SAR observations not used in the assimilation up to that point in time. The results of this comparison show that the assimilation of SAR data improves the agreement between the data sets. In order to perform the comparisons, a model counterpart to each SAR observation was derived by bilinear interpolation of model data to the observation time and locations. The statistics of model results against the inverted but not yet assimilated wave data were averaged over three separate forecast ranges, 3–12 hours, 15–24 hours, and 27–36 hours. The comparison was made between the model with and without SAR data assimilation. The results are summarized in Table 8 and as a function of forecast time in Figure 16. We obtain a reduction in the scatter index of over 3% in wave height at the start of the forecast, which diminishes to less than 1% after 36 hours. The same trend is seen in wave period and mean direction, although with a little more variation in the case of wave direction. Figure 16 also shows that wave period is affected more in the storm case than in the moderate case. The opposite is seen to be the case for wave direction.

5. Discussion and Conclusions

We have built and tested an assimilation system for SAR wave data from ERS-1. The methodology, which treats both the inversion and the assimilation components separately, has been adapted from existing techniques. The system was embedded in a copy of the Canadian operational version of the WAM model, which runs over a regional domain covering most of the North Atlantic. We have demonstrated that the

Table 7. Comparison of Significant Wave Heights With and Without Assimilation, Between Altimeter Measurements and Model Hindcast Results, for Case 1 and Case 2

	Case 1 (1103 Samples)		Case 2 (3313 Samples)	
	WAM	AWAM	WAM	AWAM
Mean (WAM)	4.32	4.45	2.68	2.94
Bias	-0.49	-0.62	-0.15	-0.41
Std	0.95	0.89	0.71	0.76
SI _D , %	23.26	21.52	27.40	27.95
Slope	0.89	0.88	0.94	0.87
Cor, %	97.74	97.90	96.87	96.79

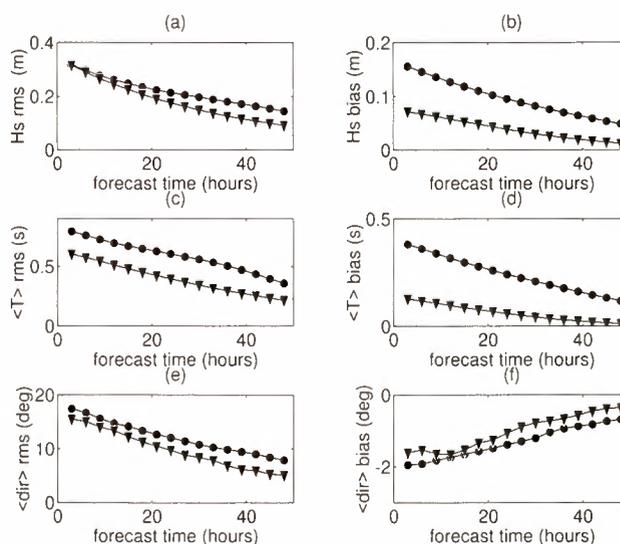


Figure 14. Plots of (left) root-mean-square difference and (right) mean difference between forecast wave parameters run from initial conditions with and without assimilation, as a function of projection time, for cases 1 (triangles) and 2 (circles). Results are given for (a, b) significant wave height, (c, d) average period, and (e, f) vector mean direction.

system can function reliably by running it on a storm event, and through 1 month of wave analyses and forecasts.

In addition to evaluating the assimilation performance in terms of the rate of successful inversion and assimilation, we have also compared both hindcasts and forecasts from the wave model with independent wave observations to determine the impact of the assimilation. In the verification, we used independent control data from the ERS-1 altimeter and non-

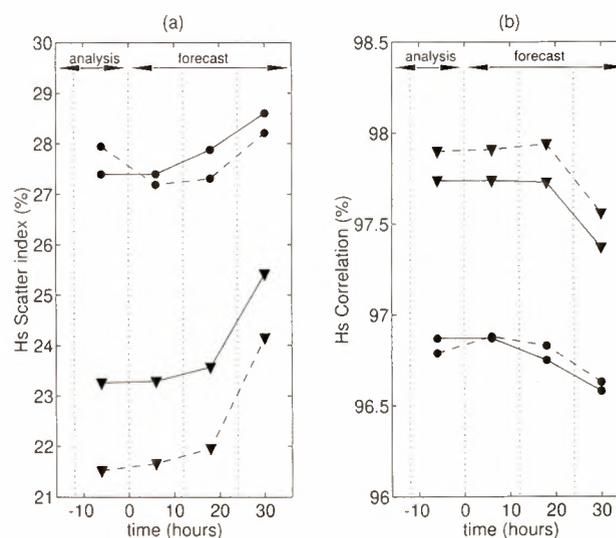


Figure 15. (a) Scatter index and (b) correlation with respect to altimeter H_s , as a function of projection time for case 1 (triangles, 1106 samples) and case 2 (circles, 1289 samples). Each data point represents results based on 12 hours of analysis or forecast. Results of the forecast following the baseline and assimilation runs are represented by solid and dashed lines, respectively.

Table 8. Statistical Comparison of First 12 Hour Forecast Results, Run From the Initial State With and Without Assimilation, Against Inverted but Not Yet Assimilated SAR Observations for Case 1 and Case 2

	Case 1 (376 Samples)		Case 2 (1717 Samples)	
	WAM	AWAM	WAM	AWAM
	H_s, m			
Mean (WAM)	4.49	4.64	2.88	3.11
Bias	0.44	0.39	0.63	0.39
Std	1.35	1.21	0.87	0.80
SI _D , %	28.95	25.77	27.39	24.21
Slope	1.05	1.03	1.23	1.15
Cor, %	96.35	97.12	97.26	97.82
	$\langle T \rangle, s$			
Mean (WAM)	10.51	10.77	8.99	9.52
Bias	-0.76	-1.02	-0.24	-0.77
Std	1.25	0.96	1.03	0.99
SI _D , %	12.32	9.33	11.58	10.81
Slope	0.92	0.90	0.97	0.99
Cor, %	99.29	99.59	99.35	99.46
	$\langle dir \rangle, deg$			
Mean (WAM)	141.59	137.60	153.3	137.60
Bias	-3.75	0.24	-8.63	-1.68
Std	42.91	40.28	55.98	42.97
SI _D , %	30.72	29.25	37.58	29.53
Slope	0.97	1.00	0.96	1.00
Cor, %	96.41	96.71	94.48	96.61

directional wave buoys. Owing partly to the lack of available independent two dimensional spectral wave observations, and also to the fact that the assimilation focuses on three summary parameters of the wave spectrum, the evaluation concentrated on the same parameters: significant wave height, average period, and wave direction. Since we could not obtain independent wave direction information, we also used not-yet-assimilated SAR data as another source of validation data, taking into consideration that the use of SAR data in the assimilation makes this a less independent data source.

Our comparison of the SAR wave data with the model wave parameters revealed that the SAR tends to give higher wave heights and longer periods than WAM and that the difference in wave height increases with higher sea states. This is in agreement with other published results.

When averaged over all the available data for a period of several days or more, the impact of the assimilation of SAR wave data was always small in magnitude. However, our comparisons with independent observations indicate that the impact was usually positive, that is, the average difference, the standard difference and the scatter index between the model and observed wave parameters usually were reduced slightly by the assimilation, while the correlation with independent observations increased slightly. There was one notable instance where a small negative impact was indicated, for the hindcast comparison with altimeter data for case 2, where both the bias and standard difference were increased by the assimilation.

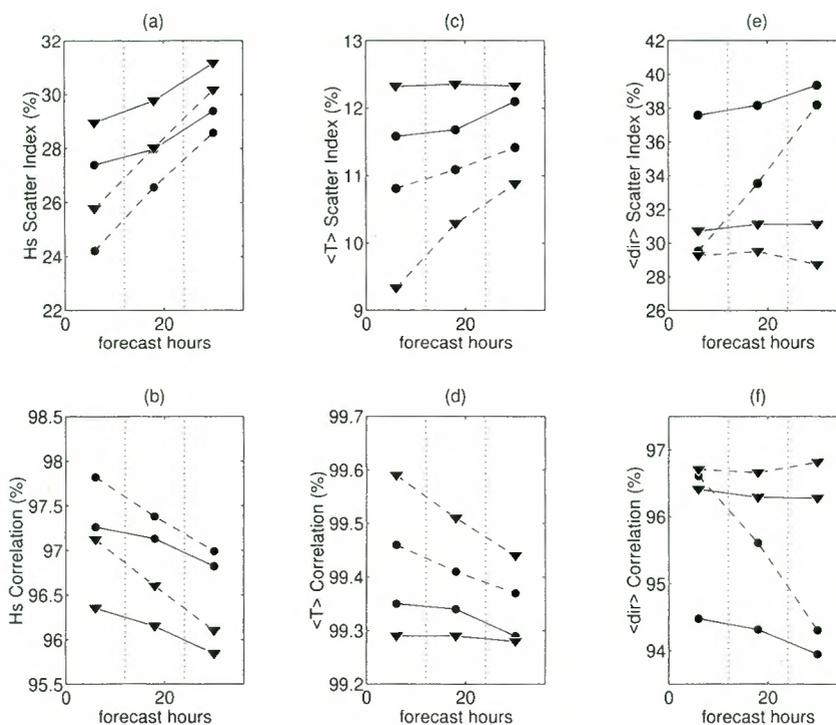


Figure 16. Scatter index and correlation with respect to wave parameters estimated from not-yet-assimilated SAR data, as a function of projection time for case 1 (triangles, 376 samples) and case 2 (circles, 749 samples). Results of the forecast following the baseline and assimilation runs are represented by solid line and dashed lines, respectively. Results are given for (a, b) significant wave height, (c, d) average period, and (e, f) vector mean direction.

The increase in negative bias in the altimeter comparisons, which also occurred for case 1, is attributed to the effect of systematic underestimating of wave heights by the ERS-1 altimeter, combined with the tendency of the SAR observations to overestimate the wave heights. We also note that the negative impact for the case 2 hindcast was reversed to a small positive impact for the forecast.

We attribute the small overall impact partly to the relative scarcity of the SAR data points compared with the number of model grid points. For the buoy data we also attribute the small overall impact to the scarcity of collocations of SAR data with the buoy observations and to the fact that the buoy observation locations tended to be upstream of the main wave propagation direction. Evaluation of the results of a single satellite pass indicated a locally larger positive impact with respect to buoy data when a near collocation occurred. The impact was most noticeable for the decaying stage of the waves and for wave period.

Our assessment of the persistence of the corrections indicated that the impact of assimilation decays with time but lasts for at least 36 hours into the forecast. This was also noted from comparisons of model results with the independent data. The corrections are relatively long lasting despite the fact that no attempt was made to adjust the wind field to agree with the modified wave field. This is consistent with the fact that corrections induced by SAR wave data mostly alter the low-frequency swell, for which the wind has little effect.

While the amount of wave data has increased dramatically to the point where routine wave data assimilation can now be considered worthwhile, data are still relatively scarce compared with what would be needed to have a really large impact on the model. The model estimates the spectrum at 2318 points over the North Atlantic, while in a 3-hour assimilation period, as many as 50 data points might be available. On average, the influence cannot be very large with that coverage rate. In fact, where there is particular interest, for example, in a storm situation, one has to be lucky to have a satellite overpass at the location and time of the storm.

Therefore, on the basis of the present results, we do not have enough evidence to conclude categorically that SAR data assimilation leads to improvement of wave analysis and forecast from WAM. Further study is needed on larger sample of collocated data. In particular, unbiased altimeter data would be helpful as an independent data source, especially if from the same satellite as SAR data. Directional buoy data will be useful for evaluation of the full spectrum. While additional validation data are needed, the inversion and assimilation methodology could also be improved in many ways. First, improvements can be made to the data quality control to lower the number of unsuccessful inversions and to help filter out non-wave field signals in the data. Second, the method of spreading corrections is very simple. The knowledge that the wave field is more highly correlated in the wave propagation direction can be built into the system through an anisotropic spreading weight function. However, it should also be feasible to move to a full optimum interpolation method, by obtaining the necessary error statistics for model and data. With the availability of increasingly large archives of higher quality wave observations, the necessary data are now becoming available. In the long run, the methodology can be optimized by using a full three-dimensional variational approach, but this will require dedication of greater amounts of computer power than is currently possible in operations.

Ultimately, the use of continuous insertion methods such as four-dimensional variational techniques will allow data from all sources to be blended with wave and atmospheric models to provide an optimal and consistent analysis of both waves and marine surface winds. For the future, we can hope that the availability of additional satellites (ERS-2, RADARSAT, and ENVISAT, for example) will result in an increase in data available to wave-modeling operations. Such an increase, along with a concurrent increase in the optimal use of the data in wave analysis and forecasting, will result in steady increases in accuracy of wave analyses and forecasts in the future.

Appendix: Statistical Parameters

Using S , f , θ , and x to refer to a spectrum, wave frequency, wave direction, and observed or modeled parameter, statistical parameters used in this paper are defined as follows.

$$\text{Mean} \quad \langle x \rangle = \frac{1}{N} \sum_{n=1}^N x_n \quad (\text{A1})$$

$$\text{Spectral mean} \quad \langle x \rangle_S = \frac{\sum_{i,j} x(f_i, \theta_j) S^{\text{WAM}}(f_i, \theta_j) \Delta f_i \Delta \theta_j}{\langle E \rangle_S} \quad (\text{A2})$$

$$\text{Mean energy} \quad \langle E \rangle_S = \sum_{i,j} S^{\text{WAM}}(f_i, \theta_j) \Delta f_i \Delta \theta_j \quad (\text{A3})$$

$$\text{Significant wave height} \quad H_s = 4 \sqrt{\langle E \rangle_S} \quad (\text{A4})$$

$$\text{Mean period} \quad \langle T \rangle_S \equiv \langle 1/f \rangle_S \quad (\text{A5})$$

$$\text{Vector mean direction} \quad \langle \text{dir} \rangle_S = \tan^{-1} \frac{\langle \sin(\theta) \rangle_S}{\langle \cos(\theta) \rangle_S} \quad (\text{A6})$$

$$\text{Mean wavelength} \quad \langle k \rangle_S = \frac{1}{g} \left(\frac{2\pi}{\langle T \rangle_S} \right)^2 \quad (\text{A7})$$

$$\text{Standard deviation} \quad \text{std} = \sqrt{\langle (x - \langle x \rangle)^2 \rangle} \quad (\text{A8})$$

$$\text{Bias} = \langle x - x_{\text{WAM}} \rangle \quad (\text{A9})$$

$$\text{Root mean square of the difference} \quad \text{rms} = \sqrt{\langle (x - x_{\text{WAM}})^2 \rangle} \quad (\text{A10})$$

$$\text{Standard deviation of the difference} \quad \text{std}_D = \sqrt{\langle (x - x_{\text{WAM}} - \text{bias})^2 \rangle} \quad (\text{A11})$$

$$\text{Scatter index} \quad \text{SI} = \frac{\text{std}}{\langle x \rangle} 100\% \quad (\text{A12})$$

Scatter index of the difference

$$Si_D = \frac{\text{std}_D}{\sqrt{\langle x \rangle \langle x_{WAM} \rangle}} 100\% \quad (\text{A13})$$

Symmetric regression (slope) coefficient

$$\langle \bar{c} \rangle = \frac{\sqrt{\langle x^2 \rangle}}{\sqrt{\langle x_{WAM}^2 \rangle}} = \frac{\text{rms}}{\text{rms}_{WAM}} \quad (\text{A14})$$

Symmetric correlation coefficient

$$\text{Cor} = \frac{\langle x x_{WAM} \rangle}{\sqrt{\langle x^2 \rangle \langle x_{WAM}^2 \rangle}} \quad (\text{A15})$$

Pattern correlation coefficient

$$K = \frac{\langle S_{SAR}^{obs} S_{WAM}^{sim} \rangle}{\sqrt{\langle (S_{SAR}^{obs})^2 \rangle \langle (S_{WAM}^{sim})^2 \rangle}} \quad (\text{A16})$$

Acknowledgments. This work was carried out with funding from the Radar Data Development Program, Canada Centre for Remote Sensing, Environment Canada, and Panel for Energy Research and Development. The WAM model and algorithms for spectral mode separation and cross-assignment of corresponding wave systems were supplied by Susanne Hasselmann at the Max-Planck-Institut für Meteorologie.

References

- Alpers, W., Monte Carlo simulations for studying the relationship between ocean wave and synthetic aperture radar image spectra, *J. Geophys. Res.*, **88**(C3), 1745–1759, 1983.
- Alpers, W. R., and C. Brüning, On the relative importance of motion-related contributions to the SAR imaging mechanism of ocean surface waves, *IEEE Trans. Geosci. Remote Sens.*, **GE-24**, 873–885, 1986.
- Bauer, E., S. Hasselmann, and K. Hasselmann, Validation and assimilation of Seasat altimeter wave heights using the WAM wave model, *J. Geophys. Res.*, **97**(C8), 12,671–12,682, 1992.
- Bauer, E., S. Hasselmann, and C. Brüning, Application of wave spectra retrieved from ERS-1 SAR wave mode spectra in a wind- and wave data assimilation system, in *Proceedings, International Geoscience and Remote Sensing Symposium (IGARSS'95)*, pp. 1643–1646, IEEE, Piscataway, N. J., 1995.
- Bao, M., C. Brüning, and W. Alpers, A generalized nonlinear integral transform and its application to ERS-1 SAR ocean wave imaging, in *Proceedings Second ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-361*, 219–224, 1994.
- Breivik, L. A., and M. Reistad, Use of ERS-1 altimeter wave products at DNMI—Evaluation of wave heights and wind speeds—Assimilation in a numerical wave model, *DNMI Report*, 23 pp., Det Norske Meteorol. Inst., Oslo, 1992.
- Brooker, G., UWA processing algorithm description, *ESA Doc. ER-TN-ESA-GS-0341*, Issue 1.1, Eur. Space Agency, Paris, Jan. 19, 1994.
- Brüning, C., and S. Hasselmann, Global statistics of integrated ocean wave parameters extracted from ERS-1 SAR wave mode image spectra, in *Proceedings, International Geoscience and Remote Sensing Symposium (IGARSS'93)*, Tokyo, pp. 351–353, IEEE, Piscataway, N. J., 1993.
- Brüning, C., W. Alpers, and K. Hasselmann, Monte-Carlo simulations of the nonlinear imaging of a two dimensional surface wave field by a synthetic aperture radar, *Int. J. Remote Sens.*, **11**, 1695–1727, 1990.
- Brüning, C., S. Hasselmann, K. Hasselmann, S. Lehner, and T. Gerling, On the extraction of ocean wave spectra from ERS-1 SAR wave mode image spectra, in *Proceedings of the First ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA-SP 359*, 747–752, 1993.
- Brüning, C., S. Hasselmann, K. Hasselmann, S. Lehner, and T. Gerling, First evaluation of ERS-1 synthetic aperture radar wave mode data, *Global Atmos. Ocean Syst.*, **2**, 61–98, 1994a.
- Brüning, C., R. Schmidt, and W. Alpers, Estimation of the ocean wave-radar modulation transfer function from synthetic aperture radar imagery, *J. Geophys. Res.*, **99**(C5), 9803–9815, 1994b.
- Cressman, G. P., An operational objective analysis system, *Mon. Weather Rev.*, **87**, 367–374, 1959.
- de las Heras, M. M., and P. A. E. M. Janssen, Data assimilation with a coupled wind-wave model, *J. Geophys. Res.*, **97**(C12), 20,261–20,270, 1992.
- Engen, G., H. Johnsen, H. E. Krogstad, and S. F. Barstow, Directional wave spectra by inversion of ERS-1 synthetic aperture radar ocean imagery, *IEEE Trans. Geosci. Remote Sens.*, **32**, 340–352, 1994.
- Francis, P. E., and P. P. Stratton, Some experiments to investigate the assimilation of Seasat altimeter wave height data into a global wave model, *Q. J. R. Meteorol. Soc.*, **116**, 1225–1251, 1990.
- Gandin, L. S., *Objective Analysis of Meteorological Fields*, 242 pp., Isr. Program for Sci. Transl., Jerusalem, 1963.
- Gerling, T. W., Partitioning sequences and arrays of directional wave spectra into component wave systems, *J. Atmos. Oceanic Technol.*, **9**, 444–458, 1992.
- Graber, H., V. Cardone, R. Jensen, S. Hasselmann, H. L. Tolman, and L. Cavaleri, The accuracy of wind field description, in *Dynamics and Modeling of Ocean Waves*, Chap. IV.3.2, pp. 285–293, Cambridge Univ. Press, New York, 1994.
- Hasselmann, K., and W. Alpers, The response of synthetic aperture radar to ocean surface waves, in *Wave Dynamics and Radio Probing of the Ocean Surface*, edited by D. M. Phillips and K. Hasselmann, pp. 393–400, Plenum, New York, 1986.
- Hasselmann, K., and S. Hasselmann, On the nonlinear mapping of an ocean wave spectrum into a synthetic aperture radar image spectrum and its inversion, *J. Geophys. Res.*, **96**, 10,713–10,729, 1991.
- Hasselmann, K., R. K. Raney, W. J. Plant, W. Alpers, R. A. Shuchman, D. R. Lyzenga, C. L. Rufenach, and M. J. Tucker, Theory of synthetic aperture radar ocean imaging: A MARSEN view, *J. Geophys. Res.*, **90**(C3), 4659–4686, 1985.
- Hasselmann, S., C. Brüning, and P. Lionello, Towards a generalized optimal interpolation method for the assimilation of ERS-1 SAR retrieved wave spectra in a wave model, in *Proceedings Second ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-361*, 21–25, 1994.
- Hasselmann, S., C. Brüning, K. Hasselmann, and P. Heimbach, An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra, *J. Geophys. Res.*, **101**(C7), 16,615–16,629, 1996.
- Jacobsen, S., and K. A. Høgda, Estimation of the real aperture radar modulation transfer function directly from synthetic aperture radar ocean wave image spectra without a priori knowledge of the ocean wave height spectrum, *J. Geophys. Res.*, **99**(C7), 14,291–14,302, 1994.
- Jain, A., SAR imaging of ocean waves, Theory, *IEEE J. Oceanic Eng.*, **OE-6**, 130–139, 1981.
- Janssen, P. A. E. M., Quasi-linear theory of wind-wave generation applied to wave forecasting, *J. Phys. Oceanogr.*, **21**, 1631–1642, 1991.
- Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen, *Dynamics and Modeling of Ocean Waves*, Cambridge University Press, New York, 1994.
- Krogstad, H. E., A simple derivation of Hasselmann's nonlinear ocean-synthetic aperture radar transform, *J. Geophys. Res.*, **97**(C2), 2421–2425, 1992.
- Krogstad, H. E., O. Samset, and P. W. Vachon, Generalization of the non-linear ocean-SAR transform and a simplified SAR inversion algorithm, *Atmos. Ocean*, **32**, 61–82, 1994.
- Lasnier, P., H. Hajji, S. Looul, D. Bonicel, and B. Chapron, Use of SAR-wave mode, altimeter and wind scatterometer data in operational wave forecasting system, in *Proceedings Second ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-361*, 21–25, 1994.
- Lionello, P., H. Günther, and P. A. E. M. Janssen, Assimilation of altimeter data in a global third-generation wave model, *J. Geophys. Res.*, **97**(C9), 14,453–14,474, 1992.
- Lionello, P., H. Günther, and B. Hansen, A sequential assimilation scheme applied to global wave analysis and prediction, *J. Mar. Syst.*, **6**, 87–107, 1995.
- Mailhot, J., R. Sarrazin, B. Bilodeau, N. Brunet, A. Methot, G. Pellerin, C. Chouinard, L. Garand, C. Girard, and R. Hogue, Changes to the Canadian regional forecast system: Description and evaluation of the 50 km version, *Atmos. Ocean*, **33**, 55–80, 1995.

- Raney, R. K., and R. T. Lowry, Oceanic wave imagery and wave spectra distortions by synthetic aperture radar, paper presented at 12th International Symposium on Remote Sensing of the Environment, Environ. Res. Inst. of Mich., Manila, April 20–26, 1978.
- Talagrand, O., and P. Courtier, Variational assimilation of meteorological observations with the adjoint vorticity equation, I, Theory, *Q. J. R. Meteorol. Soc.*, *113*, 1311–1328, 1987.
- Vachon, P. W., R. K. Raney, and W. J. Emery, A simulation for spaceborne SAR imagery of a distributed, moving scene, *IEEE Trans. Geosci. Remote Sens.*, *27*, 67–78, 1989.
- Vachon, P. W., R. B. Olsen, H. E. Krogstad, and A. K. Liu, Airborne synthetic aperture radar observations and simulations for waves in ice, *J. Geophys. Res.*, *98*(C9), 16,411–16,425, 1993.
- Vachon, P. W., H. E. Krogstad, and J. S. Paterson, Airborne and spaceborne synthetic aperture radar observations of ocean waves, *Atmos. Ocean*, *32*, 83–112, 1994.
- Wave Model Development and Implementation (WAMDI) Group (S. Hasselmann et al.), The WAM model—A third generation ocean wave prediction model, *J. Phys. Oceanogr.*, *18*, 1775–1810, 1988.
- Zambreski, L., A verification study of the global WAM model, December 1987–November 1988, *ECMWF Tech. Rep. 63*, Eur. Cent. for Medium-Range Weather Forecasts, Reading, England, 1989.
-
- E. M. Dunlap and S. De Margerie, ASA Consulting Ltd., 1888 Brunswick Street, Suite 400, Halifax, Nova Scotia, Canada B3J 3J8. (e-mail: edunlap@esa.ca)
- R. B. Olsen, Satlantic, Inc., Richmond Terminal, Pier 9, 3295 Barrington Street, Halifax, Nova Scotia, Canada B3K 5X8. (e-mail: richard@satlantic.com)
- R. Lalbeharry and L. Wilson, Atmospheric Environment Service, Environment Canada, 2121 Trans-Canada Highway, Suite 500, Dorval, Quebec, Canada H9P 1J3.

(Received January 3, 1997; revised August 27, 1997; accepted September 8, 1997.)

Tracking storm-generated waves in the northeast Pacific Ocean with ERS-1 synthetic aperture radar imagery and buoys

Benjamin Holt,¹ Antony K. Liu,² David W. Wang,³ Anand Gnanadesikan,⁴ and H. S. Chen⁵

Abstract. This paper examines the capability of synthetic aperture radar imagery from ERS-1 and buoys to track the wave field emanating from an intense storm over a several-day period. The first part of the study is a validation component that compares SAR-derived wave length and direction with buoy data from two locations over 10 different dates in late 1991 and National Oceanic and Atmospheric Administration (NOAA) wave model (WAM) wave direction results. When the SAR is linear (8 out of 10 cases), mean wavelength is within 5% of the buoy measurements and mean wave direction is within 1° of direction derived from the wave model (albeit with a large standard deviation of 27°), indicating close agreement. The wave field generated from the intense storm in late December 1991 was measured by three separate ERS-1 SAR passes over a 3-day period. A simple kinematic model was used for waves propagating from a storm. Comparing the model results with both SAR and buoy data indicates that SAR-derived peak wavelength and direction measurements can be reliably used to predict arrival times and propagation direction over a several-day period and considerable distances. The measurements can also be used to derive estimated wave generation source regions about the storm as well. Such measurements are useful for comparing with wave model results, which perform less accurately for direction than wave height for example, and for predicting hazardous conditions for ship navigation and coastal regions.

1. Introduction

The tracking of storms and their associated wave fields is of interest both scientifically, for understanding wind-wave generation physics and its inclusion into wave models, and operationally for monitoring potentially hazardous conditions for shipping routes and coastal environments. Early studies showed that swell could be accurately identified as emanating from storms thousands of kilometers away [Barber and Ursell, 1948; Munk et al., 1963; Snodgrass et al., 1966]. Current third-generation wave models (WAM) attempt to predict wave conditions based on integrated wind fields assimilated from satellite imagery and atmospheric soundings and much improved wind and wave physics [Wave Model Development and Implementation (WAMDI) Group, 1988; Komen et al., 1994]. While considerable improvements have been made in predicting wave height from severe storms and rapidly rotating winds since the report by the Sea Wave Modeling Project (SWAMP) Group [1985], problems still exist in predicting the directional wave energy distribution. Some of these problems are due to

the model resolution of propagation direction (generally 30°), which will improve with increased computer capabilities. Others are due to difficulties in determining the source regions of high storms, found to be located in the lower right quadrant of a moving storm. Waves generated from severe storms approaching landfall clearly can lead to potentially dangerous conditions for coastal inhabitants, and directional accuracy is of prime importance [Earle et al., 1984; Wang and Carolan, 1991; Mettlach et al., 1994]. The latter study evaluated both nondirectional and directional buoy spectra as well as wave model hindcasts together with a kinematic model of wave propagation to determine a possible approach for automating the estimation of the wave source regions for a large storm to help with such predictions of large waves approaching the coast. It was found that established networks of wave buoys in the northern Pacific, usually relatively close to shore, can provide valuable and near-real time data on approaching wave fields, but only under certain conditions. The use of satellite imagery especially from synthetic aperture radar (SAR) can generally improve wave directional information and prediction accuracy [Beal, 1991; Komen et al., 1994].

Satellite imagery from SAR provides a unique two-dimensional, fine-resolution (usually 25 m) view of the ocean surface, especially waves. From this view, wavelength, propagation direction, and height information can be derived. One of the key applications of SAR has been in providing valuable information on the properties of storm-generated ocean waves. When the SAR is operating linearly over the ocean, it has been shown to provide accurate measurements on the evolution of wave fields [Beal et al., 1986; Beal, 1991], the spatial properties of hurricane-generated waves [McLeish and Ross, 1983; Holt and Gonzalez, 1986; Monaldo et al., 1993], the improvement in wave information surrounding storms compared to wave mod-

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena.

²Oceans and Ice Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland.

³Applied Technology Division, Computer Sciences Corporation, Stennis Space Center, Mississippi.

⁴Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey.

⁵National Centers for Environmental Prediction, NOAA, Washington, D.C.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02567.
0148-0227/98/97JC-02567\$09.00

els [Beal, 1991; Monaldo and Beal, 1996], and the source regions of storm-generated waves [Gonzalez et al., 1987]. However, the key word here is "linear," which occurs primarily when the SAR is detecting swell or when the radar platform has a low shuttle-like orbital altitude (200–400 km). The problem is that SAR imagery of ocean waves and derived SAR wave spectra can be severely nonlinear because of the simple fact that the waves are moving at the same time the radar is precisely measuring the Doppler frequency of the moving SAR platform.

The nonlinear conditions occur primarily when waves are traveling in a direction parallel (or in azimuth traveling) to the platform flight direction. This simultaneous wave-platform motion can result in a distorted mapping of waves to imagery because the precisely timed radar returns from the waves have slightly varying Doppler frequencies and rates, termed velocity bunching. The result is either a shifting of the wave energy or actual nonimaging of wind waves moving in the azimuth direction. In addition, the SAR resolution may be degraded by coherence time limitation of the moving smaller waves [Raney, 1980]. Both conditions are alleviated by decreasing the ratio of the platform range-to-target compared to platform velocity. Many researchers have long sought to understand the nonlinearities sufficiently in order to correct for these by applying transfer functions [e.g., Hasselmann et al., 1985; Plant, 1992; Krogstad, 1992; Plant and Zurk, 1997]. Understanding has progressed sufficiently to the point that SAR image spectra from the European Space Agency's ERS systems are being used reliably, albeit at considerable computational cost, in SAR wave data assimilation schemes for global wind and wave model forecasts [Hasselmann and Hasselmann, 1991; Engen et al., 1994; Komen et al., 1994; Hasselmann et al., 1996].

This study examines the capability of spaceborne synthetic aperture radar imagery, obtained from ERS-1, and buoy data to track a wave field generated by an intense storm in the northeast Pacific over a several-day period. The SAR imagery and radar wave spectra were obtained from the Alaska SAR Facility (ASF) during September–December 1991. First, we establish the accuracy of the ERS-1 SAR spectra, produced by the ASF geophysical processor system (ASF GPS), for deriving the dominant wavelength and wave direction. This was accomplished by comparing 10 separate examples of SAR data nearly coincident in time and location to two sets of buoy data, one set operated by the National Data Buoy Center (NDBC) and another set deployed during a field experiment off Vancouver Island run by the Woods Hole Oceanographic Institution [Galbraith et al., 1994], plus NDBC buoy-modeled wave direction from the National Oceanic and Atmospheric Administration (NOAA) wave model (WAM) [Chen, 1995a]. The validation segment determined that the SAR-derived dominant wave components were largely linear, which is related to the mean climatological conditions of the northern Pacific and the relative orientation of the orbital track. From this the tracking of the storm waves could proceed with some confidence. The results from the SAR wave spectra are compared with buoy data for waves generated during a storm in late December 1991. For the case study we describe a wave kinematic model, the storm, and then track the wave field over a 3-day period.

2. Data Set Description

2.1. Radar Imagery

The ERS-1 SAR imagery was acquired in the Gulf of Alaska along the general tracks outlined on Figure 1 from September–

December 1991 during the 3-day repeat of the mission's commissioning phase. ERS-1 is in a Sun-synchronous orbit of about 98° and has a flight direction (azimuth angle) of about 196°T over the Gulf of Alaska during descending passes. The SAR operates at a frequency of 5.3 GHz (C band, 5-cm wavelength) over a fixed range of incidence angles from 20°–26°, which results in a swath width of 100 km. The resolution of the processed imagery is 30 m in both range and azimuth directions, and the data are formatted into 12.5-m pixels. The imagery was acquired, processed, and distributed by the ASF, which is located at the University of Alaska in Fairbanks.

The SAR wave spectra were generated from the full-resolution imagery using a wave product algorithm in the ASF geophysical processor system (ASF GPS), which also generated ice motion and ice classification products [Stern et al., 1994]. The wave algorithm has been previously tested using Seasat imagery [Wadhams and Holt, 1991]. For each image (100 by 100 km) the wave product algorithm performs a two-dimensional digital fast Fourier transform on 16 subscenes, each 6.4 by 6.4 km in size, which are contiguous in the azimuth direction and centered along the median range line. The resulting unsmoothed spectral density estimate has a spectral resolution of $\Delta k = (1/256) 2\pi/25 \text{ rad m}^{-1} = 0.001 \text{ rad m}^{-1}$ and a chi-square distribution with 2 degrees of freedom [Monaldo, 1991]. To reduce sampling variability, each spectrum is smoothed using a moving Gaussian filter with a full width of 21 by 21 pixels and a kernel size of 5 pixels. The smoothing increases the degrees of freedom to 164 on the basis of the effective area of the filter [Beal et al., 1986]. A peak-finding routine locates the dominant local maxima or wave peaks and determines the wavelength and direction of the primary wave components by their distance and orientation from the spectra center. The radar spectra are displayed as contour plots, as seen in Figure 2a. Wave direction has a 180° ambiguity, which is generally resolved by examining weather data. For this analysis the mean and standard deviation of the wavelength and wave direction of the dominant wave peaks for a single scene are obtained from all 16 subscenes, which assumes that the wave field is approximately homogeneous over a 100-km image frame. No other corrections have been made to the SAR imagery for either system or modulation transfer functions. As further described by Wadhams and Holt [1991], on the basis of analysis by Monaldo [1991] and the characteristics of the smoothing filter the theoretical SAR wave spectral precision is $\pm 2^\circ$ in wave direction and $\pm 0.002 \text{ rad m}^{-1}$ in wavenumber.

2.2. NDBC Buoy Data

Wave data from three separate National Data Buoy Center (NDBC) moored buoys were utilized in this study (Figure 1). Buoys 46001 and 46003 are 6-m Navy Oceanographic and Meteorological Automatic Device (NOMAD) deep ocean buoys in the Gulf of Alaska. Data used from these buoys include wave period, significant wave height (H_s), wind speed and direction (Figure 3), and barometric pressure. All of the NOAA data used in the validation portion of the study were derived from buoy 46001 since it was closest in location to the SAR orbital tracks. A third buoy used is 46042, a 3-m coastal wave buoy which also provides wave direction. All wave data are averaged over 20-min periods. The accuracies of the wave parameters are as follows: significant wave height, $\pm 0.2 \text{ m}$ or 5%; wave period, $\pm 1 \text{ s}$; and wave direction, $\pm 5^\circ$.

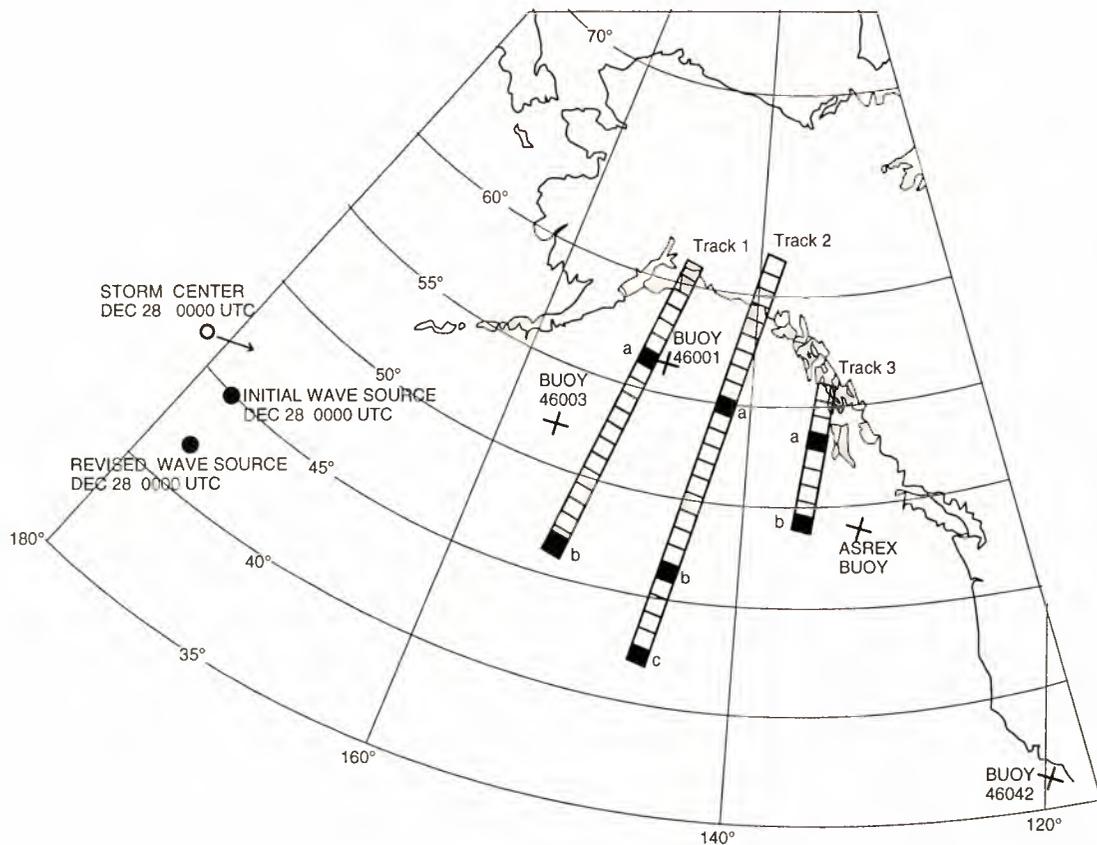


Figure 1. Map of the northeast Pacific Ocean showing the location of buoys from the National Data Buoy Center (NDBC) (46003, 46001, and 46042) and the Acoustic Surface Reverberation Experiment (ASREX) and representative synthetic aperture radar (SAR) tracks for both the validation and storm-tracking sections with solid boxes indicating SAR spectra locations.

2.3. Field Experiment Buoy Data

During the late fall and early winter of 1991 an array of buoys was deployed off the coast of Vancouver (49.2°N, 131.9°W) as part of the 1991 Acoustic Surface Reverberation Experiment (ASREX 91) (Figure 1) [Galbraith *et al.*, 1994]. Two of the moorings were waverider buoys, the first a Wavescan from Seatex AS of Norway and the other a Wavetrack from Endeco/YSI, Inc., of Marion, Massachusetts. The Seatex buoy obtains 2048 data points at 1 Hz every 3 hours, while the Endeco buoys obtained 2048 points at 2 Hz every 12 hours. The Seatex operated from November 1 to December 4, 1991, and was used in the validation segment of this study, while the Endeco buoy operated through the end of December 1991, providing data for the storm-tracking segment of the study. Also included in the mooring array was a series of meteorological measurements. A time series of wind speed and direction, significant wave height, and peak wave period derived from Seatex buoy measurements during its period of operation is shown in Figure 3.

2.4. Wave Model Data

Wave hindcast (3-hour interval) data were also obtained from the NOAA National Centers for Environmental Prediction (NCEP) using the operational NOAA WAM [Chen, 1995a, b; WAMDI Group, 1988], which provides estimates of wave direction along with the four other parameters also obtained from the NDBC deep water buoys. Wave direction from these hindcasts was used in five cases for SAR validation,

supplementing the measurements from the nondirectional buoy 46001. The NOAA WAM has been used for operational forecasting since 1994 after several years of development and comparison studies with the NDBC buoy data. The model is essentially based on the cycle 4 WAM, which incorporates both wind velocity and wave age for wave generation source functions, and a quasi-linear theory of wind-wave generation [Janssen, 1989, 1991]. The model runs twice daily to predict global ocean wave spectra for both 12-hour hindcasts and 72-hour forecasts. The model has a grid of 2.5° in latitude and longitude. The wave spectrum is represented by 25 logarithmically spaced frequencies with the ratio of frequency increment to its frequency being equal to 0.1 and wave directionality in 30° bins. The lowest sigma layer winds from the NCEP analysis and aviation version of the global forecast system are adjusted to a height of 10 m by using a logarithmic profile [Shearman and Zelenko, 1989] and are used to drive the ocean surface waves. The analysis wind is for wave hindcasts, and the aviation wind is for wave forecasts.

3. Validation of SAR-Derived Wave Measurements

3.1. SAR-Buoy Comparisons

Before using the SAR wave spectra for the storm-tracking case study an assessment of the measurement accuracy of the SAR-derived dominant wavelength and wave direction was performed using comparisons with NDBC buoy 46001 and the

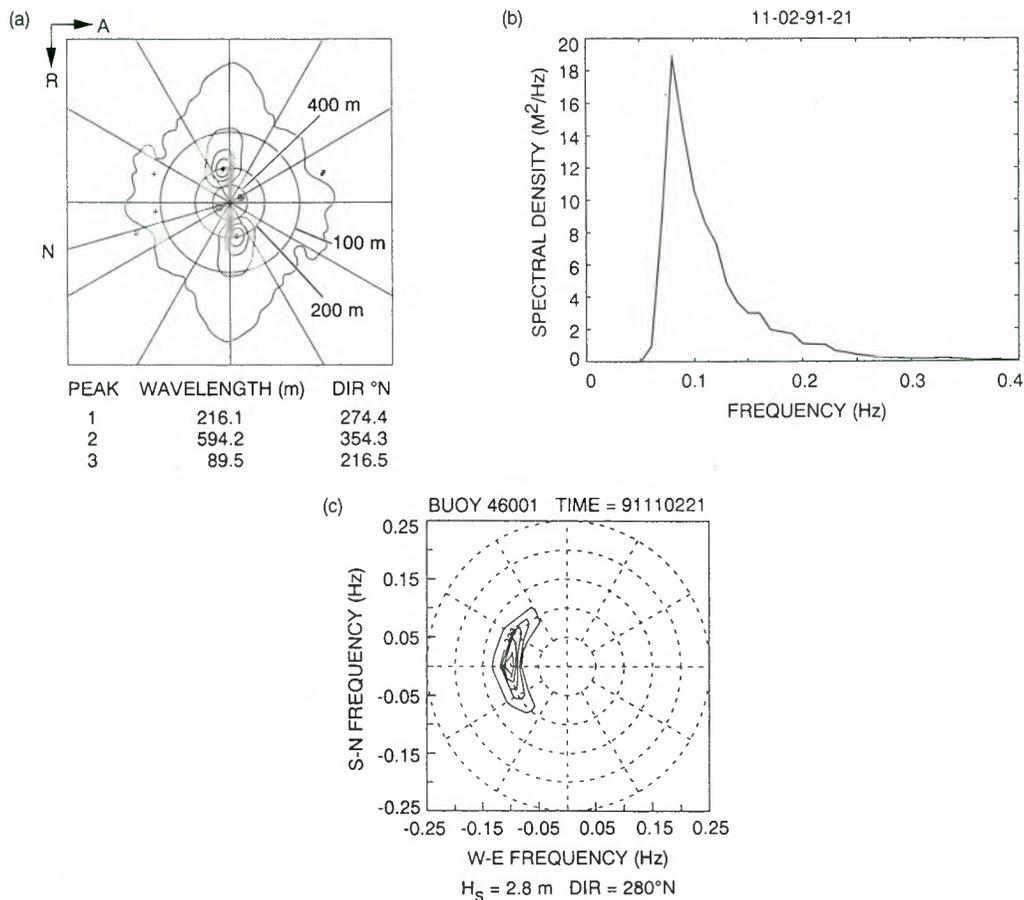


Figure 2. Comparison from November 2, 1991, at 2100 UTC of the (a) Alaska SAR Facility wave product spectrum from ERS-1 SAR and (b) buoy frequency spectrum and (c) National Oceanic and Atmospheric Administration (NOAA) wave model (WAM) results for NDBC buoy 46001.

WHOI data sets over varying environmental conditions. Ten separate buoy data sets coincident with ERS-1 SAR were utilized, five for buoy 46001 and five for the ASREX Seatex buoy. Table 1 lists the dates and time of each SAR pass. The environmental conditions for each date are detailed in Appendix A. For the comparisons with buoy 46001, buoy wave direction was derived from the WAM4 hindcasts. The SAR wave spectra were obtained from the SAR image frames closest to the two buoy locations, track 1, frame a for 46001 and, generally, track 3, frame b for Seatex (Figure 1). The distances between the buoys and the SAR frames were variable, especially for the ASREX buoys, where the closest distance was at best 135 km. The key wave parameters for the SAR and buoys are listed in Table 1. Buoy wavelength λ is derived from peak wave period T using the deep water wave dispersion relationship

$$\lambda = (T^2g)/2\pi. \quad (1)$$

Figure 2a shows a representative SAR wave spectrum collected on November 2, 1991, which has a dominant wavelength of 216 m and propagation direction from 274°. The frequency spectra from buoy 46001 is shown in Figure 2b and the hindcast results of the NOAA WAM are shown in Figure 2c. The buoy results show the dominant wave period of about 12.4 s (0.081 Hz), corresponding to a wave length of 239 m, and the WAM results show a propagation direction from 280°. For the dominant peak wave component the SAR spectrum compares reasonably well with the buoy results; however, the SAR spectral

shape is slightly distorted in the azimuth direction by the radar transfer function as compared to the model results.

Referring back to Table 1, it can be seen that of the 10 cases, 8 comparisons between the SAR measurements and the buoy measurements are favorable, while 2 cases, October 15 and November 16, show considerable differences in wavelength and wave direction. We believe these differences are due to nonlinearities in the SAR spectra resulting from azimuthal distortion and so are not considered in the following error calculations (more discussion below). Figure 4 shows the comparison of the eight linear peak SAR-derived wavelengths and buoy wave periods using the dispersion relation (1). Here we see that the SAR-derived wavelengths follow the relation but are slightly underestimated.

For the eight cases the mean fractional difference in wavelength is -5.1% with a standard deviation of 9.6% , and the mean difference is -22 m with a standard deviation of 31 m. Using the wavenumber precision of 0.002 rad m^{-1} [Monaldo, 1991] over the approximate range of measured buoy wavelengths (100–600 m), the theoretical accuracy in the wavelength measurements are between 3 and 20%, respectively. This average compares well with the mean absolute difference of 9.6% (29.4 m), as does the range of accuracies over the wavelength spread, since the largest discrepancies between buoy and SAR are generally those cases with the longest wavelengths. For direction the mean difference is -0.8° as compared with the theoretical error of $\pm 2^\circ$ [Monaldo, 1991], but

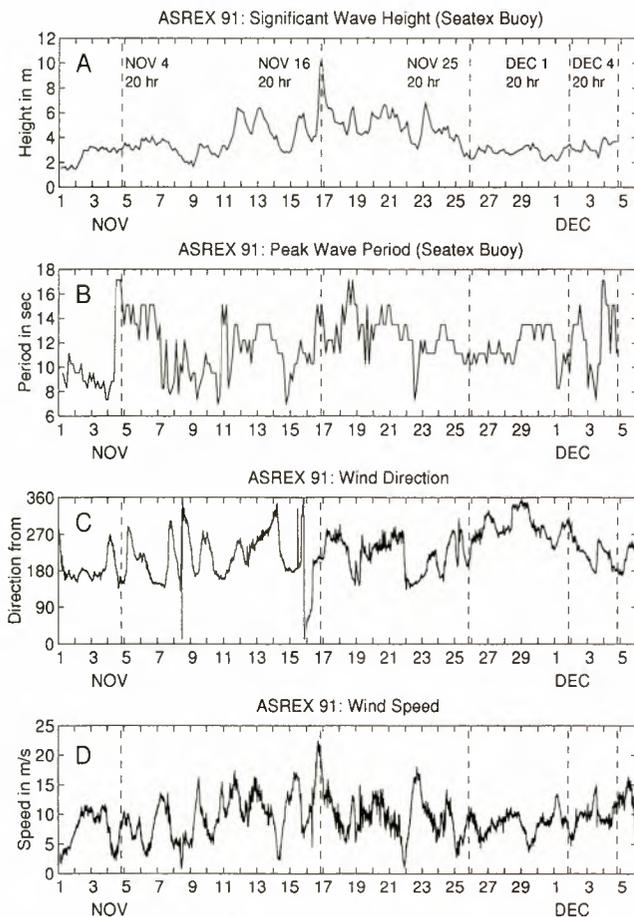


Figure 3. A time series from the ASREX Seatex buoy showing (a) significant wave height H_s , (b) peak wave period T , (c) wind direction, and (d) wind speed from November 1 to December 5, 1991. The times of the five coincident SAR data are indicated by vertical dashed lines.

there is a wide standard deviation of 27.5°. Most of the differences in direction arise from the single data point for the storm in late December, which is the storm discussed in the case study. From these two assessments the mean errors in wavelength and wave direction fall within the expected ranges when the SAR wave spectra are linear.

To assess the tendency of the SAR spectra to be linear or nonlinear in the case of the two unfavorable measurements, we use the velocity bunching parameter,

$$C = (1/4)(R/V) H_s g^{1/2} 2\pi\lambda^{-3/2} G(\theta, \phi) \cos \phi \quad (2)$$

where R is slant range from radar to the surface, V is platform velocity, H_s is significant wave height, g is the gravitational constant, and G is the platform and ocean wave geometry (θ is radar incidence angle and ϕ is the azimuth angle between the platform and wave direction) [Alpers, 1983]. Significant nonlinearities occur if $C > \pi/2$, indicating that shifts of spectral peaks toward lower wavenumbers are likely to occur. More specifically, nonlinear mapping results from a surface scatterer having a velocity component radial to the SAR look direction. The resultant displacement of the scatterer on the SAR image is in the azimuth or along-track direction (in all cases the spacecraft heading is about 195°T). Random shifts cause the image to be smeared in the azimuth direction, which causes a decrease in response for high-azimuth wavenumbers in the image spectra. For ERS-1 the R/V ratio is about 115 s. As shown in Table 1, values of C near 2 or greater are obtained for the two nonlinear cases on October 15 and November 16 because of the comparatively short wavelength and large H_s in both cases. The remaining eight cases have values generally much less than 1.5. The one other case with azimuth-traveling waves, October 24, shows reasonable comparisons in wavelength and wave direction or no obvious distortions in the wave spectra, despite a generally high C value.

Table 1. Comparison of SAR and Buoy Wave Parameters

Buoy	Track	Position	Date	Time, UTC	Distance, km	SAR Data			Buoy Data					Δ direction, deg		
						SAR λ , m	SAR Direction, deg	C	T , s	λ , m	Direction, deg	H_s , m	U , m s ⁻¹		$\Delta\lambda$, %	
46001	1	a	Sept. 18	2100	58	179	252	1.09	11	189	230	3.0	8	-5	22	
46001	1	a	Oct. 15	2100	38	157	265	2.28	10	156	140	4.0	12	
						196	134									
46001	1	a	Oct. 24	2100	38	103	195	1.39	8.3	107	180	1.0	4	-4	15	
46001	1	a	Nov. 2	2100	25	216	274	0.21	12.4	239	280	4.5	8	-10	-6	
46001	1	a	Dec. 29	2100	25	565	240	0.22	20.0	624	300	6.0	14	-10	-60	
Seatex	3	b	Nov. 4	2000	198	327	261	0.25	15.7	383	264	3.1	8	-15	-3	
Seatex	3	a	Nov. 16	2000	228	195	288	1.97	14.6	332	238	10	21	
						536	175									
Seatex	3	b	Nov. 25	2000	135	175	307	0.24	10.3	164	278	2.3	8	7	29	
Seatex	3	b	Dec. 1	2000	135	182	291	0.02	10.3	164	286	3.4	5	11	5	
Seatex	3	b	Dec. 4	2000	202	296	273	0.08	15.0	350	281	3.5	8	-15	-8	
Mean Δ															-5.1	-0.8
Standard deviation															9.6	27.5
Mean absolute Δ															9.6	18.5
Standard deviation															4.1	19.1

Track and position are of the SAR image, see Figure 1; Distance is from the buoy to the SAR image; SAR λ is the SAR-measured wavelength of dominant peak; SAR Direction is the SAR-measured dominant peak wave direction from source; C is the velocity bunching parameter, reference is Alpers [1983]; T is the buoy peak wave period; λ is derived from T using dispersion relationship (1); Direction is the wave direction from source, derived from NOAA/WAM hindcast; H_s is the significant wave height; U is the wind speed; $\Delta\lambda$ is the fractional difference in wavelength between buoy and SAR; and Δ direction is the difference in wave direction between buoy and SAR. ERS-1 flight (azimuth) direction is 196°T for all images. For Seatex direction, 21° are added for magnetic correction. NDBC 46001 location is 56.3°N, 148.3°W. WHOI ASREX Seatex location is 49.15°N, 131.89°W.

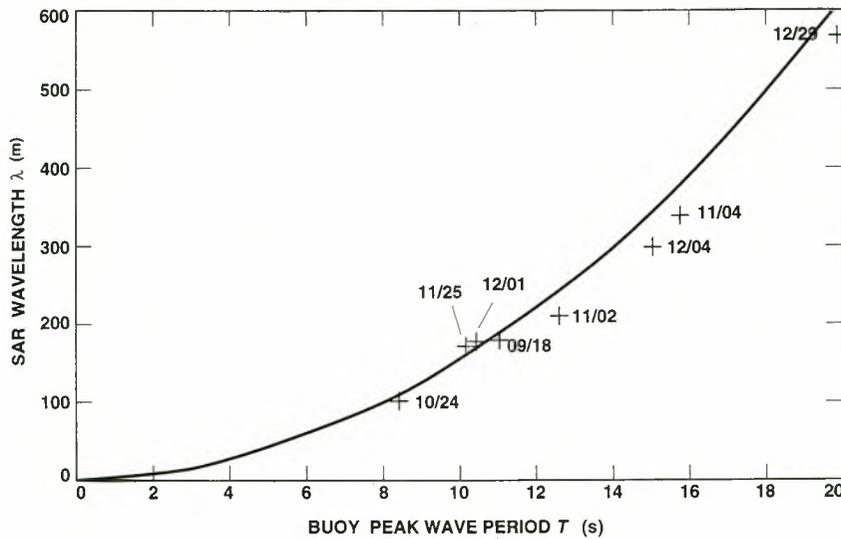


Figure 4. Comparison of measurements of peak wave period from buoy 46001 and the Seatex buoy with wavelength from ERS-1 SAR derived using the wave dispersion relation (1). See Table 1 for details.

3.2. Comparison With Other SAR Wave Validation Studies

During the early months of ERS-1 in 1991 several validation campaigns took place to assess the capability of the SAR for generating useful wave information. The Grand Banks validation program, summarized by *Dobson and Vachon* [1994], had the following key results: ERS-1 SAR accurately measured long wavelength swell, the SAR did not measure any azimuth-traveling waves shorter than 200 m, but range-traveling waves were measured as short as 50 m [*Vachon et al.*, 1994]. For the winter conditions seen in the Grand Banks area, a majority of the SAR acquisitions were obtained of azimuth-traveling waves. Quasi-linear remapping was more effective than non-linear remapping to invert the problematic SAR spectra [*Krogstad et al.*, 1994]. SAR and wave models were found to be complementary, since SAR was best at mapping swell and the wave model was best at mapping wind seas. Similar results were found in other validation experiments [*Kleijweg and Greidanus*, 1993; *Tilley and Beal*, 1994]. These results were essentially as predicted, based on years of analysis using SAR imagery from satellite, shuttle, and airborne platforms. To make full use of the ERS-1 SAR wave mode data, where small 5-km vignettes are sampled every 100–200 km on a global basis, a nonlinear inversion algorithm has been developed, where the SAR spectra are remapped in an iterative fashion together with WAM [*Hasselmann and Hasselmann*, 1991]. Results indicate that this is effective in producing accurate measurements of wavelength and wave direction as well as height [*Komen et al.*, 1994].

3.3. Validation Summary

This validation component showed that SAR wave spectra derived using a simple filter and no transfer functions were found to be reasonably accurate in measurements of peak wavelength and wave direction. This was largely fortuitous because of the meteorological conditions found in the northeast Pacific at least during the months September through December 1991, when west to east swell (or range-traveling swell relative to the ERS-1 flight direction) was predominantly measured, and, hence, accurately mapped onto the SAR wave spectra. This dominant weather pattern is attributed to the

semipermanent Aleutian Low, which is particularly intense during these early winter months [*Overland and Heister*, 1980]. One circumstance, on October 24, did accurately measure short, azimuth-traveling waves, presumably due to the low wave height. The two cases where comparisons with buoy data were poor, October 15 and November 16, were due to the presence of variable wind seas and the proximity of the buoy to the storm center, respectively. Another case, in late December, resulted in large differences in wave direction, likely due to difficulties in modeling directional data from large storms, as will be discussed in the next section. Thus, under favorable conditions, these products can be reliably used for further investigations in this general region.

4. Case Study of Storm and Swell Evolution

In this section we examine the evolution of swell across the northeastern Pacific generated by a storm using SAR, buoys, and a kinematic wave model. The data set was identified by the fortuitous combination of a strong low-pressure system, suitable SAR coverage of the emanating wave field, and buoy data both from NDBC and ASREX. We use a simple kinematic model derived by *Gonzalez et al.* [1987] for comparison with the various buoy wave measurement sources. The synoptic weather chart of the storm on December 28, 1991, at 0000 UTC is shown in Figure 5, and a location map of the ERS-1 SAR tracks and buoys is shown in Figure 1.

4.1. Kinematic Wave Model

Using SAR wave spectra of a hurricane-generated wave field obtained from a shuttle-based (and hence lower-altitude) platform and a simple kinematic wave model, *Gonzalez et al.* [1987] have shown that swell generally obeys the linear wave theory of propagation and does not seem to be affected by propagation through zones of steady wind. It is assumed that the waves originate from approximately a point source or a generation region sufficiently far away to be considered as a point source. Swell (wavenumber k) generation is defined as that point when the deep water group velocity

$$C_g = 1/2(gk)^{1/2} \quad (3)$$

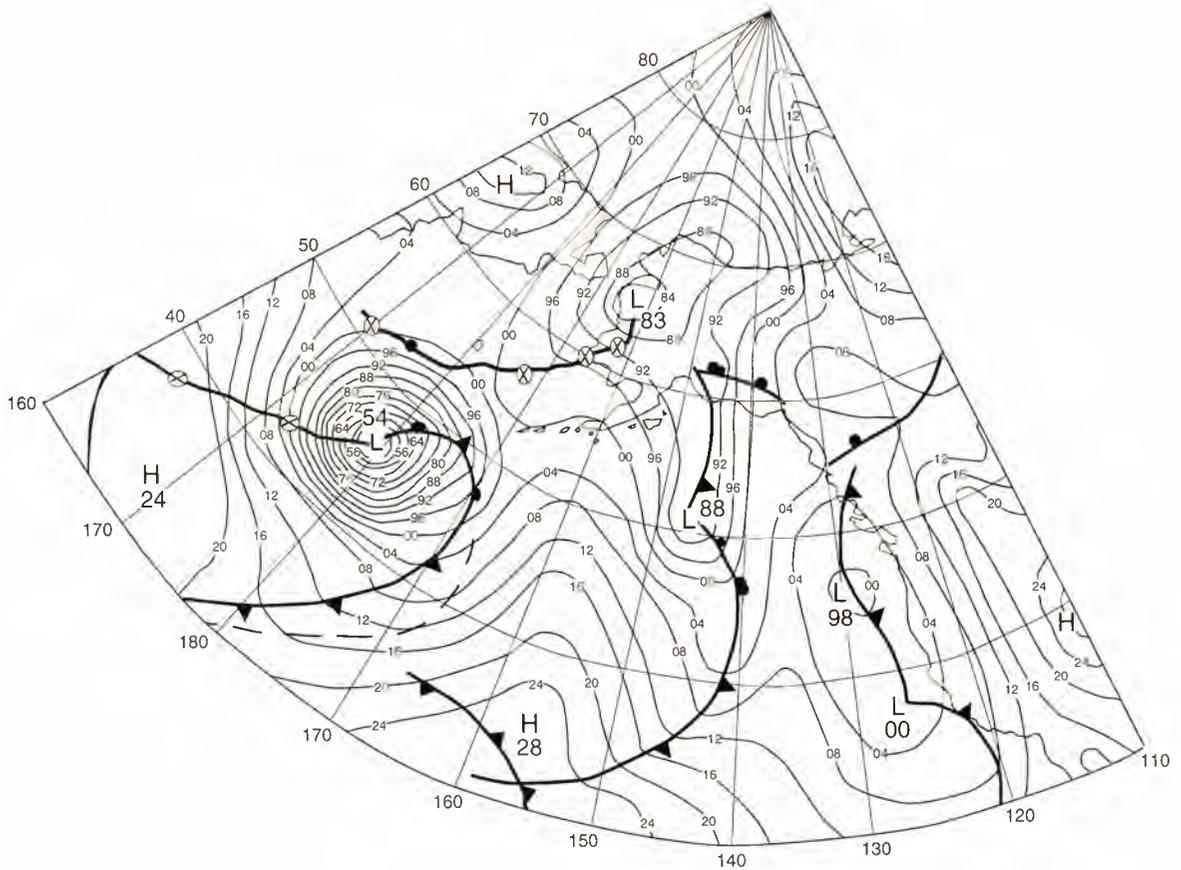


Figure 5. National Climatic Data Center (NCDC) surface analysis synoptic weather chart at 0000 UTC on December 28, 1991. The low-pressure region (954 mbar) identifies the storm used in this study.

exceeds the local wind speed component in the wave direction. The wave then ceases to be forced by the wind and begins to propagate freely as swell. By assuming such a swell system consists of free waves and is unaffected by other processes, such as wave-wave interaction and wave-current interaction, the swell directions at distant locations can be checked with the linear wave propagation model.

The geometry of the idealized storm swell kinematic model is shown in Figure 6 with the position of $S(\phi, \lambda)$ being the center of the swell generation region and P and C being observation points. The great circle distance s to an observation point is given by Snyder [1987] to be

$$\sin s = [\sin^2 \phi_s + \cos \phi_s \cos \phi_p \sin^2 \lambda_s]^{1/2}, \quad (4)$$

where s is the radian measure of the great circle distance. The wave propagation direction or azimuth angle B_p at an observation point is found by combining two angles. The first angle θ_p is the internal angle subtended by two great circles, one that extends through S and P and the other that extends through P and C . The second angle θ_r is the azimuth angle from P to C (SAR path). Spherical geometry (Figure 6) then determines the following relationships:

$$\theta_p = \cos^{-1} \dots \quad (5)$$

and

$$s = C_g(t_p - t_s) \quad (6a)$$

$$l = C_g(t_c - t_s). \quad (6b)$$

In these relationships, if the quantities of (ϕ, λ) , C_g , and t_s are

assumed to be known, then (4), (5), (6a), and (6b) can be solved for the swell propagation direction B_p and arrival time t_i for locations P and C .

4.2. Wave Data Comparison

4.2.1. Storm description. According to the surface synoptic weather analysis [*National Oceanic and Atmospheric Ad-*

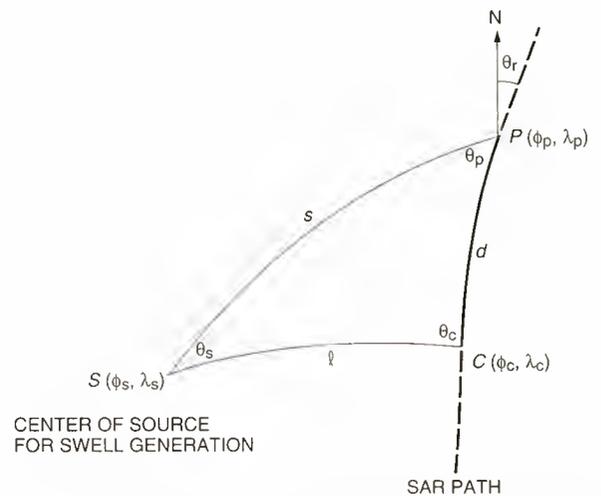


Figure 6. Geometry of the idealized storm swell kinematic model. The point S represents the source of swell generation, and P and C are points of SAR imaging locations.

Table 2. Storm-Tracking Wave Data

Buoy	SAR Orbit	Track	Position	Location		Date	Time, UTC	Distance S , km	B_p , deg	Direc- tion, deg	λ , m	H_s , m	C_g , km h ⁻¹	T_{C_g} , hour	ΔT , hour
				Latitude, °N	Longitude, °W										
NDBC 46003				51.9	155.9	Dec. 29	1000	1723	624	7	56.1	30.7	34
NDBC 46001				56.3	148.3	Dec. 29	2200	2351	624	6	56.1	41.9	44
	E2376	1	a	56.2	149.8	Dec. 29	2100	2260	250	240	565	...	53.4	42.3	45
	E2376	1	b	46.6	153.3	Dec. 29	2100	1835	264	265	388	...	44.0	41.7	45
Endeco				49.1	131.8	Dec. 30	1200	3395	624	4	56.1	60.5	60
	E2390	2	a	54.7	142.0	Dec. 30	2000	2698	263	253	373	...	43.4	62.2	69
	E2390	2	b	47.0	144.8	Dec. 30	2000	2479	278	267	316	...	40.0	62.0	69
	E2390	2	c	43.2	145.9	Dec. 30	2100	2476	286	288	317	...	40.5	61.8	69
Endeco				49.1	131.8	Dec. 31	0000	3395	433	6	46.8	72.5	72
NDBC 46042				36.8	122.4	Dec. 31	0600	4588	...	300	624	5	56.1	81.8	82
	E2404	3	a	53.5	134.2	Dec. 31	2000	3192	...	266	292	...	38.4	83.1	92

Buoy is the buoy identifier, the Endeco buoy is located at ASREX site on Figure 1; SAR Orbit (with Track and Position) is the ERS-1 (E) orbit number plus track and position of the SAR image on Figure 1; Distance S is the great circle distance from the initial estimated storm wave generation region (45°N, 177°W) to the SAR image or buoy position using (4); B_p is the great circle angle of wave propagation direction using (5); Direction is the direction from which waves are propagating measured from SAR or buoy spectra; λ is the peak wavelength measured directly from SAR spectra or derived from the buoy-measured peak wave period using the dispersion relation (1); H_s is the significant wave height measured from buoy; C_g is the group velocity derived from SAR or buoy measurement of wavelength or period; T_{C_g} is the hours of wave travel time using C_g and distance S ; ΔT is the time difference between the estimated storm wave generation time (December 28, 0000 UT) and buoy or SAR wavelength measurement. B_p for E2404 was not derived since there are only two of three points for (5).

ministration, 1992] a low-pressure system developed near Japan on December 25, 1991, and moved rapidly eastward. By December 27, at 1200 UTC, it was a 962-mbar storm near the international date line. At 0000 UTC on December 28 the intensive storm had matured to a center pressure of 954 mbar and was moving northeastward (about 60°) to a position of about 46°N, 179°E (Figure 5). The storm had started to weaken, and its forward speed began to decrease by 0000 UTC on December 29. Using the general concept that a storm's most intense wind and wave generation zone is the lower right quadrant with respect to the storm's path, this would place the intense wave generation zone on the southeast side of the northeast traveling storm. The center of this swell generation zone is initially estimated to be 45°N, 177°W on the southwest side of the center of the storm, some 300 km distance from the storm center (Figure 1). Since the swell generation zone and time can be shifted for various propagation directions within the storm's southeast quadrant, the postulated zones and time can be fine-tuned through iteration.

4.2.2. SAR data. Data from three ERS-1 SAR acquisitions along descending orbits were obtained on December 29, 30, and 31, 1991. Six SAR wave spectra were processed, at positions selected to sample the spreading of the wave field as well as to be close to the buoys (Figure 1). The dominant wavelength and wave direction from these spectra are listed in Table 2, and representative spectra from December 29 are shown in Figure 7. In all cases the wavelengths are >300 m, and the wave directions are eastward or range traveling with respect to the satellite track. No nonlinear mapping is apparent in any of the spectra. The longest waves (565 m) are from the northernmost spectra on December 29. As the available SAR data progress to the east and move later in time, wavelength decreases but wave direction maintains a general clockwise rotation from north to south. This is seen at position a on the three tracks (240°, 253°, and 266°), indicating the rotational spread of the wind source. Position b on Tracks 1 and 2 have nearly the same wave direction (265° and 267°) while being 23 hours apart, indicating that the locations of these images are well aligned with the wave propagation direction.

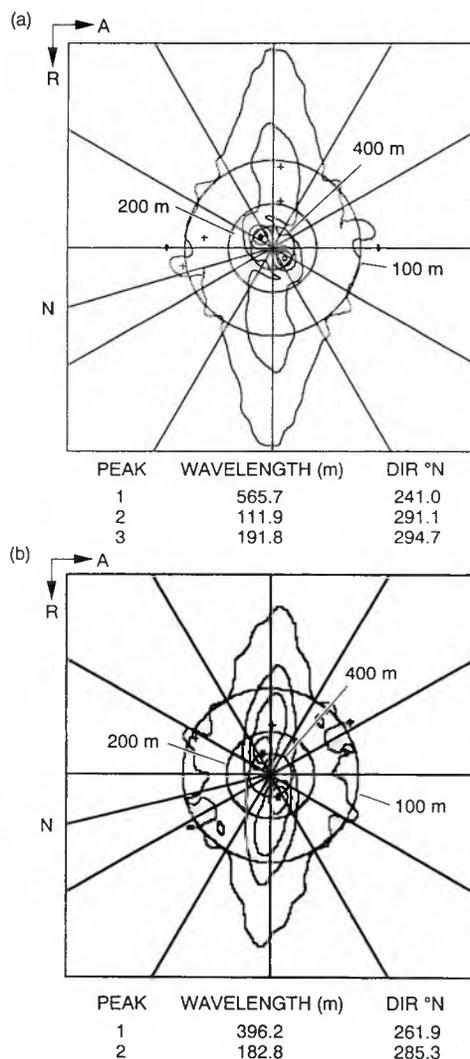


Figure 7. Alaska SAR Facility wave products of ERS-1 SAR spectra from December 29, 1991 (E2376), in the northeast Pacific from (a) track 1, frame a and (b) track 1, frame b (see Figure 1).

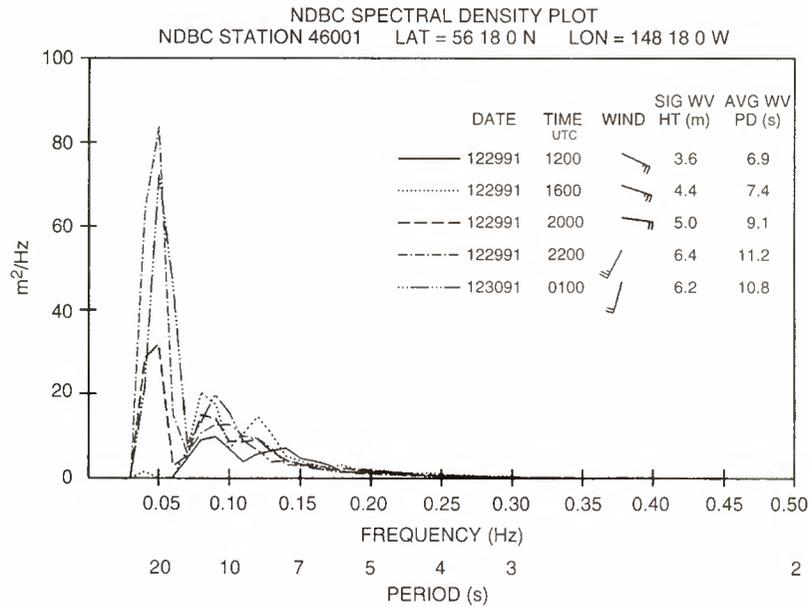


Figure 8. Time series of wave spectra from NDBC buoy 46001 on December 29, 1991.

4.2.3. Buoy wave data. For this study, NDBC buoys 46001, 46003, and 46042 have been utilized, noting again that 46042 is a directional buoy. From the WHOI ASREX program, data from the nondirectional Endeco buoy was used, since the directional Seatex buoy had already been recovered. Table 2 lists the key buoy information.

For buoy 46001 the low-frequency peak of 0.05 Hz reached a maximum density and wave height of more than 6 m on December 29 at 2200 UTC (Figure 8), very close in time and position to the nearest SAR wave spectra. The contour plots for buoys 46001 and 46003 (Figure 9) show clearly the arrival of the 0.05-Hz peak at 46003 to be about 11 hours earlier than at 46001 because of its closer proximity to the storm center. This peak persisted at each buoy for at least 6 hours after arrival (Figure 10). The Endeco buoy showed the arrival time of the wave peak at 1200 UTC on December 30 and a significant decrease in period 12 hours later. At buoy 46042 the same low (0.05 Hz) wavenumber peak arrived on December 31 at 0600 UTC, persisting for some 12 hours with a direction from 300° (Figure 11).

4.2.4. Data comparisons. It is clear from the buoy and SAR results that the peak storm wave field was captured at all four buoys at successive times and at least at the northernmost image from E2376. The remaining SAR frames sampled the storm waves after the wave field maximum had already passed through. This is primarily due to the sampling rate of the satellite being less than fortuitous (no ascending passes were acquired, for example, which are separated in time from the descending passes by about 12 hours). The southernmost image of E2376 had a shorter wavelength (388 m) than might be expected since it was only about 400 km to the south of its companion wave field, which measured over 550 m in wavelength. The smaller measurement may be caused by a slightly less favorable generation direction, which was not as parallel to the storm isobar contours. In terms of direction the southerly end of E2390 and buoy 46042 had similar directional measurements for the peak wave field 9 hours apart.

4.3. Data-Model Results and Discussion

First we consider the arrival times measured by the group velocity derived from the SAR and buoy measurements (T_{C_g}), compared with the total elapsed time (ΔT) between the observations and the estimated storm source time (December 28,

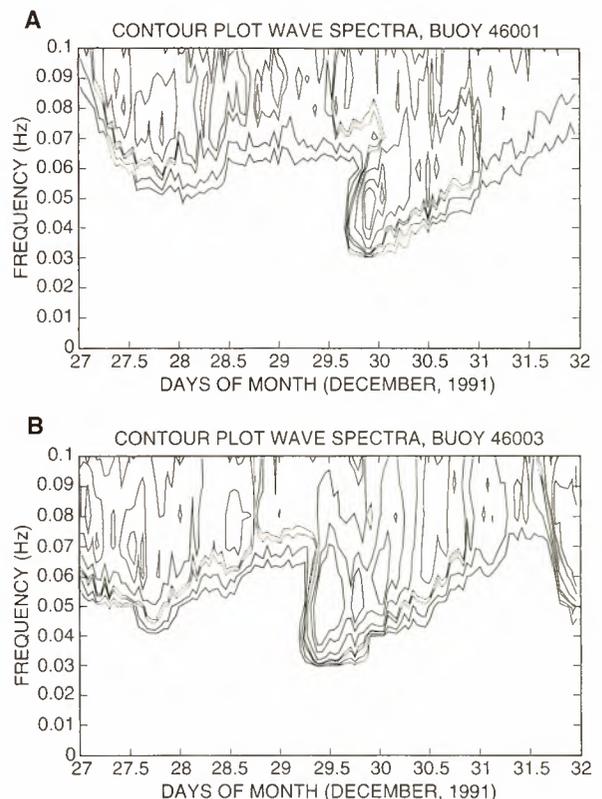


Figure 9. Contour plots of wave spectra from (a) buoy 46001 and (b) buoy 46003 during December 27–31, 1991.

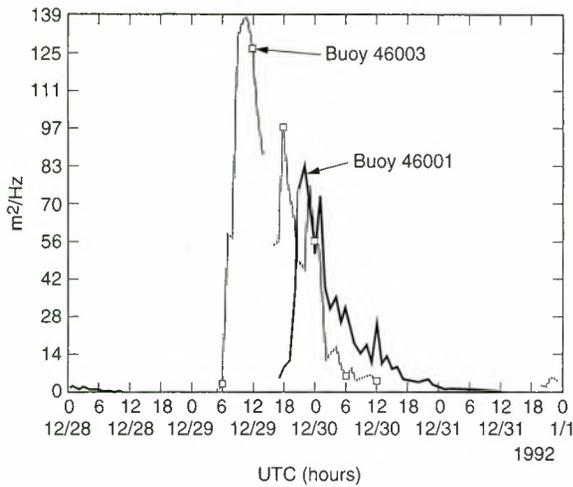


Figure 10. Time series plot of peak wave spectral density for (a) buoy 46001 and (b) buoy 46003 during December 28–31, 1991.

0000 UTC). It can be seen from Table 2 that by and large these measurements result in earlier arrival times, with the exception of the most distant buoy measurements (Endeco and 46042), where the derived arrival times are nearly equal to the elapsed times. The earlier times would indicate that the estimated source region needs updating by being moved farther away, but the close times of the distance buoys suggest that the estimated source location and time is actually quite good.

The comparisons of direction estimated by the model results (B_p) with the SAR measurements are generally favorable. We see that the two southernmost measurements are nearly the same, while the remaining three northerly SAR measurements are 10° less (or rotated counterclockwise) from the model. The differences may be due to wave-current refraction as the wave field passes through the Alaskan Coastal Current. This small set of measurements does suggest that this simple kinematic model is adequate for describing the general fan-shaped directional wave information from a strong storm.

Next, we use the SAR directional measurements to iterate the storm wave generation zone. This can be done simply by adding the equivalent distance ΔS based on C_g for the difference in time between T_{C_g} and ΔT for each measurement (Table 3). Using (4), revised source regions are calculated for the six measurements, which results in an updated mean position of about 42°N , 177°W , ~ 475 km from the storm center (46°N , 179°E) as compared to 300 km with the first estimated point source (45°N , 177°W) (Figure 1). Using the revised point source, we then recalculate T_{C_g} for the four buoys ($*T_{C_g}$ in Table 3), showing closer comparisons with ΔT . The revised source region is moved farther back into the lower right quadrant with respect to storm travel (about 60°). These results indicate that the peak waves were generated within a few hours of the original estimated time that the storm reached its lowest pressure (December 28 at 0000 UTC) and from a region some 300–500 km to the lower right of the storm with respect to its travel direction. This is compared to previous results from *Gonzalez et al.* [1987] showing waves being generated within 200–300 km of a hurricane center, a radius greater than expected from the generally 50–100-km radius of maximum hurricane winds. The larger radius for this Pacific storm is still contained within the tight isobars seen on Figure 5. *Mettlach et*

al. [1994] found differences of 5° in latitude and 4° in longitude in estimates of the wave source region about a typhoon using several methods, which are comparable to these results.

5. Summary

This study indicates that the dominant wavelength and wave direction from SAR measurements plus a simple kinematic model can be used to provide useful information on the wave field emanating from strong storms. Taken separately, the SAR spectra from either E2376 and E2390 provides sufficient details to enable some predictive capability for determining the arrival time, energetics, and direction of large swell propagating from remote storms toward coastal areas. Samples from multiple SAR passes improve this predictive capability. The SAR directional information in particular can enhance both buoy data and wave model results as well as provide improved information on the wind fields from storms, including estimating the source region and generation time.

A key to using SAR for such studies and especially in an operational application is adequate sampling. Spaceborne SAR is either data-rate limited or power limited, which prevents duty cycles (amount of time that the SAR can operate) greater than about 30%. The ERS SARs operate outside ground reception stations in its unique wave mode, producing 5-km vignettes every 200 km along-track. This is potentially a very useful mode except that the data are stored on board and downlinked to a ground station when possible. This delay is at least 90–100 min or as much as several hours depending on how often the satellite comes into view per day of a ground station designated to receive such data. Also, the spectral products from these vignettes are not yet operationally available. Spaceborne SARs have 14–16 orbits per day and limited swaths compared to operational sensors such as AVHRR. Options therefore for improving coverage are to increase swath width (usually at the expense of resolution, which is particularly important for wave imaging on SAR), such as with Can-

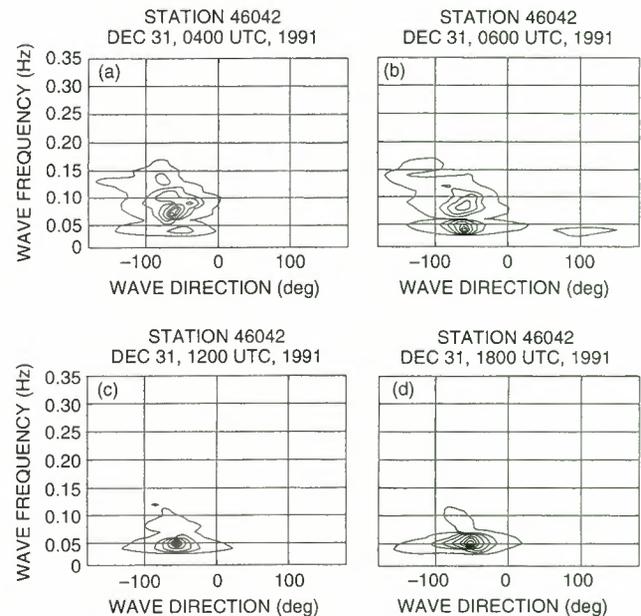


Figure 11. Directional wave spectra from buoy 46042 near Monterey, California, at (a) 0400, (b) 0600, (c) 1200, and (d) 1800 UTC on December 31, 1991.

Table 3. Revised Storm-Tracking Wave Data

Buoy	SAR Orbit	Track	Position	Distance S , km	Direction, deg	C_g , km h ⁻¹	$T_{C_g} - \Delta T$, hour	ΔS , km	Latitude, °N	Longitude	* T_{C_g} , hour	ΔT , hour
	E2376	1	a	2260	240	53.4	2.7	144	42.1	175.3°W		
	E2376	1	b	1835	265	44.0	3.3	145	42.3	177.7°W		
	E2390	2	a	2698	253	43.4	6.8	273–295	40.8	176.6°W		
	E2390	2	b	2479	267	40.0	7.0	260–280	40.5	178.0°W		
	E2390	2	c	2476	288	40.5	7.1	271–288	45.7	179.4°E		
	E2404	3	a	3192	266	38.4	8.9	342	41.4	179.7°W		
NDBC 46003											34.4	34
NDBC 46001											46.2	44
Endeco*											63	60
NDBC 46042											82.8	82

$T_{C_g} - \Delta T$ is the travel time difference from Table 2; ΔS is the change in kilometers from $(T_{C_g} - \Delta T)C_g$; Latitude and Longitude give the revised storm wave generation source region based on Direction and ΔS ; * T_{C_g} is the revised hours of wave travel time using C_g , Distance $S + \Delta S$, and mean of revised latitude and longitude (42°N, 177°W). See footnotes to Table 2.

*Date is December 30, and time is 1200 UTC.

ada's RADARSAT, or to use multiple satellites placed in orbits 1–2 days apart in their repeat pattern, such as ERS-1 and ERS-2 recently were in their tandem phase. Also, of key importance is to get rapid data reception, processing, and incorporation of results with wave model and buoy data. The real-time SAR ocean wave spectra processor flown on two shuttle flights in 1994 demonstrated the capability of onboard processing from image data to wave spectra, which also significantly reduces data storage requirements [Monaldo and Beal, 1996].

This case study indicates the feasibility of using SAR in the advanced warning of high swell, which could also be crucial to the safety of offshore operations and coastal regions. Such a system is feasible with current spaceborne SARs for demonstration purposes at least and operationally in the future with improved ocean coverage and potentially onboard processing. In addition, combining improvements in high-swell prediction with storm surge modeling could also improve the prediction of water run-up in the coastal zone [Bode and Hardy, 1997].

Appendix A: Environmental Conditions for Validation of ERS-1 SAR Data With Buoy Data

A brief summary is provided of the weather and wave environment for each of the 10 dates used in the buoy-SAR validation (Table 1) based on NMC surface charts, site descriptions [Galbraith et al., 1994], and buoy and wave model results.

A1. NDBC Buoy 46001

A1.1. September 18. On September 17 and 18 a low-pressure system to the west gradually moved northwestward and weakened later on September 18. Conditions were mild near the buoy until midday when a weak low-pressure system south and east of the buoy developed. The buoy measured wind speeds near 10 m s⁻¹ with a wind direction from about 270° at midday, which gradually shifted to easterly during September 19 because of the front. H_s was near 3 m. The wave model indicated a wave direction from 230°. The SAR spectra showed a very consistent peak wavelength at 179 m from around 250°, from the low-pressure system to the west, and a less energetic peak of about 300 m from the north, with the latter related to the closer system. This lower frequency peak was not seen on the buoy frequency spectra and may have been nonlinearly mapped to a lower wavenumber on the SAR spec-

tra since it had a strong radial component nearly parallel to the flight direction.

A1.2. October 15. No significant pressure systems were present near the buoy for October 15. Winds had dropped slightly to 12 from 15 m s⁻¹ the day before. Modeled wave direction is 140°. The frequency spectra had a dominant period of 10 s and H_s of about 4 m. The dominant peak on the SAR spectra was at 157 m from a direction of 265°. A less energetic and quite variable peak in the SAR spectra is at 196 m from 134°, which aligns with the wind direction. The 157-m wave field may have emanated from a large low-pressure system present far to the west. In summary, the lack of a significant pressure system has resulted in a variable wind sea on both the buoy and the SAR data, producing a somewhat confusing comparison. There appears to be nonlinear mapping on the SAR spectra, even though the winds are some 40° off of the flight direction, which would tend to reduce the azimuth spectral component.

A1.3. October 24. A stable high-pressure system south of Alaska resulted in a very mild period before and during October 24. Wind speeds were <5 m s⁻¹, H_s was 1 m, and wind direction was from 180°. The wave model hindcast indicated a wave direction also of 180°. The dominant peak in the SAR spectra has similar measurements of 100 m from the south. In this case, then, short wavelength azimuth-traveling waves were apparently accurately mapped on the SAR spectra, primarily because of the low H_s .

A1.4. November 2. A low-pressure system developed near the buoy during November 2, with two substantial low-pressure systems to the west and south. Wind speeds were steady before and during this day at about 10 m s⁻¹ from the west. The frequency spectra showed a stable 12-s wave field, also from the west according to the wave model (Figure 2). H_s was high at around 4.5 m. The SAR spectra showed a very steady 215-m wave field from nearly due west, comparing quite favorably with the buoy data.

A1.5. December 29. The strong low-pressure system discussed here is the same storm used in the case study discussed in section 3. According to the surface synoptic weather analysis a low-pressure system developed near Japan on December 25, 1991, and moved rapidly eastward. By December 27 at 1200 UTC it was a 962-mbar storm near the date line. At 0000 UTC on December 28 the intensive storm had matured with a center pressure of 954 mbar, shifting northeastward (about 60°) to a

position of about 46°N, 179°E (Figure 5). The storm had begun to weaken and its forward speed began to decrease by 0000 UTC on December 29, finally passing very near buoy 46001 late on December 29, with a pressure on 984 mbar. The time series frequency spectra shows a peak with an increasing density coincident to the SAR data on December 29 at 2100 UTC, indicating a peak of around 20 s (Figure 8). The model wave direction is from 300°, although the wind direction is generally from the west. The SAR spectra is quite comparable at 565 m, with a direction rotated slightly southward of the wind direction.

A2. ASREX Seatex Buoy

A2.1. November 4. A weak high-pressure system passed by the mooring in the early part of the day, when the westerly winds dropped from 10 to 3–4 m s⁻¹ by about 0800 UTC. Later in the day, a weak low-pressure system arose, which increased wind speeds to 9–10 m s⁻¹ and shifted the wind direction to southerly. Wave heights remained constant at around 3–3.5 m, with the dominant low-frequency peak emanating from a westerly direction of between 255° and 270°, having been generated from a strong low-pressure system well to the west. These waves are orthogonal or range traveling relative to the ERS-1 flight direction of 196°. Thus these waves are not subject to nonlinear SAR mapping conditions, and the buoy and SAR measurements are comparable.

A2.2. November 16. A strong storm system passed over the mooring during the middle of this day, reaching a low of about 968 mbar near midday. This was followed at 1800 UTC by peak winds of 22 m s⁻¹ and H_s of 10 m with a maximum period of about 14–15 s. Earlier in the day, there was evidence of three wave systems: a low-frequency system propagating from the west with about a 14-s period, a 12-s period from the south, and a wind sea from the southeast. With the storm, the wind sea and low-frequency peaks steadily coalesced from a direction of about 240°. NMC surface charts showed the center of the low at 49°N, 136°W at 1200 UTC and 52°N, 133°W at 0000 UTC on November 17, some 250 km to the west of the buoy.

As noted in section 2, the SAR data showed considerably different dominant wave fields. The location of the SAR data is slightly east of the estimated storm center, between the storm and the buoy. There was a 195-m wave field with a propagation direction from 288° and a 536-m wave component propagating from 175°. The westerly SAR wave field is shorter than that measured by the buoy earlier in the day, and the southerly SAR wave field is considerably longer and rotated 60° toward the south from the dominant buoy component. Also, the steep wind waves in the SAR region are undoubtedly mapped in a nonlinear fashion on the imagery due to the strong wind component in the radial direction, nearly parallel to the flight direction, which resulted in a mapping to a lower wavenumber. The westerly wave component on the SAR imagery may be remaining from the earlier westerly component seen on the buoy data, although it is much shorter in wavelength. In conclusion, the strong rotating winds in the near field to the SAR likely produced nonlinear mapping of the wave field on the SAR spectra.

A2.3. November 25. At the time of the SAR overpass the wind speeds had dropped to around 5 m s⁻¹ from the south from a high of 12 m s⁻¹ from the west late on November 24. Wave heights are relatively constant at around 2–3 m. The low-frequency waves are propagating from the west. The SAR

spectra is broad from the west, but the dominant lobe is centered at 307°. Thus, under these relatively mild conditions, the measured wavelengths are quite comparable, but the broad spectra results in a less accurate angular measurement.

A2.4. December 1. A high-pressure system has remained in the area for several days. A high wind speed of 14 m s⁻¹ occurred at 1200 UTC from the northwest. Wave height remains around 3 m. The frequency spectra at the end of the day was unimodal at about 10 s with a directional peak at about 280°. The buoy and SAR comparisons are quite good.

A2.5. December 4. A front passed over the mooring late on December 3, with wind directions shifting from southerly to westerly. During December 4 the winds shifted slowly back to the south, with the wind speeds remaining steadily at 10 m s⁻¹. Wave height remains around 3–4 m. The low-frequency swell is from the west, while the higher-frequency waves track the wind. Again, the buoy and SAR comparisons of the dominant low-frequency waves are good.

Acknowledgments. The authors wish to thank C. Y. Peng and Y. Y. Chao of GSFC for their valuable discussions and suggestions; Ron Kwok, Glenn Cunningham, Amy Rodgers, and Joanne Shimada of JPL for development of the ASF geophysical processing system; and the ASF operational staff. The JPL effort was supported by the National Aeronautics and Space Administration through a contract with the Jet Propulsion Laboratory, California Institute of Technology. The GSFC work was supported by the National Aeronautics and Space Administration and Office of Naval Research. The ASREX program was supported by the Office of Naval Research under grant N00014-91J-1891.

References

- Alpers, W. R., Monte Carlo simulations for studying the relationship between ocean waves and synthetic aperture radar image spectra, *J. Geophys. Res.*, **88**(C3), 1745–1759, 1983.
- Barber, N. F., and F. Ursell, The generation and propagation of ocean waves and swell, *Philos. Trans. R. Soc. London A*, **240**, 527–560, 1948.
- Beal, R. C. (Ed.), *Directional Ocean Wave Spectra*, 218 pp., Johns Hopkins Univ. Press, Baltimore, Md., 1991.
- Beal, R. C., T. W. Gerling, D. E. Irvine, F. M. Monaldo, and D. G. Tilley, Spatial variations of ocean wave directional spectra, *J. Geophys. Res.*, **91**(C2), 2433–2449, 1986.
- Bode, L., and D. A. Hardy, Progress and recent development in storm surge modeling, *J. Hydraul. Eng.*, **123**, 316–331, 1997.
- Chen, H. S., Ocean surface waves, *Tech. Procedures Bull.*, Ser. 426, 20 pp., Off. of Meteorol., Natl. Weather Serv., Natl. Oceanic and Atmos. Admin., 1995a.
- Chen, H. S., Evaluation of a global ocean wave model at the National Meteorological Center, in *Proceedings of the Second International Conference on Hydro-Science and -Engineering, March 22–26, 1995, Beijing, China*, pp. 1436–1443, Tsinghua Univ. Press, Beijing, 1995b.
- Dobson, F. W., and P. W. Vachon, The Grand Banks ERS-1 SAR wave spectra validation experiment, *Atmos. Ocean*, **32**(1), 7–29, 1994.
- Earle, M. D., K. A. Bush, and G. D. Hamilton, High-height long-period ocean waves generated by a severe storm in the northeast Pacific Ocean during February 1983, *J. Phys. Oceanogr.*, **14**, 1286–1299, 1984.
- Engen, G., H. Johnsen, H. E. Krogstad, and S. F. Barstow, Directional wave spectra by inversion of ERS-1 synthetic aperture radar ocean imagery, *IEEE Trans. Geosci. Remote Sens.*, **GE-32**, 340–352, 1994.
- Galbraith, N. R., A. Gnanadesikan, G. H. Tupper, and B. S. Way, Meteorological and oceanographic data collected during the ASREX 91 field experiment, *Tech. Rep. 94-1*, 114 pp., Woods Hole Oceanogr. Inst., Woods Hole, Mass., 1994.
- Gonzalez, F. I., B. Holt, and D. G. Tilley, The age and source of ocean swell observed in hurricane Josephine, Johns Hopkins *APL Tech. Dig.*, **8**, 94–99, 1987.
- Hasselmann, K., and S. Hasselmann, On the nonlinear mapping of an

- ocean wave spectrum into a synthetic aperture radar image spectrum and its inversion, *J. Geophys. Res.*, *96*(C6), 10,713–10,729, 1991.
- Hasselmann, K., R. K. Raney, W. J. Plant, W. Alpers, R. A. Shuchman, D. R. Lyzenga, C. L. Rufenach, and M. J. Tucker, Theory of synthetic aperture radar ocean imaging: A MARSEN view, *J. Geophys. Res.*, *90*(C3), 4659–4686, 1985.
- Hasselmann, S., C. Brüning, K. Hasselmann, and P. Heimbach, An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra, *J. Geophys. Res.*, *101*(C7), 16,615–16,629, 1996.
- Holt, B., and F. I. Gonzalez, SIR-B observations of dominant ocean waves near Hurricane Josephine, *J. Geophys. Res.*, *91*(C7), 8595–8598, 1986.
- Janssen, P. A. E. M., Wave-induced theory of wind-wave generation applied to wave forecasting, *J. Phys. Oceanogr.*, *19*, 745–754, 1989.
- Janssen, P. A. E. M., Quasi-linear theory of wind-wave generation applied to wave forecasting, *J. Phys. Oceanogr.*, *21*, 1631–1642, 1991.
- Kleijweg, J. C. M., and H. Greidanus, Validation of directional wave spectra from the ERS-1 during the ESA Cal/Val campaign, in *Proceedings of the First ERS-1 Symposium, Cannes, France, 4–6 November 1992*, Eur. Space Agency Spec. Publ., ESA SP-359, 31–34, 1993.
- Komen, G. J., L. Cavalieri, M. Donelan, K. Hasselman, S. Hasselmann, and P. A. E. M. Janssen, *Dynamics and Modeling of Ocean Waves*, 532 pp., Cambridge Univ. Press, New York, 1994.
- Krogstad, H. E., A simple derivation of Hasselmann's nonlinear ocean-synthetic aperture radar transform, *J. Geophys. Res.*, *97*(C2), 2421–2425, 1992.
- Krogstad, H. E., O. Samset, and P. W. Vachon, Generalizations of the non-linear ocean-SAR transform and a simplified SAR inversion algorithm, *Atmos. Ocean*, *32*(1), 61–82, 1994.
- McLeish, W., and D. B. Ross, Imaging radar observations of directional properties of ocean waves, *J. Geophys. Res.*, *88*(C7), 4407–4419, 1983.
- Mettlach, T., D. Wang, and P. Wittman, Analysis and prediction of ocean swell using instrumented buoys, *J. Atmos. Oceanic Technol.*, *11*(2), 506–524, 1994.
- Monaldo, F. M., The consequences of sampling variability on the estimation of wave number and propagation direction from spaceborne SAR image spectra, *IEEE Trans. Geosci. Remote Sens.*, *GE-29*, 113–119, 1991.
- Monaldo, F. M., and R. C. Beal, Comparison of wave parameters measured from the SIR-C on-board processor with WAM predictions in the Southern Ocean, *Rep. SIR-96-01*, 37 pp., Johns Hopkins Univ., Laurel, Md., 1996.
- Monaldo, F. M., T. G. Gerling, and D. G. Tilley, Comparison of SIR-B SAR wave image spectra with wave model predictions: Implications on the SAR modulation transfer function, *IEEE Trans. Geosci. Remote Sens.*, *GE-31*, 1199–1209, 1993.
- Munk, W. H., G. R. Miller, F. E. Snodgrass, and N. F. Barber, Directional recordings of swell from distant storms, *Philos. Trans. R. Soc. London A*, *255*, 505–583, 1963.
- National Oceanic and Atmospheric Administration, North Pacific weather: October, November, and December 1991, *Mar. Weather Log*, *36*(2), 72–85, 1992.
- Overland, J. E., and T. R. Heister, Development of a synoptic climatology for the northeast Gulf of Alaska, *J. Appl. Meteorol.*, *19*, 1–14, 1980.
- Plant, W. J., Reconciliation of theories of synthetic aperture radar imagery of ocean waves, *J. Geophys. Res.*, *97*(C5), 7493–7501, 1992.
- Plant, W. J., and L. M. Zurk, Dominant wave directions and significant wave heights from synthetic aperture radar imagery of the ocean, *J. Geophys. Res.*, *102*(C2), 3473–3482, 1997.
- Raney, R. K., SAR response to partially coherent phenomena, *IEEE Trans. Antennas Propag.*, *AP-26*, 777–787, 1980.
- Sea Wave Modeling Project (SWAMP) Group, *Ocean Wave Modeling*, 262 pp., Plenum, New York, 1985.
- Shearman, R. J., and A. A. Zelenko, Wind measurements reduction to a standard level, *Mar. Meteorol. and Relat. Oceanogr. Activ. Rep.* *22*, 30 pp., World Meteorol. Organ., Geneva, 1989.
- Snodgrass, F. E., G. W. Groves, K. F. Hasselmann, G. R. Miller, W. H. Munk, and W. H. Powers, Propagation of ocean swell across the Pacific, *Philos. Trans. R. Soc. London A*, 431–497, 1966.
- Snyder, J. P., Map projections: A working manual, *U.S. Geol. Surv. Prof. Pap.*, *1395*, 383 pp., 1987.
- Stern, H. L., D. A. Rothrock, R. Kwok, and B. Holt, The geophysical processor system: Automated analysis of ERS-1 SAR imagery, in *Proceedings of the Second ERS-1 Symposium, Hamburg, Germany, 11–14 October 1993*, Eur. Space Agency Spec. Publ., ESA SP-361, 281–286, 1994.
- Tilley, D. G., and R. C. Beal, ERS-1 and Almaz estimates of directional wave spectra conditioned by simultaneous aircraft SAR and buoy measurements, *Atmos. Ocean*, *32*(1), 113–142, 1994.
- Vachon, P. W., H. E. Krogstad, and J. S. Paterson, Airborne and spaceborne synthetic aperture radar observations of ocean waves, *Atmos. Ocean*, *32*(1), 83–112, 1994.
- Wadhams, P., and B. Holt, Waves in frazil and pancake ice and their detection in Seasat synthetic aperture radar imagery, *J. Geophys. Res.*, *96*(C5), 8835–8852, 1991.
- Wang, D. W., and R. Carolan, Estimation of swell direction by a small disc buoy in high seas, in *Proceedings of the Fifth Conference on Meteorology and Oceanography of the Coastal Zone Miami, Fla., May 6–9, 1991*, pp. 53–59, Am. Meteorol. Soc., Boston, Mass., 1991.
- Wave Model Development and Implementation (WAMDI) Group, The WAM model: A third generation ocean wave prediction model, *J. Phys. Oceanogr.*, *18*, 1775–1810, 1988.
- H. S. Chen, National Centers for Environmental Prediction, NOAA, Washington, DC 20233.
- A. Gnanadesikan, Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ 08540.
- B. Holt, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109-8099. (e-mail: ben@pacific.jpl.nasa.gov)
- A. K. Liu, Oceans and Ice Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771.
- D. W. Wang, Applied Technology Division, Computer Sciences Corporation, Stennis Space Center, MS 39529.

(Received February 5, 1997; revised August 21, 1997; accepted September 8, 1997.)

Statistical analysis and intercomparison of WAM model data with global ERS-1 SAR wave mode spectral retrievals over 3 years

P. Heimbach, S. Hasselmann, and K. Hasselmann

Max-Planck-Institut für Meteorologie, Hamburg, Germany

Abstract. Ocean wave spectra were retrieved from a set of ERS-1 synthetic aperture radar (SAR) wave mode (SWM) spectra between January 1993 and December 1995. An assessment is given of the SWM data quality and the retrieval performance as well as the operational feasibility of the retrieval algorithm. Sensitivity studies are performed to demonstrate the weak residual dependence of the retrieval on the first-guess input spectrum. The mean spectral parameters of the SWM retrievals are compared with spectral parameters from collocated wave model (WAM) spectra. The time series of SWM-retrieved and WAM-derived monthly mean significant wave heights H_s in various ocean basins show good overall agreement but with a small systematic underestimation of H_s by the WAM. A decomposition of the wave spectra into wind sea and swell reveals an average 10% overprediction of the wind sea by the WAM while swell is underpredicted by 20–30%. The positive wind-sea bias exhibits no clear wave height dependence, while the negative swell bias decreases with swell wave height. This could be due to a too strong damping in the WAM at low frequencies. Detailed regional investigations point to the existence of smaller-scale phenomena, which may not be adequately reproduced by the WAM at the present resolution of the wind forcing. Finally, an intercomparison is made of the observed and modeled azimuthal cutoff length scales, and global distributions are investigated. Ratios of the observed azimuthal cutoff wavenumber to the mean azimuthal wavenumber component indicate that about 75% of the swell can be directly resolved by the SAR, while about 70% of the wind sea lies at least partially beyond the cutoff.

1. Introduction

One of the major drivers of the first European Remote Sensing Satellite ERS-1, launched on July 17, 1991, was the applications in wave research and wave forecasting. Through the wave mode [see, e.g., *European Space Agency (ESA)*, 1993; *Brüning et al.*, 1994] of the synthetic aperture radar (SAR) [see, e.g., *Harger*, 1970; *Allan*, 1983; *Bamler and Schättler*, 1993], ERS-1 provided, for the first time, detailed spectral information on the sea state globally, continuously, and in quasi-real time, far beyond the capabilities of other wave observation systems.

The development of improved ocean wave models for operational wave forecasts and the computation of wave climatologies from past wind field data is of considerable economic benefit for numerous activities, such as ship routing, fisheries, offshore operations, and coastal protection [see, e.g., *ESA*, 1996]. Beyond these practical applications an understanding of the dynamics of waves is essential also for an improved understanding of the mechanisms that govern the transfer of momentum, sensible and latent heat, and gases such as CO_2 across the air-sea interface, an area of research that is rapidly gaining significance through the development of sophisticated global coupled ocean-atmosphere carbon cycle climate models [*Komen et al.*, 1994; see also *Weber et al.*, 1993; *Mastenbroek et*

al., 1993; *Jähne and Monahan*, 1995; *Janssen and Viterbo*, 1996; *Weisse and Alvarez*, 1997].

An accurate representation of the wind stress from air-sea bulk transfer coefficients, especially the relationship between the wind stress, wind profile, and sea state, is important for studies of the wind-driven ocean circulation. The dependence of wind stress on sea state is still an open question [see, e.g., *Geernaert*, 1990; *Komen et al.*, 1998]. Although recent analyses of Humidity Exchange Over the Sea Main Experiment (HEXMAX) (1986) data from the Humidity Exchange Over the Sea Experiment (HEXOS) project [see *Smith et al.*, 1992; *Janssen*, 1997] suggest only a limited sea state dependence, *Rieder et al.* [1994] and *Fairall et al.* [1996], from the analysis of data from the Surface Waves and Processes Program (SWAPP) campaign (1990) and the Tropical Ocean-Global Atmosphere Coupled Ocean-Atmosphere Response Experiment (TOGA COARE) project, respectively, report deviations between the directions of the wind stress and the wind velocity under the influence of swell. The investigations of data from the SWAPP campaign (1990) by *Large et al.* [1995] and from the Ocean Storm Experiment (1987) by *Rieder* [1997] also indicate a complex dependence of the sea surface drag on the wind profile and the wind-sea and swell components of the sea state. The new information from ERS-1/2, including both wave spectral data and scatterometer wind measurements, could aid significantly in shedding light on the relevant mechanisms. Long-term climatologies based on continuous global wave field observations would furthermore be valuable for assessing possible changes of wave climate due to global greenhouse warming or other climate changes [see, e.g., *Bacon and Carter*,

Copyright 1998 by the American Geophysical Union.

Paper number 97JC03203.
0148-0227/98/97JC-03203\$09.00

1991; von Storch *et al.*, 1993; *Wave and Storms in the North Atlantic (WASA) Group*, 1995; Bouws *et al.*, 1996].

The potential of forthcoming ERS-1 wave data has been a strong motivation for significant advances in wave modeling and methods of satellite data assimilation (compare the detailed presentation in work by Komen *et al.* [1994]). However, we shall not discuss the various proposed wind and wave data assimilation schemes for modern third-generation wave models in this paper [see, e.g., Hasselmann *et al.*, 1988; Lionello *et al.*, 1992, 1995; Hasselmann *et al.*, 1994; Bauer *et al.*, 1996b; Hasselmann *et al.*, 1997; Breivik *et al.*, this issue], but will focus rather on the analysis of the imaged spectra provided by the ERS-1 SAR wave mode (SWM) data and the applicability of the data for wave model validation.

The period of our analysis extends over the 3-year period January 1, 1993, to December 31, 1995, for which reliable ERS-1 SWM data were available. Prior to January 1, 1993, the SWM product was still in a process of modification. On April 21, 1995, ERS-2 was launched. It has now replaced ERS-1 in providing near-real-time global SWM data. It is hoped that the continuity and operational application of global SAR wave data will be ensured in the future through the launch of ENVISAT in 1999, for which an advanced SAR (ASAR) is foreseen [ESA, 1997].

The retrieval of ocean wave spectra from SAR imaged spectra is not a simple exercise. The SAR imaging mechanism is strongly nonlinear through the distortions induced by the wave orbital motions (the "velocity bunching mechanism," compare the MARSEN review in work by K. Hasselmann *et al.* [1985]). This results, among other effects, in image smearing and a loss of information beyond the so-called azimuthal cutoff wavenumber, corresponding typically to wavelengths shorter than about 100–200 m in the satellite flight direction. In addition, ocean wave spectra from satellite SAR images suffer from a basic 180° frozen-image ambiguity. This shortcoming can in principle be overcome by making use of the information on the wave propagation direction contained in successive components of multilook images [e.g., Young *et al.*, 1985; Vachon and Raney, 1991; Johnsen, 1992; Bao and Alpers., 1998; Engen and Johnsen, 1995]. However, although the feasibility has been demonstrated, the technique has not yet been implemented operationally. Nevertheless, even with these limitations, Hasselmann and Hasselmann [1991] (hereinafter referred to as HH) were able to develop an efficient inversion algorithm enabling a reliable retrieval of ocean wave spectra from SAR spectra within the computational constraints of real-time operational applications [see also Krogstad, 1992; Bao *et al.*, 1994; Krogstad *et al.*, 1994]. An improved algorithm has recently been proposed by S. Hasselmann *et al.* [1996] (hereinafter referred to as HBHH). This retrieval algorithm has been used in the present study.

The WAM model plays a central role in our analysis [*Wave Model Development and Implementation (WAMDI) Group*, 1988]. It is used to provide the first guess for the wave spectral retrievals, and we use the WAM model predictions again in intercomparing the modeled and retrieved wave spectra. However, the approach is not as circular as this may appear, since it has been demonstrated by HBHH and is further verified in this paper that the improved retrieval algorithm used here is insensitive to the initial input spectrum. The first-guess information is used essentially only to remove the 180° ambiguity and augment the retrieved spectrum beyond the azimuthal cutoff.

The WAM model now runs routinely at most operational forecasting centers and has been implemented at more than 100 research institutions. It has been extensively validated through case studies and field data, including fetch-limited wave growth conditions and extreme storm conditions, and on a statistical basis against buoy data [see Komen *et al.*, 1994].

Validation of the WAM model using spaceborne data has so far been limited to H_s data derived from satellite altimeters. A comprehensive compilation of WAM validation studies using spaceborne altimeter data from Seasat (1978), Geosat (1985–1989), ERS-1 (1991–1996), and TOPEX/POSEIDON (since 1992) as well as a global comparison between collocated ERS-1 and TOPEX altimeter significant wave heights is given by Bauer and Staabs [1998].

Most validation exercises indicate a good overall agreement between the model and observations, but significant differences are nevertheless occasionally found. The discrepancies tend to differ from study to study and have usually been attributed to errors in the forcing wind field or, in some cases, to inadequate model resolution. Small systematic deviations of the model from a larger ensemble of altimeter wave height data have also been partly attributed to differences in the instrument calibrations or in the sensor algorithms or to differences in the spatial and temporal sampling. In their intercomparison of altimeter data from different satellites, Bauer and Staabs [1998] state that no definite conclusion can yet be drawn regarding which altimeter wave height data set is closest to the true sea state. Nevertheless, such intercomparisons, in conjunction with model data, are helpful in clarifying the characteristics of the different data sets and identifying possible error sources in either the data or the model.

Although providing continuous global coverage, altimeter data suffer from the basic limitation of yielding only significant wave heights. When trying to detect detailed deficiencies in the spectral properties of a wave model, validation studies based on altimeter data alone are therefore of restricted value. The SAR is at present the only instrument able to provide the required detailed two-dimensional spectral data with continuous global coverage. Following demonstrations of the feasibility of retrieving wave spectra from SAR image spectra for aircraft and the first ocean satellite Seasat [Hasselmann *et al.*, 1991], a first validation of the ERS-1 SAR wave mode was carried out for a limited 3-day data set in the Atlantic [Brüning *et al.*, 1993, 1994]. Various aspects of the ERS-1 SWM product have been investigated also by other authors [e.g., Krogstad, 1994; Hansen *et al.*, 1994; Chapron *et al.*, 1995; Breivik *et al.*, 1995; K. Hasselmann *et al.*, 1996; Heimbach *et al.*, 1996; Kerbaol and Chapron, 1996; E. Bauer and P. Heimbach, manuscript in preparation, 1998]. However, the present study represents the first attempt at a global assessment of the complete ERS-1 SWM data set.

In our intercomparison of retrieved SAR wave spectra with the WAM model over the 3-year period effectively covered by ERS-1, we shall be able to validate not only the SAR wave mode data product and retrieval algorithm but also the WAM model at a higher level of spectral detail than was previously feasible. However, in considering the full ERS-1 data set we have necessarily had to restrict the analysis to an essentially statistical investigation. We hope that the general results presented here will nonetheless reveal a number of interesting features that will stimulate further detailed case studies addressing specific dynamical questions.

The paper is organized as follows. Section 2 briefly reviews

the basic inputs to the SAR spectral retrieval algorithm, describing the SAR wave mode data (section 2.1), the nonlinear wave-to-SAR spectral mapping relation and inversion technique (section 2.2), and the WAM model (section 2.3). The retrieval process is assessed in section 3 with respect to the quality of the input SWM data (section 3.1), the fidelity of the retrieval, as expressed by the correlation between the observed and simulated SAR spectra (section 3.2), and the sensitivity of the retrieval with respect to the first guess (section 3.3). In section 4 the spectral retrievals are intercompared with the WAM model spectra. Time series of monthly mean significant wave heights and regression coefficients are presented for the total wave field in section 4.1 and for separate wind-sea and swell contributions in section 4.2. More detailed regional and zonal intercomparisons of wind-sea and swell wave heights, wavelengths, and directions are presented in section 4.3. Case studies for the South Pacific Ocean, a meridional data section through the mid-Pacific, and a global analysis of long-swell components are discussed in sections 4.4, 4.5, and 4.6, respectively. In section 5 the observed and simulated azimuthal cutoff parameters characterizing the nonlinearity of the SAR imaging process are intercompared, the impact of the cutoff on the retrieval of swell and wind-sea systems is investigated, and a proposed dependence of the cutoff on the wind speed is tested. The principal conclusions of our investigation are summarized in section 6.

2. Derivation of Wave Spectra From ERS-1 SAR Wave Mode Data

2.1. ERS-1 SAR Wave Mode Fast Delivery Product

The operation of the ERS-1 synthetic aperture radar (SAR) in the intermittent sampling wave mode provided for the first time global two-dimensional wave spectral data in near-real time. In contrast to the $100\text{-} \times \text{ } 100\text{-km}$ full-swath SAR image mode, which can be operated only during maximally 10% of the orbit and while in line-of-sight of a ground station, the SAR wave mode (SWM) $10\text{-} \times \text{ } 5\text{-km}$ snapshot imagettes are recorded every 200 km along the satellite track. The data are stored on board and transmitted once per orbit to the ERS ground stations. This yields a daily coverage of the global wave spectral field (with occasional gaps when the SAR was operated in full-swath mode) at an along-track resolution comparable with the model resolution, but with a lower cross-track resolution of order $1000\text{--}2000$ km. The imagettes are fast Fourier transformed to 512×512 wavenumber spectra, which are subsequently averaged and converted to 12×12 polar wavenumber coordinates (see Table 3; detailed descriptions are given in work by ESA [1993], Brooker [1995], and Hasselmann et al. [1998]).

The European Space Research Institute (ESRIN) at Frascati, Italy, distributes the European Space Agency (ESA) fast delivery product (FDP) in quasi-real time to the major operational forecasting centers. We received the FDP via the United Kingdom Meteorological Office (UKMO) at Bracknell, England, and the Deutscher Wetterdienst (DWD) at Offenbach, Germany.

2.2. Retrieval of Wave Spectra From SAR Wave Mode Spectra

2.2.1. Closed nonlinear spectral integral transform. The nonlinear imaging theory of ocean waves by a SAR is by now well understood (compare the Marine Remote Sensing Exper-

iment (MARSEN) review by K. Hasselmann et al. [1985]). The imaging is produced by two mechanisms: the linear real aperture radar (RAR) modulation of the backscatter cross section and the orbital motion effects that induce Doppler shifts modifying the phase history of the backscattering signals used to reconstruct the SAR image. The RAR modulation is caused by hydrodynamic interactions between the short backscattering Bragg waves and long waves and the tilting of the backscattering surface by the long waves. The orbital motion Doppler shifts produce azimuthal displacements of the backscattering elements in the SAR image plane, leading to a redistribution of the apparent backscattering density ("velocity bunching"). If the azimuthal displacements become comparable with a typical ocean wavelength, the image becomes nonlinearly distorted. For large displacements compared with a wavelength the image is effectively smeared out. This occurs at the so-called high-wavenumber azimuthal cutoff of the SAR image spectrum.

First simulations of the fully nonlinear velocity bunching mechanism were achieved using Monte Carlo methods [see Alpers, 1983; Alpers et al., 1986; Brüning et al., 1990]. This technique, however, requires considerable computing time and does not lend itself readily to an inversion, as required for the construction of a retrieval algorithm.

The inversion problem became tractable through the derivation of a closed nonlinear integral expression for the mapping of the ocean wave spectrum into a SAR image spectrum [HH; see also Krogstad, 1992; Bao et al., 1994; Krogstad et al., 1994]. The nonlinear integral can be represented as the Fourier transform of a nonlinear function of the auto- and cross-covariance functions $f^R(\mathbf{r})$, $f^V(\mathbf{r})$ and $f^{RV}(\mathbf{r})$ of the RAR image intensity $I^{RAR}(\mathbf{r})$, and the orbital velocity $v(\mathbf{r})$. It contains an exponential cutoff factor $\exp[-k_x^2 \xi^2]$ that describes the azimuthal falloff of the image spectrum beyond an azimuthal cutoff wavenumber k_x , where $\xi = v \cdot R/U$ is the root-mean-square (rms) azimuthal displacement, given by the product of the rms orbital (range) velocity v and the range-to-platform velocity ratio $\beta = R/U$ (see HH).

The imaging theory and nonlinear mapping relation have been verified in a number of field experiments with airborne SARs, such as the Labrador Extreme Waves Experiment (LEWEX) [see Beal, 1991] and the Synthetic Aperture Radar and X Band Ocean Nonlinearities—Forschungsplattform Nordsee (SAXON—FPN) campaign [see Plant and Alpers, 1994] as well as with spaceborne SARs on board Seasat (see the special Seasat issue of the *Journal of Geophysical Research* (volume 88, number C3) and Hasselmann et al. [1991]), the space shuttle missions shuttle imaging radar B (SIR-B) [e.g., Alpers et al., 1986; Monaldo and Lyzenga, 1988; Brüning et al., 1988], and the Grand Banks ERS-1 SAR Wave Spectra Validation Experiment, which also included a validation against airborne SARs (see the special issue of *Atmosphere Ocean* (volume 32, number 1) and the ALMAZ-1 satellite mission [Wilde et al., 1994]).

For aircraft SARs the quasi-linear approximation of the SAR imaging expression, obtained by terminating the Fourier series expansion after the first linear terms but retaining the nonlinear exponential azimuthal cutoff factor, is often an acceptable approximation [see e.g., Monaldo and Lyzenga, 1986; Beal et al., 1991; Monaldo and Beal, 1995, and references therein]. However, the quasi-linear approximation is unable to account for the nonlinear image distortions inherent in the imaging process for the high range-to-platform velocity ratios

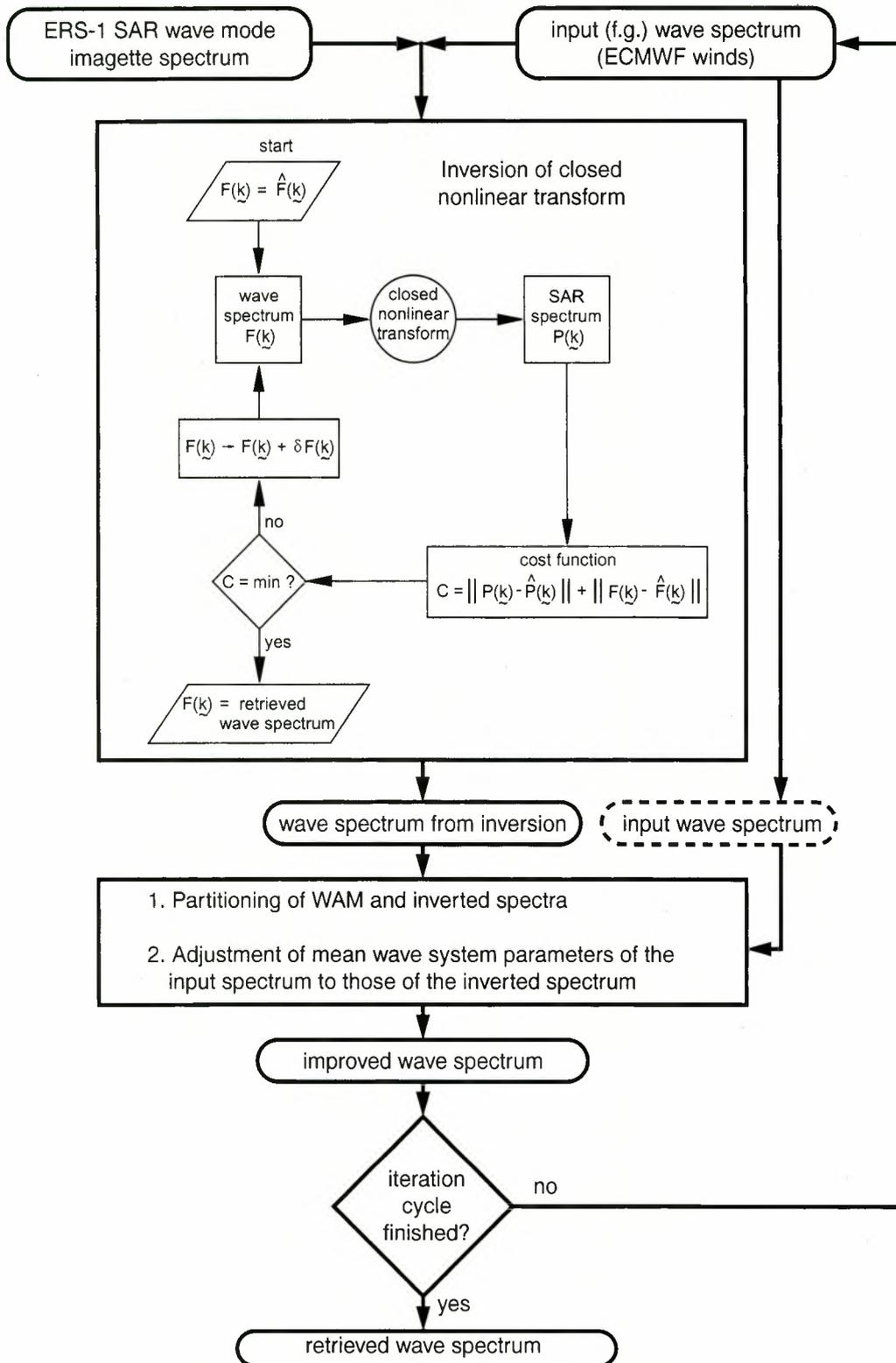


Figure 1. Flow chart of the retrieval algorithm showing the inner loop iterative inversion procedure of the closed nonlinear transform and the outer loop iterative updating of the input spectrum using a spectral wave partitioning scheme.

typical of satellite SARs (for ERS-1, $R/U = 112$ s, which is much larger than typical R/U ratios for low-flying aircrafts).

2.2.2. Retrieval algorithm. Subsequent to their derivation of the nonlinear integral transform, HH presented an algorithm for inverting the spectral mapping relation. Their

retrieval algorithm has since been extended (HBHH) by adding an explicit cutoff adjustment term to the cost function and a second iteration loop that successively modifies the first-guess input spectrum (see flow chart, Figure 1 and section 3.3). The latter extension effectively decouples the retrieval from

the original first-guess spectrum. As one of the aims of this paper is to assess the performance of the improved retrieval algorithm, we briefly review the main features of the technique.

2.2.2.1. Inversion of the nonlinear transform: The first step of the retrieval algorithm is the inversion of the nonlinear spectral mapping relation for a given first-guess input spectrum by minimizing a suitably defined error cost function (inner iteration loop in Figure 1). A first guess is required for two reasons.

1. Since the standard SAR product is an image of the instantaneous sea surface, it contains a 180° directional ambiguity. The ambiguity can be resolved only through additional information from a first-guess spectrum.

2. The nonlinear velocity bunching mechanism smears out the short wave components beyond an azimuthal cutoff wavenumber dependent on the rms orbital velocity. The missing information must again be augmented by a first-guess wave spectrum. However, this limitation is partially overcome by HBHH by including an explicit penalty term in the cost function for errors in the azimuthal cutoff (not shown in Figure 1), which is sensitive to short waves beyond the cutoff (see section 5 for a validation of the cutoff fitting procedure).

A valuable feature of the SAR spectral retrieval algorithm is the availability of an internal calibration based on the level of the background clutter spectrum. Thus the retrieved spectra can be calibrated in absolute wave height units without reference to the SAR instrument calibration or measurements of the absolute backscattering cross section [see *Alpers and Haselmann, 1982; Brüning et al., 1994*].

2.2.2.2. Spectral partitioning and iterative updating of the input spectrum: Although the overall level of the high wavenumber part of the spectrum is adjusted through the inclusion of an explicit azimuthal cutoff penalty term in the cost function, the inversion method modifies the detailed form of the spectrum only in the main part of the spectrum for which direct SAR spectral information is available. A straightforward inversion of the SAR spectrum therefore normally leads to unrealistic discontinuities in the transition zone near the azimuthal cutoff wavenumber separating the spectral regions with and without direct SAR information (HH).

This difficulty is overcome in the revised algorithm of HBHH by introducing a spectral partitioning scheme into the additional iteration loop that updates the input spectrum (compare Figure 1). The iteration consists of repeatedly replacing the input spectrum used for the inversion for the iteration cycle $k + 1$ by the wave spectrum derived from the inversion in the previous iteration cycle k . However, instead of simply replacing the spectra wavenumber by wavenumber, the transfer is implemented via a spectral partitioning scheme (a modified form of the original partitioning algorithm of *Gerling [1992]*; see *Brüning et al. [1994]* and HBHH). This effectively extends the available SAR information below the azimuthal cutoff into higher wavenumbers beyond the cutoff. The new input spectrum retains the continuity properties of the original input spectrum, but the scales and propagation directions of the wave systems of the new spectrum are adjusted to the inverted spectrum. For details we refer to HBHH.

In the present study, five iterative updates of the input spectrum were carried out. As discussed by HBHH, the iteration leads to a closer agreement between the observed and simulated SAR spectrum for the first iteration cycles but then often begins to diverge. We have accordingly selected as the best retrieval the solution that yields the smallest dimensionless square error ε between the SAR spectrum computed from the

Iterative wave spectral retrieval		
0-th cycle	best cycle	Inversion and partitioning cycle
		input wave spectrum
		SAR spectrum computed from input wave spectrum
		observed SAR spectrum
		SAR spectrum computed from inversion
		wave spectrum from inversion
		retrieved wave spectrum after partitioning operation

Figure 2. Example of the iterative wave spectral retrieval operation on a Cartesian 128×128 wavenumber grid.

input wave spectrum and the observed SAR spectrum (compare HBHH and Figure 2).

2.3. WAM Model

The wave model spectra used as first-guess input for our retrievals (and for the later retrieval-model intercomparisons) were computed with the ECMWF operational wave forecast model WAM cycle 4 every 6 hours on a global $3^\circ \times 3^\circ$ (since July 1994, $1.5^\circ \times 1.5^\circ$) latitude-longitude grid [*Janssen et al., 1996*]. As wind forcing, we used the ECMWF analyzed wind field u_{10} at 10-m height. For the period of the present analysis, ERS-1 scatterometer winds were not yet being assimilated into the analysis [see *European Centre for Medium-Range Weather Forecasts (ECMWF), 1995*].

The WAM model solves the energy balance equation

$$\begin{aligned} \frac{D}{Dt} F &= \frac{\partial}{\partial t} F + \frac{1}{\cos \varphi} \frac{\partial}{\partial \varphi} (\dot{\varphi} \cos \varphi F) + \frac{\partial}{\partial \psi} (\dot{\psi} F) + \frac{\partial}{\partial \theta} (\dot{\theta} F) \\ &= S_m + S_{nl} + S_{ds} \end{aligned} \quad (1)$$

for the two-dimensional wave spectrum $F = F(f, \theta; \varphi, \psi, t)$ as a function of frequency f and direction θ on a spherical coordinate grid φ, ψ (see appendix, section A2). The source terms on the right-hand side are integrated using an implicit integration scheme with a time step of 15 min, while the integra-

tion of the advection terms on the left-hand side is performed using a first-order upwind scheme with a time step of 30 min.

As a third-generation model (as opposed to second-generation models), the WAM model introduces no prior assumptions on the shape of the spectrum but determines the evolution of the spectrum by direct integration of the transport equation from first principles, as expressed by the structure of the three source terms as follows [WAMDI Group, 1988; Günther *et al.*, 1992; Komen *et al.*, 1994].

1. The input term S_{in} describing the wind forcing is proportional to the wave spectrum. In WAM cycle 4 used in the present study the wind input is based on Miles' [1957] critical layer mechanism for laminar flows as adopted by Snyder *et al.* [1981], with modifications by Janssen [1989, 1991, 1992] to include the effect of a wave-induced stress τ_w .

2. The nonlinear transfer term S_{nl} is represented by the so-called direct interaction approximation of the full Boltzmann collision integral expression describing the conservative spectral energy transfer due to resonant weak nonlinear wave-wave interactions [see Hasselmann and Hasselmann, 1985; S. Hasselmann *et al.*, 1985, and references therein].

3. The dissipation term S_{ds} is the least well known. The dominant dissipation process is generally regarded as white-capping, which involves many scales of breaking waves. However, as a process which is weak in the mean, it follows from general principles [Hasselmann, 1974] that it can normally be represented as a quasi-linear process. The form implemented in WAM cycle 4 is based on a parameterization explored by Komen *et al.* [1984] and extended by Janssen [1991].

3. Retrieval Assessment

An assessment of the overall performance of the combined ERS-1 SAR wave mode system and retrieval procedure involves several complementary aspects.

1. The first aspect is data quality, which concerns the quality of the SWM product in terms of signal-to-noise and contamination with nonwavelike features [HH; Brüning *et al.*, 1994]. Nonlinear image distortions, on the other hand, as an inherent property of the measurement mechanism, are not regarded as a quality reduction of the SWM product.

2. The second aspect is retrieval performance, which refers to the ability of the retrieved wave spectrum to reproduce the observed SAR spectrum. If the SAR spectrum computed from the retrieved wave spectrum via the full nonlinear forward transform agrees closely with the observed wave spectrum, as measured, for example, by the normalized spectral error, the retrieval has high fidelity. This is clearly the first goal of a retrieval algorithm. A retrieval with high fidelity has extracted all the information that is available from the SAR spectrum [see HH; Brüning *et al.*, 1994; HBHH]. An assessment of the retrieval fidelity or performance together with the data quality is given in section 3.1.

3. The third aspect is retrieval validation, which assesses the degree of agreement of the retrieved spectrum with the true wave spectrum. We shall address this problem in later sections of this paper by intercomparing the retrievals with model data. A comparison with model data has the advantage that both data sets are available globally for the complete two-dimensional wave spectra and throughout the entire 3-year analysis period. But the approach naturally raises the question whether model spectra can be regarded as reasonable approximations of the true wave spectra. However, this is a basic shortcoming of all intercomparisons. Normally, it is not possible

to regard either of the two data sets of an intercomparison as a faithful replica of the true state. All one can do is cross-validate the data in an attempt to gain a synthetic assessment of both data sets, taking into account their individual shortcomings.

Validation of ERS-1 SAR retrievals against other wave observations has been undertaken to date only for wave height data from satellite altimeters. Unfortunately, although satellite altimeters also provide global continuous coverage, the lack of detailed spectral information limits the value of such validation exercises. Nevertheless, a global intercomparison of a 1-year global data set of significant wave heights retrieved from ERS-1 SWM with collocated TOPEX/POSEIDON altimeter data has recently been performed by E. Bauer and P. Heimbach (manuscript in preparation, 1998). A regional stratification of the data similar to that introduced here (see section 4) exhibited good overall agreement between SWM-retrieved and altimeter wave heights but indicated some differences between the altimeter wave heights. The ERS-1 altimeter H_s data were found to be generally lower by an average of 0.2 m than the other two data sets. The TOPEX data agreed to within 5–10% with the SWM data in the tropics throughout the year and in the extratropics during summer, i.e., generally for low sea states. However, they were systematically higher by about 15% in the midlatitude winter, i.e., for high sea states. In general, the validation exercise indicated a satisfactory overall accuracy of SWM-retrieved significant wave heights.

4. The last aspect is operational efficiency. Three considerations are relevant for the assessment of operational efficiency: (1) the number of SWM spectra delivered to the operational forecast center (in the present case, ECMWF) in quasi-real time per month, or, equivalently, the percentage of time that the end-to-end SWM system was functioning satisfactorily in the operational quasi real-time mode; (2) the percentage of the received SWM spectra for which the operational center was able to provide collocated first-guess wave spectra as input for the retrievals; failure to provide an input spectrum can be either due to reception of the SWM product outside the relevant operational time window (which is classed as a failure of the delivery system), or to system interruptions at ECMWF; and (3) the computational load of the operational production of global SWM spectral retrievals; at a peak frequency of one SWM spectrum every 30 s it is imperative to keep central processing unit (cpu) time low; this requires a trade-off between the number of iterative cycles of the input wave spectrum and cpu time.

We shall study the first two considerations regarding the production reliability in section 3.1 and study the third consideration of computational efficiency in section 3.2. A sensitivity analysis of the retrieval with respect to the first-guess input spectrum is given in section 3.3.

3.1. Data Quality and Retrieval Performance

The 3-year set of ERS-1 SWM data between January 1993 and December 1995 comprises a total of some 1.2 million wave spectra with an average daily output of 1100 spectra. The number of monthly available SWM spectra is shown in Figure 3. The dark column segments indicate the number of fast delivery product (FDP) spectra for which no first-guess model spectra were available. The envelope of the FDP columns shows a high value from the outset. A low FDP content was found during the summer seasons 1993 and 1994. For summer 1993 this can be explained by a change in the setup of the operational wave-forecasting facilities; the origin of the second minimum is not clear.

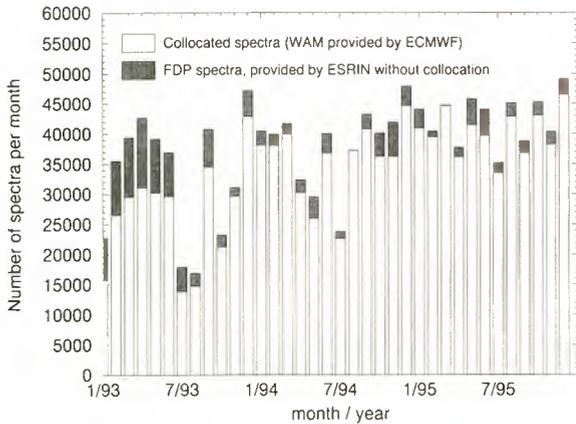


Figure 3. Number of monthly available ERS-1 synthetic aperture radar (SAR) wave model (SWM) data. SWM spectra that could be collocated with wave model (WAM) spectra are represented by the light lower column segments; those for which no first-guess model spectra were available are represented by the dark upper segments.

The retrieval process incorporated two quality tests: a data quality test for the input SWM data and a fidelity test for the retrieval. The input SWM data was rejected if the signal-to-noise ratio of the image spectrum (defined as the ratio of the spectral peak to the clutter noise background) was less than 3 dB or if the first-guess wave height was below 0.1 m. The retrieved spectra were discarded if the cost function was not reduced to less than half of its initial value. Furthermore, all spectral information beyond 800 m wavelength was excluded. These frequently represent nonwavelike sea surface features like slicks, wind rows, etc. [Brüning *et al.*, 1994]. Statistical analyses (compare section 4) indicate that swell of extreme wavelengths beyond 800 m are rare.

Figure 4 shows that with the beginning of the quasi real-time distribution of the FDP through the European Space Research Institute (ESRIN) in February 1993, maximally 2% of the data

showed low signal-to-noise ratios. Another 10% of the data were rejected because the first-guess wave height was too low ($H_s < 0.1$ m).

For roughly 85% of the remaining data the retrieval algorithm achieved an acceptable performance, with cost function reductions to <0.5 of the initial value, while for one third a reduction factor of <0.1 was achieved. Spectra with a cost function reduction factor lower than 0.5 for which the azimuthal clutter cutoff adjustment or the convergence of the iterative inversion procedure was poor were nevertheless retained in the subsequent analysis, as their fidelity measured in terms of the pattern correlation coefficient was generally still acceptable. The remaining 15% inversions, for which the ratio of the final to initial cost could not be reduced below 0.5, generally exhibited relatively low pattern correlations between the simulated and observed SAR spectra and were discarded from the further analysis. The number of monthly available SWM-retrieved spectra that passed both tests is geographically relatively homogeneous and is varied with the size of the ocean basin from 1500 in the North Atlantic to 5000 in the tropical Pacific.

The retrieval performance is represented in Figure 5 in terms of the distributions of the pattern correlation coefficients for the simulated and observed SAR spectra for the summer 1994 and winter 1994–1995 seasons in the northern and southern hemisphere. All distributions show a rather narrow peak around 0.9, indicating that for the major part of the data successful retrievals of high fidelity were achieved. The winter distributions in both hemispheres appear to be peaked toward slightly higher correlations than the summer distributions, suggesting that the retrieval performance is generally higher for high sea states (as may be expected from the generally higher signal-to-noise ratios for higher sea states).

3.2. Iteration Procedure

As described in section 2.2.2, five iterations of the input wave spectrum were carried out, and the inversion that yielded the smallest error between the observed and simulated SAR spectrum was then selected as the final retrieval. A time series of

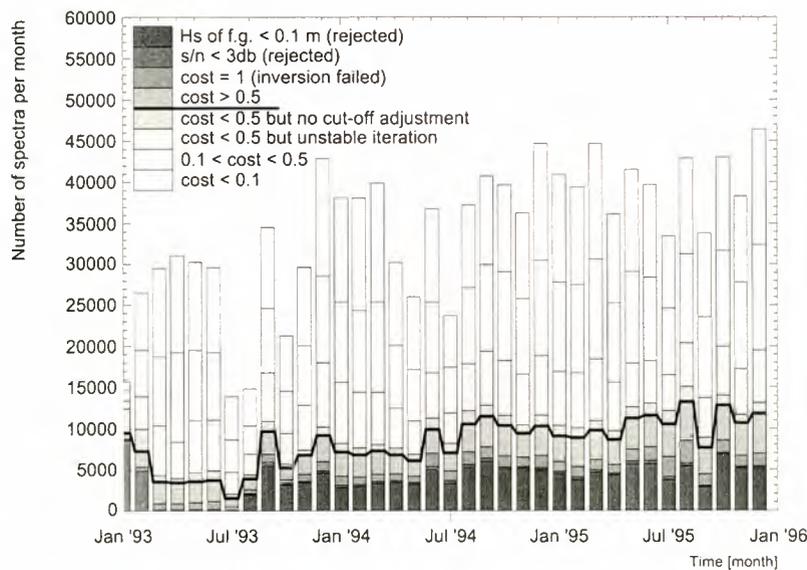


Figure 4. Data quality and performance of the retrieval algorithm measured in terms of the cost function reduction factor. Retrievals with a cost function reduction below 0.5 are regarded as reliable (above thick solid line). For data of too poor quality, no retrieval was attempted (below thick solid line).

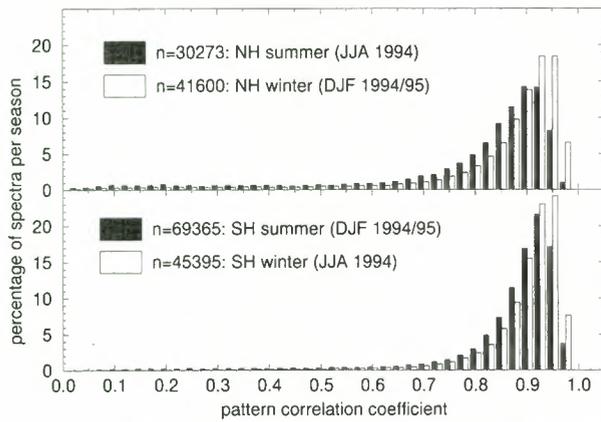


Figure 5. Fidelity of the retrieval algorithm measured in terms of the pattern correlation coefficient between the observed and simulated SAR image spectrum in the (top) northern hemisphere and (bottom) southern hemisphere.

the monthly distribution of the number of iterations yielding the best retrieval revealed a stable behavior throughout the 3 years. Table 1 shows the average percentage of spectra for which N iterations of the input spectrum yielded an optimal retrieval. For about 92% of the retrievals an iteration of the input spectrum yielded an improvement, 75% showing a further improvement for more than one iteration. For operational applications a trade-off must be sought between computing time and the improvement in retrieval with increasing number of iterations. We discuss the trade-off first with respect to the impact on the retrieval fidelity, as expressed by the pattern correlation between the simulated and observed spectrum, and subsequently with respect to actually retrieved wave parameters.

The rate at which the error between the simulated and observed image spectrum decreases with the number of iterations, given the number of iterations N for an optimal retrieval, is illustrated in Figures 6a and 6b, which show the histograms of the pattern correlation for August 1995 northern and southern hemisphere, respectively. The total set of spectra for each hemisphere is divided into five samples S_0, \dots, S_4 , where S_N denotes the set of spectra for which the optimal retrieval was achieved after N iterations (the last iteration set S_5 is not considered in this analysis because it also contains spectra that would have required more than five iterations for an optimal retrieval). All histograms are normalized to unit integral values. The graphs correspond, from bottom to top, to spectra requiring zero, one, two, etc. iterations for optimal retrieval, and for each sample S_N the evolution of the histogram from iteration to iteration is depicted by $N + 1$ distribution curves.

Table 1. Percentage of Spectra for Which N Iterations Yielded an Optimal Retrieval

Number of Iterations (N)	Number of Spectra, %
0	8
1	17
2	17
3	15
4	13
≥ 5	30

The behavior is similar in both hemispheres. The largest impact on the retrieval is found in performing the first iteration. The impact of subsequent iterations diminishes with increasing N .

As an alternative way of assessing the impact of the number of iterations on the retrieval performance, we used the same data sample of August 1995 to carry out a series of retrieval experiments in which the number of input iterations was set at a different fixed number M in each experiment. The best of the M iterations, defined in terms of the dimensionless error, was then taken as the retrieval. For each experiment this yielded a distribution curve of the pattern correlation index, shown in Figure 7 for the northern and southern hemispheres. The distributions are qualitatively similar to Figure 6. Both results suggest that at least one iteration should be carried out and that two iterations may represent an acceptable trade-off of retrieval improvement against computing time; the additional improvements beyond $M = 2$ tend to be rather small.

To assess the impact of the number of iterations not only on the fidelity measure but also on the retrieved wave parameters, a regression was carried out between the significant wave heights H_s of the retrieved spectra with and without iteration of the input spectrum. The regression parameters listed in Table 2 confirm that the retrieved wave heights do indeed converge with increasing iteration. The largest change is found as before for the first iteration, but the step from one to two iterations is also nonnegligible, supporting the previous conclusion that it appears advisable to carry out at least one but preferably two iterations.

3.3. Sensitivity of the Retrieval With Respect to the First Guess

Through the inclusion of the deviation between the first-guess (fg) input spectrum and the inverted spectrum as a (weakly weighted) regularization term in the cost function, the retrieved spectrum tends to be biased toward the input spectrum. However, the bias is systematically reduced in the HBHH scheme through successive iterations of the first-guess input. In this section we investigate the sensitivity of the retrieval with respect to the first-guess input spectrum.

Three forms of synthetic modifications of the input spectrum were investigated: an increase or decrease in energy, a frequency scale change, and a rotation. In all cases the modifications were largely eliminated in the retrieved spectra. We discuss in the following in detail only the impact of doubling the wave height or increasing the energy by a factor of 4. The impact is illustrated first for four typical wave spectra corresponding to the various combinations of azimuth or range traveling waves in strong or light winds.

Plates 1a–1d depict the various retrieval steps of the four spectra. In all cases it is apparent that the first inversion already removes most of the artificial error introduced into the modified wave spectra, but the error is suppressed still further in the subsequent iteration of the input spectrum. Table 3 summarizes the impact of the retrieval on characteristic wave parameters of the four spectra. In all four cases the significant wave heights retrieved from ERS-1 SWM data using modified WAM fg spectra H_s (modified (mod), retrievals (re)) agree closely to those H_s (unmodified (orig), re) retrieved using the original WAM fg spectra. The azimuthal cutoff wavelengths $\lambda_{az}(\dots, re)$ could be reliably adjusted to the observed (obs) values $\lambda_{az}(obs)$. Plates 1a–1d clearly illustrate the relationship between the cutoff length scale and the high-frequency energy

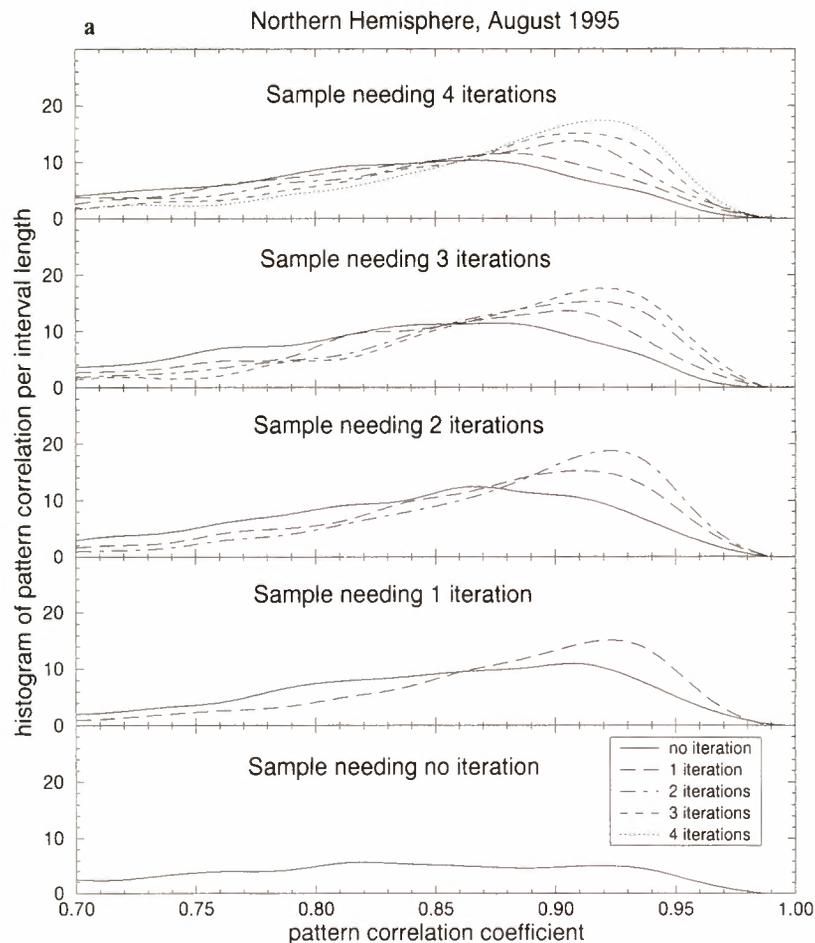


Figure 6. Distribution of pattern correlation in the (a) northern hemisphere and (b) southern hemisphere. The data sets of each panel consist of all spectra for which exactly N iterations produced the optimal retrieval (see Table 1). The curves within each panel indicate the improvement in pattern correlation with increasing number of iterations.

content (see, e.g., the SAR spectrum simulated from the modified input spectrum in Plates 1a and 1d). While for the original WAM fg spectra less than five iterations were often sufficient to find the best estimate, the modified WAM fg spectra required all five iterations to adjust to the best solutions. Table 3 shows that the initial errors $\varepsilon(\cdot, \cdot, fg)$ between the observed and simulated SAR spectrum are reduced considerably through the retrieval process.

To assess the sensitivity of the retrieval for a larger ensemble of cases, we have repeated the retrieval using WAM fg spectra similarly increased by a factor of 4 for a 1-day set of some 1100 ERS-1 SWM spectra on June 7, 1995. For the analysis we have considered only retrievals for which both the original and the modified retrievals yielded an acceptable fidelity level, corresponding to a normalized SAR spectral error $\varepsilon < 0.3$. As can be expected, this eliminated some of the cases that passed the quality test for the original spectra but not for the modified spectra.

Figure 8 shows the regression between the significant wave heights retrieved using the original and modified WAM fg input spectra. The regression coefficient is 1.15. Thus the retrieval reduced the relative error of the modified fg wave heights on average by a factor of 6, from 1 to 0.15. The error would have been reduced still further if more iterations of the input spectrum had been carried out. However, typical errors

of the first-guess wave fields are significantly smaller than the synthetic errors considered here, so that five iterations are normally quite adequate (see section 3.2).

4. Statistical Intercomparisons

Global and regional intercomparisons between the collocated spectral wave parameters retrieved from ERS-1 SWM spectra and computed with the WAM model were carried out for the 3-year period January 1, 1993, to December 31, 1995. The collocation was performed using the nearest model grid point and nearest 6-hour model synoptic time to the observation location and time.

In a first analysis the data were divided into eight ocean basin regions, namely, the North, tropical, and South Pacific; the North, tropical, and South Atlantic; and the tropical and South Indian Ocean, with tropical boundaries at 25°N–25°S. Linear regressions were performed of the collocated data for monthly, seasonal, and the full-period data sets (see appendix, section A3). Figure 9 shows as an example the seasonal regression between WAM model and SWM-retrieved H_s values in the northern hemisphere extratropics, the tropics, and the southern hemisphere extratropics for the northern hemisphere winter December, January, and February (DJF) 1993–1994.

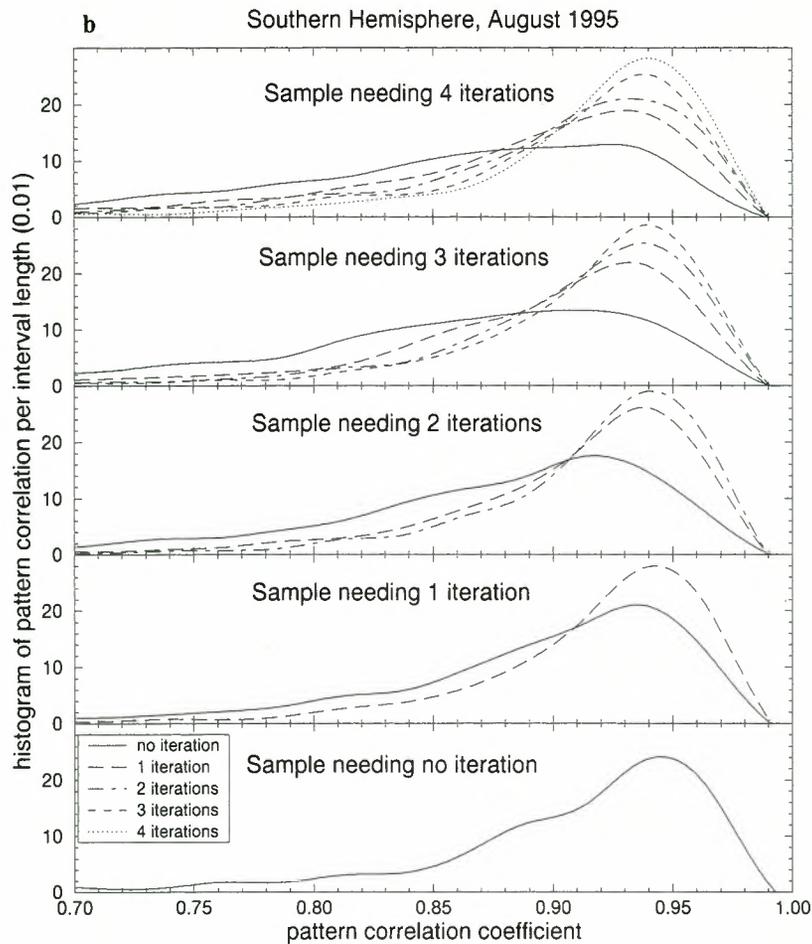


Figure 6. (continued)

4.1. Comparison of Integral Properties

4.1.1. Monthly mean H_s . Figure 10 compares the time series of monthly mean significant wave height in different regions of the Atlantic, Pacific, and Indian Ocean retrieved from ERS-1 SWM (dashed line) and computed by the WAM model (dotted line). The overall agreement is satisfactory. There is a more pronounced annual cycle in the northern hemisphere than in the southern hemisphere, the summer wave heights decreasing more strongly to about 1.5 m in the north as compared to 2.5 m in the south. This reflects the higher abundance of low-pressure wind systems that still prevail in the southern midlatitude zonal belt during the summer season. In the northern hemisphere the monthly mean H_s values are higher (by about 10%) in the Atlantic than in the Pacific, while the opposite holds for the southern hemisphere (20% higher values in the Pacific than in the Atlantic). The maxima of the South Atlantic winter are also less strongly pronounced than in the other southern and northern basins for the winter seasons. The tropical Atlantic and Pacific show essentially no annual cycle, with monthly mean H_s values fluctuating about 2 m in the tropical Pacific and 1.75 m in the tropical Atlantic. In contrast, the tropical Indian Ocean exhibits a marked seasonal cycle reflecting the monsoons. Pronounced maxima comparable to those in the North Atlantic (although somewhat broader) are found also in the Southern Indian Ocean.

4.1.2. Bias. Despite the satisfactory overall agreement between modeled and retrieved wave heights, a small but systematic underestimation of H_s by the WAM model relative to the ERS-1 SWM data is seen in Figure 11 (solid line). The bias is more pronounced for low sea state summer conditions. A different behavior for different ocean basins is again apparent. In the northern hemisphere the negative bias shows a pronounced seasonal cycle, with larger values during summer (low sea states) than winter (high sea states). The bias in the tropical Pacific and Atlantic oceans also exhibits a weak seasonal cycle that follows the northern hemisphere, in contrast with the bias of the tropical Indian Ocean, which follows the Southern Indian Ocean. The southern hemisphere basins, finally, exhibit an approximately constant negative bias without a significant seasonal modulation.

4.1.3. Regression line slope. In addition to the bias a second useful parameter characterizing the statistical relation between modeled and retrieved wave heights is the slope b of the regression line (we have defined the regression line here as passing through the zero point of the axes, for details see appendix, section A3). Figure 12 (solid line) shows time series of b for the monthly data sets. Values greater (smaller) than 1 represent under(over)-estimated model values compared to observation. The time series for b are seen to follow the behavior of the bias, as expected.

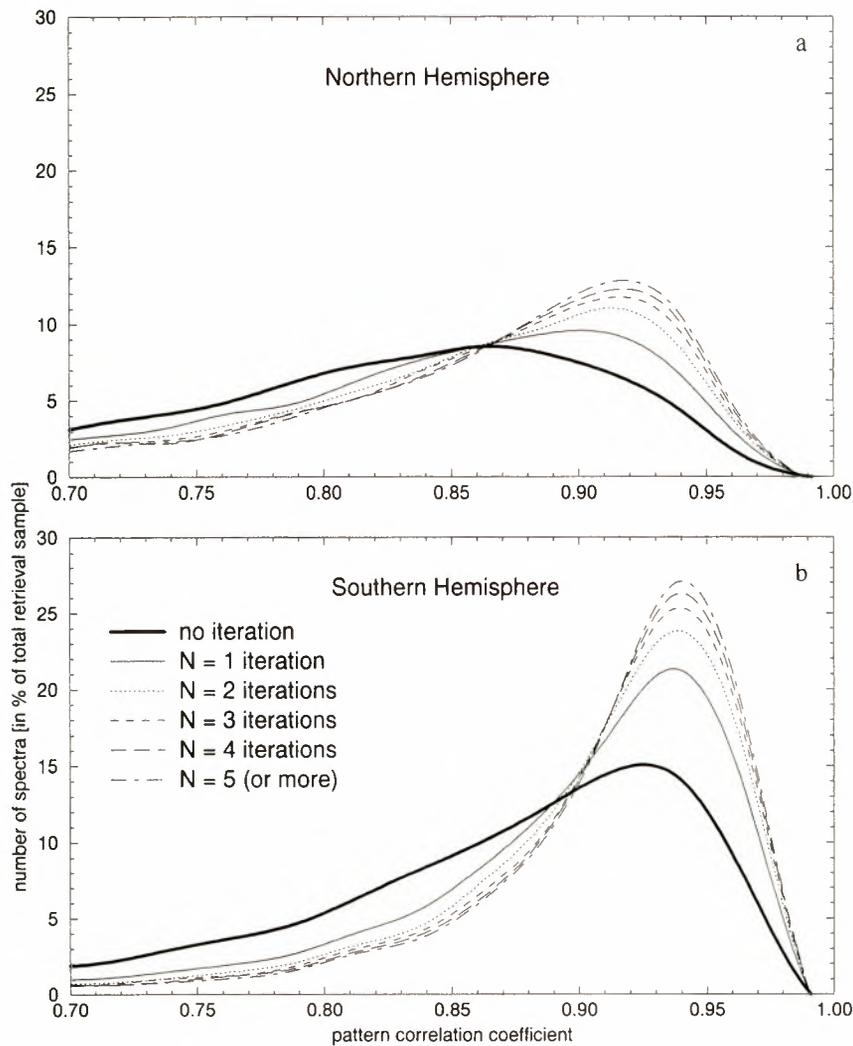


Figure 7. Distribution of pattern correlation in the (a) northern hemisphere and (b) southern hemisphere. In contrast to Figure 6, each panel is based on the same set. The data set is subjected to a series of retrieval runs, each retrieval differing in the prescribed maximum number of iterations allowed. The retrieval fidelity increases with increasing maximum permitted number of iterations.

4.1.4. Correlation. Time series of monthly correlation coefficients obtained from the regression are shown in Figure 13 (solid line). In the extratropical northern and southern hemisphere, fairly constant values of the order of 0.8 are found, except for a weak seasonal cycle in the North Pacific, with minima corresponding to low northern hemisphere summer sea states. The tropical Atlantic and Pacific exhibit somewhat lower correlation levels between 0.5 and 0.7. Winds in the tropics are generally low, so that the wave spectra are domi-

nated by swell (see section 4.2). This suggests that the degree of agreement between model and observation is generally higher for locally generated wind seas than for swell that has traveled to the observation location from distant sources. This interpretation will be supported later in section 4.2 by the correlation time series of partitioned wind-sea and swell systems.

4.2. Comparison of Spectral Partitionings

A more detailed intercomparison of modeled and retrieved wave spectra can be obtained by partitioning the spectra into wave systems (compare section 2.2.2; for details of the partitioning system, the method for cross-assigning individual wave systems of the collocated modeled and retrieved wave spectra, and the criteria applied in the discrimination between young and old wind sea, mixed wind sea–swell, and pure swell systems, we refer to HBHH). For the present analysis we have included in the wind-sea category young and old wind-sea systems (as defined by HBHH), while swell refers to “pure” swell. Mixed wind-sea–swell systems are not included in either category and are also not presented as a separate wave system

Table 2. Change in Mean Retrieved H_s and Regression Parameters With Increasing Number of Iterations for a 1-Month Global Sample of ERS-1 SWM Data

	Iteration					
	0	1	2	3	4	5
H_s , m	3.01	3.17	3.25	3.29	3.31	3.33
Slope	...	0.993	1.011	1.020	1.028	1.036
Correlation	...	0.942	0.897	0.874	0.859	0.846

There are 19,129 entries.

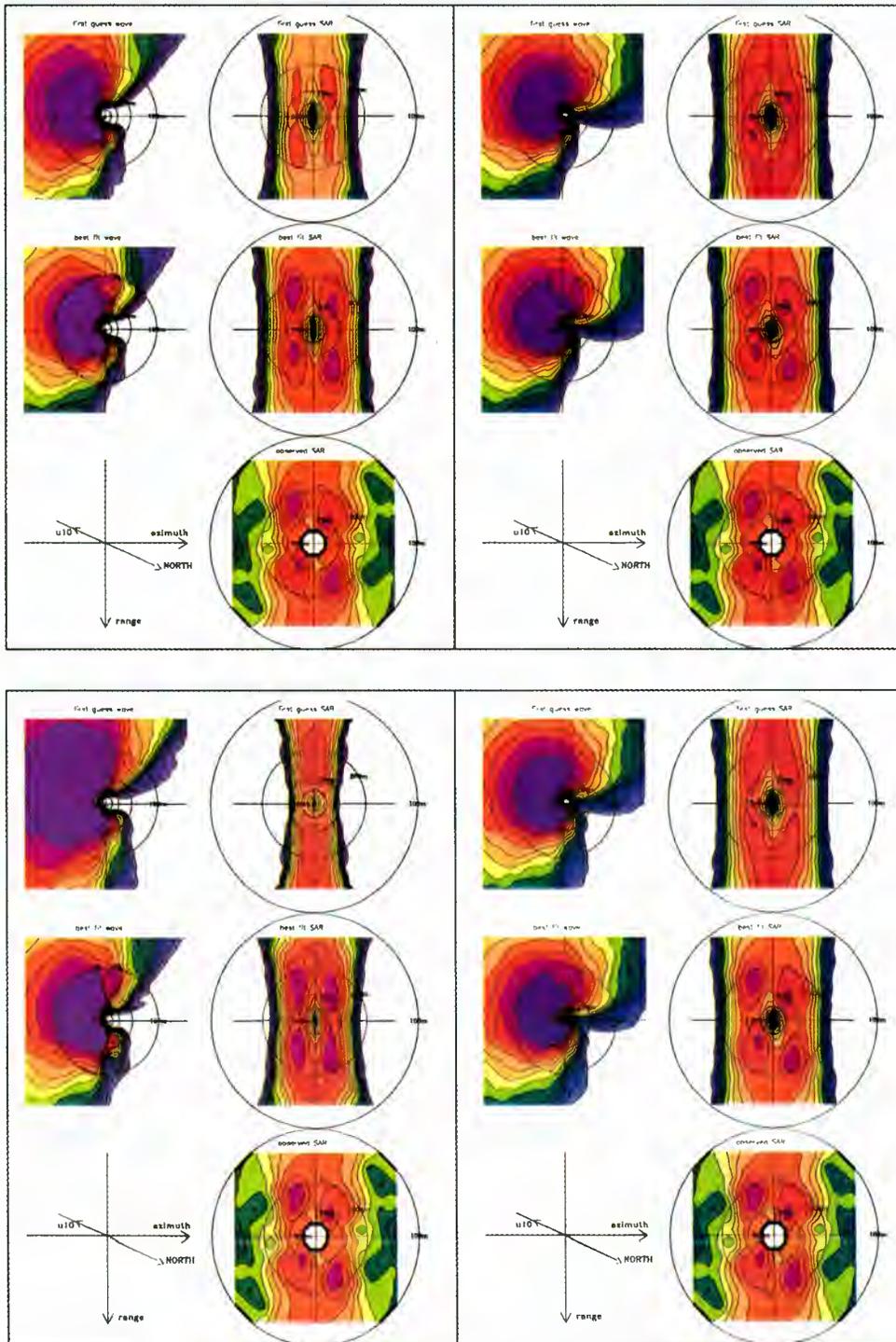


Plate 1a. Input and retrieved wave spectra with associated SAR spectra for azimuthally traveling waves under strong winds. Top two quadrants show unmodified input spectrum; bottom two quadrants show input spectrum increased by a factor of 4. Left quadrants show straightforward inversion without iteration of the input spectrum; right quadrants show the best retrieval of five iterations of the input spectrum. First row in each quadrant shows the (left) input spectrum (in the case of the optimally iterated input spectrum this is identical to the retrieval) and (right) associated computed SAR spectrum; second row in each quadrant shows the inverted wave spectrum (prior to application of smoothing and interpolation by wave system partitioning) and associated computed SAR spectrum; bottom row, right side in each quadrant shows observed SAR spectrum.

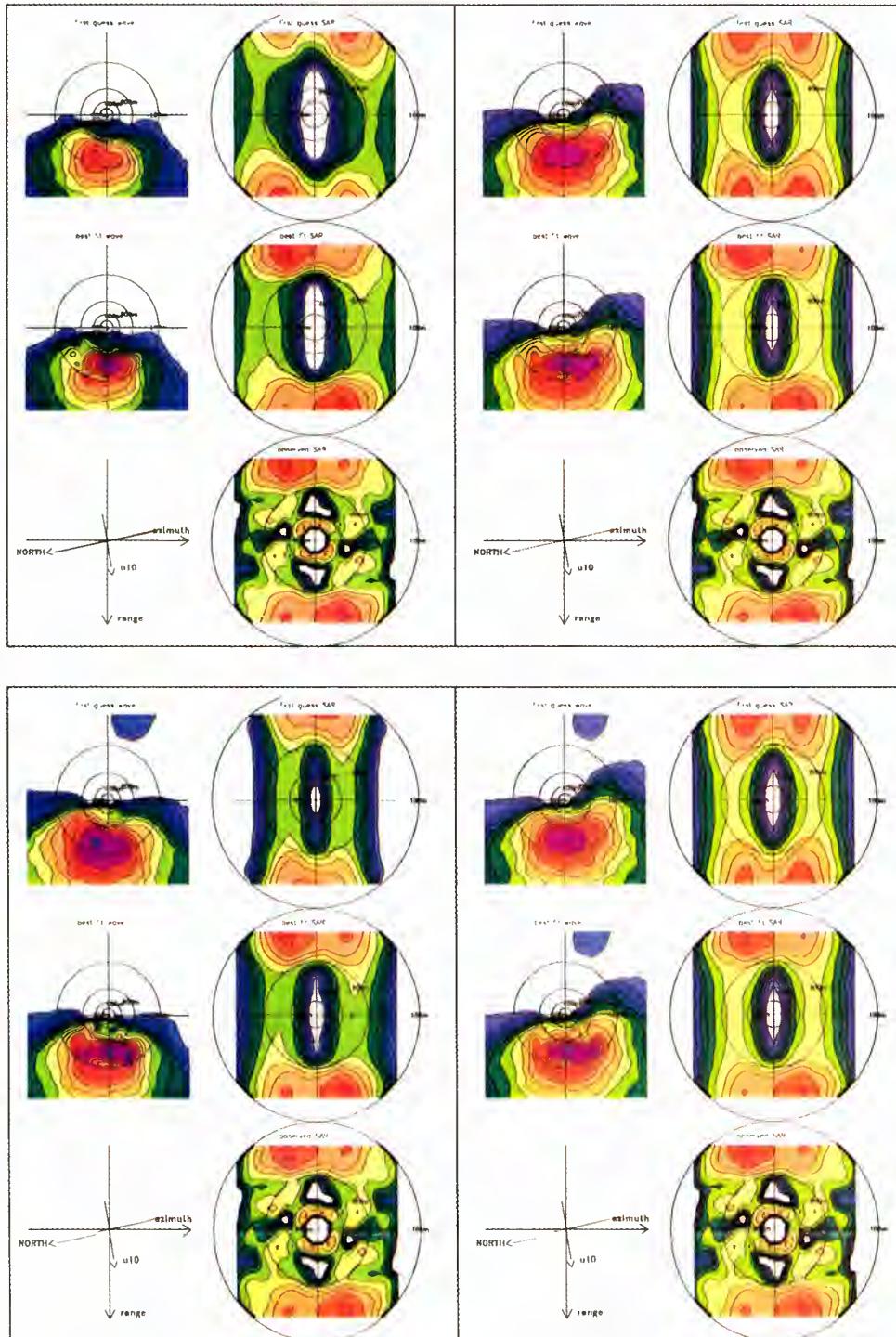


Plate 1b. Same as Plate 1a but for range traveling waves under light winds.

class since their contribute to the total wave energy is relatively small.

Plate 2 shows as an example the global distributions of the modeled and observed average number of swell systems per spectrum for the northern hemisphere winter DJF 1993–1994. The largest number of swell systems is found in the tropics, which receive swell from the midlatitude high wind regimes in both hemispheres as well as locally generated swell by the fluctuating trade winds. The ability of both the model and the SAR system to distinguish between a relatively large number of

different wave systems and the agreement between the simulated and observed distributions of this characteristic structural spectral parameter is encouraging.

The corresponding distributions for wind-sea systems are not shown. Theoretically, they should be nearly uniform at a value ≈ 1 , since a wind sea is always present as soon as a light breeze occurs. However, for the operational WAM model, which was integrated with a resolution of $3^\circ \times 3^\circ$ (later $1.5^\circ \times 1.5^\circ$), such light winds cases, as well as typical wind-sea situations with changing wind directions, yielded wave systems that

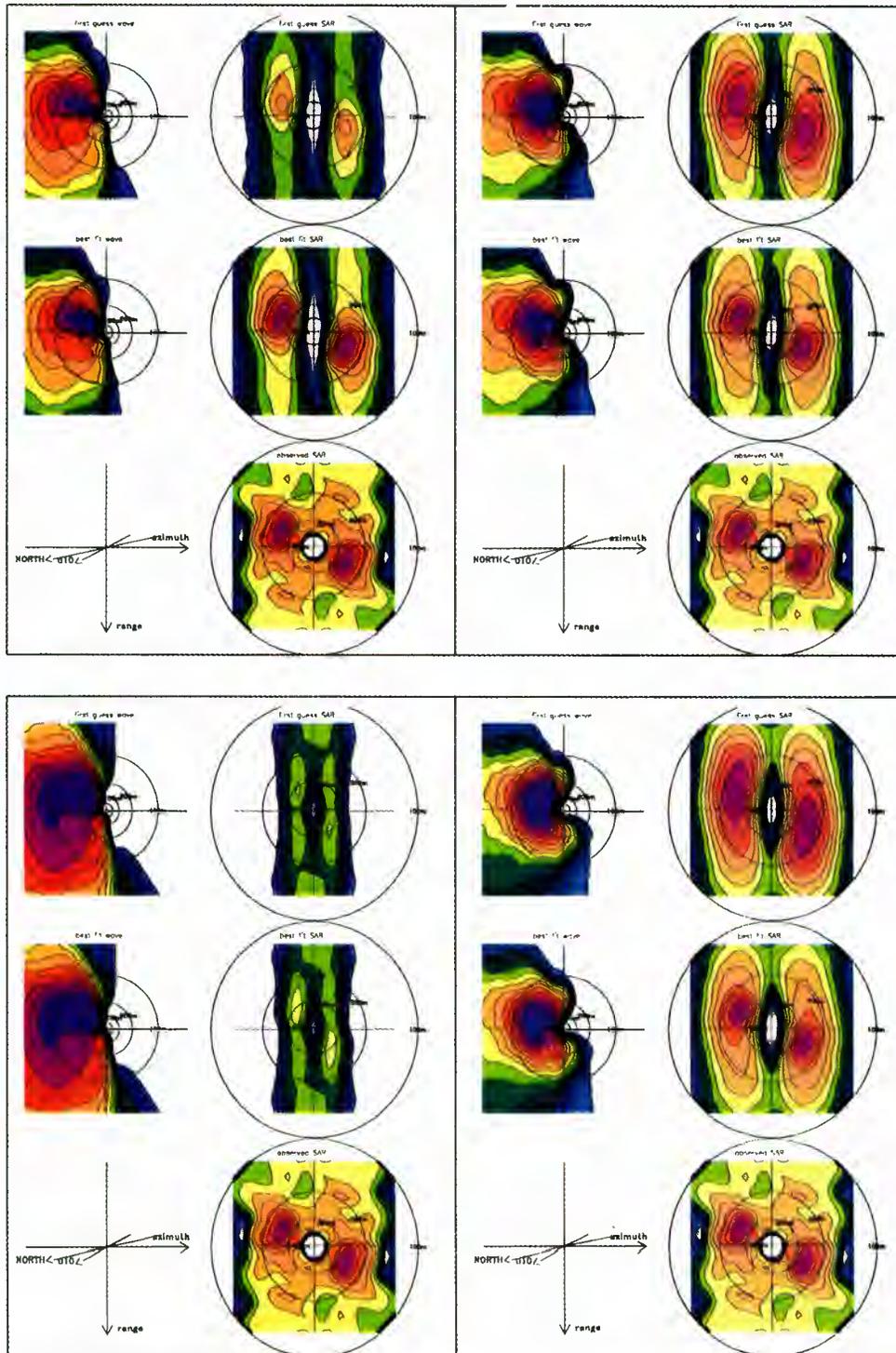


Plate 1c. Same as Plate 1a but for azimuthally traveling waves under light winds.

were normally classed as mixed wind-sea–swell systems rather than pure wind sea (compare HBHH). The SAR is similarly unable to detect or adequately resolve wind-sea systems for low to moderate wind speeds because of the azimuthal high wavenumber cutoff (although, as explained above, partially resolved wind seas are reconstructed in the retrieval system of HBHH using the information on the azimuthal cutoff in combination with the partitioning scheme incorporated in the iterative updating of the input spectrum). The data on wind-sea systems presented in the following refer therefore only to wind seas generated by fairly

large, moderate to strong wind fields that produce spectra satisfying the pure young or old wind-sea criteria of HBHH.

The time series of the monthly mean swell wave heights H_{sw} (defined as the significant wave height corresponding to the total swell energy summed over all swell systems) are shown in Figure 14 for the modeled (dashed lines) and retrieved (long-dashed lines) swell in different basins of the Atlantic, Pacific, and Indian Ocean. Also shown are the simulated and observed monthly mean wind-sea wave heights H_{ws} (dotted and dotted-dashed lines, respectively). The general agreement between

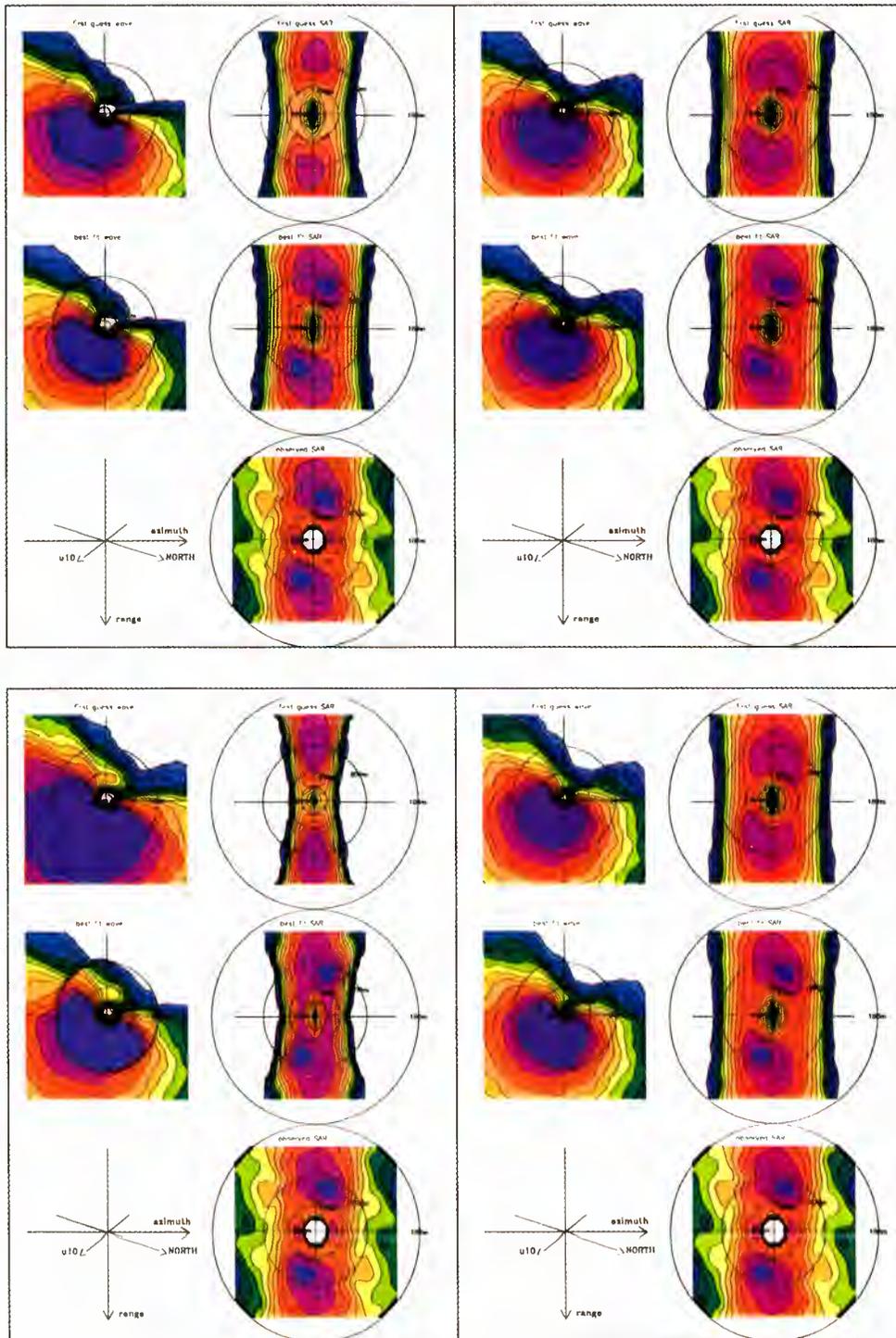


Plate 1d. Same as Plate 1a but for near-range traveling waves under strong winds.

model and observation and the distinctive annual cycles is again confirmed.

However, a discrimination is now possible with respect to the small systematic deviations noted earlier between the simulated and observed total wave heights. Whereas the WAM wind sea tends to be slightly overpredicted by about 10% relative to the retrieved data, with larger relative deviations for high sea states corresponding to winter conditions, the swell wave heights are systematically underestimated by about 20–30%.

The overestimation of wind sea is most pronounced for high

monthly mean H_{wi} during winter seasons, whereas the underestimation of swell, especially in the northern hemisphere, is seen most clearly for low swell, i.e., during summer. The biases for wind sea and swell in the northern hemisphere exhibit strong seasonal cycles, whereas in the southern hemisphere they remain fairly constant. We note that the swell biases exhibit a slight downward trend that is not present in the wind-sea biases. The time series of the slope of the monthly regression lines for wind sea and swell (Figure 12, dotted and dashed lines, respectively) follow the bias lines.

Table 3. Retrieval Impact on Characteristic Wave Parameters of Four ERS-1 SWM Spectra

	Spectrum Number			
	1	2	3	4
Date	June 2, 1995	June 6, 1995	June 7, 1995	June 8, 1995
Time, UTC	0128	1003	0726	1907
u_{10} , m/s	19.0	4.0	7.6	20.5
θ_{wind} , deg	180	272	345	126
ϕ_{az} , deg	10	99	3	126
H_{sw} , m				
orig, re	5.3	1.7	2.3	5.5
mod, re	5.5	1.6	2.8	5.6
λ_{az} , m				
obs	259	191	181	239
orig, fg	266	185	188	259
orig, re	258	187	180	242
mod, fg	299	197	>400	309
mod, re	259	188	178	242
n_{iter}				
orig	5	5	3	2
mod	5	5	5	5
ϵ				
orig, fg	0.68	0.48	2.46	0.66
orig, re	0.10	0.09	0.17	0.08
mod, fg	0.77	0.41	6.55	0.52
mod, re	0.11	0.09	0.45	0.10

Retrievals (re); unmodified (orig) and modified (mod) WAM first-guess (fg) spectra.

As the magnitudes of the wind-sea and swell wave heights are of the same order, the different features of the wind-sea and swell biases cannot be explained simply by a dependence on wave height but must be attributed to the different spectral and dynamical properties of wind sea and swell. We have considered four possible causes for the differences.

1. The hydrodynamic transfer function occurring in the SAR mapping relation is poorly known. The present nonlinear forward mapping uses the theoretical modulation transfer function (MTF) proposed by Feindt [1985] without the wind dependent terms (see HH). Experiments with a wind dependent MTF showed little influence on the retrieval in the case of ERS-1 SWM spectra (W. R. Plant, private communication, 1995). This is to be expected for spaceborne SARs for which the velocity bunching mechanism is the dominant mapping contribution due to the high R/V ratio [see Alpers *et al.*, 1981, Figure 6; Br uning, 1994]. In the limited domain of linear mapping for range traveling waves, where the RAR MTF becomes important, it contains also a significant contribution from the geometrically well-defined tilt modulation MTF. Errors in the hydrodynamic MTF should appear as systematic errors in only a small subset of data representing range traveling waves, rather than in a difference between wind sea and swell distributed rather uniformly over all data. Thus the hypothesis of errors in the hydrodynamic MTF causing differences in the statistics of wind-sea and swell errors appears improbable.

2. Recent results from wave data assimilation exercises in which both the wave field and the generating wind field were updated suggest that in certain areas of the Atlantic and Pacific the ECMWF-analyzed wind fields produce too strong winds and thus too strong wind seas [see, e.g., Bauer *et al.*, 1996a, b; Hasselmann *et al.*, 1997; Heimbach *et al.*, 1996]. However, if this were the only error, an overestimation of the wind sea would necessarily lead also to an overestimation of the swell that

emerges from the storm region, which is not found. Another source of error must therefore be sought.

3. The combination of an overestimation of the wind sea together with an underestimation of swell suggests a possible error in the balance between wind-sea and swell dissipation. The balance could be modified to conform with the SAR data either by increasing the exponent characterizing the frequency dependence of the dissipation source function, which would decrease the ratio of swell to wind-sea dissipation, or by increasing the nonlinearity dependence, which would similarly result in a relatively weaker dissipation of long swell, which is characterized by smaller wave slopes.

4. The transition regime between wind sea and swell in the area close to but outside the immediate wind generation region, where the wind forcing has already decreased but the nonlinear interactions are still important, plays a critical role in determining the ratio of wind-sea to swell energy [see, e.g., Snodgrass *et al.*, 1966; S. Hasselmann *et al.*, 1985; Young *et al.*, 1987]. It is conceivable that the strongly simplified discrete interaction parametrization of the nonlinear energy transfer implemented in the WAM model is inadequate for a proper representation of this process. Deficiencies in the parametrization of the nonlinear energy transfer may thus be considered as another candidate for explaining the different biases.

We have also considered numerical errors associated with the first-order upwind propagation scheme employed in the WAM model. In developing the WAM model, two schemes were tested: a first-order upwind scheme and a second-order leapfrog scheme [WAMDI Group, 1988]. While the first-order scheme produced a higher numerical dispersion, a diffusion term had to be introduced in the second-order scheme to suppress negative energies occurring in regions of high gradients, which resulted in a comparable effective diffusion. Because of the relatively small differences between the two schemes the simpler first-order scheme was chosen. A third-order scheme has also been proposed [Tolman, 1992, 1995]. However, numerical errors in the propagation scheme can be excluded in principle as the origin of the spatially and temporally averaged mean global bias in swell energy. Provided the propagation schemes satisfy the standard requirement of energy conservation, all schemes will necessarily yield the same globally averaged swell energy, even though they yield different spatial distributions for individual swell cases.

We regard the third candidate of a too strong damping for low-frequency swell as the most likely explanation for the negative swell bias. This is supported by a plot of the relative bias of the WAM wave heights relative to the SAR as a function of wave height for both wind sea and swell, shown in Figure 15 for the data sample June, July, and August (JJA) 1994. The wind-sea graphs show a fairly constant relative bias of 15–25% for wave heights $H_{wi} > 1$ m, except for a slight positive slope in the northern hemisphere. In contrast, all swell graphs exhibit a more or less pronounced change in relative bias from negative values at low H_{sw} (–35% in the North Pacific (NP), North Atlantic (NA), tropical Pacific (TP), and tropical Atlantic (TA) and –15% in the South Pacific (SP), South Atlantic (SA), South Indian (SI), and tropical Indian (TI)) to positive values at high H_{sw} (except in the SP, where a –10% bias remains).

This can be understood by the dependence of swell height on propagation distance. The farther a swell system has propagated, the lower its energy level due to geometrical dispersion (and, to a smaller extent, dissipation; compare Snodgrass *et al.* [1966]). At the same time, if the dissipation is too large, as

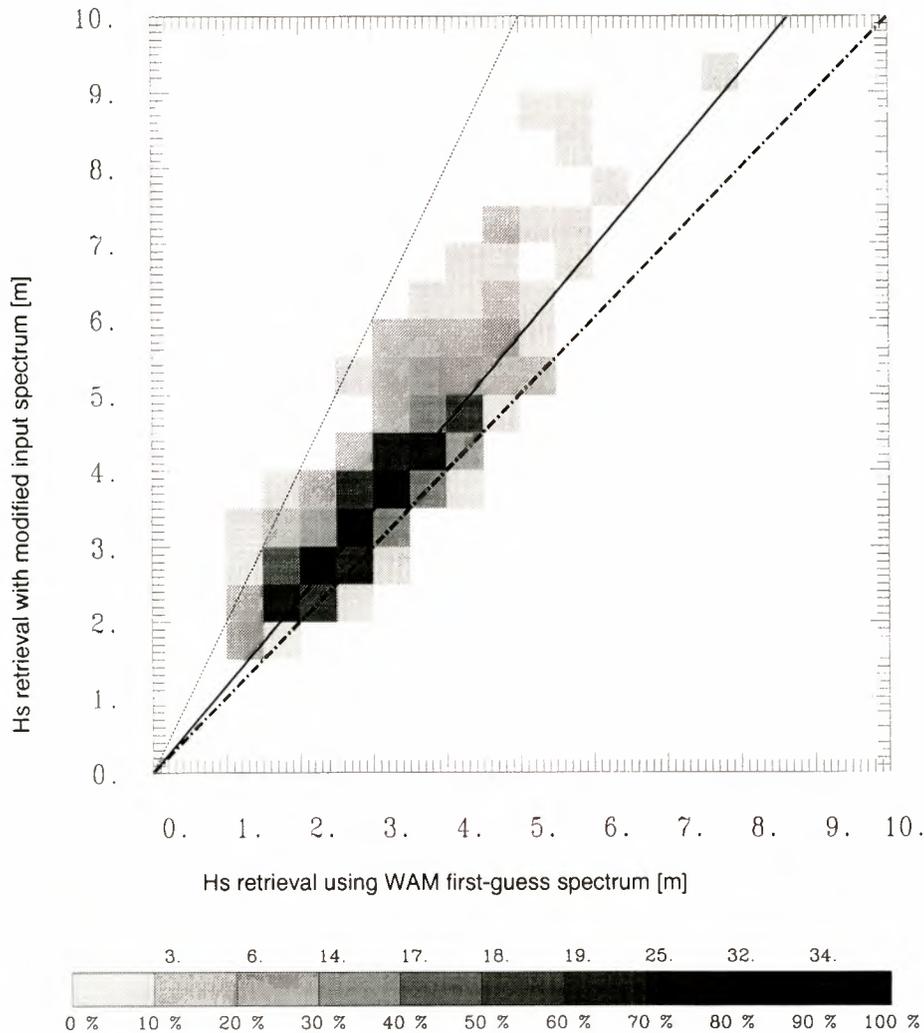


Figure 8. Regression between significant wave heights retrieved using the original and modified WAM first-guess input spectra for a 1-day set of ERS-1 SWM spectra on June 7, 1995.

hypothesized, the farther the swell has propagated, the larger the negative bias of the swell relative to the observations. Thus we should expect a direct correlation between the bias and the swell wave height, as seen in Figure 15.

The impact of the swell propagation distance on the bias could be demonstrated more directly if the swell propagation distance or, equivalently, swell age was computed by the WAM model. Unfortunately, this is not one of the output quantities of the standard WAM model. However, *Bauer et al.* [1996b] have developed a WAM model version for their Green function spectral data assimilation scheme which provides this information, and E. Bauer (personal communication, 1996) has kindly made available wave age data in the Atlantic for the period November 1992 with which we could test the assumed correlation between wave age and wave height. (The term wave age is meant here and in the following literally as the elapsed time since the waves last experienced a significant wind input, in contrast to the alternative usage of the term in wind-sea studies as the ratio of the phase velocity to the wind speed.) Figure 16 shows the average wave age over this period as a function of H_s for the North, tropical, and South Atlantic. The plots confirm that low H_s is associated on the average with high wave ages, especially in regions of low wind activity. Thus, in

the statistical average, it appears justified to regard the wave height in Figure 15 as a proxy for the wave age.

A remark needs to be made concerning the operational WAM model data. Since August 1993, ERS-1 altimeter wave height data have been assimilated by ECMWF into the WAM model using an optimal interpolation scheme, thereby influencing the first-guess wave field through the analysis 6 hours earlier. The swell wave field is thus continuously updated by the assimilation of new wave height data, thereby reducing the impact of systematic model errors in the propagation of swell over large distances.

However, an investigation of altimeter wave heights suggests that the impact on the first-guess data available for the present study is probably rather limited [see *Bauer and Staabs*, 1998; E. Bauer and P. Heimbach, manuscript in preparation, 1998]. First, no marked change in the monthly mean significant wave height time series can be seen for the northern hemisphere summer 1993 at the time when operational assimilation of altimeter H_s into the WAM model was introduced (Figure 10). Furthermore, the careful study by *Bauer and Staabs* [1998] revealed a systematic shift in the monthly mean wave height time series of the ERS-1 altimeter beginning in January 1994 relative to 1993 [see *Bauer and Staabs*, 1998, Figure 8], the

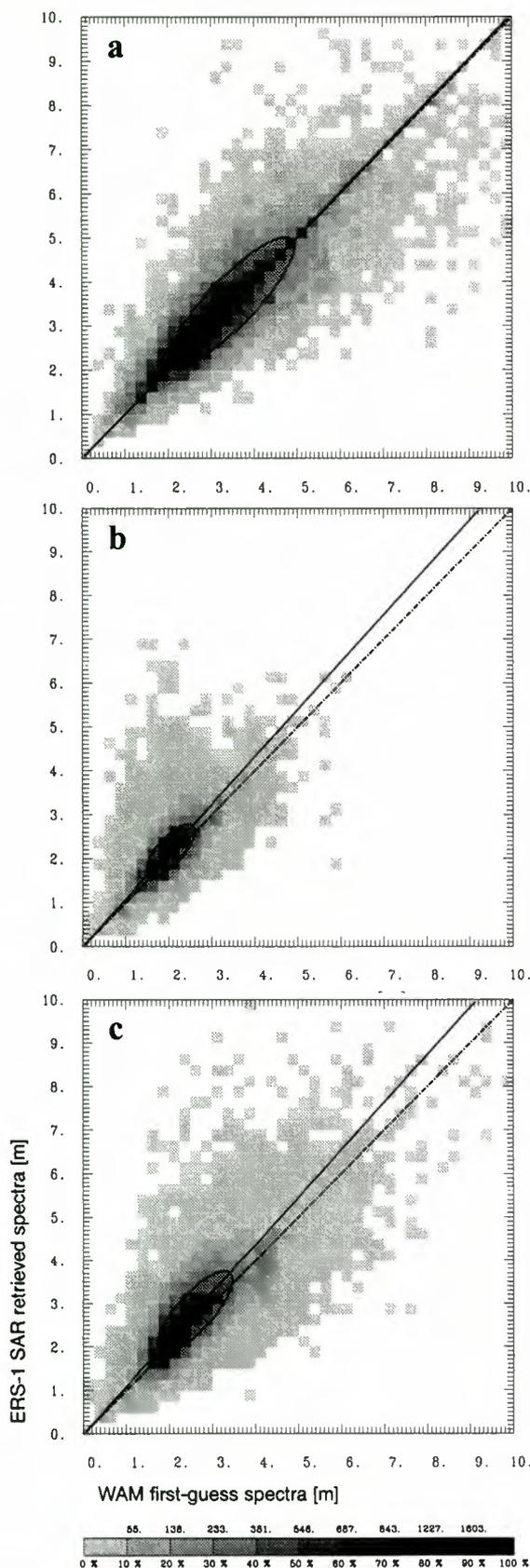


Figure 9. Comparison between ERS-1 SWM-retrieved and WAM model wave heights for DJF 1993–1994 in (a) the northern hemisphere extratropics, (b) the tropics (between 25°N and 25°S), and (c) the southern hemisphere extratropics. Solid lines represent the regression line and covariance ellipse; the dashed-dotted line represents the diagonal.

values for 1994 being reduced by as much as 0.6 m in the northern hemisphere extratropics and in the tropics. This results in a systematic shift in the bias between WAM and the ERS-1 altimeter wave heights in the transition from 1993 to 1994. If the assimilation of altimeter data had a significant impact on the first-guess data, a shift in the H_s time series for the WAM toward lower values should have been noticeable both in the summer of 1993, when the assimilation scheme was introduced, and at the end of 1993, when the ERS-1 altimeter algorithm was apparently modified. Although some indication of this effect may be apparent in the northern hemisphere extratropics, no influence is seen in the tropics, where the swell contribution is largest. We are therefore reasonably confident that although our model data are undoubtedly affected by the assimilation of altimeter data, the dynamical aspects we discuss later are still apparent in the data set.

A more quantitative analysis of the wind-sea and swell bias requires a regional and directional stratification of the data, as discussed in section 4.4. A still more illuminating approach would probably be to study the propagation of individual wave systems, using the WAM model without data assimilation. However, this is beyond the scope of the present statistical analysis.

The time series of the correlation coefficient obtained from the regression analysis show another interesting difference between wind-sea and swell systems (Figure 13, dotted and dashed lines, respectively). Whereas the correlation for wind sea maintains a fairly constant high level of 0.9, the correlation for swell is lower and undergoes stronger variations between 0.6 and 0.8 in the northern hemisphere and in the tropics. Only the southern hemisphere exhibits a fairly constant correlation for swell of 0.8. This can be explained by the large and variable wave ages (propagation distances) of swell systems. Accepting the interpretation of the observed swell bias as resulting from an overestimation of the swell dissipation, the ratio of the predicted to measured swell energy will be a function of the propagation distance. The large range of propagation distances occurring in the ocean will therefore reduce the correlation between the modeled and observed swell wave heights and will lead also to variations in the correlation level depending on the season, i.e., the location of the storm areas.

4.3. Regional and Seasonal Intercomparison of Modeled and Retrieved Wind Sea and Swell

Further insight can be gained by looking at regionally stratified data as a function of season. To this end we have mapped all collocated WAM and SAR wave system data, subdivided again into wind sea and swell, for four consecutive seasons from winter (DJF) 1993–1994 until fall (September, October, and November (SON)) 1994 on a $5^\circ \times 5^\circ$ global ocean grid. Plates 3a and 3b and Plates 4a and 4b show the distributions for the wind-sea and swell wave heights, respectively, while the associated “directional wave roses” representing the energy-weighted distributions of the wave propagation directions in the different ocean basins are reproduced in Figures 17a and 17b. The directional wave roses were computed as the square root of the mean energy for each 30° angular sector.

4.3.1. Wind-sea wave heights. The wind-sea distributions (Plates 3a and 3b) reflect the seasonal properties of the atmospheric circulation [see, e.g., *Peixoto and Oort, 1992*]. In the northern and southern hemisphere extratropics the wind-sea wave heights are governed by the midlatitude westerlies. The regions of strongest wind-sea activity in DJF (northern hemi-

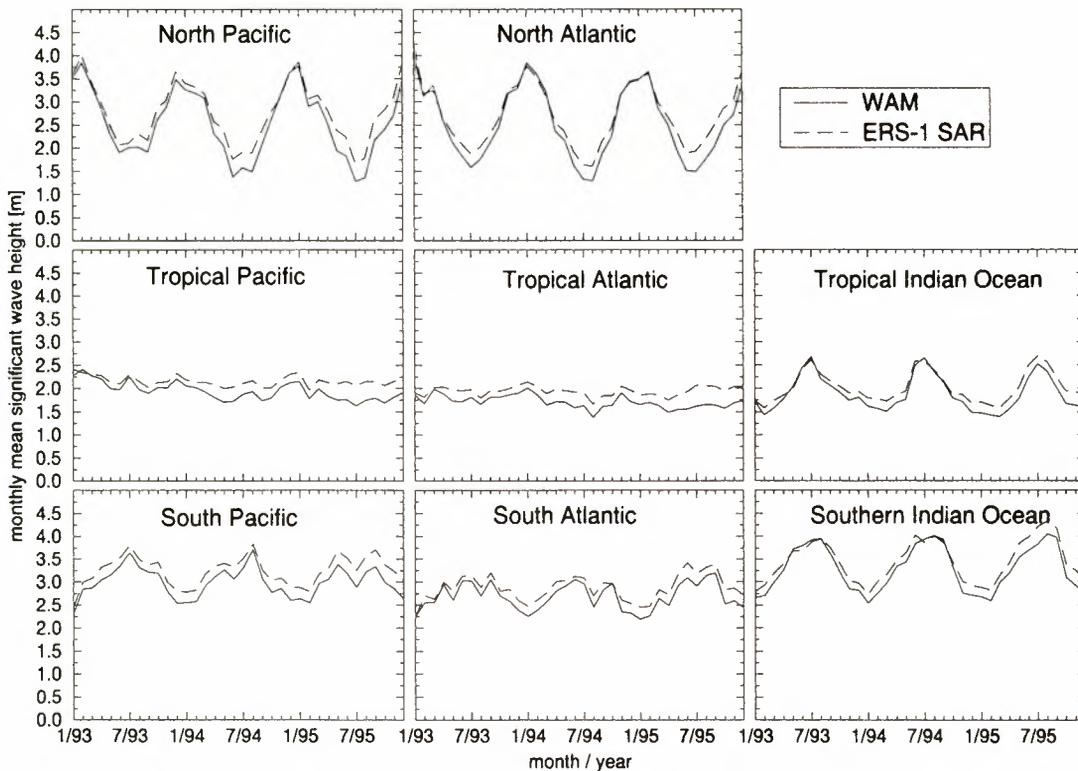


Figure 10. Time series of monthly mean significant wave heights between January 1993 and December 1995 for different ocean basins. Dotted lines refer to WAM model values, and dashed lines refer to ERS-1 SWM values.

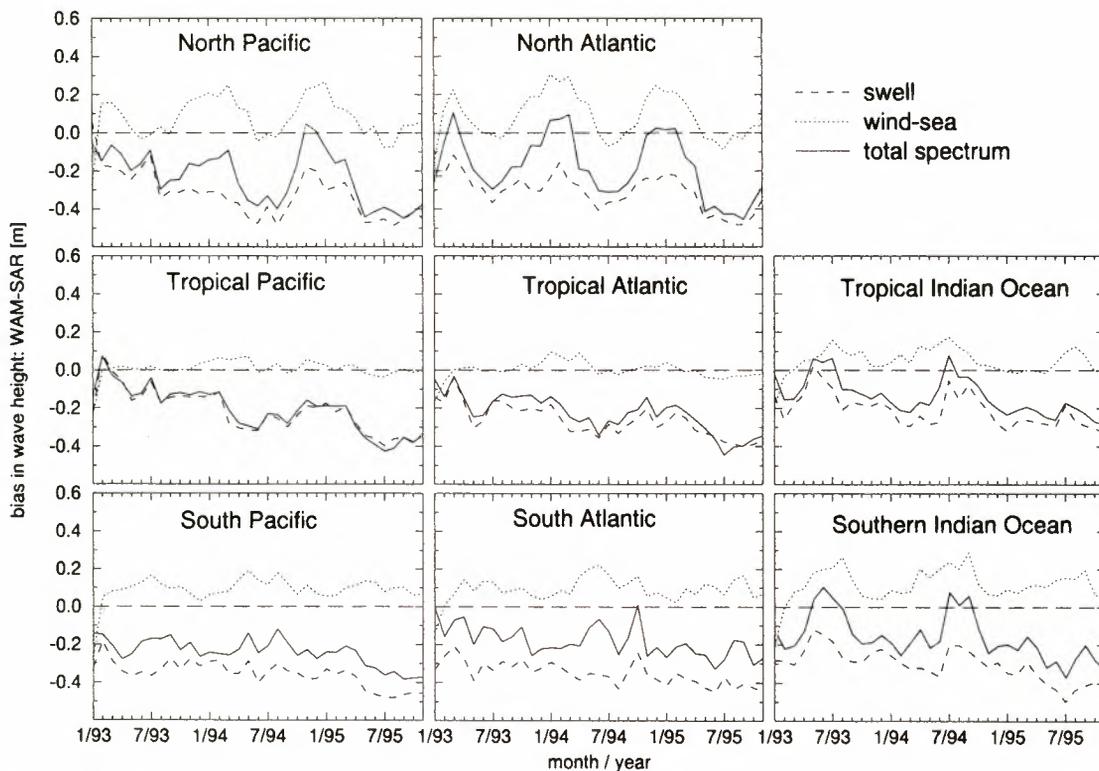


Figure 11. Same as Figure 10, but for the bias between monthly WAM model and ERS-1 SWM significant wave heights (solid line), wind-sea wave heights (dotted line), and swell wave heights (dashed line).

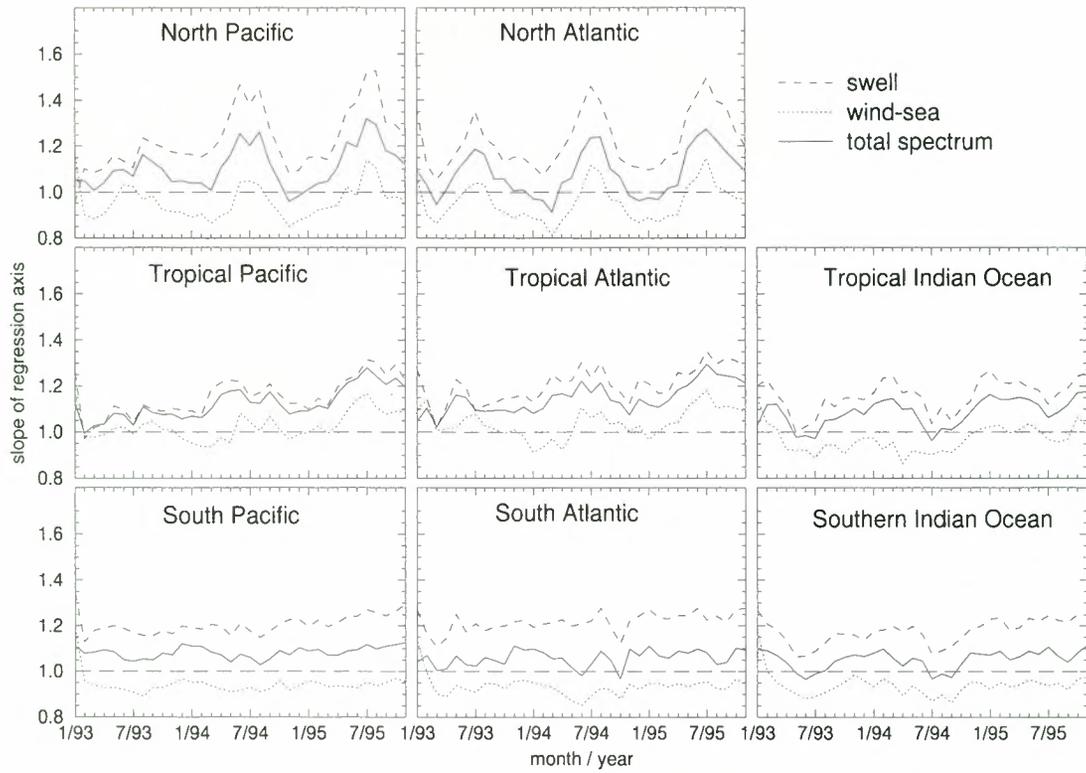


Figure 12. Same as Figure 10, but for the slope of the monthly regression between WAM model and ERS-1 SWM significant wave heights (solid line), wind-sea wave heights (dotted line), and swell wave heights (dashed line).

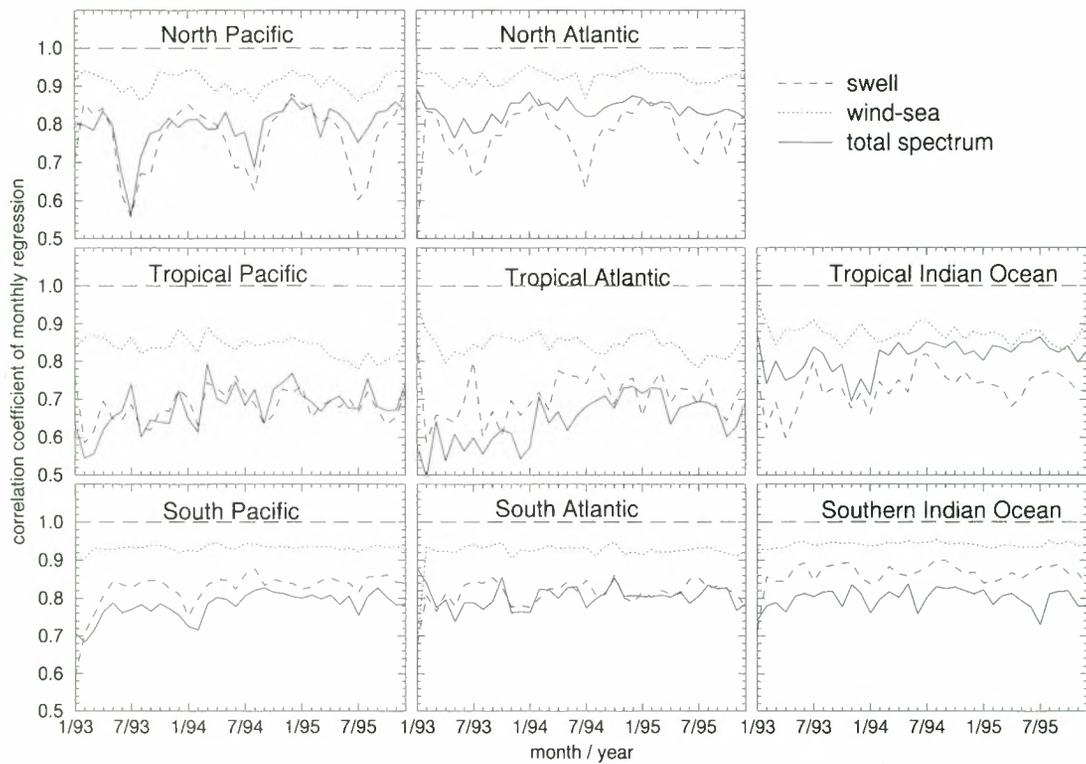


Figure 13. Same as Figure 10, but for the correlation of the monthly regression between WAM model and ERS-1 SWM significant wave heights (solid line), wind-sea wave heights (dotted line), and swell wave heights (dashed line).

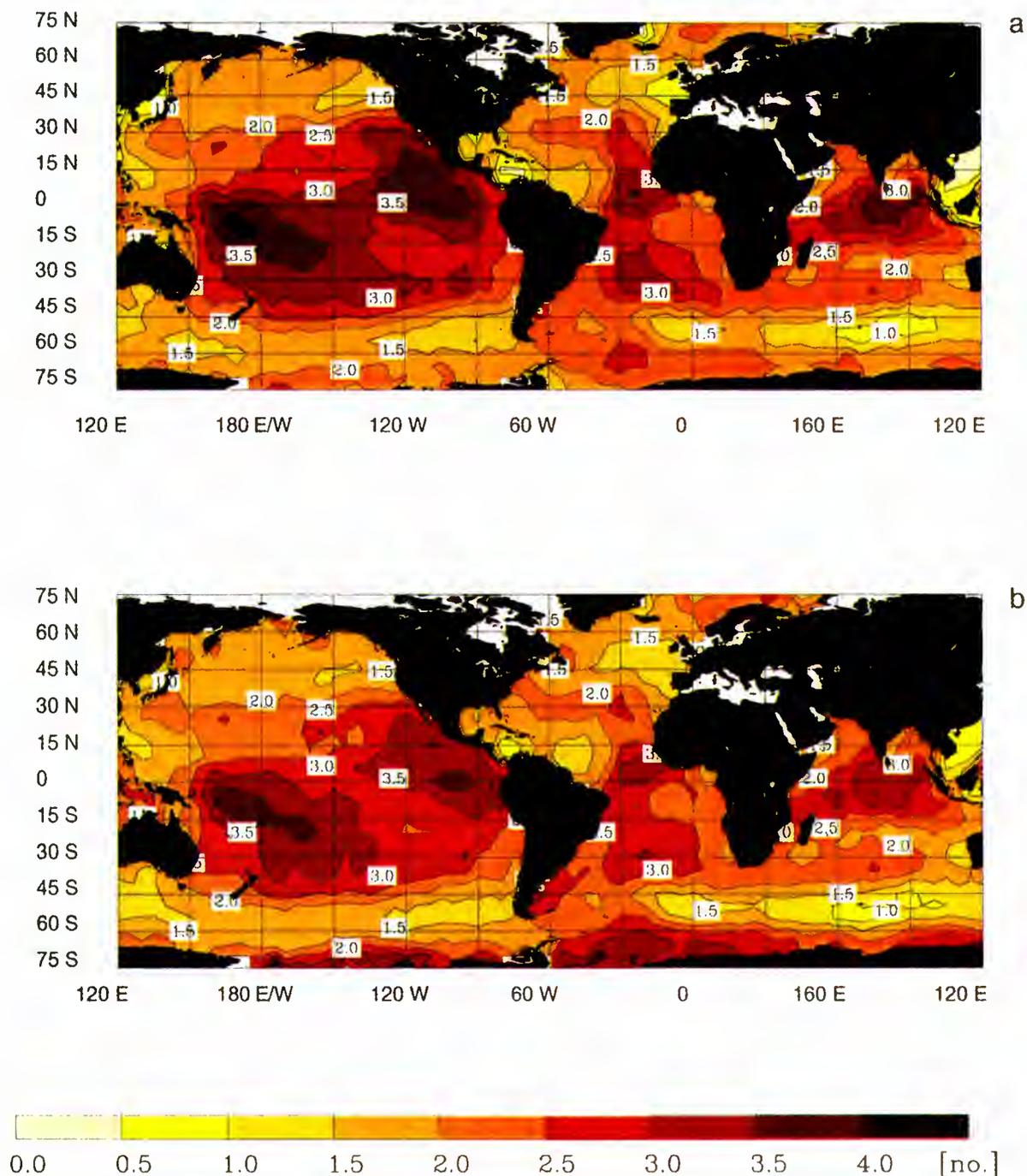


Plate 2. Global map of the average number of swell partitionings per spectrum for DJF 1993–1994. (a) WAM model; (b) ERS-1 SWM.

sphere winter) are found in the vicinity of the Aleutian low in the North Pacific and the Icelandic low in the North Atlantic. A belt of moderate H_{wv} is found in the southern hemisphere midlatitudes, driven by the summer westerly wind belt around the Antarctic.

In the southern hemisphere winter JJA the highest wind seas are found in the midlatitude westerly wind belt in the southern hemisphere. Strong monsoon-driven wind-sea systems propagating to the NE and NW are found also in the Arabian Sea. In the tropics the dominant wave propagation direction is determined by the easterly trade winds, which are generally stronger in the winter seasons.

An alternative view of the global sea state structure is pro-

vided by the latitudinal distribution of the zonally averaged wave heights and directions for wind sea and swell (Figures 18a and 18b, respectively). The following features can be recognized: (1) low wave heights in the seasonally shifting equatorial Intertropical Convergence Zone (ITCZ) due to weak surface easterlies; (2) northwestward (in the southern hemisphere) to southwestward (in the northern hemisphere) propagating waves of moderate wave height in the trade wind regimes on either side of the ITCZ; (3) low wave heights in the region of the subtropical semipermanent high-pressure cells around 30°N to 30°S (the horse latitudes); (4) strong wind-sea systems in the midlatitude west wind belts, particularly in the roaring forties, fighting fifties, and screaming sixties in the winter

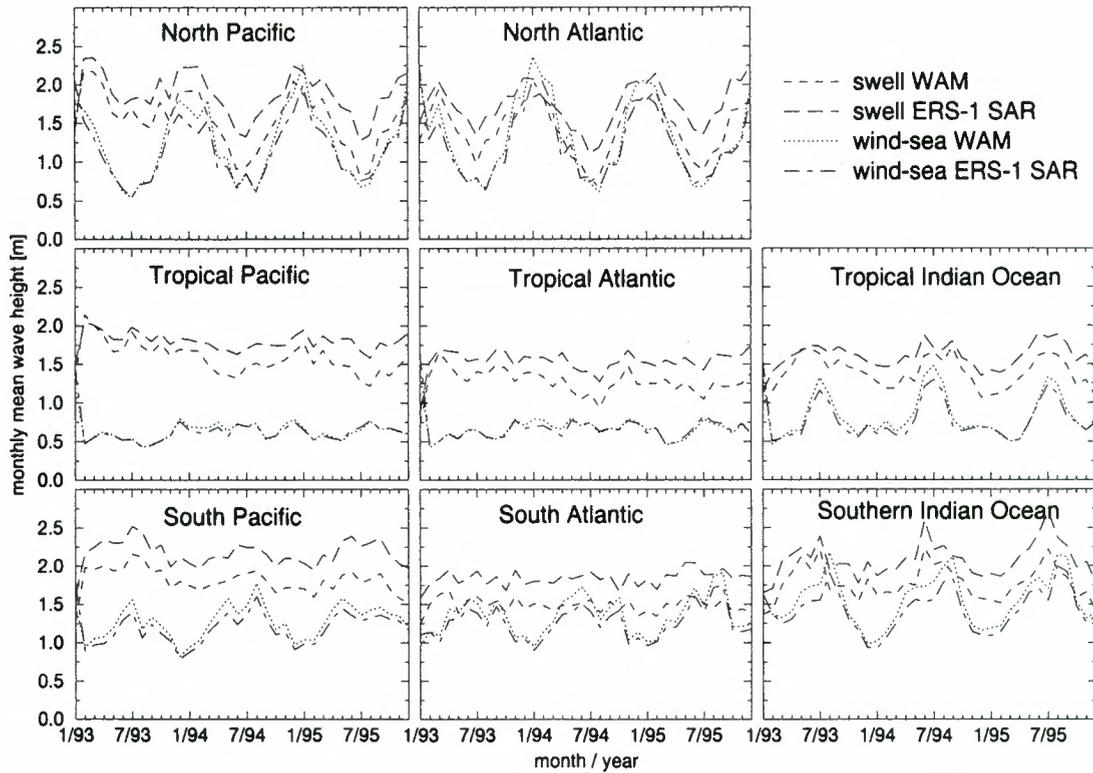


Figure 14. Same as Figure 10, but for the monthly mean wind-sea and swell wave heights.

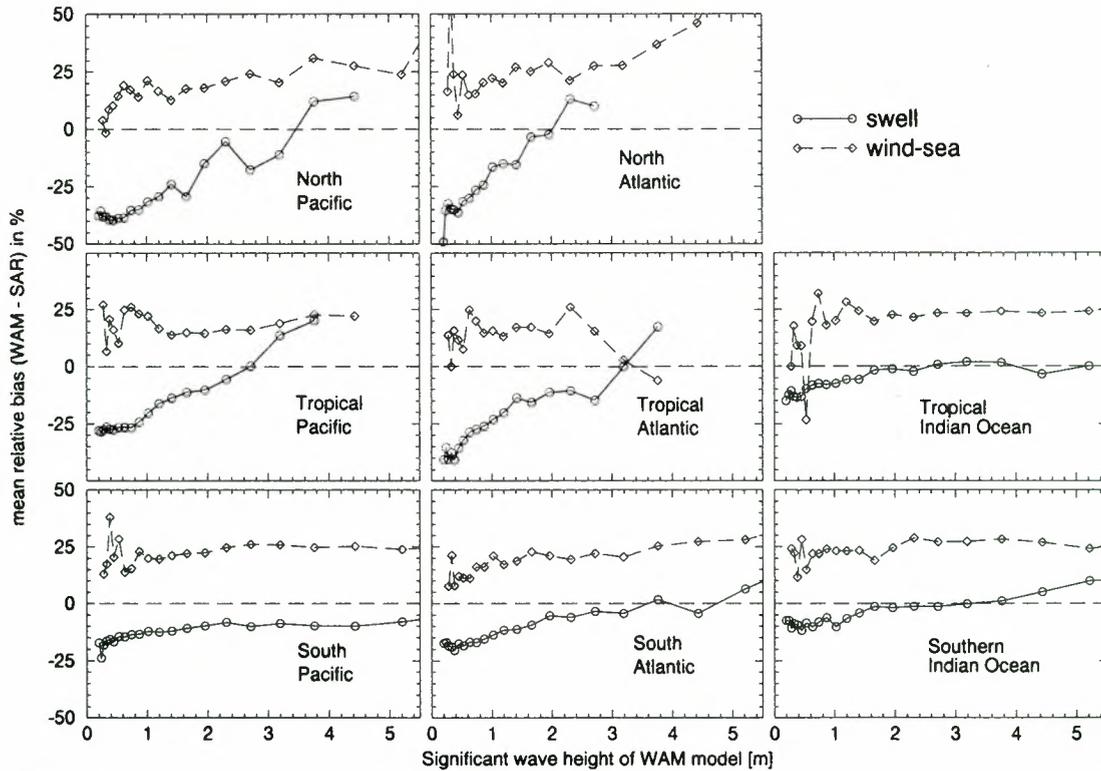


Figure 15. Mean relative bias between WAM and ERS-1 SWM as a function of the WAM model wave height for June, July, and August (JJA) 1994. Solid lines refer to swell, and dashed lines refer to wind sea.

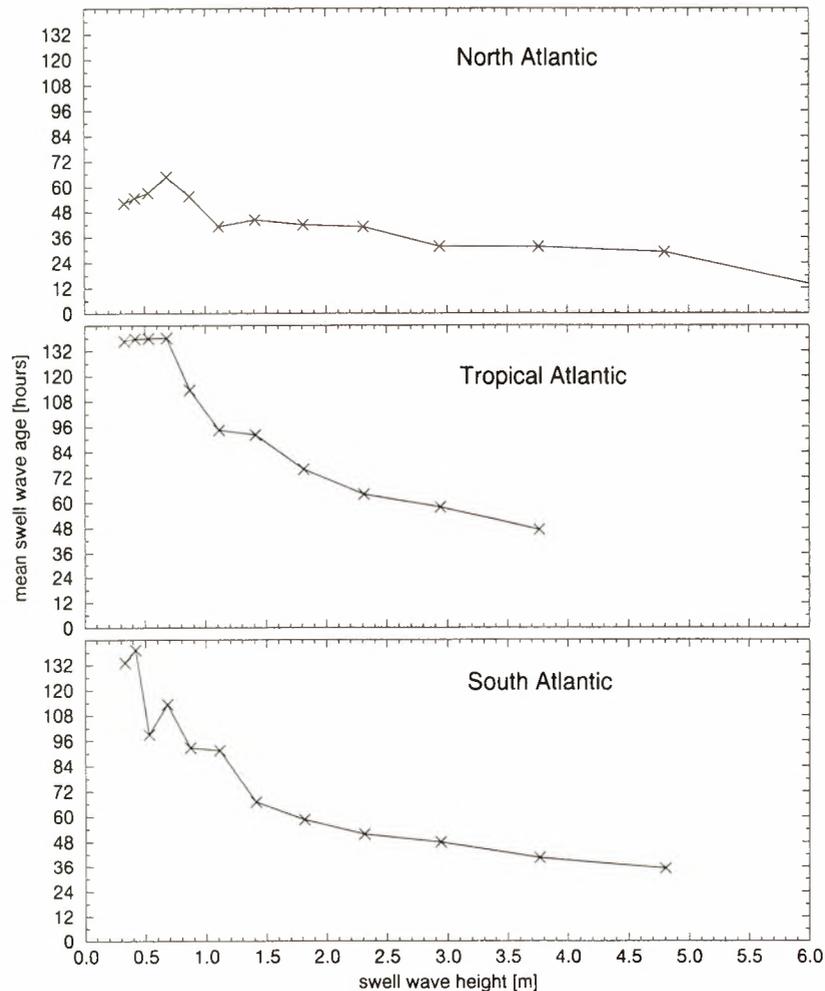


Figure 16. Average wave age as a function of the wave height of individual swell partitionings for a data set in the Atlantic for November 1992.

southern hemisphere; the stronger seasonal cycle in the northern hemisphere noted already in the H_{wi} time series is also seen; (5) the influence of the easterly winds from the polar high-pressure cells south of 65°S in the southern hemisphere; and (6) pronounced seasonal differences in zonally averaged swell wave height around 25°N in the Indian Ocean, caused by the seasonal monsoon cycle, and below 60°S near the Antarctic, presumably caused by changes in sea ice extent.

The seasonal maps and zonal averages of wave heights and directional distributions confirm the previous findings of a generally overestimated wind sea by the WAM model relative to the SWM retrievals. The effect is strongest in high sea states occurring in midlatitudes during winter. However, the modeled wind sea is slightly underestimated in the trade wind regions during JJA and SON.

The modeled and retrieved directional distributions in both Figures 17a and 17b and Figures 18a and 18b are remarkably similar. However, WAM shows a rather more symmetrical distribution of the meridional component about the west-east axis in the southern hemisphere extratropics, with a maximum directed to the east (90°), whereas the SWM retrieved data are more asymmetric with maxima directed toward ENE (60°). Discrepancies can also be seen in the directional distribution of east to northward traveling waves in the tropical Indian Ocean during the summer monsoon.

4.3.2. Swell wave heights. The seasonal maps of swell wave heights and directional distributions are less sensitive to local atmospheric conditions, but the principal features seen in Figures 18a–18b can again be largely attributed to the synoptic conditions that produced the wind sea that is later transformed into swell. Thus the predominance of eastward and equatorward propagating swell in the midlatitudes can be explained by the strong westerly cyclones in that belt. The significant meridional swell components in the tropics, characterized by a marked seasonal variation in the directional distribution, presumably also originate in the midlatitude cyclones. In addition, the tropical distributions exhibit a significant westward component due to the trade winds.

The previously noted underestimation of swell by the WAM model is again apparent, both from the directional distributions of Figure 17b and the zonal averages of Figures 18a–18b. In the extratropics the strongest bias is seen for eastward propagating swell; in the tropics it is seen for westward propagation. A more detailed analysis of the conditions for the South Pacific Ocean is given in section 4.4.

4.3.3. Mean wavelengths. In addition to the significant wave height and mean propagation direction, the third parameter normally used to characterize a wave spectrum or, in the case of a partitioned spectrum, a wave system is the energy-weighted mean frequency \bar{f} . Equivalently, we present in the

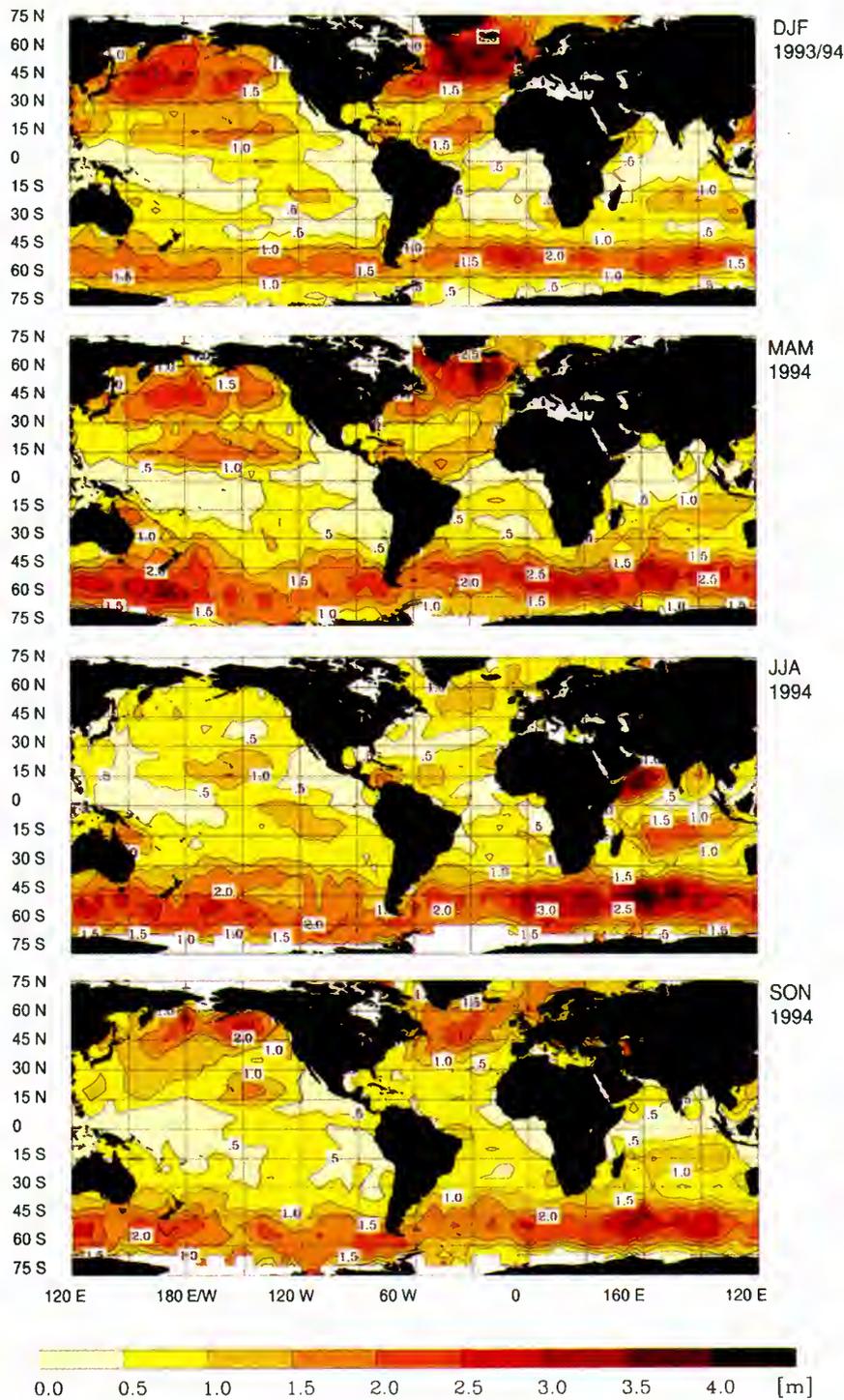


Plate 3a. Global maps of the seasonal mean wind-sea wave height for the four seasons (top two panels) DJF 1993–1994 to (bottom two panels) SON 1994 for the WAM model.

following the mean wavelength $\bar{\lambda} = g/2\pi f^2$. Plates 5 and 6 show the distribution of seasonal mean wavelengths for wind sea and swell, respectively, for the modeled and SWM-retrieved spectra for the winter season DJF 1993–1994 (upper two panels) and summer season JJA 1994 (lower two panels).

The retrieved swell wavelengths agree rather closely with the modeled wavelengths but are generally slightly larger. This is consistent with the higher retrieved swell wave heights discussed in section 4.2. Too much significance should not be

attached to the retrieved wind-sea wavelengths. Because of the azimuthal cutoff the wind-sea part of the wave spectrum is normally only partially resolved by the SAR (compare section 5), the full wind-sea spectrum being reconstructed from the SAR data by fitting the azimuthal cutoff parameter and rescaling the parameters of the wind-sea wave system provided by the WAM model.

The effect of fetch can be clearly seen in Plates 5 and 6. The shortest wavelengths are found in the lee of the continents, in

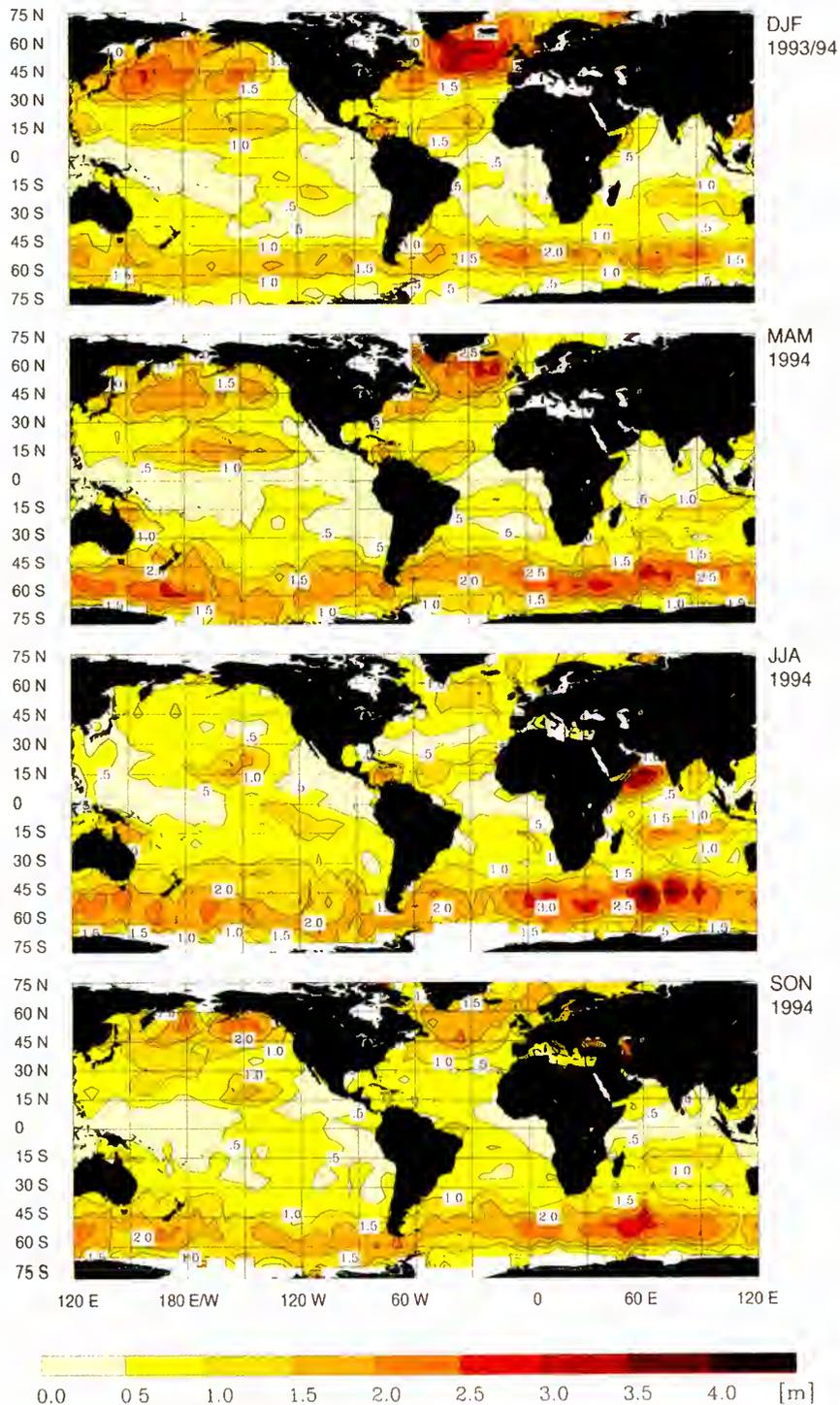


Plate 3b. Same as Plate 3a, but for the ERS-1 SWM.

midlatitudes off the east coasts, and the longest wavelengths are found in the eastern parts of the ocean basins and in the open Southern Ocean. The increase in wavelength with the higher wind speeds in winter is also clearly visible, particularly in the northern hemisphere, which has a more pronounced seasonal cycle (see section 4.1).

An interesting feature is the occurrence of long wavelength swell originating in the high-wind midlatitude regions in the eastern equatorial oceans; the swell wavelength isolines exhibit a distinct eastward-equatorward slant, corresponding to the main

direction of propagation from the extratropics, in contrast to the predominantly east-west oriented isolines of the wind-sea wavelengths, corresponding to the mainly zonally directed winds.

The local impact of wind speed is seen most clearly, as expected, in the wind-sea wavelengths, which correlate well with the wind-sea wave heights, compare Plates 3a, 3b, and 5. However, for wind sea both the wind speed and the fetch are important, as demonstrated by the pronounced impact of both coastal shadowing and the high-wind regions on the wind-sea wavelength distributions.

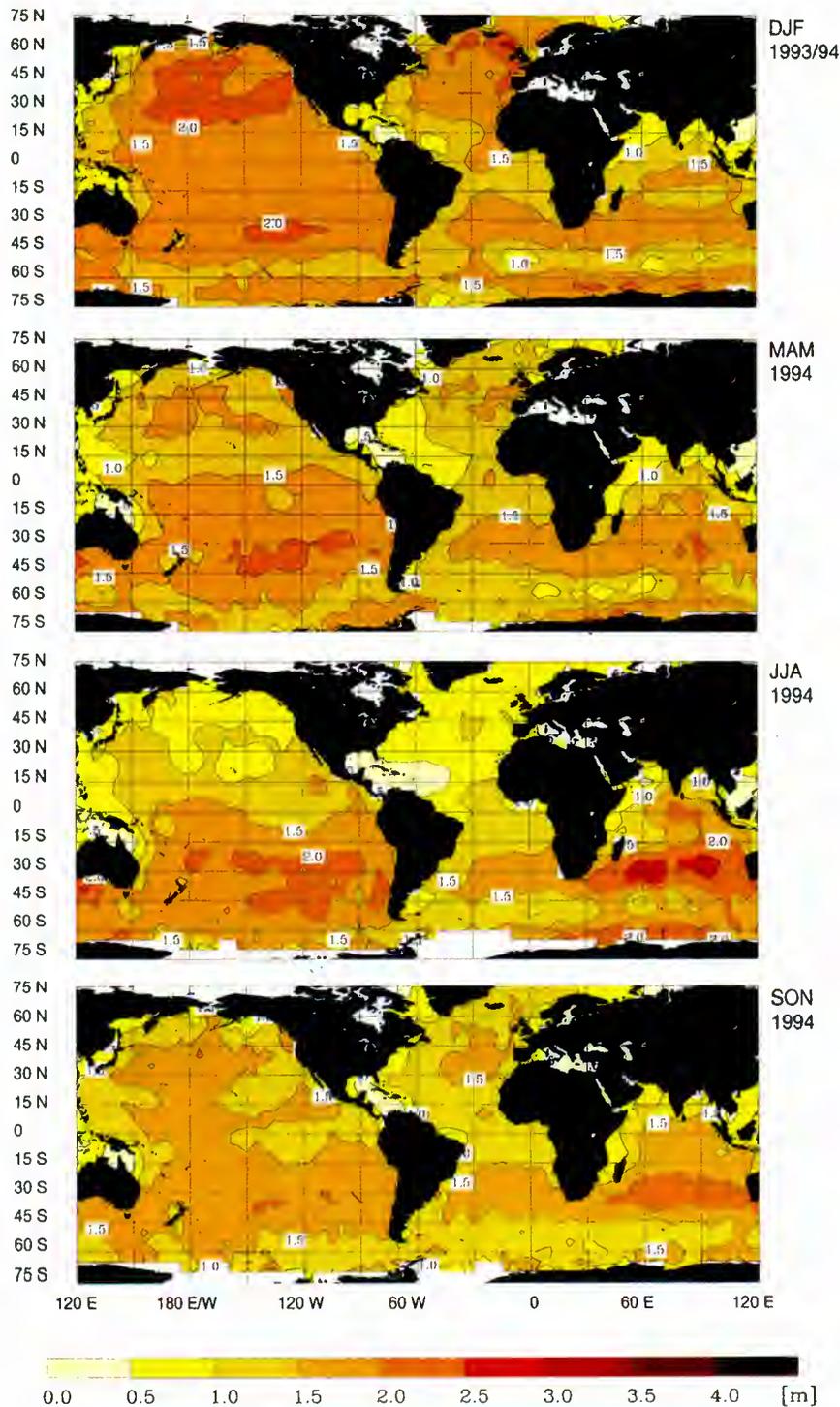


Plate 4a. Same as Plate 3a, but for the seasonal mean swell wave height.

4.4. Regional Analysis for the South Pacific Ocean

As an example illustrating the detailed spectral and regional information contained in the ERS-1 SWM data, we present in Figure 19 the retrieved and modeled wave roses for wind sea (Figures 19a and 19b, for DJF 1993–1994 and JJA 1994, respectively) and swell (Figures 19c and 19d, for DJF 1993–1994 and JJA 1994, respectively) for nine regions in the South Pacific.

4.4.1. Wind sea. The general zonal wind characteristics are reproduced by both the modeled and retrieved data, which

show also significant deviations from the mean wind directions that appear as side lobes in the directional distributions. In the sample areas between 0° and 30°S the wind sea is dominated by west to northwest traveling waves generated by the trade winds. The WAM wind sea tends to be turned more equatorward than the retrieved wind sea, suggesting a possible shortcoming in the angular distribution of the ECMWF trade wind field.

Between 30° and 60°S the dominant wind-sea direction is eastward, reflecting the midlatitude westerlies. However, other

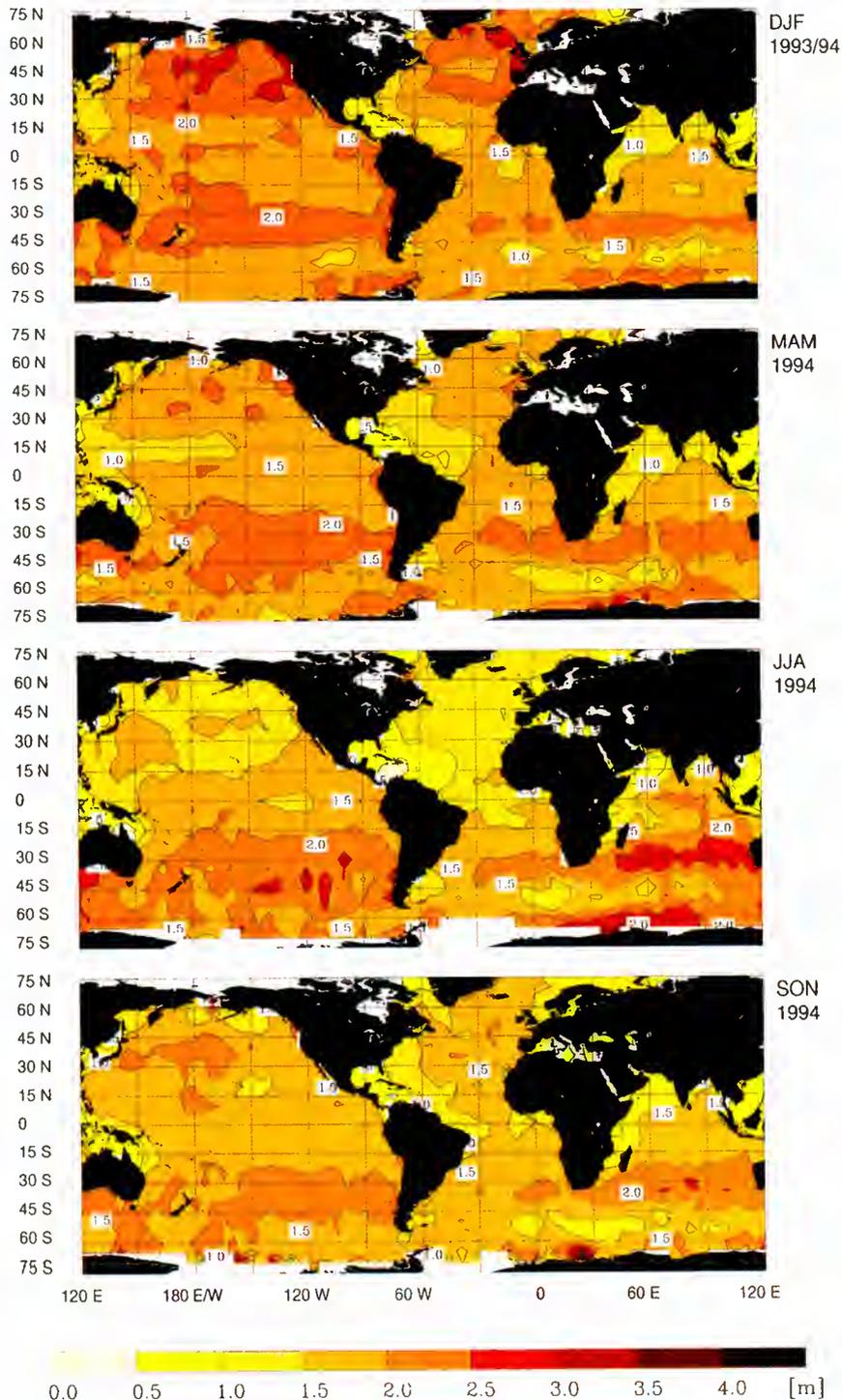


Plate 4b. Same as Plate 3b, but for the seasonal mean swell wave height.

directions also occur with significant probability, for example, toward the NW, as a result of the cyclonic disturbances characteristic of these latitudes. The WAM model overestimates mainly the east and southeast traveling components, suggesting again that the model errors must be attributed at least in part to wind field errors.

In the band south of 60°S the wind-sea distribution is similarly dominated by eastward traveling waves, and the WAM model again tends to overestimate waves traveling east to

southeast. The shadowing effect of Antarctica is clearly evident in the most westerly segment.

4.4.2. Swell. The directional distributions of swell in the tropical latitudinal band between 0° and 30°S are determined by two principal sources: the easterly trades on either side of the ITCZ, producing a broad westward propagating swell sector, and the midlatitude westerlies. The dependence on fetch leads to an increase of the trade wind swell from east to west, while the southeast traveling waves from the midlatitude west-

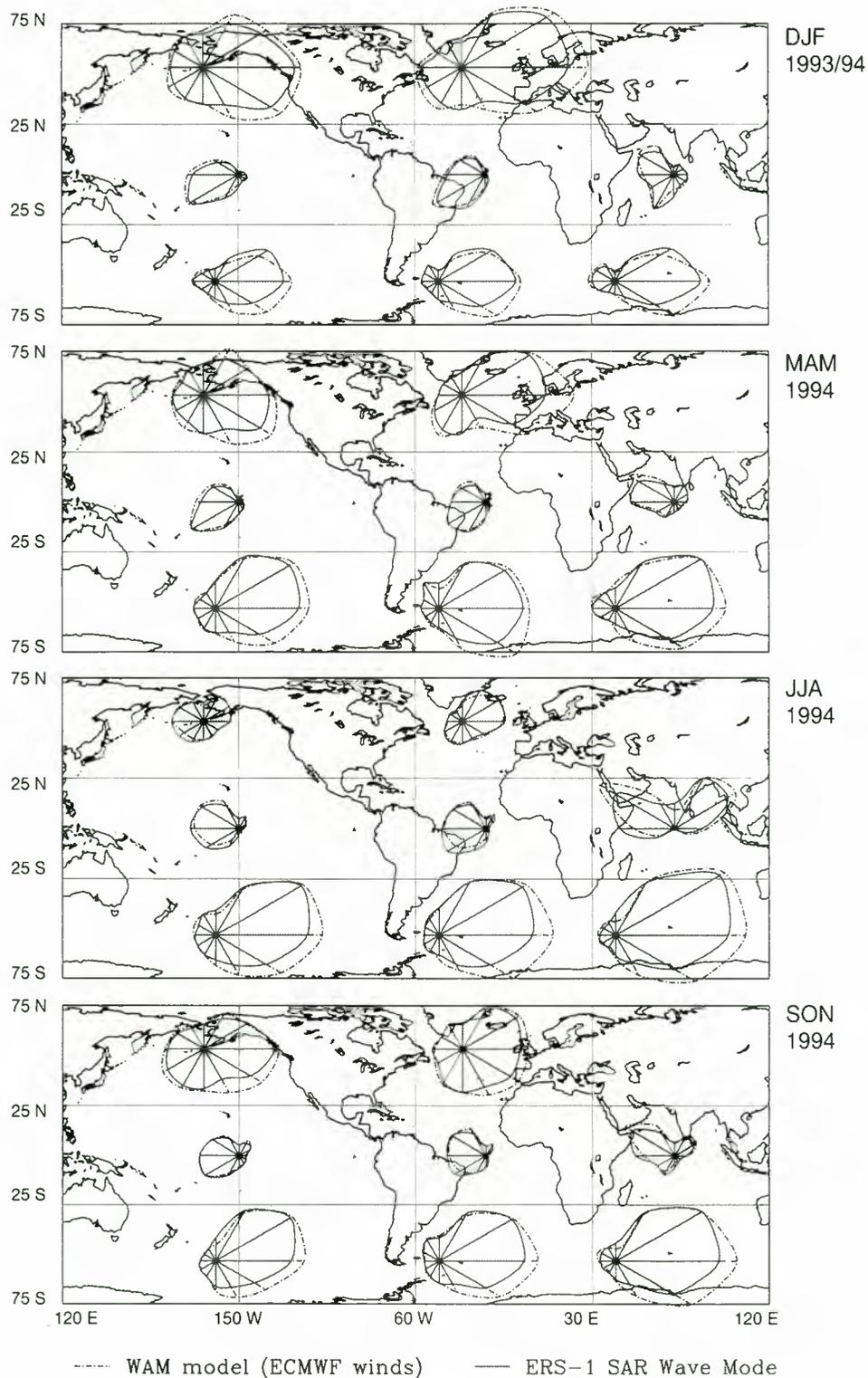


Figure 17a. Seasonal mean directional wave height distribution in different ocean basins for the four seasons (top) DJF 1993–1994 to (bottom) SON 1994. Dashed-dotted lines, modeled; solid lines, retrieved from ERS-1 SWM. For wind sea.

erlies to the south increase from west to east (see also the global distributions of wavelengths for JJA 1994, Plate 6). The WAM model tends to underestimate particularly the trade wind swell components.

In the remaining two latitudinal bands south of 30°S the swell is dominated by the midlatitude westerlies. The distribu-

tions become more isotropic, reflecting the cyclonic character of the generating storms. With a few exceptions, the WAM model underestimates the swell in all directions.

Particularly noticeable is the underestimation of the east and northeast traveling swell south of 60°S. This presumably again points to errors in the wind field. The area is bounded from the

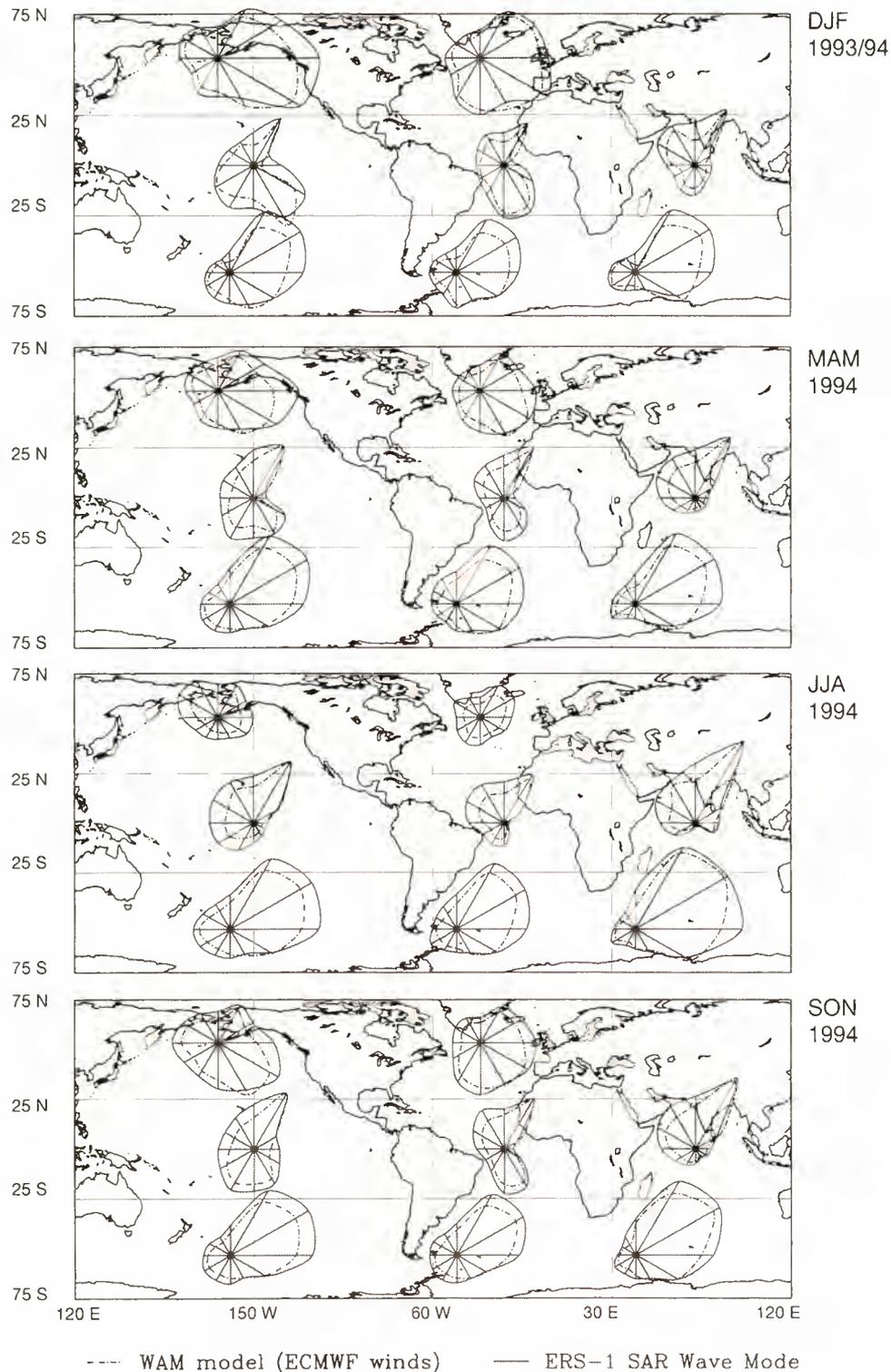


Figure 17b. Same as Figure 17a, but for swell.

south by Antarctic sea ice and thus has rather limited fetch to produce swell with a significant northward component. The existence of such components therefore implies rather strong northward or northeastward blowing winds. Katabatic winds could be a candidate, particularly as they are notoriously difficult to capture adequately in global circulation models. However, the predominantly southeasterly component of katabatic winds is not consistent with the easterly propagation direction

of the swell. We have no ready explanation for this discrepancy.

4.5. North-South Distributions Through the Pacific

As a further example of the detailed spatial-spectral information provided by the ERS-1 SWM, we have investigated the distribution of wave parameters in five areas along a meridional section of the Pacific for the boreal winter DJF 1993–1994

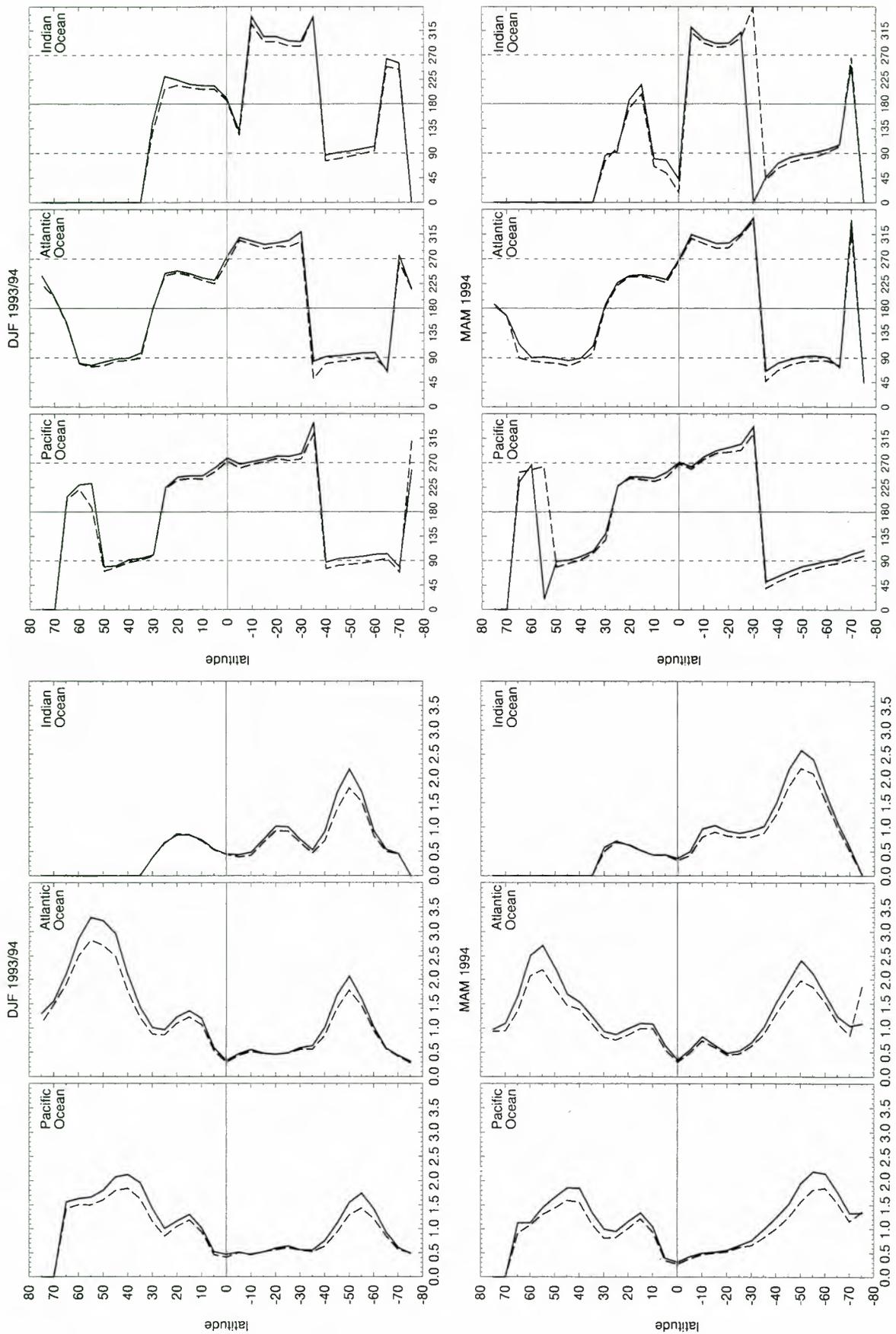


Figure 18a. Seasonal zonal (left) mean wave heights and (right) mean wave heights and (right) directions in the Pacific, Atlantic, and Indian Ocean for the four seasons from DJF 1993–1994 to SON 1994. Solid lines, modeled; dashed lines, retrieved from ERS-1 SWM. For wind sea.

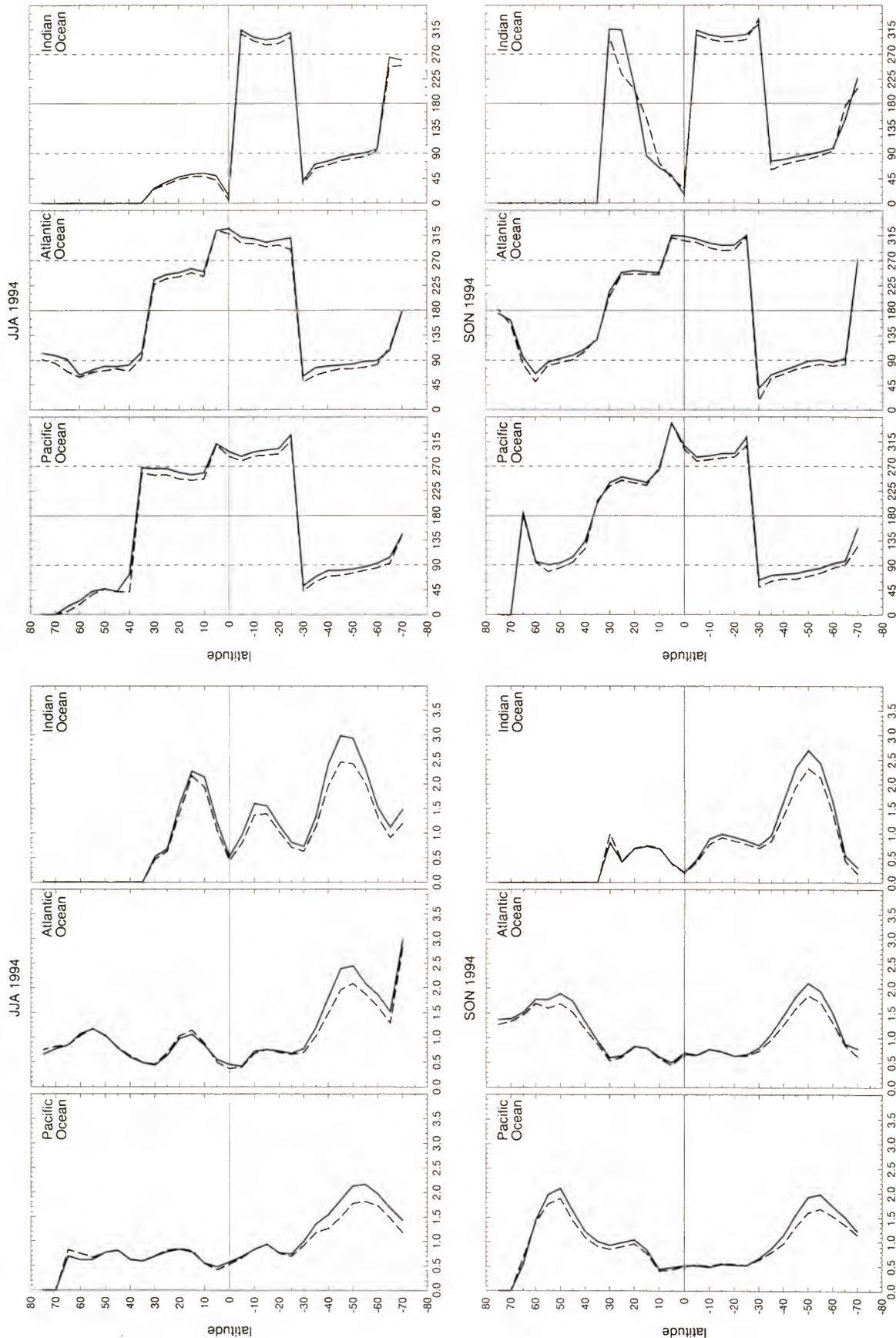


Figure 18a. (continued)

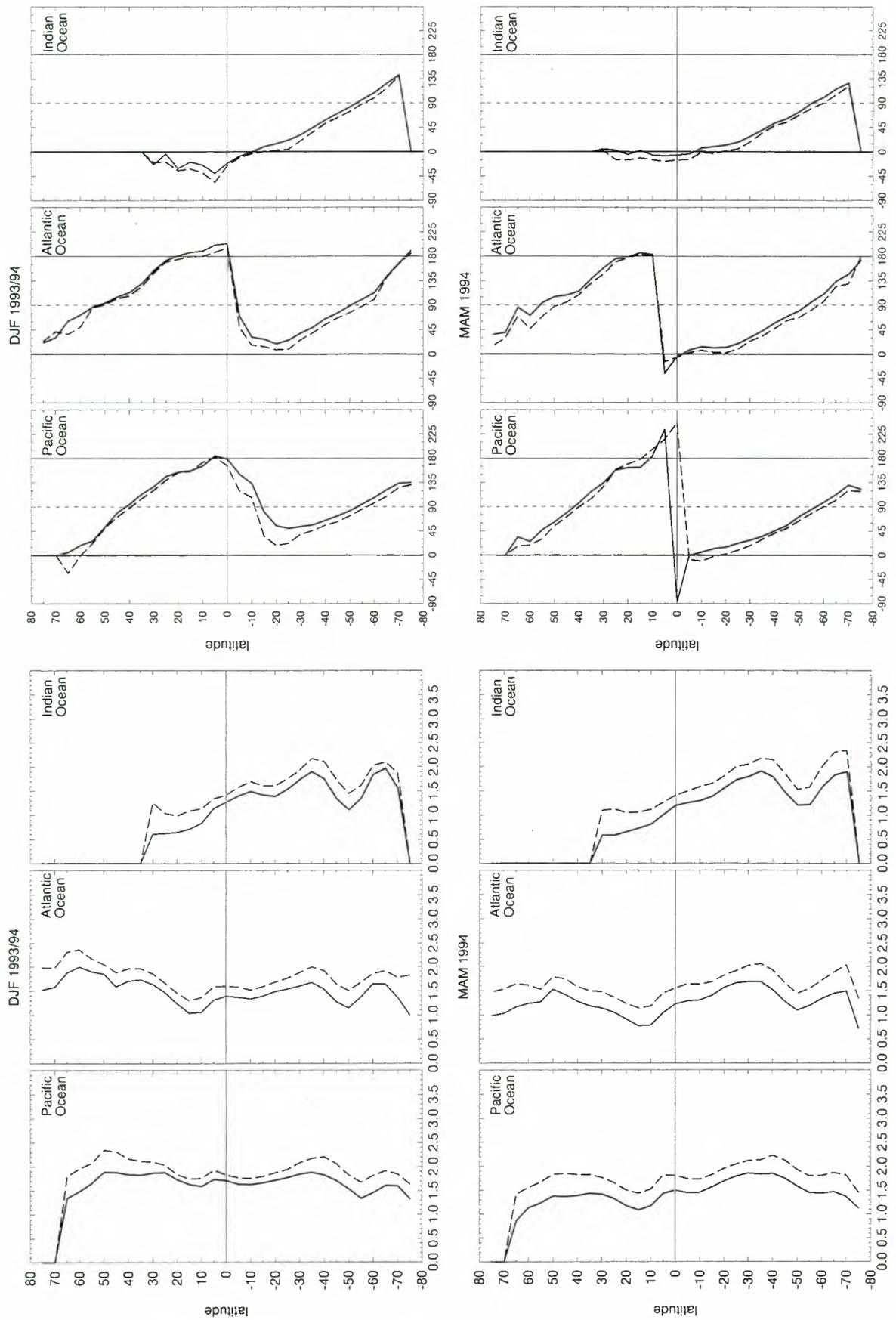


Figure 18b. Same as Figure 18a, but for swell.

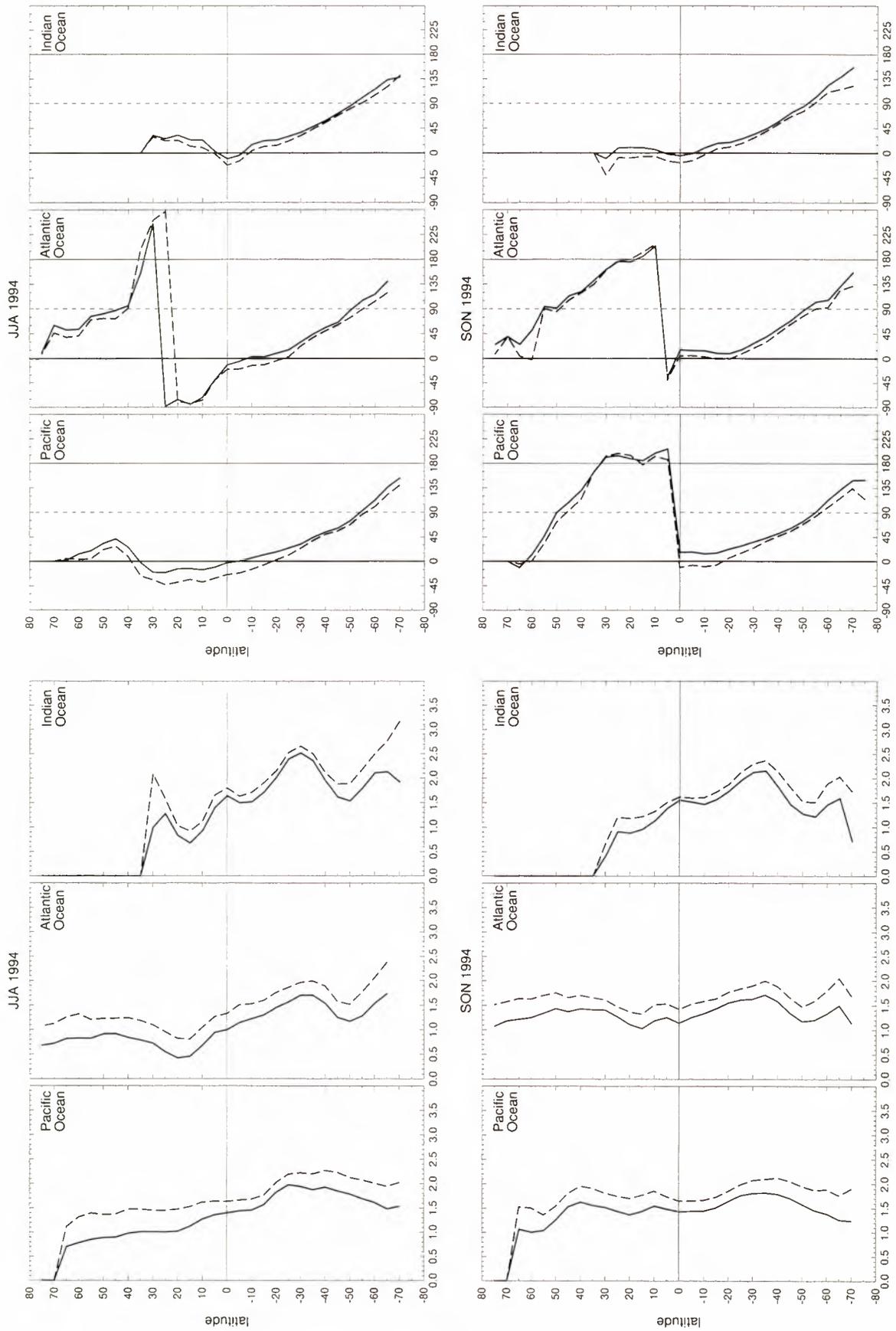


Figure 18b. (continued)

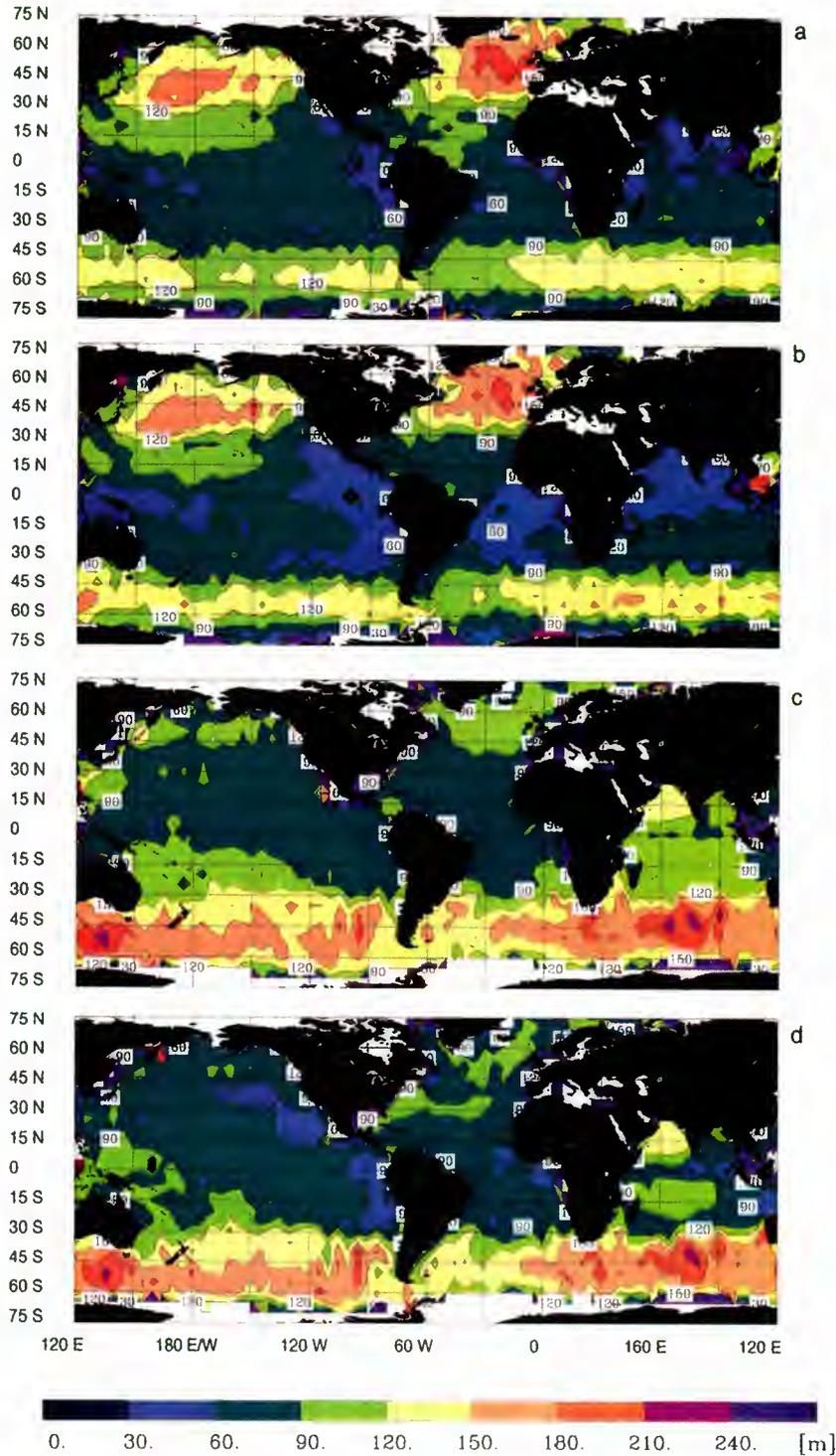


Plate 5. Same as Plates 3a and 3b for the seasonal mean wind-sea wavelength. (a) DJF 1993–1994 WAM, (b) DJF 1993–1994 ERS-1 SWM, (c) JJA 1994 WAM, and (d) JJA 1994 ERS-1 SWM.

and summer JJA 1994 (Figures 20a and 20b). The areas are representative of the northern and southern tropics and mid-latitudes and the Ross Sea in the Antarctic Ocean. Figures 20a and 20b depict the polar wave rose plots of the directional distributions for swell, while Figure 21 shows the relative distributions of wind sea, swell, and mixed wind-sea–swell wave heights.

The directional distributions exhibit considerable structure, with a pronounced seasonal dependence. They can again be

explained in terms of the various wind systems characterizing the different latitudinal zones of the Pacific.

As expected, the distribution curves of retrieved swell wave heights (dashed curves) are shifted to higher values than the corresponding model distributions, while the opposite holds, although less pronounced, for the wind-sea curves. A noticeable feature of the distributions (apparent also in the spatial distributions shown earlier) is that the mean significant wave heights are considerably higher for swell than for wind sea.

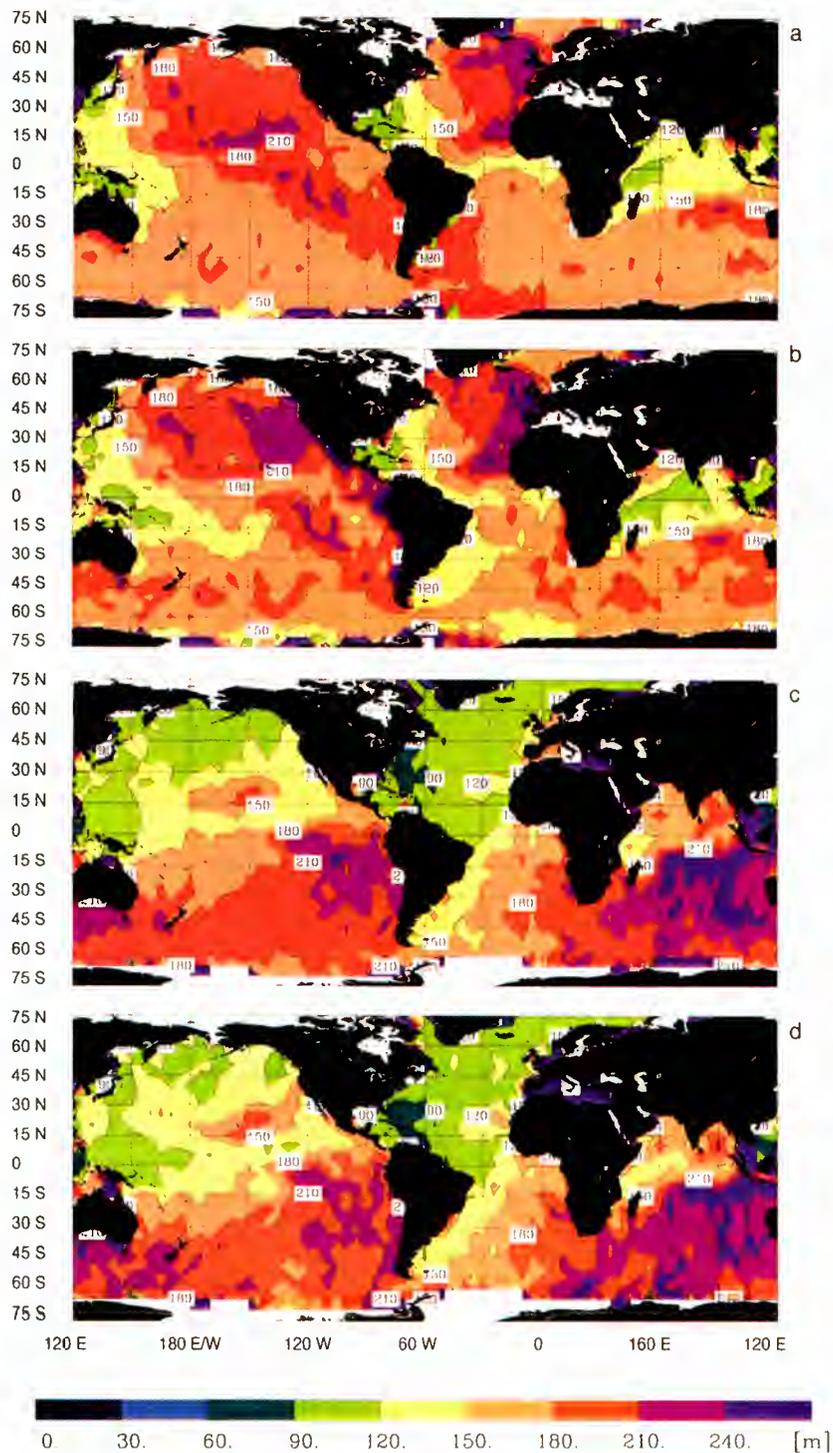


Plate 6. Same as Plate 5, but for swell.

This is because a given spectral wave component, although attaining its maximum energy at the end of its relatively short wind-sea growth period, subsequently propagates through the ocean for a significantly longer period as swell.

We note that while the modeled and retrieved directional distributions for swell are generally consistent with the hypothesis that the WAM model underestimates swell because of a too large swell dissipation, individual features of the polar distributions do not appear to support this interpretation. In particular, the NNE and SSE traveling swell components in

northern hemisphere and southern hemisphere tropics are overestimated by the WAM model. In the next section we analyze this phenomenon in more detail in the spectral domain.

4.6. Stratification With Respect to Swell System Wavelength

For a clearer insight into the nature of the differences between the modeled and retrieved wave spectra we need to stratify the partitioned wave systems with respect to wave-

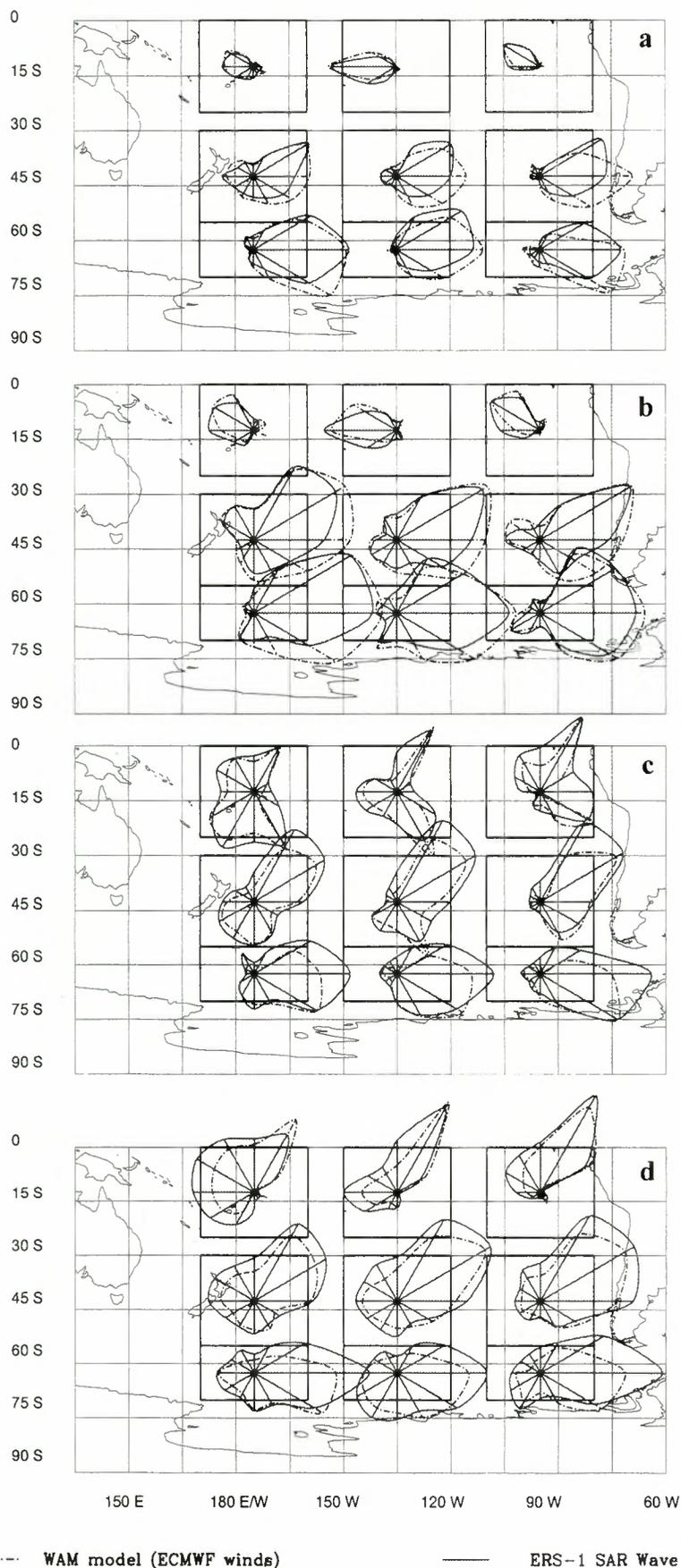


Figure 19. Seasonal directional wave height distribution of the wind sea and swell in selected areas of the South Pacific. (a) DJF 1993-1994 wind sea, (b) JJA 1994 wind sea, (c) DJF 1993-1994 swell, and (d) JJA 1994 swell. Dashed-dotted lines, modeled; solid lines, retrieved.

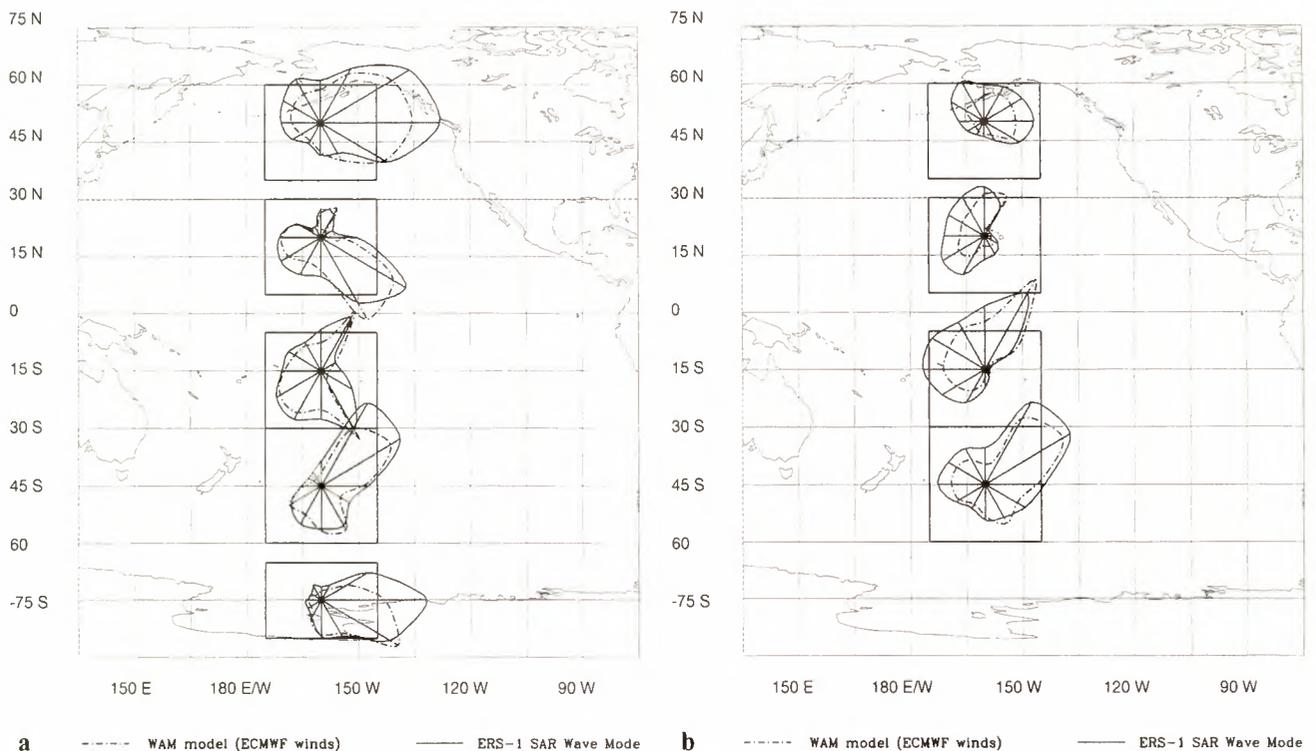


Figure 20. Seasonal directional wave height distribution of the swell in five meridionally distinct boxes of the Pacific Ocean zonally bounded between 170°E and 150°W for (a) DJF 1993–1994 and (b) JJA 1994. Dashed-dotted lines refer to WAM model values, and solid lines refer to ERS-1 SWM values.

length as well as direction. Plate 7 shows the global distributions of mean wave height for the same winter and summer seasons DJF 1993–1994 (Plates 7a and 7b for WAM and ERS-1 SWM, respectively) and JJA 1994 (Plates 7c and 7d for WAM and ERS-1 SWM, respectively) considered previously, but now only for the longest wavelength swell system in each spectrum. To be included in the statistic, however, the swell wavelength was required to exceed a lower threshold, set in Plate 7 at 250 m. Consistent with our general findings for the total swell, the SWM-derived wave heights for these swell systems are again higher than the WAM wave heights. However, in contrast to the maps of the total swell wave height (defined as the rms value of the energy summed over all partitioned swell systems), the distributions for individual long wavelength wave systems are seen to be very similar to the maps of mean swell wavelengths. One can thereby more easily identify the regions into which the long wavelength swell is radiated from the midlatitude source regions.

Figures 22a and 22b show the associated directional swell wave height distributions in different sampling areas in the South Pacific for different lower long wavelength thresholds. With increasing wavelength threshold, the wave heights diminish, as expected. The model wave heights generally decrease faster than the SWM-retrieved wave heights, as again expected if the underestimation of swell by the model is due to a too strong model dissipation acting selectively on the longer swell components. However, the discrepancies in the details of the directional distribution increase with the swell wavelength, and not all of these can be explained simply by an incorrect dissipation expression. For example, the WAM model predicts significant SE traveling swell components in the southern hemisphere in the northern winter, presumably radiating from

northern midlatitude storms, which are not seen in the retrieved swell.

The processes involved in the transformation of wind sea into swell and the subsequent propagation of swell over large distances in the ocean are clearly complex and cannot be adequately analyzed within the framework of a general statistical study. A more appropriate approach for a process analysis would be a detailed investigation of a set of individual events. The ERS SWM measurements provide a valuable comprehensive database for such investigations, but these lie outside the scope of the present exploratory study.

5. Analysis of the rms Orbital Velocity

5.1. General Considerations

As already pointed out, a basic limitation of SAR ocean wave images is the azimuthal cutoff of the image spectrum caused by the nonlinear velocity bunching mechanism (HH; HBHH). The cutoff normally lies within the wind-sea part of the spectrum and suppresses the direct spectral information provided by the SAR on short to moderate wavelength waves with significant components of propagation in the azimuthal direction. However, the azimuthal cutoff wavenumber also provides a direct measure of an important integral parameter of the wave field, the rms orbital velocity component in the range direction. This information is used in the retrieval algorithm to adjust the rms orbital velocity to reproduce the observed cutoff, thereby partially replacing the missing direct spectral information at high wavenumbers, as discussed in section 2.2.2. However, the cutoff wavenumber, as a measure of the orbital velocity, is also useful data in its own right.

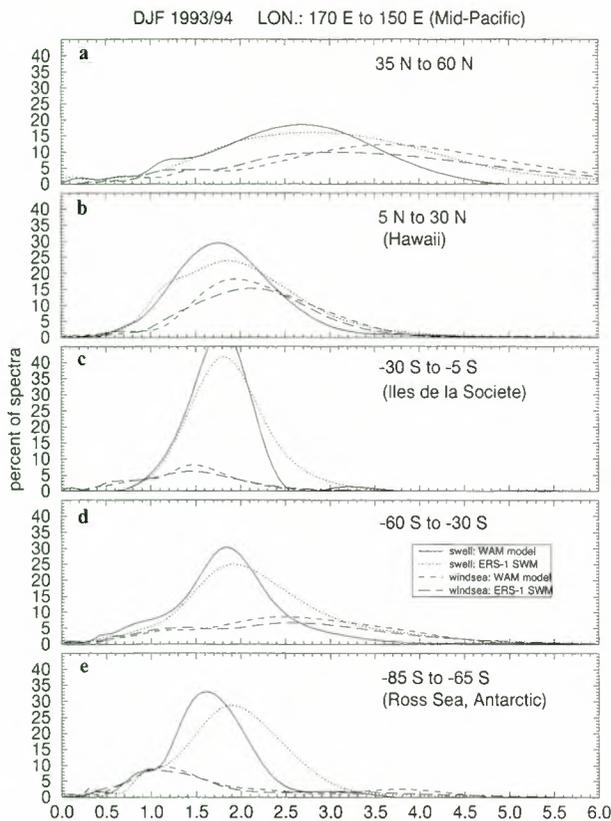


Figure 21. Distribution of seasonal wind sea, swell, and mixed wind-sea-swell wave heights for the five meridionally distinct mid-Pacific boxes of Figure 20 for DJF 1993–1994. Solid lines refer to WAM model values, and dashed lines refer to ERS-1 SWM values. Figure 21b corresponds to an area around Hawaii, Figure 21c corresponds to an area around the Îles de la Société, and Figure 21e corresponds to an area of the Ross Sea in the Antarctic.

5.2. Comparison of Observed Versus Simulated Cutoff

Figure 23 shows scatter diagrams and regression lines for the observed and simulated azimuthal cutoff wavelengths for the boreal winter DJF 1993–1994 in the tropics and the northern and southern hemisphere extratropics. The observed azimuthal cutoff wavelengths were determined directly from the SAR image spectra, while the simulated cutoff wavelengths were inferred, using the same algorithm, from the SAR image spectra computed from the retrieved wave spectra. Regression line slopes close to 1 and correlation coefficients of 0.9 and higher in all three data sets indicate that the adjustment of the cutoff parameter in the retrieval algorithm was successful.

5.3. Spectral Resolution Imposed by the Cutoff

To estimate the impact of the azimuthal cutoff on the retrieval of wind-sea and swell systems, we have computed the ratio k_c/k_x of the observed azimuthal cutoff wavenumber to the mean azimuthal wavenumber component for an ensemble of wind-sea and swell systems. Figure 24 shows the resulting histograms for three wind-sea (Figure 24a) and swell (Figure 24b) data sets for the period DJF 1995–1996 for the northern hemisphere extratropics, the tropics between 25°N and 25°S, and the southern hemisphere extratropics. The medians of the histograms are listed in Table 4, together with the percentage of wave systems with ratios $k_c/k_x > 1$, representing waves that

can be explicitly imaged by the SAR. The values clearly demonstrate the ability of the ERS-1 SAR to detect a substantial fraction of the swell systems encountered in the ocean, despite the high R/v ratio. A fair fraction of 30% of the wind-sea systems can also be directly imaged by the SAR, but short wavelength wind seas generated by light winds cannot be seen at all by the SAR.

5.4. Global Analysis of the Cutoff

A global map of the mean azimuthal cutoff wavelength λ_c for DJF 1995–1996 is shown in Plate 8. High cutoff wavelengths corresponding to high rms orbital velocities are found, as expected, in the regions of high winds in midlatitudes. In addition, a pronounced fetch effect is seen. Young wind seas have higher energies at shorter wavelengths and therefore higher rms orbital velocities than well-developed wind seas [cf. *Komen et al.*, 1994]. Thus the cutoff wavelength is significantly higher in the western parts of the midlatitude ocean basins and in enclosed or partially enclosed seas such as the Baltic Sea, Mediterranean Sea, Gulf of Mexico, Caribbean Sea, Bering Sea, or the seas south of New Guinea.

In view of the strong dependence of the orbital velocity on fetch a general expression relating λ_c alone to the wind velocity u_{10} at 10-m height, for example, in the form $\lambda_c = 25u_{10}$ [*Chapron et al.*, 1995], cannot be expected to be generally valid. Although dimensionally consistent and therefore presumably appropriate for fully developed seas, the expression yielded a low correlation of 0.3–0.4 between the observed and predicted cutoff wavelengths for the global data set of DJF 1995–1996.

6. Summary and Conclusions

The ERS-1 SAR wave mode (SWM) data represent the first data set to provide detailed spectral information of ocean surface waves with continuous global coverage over several years. In the first part of this paper we have carried out an assessment of the SWM data quality and the performance (fidelity) of the ocean wave spectral retrieval algorithm by HBHH over the 3-year period January 1993 to December 1995 for which reliable SWM data are available. In the second part we investigated the statistical properties of the retrieved wave spectra and intercompared the retrievals with wave spectra computed by the operational global wave model WAM of the European Centre for Medium-Range Forecasts.

Approximately 75% of the available SWM data yielded successful retrievals. This net return factor represented the product of two quality tests. About 12% of the SWM data were rejected for data quality reasons related to too low signal-to-noise ratios or contaminations such as slicks or wind rows. The retrieval performance (defined in terms of the level of agreement between the observed SAR spectrum and the SAR spectrum computed from the retrieved wave spectrum) was accepted for roughly 85% of the SWM data that passed the first test.

Iteration of the first-guess spectrum used as input for the inversion was found to have a significant positive impact on the fidelity of the retrievals. Tests with different numbers of iteration cycles suggest that one or preferably two iterations may represent a reasonable compromise in operational applications between computing costs and retrieval fidelity. For the present study we carried out five iterations and chose as the retrieval the inversion with the highest fidelity.

In the statistical investigations in the second part of the

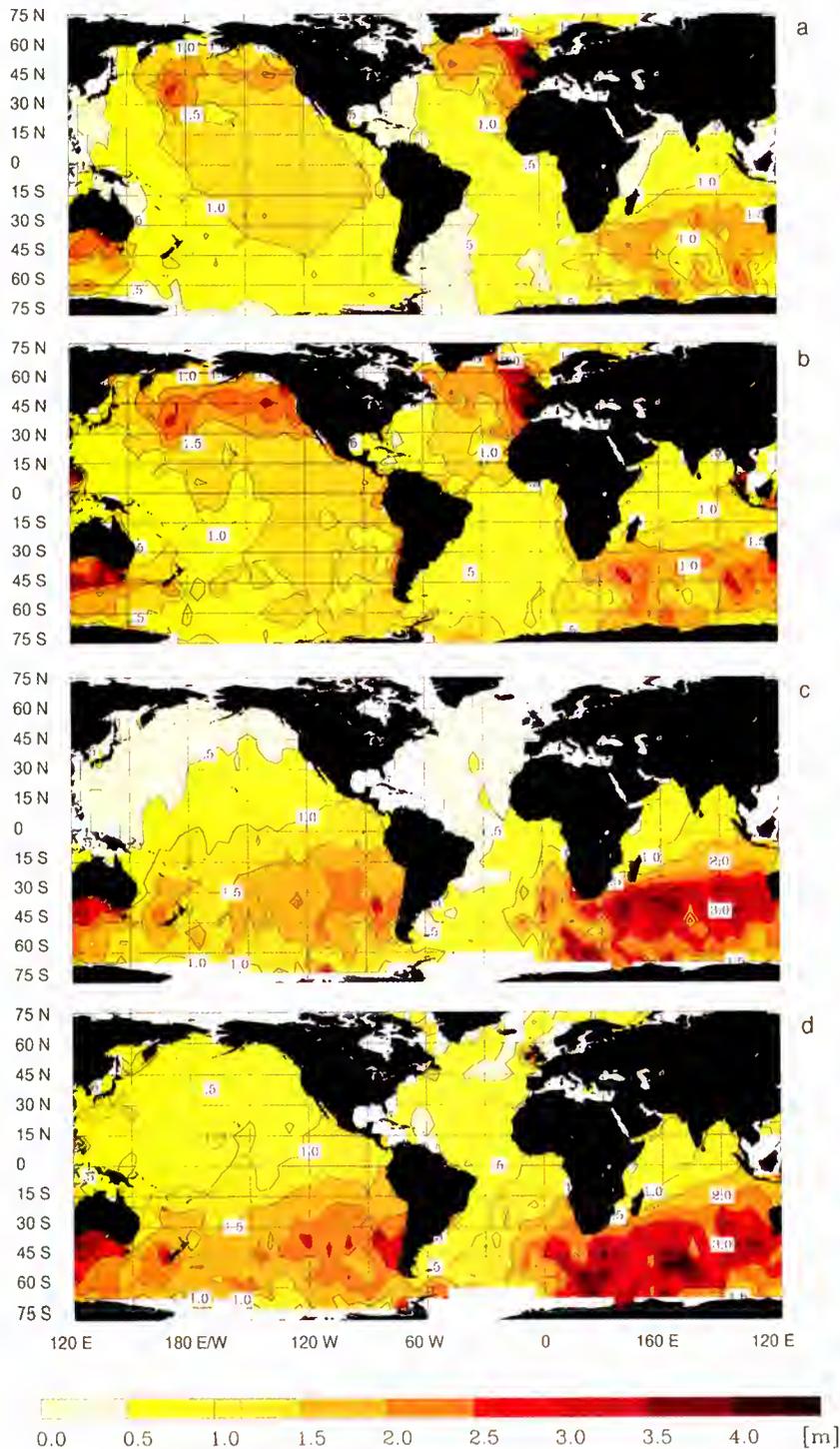


Plate 7. Global maps of seasonal mean wave height distributions of individual largest wavelength swell systems of wavelength exceeding 250 m. Plates 7a–7d same as Plates 5a–5d, respectively.

paper the SWM-retrieved data were analyzed together with the corresponding data computed with the ECMWF WAM model. The data were subdivided into wind-sea and swell constituents and were presented in the form of time series for different ocean basins, global maps for different seasons, and polar “wave rose” plots of directional energy distributions. In addition, results were presented for the dominant long wavelength swell components and for the cutoff wavelengths of the SWM

spectra (which provide a direct measure of the rms orbital velocity).

In the discussion of the intercomparison results, it was assumed that the retrieved spectra represented the “truth”; all discrepancies between the retrieved and modeled data were attributed to the model. This assumption will need to be scrutinized more carefully as further data become available, but for the present state of knowledge it appears a reasonable assump-

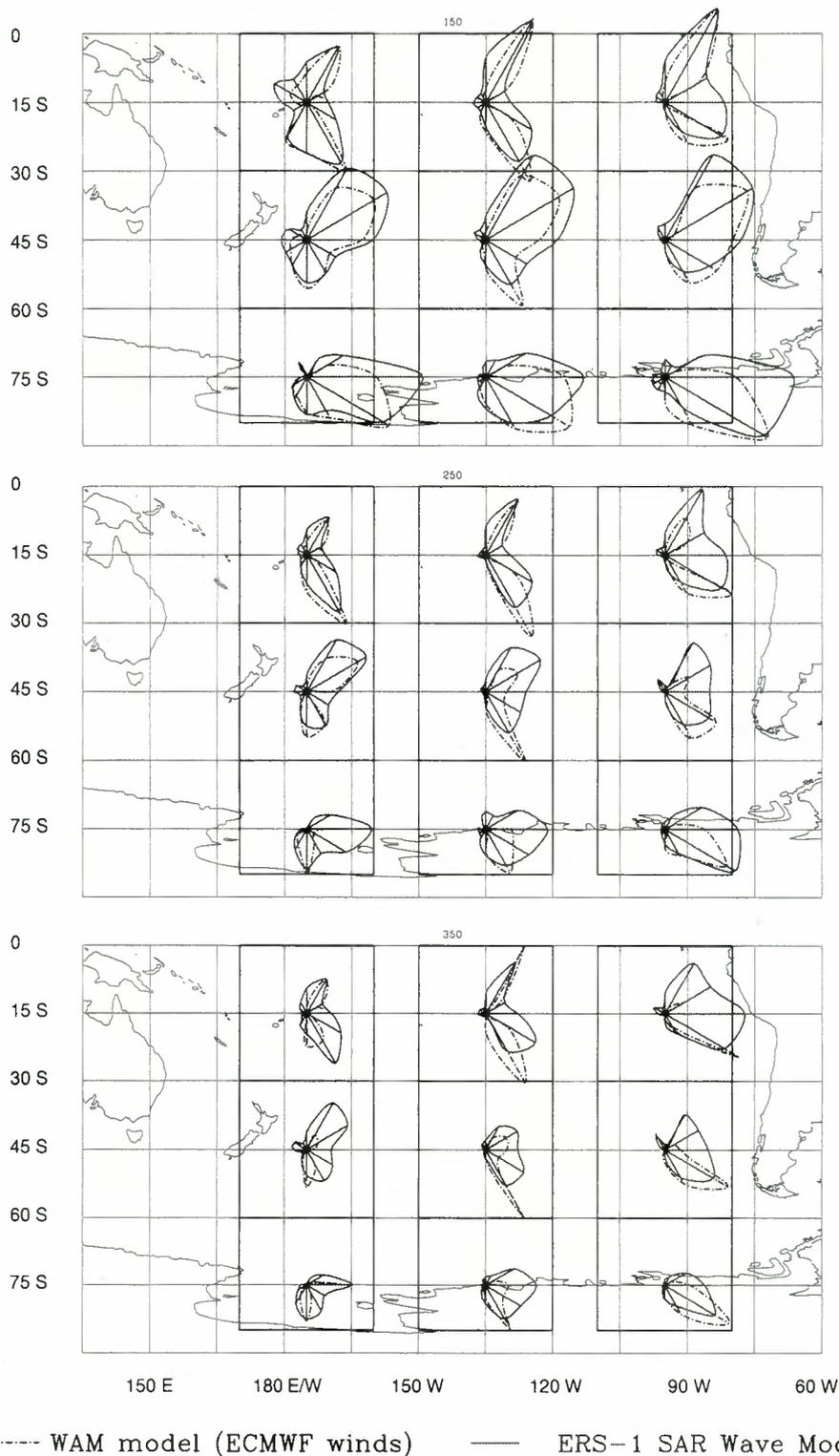


Figure 22a. Directional wave height distributions of individual longest wavelength swell systems in various sampling areas of the South Pacific for different lower wavelength thresholds; (top) $\lambda > 150$ m, (middle) $\lambda > 250$ m, and (bottom) $\lambda > 350$ m. For DJF 1993–1994.

tion; it would be difficult to explain the various types of discrepancies for wind sea and swell, different ocean regions, and different propagation directions, in terms of a common error source in the SAR retrievals. Also, validation studies of wave spectra retrieved from SAR spectra (e.g., HH; HBHH; E. Bauer and P. Heimbach, manuscript in preparation, 1998)

have revealed no obvious significant systematic errors in the SAR retrievals.

The retrieved and modeled data both confirm on a global scale the complexity of the spatial, temporal, and spectral distributions of ocean wave energy found previously by Brüning *et al.* [1994] in their first analysis of a 3-day period of SWM

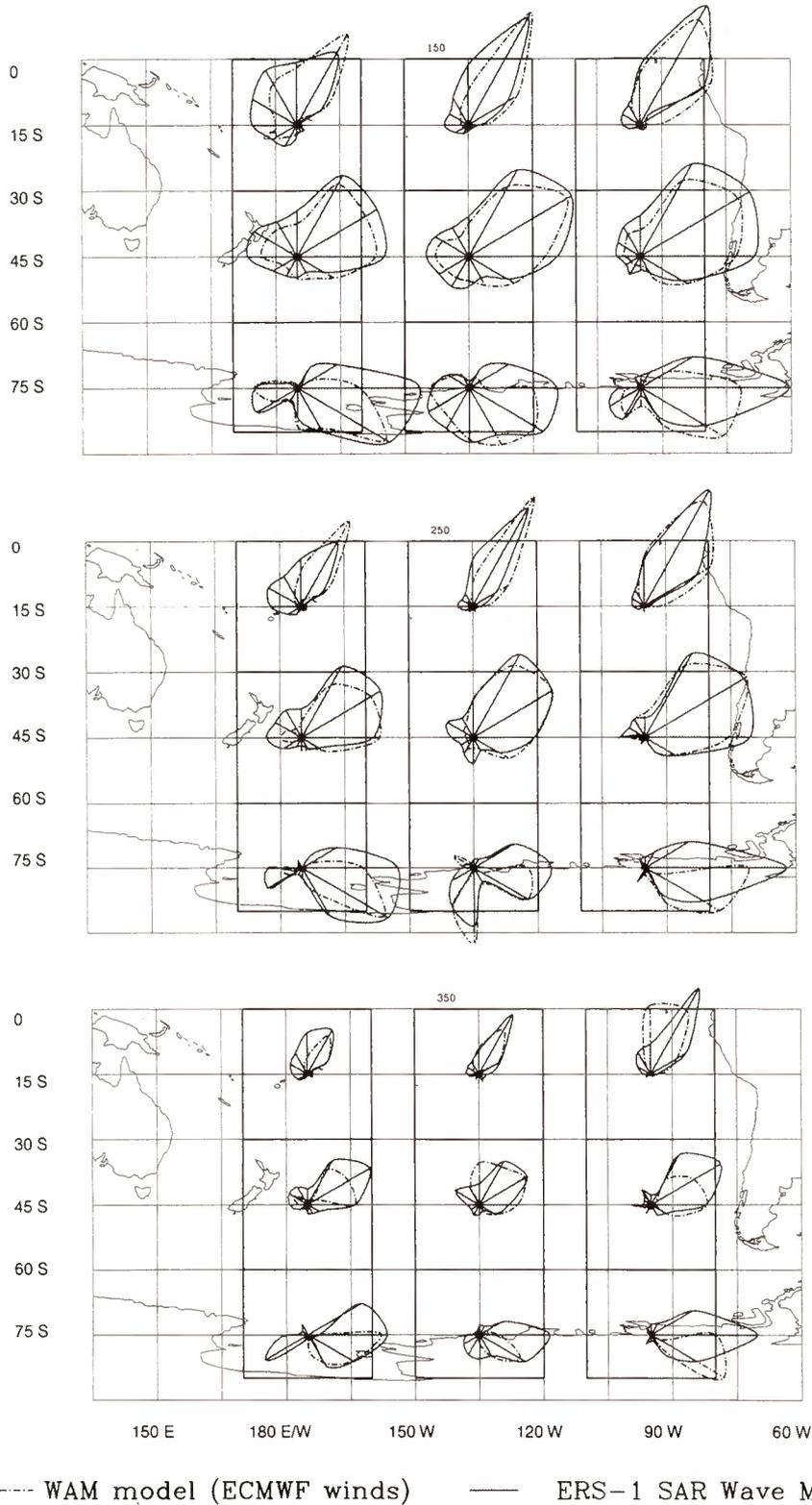


Figure 22b. Same as Figure 22a, but for JJA 1994.

spectral data for the Atlantic. In view of this complexity the general structural agreement between the modeled and retrieved spectra is encouraging. It has been demonstrated by HH and HBHH that this cannot be attributed to the use of a first-guess model spectrum in the retrieval algorithm; the retrieval is largely decoupled from the input spectrum, particularly with the introduction of an iterative modification of the

input spectrum in the improved algorithm by HBHH. Their result is confirmed by further detailed sensitivity tests of the present data set.

The availability of global spectral data has made possible a more detailed analysis of the performance of the WAM model than has previously been feasible. However, the present statistical analysis summarized below is necessarily restricted in

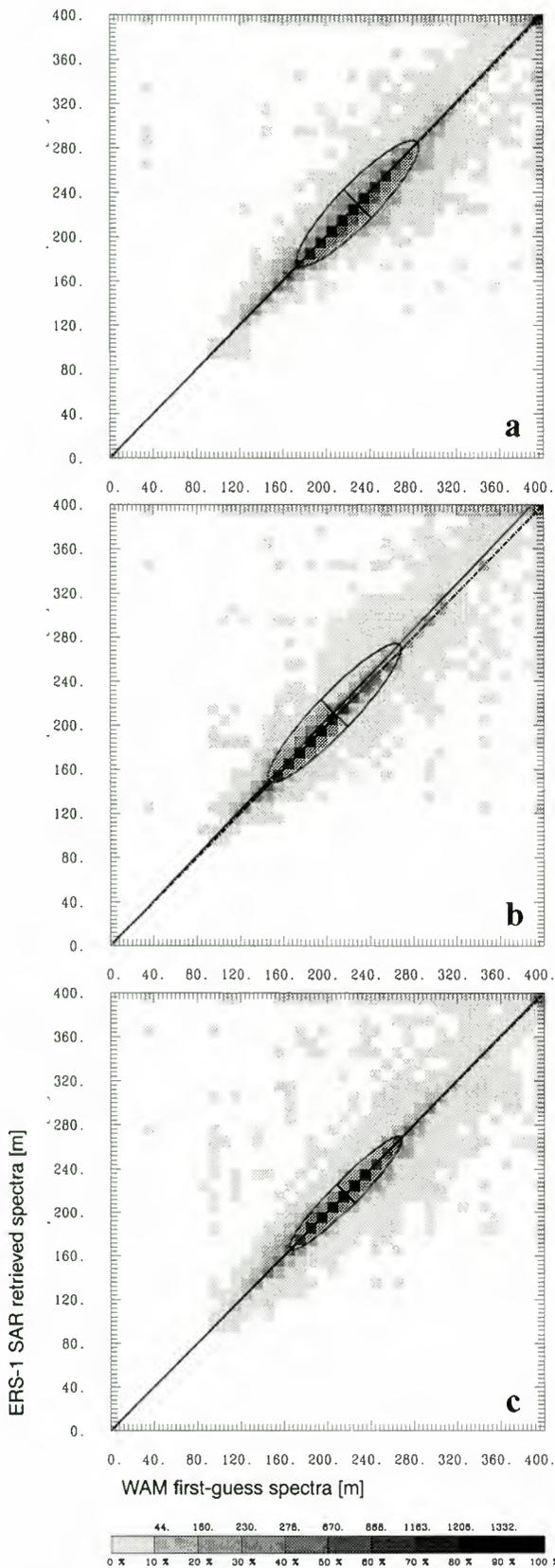


Figure 23. Regression between observed and simulated azimuthal cutoff wavelengths for (a) northern hemisphere extratropics, (b) tropics, and (c) southern hemisphere extratropics.

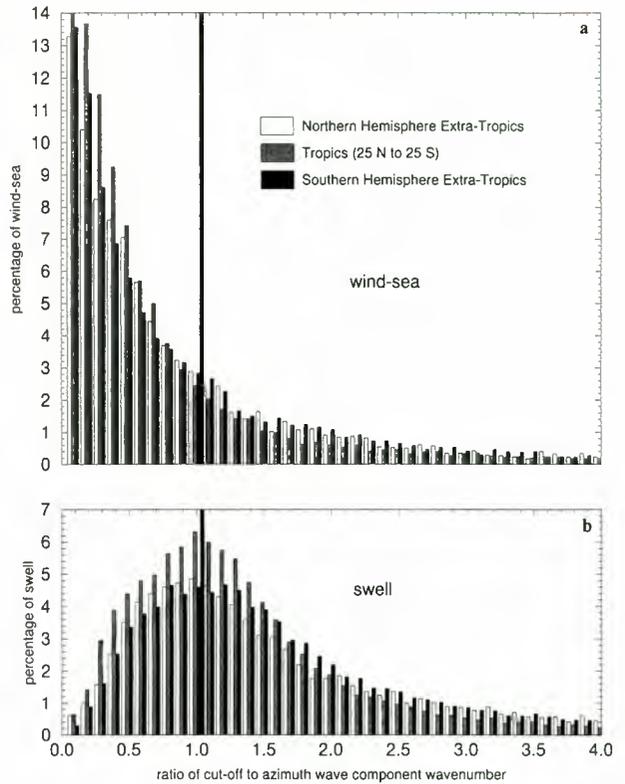


Figure 24. (a) Histogram of (a) wind-sea and (b) swell systems with respect to the ratio k_c/k_x of the observed azimuthal cutoff wavenumber to the mean azimuthal wavenumber component. The data for the period DJF 1995–1996 were divided into three ensembles for the northern hemisphere extratropics, the tropics between 25°N and 25°S, and the southern hemisphere extratropics.

scope. Moreover, a quantitative assessment of the dynamical effects identified by our statistical approach is made difficult by the continual assimilation of observed altimeter data in the ECMWF operational model. This attenuates systematic model errors, for example, in swell propagation. Specific investigations of a larger ensemble of individual wave events and the application of the WAM model without wave data assimilation can be expected to yield significantly more insight into the basic wave dynamical processes governing the general statistical results presented here. Nevertheless, the following general conclusions can be drawn from our analysis.

1. The time series and regression parameters of significant wave height indicated a good overall agreement between mod-

Table 4. Characteristics of Histograms of Observed Azimuthal Cutoff Wavenumber to Wavenumber Component in the Azimuth Direction

	Median		Percent of Waves With Ratio $k_c/k_x > 1$	
	Wind Sea	Swell	Wind Sea	Swell
Northern hemisphere	0.65	1.4	34	72
Tropics	0.5	1.85	28	78
Southern hemisphere	0.65	1.5	35	74

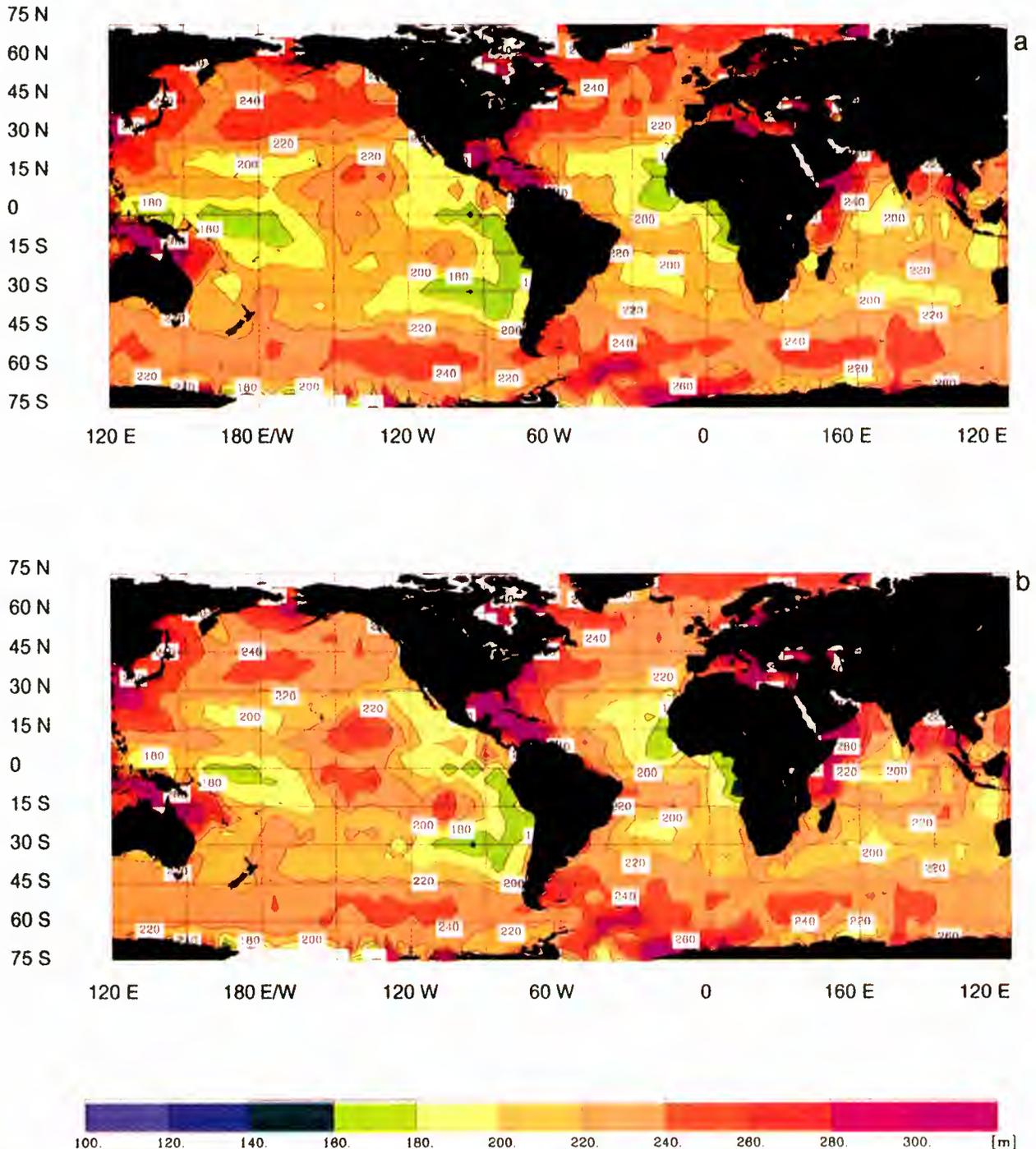


Plate 8. Global maps of the seasonal mean azimuthal cutoff wavelength for DJF 1995–1996. (a) Cutoff simulated from retrieved wave spectrum; (b) cutoff from observed SWM imagette spectrum.

eled and retrieved spectra, with a small but systematic underestimation of H_s by the WAM model in all ocean basins.

2. The decomposition of the wave energy into wind-sea and swell components revealed that the net wave height underestimation by the model is due to an underprediction of the swell components by 20–30%, the wind sea being slightly overpredicted on average by about 10%. This holds with few exceptions throughout the 3-year period and for all ocean basins and seasons.

3. The relative wind-sea overprediction was found to be

fairly constant, but slightly higher for high sea states in the midlatitude winters, suggesting that either the wind dependence of the WAM input source term should be slightly reduced or that the analyzed wind fields used to drive the model are too strong.

4. The relative swell bias is fairly uniformly distributed but exhibits maxima in the midlatitude summers. In contrast to the fairly constant relative wind-sea bias, the negative swell bias increases in relative magnitude with decreasing swell wave height. We interpret this as reflecting a dependence of the

swell underprediction on the wave age, larger wave ages being associated statistically with smaller wave heights (wave age refers here literally to the travel time of the swell since emerging from the wind-sea generation region). A model error with this characteristic may be explained by a dissipation source term which is too large for low frequencies or not sufficiently nonlinear. Another possible explanation is an inadequate representation of the nonlinear transition from wind sea to swell in the strongly simplified discrete interaction parametrization. However, this would not readily explain the increase of the swell error with decreasing wave height.

5. Modeled wavelengths were generally found to be in close agreement with the retrieved wavelengths, also favoring the hypothesis of errors in the swell dissipation rather than in the representation of the nonlinear wind-sea-swell transition process.

6. Most of the temporal and spatial features of the wave field distributions could be explained by the general properties of the seasonally varying global wind system. The data for both wind sea and swell exhibited a pronounced fetch dependence extending across entire ocean basins. This information is valuable for determining the transition of a growing wind sea to a fully developed state, for which little data is otherwise available. Intercomparisons of retrieved and modeled fetch dependences suggest no systematic errors of the WAM model in this respect. Global maps of mean swell wavelengths and the wave heights of long swell, together with polar "wave rose" distributions of wave propagation directions, were useful in identifying the radiation emanating from the midlatitude storm areas into distant regions. Some discrepancies between modeled and retrieved swell data were detected in the radiation of long swell from the northern into the southern hemisphere, which appeared to be more effective in the model than observed.

In summary we emphasize again that our conclusions from this first exploratory study of the 3-year ERS-1 SAR wave mode data set were based solely on statistical analyses. These have provided an interesting overview of the seasonally varying wind-sea and swell distributions of the global ocean and have demonstrated an encouraging overall correspondence between the complex wave systems inferred from the ERS SAR wave mode data and computed by the ECMWF operational wave model WAM. The statistical analyses also enabled the identification of a number of discrepancies between the two data sets, pointing to various features and possible shortcomings of the WAM model or the forcing wind fields, which require closer investigation. However, we have not undertaken investigations of individual events, which would have presumably shed more light on the underlying wave dynamical processes responsible for the mean distributions and observed discrepancies. It is hoped that this first investigation will motivate such dynamical event studies, for which the ERS-1/2 SWM system has provided and is still producing a unique comprehensive database.

Appendix

A1. Orbit, Instrument, and Data Specifications for the ERS-1 SWM

Tables 5 and 6 list data on the ERS-1 satellite, its orbit, the SAR instrument, the wave mode fast delivery product, and the repeat cycles between July 1992 and May 1996.

Table 5. Some ERS-1 Satellite and SAR Instrument Specifications

Aspect	Description
Start	July 17, 1991
Altitude (nadir)	782–785-km polar orbit (Sun synchronous)
Inclination	98.52°
Orbit period	about 100 min (≈ 14.3 orbits per day)
Platform velocity	7455 m/s
Daily coverage	about 1100 spectra per day on average
Incidence angle	until June 1995, 19.9° at center of imagette (right looking); after June 1995, 23.5°
Polarization	linear vertical (VV)
Radar Bragg frequency	5.3 GHz, C band (5.6 cm)
Imagette size	10 km (range) \times 5 km (azimuth)
Product spacing	200 km along track (every 30 s)
Spectral grid	12 wavenumbers in logarithmic spacing corresponding to wavelengths between 100 and 1000 m
	12 directional sectors of 15° width (between 0° and 180°, anticlockwise relative to azimuth)
Intensity	8 bit (0–255) relative to maximum
Number of looks	three-look incoherent averaging of single-look amplitude
SAR integration time	1.8 s
R/v ratio	112 s (R , distance SAR – target; v , platform velocity in orbit)

A2. SWM, WAM, and Cost Function Grid

Three different Cartesian wavenumber, polar wavenumber, or frequency/direction grids are involved in the inversion process: (1) the ERS-1 SAR Wave Mode grid (k , θ); the SWM spectra of the FDP are given on a polar wavenumber grid on a semidisc and are composed of 12 15° angular sectors between 0° and 180° running counterclockwise relative to the satellite flight direction (azimuth) and $N_{\text{swm}} = 12$ wavenumbers corresponding to wavelengths between 100 and 1000 m on a logarithmic scale

$$k_{j+1} = 10^{1/(N_{\text{swm}}-1)} k_j \quad j = 1, \dots, N_{\text{swm}}$$

with

$$k_1 = \frac{2\pi}{100} \text{ m}^{-1}$$

(2) the WAM model grid (f , θ); the first-guess WAM and the retrieved wave spectra are defined on a polar frequency/direction grid composed of 12 30° angular sectors between 0° and 360°, running clockwise relative to north, and $N_{\text{wam}} = 25$

Table 6. ERS-1 Mission Phases and Repeat Cycles Between July 1992 and May 1996

Phase	Start	End	Repeat Cycle, days	Orbits Per Cycle
C	April 14, 1992	December 21, 1993	35	501
D	December 24, 1993	April 10, 1994	3	43
E	April 10, 1994	September 28, 1994	168	2408
F	September 28, 1994	March 21, 1995	168	2408
G	March 21, 1992	end	35	501

The satellite was switched off in June 1996 but is parked in orbit in a "sleeping" mode.

frequencies corresponding to logarithmically spaced wavelengths between 9 and 895 m, or frequencies

$$f_{j+1} = 1.1f_j \quad j = 1, \dots, N_{\text{wam}}$$

with

$$f_1 = 0.04177 \text{ Hz}$$

(3) the cost function grid (k_x, k_y) ; both the observed SWM image spectrum and the first-guess wave spectrum are interpolated to a $N_{\text{cart}} \times N_{\text{cart}}$ cartesian wavenumber grid with $N_{\text{cart}} = 128$ and

$$k_x(j), k_y(j) = j\Delta k$$

with

$$-\frac{N_{\text{cart}}}{2} < j \leq \frac{N_{\text{cart}}}{2} \quad \Delta k = \frac{2\pi}{32} \text{ m}^{-1}$$

on which all computations (nonlinear mapping, cost function, etc.) are performed.

To properly account for the spectral smoothing incurred through the processing of the SWM image to the low-resolution polar wavenumber grid of the SWM FDP image spectrum, a smoothing filter is introduced whenever a SAR image spectrum is computed from a model spectrum. The filter is obtained by transforming the computed SAR image spectrum from the fine Cartesian grid (k_x, k_y) to the coarse polar grid (k, θ) and subsequently interpolating back to the (k_x, k_y) grid, yielding a net smoothing of the simulated high-resolution SAR spectrum analogous to the treatment of the FDP.

A3. Regression Analysis

In standard regression analysis a straight line $y(x) = mx + a$ is drawn through a scatterplot of data pairs (x_i, y_i) with slope m and y -axis intercept a , chosen such that the sum of the square differences $\sum_i [y_i - y(x_i)]^2$ is minimized. However, the use of two regression parameters makes it more cumbersome to intercompare the effective calibration of two data sets, and we have therefore restricted the regression line in our analyses to pass through the origin, $a = 0$ (comparisons with adjustable intercepts have shown that these in fact have little influence).

Instead of the asymmetric definition of the square error in terms of the ordinate errors, we have then fitted our regression line

$$y_i = bx_i$$

through the origin by minimizing the sum of the orthogonal distances of the data points from the regression line. The extreme values of this quantity yield the major and minor axes of the covariance ellipse. The method is not only symmetric with respect to an interchange of x and y but is also invariant with respect to rotations in the x, y plane [see, e.g., Preisendorfer, 1988; Bauer et al., 1992; Bauer and Staabs, 1998]. The slope b is given by

$$b = \tan \phi \quad \tan 2\phi = \frac{\langle 2xy \rangle}{\langle x^2 \rangle - \langle y^2 \rangle}$$

and $\langle \rangle$ denotes the average over the data ensemble. The angle ϕ as defined above is actually not invariant with respect to an interchange of x and y , which yields $\phi \rightarrow \phi + 90^\circ$. The correct solution yielding the major rather than the minor ellipse axis can be determined by comparing the square errors

for the two solutions. (The only exception to our symmetrical regression approach is in the scatter diagram of Figure 8, comparing the retrievals for a modified and nonmodified input spectrum. Here we resorted to the standard nonsymmetrical regression analysis, as this yielded a more realistic-looking line passing through the maxima of the distribution.)

Acknowledgments. This work is a contribution to the ESA Pilot Project PP2-D1 "Derivation and assimilation of ERS-1 wind and wave data" coordinated by W. Alpers of the Institut für Meereskunde (IfM), Universität Hamburg. P. Heimbach was funded within the project SFB 318/TP C-1 "Klimarelevante Prozesse im System Ozean-Atmosphäre-Kryosphäre" of the Deutsche Forschungsgemeinschaft (DFG). S. Hasselmann was supported by the Office of Naval Research (ONR) under grants N00014-92-J-1840 and N00014-1-0541. The WAM model spectra were provided by Björn Hansen from ECMWF (Reading, England). Part of the ERS-1 SWM fast delivery product was decoded by Christian Melsheimer at the IfM. The retrieval runs were performed on a CRAY C916 at the Deutsches Klimarechenzentrum (DKRZ). We are grateful to Eva Bauer for fruitful discussions as well as to Renate Brokopf, Christian Bennefeld, Maurice Toporek, and the staff of the DKRZ for their technical assistance.

References

- Allan, T. D. (Ed.), *Satellite Microwave Remote Sensing*, 526 pp., Ellis Horwood, Chichester, England, 1983.
- Alpers, W., Monte Carlo simulations for studying the relationship between ocean wave and synthetic aperture radar image spectra, *J. Geophys. Res.*, **88**, 1745–1759, 1983.
- Alpers, W., and K. Hasselmann, Spectral signal-to-clutter and thermal noise properties of ocean wave imaging synthetic aperture radars, *Int. J. Remote Sens.*, **3**, 423–446, 1982.
- Alpers, W., D. B. Ross, and C. L. Rufenach, On the detectability of ocean surface waves by real and synthetic aperture radar, *J. Geophys. Res.*, **86**, 6481–6498, 1981.
- Alpers, W., C. Brüning, and K. Richter, Comparison of simulated and measured synthetic aperture radar image spectra with buoy-derived ocean wave spectra during the shuttle imaging radar-B mission, *IEEE Trans. Geosci. Remote Sens.*, **24**, 559–566, 1986.
- Bacon, S., and D. J. T. Carter, Wave climate changes in the North Atlantic and North Sea, *Int. J. Climatol.*, **11**, 545–558, 1991.
- Bamler, R., and B. Schättler, SAR geocoding: Data and systems, in *SAR Data Acquisition and Image Formation*, edited by G. Schreier, pp. 53–102, Wichmann, Heidelberg, Germany, 1993.
- Bao, M., and W. Alpers, On the cross spectrum between single-look synthetic aperture radar images of ocean waves, *IEEE Trans. Geosci. Remote Sens.*, in press, 1998.
- Bao, M., C. Brüning, and W. Alpers, A generalized nonlinear ocean wave-SAR spectral integral transform and its application of ERS-1 SAR ocean wave imaging, in *Proceedings of the Second ERS-1 Symposium, Hamburg, Germany, 11–14 October 1993*, Eur. Space Agency Spec. Publ., ESA SP-361, 219–224, 1994.
- Bauer, E., and C. Staabs, Statistical properties of global significant wave heights and their use for validation, *J. Geophys. Res.*, **103**, 1153–1166, 1998.
- Bauer, E., S. Hasselmann, K. Hasselmann, and H. C. Graber, Validation and assimilation of Seasat altimeter wave heights using the WAM wave model, *J. Geophys. Res.*, **97**, 12,671–12,682, 1992.
- Bauer, E., S. Hasselmann, and K. Hasselmann, An operational wave forecast system using wind and wave data, in *Proceedings of the Second ERS Applications Workshop, London, 6–8 December 1995*, Eur. Space Agency Spec. Publ., ESA SP-383, 309–312, 1996a.
- Bauer, E., K. Hasselmann, I. R. Young, and S. Hasselmann, Assimilation of wave data into the wave model WAM using an impulse response function, *J. Geophys. Res.*, **101**, 3801–3816, 1996b.
- Beal, R. C. (Ed.), *Directional Ocean Wave Spectra*, Johns Hopkins Univ. Press, Baltimore, Md., 1991.
- Beal, R. C., T. W. Gerling, F. M. Monaldo, and D. G. Tilley, Measuring ocean waves from space, *Int. J. Remote Sens.*, **12**, 1713–1722, 1991.
- Bouws, E., D. Jannink, and G. J. Komen, The increasing wave height in the North Atlantic Ocean, *Bull. Am. Meteorol. Soc.*, **77**, 2275–2277, 1996.

- Breivik, L.-A., B. Haugse, H. Johnsen, H. E. Krogstad, and M. Reistad, Evaluation of SAR wave products, *Res. Rep. 2, Norw. Meteorol. Inst., Oslo*, January 1995.
- Breivik, L.-A., M. Reistad, H. Schyberg, J. Sunde, H. E. Krogstad, and H. Johnsen, Assimilation of ERS SAR wave spectra in an operational wave model, *J. Geophys. Res.*, this issue.
- Brooker, G., UWA processing algorithm specification, *Rep. ER-TN-ESA-GS-0342*, version 2.0, 26 pp., Eur. Space Res. and Technol. Cent., Noordwijk, Netherlands, 1995.
- Brüning, C., On the relative importance of the ocean wave-radar modulation transfer function for inverting ERS-1 SAR image spectra into ocean wave spectra, in *Proceedings of the Second ERS-1 Symposium, Hamburg, Germany, 11-14 October 1993*, Eur. Space Agency Spec. Publ., ESA SP-361, 233-238, 1994.
- Brüning, C., W. Alpers, L. F. Zambresky, and D. G. Tilley, Validation of a SAR ocean wave imaging theory by the shuttle imaging radar-B experiment over the North Sea, *J. Geophys. Res.*, 93, 15,403-15,425, 1988.
- Brüning, C., W. Alpers, and K. Hasselmann, Monte-Carlo simulation studies of the nonlinear imaging of a two-dimensional surface wave field by a synthetic aperture radar, *Int. J. Remote Sens.*, 90, 1695-1727, 1990.
- Brüning, C., S. Hasselmann, K. Hasselmann, S. Lehner, and T. Gerling, On the extraction of ocean wave spectra from ERS-1 SAR wave mode image spectra, in *Proceedings of the First ERS-1 Symposium, Cannes, France, 4-6 November 1992*, Eur. Space Agency Spec. Publ., ESA SP-359, 747-752, 1993.
- Brüning, C., S. Hasselmann, K. Hasselmann, S. Lehner, and T. Gerling, A first evaluation of ERS-1 synthetic aperture radar wave mode data, *Global Atmos. Ocean Syst.*, 2, 61-98, 1994.
- Chapron, B., T. Elfouhaily, and V. Kerbaol, Calibration and validation of ERS wave mode products, *Doc. DRO/OS/95-02*, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, March 1995.
- Egen, G., and H. Johnsen, SAR ocean wave inversion using image cross spectra, *IEEE Trans. Geosci. Remote Sens.*, 33, 1047-1056, 1995.
- European Centre for Medium-Range Weather Forecasts, (ECMWF), User guide to ECMWF products, *Meteorol. Bull. M3.2*, edition 2.1, Reading, England, 1995.
- European Space Agency (ESA), ERS user handbook, *Eur. Space Agency Spec. Publ., ESA SP-1148*, 30-31, 1993.
- European Space Agency (ESA), Application achievements of ERS-1, *Eur. Space Agency Spec. Publ., ESA SP-1176/II*, pp. 15-18, 1996.
- European Space Agency (ESA), ENVISAT-1, Mission and system summary, report, Eur. Space Res. and Technol. Cent., Noordwijk, Netherlands, March 1997.
- Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young, Bulk parameterization of air-sea fluxes for Tropical Ocean-Global Atmosphere Coupled Ocean-Atmosphere Response Experiment, *J. Geophys. Res.*, 101, 3747-3764, 1996.
- Feindt, F., Radar-Rückstreuxperimente am Wind-Wellen-Kanal bei sauberer und filmbedeckter Wasseroberfläche, Ph.D. thesis, Fachbereich Geowiss., Univ. of Hamburg, Germany, 1985.
- Geernaert, G. L., Bulk parameterizations for the wind stress heat fluxes, in *Surface Waves and Fluxes*, vol. 1, edited by G. L. Geernaert and W. L. Plant, chap. 5, pp. 91-172, Kluwer Acad., Norwell, Mass., 1990.
- Gerling, T. W., Partitioning sequences and arrays of directional wave spectra into component wave systems, *J. Atmos. Oceanic Technol.*, 9, 444-458, 1992.
- Günther, H., S. Hasselmann, and P. A. E. M. Janssen, WAModel Cycle 4 (revised version), *Tech. Rep. 4*, Ger. Clim. Comput. Cent., Hamburg, Germany, 1992.
- Hansen, B., C. Brüning, and C. Staabs, Global comparison of significant wave heights derived from ERS-1 SAR wave mode, ERS-1 altimeter and TOPEX altimeter data, in *Proceedings of the Second ERS-1 Symposium, Hamburg, Germany, 11-14 October 1993*, Eur. Space Agency Spec. Publ., ESA SP-361, 33-36, 1994.
- Harger, R. O., *Synthetic Aperture Radars: Theory and Design*, Academic, San Diego, Calif., 1970.
- Hasselmann, K., On the spectral dissipation of ocean waves due to white capping, *Boundary Layer Meteorol.*, 6, 107-127, 1974.
- Hasselmann, K., and S. Hasselmann, On the nonlinear mapping of an ocean wave spectrum into a synthetic aperture radar image spectrum and its inversion, *J. Geophys. Res.*, 96, 10,713-10,729, 1991.
- Hasselmann, K., R. K. Raney, W. J. Plant, W. Alpers, R. A. Shuchman, D. R. Lyzenga, C. L. Rufenach, and M. J. Tucker, Theory of synthetic aperture radar ocean imaging, A MARSEN view, *J. Geophys. Res.*, 90, 4659-4686, 1985.
- Hasselmann, K., S. Hasselmann, E. Bauer, C. Brüning, S. Lehner, H. Gaber, and P. Lionello, Development of a satellite SAR image spectra and altimeter wave height data assimilation system for ERS-1, *Rep. 19*, Max-Planck-Inst. für Meteorol., Hamburg, Germany, July 1988.
- Hasselmann, K., S. Hasselmann, C. Brüning, and A. Speidel, Interpretation and application of SAR wave image spectra in wave models, in *Directional Ocean Wave Spectra*, edited by R. C. Beal, pp. 117-124, Johns Hopkins Univ. Press, Baltimore, Md., 1991.
- Hasselmann, S., and K. Hasselmann, Computations and parameterizations of the nonlinear energy transfer in a gravity wave spectrum, I, A new method for efficient computations of the exact nonlinear energy transfer integral, *J. Phys. Oceanogr.*, 15, 1369-1377, 1985.
- Hasselmann, S., K. Hasselmann, J. H. Allender, and T. P. Barnett, Computations and parameterizations of the nonlinear energy transfer in a gravity wave spectrum, II, Parameterizations of the nonlinear energy transfer for application in wave models, *J. Phys. Oceanogr.*, 15, 1378-1391, 1985.
- Hasselmann, S., C. Brüning, and P. Lionello, Towards a generalized optimal interpolation method for the assimilation of ERS-1 SAR retrieved wave spectra in a wave model, in *Proceedings of the Second ERS-1 Symposium, Hamburg, Germany, 11-14 October 1993*, Eur. Space Agency Spec. Publ., ESA SP-361, 21-25, 1994.
- Hasselmann, S., C. Brüning, K. Hasselmann, and P. Heimbach, An improved algorithm for the retrieval of ocean wave spectra from SAR image spectra, *J. Geophys. Res.*, 101, 16,615-16,629, 1996.
- Hasselmann, S., P. Lionello, and K. Hasselmann, An optimal interpolation scheme for the assimilation of spectral wave data, *J. Geophys. Res.*, 102, 15,823-15,836, 1997.
- Hasselmann, S., P. Heimbach, and C. Bennefeld, Algorithm for retrieving ocean wave spectra from SAR image spectra, technical report, Ger. Clim. Comput. Cent., Hamburg, Germany, 1998.
- Heimbach, P., S. Hasselmann, C. Brüning, and K. Hasselmann, Application of wave spectral retrievals from ERS-1 SAR wave mode data for improved wind and wave field analyses, in *Proceedings of the Second ERS Applications Workshop, London, 6-8 December 1995*, Eur. Space Agency Spec. Publ., ESA SP-383, 303-308, February 1996.
- Jähne, B., and E. C. Monahan (Eds.), *Air-Water Gas Transfer*, AEON, Hanau, Germany, 1995.
- Janssen, J. A. M., Does wind stress depend on sea-state or not?—A statistical error analysis of HEXMAX data, *Boundary Layer Meteorol.*, 83, 479-503, 1997.
- Janssen, P. A. E. M., Wave-induced stress and the drag of the air flow over sea waves, *J. Phys. Oceanogr.*, 19, 745-754, 1989.
- Janssen, P. A. E. M., Quasi-linear theory of wind-wave generation applied to wave forecasting, *J. Phys. Oceanogr.*, 21, 1631-1391, 1991.
- Janssen, P. A. E. M., Experimental evidence of the effect of surface waves on the airflow, *J. Phys. Oceanogr.*, 22, 1600-1604, 1992.
- Janssen, P. A. E. M., and P. Viterbo, Ocean waves and the atmospheric climate, *J. Clim.*, 9, 1269-1287, 1996.
- Janssen, P. A. E. M., B. Hansen, and J. Bidlot, Verification of the ECMWF wave forecasting system against buoy and altimeter data, *Tech. Memo. 229*, Eur. Cent. for Medium Range Weather Forecasts, Reading, England, July 1996.
- Johnsen, H., Multi-look versus single-look processing of synthetic aperture radar images with respect to ocean wave spectra estimation, *Int. J. Remote Sens.*, 13, 1627-1643, 1992.
- Kerbaol, V., and B. Chapron, Calibration and validation of ERS-1/2 wave mode products, report, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, October 1996.
- Komen, G. J., L. Cavaleri, M. Donelan, K. Hasselmann, S. Hasselmann, and P. A. E. M. Janssen, *Dynamics and Modelling of Ocean Waves*, 560 pp., Cambridge Univ. Press, New York, 1994.
- Komen, G. J., P. A. E. M. Janssen, V. K. Makin, C. Mastenbroek, and W. A. Oost, On the sea state dependence of the Charnock parameter, *Global Atmos. Ocean Syst.*, in press, 1998.
- Krogstad, H. E., A simple derivation of Hasselmann's nonlinear ocean-synthetic aperture radar transform, *J. Geophys. Res.*, 97, 2421-2425, 1992.
- Krogstad, H. E., The speckle spectrum of the ERS-1 wave mode product, *SINTEF Industrial Mathematics Rep. STF10 A94008*, Trondheim, Norway, 1994.
- Krogstad, H. E., O. Samset, and P. W. Vachon, Generalization of the

- non-linear ocean-SAR transform and a simplified SAR inversion algorithm, *Atmos. Ocean*, 32, 61–82, 1994.
- Large, W. G., J. Morzel, and G. B. Crawford, Accounting for surface wave distortion of the marine wind profile in low-level ocean storm wind measurements, *J. Phys. Oceanogr.*, 25, 2959–2971, 1995.
- Lionello, P., H. Günther, and P. A. E. M. Janssen, Assimilation of altimeter data in a global third-generation wave model, *J. Geophys. Res.*, 97, 14,453–14,474, 1992.
- Lionello, P., H. Günther, and B. Hansen, A sequential assimilation scheme applied to global wave analysis and prediction, *J. Mar. Syst.*, 6, 87–107, 1995.
- Mastenbroek, C., G. Burgers, and P. A. E. M. Janssen, The dynamical coupling of a wave model and a storm surge model through the atmospheric boundary layer, *J. Phys. Oceanogr.*, 23, 1856–1866, 1993.
- Miles, J., On the generation of surface waves by shear flows, *J. Fluid Mech.*, 3, 185–204, 1957.
- Monaldo, F. M., and R. C. Beal, Real-time observations of Southern Ocean wave fields from the shuttle imaging radar, *IEEE Trans. Geosci. Remote Sens.*, 33, 942–949, 1995.
- Monaldo, F. M., and D. R. Lyzenga, On the estimation of wave slope and height variance spectra from SAR imagery, *IEEE Trans. Geosci. Remote Sens.*, 24, 543–551, 1986.
- Monaldo, F. M., and D. R. Lyzenga, Comparison of shuttle imaging radar-B ocean wave spectra with linear model predictions based on aircraft measurements, *J. Geophys. Res.*, 93(C12), 15,374–15,388, 1988.
- Peixoto, J. P., and A. H. Oort, *The Physics of Climate*, Am. Inst. of Phys., New York, 1992.
- Plant, W. J., and W. Alpers, An introduction to SAXON-FPN, *J. Geophys. Res.*, 99, 9699–9703, 1994.
- Preisendorfer, R. P., *Principal Component Analysis*, Elsevier, New York, 1988.
- Rieder, K. F., Analysis of sea-surface drag parameterizations in open ocean conditions, *Boundary Layer Meteorol.*, 83, 355–377, 1997.
- Rieder, K. F., J. A. Smith, and R. A. Weller, Observed directional characteristics of the wind, wind stress, and surface waves on the open ocean, *J. Geophys. Res.*, 99, 22,589–22,596, 1994.
- Smith, S. D., et al., Sea surface wind stress and drag coefficients: The HEXOS results, *Boundary Layer Meteorol.*, 60, 109–142, 1992.
- Snodgrass, F. E., G. W. Groves, K. Hasselmann, G. R. Miller, W. H. Munk, and W. H. Powers, Propagation of ocean swell across the Pacific, *Philos. Trans. R. Soc. London A*, 249, 431–497, 1966.
- Snyder, R. L., F. W. Dobson, J. A. Elliott, and R. B. Long, Array measurements of atmospheric pressure fluctuations above surface gravity waves, *J. Fluid Mech.*, 102, 1–59, 1981.
- Tolman, H. L., Effects of numerics on the physics in a third-generation wind-wave model, *J. Phys. Oceanogr.*, 22, 1095–1111, 1992.
- Tolman, H. L., On the selection of propagation schemes for a spectral wind-wave model, *Off. Note 411*, Natl. Weather Serv., La Jolla, Calif., 1995.
- Vachon, P. W., and R. K. Raney, Resolution of the ocean wave propagation direction in SAR imagery, *IEEE Trans. Geosci. Remote Sens.*, 29, 105–112, 1991.
- von Storch, H., J. Guddal, K. A. Iden, T. Jonson, J. Perlwitz, M. Reistad, J. de Ronde, H. Schmidt, and E. Zorita, Changing statistics of storms in the North Atlantic?, *Rep. 116*, Max-Planck-Inst. für Meteorol., Hamburg, Germany, 1993.
- Wave Model Development and Implementation (WAMDI) Group, The WAM model—A third generation ocean wave prediction model, *J. Phys. Oceanogr.*, 18, 1775–1810, 1988.
- Wave and Storms in the North Atlantic (WASA) Group, The WASA project: Changing storm and wave climate in the northeast Atlantic and adjacent seas?, paper presented at the Fourth International Workshop on Wave Hindcasting and Forecasting, Environ. Can., Banff, Alberta, Canada, Oct. 16–20, 1995.
- Weber, S. L., H. von Storch, P. Viterbo, and L. Zambresky, Coupling an ocean wave model to an atmospheric general circulation model, *Clim. Dyn.*, 9, 63–69, 1993.
- Weisse, R., and E. F. Alvarez, The European Coupled Atmosphere Wave Ocean Model ECAWOM, *Rep. 238*, Max-Planck-Inst. für Meteorol., Hamburg, Germany, July 1997.
- Wilde, A., C. Brüning, W. Alpers, V. Etkin, K. Litovchenko, A. Ivanov, and V. Zajtsev, Comparison of ocean wave imaging by ERS-1 and ALMAZ-1 synthetic aperture radar, in *Proceedings of the Second ERS-1 Symposium, Hamburg, Germany, 11–14 October 1993*, Eur. Space Agency Spec. Publ., ESA SP-361, 239–245, 1994.
- Young, I. R., W. Rosenthal, and F. Ziemer, A three-dimensional analysis of marine radar images for the determination of ocean wave directionality and surface currents, *J. Geophys. Res.*, 90, 1049–1059, 1985.
- Young, I. R., S. Hasselmann, and K. Hasselmann, Computations of the response of a wave spectrum to a sudden change in wind direction, *J. Phys. Oceanogr.*, 17, 1317–1338, 1987.

K. Hasselmann, S. Hasselmann, and P. Heimbach, Max-Planck-Institut für Meteorologie, Bundesstrasse 55, 20146 Hamburg, Germany. (e-mail: klaus.hasselmann@dkrz.de; susanne.hasselmann@dkrz.de; heimbach@dkrz.de)

(Received March 20, 1997; revised October 21, 1997; accepted November 4, 1997.)

Coastal lee waves on ERS-1 SAR images

Quanan Zheng, Xiao-Hai Yan, and Vic Klemas

Graduate College of Marine Studies, University of Delaware, Newark

Chung-Ru Ho and Nan-Jung Kuo

Department of Oceanography, National Taiwan Ocean University, Keelung, Taiwan

Zongming Wang

Texaco, Houston, Texas

Abstract. Alternative dark-bright patterns on ERS-1 synthetic aperture radar (SAR) images of the west side of the Taiwan Strait taken on December 8, 1994, were recognized to be the sea surface signature of a coastal lee wave. Such waves are called coastal lee waves because they occur along the lee side of the coast. The coastal lee waves appeared in the form of a wave packet distributed within an offshore band 20–40 km wide. The first packet, which occurred in the northern portion of the observed area, contained six waves with variable wavelengths (defined as the spatial separation between two waves) from 1.7 to 2.7 km. The second packet, in the middle, contained 10 waves with a relatively uniform wavelength of 4.2 km. The third packet, in the southern portion, contained 17 waves with an average wavelength of 2.0 km. The crest lengths were from 20 to 80 km. Local meteorologic parameters observed simultaneously at Fuzhou, China, close to the imaged area, showed an offshore wind of 1.5–3.5 m/s and a land surface air temperature of 19°C, which was 4°C lower than the sea surface temperature (SST). Thus the lower atmospheric boundary conditions at imaging time were very favorable both for generating the land breeze circulation and small wind waves on the sea surface, which are in the Bragg-scattering wavelength band of the C band ERS-1 SAR. A physical model of a three-layer atmosphere was developed in order to explain how the land breeze circulation can generate the coastal lee waves. The results showed that the vertical velocity disturbance caused by the wind convergence at the land breeze frontal zone is of vital importance for the generation of coastal lee waves, and the model gave very good estimates of the processes observed. The SAR imaging mechanisms of the waves were analyzed in detail. The differences between coastal lee waves and ocean internal waves, which appear as similar alternative dark-bright patterns on SAR images, were also discussed.

1. Introduction

Lee waves are a kind of atmospheric internal gravity waves, which occur in the lower layer of the atmosphere. It is believed that an obstacle to the airflow is a common reason for generation of lee waves. Mountain lee waves generated by isolated mountain obstacles on the lee side of mountains are a typical example of atmospheric lee waves and have been intensively studied. Observations showed that mountain lee waves are stationary waves and have a typical wavelength of 10 km, ranging from 2–3 to 40–45 km, and a vertical amplitude

of the same order of height as the mountain generating them [Gossard and Hooke, 1975; Panchev, 1985]. Lee waves, which are analogous to mountain lee waves but caused by an isolated linear island, have recently been observed by Vachon *et al.* [1994] using a European Remote Sensing Satellite (ERS-1) synthetic aperture radar (SAR) image. They found that a packet of six atmospheric lee waves with a wavelength of 7.6 km occurred on the lee side of the island Hopen in the Barents Sea, which is located near 76.5°N, 25.0°E, southeast of Svalbard, Norway. From another ERS-1 SAR image a nonlinear wave disturbance with an average wavelength of 2.3 km was identified in the marine atmosphere 30 km offshore in the German Bight of the North Sea by Alpers and Stille [1996].

In this study we will investigate another kind of lee waves, which is different from the above mentioned ones. The waves were observed on ERS-1 SAR images

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02176.
0148-0227/98/97JC-02176\$09.00

of the Taiwan Strait and characterized by their occurrence being associated with the coastal environment, i.e., the transition region between marine and terrigenous air masses. The waves occurred in the form of wave packets along the lee side of the coast. The wave crest lines were nearly parallel to the coastline and reached 80 km long. According to this close association of its occurrence with the coast, a unique feature, we define this kind of lee wave as a coastal lee wave.

This paper is organized as follows. In section 2 we present observed facts including description of SAR images and interpretation results and simultaneous atmospheric and oceanic conditions. A physical model, which is used to explain how a land breeze circulation can generate the coastal lee waves, is described in section 3. In section 4 we discuss the SAR imaging mechanisms of the coastal lee waves at low wind conditions.

Conclusions and discussion follow in section 5.

2. Observations

2.1. ERS-1 SAR Images of the Taiwan Strait

The patterns of atmospheric waves under investigation were recognized on two sequential ERS-1 SAR images provided by the European Space Agency (ESA). The two images were taken at 0611:00 and 0626:43 UT on December 8, 1994, along an ascending orbit of the ERS-1 satellite. The first image in the south is centered at 24°16'28"N, 119°57'00"E, and the second image in the north is centered at 25°38'27"N, 119°38'41"E. The coverage of the images is a square area with one side 100 km wide. The resolution of 12.5 × 12.5 m of original digital images was reduced to 100 m × 100 m for producing the images we use in this study in order to save the storage space and computing time. We believe that the reduced resolution is still high enough for observing the coastal lee waves with the average wavelength of 3 km.

On the original images the imageries of both land and ocean surfaces are a little vague because of unknown reasons. We analyzed the image and found that the images contained spatially periodical noise distributed over the entire scene. In order to improve the quality of the imagery of the atmospheric waves, in which we are interested, a filter was employed to remove the noise during the image processing. An enhanced image of the second scene is shown in Figure 1. One can see that the major features, particularly alternative dark-bright patterns on the ocean portion of the image, are quite clear. Among the two images we analyzed, this one contains >90% of the atmospheric wave imagery, while the first one in the south contains <10% of the imagery only on the left top corner. In the following sections, therefore, we will use this image as an example.

2.2. Interpretation

Consulting the map, we recognized that the land area shown on the image of Figure 1 is the northern portion

of the coastal region of Fujian Province, China, and the ocean area is the northwest portion of the Taiwan Strait. On the ocean side, there are three groups of the wave-like, alternative dark-bright patterns distributed in a band with an offshore width of 20–40 km. The first packet on the upper right corner of Figure 1 contains six waves with a spatially variable wavelength, the average wavelength of 1.7 km measured along the line marked by A-A' and 2.7 km along line B-B'. The second packet, which is close to the central right, contains 10 waves with the average wavelength of 4.2 km measured along line C-C'. The third packet on the lower left corner contains 17 waves with the average wavelength of 2.0 km measured along line D-D'. The crest length of the waves in the first and second packets is about 20–40 km and reaches 80 km in the third packet as the imagery on the southern image is counted in. The total extent of the distribution of the waves imaged on the two images reaches as long as 150 km along the western side of the Taiwan Strait. An interpretation map is shown in Figure 2.

According to Vachon *et al.* [1994] and Alpers and Strilke [1996] the radar cross section over the ocean in original ERS-1 SAR images with satellite-viewing geometry can be used for retrieving the sea surface wind speed. We calculated the wind speed along lines A-A', B-B', C-C', and D-D' denoted in Figure 1 using the original radar cross-section data and the C band model (CMOD5) [Institut Français de Recherche et d'Exploitation de la Mer, 1993]. The results obtained with 16-point running average (corresponding to a resolution of 200 m) are shown in Figure 3. The wind speed along line A-A' has an average of 2.7 m/s and an amplitude of variability of 0.9 m/s; along B-B', 1.6 m/s with the variability of 0.6 m/s; along C-C', 2.7 m/s with the variability of 1.1 m/s; and along D-D', 2.3 m/s with the variability of 0.5 m/s.

2.3. Atmospheric and Oceanic Conditions

In order to understand the generation mechanisms and the SAR imaging mechanisms of the atmospheric waves shown on the SAR images it is necessary to collect the complementary atmospheric and oceanic observations in as much detail as possible.

Large-scale atmospheric conditions were obtained by analyzing the weather charts. The surface weather charts of east Asia and the northwestern Pacific at 0000 and 0600 UT (Figure 4) on December 8, 1994, showed that the weather situation in the study area was controlled by two major systems. A high-pressure cell (1025 mbar) was located in the southern Japan Sea at 0000 UT, and moving northeastward, a low-pressure center (1000 mbar) was near stationary over the Yunnan-Guizhou region of southwest China. The distance between the isobars with an interval of 5 mbar over the East China Sea and the South China Sea reached about 2000 km, implying a weak pressure gradient. Using the pressure field on the weather chart of

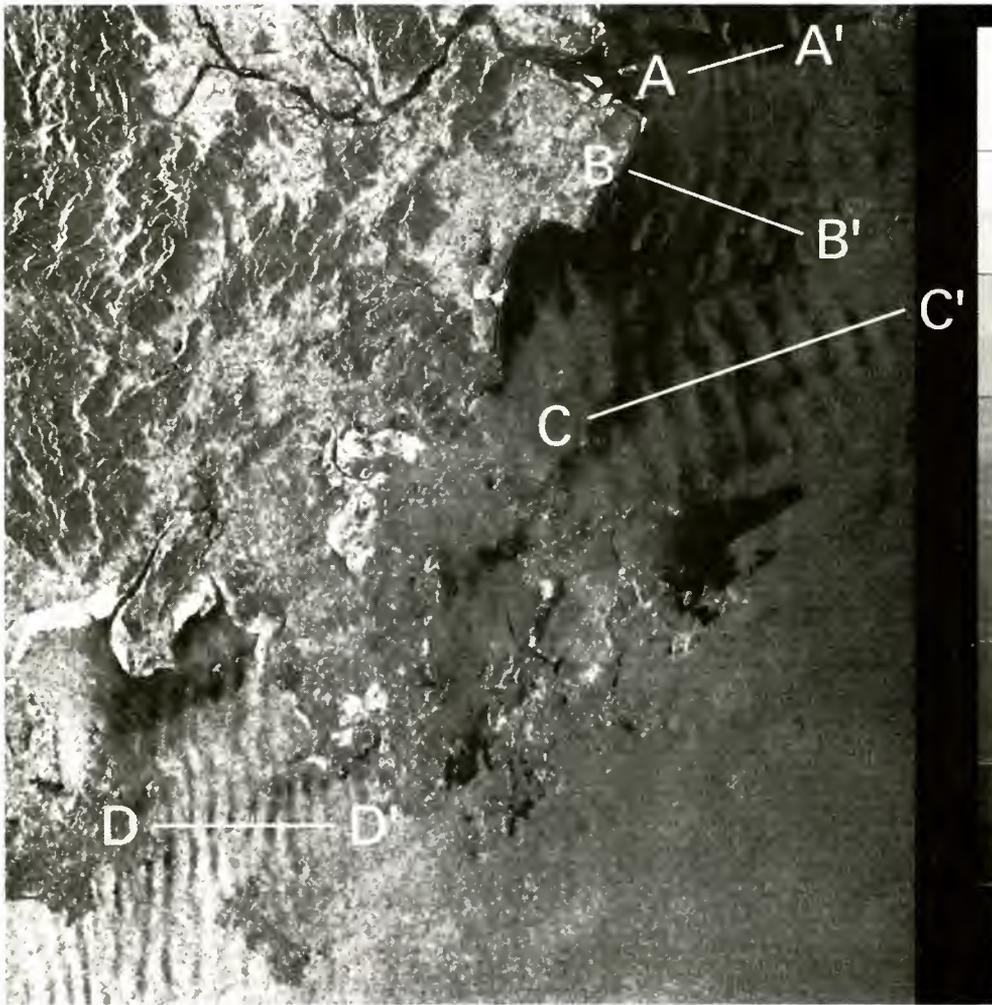


Figure 1. Enhanced ERS-1 synthetic aperture radar (SAR) image of the western side of Taiwan Strait acquired at 0626:43 UT on December 8, 1994. The image covers an area of 100×100 km, centered at $25^{\circ}38'27''\text{N}$, $119^{\circ}38'41''\text{E}$ with a scale of 1:560,000. Alternative dark-bright patterns along the coast represent coastal lee waves (copyright European Space Agency).

Figure 4 and a geostrophic approximation, we calculated the sea surface wind in the open water near the studied area; the result gave a southeasterly wind of 3.6 m/s.

Comparison of these observations of the wind speed to remotely sensed wind speeds given above indicated that they were in a good agreement. Both of them represented a low wind condition. This fact implies that the CMOD5 model works well even at the very low wind conditions.

The weather parameters at Fuzhou, which is the nearest meteorological station to the study area as shown in Figure 2, at 0600 UT on December 8, 1994, which is nearly simultaneous with the imaging time, were as follows: the land surface air temperature, 19°C , the wind speed, 1.5–3.5 m/s; the wind direction, 240° ; the dew-point, 12°C , the cloud type, Cu; the barometric pressure, 1015.8 mbar; and the barometric tendency, -2.0 mbar. The weather in the past 6 hours was intermittent rain, and precipitation in the past 6 hours was 10 mm,

implying that a cold front passed. An infrared (IR) image taken by the Geostationary Meteorologic Satellite (GMS, launched by Japan) at 0546 UT on the same day (National Oceanic and Atmospheric Administration (NOAA) advanced very high resolution radiometer (AVHRR) was unavailable at that time) is shown in Figure 5. One can see that there was a large extent of cumulus over east and south China, the Yellow Sea, the East China Sea, and the northern South China Sea. The Taiwan Strait, the study area, was also covered by the clouds; there was no ocean phenomenon showing up on this image.

The thermal structures of the ocean, including the sea surface temperature (SST) and the vertical temperature distribution, are also important parameters for understanding the studied process. The best way to acquire these data is direct measurement. In our case, however, there are no simultaneous data available. Fortunately, we found some valuable oceanographic data measured in the study area and in some analytical articles. We

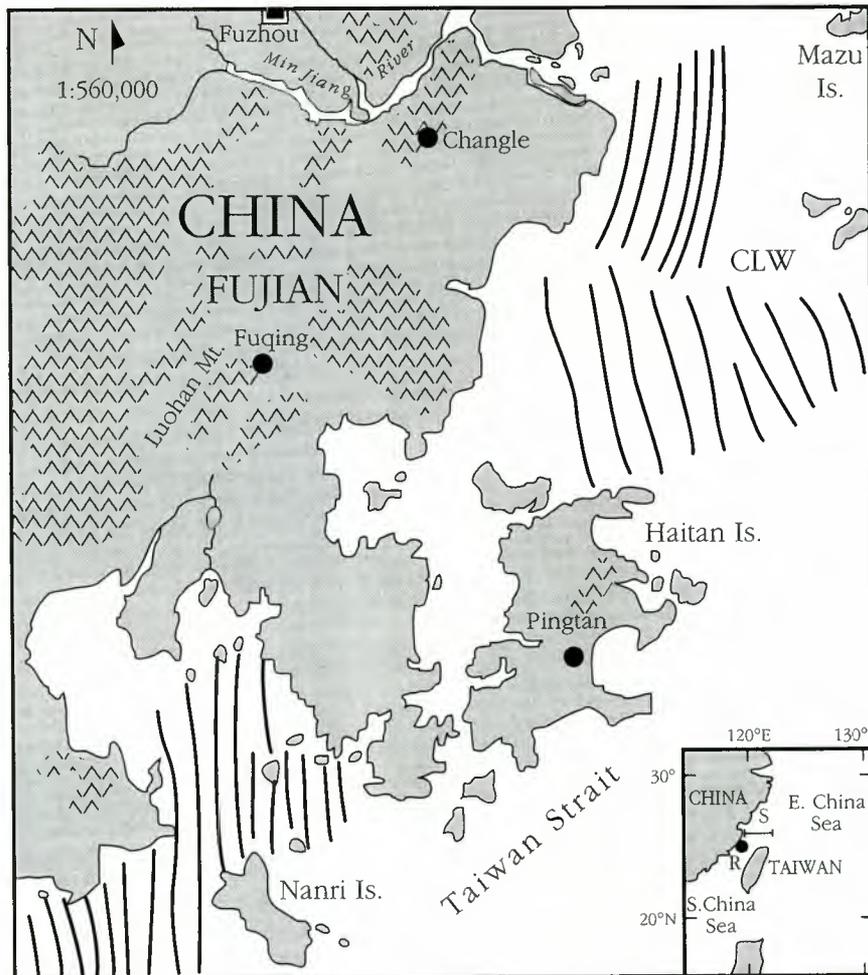


Figure 2. An interpretation map of Figure 1. The location of the center of the imaged area is denoted by R in the inset. Three packets of coastal lee waves are shown as three groups of heavy lines.

used these references and objective analysis method to estimate the thermal structure at the time the SAR images were taken. The imaged area is located at the northern portion of the Taiwan Strait, which serves as a channel for water exchange between the East China Sea and the South China Sea. The Taiwan warm current plays a dominant role in the thermal balance in the Taiwan Strait. This current originates from the South China Sea as a branch of the Kuroshio warm current and flows northward through the Strait all year round. Because of the influence of the Taiwan warm current, the isotherms in the Strait are generally parallel to the coastline [Zheng and Klemas, 1982; Zheng and Xiu, 1992], and the ocean water temperature in the central strait is always higher than that of the surrounding waters. Field measurements made from a cross section along the latitude line of 26°N across the northern outlet of the Strait (see segment S in the small scale inset on the lower right corner of Figure 2) during 1987–1988 showed a seasonal variability of the water temperature [Pan *et al.*, 1990, 1992]. The sea water temperature profile measured during a cruise from December 8 to

10, 1987, is shown in Figure 6 [Pan *et al.*, 1992]. One can see that the SST varied from 21° to 24°C from near coast to the open water. Hence, at the SAR imaging time, the same season as that of the field measurements, 23°C should be a reasonable estimate for the SST in the central Strait. Thus the SST was 4°C higher than the land surface air temperature in our case. Figure 6 also shows a vertically uniform distribution of water temperature from the sea surface to the depth of 60 m, which is near the bottom. This unstratified thermal structure eliminates the possibility for the occurrence of oceanic internal waves.

3. Generation Mechanisms

3.1. Sea and Land Breeze Circulation

It is well documented that the temperature difference between the land and the sea can drive an air circulation traversing the coastline. The circulation which flows landward near the surface is called the sea breeze, and the reversed one is called the land breeze [Simpson,

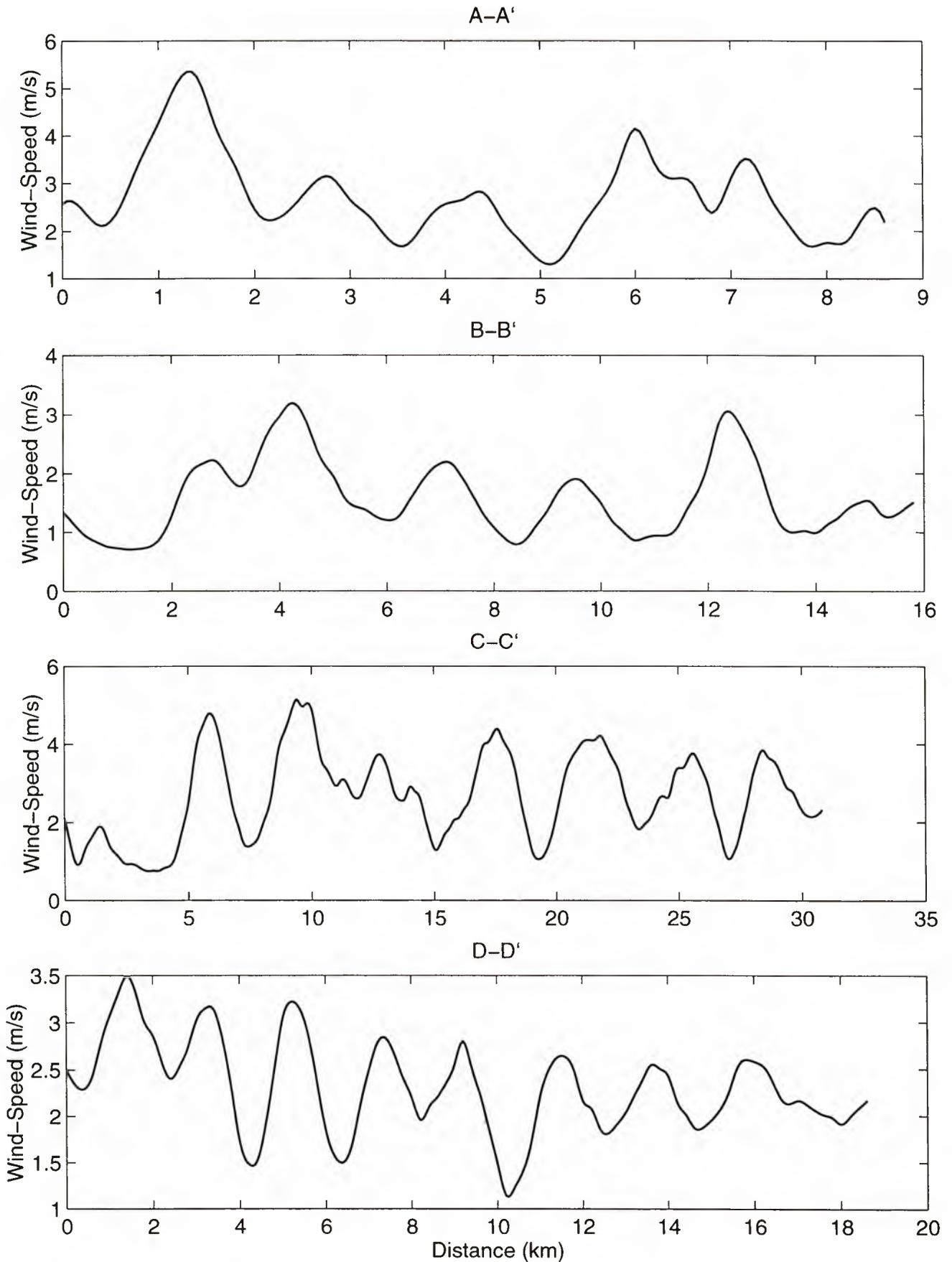


Figure 3. Sea surface wind speeds along the lines A-A', B-B', C-C', and D-D' marked in Figure 1 calculated from original radar section data by using the C band model (CMOD5) [Institut Français de Recherche et d'Exploitation de la Mer, 1993].

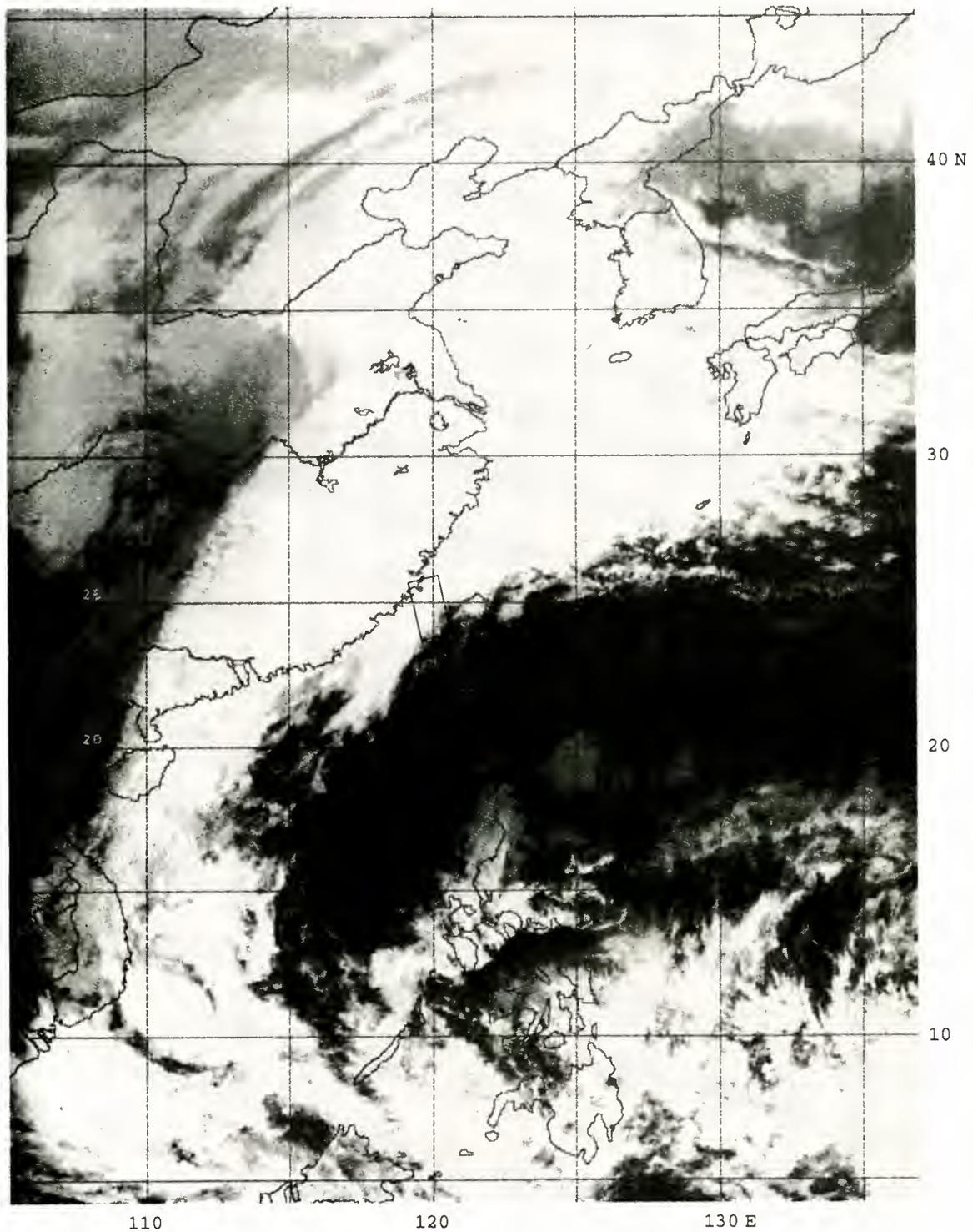


Figure 5. A GMS IR image of east Asia and the northwestern Pacific taken at 0546 UT (1346 LT) on December 8, 1994. The spatial resolution of image is 5 km. The approximate extent of SAR images used for observing the coastal lee wave is shown as a rectangular box in the Taiwan Strait.

air temperature, so that it constitutes a seasonal thermodynamic condition, which is favorable for land breeze establishment.

The cold air intrusion into the warm air driven by the sea and land breeze circulations may cause some interesting phenomena to happen. The convergence of the

winds near the breeze fronts must cause the warm air to rise and to form an air rising zone. The rising air condenses and forms cumulus clouds, which have easily been imaged by cameras on the ground and by sensors on-board satellites [Simpson, 1994]. Meanwhile, the temperature difference between the cold air and warm

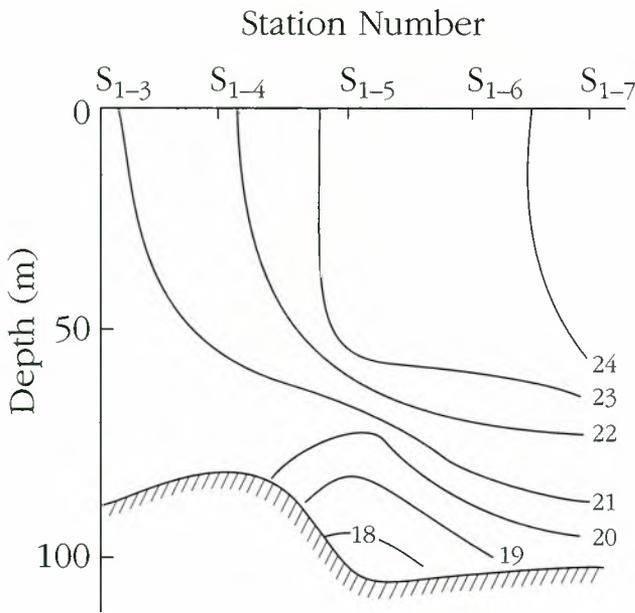


Figure 6. Two-dimensional ocean temperature section measured along 26°N latitude near the northern edge of the SAR imaged area (shown as segment S in the inset of Figure 2) during a cruise from December 8 to 10, 1987. The numbers on the curves are in degree centigrade.

air causes a vertical density stratification. Along the interface of two layers with different densities, wavelike Kelvin-Helmholtz billows may be generated, and the leading edge of the breeze front may be divided into a series of lobes and clefts due to the overrun of the less dense air (see Figure 3.2 of *Simpson* [1994] for details). This lobes-clefts structure may also be caused by the nonuniformity of offshore wind fields along the coast.

The vertical density stratification of the atmosphere caused by sea and land breeze circulations and the disturbance in the air rising zone constitute sufficient conditions for the generation of atmospheric gravity wave. We suggest that the coastal lee wave is generated by the land breeze circulation. The following physical model will dynamically confirm this hypothesis.

3.2. Physical Model

From section 2.3 we know that the land-sea temperature difference was 4°C and the average wind speed was 2.5 m/s at imaging time. On the basis of (1) we obtain the sea breeze index of 1.6, which is much smaller than 10, a critical value of the index. It was a favorable condition for the establishment of the land breeze circulation. In fact, an offshore wind of 1.5–3.5 m/s confirmed that the land breeze had occurred. Meanwhile, in the offshore open water, the wind speed was 3.6 m/s, and the wind direction was southeasterly. This implies that the onshore wind and the offshore wind had a close speed but nearly opposite directions. The two wind systems might collide and generate a front. This process can be described with a physical model schematically illustrated in Figure 7.

This is a two-dimensional model. The x axis is perpendicular to the coastline positive shoreward. The z axis is perpendicular to the sea surface positive upward. The origin is located at the land breeze front. The system is horizontally divided into two regions from the front at $x = 0$. The land breeze and vertical stratification occur only in the region for $x > 0$. Therefore only the dynamical processes in that region will be analyzed hereafter. The upper layer is assumed to have an infinite depth, and the lower layer, the land breeze, has a depth of H . There is a thin transition layer with a thickness of $2\Delta H$ sandwiched between the two layers. The amplitudes of the wind speeds in the two layers are supposed to be the same, u_0 . The wind field in the upper layer is supposed to be vertically and horizontally uniform. The lower layer is characterized by a vertically and horizontally uniform wind speed beyond the frontal zone but is characterized by a variable wind speed within the frontal zone. The horizontal variation of wind speed is assumed to be a hyperbolic tangent function

$$u(x, 0) = -u_0 \tanh \frac{x}{L} \quad x \geq 0 \quad (2)$$

where L is defined as a characteristic length of the front.

The boundary conditions are

$$\begin{aligned} u &= -u_0 & \rho &= \rho_0 & z &< H \\ u &= -u_0 + \beta(z - H) & \rho &= \rho_0 - \Delta\rho/2 & H < z < H + 2\Delta H \\ u &= u_0 & \rho &= \rho_0 - \Delta\rho & z > H + 2\Delta H \end{aligned}$$

where ρ_0 is the air density in the lower layer and $\Delta\rho$ represents the density difference between the lower and upper layers. If we consider the model as a linear system, then the wave equation and continuity equation are [*Gossard and Hooke*, 1975]

$$\frac{\partial^2 w}{\partial z^2} - k^2 w = 0 \quad (3)$$

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0 \quad (4)$$

where u and w represent the horizontal and vertical components of the wind velocity, respectively, and k ($= 2\pi/\lambda$, in which λ is the wavelength) represents the wavenumber. This boundary value problem has been solved by *Gossard* [1974]. The solution in the lower layer is

$$w = \frac{w_H \sinh kz}{\sinh kH} \exp[i(kx - \sigma t)] \quad z < H \quad (5)$$

where the real part is to be taken. The dispersion relation of the waves can be derived from a quartic equation [*Gossard and Hooke*, 1975]

$$\begin{aligned} \left(\frac{C}{u_0}\right)^4 - \left[1 + \frac{(2\alpha - 1)^2 - e^{-4\alpha}}{4\alpha^2} + \frac{R}{\alpha}\right] \left(\frac{C}{u_0}\right)^2 \\ + \left(\frac{1+R}{2\alpha} - 1\right)^2 - \left(\frac{1+R}{2\alpha}\right)^2 e^{-4\alpha} = 0 \end{aligned} \quad (6)$$

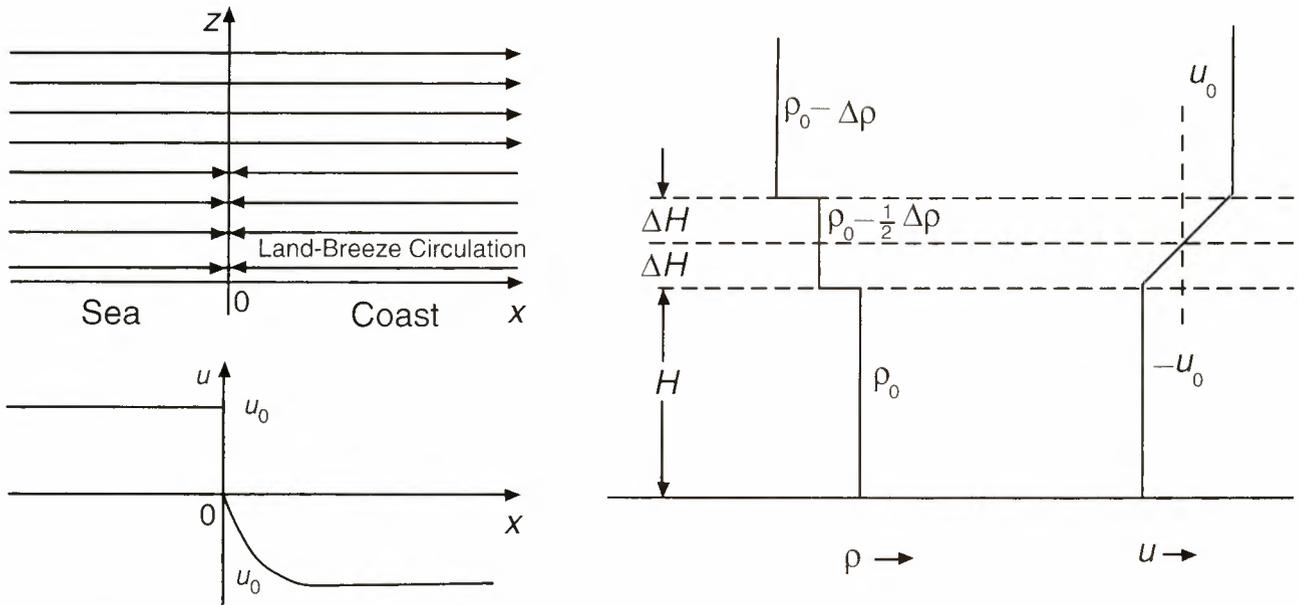


Figure 7. A schematic diagram of a physical model used for analyzing the generation mechanism of the coastal lee wave.

where $\alpha = k\Delta H$ and $R = \Delta\rho g\Delta H/\rho_0 u_0^2$ defined as the Richardson number.

Examining this solution, we find that the existence of a nonzero wave solution depends completely on the existence of a vertical velocity disturbance. In our case the vertical velocity disturbance is generated by the convergence of the horizontal wind within the frontal zone. From (2) and (4) we derive this vertical velocity disturbance as

$$w(x, z) = \frac{u_0}{L} \operatorname{sech}^2\left(\frac{x}{L}\right) z \quad (7)$$

According to our model, this vertical disturbance occurs only in the lower layer. Therefore, the waves are generated in this layer. Thus, a full expression of the waves is

$$w = \frac{u_0 H \sinh kz}{L \sinh kH} \operatorname{sech}^2\left(\frac{x}{L}\right) \exp[i(kx - \sigma t)] \quad (8)$$

This solution is illustrated in Figures 8a and 8b. One can see that the modeled waves are comparable to the observations shown in Figure 3.

3.3. Dynamic Features of Model Waves and Comparison With Observations

Equation (8) reveals the following features of the model waves.

1. The waves are linear waves and are generated by the frontogenesis of the land breeze on the lee side of the coast. This agrees with SAR observations in the Taiwan Strait; the coastal lee waves occurred on the lee

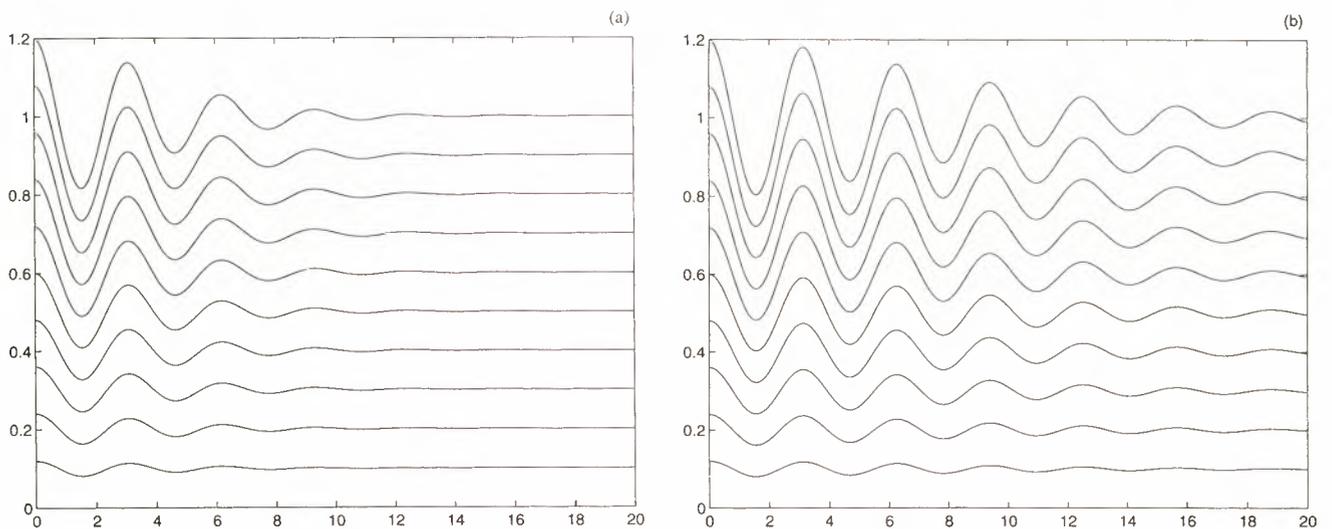


Figure 8. A schematic diagram of the wave solution in the lower layer of the physical model shown in Figure 7. (a) L is chosen as 5 km. (b) L is chosen as 10 km. The amplitude of the waves is in $u_0 H/L \sinh kH$.

side of the coast during land breeze establishment, and the waves appeared to have a good linearity as shown in Figure 3.

2. The waves can only exist within the frontal zone. Beyond the frontal zone the amplitudes of the waves decay quickly, implying that the waves appear in the form of a single solitary packet of linear waves. The number of the waves in the package depends on a ratio between the characteristic length of the frontal zone and the wavelength. The weaker the front (larger L), the more waves. SAR observations showed that coastal lee waves appeared in the form of a single solitary packet with the offshore extent of 20–40 km, which is within the extent of general scales of the land breeze front; there were three separate packets with different numbers of waves and different offshore extent, implying the longshore variation of the wind field and front.

3. From (6) approximate expressions of the dispersion relations of the waves can be derived under a condition $\alpha = k\Delta H \ll 1$. In this case we have four solutions. One group is

$$C_{1,2} = \pm u_0 \sqrt{1 + \frac{R}{\alpha}} - R \quad (9)$$

This group of solutions represents a dispersive wave with two travel directions, i.e., either offshore or onshore. The other group is

$$C_{3,4} = \pm u_0 \sqrt{1 + R} \quad (10)$$

This group of solutions represents a nondispersive wave with the same travel direction possibilities. Furthermore, from (9), we obtained the group velocity

$$C_g = \pm \left(\frac{C^2 + u_0^2}{C} \right) \quad (11)$$

In order to estimate these speeds in our case quantitatively, the parameters R and R/α should be determined from the following observations: the relative density difference, $\Delta\rho/\rho_0 = 0.013$; the horizontal wind speed, $u_0 = 2.5$ m/s; the mean wavelength, $\lambda = 3$ km; and the gravitational acceleration, $g = 9.8$ m/s².

$$\frac{R}{\alpha} = \left(\frac{\Delta\rho}{\rho_0} \right) \frac{g}{ku_0^2} = 9.6 \quad (12)$$

and the Richardson number $R = O(1)$ for $\Delta H = O(100\text{m})$. Therefore we have

$$C_{1,2} = \pm 3.0u_0 \quad (13)$$

$$C_{3,4} = \pm 1.4u_0 \quad (14)$$

$$C_g = \pm 3.3u_0 \quad (15)$$

These solutions indicate that the coastal lee waves are a type of propagating wave, but they are trapped within the frontal zone because of a decrease in wind speed to zero at the front. We cannot measure these speeds

using SAR images only, or snapshot pictures, to verify these results directly. But, we can indeed find some signatures of propagating waves from the images. For example, the shoreward curved crest lines in the first packet in Figure 1 indicate an offshore propagation of the coastal lee waves.

These results were derived under the condition $kH \ll 1$, i.e., $H \ll \lambda$. H cannot be determined only with SAR images or conventional weather observations but can be measured with specially designed experiments. Historically, Meyer [1971] measured the depth of land breeze at Wallops Island, which was a layer of cold air < 100 m deep. In our case, there is no measurement of H available, but we may suppose that our case is similar to the case of Wallops Island and use Meyer's measurement as a scale estimate, i.e., $H = O(100\text{ m})$. We have measured the average wavelength to be 3 km; therefore, the condition $kH \ll 1$ is satisfied.

4. From (4) and (8) an amplitude of horizontal wind fluctuation at the sea surface caused by coastal lee waves can be derived. In order to have a reasonable estimate of this amplitude and to avoid complicated calculations we used a step function, i.e., $S = 1$ for $0 > x > 2L$, and $S = 0$ for $x > 2L$, to approximate the function $\text{sech}^2(x/L)$ in (8) for the integral calculation. Then we have

$$u_a = \frac{Hu_0}{L \sinh kH} \quad (16)$$

Here u_0 can be determined with field observations or calculated using the radar cross-section measurements, k can be calculated using the wavelength of coastal lee waves measured directly from SAR images, L can be estimated using the distribution extent of coastal lee waves, and H satisfies the condition $kH \ll 1$. In this case, (16) is simplified as

$$u_a = \frac{u_0}{2\pi} \left(\frac{\lambda}{L} \right) \quad (17)$$

One can see that actually u_a is independent of H . Furthermore, if we choose $2L$ as the distribution extent of coastal lee waves, we have

$$u_a = \frac{u_0}{N\pi} \quad (18)$$

where N is the number of coastal lee waves in the packet.

In our case, if $u_0 = 2.5$ m/s and $N = 10$, (18) gives u_a of the order of 0.1 m/s. This value is much smaller than values from 0.5 to 1.1 m/s, which were calculated from radar cross section using the CMOD5 model. There are two possibilities to cause this departure. The first one is that our model wind profile may deduce the amplitude of variability of the horizontal wind. The second is that the CMOD5 model may overestimate the wind variability when the wind speed is close to the threshold value to be discussed in section 4. Using more complicated model wind profiles or improving the wind retrieval al-

gorithms may help to fill the gap but are beyond the scope of this study. We believe that the physics of the phenomenon under investigation has been clarified by our analysis, even though the physical model we used seems to be rather simple.

5. From (8), we derived a maximum amplitude of vertical replacement of the coastal lee waves for $z = H$, and $x = 0$. Substituting $\sigma = Ck$, in which $C = \beta u_0$ and β is a constant, and $kL = N\pi$ yields

$$\eta(O, H) = \frac{u_0 H}{\sigma L} \quad (19)$$

$$\eta(O, H) = \frac{H}{N\pi\beta} \quad (20)$$

One can see that the depth of the land breeze, H , is a major factor in deciding the amplitude of coastal lee waves. In our case the amplitude is estimated at 2–10 m, if H is within an extent from 100 to 500 m.

4. SAR Imaging Mechanism Analysis

We have interpreted alternative dark-bright patterns in Figure 1 as the coastal lee waves from the point of view of atmospheric wave dynamics without explaining their SAR imaging mechanisms. However, the mechanisms by which the SAR images these patterns should also be analyzed. In this section we present the results of our analysis.

The microwave pulses of the C band ERS-1 SAR can penetrate the cloud cover, implying that the clouds are transparent. The alternative dark-bright patterns in the image therefore cannot be the patterns of clouds, unlike the imagery in the visible and infrared satellite images. Instead of the clouds, the SAR is sensitive to variabilities of sea surface roughness and physical parameters [Valenzuela, 1978; Plant, 1990; Zheng *et al.*, 1995]. Holliday *et al.* [1986] revealed that under low wind conditions (wind speed < 10 m/s), such as the case we are studying, the Bragg scattering is a dominant mechanism for generating the radar return signals from the sea surface. For the C band ERS-1 SAR the frequency is 5.3 GHz, the wavelength is 5.66 cm, and the average incidence angle is 23° (varies from 20.1° at the near side of swath to 25.8° at the far side). The Bragg-scattering ocean wavelength is 7.2 cm, which is referred to as a gravity-capillary wave. From the dispersion relation

$$\omega^2 = gk + Tk^3/\rho \quad (21)$$

where ω is the angular frequency, g is the gravitational acceleration, k is the wavenumber, T is the surface tension (about 75 dyn/cm² for the sea water), and ρ is the density of sea water, the frequency of the gravity-capillary wave can be calculated. For the wavelength of 7.2 cm the frequency is 4.8 Hz. The gray value of the pixel in the SAR ocean image represents the radar return power from the resolution cell of the ocean sur-

face, which depends on the normalized radar cross section σ_0 . According to Plant [1990], σ_0 is dependent on the wavenumber and the wavenumber spectral density of the Bragg waves, $\Psi(k)$, given by

$$\sigma_0(\theta)_{ij} = 16\pi k^4 |g_{ij}(\theta)|^2 \Psi(0, 2k \sin \theta), \quad (22)$$

where the indices ij denote the polarizations of the incident and backscattered radiation, respectively, and $g_{ij}(\theta)$ are the first-order scattering coefficients.

However, since ocean wave spectra are usually given in the form of a frequency spectrum [Huang *et al.* 1990], the wavenumber spectrum should be transformed into the frequency domain if (22) is used for calculating the normalized radar cross section. This transformation can be performed using the dispersion relation (21) and definitions for the wavenumber spectrum

$$\Psi(K) = \int_n X(k, n) dn \quad (23)$$

and for the frequency spectrum

$$\Phi(n) = 2 \int_k X(k, n) dk \quad (24)$$

where $X(k, n)$ is the energy spectrum and n is the frequency. Hence we have

$$\sigma_0(n)_{ij} = F_1(n) |g_{ij}|^2 \Phi(n) \quad (25)$$

where $F_1(n)$ is a function of the frequency and $\Phi(n)$ is a function of wind speed and other parameters. The alternative dark-bright patterns observed therefore reflect spatially periodical distributions of the 4.8 Hz ocean waves. These gravity-capillary waves are generated by the local wind and the wave-wave interaction. In our case the sea surface wind speed was estimated within a range of 2.5–5 m/s, which is corresponding to the frictional velocity u_* being within a range of 10 cm/s $< u_* < 20$ cm/s. According to the results of Long and Huang [1976], this range of wind is extremely important for wind wave generation and growth. Their laboratory measurements revealed two facts, which are related to the phenomena observed in the SAR image. The first one is that the sea surface roughness represented by the mean squared slope of the ocean surface, $(\nabla\zeta)^2$, which is defined as an integrated quantity of the slope spectrum

$$(\nabla\zeta)^2 = \int_k S(k) dk \quad (26)$$

where $S(k)$ is the slope spectrum, starts occurring at the frictional velocity of about 15 cm/s (corresponding to the sea surface wind of 3.5 m/s). This implies that there are no wind-driven waves generated below this threshold wind speed, and the sea surface roughness increases drastically around 15 cm/s $< u_* < 20$ cm/s, then slowly increases beyond $u_* = 20$ cm/s as shown in Figure 9a. The second fact is that the slope spec-

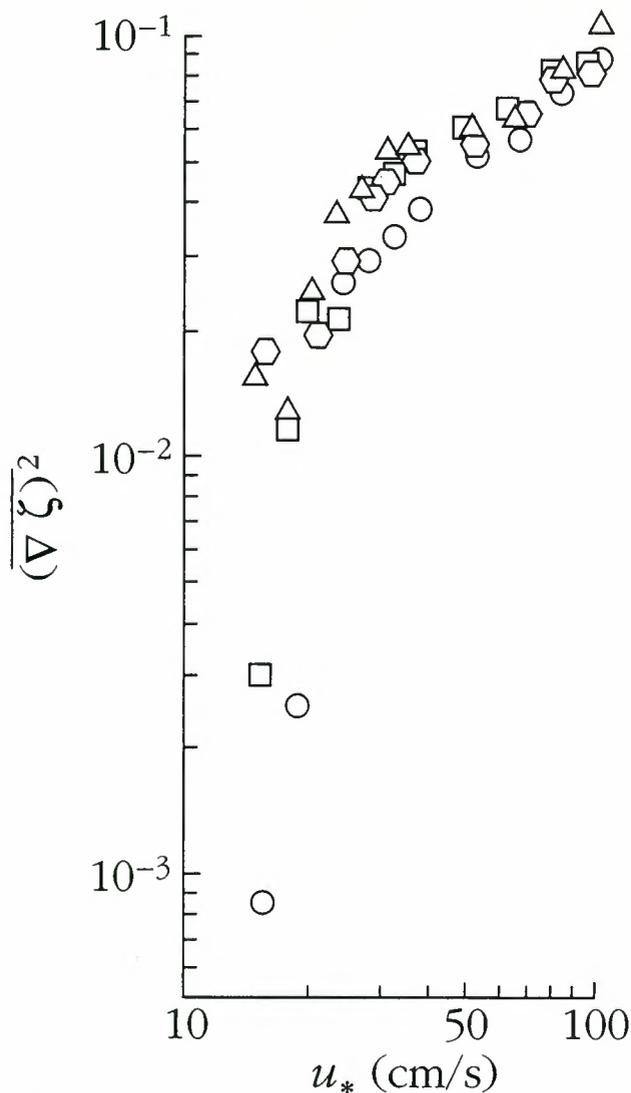


Figure 9a. The growth of sea surface roughness versus the friction velocity (the data were taken from Figure 9 of *Long and Huang, [1976]*). There is a threshold velocity at about 15 cm/s corresponding to the sea surface wind of 3.5 m/s.

trum density of gravity-capillary waves, $S_{11}(n)$, which is defined as

$$S_{11} = F_2(\mu) \left(\frac{g^3}{Tn^4} \right)^{\frac{1}{4}M(\mu)} n^{-1} \quad (27)$$

where $F_2(\mu)$ is a dimensional coefficient function for a nondimensional parameter

$$\mu = (Tg)^{\frac{1}{4}} / u_* \quad (28)$$

and M is a parameter to be determined, increases dramatically with the frictional velocity as shown in Figure 9b. One can see a situation similar to that of sea surface roughness. The slope spectrum density has a sharp jump around 15 cm/s $< u_* < 20$ cm/s.

The sensitivity of the slope spectrum density to wind decreases dramatically beyond $u_* = 30$ cm/s (8 m/s).

Quantitatively, a 1.5 m/s change in the wind speed from 3.5 to 5 m/s (or 16 cm/s $< u_* < 20$ cm/s) results in a 6.5 times increase in the slope spectrum density (from 2×10^{-3} to 1.3×10^{-2}). These results imply that a small change in the wind speed within the range from 3.5 to 5 m/s will result in a big change in σ_0 because of the equivalence of S_{11} to $\Phi(n)$ in (25). Namely, there is an "amplification effect" of the wind-wave response to the variability of the low winds within the range from 3.5 to 5 m/s and 3.5 m/s once again emerges to be the threshold wind speed, below which insignificant radar return will be produced. These mechanisms can be used to explain the ERS-1 SAR image shown in Figure 1. The plume-like imagery covering the ocean surface was produced by the 4.8 Hz ocean waves. The alternative dark-bright patterns were originally generated by a periodical modulation process applied to these waves as illustrated in Figure 10. According to the geographic distribution and characteristics of the patterns we concluded that the modulation process is the coastal lee wave, in other words, the alternative dark-bright patterns are "footprints" of the coastal lee wave. The dark areas look like the shadows of the coast and the islands represent the calm ocean surface, where the wind speeds were far below the threshold wind speed of 3.5 m/s.

Above mentioned measurements of *Long and Huang [1976]* explain the SAR imaging mechanisms of the coastal lee wave very well, though their laboratory data were measured under a neutral condition. In our case the air transported by the land breeze from the land to the sea had a temperature of 4°C lower than the sea surface temperature, implying an unstable condition.

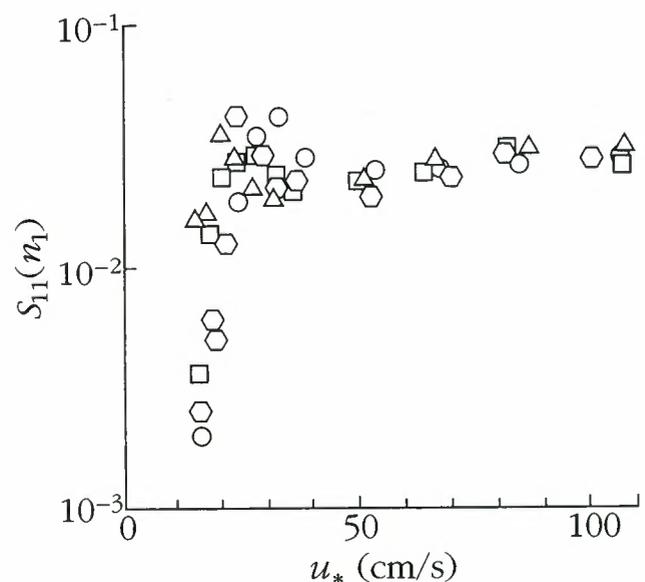


Figure 9b. The growth of slope spectral components $S_{11}(n)$ for gravity-capillary waves with a frequency of 5 Hz (the data were taken from Figure 10 of *Long and Huang, [1976]*). There is a surprising jump from 15 to 20 cm/s for the friction velocity corresponding to sea surface winds from 3.5 to 5 m/s.

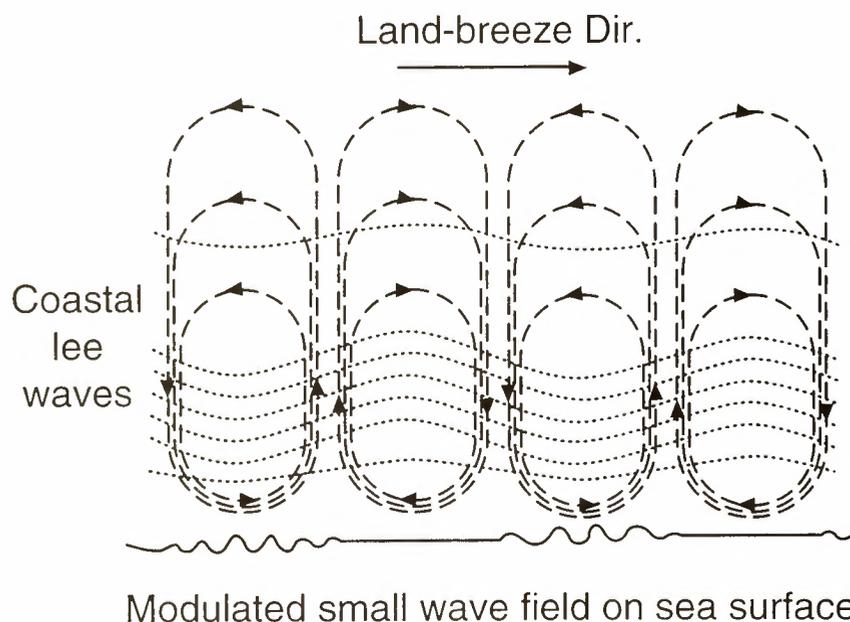


Figure 10. A schematic diagram of sea surface small wave modulation induced by coastal lee waves.

Under this unstable condition the dependence of the sea surface radar cross section on the wind speed may vary. In fact, *Wu* [1991] has formulated this variation with an empirical model

$$\sigma_u = \sigma_n e^{-0.30\Delta T/\bar{U}_{10}} \quad (29)$$

where σ_u represents sea surface radar cross sections under the unstable condition and σ_n represents sea surface radar cross sections under the neutral condition, respectively, ΔT is the air-sea temperature difference, and \bar{U}_{10} is the wind speed measured at a height of 10 m above the sea level. Thus we derived the relative variation of σ_u with respect to \bar{U}_{10} .

$$\frac{\Delta\sigma_u}{\sigma_u} = \frac{0.30\Delta T}{\bar{U}_{10}} \left(\frac{\Delta\bar{U}_{10}}{\bar{U}_{10}} \right) + \frac{\Delta\sigma_n}{\sigma_n} \quad (30)$$

One can see that the relative variation of the sea surface radar cross section under the unstable condition is larger than that under the neutral condition, and the increase is formulated by the first term of the right hand side of (30). In our case, $\Delta T = 4^\circ\text{C}$, the relative radar cross section will increase an additional 13% more than that of the neutral condition as the wind speed U_{10} increases from 2.5 to 3.5 m/s. In other words, the radar return from the sea surface is more sensitive to the small change in the wind speed, and the above mentioned amplification effect will be further enhanced.

5. Conclusions and Discussion

This paper deals with coastal lee waves, which are defined as a kind of atmospheric gravity wave occurring along the lee side of the coast. The case under investi-

gation was observed from space by the ERS-1 SAR. The coastal lee waves exhibit the following general characteristics. (1) The average wavelength is 3 km ranging from 1.7 to 4.2 km. The crest lines, with lengths from 20 to 80 km, are generally parallel to the coastline. The horizontal distribution range is a band 20–40 km wide from near the coast to 40 km offshore. The vertical disturbance extent reaches from the sea surface to a height of 100 m. (2) The physical model we developed shows that the land breeze circulation, which is driven by the land-sea temperature difference, is a major mechanism for generating coastal lee waves. The offshore distribution and the spatial structure of coastal lee waves in the observed case coincide with features of the land breeze described by *Simpson* [1994]. The observed case occurred in the middle latitudes (around 25°N latitude) during the local early winter. The land-sea temperature difference reaches the maximum value during that season and constitutes favorable conditions for generating the seasonal land breeze, providing further evidence for the proposed generation mechanism. (3) The coastal lee wave is a type of propagating wave, but the phase speed gradually tends toward null at the land breeze front. Thus the coastal lee wave should never propagate far from the coast, and it appears to be “trapped” by the coast.

The reason why ERS-1 SAR can image this type of atmospheric wave was examined. On the basis of the results by previous investigators [e.g., *Long and Huang*, 1976; *Wu*, 1991] the threshold and the amplification effect of the gravity-capillary wave response to the variability of low winds are considered as major mechanisms. A 1.5 m/s change in the wind speed from 3.5 to 5 m/s results in a 6.5 times increase in the slope spec-

trum density. This means that the normalized radar cross section (NRCS) will surprisingly increase 8 dB under the neutral condition and 9 dB under the unstable condition. In other words, at low wind conditions, the SAR is extremely sensitive to the variability of wind speed and ocean phenomena caused by wind modulation processes, even though the modulation is traditionally considered to be quite weak. In our case it is the coastal lee wave in the lower atmosphere which modulates the sea surface wind stress and in turn produces the wavelike dark-bright patterns on ERS-1 SAR images. We have noted that the nonlinear wave disturbance in the North Sea observed by *Alpers and Stilke* [1996] was also imaged by the ERS-1 SAR at low sea surface wind speeds between 1 and 5 m/s. *Johannessen et al.* [1995, 1996] have reached a similar conclusion, i.e., that low sea surface winds between 3 and 10 m/s are favorable for surface current feature detection by the SAR. These cases indicate that studies of sea state and air-sea coupling at low wind conditions are significant for SAR image interpretation and application.

In addition, we would like to emphasize that in order to avoid misinterpretation it is necessary to distinguish coastal lee waves from ocean internal waves, which have similar wavelike patterns on SAR images (see SAR images shown in *Alpers and Salusti*, [1983]; *Apel and Gonzales* [1983]; *Fu and Holt* [1984]). In fact, it is not difficult to identify them. Physically, they are generated by different mechanisms. As analyzed above, the coastal lee wave is a type of atmospheric wave generated by the land breeze circulation, which is caused by the temperature difference between the land and the sea. The ocean internal wave is a kind of oceanic wave generated by the internal tides in the stratified upper oceans. These differences in nature define their different behaviors. Coastal lee waves may appear only near the coast, but internal waves may appear anywhere in the stratified upper oceans. Coastal lee waves may appear only in the mornings of summer, when the land breeze forms, or during the local early winter, when the strongly seasonal land breeze is established, but the internal waves may appear at most times of the year in the middle latitudes. Statistically, they have different characteristic scales. The average wavelength of the coastal lee wave is 3 km, and the crest lines can reach more than 80 km as we observed. Accordingly, the average wavelength of internal waves in the shallow coastal ocean is 0.5–1 km [*Apel*, 1981; *Fu and Holt*, 1982; *Alpers and Salusti*, 1983; *Howell and Brown*, 1985; *Zheng et al.*, 1989, 1993], and few crest lines in a packet reach 50 km [*Zheng et al.*, 1993]. Our experience in SAR image interpretation also indicates that the imagery of coastal lee waves on SAR images is much broader than that of internal waves, and the wavelengths of coastal lee waves in a packet change little but that of internal waves in a packet decrease quickly from the front to the rear. In our case the possibility of interpreting the wavelike patterns in Figure 1 to be ocean internal waves can be

completely excluded because the nonstratified thermal structure in the imaged coastal ocean, as shown in Figure 6, excludes the possibility of the existence or generation of ocean internal waves. Furthermore, coastal lee waves have also been observed in the Landsat images and space shuttle photographs as visible images of clouds (not included in this paper), further confirming the coastal lee wave as a newly found atmospheric wave. Interpreting these visible images and revealing more detailed features of the coastal lee wave will be our future effort.

Acknowledgments. The authors would like to express their thanks to Zhouwen Yu and Runheng Huang for providing weather maps and GMS satellite images, to Norden E. Huang for valuable discussions on small wave spectra, and to Harald Frick for assistance in data processing. The work is supported by the ONR Physical Oceanography and Environmental Optics Program and Physical Oceanography Program under grants N00014-95-1-0065 and N00014-97-1-0648 (VK, ZW, and QZ), partly supported by the NOAA Sea Grant Program through grants NA16RG0162-03 and NA56RG0147 (XY and QZ), the Chinese National Science Foundation (QZ), and the grant Taiwan NSC 85-NSPO(A)-OCI-09-02 (CH and NK). We appreciate the comments of two anonymous reviewers, who helped generate this final version of our paper.

References

- Alpers, W., and E. Salusti, Scylla and Charybdis observed from space, *J. Geophys. Res.*, *88*, 1800–1808, 1983.
- Alpers, W., and G. Stilke, Observation of a nonlinear wave disturbance in the marine atmosphere by the synthetic aperture radar aboard the ERS 1 satellite, *J. Geophys. Res.*, *101*, 6513–6525, 1996.
- Apel, J. R., Nonlinear features of internal waves as derived from the Seasat imaging radar, in *Oceanography From Space*, edited by J. F. R. Gower, pp. 525–533, Plenum, New York, 1981.
- Apel, J. R., and F. I. Gonzales, Nonlinear features of internal waves off Baja California as observed from the Seasat imaging radar, *J. Geophys. Res.*, *88*, 4459–4466, 1983.
- Fu, L.-L., and B. Holt, Seasat views oceans and sea ice with synthetic aperture radar, *JPL Publ. 81-120 26-41*, pp. 525–533, 1982.
- Fu, L.-L., and B. Holt, Internal waves in the Gulf of California: Observations from a spaceborne radar, *J. Geophys. Res.*, *89*, 2053–2060, 1984.
- Gossard, E. E., Dynamic stability of an isentropic shear layer in a statically stable medium, *J. Atmos. Sci.*, *31*, 488–492, 1974.
- Gossard, E. E., and W. H. Hooke, *Waves in the Atmosphere*, pp. 1–46, Elsevier Sci., New York, 1975.
- Holliday, D., G. St-Cyr, and N. E. Woods, 1986. A radar ocean imaging model for small to moderate incidence angles, *Int. J. Remote Sens.*, *7*, 1809–1834, 1986.
- Howell, T. L., and W. S. Brown, Nonlinear internal waves on the California continental shelf, *J. Geophys. Res.*, *90*, 7256–7264, 1985.
- Huang, N. E., C.-C. Tung, and S. R. Long, 1990, Wave spectra, in *Ocean Engineering Science, The Sea*, vol. 9, part A, edited by B. L. Mehaute and D. M. Hanes, pp. 197–238, John Wiley, New York, 1990.
- Institut Français de Recherche et d'Exploitation de la Mer

- (IFREMER), ERS 1 off-line wind scatterometer products, *IFREMER Tech. Rep., ERS-SCAT/IOA/DOS-01*, Brest, France, 1993.
- Johannessen, J. A., R. A. Shuchman, and O. M. Johannessen, Mesoscale variability with ERS-1 SAR, in *Oceanographic Application of Remote Sensing*, edited by M. Ikeda and F. Dobson, pp. 27–44, CRC Press, Boca Raton, Fla., 1995.
- Johannessen, J. A., R. A. Shuchman, G. Digranes, D. R. Lyzenga, C. Wackerman, O. M. Johannessen, and P. W. Vachon, Coastal ocean fronts and eddies imaged with ERS-1 synthetic aperture radar, *J. Geophys. Res.*, **101**, 6651–6667, 1996.
- Lewellen, W. S. and M. E. Teske, A second-order closure model of turbulent transport in the coastal planetary boundary layer, in *Conference on Coastal Meteorology*, pp. 118–123, Am. Meteorol. Soc., Boston, Mass., 1976.
- Long, S. R., and N. E. Huang, On the variation and growth of wave-slope spectra in the capillary-gravity range with increasing wind, *J. Fluid Mech.*, **67**, 209–228, 1976.
- Meyer, J. H., Radar observations of land breeze fronts, *J. Appl. Meteorol.*, **10**, 1224–1232, 1971.
- Pan, Y., J. Su, and D. Xu, Oceanographic features on the northern shelf of Taiwan in summer and winter, *Paper Collection of Kuroshio Investigation II*, pp. 126–135, Ocean Press, Beijing, 1990.
- Pan, Y., J. Su, and D. Xu, Seasonal hydrographic features of the bottom water on the southern shelf in the east China Sea between 50–200 m isobath, in *Paper Collection of Kuroshio Investigation IV*, pp. 211–219, Ocean Press, Beijing, 1992.
- Panchev, S., *Dynamic Meteorology*, pp. 1–360, D. Reidel, Norwell, Mass., 1985.
- Plant, W. J., Bragg scattering of electromagnetic waves from the air/sea interface, in *Surface Waves and Fluxes, vol. 2, Remote Sensing*, edited by G. L. Geernaert and W. J. Plant, pp. 41–108, Kluwer Acad., Norwell, Mass., 1990.
- Simpson, J. E., *Sea Breeze and Local Winds*, pp. 1–234, Cambridge Univ. Press, New York, 1994.
- Tingle, A. G., and D. A. Dieterle, A numerical study of pollutant distributions in sea breeze circulations, in *Conference on Coastal Meteorology*, pp. 114–117, Am. Meteorol. Soc., Boston, Mass., 1976.
- Vachon, P. W., O. M. Johannessen, and J. A. Johannessen, An ERS 1 synthetic aperture radar image of atmospheric lee waves, *J. Geophys. Res.*, **99**, 22,483–22,490, 1994.
- Valenzuela, G. R., Theories for the interaction of electromagnetic and ocean waves: A review, *Boundary Layer Meteorol.*, **13**, 61–85, 1978.
- Waters, A. J., Radar image: 19 October 1990 at 0600 UTC, *Meteorol. Mag.*, **119**, 272, 1990.
- Wu, J., Effects of atmospheric stability on ocean ripples: A comparison between optical and microwave measurements, *J. Geophys. Res.*, **96**, 7265–7269, 1991.
- Zheng, Q., and V. Klemas, Determination of winter temperature patterns, fronts, and surface currents in the Yellow Sea and East China Sea from satellite imagery, *Rem. Sens. Environ.*, **12**, 201–218, 1982.
- Zheng, Q., and S. Xiu, Studies on satellite remote sensing information of the Kuroshio system, I, Meso-scale dynamics of the Kuroshio of the East China Sea in spring, in *Paper Collection of Kuroshio Investigation IV*, pp. 79–85, Ocean Press, Beijing, 1992.
- Zheng, Q., L. Wu, D. Zhang, and X. Li, A study on SLAR image signatures of internal waves in the sea area nearby the Laoshan Bay, *Oceanol. Limnol. Sin.*, **20**, 281–287, 1989.
- Zheng, Q., X.-H. Yan, and V. Klemas, Statistical and dynamical analysis of internal waves on the continental shelf of the Middle Atlantic Bight from space shuttle photographs, *J. Geophys. Res.*, **98**, 8495–8504, 1993.
- Zheng, Q., X.-H. Yan, V. Klemas, N. E. Huang, J. Pan, and Y. Yuan, Laboratory measurements of the effects of viscosity on short water wave spectra and implication for radar remote sensing of the ocean surface, *Chin. J. Oceanol. Limnol.* **13**, 193–205, 1995.
- V. Klemas, X.-H. Yan, and Q. Zheng, Graduate College of Marine Studies, University of Delaware, Newark, DE 19716. (e-mail: zheng@mingming.cms.udel.edu; xiaohai@mingming.cms.udel.edu; klemas@mingming.cms.udel.edu)
- C.-R. Ho and N.-J. Kuo, Department of Oceanography, National Taiwan Ocean University, Keelung 2024, Taiwan.
- Z. Wang, Texaco, 3901 Briarpark, WP-517, Houston, TX 77042.

(Received November 6, 1996; revised June 9, 1997; accepted July 30, 1997.)

Evolution of nonlinear internal waves in the East and South China Seas

Antony K. Liu and Y. Steve Chang

Oceans and Ice Branch, NASA Goddard Space Flight Center, Greenbelt, Maryland

Ming-K. Hsu

Department of Oceanography, National Taiwan Ocean University, Keelung

Nai K. Liang

Institute of Oceanography, National Taiwan University, Taipei

Abstract. Synthetic Aperture Radar (SAR) images from ERS-1 have been used to study the characteristics of internal waves northeast and south of Taiwan in the East China Sea, and east of Hainan Island in the South China Sea. Rank-ordered packets of internal solitons propagating shoreward from the edge of the continental shelf were observed in the SAR images. On the basis of the assumption of a semidiurnal tidal origin, the wave speed can be estimated and is consistent with the internal wave theory. By using the SAR images and hydrographic data, internal waves of elevation have been identified in shallow water by a thicker mixed layer as compared with the bottom layer on the continental shelf. The generation mechanism includes the influences of the tide and the Kuroshio intrusion across the continental shelf for the formations of elevation internal waves. The effects of water depth on the evolution of solitons and wave packets are modeled by the nonlinear Kortweg-deVries (KdV) type equation and linked to satellite image observations. The numerical calculations of internal wave evolution on the continental shelf have been performed and compared with the SAR observations. For a case of depression waves in deep water, the solitons first disintegrate into dispersive wave trains and then evolve to a packet of elevation waves in the shallow water area after they pass through a “turning point” of approximately equal layer depths that has been observed in the SAR image and simulated by the numerical model. The importance of the dissipation effect in the coastal area is also discussed and demonstrated.

1. Introduction

The tidal flow over topographic features such as a sill or continental shelf in a stratified ocean can produce nonlinear internal waves of tidal frequency and has been studied by many investigators [Sandstrom and Elliott, 1984; Apel *et al.*, 1985; Apel, 1995]. Their observations provide insight into the soliton generation process and explain the role they play in the transfer of energy from tides to ocean mixing. However, almost all of the nonlinear internal waves observed in nature previously were mode one depression waves. Salusti *et al.* [1989] first observed two moving internal wave packets consisting of elevation waves and depression waves separated by a 12-hour period, using a thermistor chain in the eastern Mediterranean Sea during a pilot experiment in the Rio-Antirio strait, western Greece, in July 1986. The change of polarity in internal waves is caused by the change of the mean thermocline depth. Similar change of internal wave polarity was observed by A. N. Serebryany (private communication, 1995) at the shelf of the Sea of Japan in September 1982. These nonlinear internal waves were apparently generated by internal turbulent mixing or baroclinic shear instability over bottom features.

The East China Sea is rich in resources, which have been

exploited extensively. The Kuroshio, the major western boundary current of the Pacific, forms the eastern boundary of the East China Sea as it skirts the shelf edge in the Okinawa Trough. The Kuroshio Edge Exchange Process (KEEP) project studies a major site for the exchange of material between the East China Sea and the Kuroshio at a permanent upwelling region northeast of Taiwan [Liu *et al.*, 1992]. The upwelling is induced by the intrusion of the Kuroshio across the continental shelf [Hsueh *et al.*, 1993]. The Kuroshio fronts and cold eddies in the upwelling region have been observed by advanced very high resolution radiometer (AVHRR) images [Liu *et al.*, 1992]. KEEP-II, a 5-year field program started in 1994, has made direct observations northeast of Taiwan where the Kuroshio collides with the continental shelf. This unique data set may provide some ground truth for the internal wave study in this area. However, owing to the low sampling rate, the data collected during KEEP did not show any high-frequency internal waves. Further collections of internal wave data with a higher sampling rate (every half minute) have been planned and are under way. Internal wave packets were also observed by the crew of the space shuttle *Challenger* as they orbited over Hainan Island in the South China Sea. The Kuroshio branches out near the south tip of Taiwan, and part of the Kuroshio intrudes into the South China Sea through the Bashi Channel and the Luzon Strait. The internal tides and internal waves are probably generated by the shallow ridges in the Luzon Strait.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC01918.
0148-0227/98/97JC-01918\$09.00

It has been known for over 2 decades that internal waves have surface signatures recognizable in satellite images of sea surface [Fu and Holt, 1982]. The synthetic aperture radar (SAR) images from the First European Remote Sensing Satellite (ERS-1) have been used to study the characteristics of internal waves northeast of Taiwan [Liang et al., 1995] and in the Strait of Gibraltar [Brandt et al., 1996]. Rank-ordered packets of internal solitons propagating shoreward from the edge of the continental shelf were observed in many SAR images. By using the SAR images and hydrographic data, internal waves of elevation can be identified from a thicker mixed layer as compared with the bottom layer on the continental shelf. The effects of water depth on the parameters of solitons and wave packets can be linked to the observations from SAR images.

A solitary wave theory that describes the evolution of nonlinear internal waves has been developed and expanded to include effects of vertical shear, variable bottom topography, radial spreading, and dissipation by Liu et al. [1985] for the Sulu Sea internal soliton study [Apel et al., 1985]. Internal solitary waves on a shelf with shoaling effects have been studied by Liu [1988] in the New York Bight. Another mechanism of internal solitary wave attenuation is caused by shear-induced dissipation, which is associated with turbulent mixing and widening of the pycnocline [Bogucki and Garrett, 1993]. All these mechanisms of wave evolution can be simulated numerically by solving the nonlinear Kortweg-deVries (KdV) type equation with varying coefficients corresponded to the changing environments as demonstrated by Liu et al. [1985] and Liu [1988]. For the case of depression waves, the disintegration of solitons into dispersive wave packets after they pass through a "turning point" of approximately equal layer depths (critical depth) has been studied numerically by Helfrich et al. [1984].

In this paper, the internal wave evolution at northeast and south of Taiwan in the East China Sea and east of Hainan Island in the South China Sea is studied on the basis of the ERS-1 SAR data for addressing a wide range of processes. The nonlinear wave evolution on a shelf is formulated by a KdV type equation to include changing depth and dissipation effects in the next section. The selective sets of SAR images in the East and South China Seas are presented in section 3. Section 4 presents the numerical study of the evolution of nonlinear depression waves through the critical depth and the disintegration of solitons into internal wave packets with variable bottom topography and dissipation. Finally, this nonlinear internal wave analysis of the ERS-1 SAR data is summarized and discussed in section 5.

2. Model of Nonlinear Internal Wave Evolution on a Shelf

The evolution of nonlinear internal wave trains on a continental shelf has been formulated by Liu [1988]. The dissipation effects on solitary wave evolution are considered to be important in the shallow water owing to internal wave breaking and strong turbulent mixing. The evolution equation of wave amplitude $A(x, t)$ with variable coefficients is

$$A_x + C_0 A_x + \alpha A A_x + \kappa A^2 A_x + \beta A_{xxx} + \gamma A - \frac{1}{2} \varepsilon A_{xx} = 0, \quad (1)$$

where the parameters α , κ , β , γ , and ε are the coefficients for the nonlinear, higher-order nonlinear, dispersion, shoaling, and dissipation effects.

For a two-layer system with H_1 and H_2 , which are mixed

layer and bottom layer thickness, the nonlinear and dispersion coefficients are

$$\alpha = \frac{3}{2} \frac{H_1 - H_2}{H_1 H_2} C_0, \quad \beta = \frac{1}{6} H_1 H_2 C_0, \quad (2)$$

and the linear wave speed is given by

$$C_0 = [\Delta \rho g H_1 H_2 / \rho (H_1 + H_2)]^{1/2} \quad (3)$$

where g is the gravity constant, ρ is the density of water, and $\Delta \rho$ is the density difference between two layers. The evolution of solitons is based upon the balance of nonlinear effects with the dispersive effect. Note that α changes sign when the water depth is across the critical depth, where $H_1 = H_2$ (turning point). The sign of the nonlinear term depends on the wave amplitude; depression wave has the opposite sign of elevation wave in amplitude. When the mixed layer depth is thinner than the bottom layer, $H_1 < H_2$, only depression waves can be evolved. While the mixed layer is thicker than the bottom layer, $H_1 > H_2$, only the elevation waves can be evolved. Near the critical depth, the higher-order nonlinear coefficient κ is important and is given by [Helfrich et al., 1984]

$$\kappa = -3 \frac{H_1^2 - H_1 H_2 + H_2^2}{(H_1 H_2)^2} C_0. \quad (4)$$

The coefficient for the shoaling effect, γ , is in the order of $-2 \times 10^{-5} \text{ s}^{-1}$ [Liu, 1988]. In this study we will neglect the shoaling effect, since it is small and can be compensated by the dissipation effect. Liu et al. [1985] and Liu [1988] reported an effective horizontal eddy viscosity for solitons of $\varepsilon = 1\text{--}10 \text{ m}^2/\text{s}$. Because eddy viscosity is not a property of a fluid, its value may vary with location and water depth. It is possible that local, incipient shear flow instability or wave breaking could be a cause leading to an eddy viscosity of such value. For shallow water the bottom friction and induced mixing can be another dissipation mechanism.

A numerical approach using Fornberg's pseudo-spectral method [Liu et al., 1985] has been developed to solve the evolution equation (1). A fast Fourier transform (FFT) algorithm is used in the spatial coordinate, and the split-step method is used for time derivatives. The period and mesh size have to be chosen with care in order to obtain an accurate numerical solution. The time step was chosen in order to maintain numerical stability; the computational reference frame was chosen to move in the direction of wave propagation at a certain constant speed such that the wave train remains in the computational domain. Thus changes in wave speed as well as shape will become apparent in a space-time evolution plot. Also, the Hanning window is used to filter out any waves entering the computational domain from the adjacent domain.

Numerical simulations can be performed by using the observed internal wave field near the generation area as an initial condition to produce the wave evolution on the continental shelf and compare with the observations downstream. A parametric study for various environmental conditions can be carried out to demonstrate and to assess the nonlinear effects such as bottom topography, shoaling (across critical depth), and dissipation and mixing on internal wave evolution. The inclusion of these physical processes is essential to improve quantitative understanding of the coastal dynamics.

3. ERS-1 SAR Observations

The map of Taiwan and the study area of the East China Sea are shown in Figure 1 superimposed on the bathymetry for

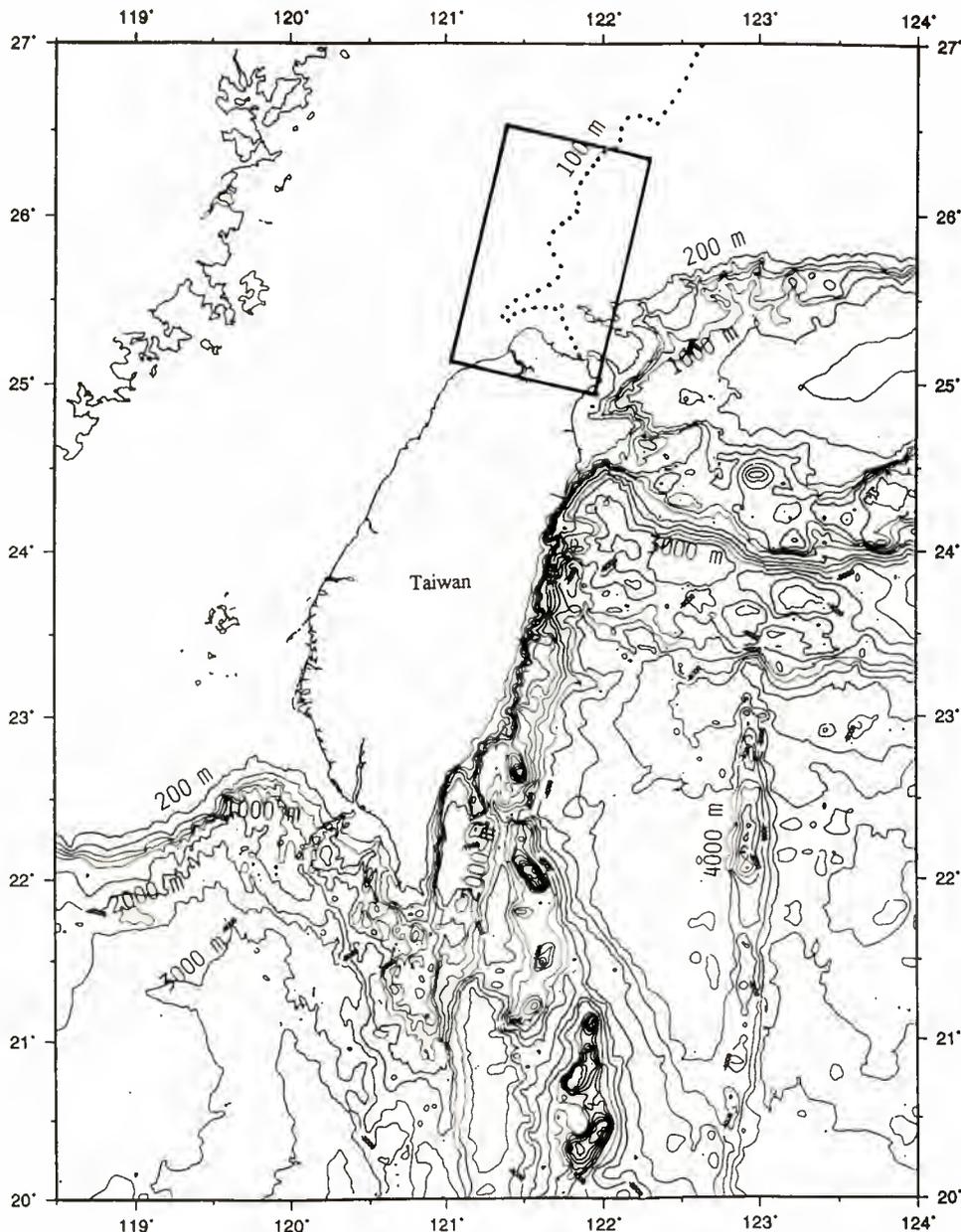


Figure 1. Map of Taiwan and the East China Sea superimposed on the bathymetry. The SAR coverage area is indicated by the box. The 100-m and 200-m depth contours are plotted as dashed and solid lines. Depth contours of 1000 m, 2000 m, 3000 m, and 4000 m are specified with 500-m interval.

reference. The SAR coverage area is indicated by the box. The 100-m and 200-m depth contours are plotted as dashed and solid lines. Depth contours of 1000 m, 2000 m, 3000 m, and 4000 m are specified with 500-m interval. We have used SAR images from ERS-1 to study the characteristics of internal waves northeast of Taiwan. In all SAR images the flight direction is indicated by an arrow. An ERS-1 SAR image (100×137 km) collected on November 10, 1993, shows a complicated internal wave pattern with wave generation, wave refraction, and wave-current interaction [Liu *et al.*, 1994] in Figure 2. Taiwan is located near the bottom of Figure 2, with the city of Taipei clearly identifiable at the center of the bottom edge. The dark area northeast of Taiwan is an upwelling induced by the Kuroshio intrusion on the continental shelf. Note that the internal wave packets are propagating in both onshore and offshore directions.

Rank-ordered packets of internal solitons propagating shoreward from the edge of the continental shelf were observed in the SAR images collected during September–October 1992. Two internal wave packets separated by approximately 30 km were identified from a SAR image of northeast of Taiwan collected on October 21, 1992 [Liang *et al.*, 1995, Figure 2]. Each wave is characterized by a dark band followed immediately by a bright band that is different from the previous observations of nonlinear depression waves (a bright band followed immediately by a dark band) in the New York Bight [Liu, 1988]. The reversed pattern in Figure 2 of Liang *et al.* [1995] indicates the existence of underlying nonlinear elevation internal waves. Based on the assumption of a semidiurnal tidal origin (12.4 hours), the wave speed can be estimated to be 0.66 cm s^{-1} and is consistent with that obtained from the internal wave theory. By using the SAR images and hydrographic data

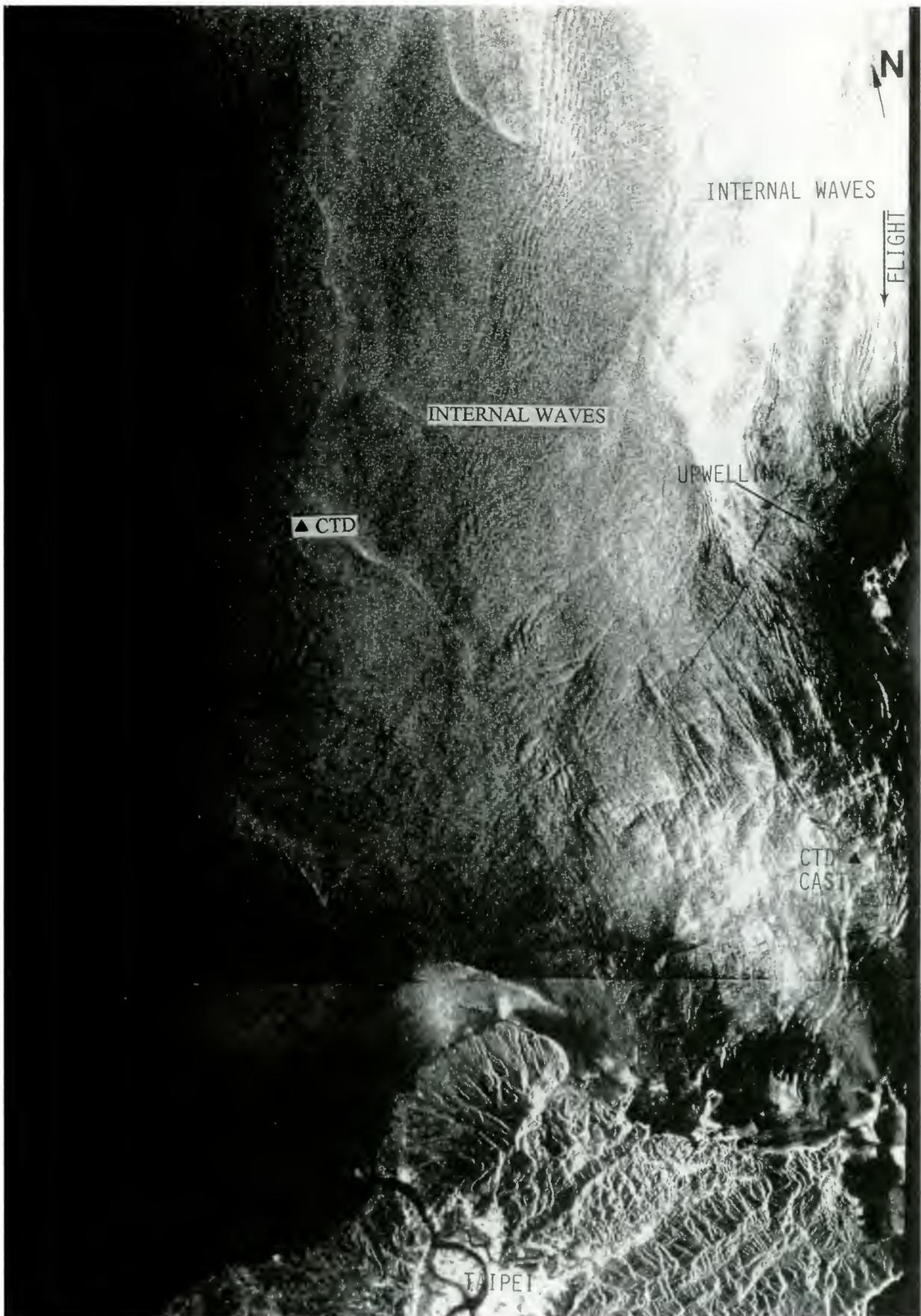


Figure 2. ERS-1 SAR image (100×137 km) collected northeast of Taiwan on November 10, 1993, showing a complicated internal wave pattern (copyright ESA 1993). The locations of upwelling, internal waves, and conductivity-temperature-depth (CTD) casts are indicated by letters.

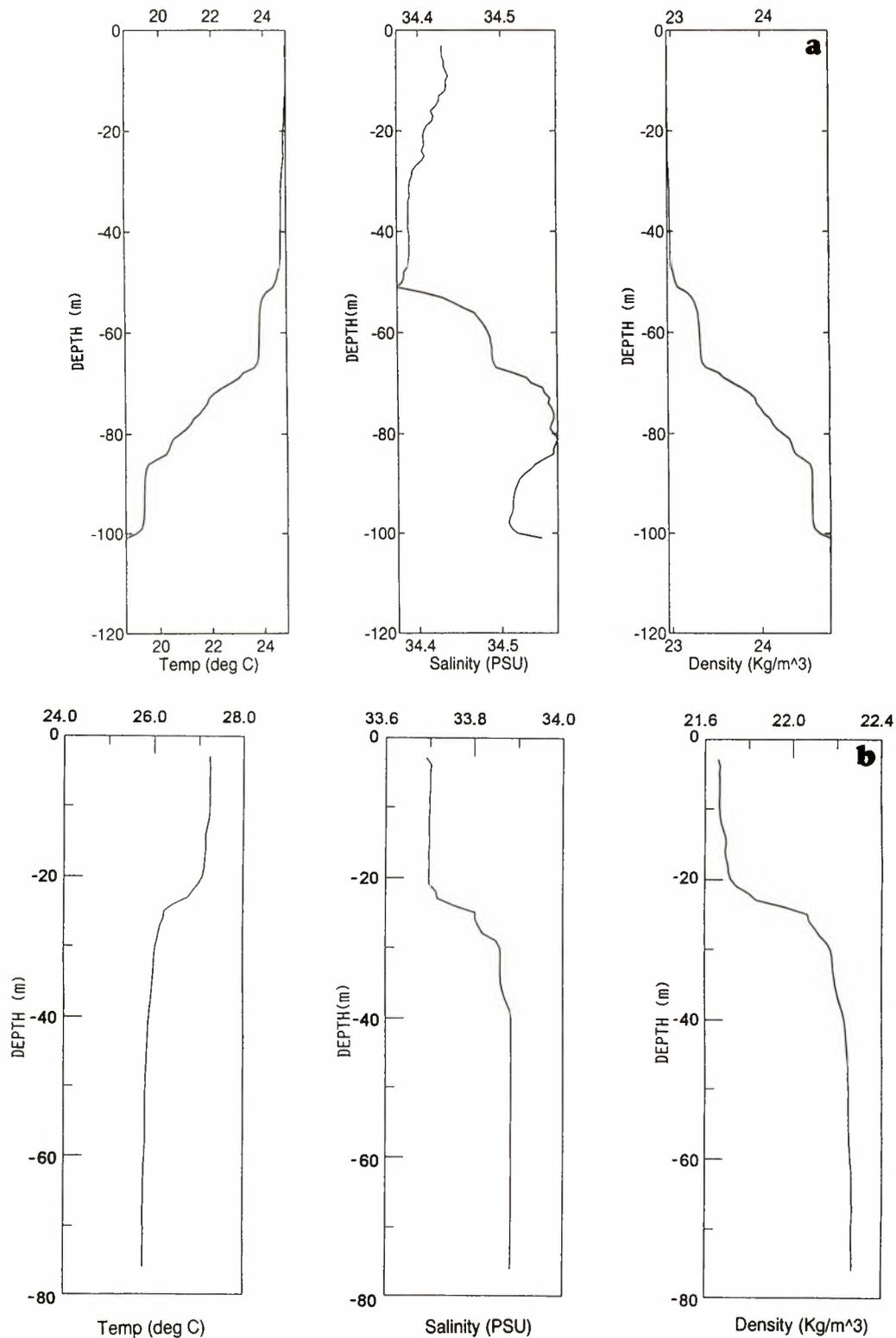


Figure 3. Temperature, salinity, and density profiles from CTD casts (a) on November 12, 1995 at 25°42'N, 122°29'E, northeast of Taiwan and (b) on July 3, 1996, at 26°19'N, 121°15'E, north of Taiwan as indicated in Figure 2.

[Hsueh, 1993], internal waves of elevation have been identified because of a thicker mixed layer as compared with the bottom layer on the continental shelf. The temperature, salinity, and density profiles from a conductivity-temperature-depth (CTD) cast on November 12, 1995, at coordinate 25°42'N, 122°29'E, which is at the vicinity of internal waves observed in the SAR

image, are shown in Figure 3a for reference. Notice that the mixed layer thickness is about 65 m in the water of 100-m depth, which represents typical late autumn condition. Another CTD cast representing the summer condition with a shallow mixed layer collected on July 3, 1996, at 26°19'N, 121°15'E, north of Taiwan is shown in Figure 3b for comparison.

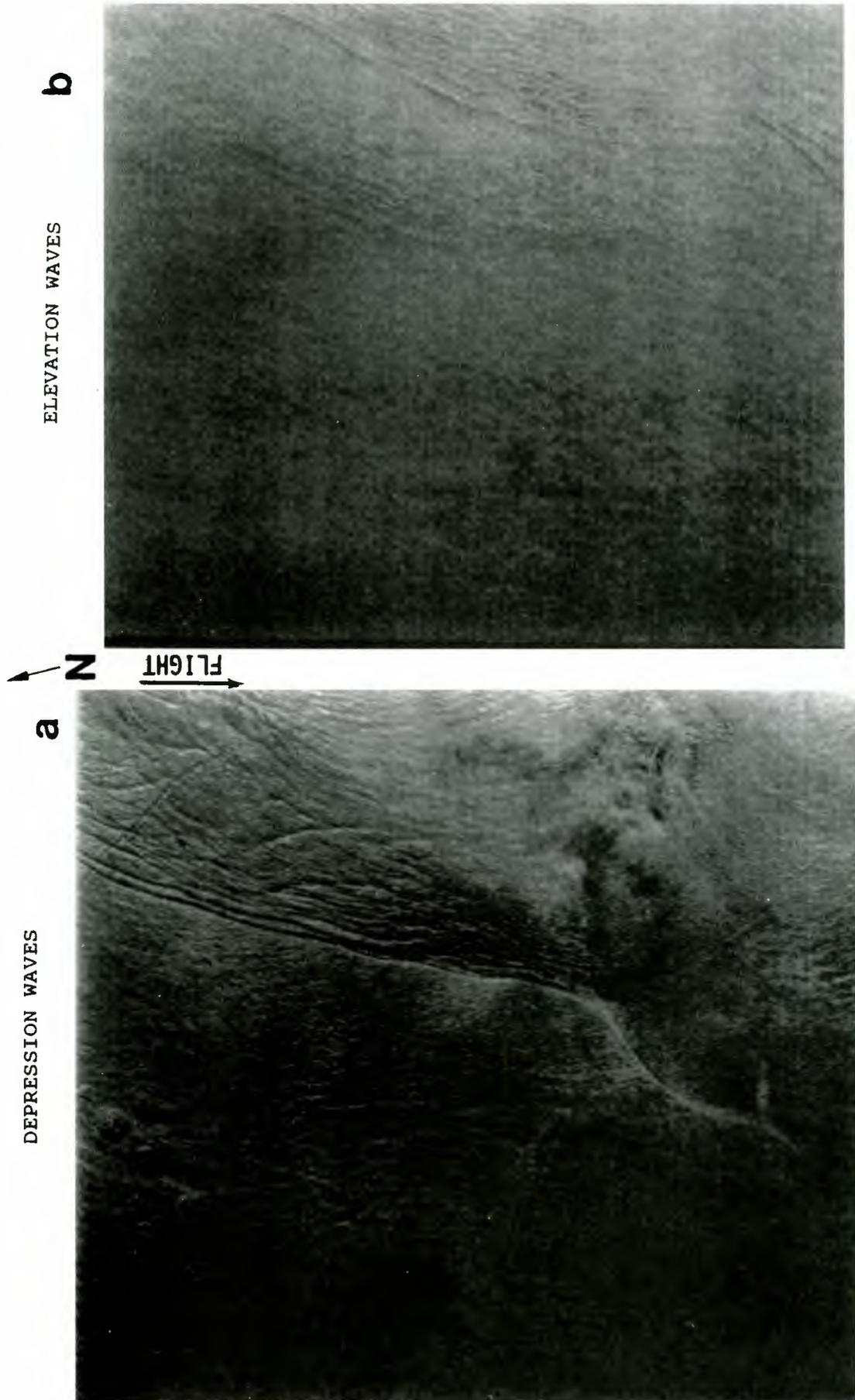


Figure 4. Surface signature patterns of two ERS-1 SAR images (100×100 km) collected east of Hainan Island in the South China Sea in June and April 1993, indicating the existence of underlying (a) depression and (b) elevation internal waves, respectively (copyright ESA 1993).

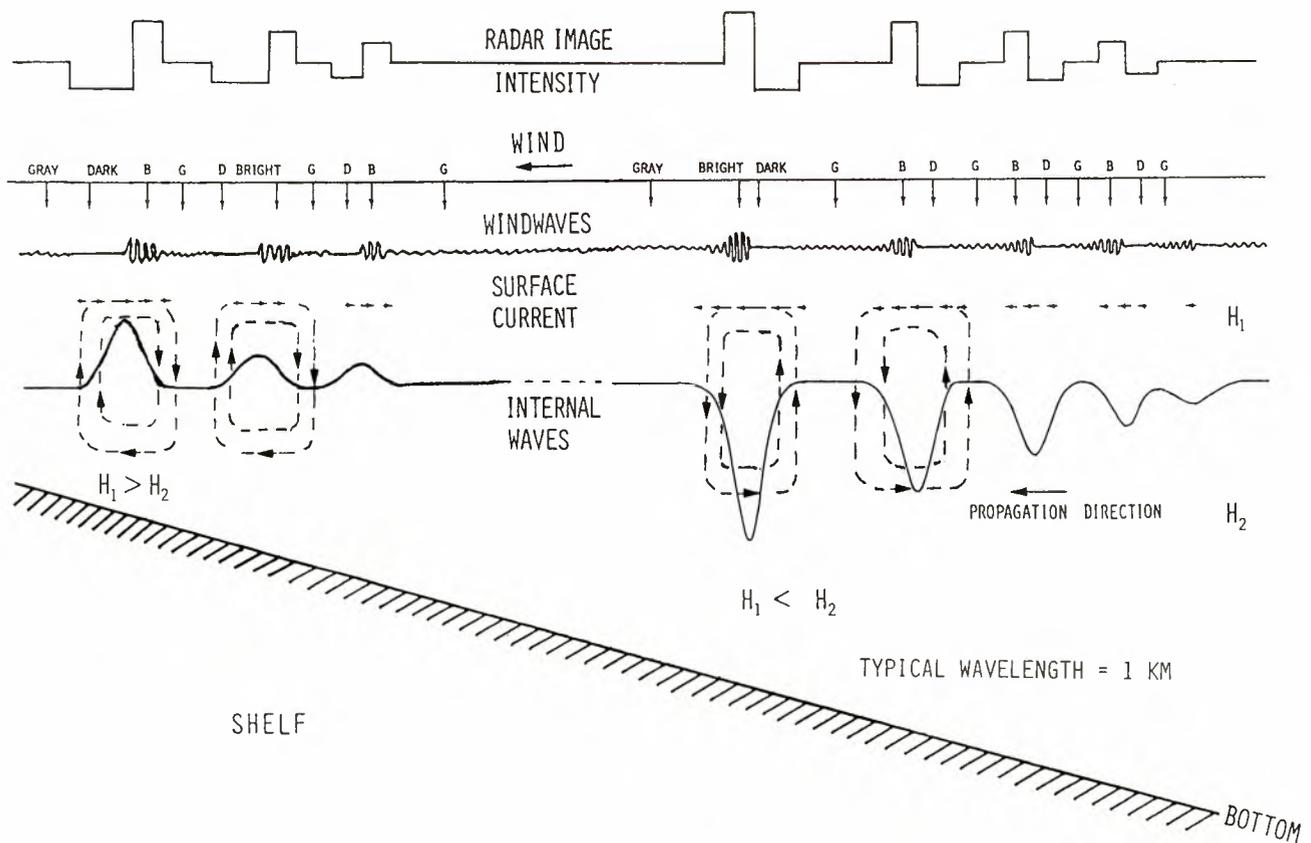


Figure 5. Schematic diagram of interaction of internal waves, surface current field, wind waves, and the resultant SAR image intensity variation.

The elevation internal waves were also observed on the east shelf of Hainan Island in the South China Sea. Surface signature patterns of two ERS-1 SAR images (100×100 km) of east of Hainan Island in the South China Sea (center coordinate at 21.35°N and 115.6°E) collected on June 16 and April 7, 1993, indicate the existence of underlying depression and elevation internal waves, respectively (Figure 4a and 4b) due to the change of mixed layer depth in different seasons. During summer the mixed layer is shallow, and its thickness is thinner than the bottom layer. During winter–spring the mixed layer deepens as a result of strong winds, and its thickness is thicker than the bottom layer. A schematic diagram has been established to describe the interaction of internal waves, surface current field, wind waves, and the resultant SAR image intensity variation as shown in Figure 5. When the mixed layer depth is thinner than the bottom layer, $H_1 < H_2$, only depression waves can be evolved (Figure 4a) as indicated in equation (2). While the mixed layer is thicker than the bottom layer, $H_1 > H_2$, only the elevation waves can be evolved (Figure 4b). The elevation waves induce a reversed flow field as compared with the depression waves. Thus the surface pattern of elevation internal waves observed in the SAR image is reversed as compared with that of the depression waves.

The surface signature of internal waves can be different when the surface slicks or films are abundant owing to active biological processes in the coastal water such as in the Taiwan coastal area. The reason that oil slicks are detected on radar images is that oil films have a dampening effect on short surface waves [Ermaikov and Pelinovsky, 1984; Alpers and Hü-

nerfuss, 1989]. The dark appearance of surface slicks on radar images is due to the smoothing of the ocean surface caused by the damping of short backscattering waves. Therefore the convergence of surface slicks induced by the underlying internal waves will show a series of dark bands in the SAR images (12.5×12.5 km) as indicated in the right part of Figure 6. In the left part of SAR image (Figure 6), a depression internal wave packet can be identified as a bright band followed immediately by a dark band as induced by the surface current (strain rate). Therefore the surface signatures of internal waves caused by surface slicks (dark bands only) are different from those due to surface current strain rate (a bright band followed immediately by a dark band).

The generation mechanisms include influences of the tide and the Kuroshio for the formations of both elevation internal waves and depression waves in different ocean environment conditions. Near northeastern Taiwan, the Kuroshio intrudes onto the continental shelf immediately after passing north of Taiwan. A cold water anomaly, manifested as upwelling of the subsurface Kuroshio water, has been frequently observed at the shelf break of the East China Sea to the north of Taiwan. A schematic diagram of internal elevation wave generation process is proposed and shown in Figure 7. The upwelling induced by the Kuroshio intrusion at the shelf break may have a dooming effect on the mixed layer bottom. The disturbance of the doomed area is then driven by the semidiurnal tide onto the shelf and evolves into a rank-ordered elevation wave packet in the shallow area or a rank-ordered depression wave packet in the deep area depending on the mixed layer depth.

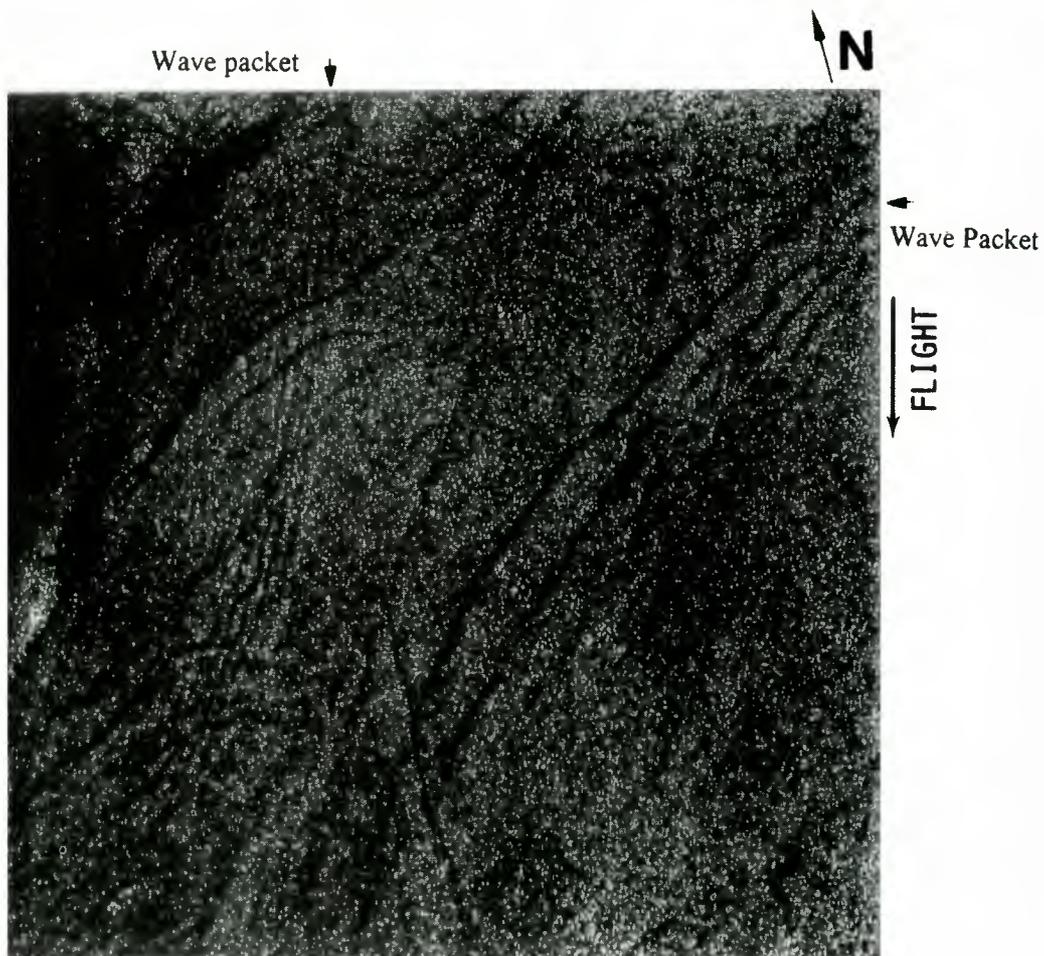


Figure 6. Subscene of ERS-1 SAR image (12.5×12.5 km) collected northeast of Taiwan on May 31, 1995, showing internal wave patterns caused by surface slicks (indicated by the arrow on the right side) and by surface current (indicated by the arrow at the top). Copyright ESA 1993.

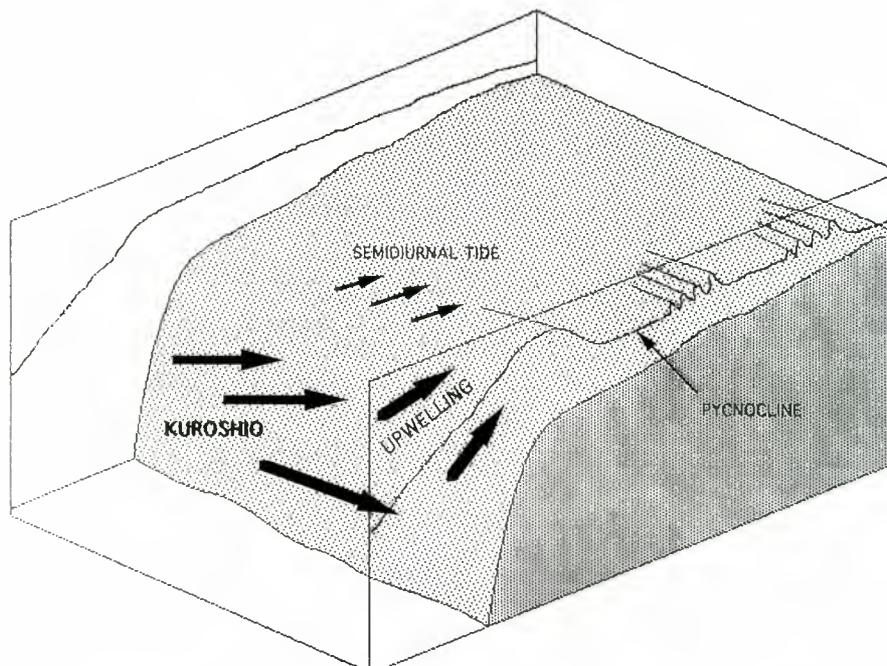


Figure 7. Schematic diagram of internal elevation wave generation process with upwelling induced by the Kuroshio intrusion at the continental shelf break.

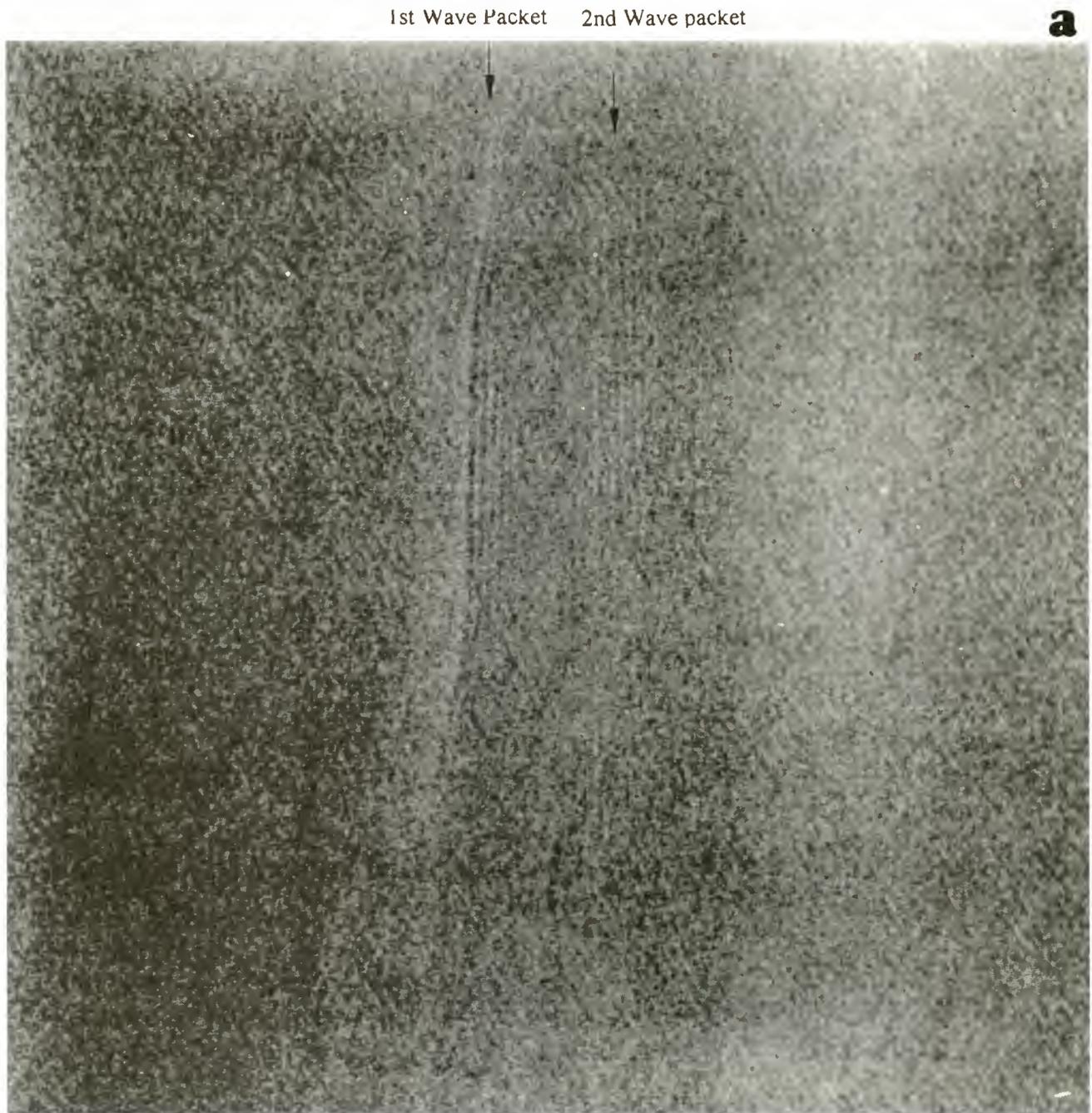


Figure 8. (a) Subscene of the ERS-1 SAR image (25×25 km) collected northeast of Taiwan on September 16, 1992, showing two short internal wave packets produced by the disintegration of solitons (copyright ESA 1992). (b) Two-dimensional wave spectrum of the SAR subscene showing the internal waves and background swell system.

The tidal flow over topographic features such as a valley, island or continental shelf in a stratified ocean can produce depression type disturbances in deep water, as has also been observed in the Taiwan area (Figure 2). The evolution of solitons is based upon the balance of nonlinear effect with the dispersive effect. The sign of nonlinear term depends on the wave amplitude; a depression wave has the opposite sign of an elevation wave in amplitude. Therefore only the elevation waves can be evolved nearshore because the bottom layer thickness is thinner than the mixed layer depth over the sloping shelf. Then the depression solitons from the shelf break will disintegrate into dispersive wave packets in the shallow water

area after they pass through a “turning point” of approximately equal layer depths (critical depth where the nonlinear term is zero and changes sign). Two disintegrated internal wave packets have been observed in the SAR subscene (25×25 km) image collected on September 16, 1992 (Figure 8a) in a shallow water area on the shelf north of Taiwan. The wavelength of internal waves in wave packets is about 260 m, which is almost the same wavelength as the swell system propagating in the southwest direction that can be seen in the background. Figure 8b shows the two-dimensional wave spectrum of the SAR scene. Note that the swell of 289-m wavelength is coming from 46° (with respect to north) and the internal waves are coming

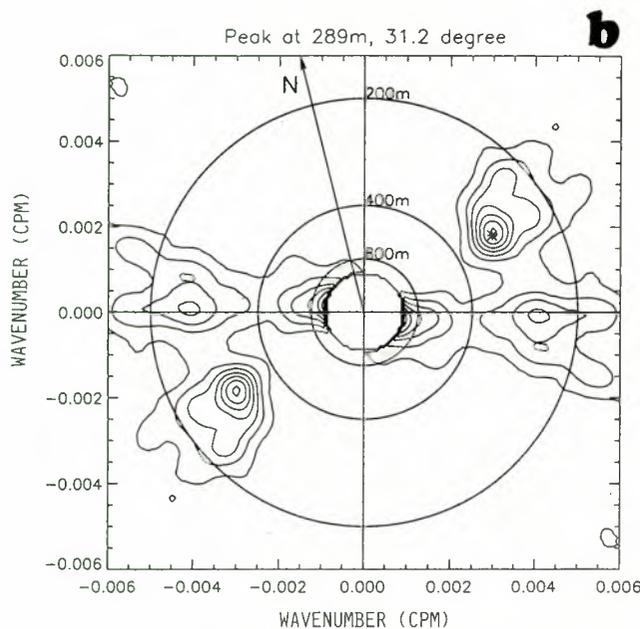


Figure 8. (continued)

from the east as indicated in the spectrum. The wave packet size is about 1 km, which is the size of a depression soliton observed before near the shelf break.

The Kuroshio moving north from Philippine Basin branches out near the south tip of Taiwan. A part of the Kuroshio intrudes into the South China Sea through the Bashi Channel and the Luzon Strait. The internal tides and internal waves have been generated by the shallow ridges (200–300 m) in the Luzon Strait. The surface signature pattern of huge internal soliton packets has been observed in the ERS-1 SAR image (100 × 200 km) collected on June 16, 1995, as shown in Figure 9. The soliton crest is >200 km long, and each packet contains >10 rank-ordered solitons with a packet width of 25 km. Within a wave packet, the wavelengths appear to be monotonically decreasing, front to rear, from 5 km to 500 m. These are the largest internal waves that have been observed in this area. The internal wave amplitude is larger than 100 m on the basis of the CTD casts from Taiwan's research ship during their South China Sea expedition (J. Wang, private communication, 1995). These huge wave packets propagate and evolve into the deep South China Sea and will reach the continental shelf of southern China. It is possible that the internal wave packets observed in the South China Sea east of Hainan are due to the evolution of these waves. The distance and degree of intrusion of high-temperature and high-salinity Kuroshio water into the Taiwan Strait and the South China Sea with heat flux and momentum flux are still open issues.

4. Numerical Study and Model Results

A series of numerical experiments were performed by solving the initial value problem described by equation (1) with a well developed soliton solution as an initial condition. The steady state depression wave solution of the KdV equation is given by

$$A(x, t) = -A_0 \operatorname{sech}^2 [(x - Ct)/l], \quad (5)$$

where A_0 is the wave amplitude, the wave speed C is

$$C = C_0 \left(1 + \frac{A_0(H_2 - H_1)}{2H_1H_2} \right), \quad (6)$$

and the half width l is

$$l = 2H_1H_2[3A_0(H_2 - H_1)]^{-1/2} \quad (7)$$

In the north of Taiwan area, the density contrast $\Delta\rho/\rho = 10^{-3}$. The first case was made for a initial profile corresponding to a single depression soliton in finite/deep water of depth H propagating towards a cosine-shaped transition to shallow water:

$$H_2 = H_{20}, \quad x < 0, \quad (8a)$$

$$H_2 = H_{20} + 0.5(H_{20} - H_{2c})[\cos(\pi^*x/L) - 1], \quad (8b)$$

$$0 < x < L,$$

$$H_2 = H_{2c}, \quad x > L, \quad (8c)$$

where L is a characteristic length of the depth variation.

The numerical simulations of nonlinear internal wave evolution on the shelf show the change of polarity through critical depth (Figure 10). A solitary depression wave is used as the initial condition at water depth of 160 m with $A_0 = 10$ m. The mixed layer thickness, $H_1 = 60$ m, the initial bottom layer thickness $H_{20} = 100$ m, the final bottom layer thickness $H_{2c} = 40$ m, and the characteristic length $L = 20$ km (with steep slope) are chosen for Figure 10a. Note that the critical depth is reached after 6 hours (at $x = 0.6L$) approximately. However, the solitary wave disintegrates into a wave packet after 10 hours (near $x = L$), and it continues to evolve as a train of rank-ordered elevation internal solitons after 22 hours. It is found that from a single depression soliton, more than five elevation solitons can merge. For sensitivity analysis, a series of parametric study have been performed with variable wave amplitude, water depth, bottom slope, and dissipation. In general, all numerical results show the change of polarity through critical depth. Two typical cases are shown in Figure 10b for mild bottom slope, and in Figure 10c for steep slope with dissipation. For mild bottom slope with characteristic length of 100 km (Figure 10b), the evolution is much slower, and the solitary wave disintegrates into a wave packet after 30 hours (near $x = L$) and continues to evolve as a train of rank-ordered elevation internal solitons after 55 hours. For steep bottom slope with dissipation coefficient $\varepsilon = 1$ m²/s (Figure 10c), the solitary wave disintegrates into a wave packet after 10 hours, but its amplitudes damp to only the half of no-dissipation case in Figure 10a.

Next, the case of two rank-ordered depression solitons ($A_0 = 10$ and 7 m) separated by 25 km is studied. Two wave packets start to evolve and merge after 6 hours near the critical depth. However, without dissipation they are well overlapped after 11 hours, as is shown in Figure 11a. The dissipation effects are expected to be important in this situation, since the interaction of wave packets may steepen the wave amplitude. Thus the eddy viscosity could be eroding the sharp peaks of the large solitons, reducing their amplitudes and increasing their half-widths at the same time. Figure 11b shows the evolution of two depression solitons through the critical depth with a dissipation coefficient $\varepsilon = 1$ m²/s. Notice that two wave packets are well separated with much shorter wavelengths, which is quite similar to the observation from SAR image in Figure 8. Although the in situ measurements are not available for this situation, the numerical calculations describe the potential in-

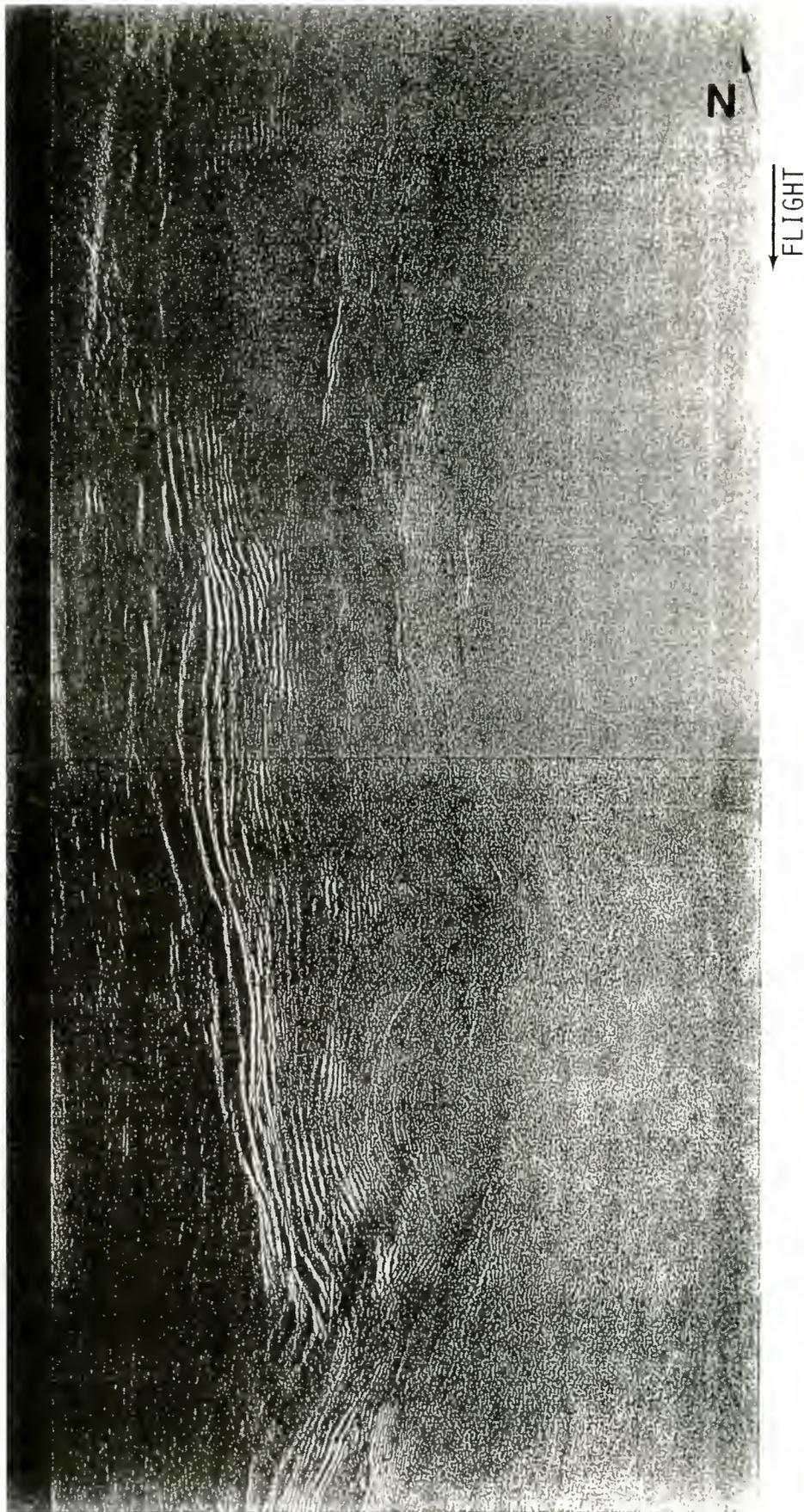


Figure 9. Mosaic of ERS-1 SAR images (100×200 km) collected south of Taiwan in Luzon Strait on June 16, 1995, showing huge internal soliton packet (copyright ESA 1995).

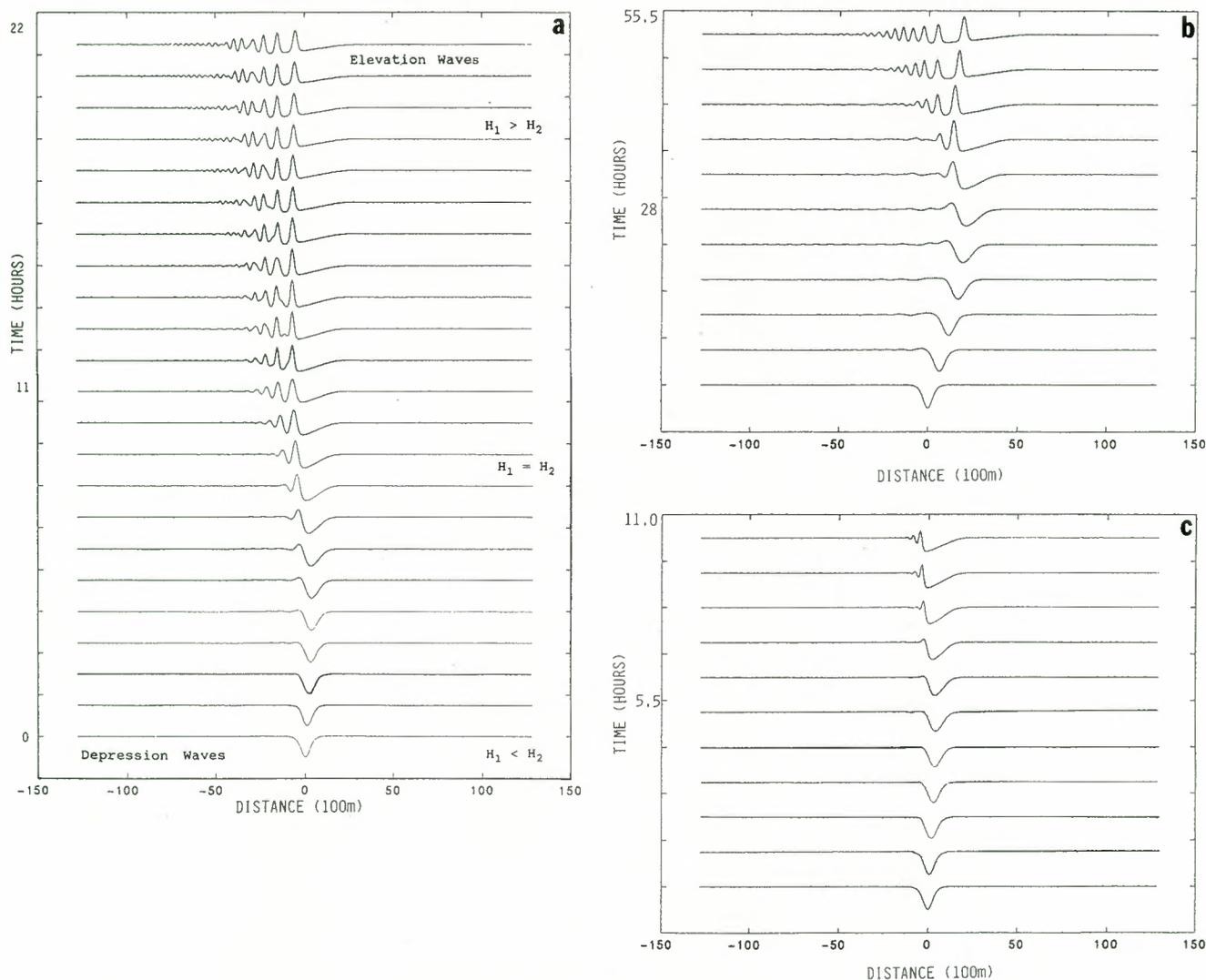


Figure 10. Numerical simulations of nonlinear internal wave evolution on the shelf showing the change of polarity through critical depth for (a) steep slope, (b) mild slope, and (c) steep slope with dissipation $\varepsilon = 1 \text{ m}^2/\text{s}$.

terpretation of SAR observations. These types of processes, by which depression solitons disintegrate into wave packets and then evolve to elevation solitons, have never been observed and reported before. While further in situ measurements are needed to verify the SAR observations, the numerical predictions developed here should remain as an interesting topic.

The essential element of the surface effects is the interaction between the internal-wave-induced surface current field and the wind-driven ocean surface waves. This interaction has been studied, notably by Hughes [1978] and Holliday *et al.* [1987]. Basically, the analysis is based on a near-equilibrium spectral transport model to estimate the roughness modulation by a variable surface current. The effect of the surface current is to alter the spectrum from its equilibrium value, while the natural processes of wave energy input from the wind, wave breaking, and other nonconservative processes act to restore the ambient equilibrium spectrum. According to a first-order radar-imaging theory [Alpers, 1985], the relative variation of the normalized radar cross section (relative modulation) associated with internal waves is related to the horizontal gradient of the surface velocity, the strain rate [Liu, 1988]. The proportional coefficient

depends on radar wavelength, radar incidence angle, angle between the radar look direction and the internal wave propagation direction, azimuth angle, and wind velocity. Thus to first order (for a linear SAR system), the variation of the SAR image intensity is proportional to the gradient of the surface velocity, or the strain rate. Figure 12a shows the relative modulation of a cross section in the SAR image along the wave packet propagation direction in Figure 8a with two packets as indicated by arrows. The strain rate has been calculated for the final numerical results in Figure 11a as shown in Figure 12b. The strain rate values are of the order of 10^{-2} s^{-1} , which are consistent with the observed data from New York Bight internal waves [Liu, 1988]. The patterns of two wave packets are similar for the SAR observation in Figure 12a and for the numerical results in Figure 12b. The wavelength of the wave packets in Figure 12b is found to be 340 m, which is also close to the wavelength of 260 m from the SAR image in Figure 12a. Since we do not have in situ measurements for the numerical simulation, this qualitative comparison demonstrates the connection between the numerical model and the SAR radar cross section observation of internal waves.

5. Discussion

Many SAR images from ERS-1 have been collected and used to study the characteristics of internal waves northeast and south of Taiwan in the East China Sea and east of Hainan Island in the South China Sea. Rank-ordered packets of internal solitons propagating shoreward from the edge of the continental shelf were observed in SAR images. The wave speed can be estimated on the basis of assumption of a semidiurnal tidal origin and is consistent with the internal wave theory. By using the SAR images and hydrographic data, internal waves of elevation have been identified in shallow water due to a thicker mixed layer as compared with the bottom layer on the continental shelf. The generation mechanisms, including the influences of the tide and the Kuroshio, for the formations of both elevation internal waves and depression waves under different ocean conditions have been proposed. The effects of water depth on the evolution of solitons and wave packets are modeled by a KdV type equation and linked to the satellite SAR observations of elevation internal waves and disintegrated internal wave packets. For a case of depression waves in deep

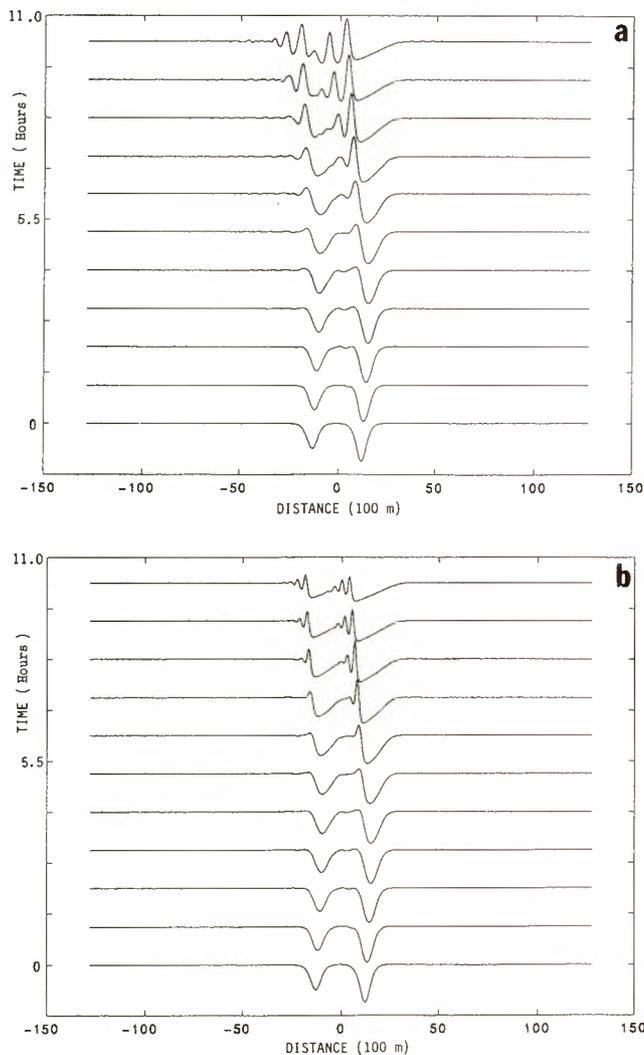


Figure 11. Numerical simulations of nonlinear internal wave evolution on the shelf showing the disintegration of two solitons into two wave packets for the cases (a) with no dissipation and (b) with dissipation $\varepsilon = 1 \text{ m}^2/\text{s}$.

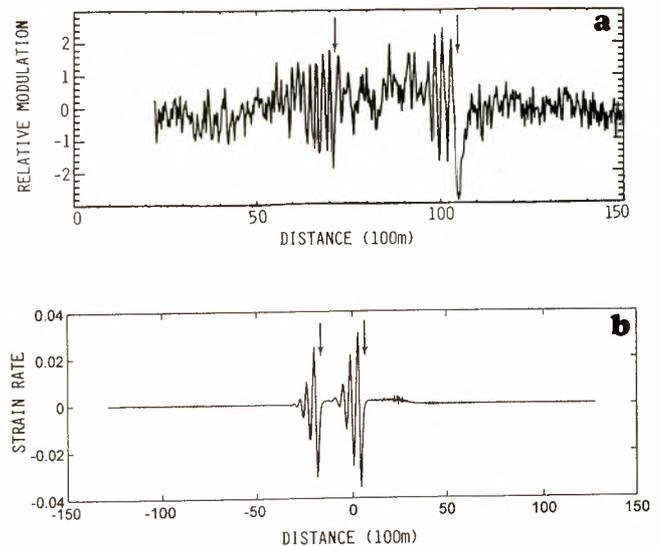


Figure 12. (a) The relative modulation of a cross section in the SAR image (Figure 8a) along the wave packet propagation direction, and (b) the surface strain rate from the final results of Figure 11a with two internal wave packets indicated by arrows.

water, the solitons first disintegrate into dispersive wave trains and then evolve into a packet of elevation waves in the shallow water area after they pass through a “turning point” of critical depth has been simulated by a numerical model. Examples of the numerical model results of nonlinear internal wave evolution on the shelf with topographic and dissipation effects are presented in this paper to interpret the observations in the SAR images.

The essential element of the surface effects is the interaction between the internal-wave-induced surface current field and the wind-driven ocean surface waves. For a linear SAR system the variation of the SAR image intensity is proportional to the gradient of the surface velocity, or to the strain rate. The proportionality depends on radar wavelength, radar incidence angle, angle between the radar look direction and the internal wave propagation direction, azimuth angle, and wind velocity. For the high-wind-speed condition, the internal wave signal may be too weak to be observed by radar owing to low signal-to-noise ratio. When the internal waves propagate in the cross-wind direction, the wave-current interaction is also relatively weak, and so is the radar backscattering for SAR observation. The strain rates have been calculated for the internal wave packets and their values are consistent with the observed data from New York Bight internal waves. Without in situ field measurements for detailed numerical simulations, the comparison of relative modulation from SAR data and strain rate from model are reasonable qualitatively.

It is clear that these internal wave observations in the East and South China Seas provide a unique resource for addressing a wide range of processes, including the following: generation of elevation internal waves by upwelling due to the Kuroshio intrusion across the continental shelf, evolution of depression waves through the critical depth, disintegration of solitons into internal wave packets, dissipation effects of internal wave breaking and turbulent mixing on wave propagation, shoaling effects of variable bottom topography on wave evolution, and wave packet interaction. These types of processes

by which the depression solitons disintegrate into wave packets and then evolve into elevation solitons have never been observed and reported before. Moderate elevation internal waves could be generated in waters where the mixed layer is still shallow in comparison with the lower bottom layer. However, owing to the nonlinear effects, these elevation internal waves will not survive and will evolve to a dispersive wave train. Further in situ measurements are needed to verify the SAR observations. A SAR image provides only a snapshot of the internal wave evolution, but a large spatial coverage over the field measurement area. However, the repeat cycle for ERS-1 SAR is 35 days for general purposes. Therefore the entire evolution from depression to elevation internal waves cannot be observed by SAR instantaneously. Based on the SAR image for field test planning, two moorings at upstream and downstream of "critical depth" can be deployed to verify the numerical results. Numerical simulations can be performed by using the observed internal wave field near the generation area as an initial condition to produce the wave evolution on the continental shelf and then to compare with the observations. On the basis of observations of internal wave evolution from SAR images and model predictions, the ratio of mixed layer depth to the bottom layer depth can be estimated. Furthermore, the mixed layer depth can be derived also from the internal wave speed, which is estimated from the distance of sequential wave packets in the SAR images (based on the assumption of semidiurnal tidal generation).

Acknowledgments. The authors wish to thank C. T. Liu of National Taiwan University for his valuable discussions and suggestions. T.-Y. Tang of National Taiwan University provided the CTD data in Figure 3. The ERS-1 SAR images collected by ESA satellite and processed by ESA and Taiwan Ground Station are also acknowledged. GSFC work was supported by the National Aeronautics and Space Administration and Office of Naval Research.

References

- Alpers, W., Theory of radar imaging of internal waves, *Nature*, 314, 245–247, 1985.
- Alpers W., and H. Hühnerfuss, The damping of ocean waves by surface films: A new look at an old problem, *J. Geophys. Res.*, 94, 6251–6265, 1989.
- Apel, J. R., Linear and nonlinear internal waves in coastal and marginal seas, in *Oceanographic Applications of Remote Sensing*, edited by M. Ikeda and F. Dobson, pp. 57–78, CRC Press, Boca Raton, Fla., 1995.
- Apel, J. R., J. R. Holbrook, A. K. Liu, and J. Tsai, The Sulu Sea internal soliton experiment, *J. Phys. Oceanogr.*, 15, 1625–1651, 1985.
- Bogucki, D., and C. Garrett, A simple model for the shear-induced decay of an internal solitary wave, *J. Phys. Oceanogr.*, 23, 1767–1776, 1993.
- Brandt, P., W. Alpers, and J. O. Backhaus, Study of the generation and propagation of internal waves in the Strait of Gibraltar using a numerical model and synthetic aperture radar images of the European ERS-1 satellite, *J. Geophys. Res.*, 101, 14,237–14,252, 1996.
- Ermakov, S. A., and E. N. Pelinovsky, Variations of the spectrum of wind ripples on coastal waters under the action of internal waves, *Dyn. Atmos. Oceans*, 8, 95–100, 1984.
- Fu, L.-L., and B. Holt, Seasat view oceans and sea ice with synthetic aperture radar, *JPL Publ. 81-120*, 200 pp., Pasadena, Calif., 1982.
- Helfrich, K. R., W. K. Melville, and J. Miles, On interfacial solitary waves over slowly varying topography, *J. Fluid Mech.*, 149, 305–317, 1984.
- Holliday, D., G. St.-Cyr, and N. E. Woods, Comparison of a new radar ocean imaging model with SARSEX internal wave image data, *Int. J. Remote Sens.*, 8, 1423–1430, 1987.
- Hsueh, Y., C.-S. Chern, and J. Wang, Blocking of the Kuroshio by the continental shelf northeast of Taiwan, *J. Geophys. Res.*, 98, 12,351–12,360, 1993.
- Hughes, B. A., The effect of internal waves on surface wind waves, 2, Theoretical analysis, *J. Geophys. Res.*, 83, 455–465, 1978.
- Liang, N. K., A. K. Liu, and C. Y. Peng, A preliminary study of SAR imagery on Taiwan coastal water, *Acta Oceanogr. Taiwan.*, 34, 17–28, 1995.
- Lin, C.-Y., C.-Z. Shyu, and W.-H. Shih, The Kuroshio fronts and cold eddies off northeast Taiwan observed by NOAA-AVHRR imageries, *Terr. Atmos. Oceanic Sci.*, 3, 225–242, 1992.
- Liu, A. K., Analysis of nonlinear internal waves in the New York Bight, *J. Geophys. Res.*, 93, 12,317–12,329, 1988.
- Liu, A. K., and C. Y. Peng, Ocean wave study using SAR, in *Satellite Remote Sensing of Oceanic Environment*, edited by I. S. F. Jones, R. W. Stewart, and Y. Sugimori, pp. 280–285, Seibutsu Kenkyusha, Tokyo, 1994.
- Liu, A. K., J. R. Apel, and J. R. Holbrook, Nonlinear internal wave evolution in the Sulu Sea, *J. Phys. Oceanogr.*, 15, 1613–1624, 1985.
- Liu, A. K., G.-C. Gong, C.-Z. Shyu, S.-C. Pai, C.-L. Wei, and S.-Y. Chao, Response of Kuroshio upwelling to the onset of the northeast monsoon in the sea north of Taiwan: Observations and a numerical simulation, *J. Geophys. Res.*, 97, 12,511–12,526, 1992.
- Liu, A. K., C. Y. Peng, and J. D. Schumacher, Wave-current interaction study in the Gulf of Alaska for detection of eddies by SAR, *J. Geophys. Res.*, 99, 10,075–10,085, 1994.
- Salusti, E., A. Lascaratos, and K. Nittis, Changes of polarity in marine internal waves: Field evidence in eastern Mediterranean Sea, *Ocean Modell.* 82, pp. 10–11, Hooke Inst., Univ. of Oxford, Oxford, England, 1989.
- Sandstrom, H., and J. A. Elliott, Internal tide and solitons on the Scotian Shelf: A nutrient pump at work, *J. Geophys. Res.*, 89, 6415–6426, 1984.
- Y. S. Chang and A. T. Liu, Oceans and Ice Branch, NASA Goddard Space Flight Center, Greenbelt, MD 20771. (e-mail: liu@neptune.gsfc.nasa.gov)
- M.-K. Hsu, Department of Oceanography, National Taiwan Ocean University, Keelung, Taiwan.
- N. K. Liang, Institute of Oceanography, National Taiwan University, P.O. Box 23-13, Taipei, Taiwan.

(Received December 2, 1996; revised June 20, 1997; accepted July 3, 1997.)

Role of surface films in ERS SAR signatures of internal waves on the shelf

1. Short-period internal waves

J. C. B. da Silva,¹ S. A. Ermakov,² I. S. Robinson,¹ D. R. G. Jeans,^{3,4} and S. V. Kijashko²

Abstract. Analysis of synthetic aperture radar (SAR) images of internal waves (IWs) on the Iberian shelf and classification of the IW surface signatures is presented. There are three types of IW signatures in the form of bright/dark, dark, and bright bands that correspond to positive/negative, negative, or positive variations of radar backscatter, respectively. Strong positive sign IW signatures occur at very low wind velocities (less than 2 m/s). At winds higher than 2 m/s both positive/negative and negative signatures are observed for range-propagating IWs. For azimuth-propagating IWs, negative signatures prevail. Two groups of experiments carried out in the frame of the Multidisciplinary Oceanographic Research in the Eastern North Atlantic (MORENA) project on the Iberian shelf in August 1994 are described: (1) IW measurements simultaneous with ERS-1 SAR overpasses and (2) IW measurements, visual observations of the surface, and sampling of films from the sea surface. IW manifestations at low to moderate winds took the form of slicks located over IW troughs; at the near-threshold wind velocity they were of the form of intensified decimeter-scale waves (antislicks) located over IW crests. Measurements of wave damping due to films collected from the slick and nonslick areas showed the concentration of surfactants and retrieved film elasticity to be higher in the slicks than in the nonslick areas. A theoretical model of the surface wave modulation by IWs is developed to include surface wave straining by the IW current and surface wave damping due to surfactant films. Pressure-area curves for real marine films are used in the model. The relation between the film and the straining effects on the surface wave spectrum in the IW field is shown to depend on film parameters, surface wavelength, and wind velocity. The model predicts strong damping of centimeter-scale waves over IW troughs mainly due to the film effect and intensification and depression of decimeter-scale waves due to both straining and film effects. Dependencies of the model spectrum variations as a function of wind velocity for range and azimuth propagating IWs are obtained, the negative sign contrast being shown to predominate for azimuth propagating IWs. Model variations of the spectrum of decimeter-scale waves are shown to be significantly sensitive to surfactants (film pressure), the negative sign contrasts due to films being superimposed on the straining effects, and these are eventually capable of suppressing the positive sign contrasts. It is concluded that the occurrence of the different classes of IW signatures in the SAR images can be explained by the action of surface-active films.

1. Introduction

It is well known that oceanic internal waves are commonly observed in synthetic aperture radar (SAR) images as systems of quasi-periodic parallel bands [see, e.g., Vesecky and Stewart, 1982; Apel and Gonzalez, 1983; Alpers, 1985;

Thompson and Gasparovic, 1986; Onstott and Rufenach, 1992]. These are frequently discussed in the literature as bright and dark bands on a grey background [see, e.g., Alpers, 1985; Onstott and Rufenach, 1992]. On some occasions, however, they are seen as dark bands, the expected associated bright bands being either absent or strongly attenuated. Additionally, there is another type of signature that will be discussed in this paper and which has received very little attention in the literature. This is when internal waves appear as bright bands in a very low (dark) SAR background level. Such single bright features have been observed when there is very little backscatter from the ambient sea, presumably due to low wind speeds [see Gasparovic *et al.*, 1986].

Radar signatures of internal waves (IWs) result from the modulation of wind-generated surface waves by the surface currents associated with the IWs. Such currents produce horizontal convergence and divergence, which changes the roughness at the sea surface. The change in roughness can be

¹Department of Oceanography, Southampton Oceanography Centre, Southampton, England, United Kingdom.

²Institute of Applied Physics, Russian Academy of Sciences, Nizhny Novgorod, Russia.

³Unit for Coastal and Estuarine Studies, Marine Science Laboratories, University College of North Wales, Anglesey, Wales, United Kingdom.

⁴Now at GEOS Limited, Swindon, England, United Kingdom.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02725.
0148-0227/98/97JC-02725\$09.00

caused by straining of the surface waves in the gradients of the surface velocity field and/or by wave damping due to films of varying thickness when periodically compacted and expanded in the internal wave field. The imaging of IWs by L-band SAR in the form of bright/dark bands was discussed by *Alpers* [1985], who suggested that the signatures could be satisfactorily explained by the hydrodynamic modulation theory on the basis of the straining of surface waves due to inhomogeneous currents induced by IWs. It was also assumed that the dark SAR signatures of IWs are observed presumably in coastal waters and related to surfactant films on the sea surface. The role of a contaminating surface film of organic composition in the action of oceanic internal waves on surface waves has been discussed since the 1950s [see, e.g., *Ewing*, 1950]. Under light winds, such surface active materials (surfactants) can inhibit the formation of ripples and damp those already formed, causing visible differences in the short-scale roughness of the water. Quantitatively, the film effect of modulation of short surface waves by IWs was analyzed by *Ermakov and Pelinovsky* [1984] and by *Ermakov et al.*, [1992], who also presented experimental evidence of the film mechanism.

Whereas theories have been advanced to interpret roughness patterns in terms of either straining of short surface waves or their variation due to films, the combined effect of both mechanisms, which can be important for SAR imaging of IWs, has not been studied so far. In particular, no quantitative analysis has been provided to account for the extent to which surface films may transform one type of signature into the other. In this paper we reexamine the role of surface films in accounting for the SAR signatures of short-scale internal waves (typically of order 0.1-1 km). It will be demonstrated that in some circumstances both straining and film mechanisms are important and appear to be necessary to explain the different classes of the signatures of IWs that are observed in the radar images. Results of an analysis of a set of ERS-1 SAR images (C-band radar) are presented, demonstrating the existence of three different types of IW signatures and their characteristics (section 2). The circumstances under which we observe each one of the mentioned three character types of radar signatures will be analyzed.

The analysis is applied to SAR images of IWs occurring in the northeast Atlantic on the Portuguese continental shelf (Figure 1). The Iberian margin is a good example of an area

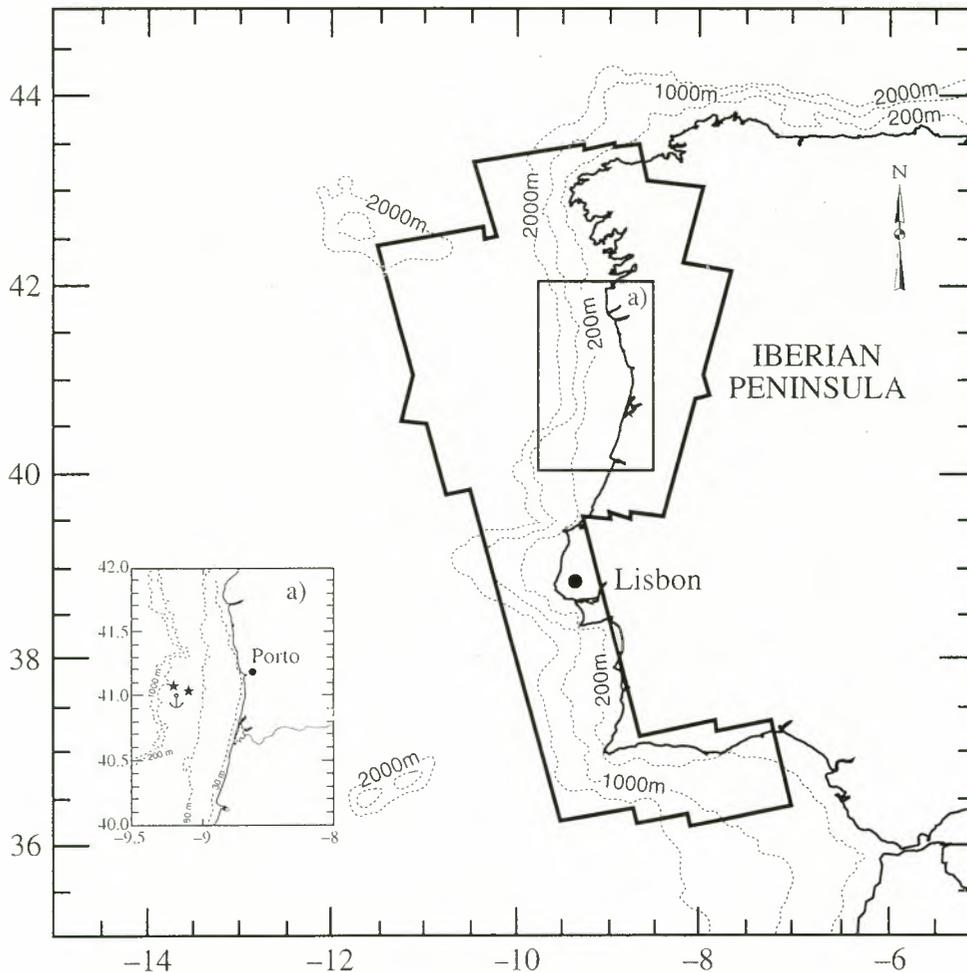


Figure 1. Area covered by ERS-1 synthetic aperture radar (SAR) images over a period of more than 2 years revealing surface manifestations of internal waves (IW) (closed polygon across the western and part of the southern Iberian margin). The bathymetry of the western Iberian margin is overlapped. The location of the in situ measurements performed in August 1994 is shown in the inset. The anchor shows the position where a moored buoy was deployed and recorded IWs concurrent with ERS-1 SAR overpasses. The star shows the positions where visual observations of IW surface signatures were made coincident with sampling of surface films and measurements of surface currents as well as the vertical structure of the water temperature.

where surface manifestations of short-scale internal waves occur. The first observations of IWs over the Iberian Shelf were made by U.S. astronauts during the Apollo-Soyuz mission in 1975 [Apel, 1979]. These were confirmed by SAR observations in 1978 during the Seasat mission [see, e.g., Alpers, 1985]. Such short-scale internal waves, commonly termed solitons in the literature [Ostrovsky and Stepanyants, 1989], propagate in groups or packets. A packet of solitons appears to originate from near the shelf break almost every tidal cycle during the summer. These wave packets are thought to be generated as part of the baroclinic (internal) tide, which results from the interaction of the barotropic tide with the shelf edge topography [see, e.g., Maxworthy, 1979; Baines, 1982]. The bathymetry of the study area is characterized by a relatively narrow continental shelf, the 200-m contour lying about 50 km offshore, and a steep continental slope. The internal tide is usually associated with longer-scale internal waves (typically of order 20 km) [Pingree et al., 1986], and the influence of such waves on the SAR signature of short-scale IWs is not considered in the present study.

The satellite observations of internal waves on the Iberian shelf prompted a detailed program of in situ measurements on the shelf near 41°N (Figure 1) in August 1994. These measurements provided valuable information about the internal waves in the region and their surface manifestations, including the distribution of surface films within the internal wave field. Some of these results are discussed in section 3. Section 4 presents a theoretical model which predicts the surface roughness modulations caused by both hydrodynamic straining and differential film damping resulting from internal waves. The model results and the in situ observations are compared, suggesting that the dual-mechanism model is the only approach which can satisfactorily account for the satellite observations, and we conclude in the final section that film damping is a process which cannot be neglected in the analysis and interpretation of SAR signatures of internal waves.

2. Character of Internal Wave Signatures on ERS-1 SAR Images

Since ERS-1 was launched in 1991, we have studied over 60 images from its C-band SAR, which confirm the ubiquity of internal waves propagating toward the coast across the continental shelf of the Portuguese coast (see Figure 1). They are imaged in a wide variety of wind conditions from April to September over several years, but no internal wave signatures were found in SAR images obtained during the winter months. This is presumably due to the development of thermal stratification during spring and summer, creating suitable conditions for the propagation of internal waves.

2.1. Examples of Double and Single Negative and Positive Signatures

A selection of ERS-1 SAR images showing IW signatures over the shelf off western Iberia are shown in Figure 2. These images are extracts from the original full SAR frames and are $\sim 12 \times 12$ km. They were enhanced with a histogram equalization for display purposes, and the spatial resolution is 25 m² (nominal full resolution for ERS-1 SAR). Radar images of IWs usually present bright and dark bands as compared to the unmodulated background clutter. Figure 2a is a typical example of such a situation, and we will refer to this signature

as being of double sign (+/-). On some occasions the dark bands are wider and much more intense than the accompanying bright bands, which may even be absent from the imagery. Examples of these cases are shown in Figures 2b and 2c, and we refer to such signatures as a single negative sign (-). Another type of signature, although more rare, is when the SAR backscatter is very low and the IWs appear as bright bands in a homogeneously dark background (Figure 2d). This signature we describe as a single positive sign (+).

In order to study these signatures quantitatively and to make comparisons with model results (section 4) the IW signature, $\delta I/I_0$, is defined as

$$\delta I/I_0 = (I - I_0)/I_0, \quad (1)$$

where I is the intensity within the IW profile (taken in the direction of the IW propagation) and I_0 is the intensity of the image background. The background is taken from a homogeneous area in the image away from the IW field and is a square of side 2.5 km (200×200 original pixels of size 12.5 \times 12.5 m). To obtain a meaningful measure of the intensity and to reduce speckle to a negligible magnitude, it is necessary to average over at least 500 pixels [Laur, 1992]. This condition was satisfied by averaging the profiles over a line of 60 pixels perpendicular to the IW propagation direction and at least 8 pixels in the IW propagation direction.

The signatures corresponding to the images of Figure 2 are shown in Figure 3. The different signature character types discussed above can be recognized in the profiles. Low wind conditions prevailed in all cases. In the first case (Figure 3a) the wind speed was between 4 and 5 m/s. For the two following cases (Figures 3b and 3c) the wind speed was between 2 and 3 m/s. For the last case (Figures 3d), which shows a single positive signature, significantly calmer conditions prevailed with wind speeds lower than 2 m/s.

It was possible to perform a statistical analysis of the IW train profiles since a relatively large data set of SAR images was processed and in each case several IW trains were usually found. The profiles were computed from a total of 62 ERS-1 SAR images, corresponding to the period from February 1992 to September 1994. The profiles were categorized according to the internal wave propagation direction (normal to the crests) relative to the radar look direction. Waves propagating in the radar look direction were termed range-propagating waves, while those propagating in the satellite flight direction were termed azimuth-propagating waves. For polar orbiting satellites, such as ERS-1, the range direction is approximately east-west, and in the case of the western Iberian peninsula this corresponds to across-shelf propagating waves. Profiles such as those shown in Figure 3 were obtained for 296 solitons propagating in the range direction but only 157 solitons traveling in the azimuth direction, suggesting that in this region, IWs predominantly propagate over the shelf toward the coast.

The normalized radar backscatter cross-section, σ (decibels) can be calculated from the intensity I , following Laur [1992]. Therefore the backscatter contrast profile $\delta\sigma = \sigma - \sigma_0$, where σ_0 is the background backscatter cross-section determined from the same homogeneous areas as I_0 , can be retrieved directly from the image intensity profiles. Within the short distances spanned by the high-frequency IW profiles it can be written as

$$\delta\sigma \approx 10 \log_{10}(\delta I / I_0 + 1). \quad (2)$$

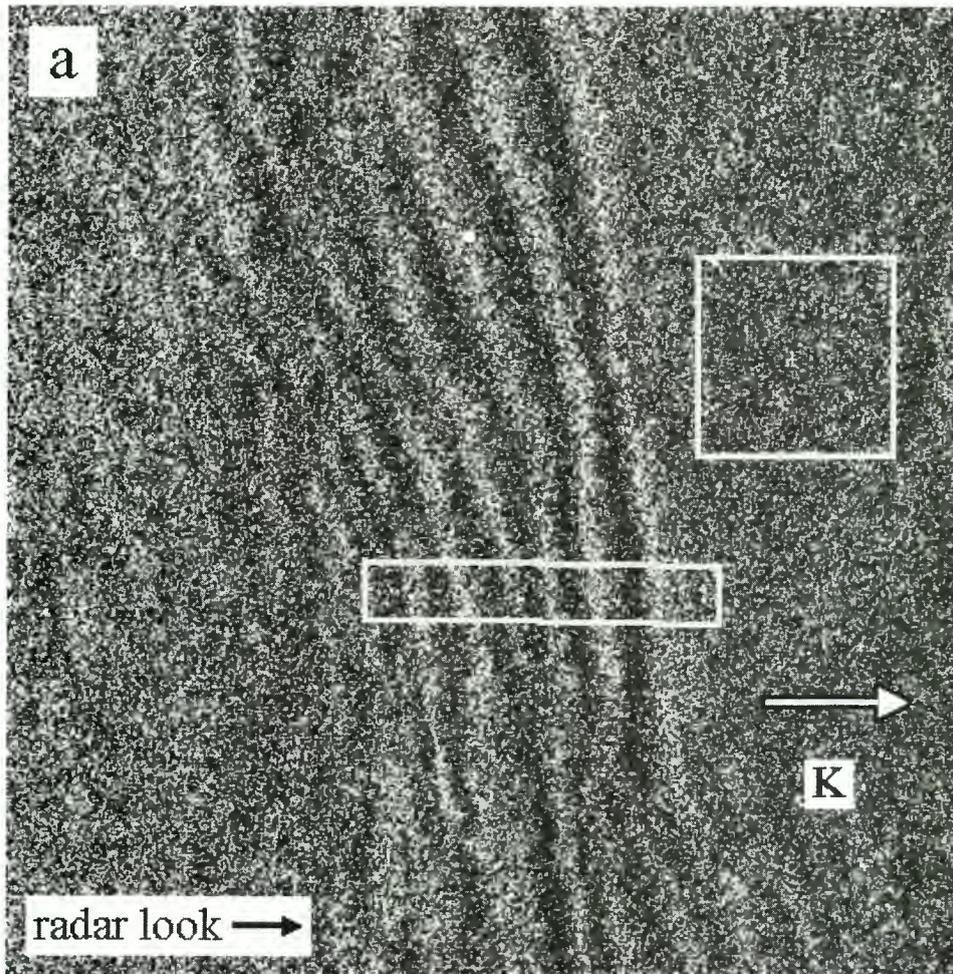


Figure 2. A selection of ERS-1 SAR images showing three types of characteristic IW signatures in the area of study: (a) double sign signature, (b,c) single negative signature, and (d) single positive signature. The squares represent the areas from which the background intensity was taken, and the rectangles represent the area over which the IW signatures were calculated. The radar look direction and IW propagation direction (wave vector \mathbf{K}) are indicated.

Histograms of backscatter contrasts for range-propagating (radar look $\kappa //$ IW wave vector \mathbf{K}) and azimuth-propagating ($\kappa \perp \mathbf{K}$) IWs are presented in Figure 4 for different wind speed conditions. The histograms show the probability density distribution of backscatter contrasts, where for each soliton the maximum and minimum values of the contrast are depicted in different shades of grey. For azimuth-propagating IWs the maximum values of contrast are biased toward zero in relation to the minimum values, indicating a dominance of single negative signature types. This is not observed for range-propagating IWs where the maximum values of contrast are generally higher, indicating a higher proportion of double sign signature types. Single positive signatures, where the minimum is greater than or equal to zero, are only observed at very low wind speeds ($V < 2\text{m/s}$) and for range-propagating IWs. For both range- and azimuth-propagating internal waves the peak of the distribution is biased to smaller absolute values as the wind speed increases. This indicates a dependence of contrasts on wind speed, which is discussed in detail in the next section. Note also that the range of the distribution for all wind speeds is wider for range-propagating waves.

2.2. Contrast Dependence on Wind Speed

The mean maximum and mean minimum contrasts were calculated as a function of wind speed for range- and azimuth-propagating IWs (Figure 5). The means were based on all IW profiles lying within wind speed intervals of 1 m/s. The wind speed was estimated from the background backscatter values using an adapted version of the CMOD4 scatterometer model [see *Scoon et al.*, 1996] for ERS-1 SAR imagery. This required an estimate of the wind direction, which was obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric model and, whenever possible, from ships of opportunity. Data from the nearest ground meteorological stations were used to confirm the model results and ship measurements. When disagreement was found, an average direction was estimated from all data available.

The absolute values of the mean maximum and minimum contrasts decrease as wind velocity increases. These absolute values are smaller for azimuth-propagating IWs than they are for range-propagating IWs. The mean maximum contrast is

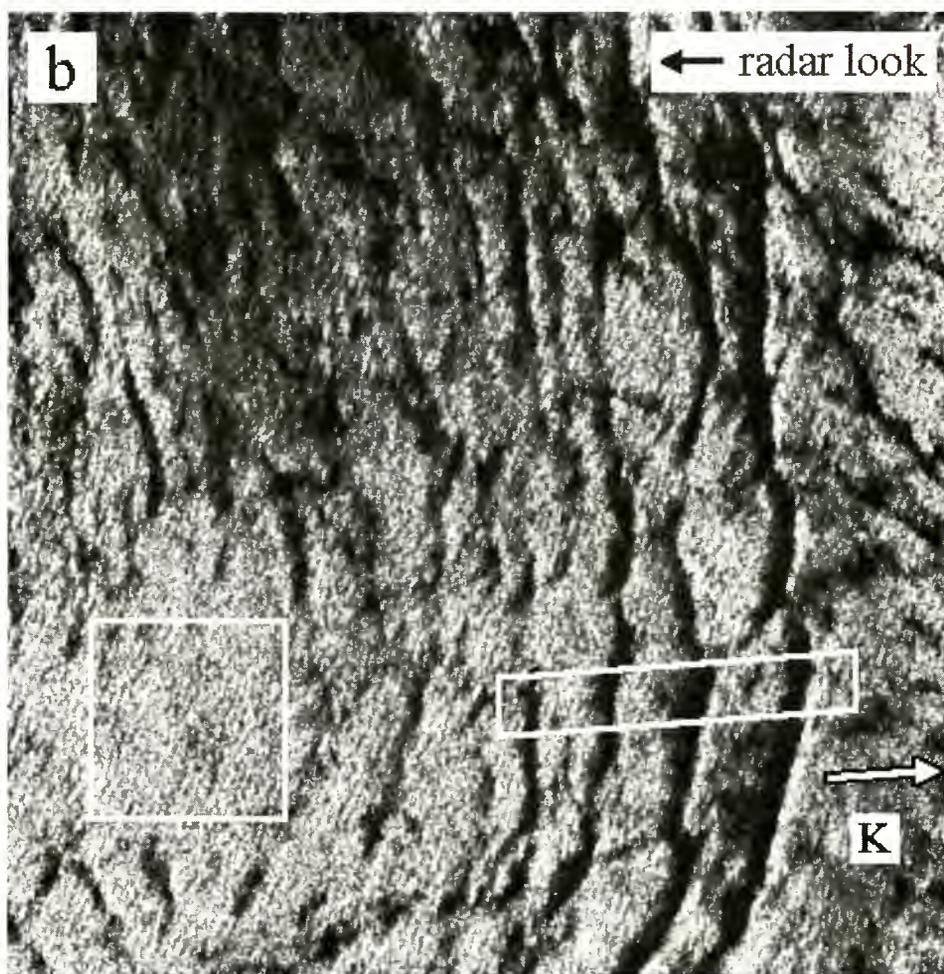


Figure 2. (continued)

always close to zero for azimuth-propagating IWs, confirming the dominance of the single negative signature type. The mean maximum contrast at wind speeds lower than 2 m/s for range-propagating waves is significantly higher than any other value. This is due to the existence of the single positive signature type, which results in far greater contrasts than the other signature types thus biasing the mean contrast. Much higher contrasts are expected for this case since I_0 is very small, being in the denominator in (1). Note, however, that the mean minimum contrast for range-propagating IWs is still lower at wind speeds less than 2 m/s than it is at higher wind speeds. This is due to the fact that single negative and double sign signatures are also observed at such low wind speeds. Finally, it should be noted that the contrast dependencies on wind are not symmetrical; that is, for a given wind speed the absolute minimum contrast is generally greater than the absolute maximum.

3. In Situ Observations of Internal Wave Surface Signatures

An oceanographic cruise was conducted in the summer of 1994 on the University of Bergen's 47-m oceangoing research vessel (R/V) *Håkon Mosby*. This was part of the European Commission (EC) Marine Sciences and Technologies (MAST) Multidisciplinary Oceanographic Research on the Eastern boundary of the North Atlantic (MORENA) project.

One of the purposes of the cruise was to make in situ measurements of IWs on the shelf near 41°N (Figure 1). Specific aims were to measure the isotherm displacements due to IWs and simultaneously to sample any associated surface films. A thermistor chain was deployed from a lightweight moored buoy before the IWs arrived and was retrieved after the largest waves had passed. In addition, the ship was held on station nearby, and profiles of the water column were made with an acoustic doppler current profiler (ADCP) and a conductivity-temperature-depth (CTD) probe. When possible, a small boat was launched to sample surface films near the moored buoy.

A fast-delivery ERS-1 SAR image from August 8 was used in near-real time to predict the location of the IW trains [see *da Silva et al.*, 1997], as an aid to planning the in situ measurements performed on August 11, 12 and 13. The surface manifestation of the waves was clearly visible from the ship during the first 2 days when there were very low wind speeds ($V < 3$ m/s). This allowed surface films to be sampled from the small boat. The wind raised in strength on August 13 to speeds typical for a northerly upwelling regime ($V \approx 10$ m/s), making it impossible to lower the small boat or easily see the surface manifestations of the IWs.

3.1. Internal Wave Measurements and Simultaneous Visual Observations

On August 11, the IW surface manifestations appeared from the ship as slick bands in a gently ruffled sea. When

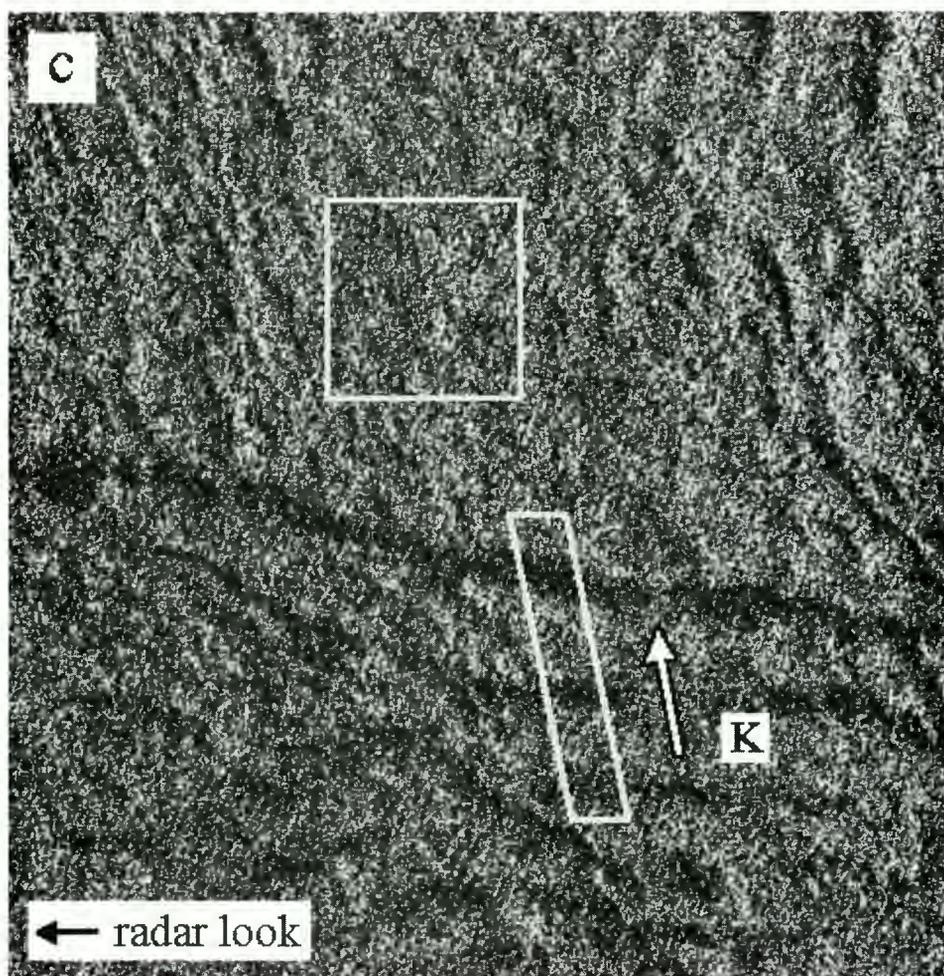


Figure 2. (continued)

measurements were performed, the wind speed average was $\langle V \rangle = 1.9 \pm 0.5$ m/s from the south, as measured on the ship. On August 12 conditions were much calmer, and the wind speed had dropped to $\langle V \rangle = 1.2 \pm 0.4$ m/s from the west. The sea surface was extremely smooth, with no discernible ripples. Because the entire sea surface now resembled a slick, it was impossible to identify IW surface manifestations with confidence. The only differentiating features were bands of very gently ruffled water, orientated in approximately the same direction as the slicks observed the previous day. These features will be referred to as "antislicks" hereafter. These antislicks proved to be the IW surface manifestations, because isotherm displacements were observed as they propagated past the observation point. The antislick bands were about 100-200 m wide and contained small wavelets, trochoidal in shape and quasi-periodic. From on board the small boat the wavelength of these wavelets was estimated to be between 20 and 30 cm and their trough to crest heights of about 2.5 cm.

The forms of the IWs observed on August 11 and 12 are represented by the depth of the isotherm which underwent the largest depression, derived from the thermistor chain data (Figure 6). In this region, amplitudes of up to 40 m were frequently observed, and typical periods of IWs are about 20 min. Individual visual observations of the sea surface roughness are indicated in Figure 6, and these are extrapolated to give a continuous record (overlying the isotherm depressions).

Slicks are generally observed over the IW troughs, consistent with previous observations explained by the film damping effect by *Ermakov et al.*, [1992]. However, this contrasts with observations in which the leading slope of the IW depression corresponded to rough water and the rear slope to smooth water [see, e.g., *Osborne and Burch*, 1980; *Hughes and Grant*, 1978], which can be explained by the straining effect.

3.2. Internal Wave Measurements and Simultaneous SAR Observations

There were two ERS-1 overpasses covering the study area on 8, August, 1994, and on August 19, 1994, i.e., before and after the field observations described in the previous section. The thermocline oscillations simultaneous with ERS-1 overpasses were recorded by a moored buoy located in the study area. The buoy position is $40^{\circ}59.92'N$ and $9^{\circ}14.54'W$, the water depth is 149 m.

On August 8 1994, at 1122 UTC, ERS-1 acquired a SAR image which covered our study area. The buoy recorded a train of large-amplitude IW solitons, which arrived at the thermistor chain at 0945 UTC (time of arrival of largest oscillation). Since the overpass occurred just about 1.5 hours after the leading soliton was recorded by the thermistor chain, we can assume that the waveforms did not undergo any significant changes during that time and that they can be used for model predictions (see below). Furthermore, we could accu-

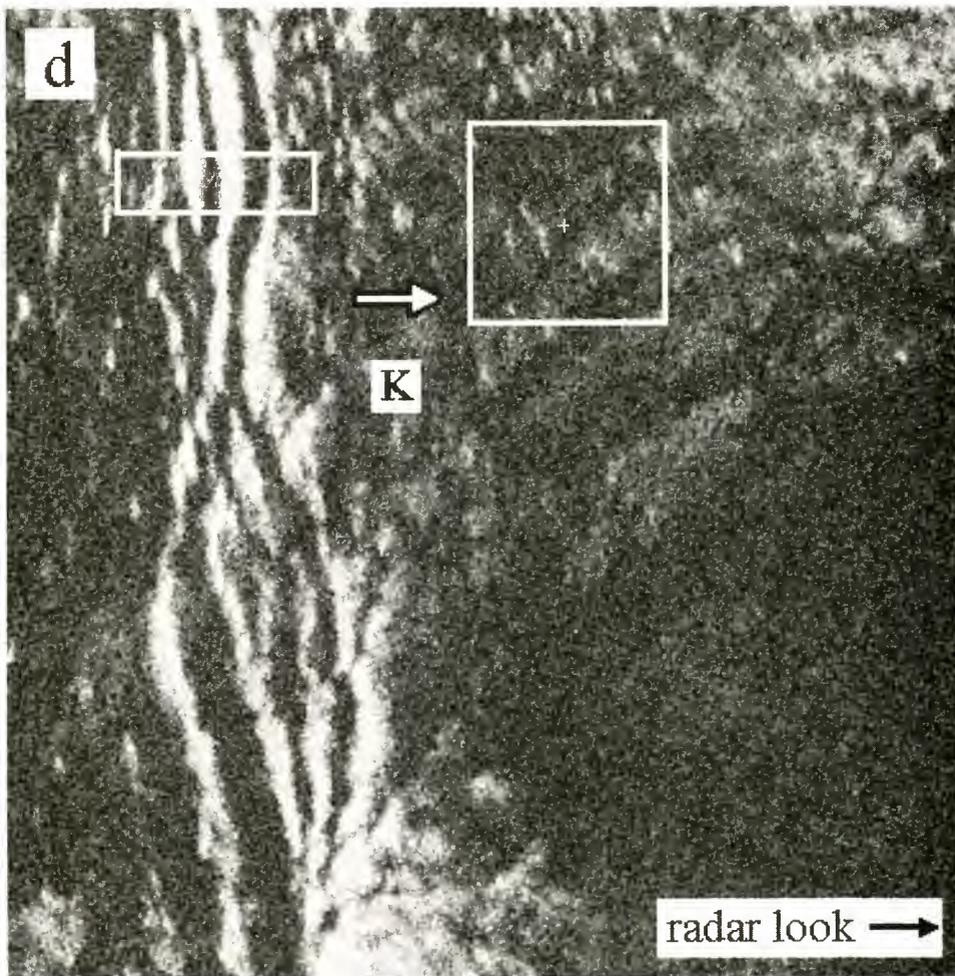


Figure 2. (continued)

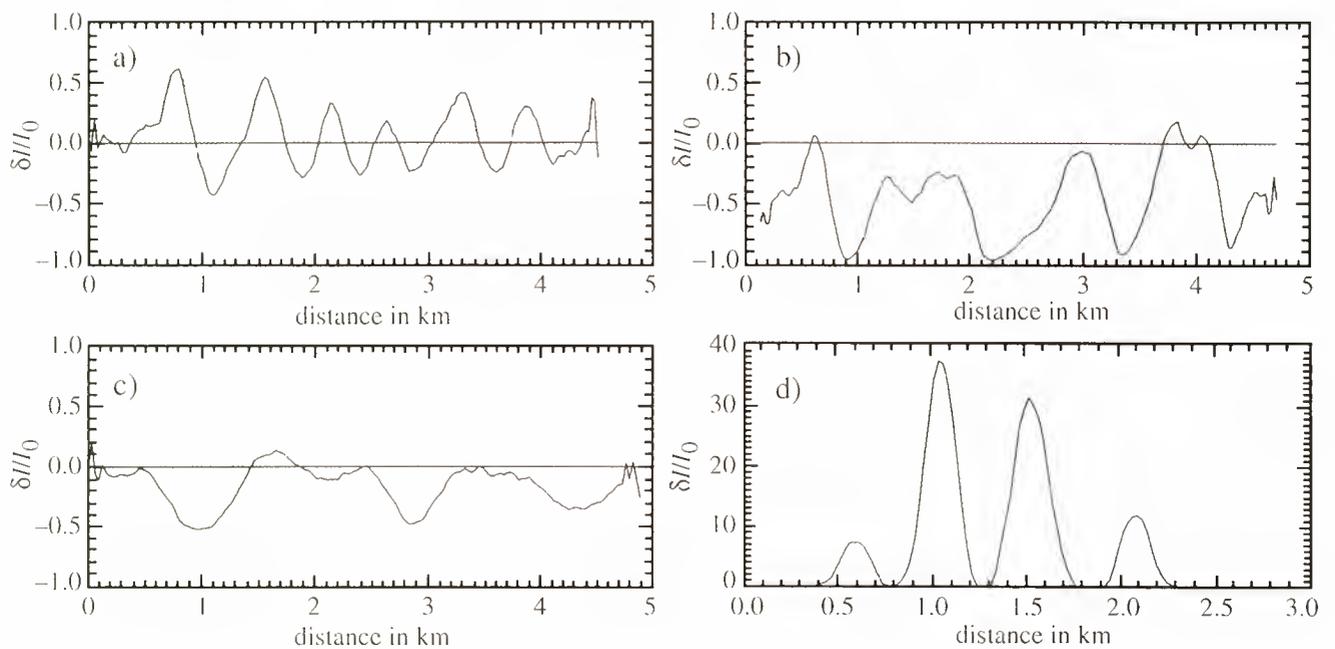


Figure 3. IW signatures corresponding to each of the images shown in Figure 2. (a) Double sign signature, (b,c) single negative signature, and (d) single positive signature. The profiles were obtained as described in section 2.1 and represent signatures taken from the rectangular boxes shown in Figure 2. Direction of IW propagation is from right to left for all cases.

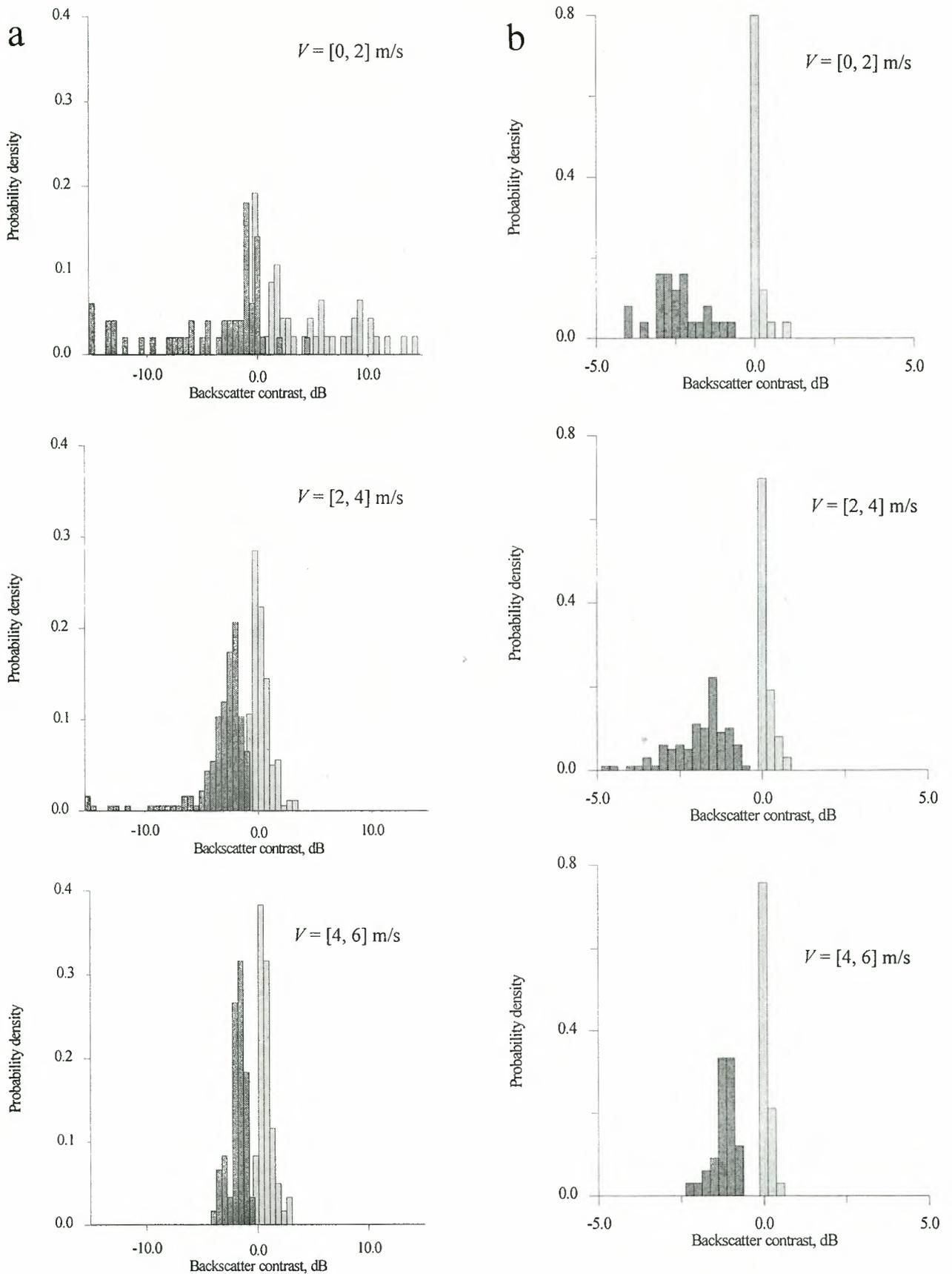


Figure 4. Distribution of backscatter contrasts for (a) range- and (b) azimuth-propagating IWs. For each soliton, maximum and minimum are computed and entered into the histograms according to the wind speed interval they represent. The probability density peak of maximums is biased toward zero for azimuth propagating waves (Figure 4b), indicating predominance of single negative signatures for these waves.

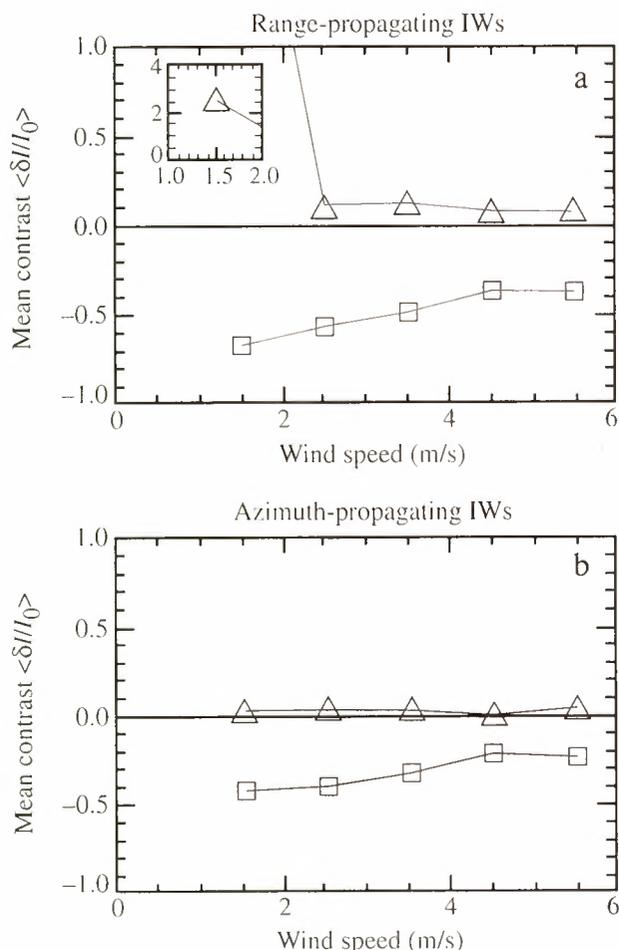


Figure 5. Mean maximum (triangles) and mean minimum (squares) contrasts as a function of wind speed for (a) range- and (b) azimuth-propagating IWs.

rately estimate the phase speed of the IW train, when measuring distances in the SAR image and locating the buoy position (the image was navigated first using reference control points). The value of the phase speed thus obtained is $C_p = 0.55$ m/s. Note that this value is in good agreement with a similar estimation based on the arrival time of the IW train on August 11, recorded by another buoy deployed at approximately the same position during the measurements described in section 3.1. This is not surprising if we assume the IWs to be generated by a tidal flow over the shelf break and if the tidal flow and stratification did not change significantly between August 8 and 11.

The SAR image from August 19 was acquired at 2250 UTC. The sea was very calm; wind speed was lower than $V = 2$ m/s. The time delay between the SAR overpass and the IW train arrival time to the moored buoy was 90 min. We can therefore assume also in this case that the waveforms have not undergone significant changes by the time they were imaged by the ERS-1 SAR.

The relevant fragment of the image from August 8, 1994, is given in Figures 7a-7c, showing that the IW signature is a double sign contrast feature for the leading wave of the IW train and single negative for the remaining two solitons. The image from August 19, 1994, presented in Figures 7d-7f demonstrates a single positive sign feature. The buoy position is marked in Figure 7 by an anchor symbol.

3.3. Film Sampling and Damping Measurements of Capillary-Gravity Waves

An attempt to sample marine films during the observations of IW slicks described above was made in order to determine whether surface films were present in the IW slicks and to estimate wave damping by films. Although highly desirable, it was impossible to perform simultaneous in situ and satellite observations because of the constraints of ship time and ERS-1 orbit restrictions.

Small circular nylon nets, with a diameter of 7 cm, were used to sample the microlayer (the thread diameter was 260 μ m; the net mesh was 0.710 mm). The nets, which were previously cleaned with pure ethanol (free of surfactant substances), were installed as a "bottom" in a plastic tube and were sunk a few millimeters in the water in the horizontal position. Then the nets were detached from the tube and placed in a precleaned container for further measurements. At the same time, a thin layer of near-surface water was sampled (about 150 mL) by a device designed to collect oil films in surface waters. The water samples were used to measure dissolved surfactants, but results are not presented here as no significant differences in wave damping were found between distilled water and water sampled from slick and nonslick areas. Both net and water samples were conserved with a saturated solution of mercuric chloride, free of any fatty organic material. The nets were also dried as part of the conservation procedure.

The ability of the nylon nets to collect surfactants from the sea surface can be characterized by the net transfer coefficient K_n , defined as the ratio of the surfactant mass M_s taken by the net and the initial surfactant mass M_o on the sea surface corresponding to the net area, i.e.,

$$K_n = M_s / M_o. \quad (3)$$

The net transfer coefficient was estimated in the laboratory as follows. First, the dependence of the wave damping coefficient for surface waves excited in a small container at a particular frequency was measured as a function of surface concentration of an artificial surfactant (oleic acid). This dependence is characterized by a typical maximum at comparably low surfactant concentrations and a small decrease of wave damping at higher concentrations (see, S.A. Ermakov et al., internal report, 1995). The initial surface concentration value before laboratory film sampling was chosen to be slightly higher than the concentration corresponding to the maximum. After film sampling with the net the damping coefficient was measured, and the amount of surfactant removed by the net was determined. The net transfer coefficient value thus estimated was 0.86.

Measurements of the damping of capillary-gravity waves were carried out after the cruise in the laboratory in a small rectangular tank (4×14.2 cm). The tank was filled with distilled water (the water depth was 1.5 cm), and damping of standing waves of longitudinal modes with frequencies 10, 15, 20 and 25 Hz were measured. The damping coefficient of surface waves due to films was taken as the difference between the measured damping and the theoretically calculated damping due to the walls and bottom of the tank (see S.A. Ermakov et al., internal report, 1995).

The preserved surfactant samples were dissolved by pouring 2.5-3 mL of purified ethanol into each container with the nets. Then the surfactant solutions were spread onto the

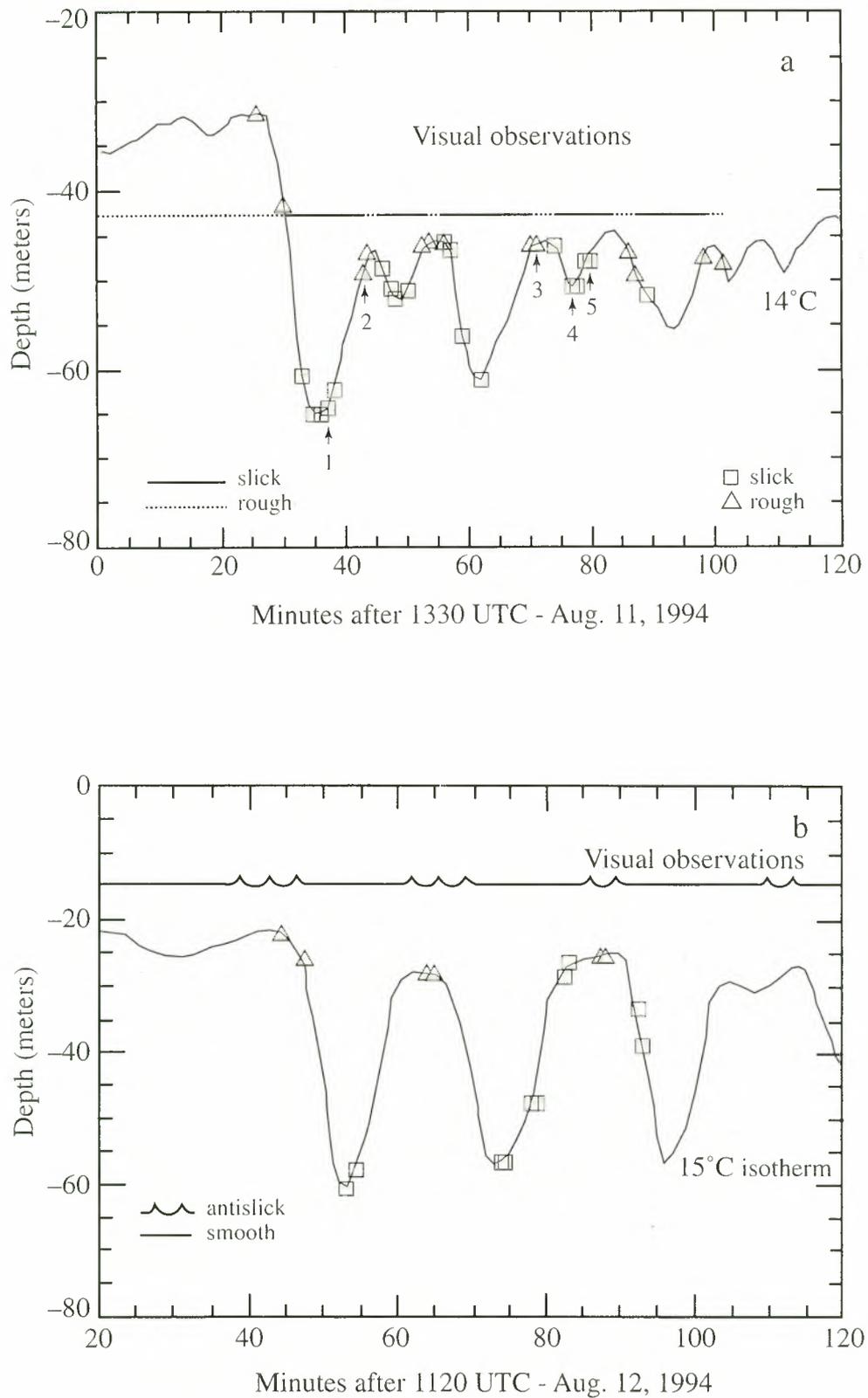


Figure 6. Isotherm depressions and surface roughness observations corresponding to the IW packets observed on (a) August 11 and (b) August 12. The triangles and squares indicate individual observations of surface roughness from which the horizontal overlaid lines were derived. The numbered arrows indicate the positions where surface film samples were taken.

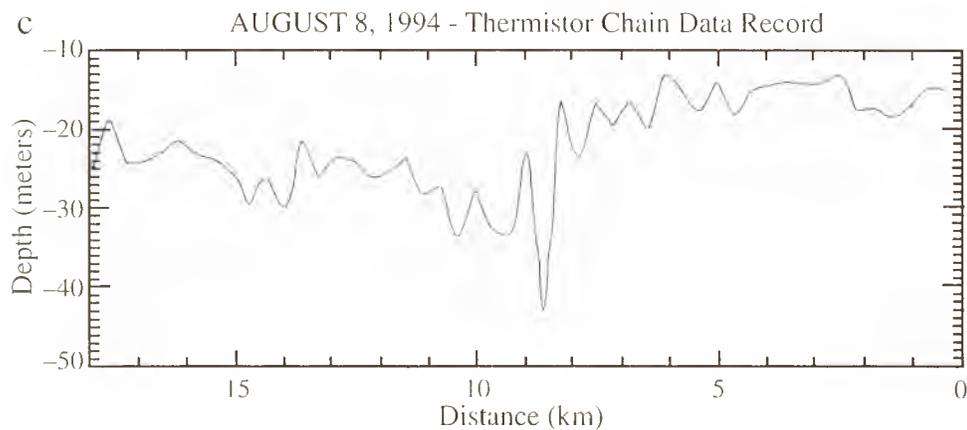
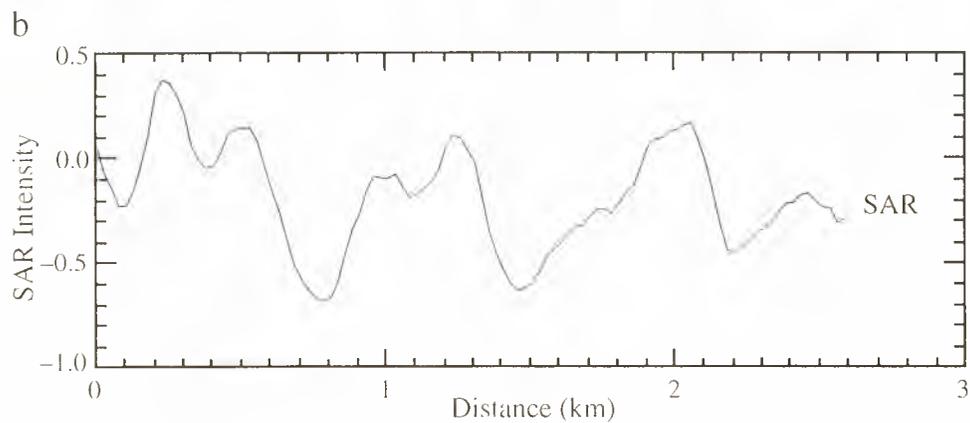
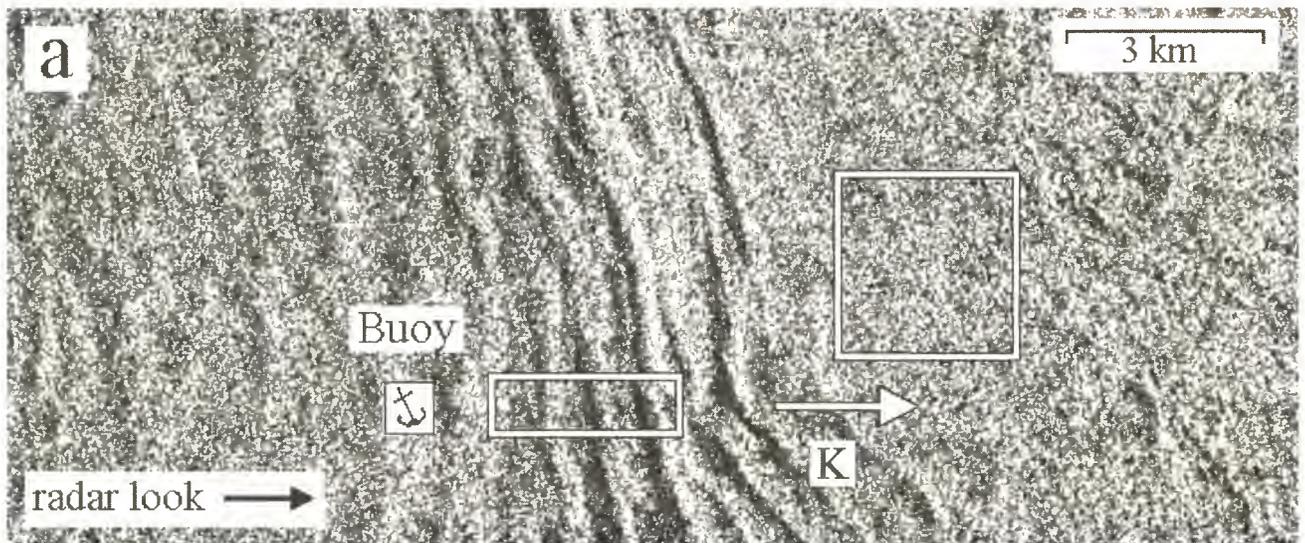


Figure 7. (a) ERS-1 SAR image acquired on August 8, 1994 (orbit: 16020, frame: 2781). The position of the moored buoy is shown by an anchor. The rectangle shows the position where the image profile shown in Figure 7b was taken, and the square is the chosen area away from the IWs to represent the background of the image. (b) The leading soliton corresponds to a double sign IW signature. (c) Thermistor chain records corresponding to the fraction of the SAR image shown in Figure 7a extrapolated to 18 km in the IW propagation direction (same scale as the SAR image). (d) ERS-1 SAR image acquired 11 days later (orbit: 16185, frame 819) during calm conditions. (e) Surface manifestations of IWs as single positive signature. The position of the buoy is indicated by the arrow, showing that the main internal wave train was “crossing” the buoy at the time of satellite acquisition. (f) The corresponding thermistor records extrapolated to the location of the image (again 18 km in length).

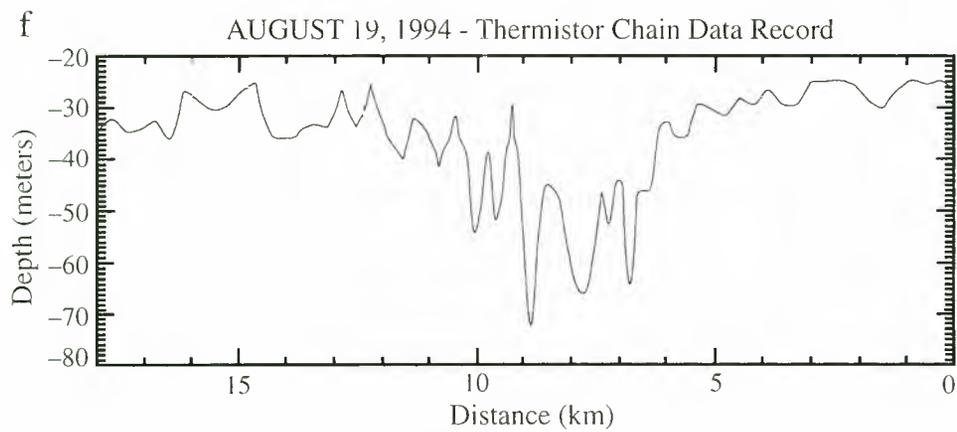
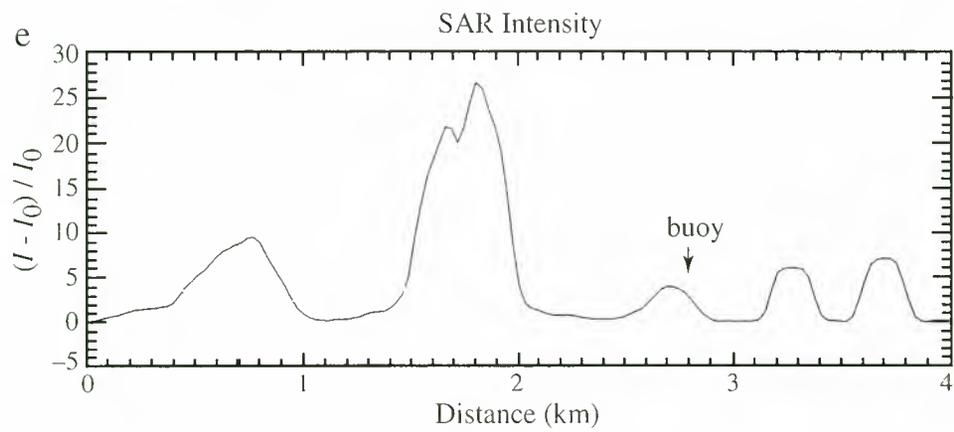
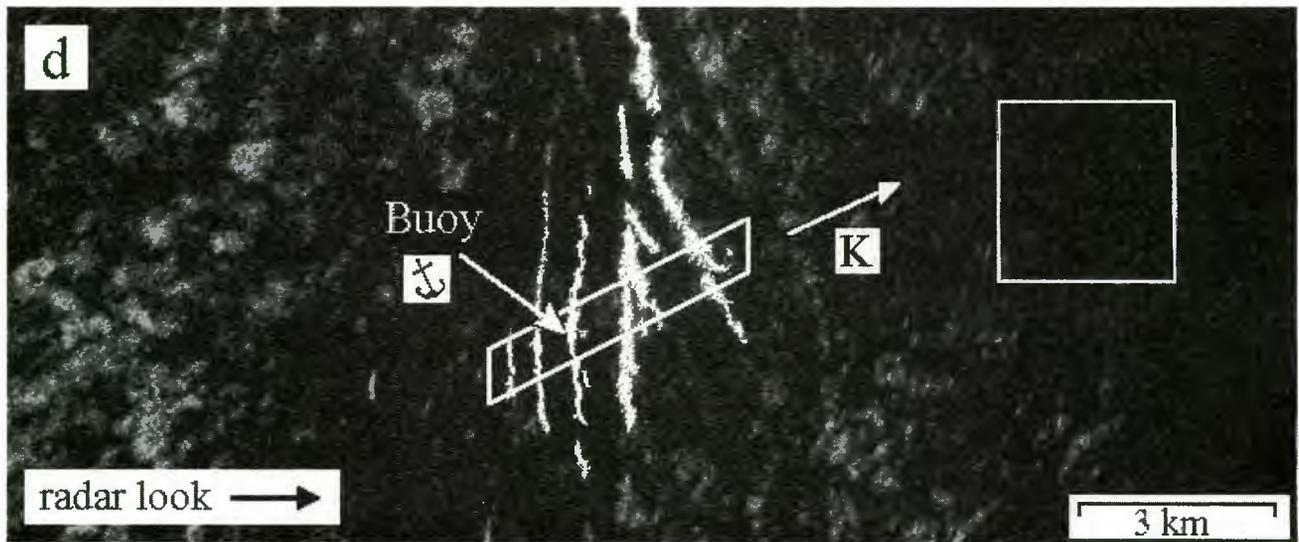


Figure 7. (continued)

tank's water surface in fractions with a pipette so that wave damping was measured at different relative surfactant concentrations. Note that ethanol dissolved in water changes the molecular viscosity coefficient, and wave damping can be enhanced even without surface-active films. Therefore the amount of ethanol used to dissolve the surfactants was chosen to be such that this effect was negligible. Special measurements of wave damping showed that the enhancing of wave damping due to ethanol did not exceed 10%.

Dependencies of the relative damping coefficient (the measured damping coefficient due to films divided by the damping coefficient due to molecular viscosity on the clean water surface) on relative surfactant concentrations are presented in Figure 8. Unit relative concentration in Figure 8 corresponds to all the surfactant amount collected by each net. After the whole amount of the surfactant solution was poured on the water surface some films were compressed with a barrier to measure wave damping at concentrations higher than unity. Taking into account the value of the net transfer coefficient and the ratio between the net and vessel areas, which is about 1/1.5, we estimated that the relative concentration value of 1.8 corresponds roughly to the original concentration of the *in situ* marine film.

The damping coefficients at frequency 10 Hz for film samples 1, 4, and 5 taken from slick areas (as indicated in Figure 6a) are presented in Figure 8a. The dependencies in Figure 8a exhibit occurrence of a maximum at low surface concentrations (lower than unity) and saturation of the damping coefficient at higher concentrations. This behavior is similar to the case of monomolecular films for pure organic substances [see, e.g., *Lucassen-Reynders and Lucassen, 1969; S.A. Ermakov et al., internal report, 1995*]. Samples 1, 4, and 5 can therefore be considered as analogous to saturated monomolecular films. Damping coefficient curves for samples taken in nonslick areas (samples 2 and 3, Figure 6a) are shown in Figure 8b. Although sample 3 also can be considered as corresponding to a saturated film, its damping coefficient maximum is reached at a relative concentration value that is higher than the two following samples taken inside the

"next" slick band (samples 4 and 5). The damping coefficient dependence for nonslick sample 2 is monotonic, showing that surfactant concentration at the point of this sample was considerably smaller than for the other mentioned samples. The results qualitatively indicate that the samples taken inside slicks correspond to higher surfactant concentrations than those outside slicks.

For the studied film samples an attempt was made to retrieve the values of the film elasticity parameter corresponding to the damping coefficient values. We assumed the model of purely elastic films, which seems to be a realistic approximation for many insoluble surfactants. The damping coefficient γ_s is given by the following expression [see, e.g., *Levich, 1962; Lucassen-Reynders and Lucassen, 1969*]:

$$\gamma_s = 2\gamma_0 \frac{1 - \frac{Ek^2}{\rho\sqrt{2\nu}\omega^{3/2}} + \frac{E^2k^3}{4\sqrt{2\nu}^{3/2}\omega^{5/2}\rho^2}}{1 - 2\frac{Ek^2}{\rho\sqrt{2\nu}\omega^{3/2}} + 2\left(\frac{Ek^2}{\rho\sqrt{2\nu}\omega^{3/2}}\right)^2} \quad (4)$$

Here $\gamma_0 = 2\nu k^2$ is the damping coefficient for clean water of ripples with wavenumber k , ν is the kinematic viscosity of water, E is the film elasticity modulus, ρ denotes water density, and $\omega(k) = [gk + (\sigma/\rho)k^3]^{1/2}$ is the eigenfrequency of gravity-capillary waves; note that here σ denotes the surface tension of water and not backscatter cross section.

When trying to retrieve the film elasticity from the measured damping coefficients, it was found, however, that the elasticity values were different at different frequencies, generally increasing with frequency. This means that the model for purely elastic films is not quite adequate to describe marine films and can be used only as a crude approximation. It is perhaps more accurate to describe marine films as viscoelastic. Estimated film elasticity values at relative concentration of about 2 are given in Table 1.

It is seen from Table 1 that the retrieved elasticities for slick sample 1 exceed the elasticity values for the nonslick

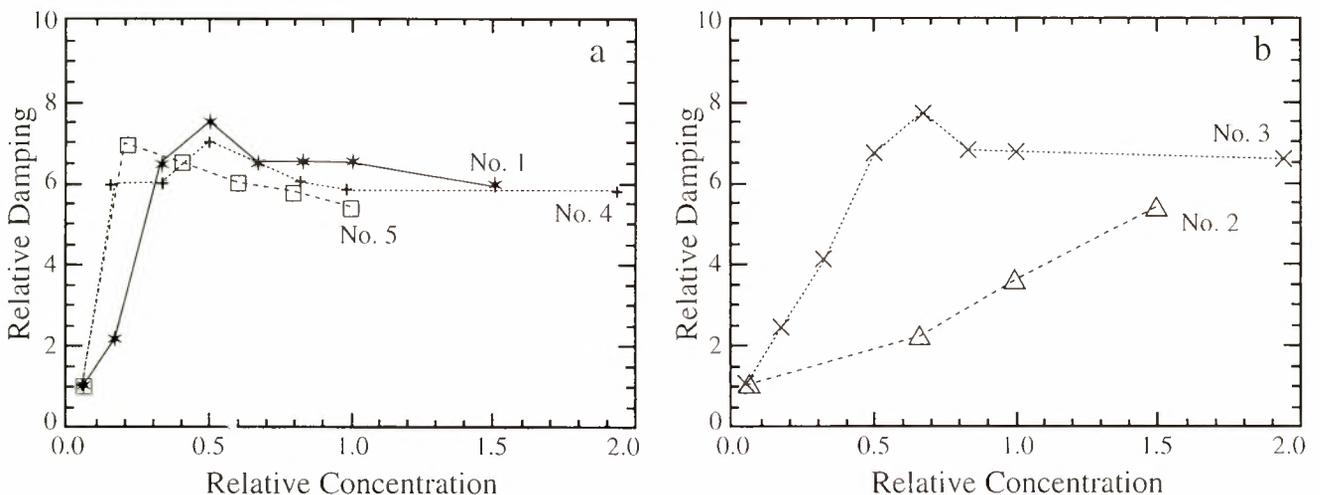


Figure 8. Results of the laboratory measurements of relative wave damping as a function of relative concentration for samples of marine films taken in the field of IWs (frequency 10 Hz). The relative concentration corresponding to the concentration in the field is about 1.8. Dependencies for films taken (a) inside slicks and (b) outside slick samples. It is seen that samples taken from slick areas present saturated dependencies. Sample numbers are as in Figure 6a.

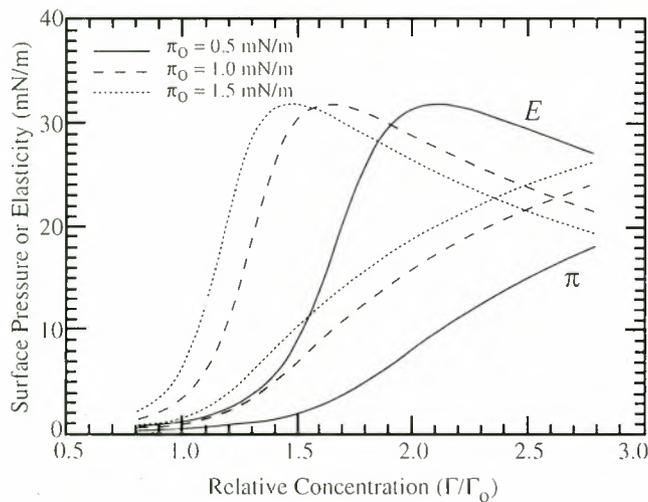


Figure 9. Dependence of elasticity E (and surface pressure π) on relative surfactant concentration variations for three different values of the unperturbed surface pressure π_0 . Typically, for the IWs measured the maximum relative surfactant concentration does not exceed 3 times its unperturbed value away from the IW field. It can be seen that for contaminated surfaces (high values of unperturbed surface pressure $\pi_0 = 1.5$ mN/m) elasticity reaches its maximum at lower relative surfactant concentrations.

waves by wind under a wide range of conditions. It proved to be adequate to describe excitation of waves at centimeter-decimeter scale at very low wind speeds, and therefore it was adopted for this model. The directional dependence of β on wind is not known and in (14) is taken, quite simply, to be the component of u_s in the direction of propagation of the surface waves.

5. Model Results and Discussion

5.1. Variability of Film Parameters Over IW Trains

One of the inputs to the model is the surface current U . Measurements of the subsurface velocity field were only available for the wave packet observed on August 12, 1994, so that for the other cases, U was estimated from the records of thermocline oscillations. For IWs with wavelengths essentially exceeding the depth of a thermocline a simple two-layer model of stratification can be used which gives

$$u = C_p \frac{\eta}{h_1 + \eta}, \quad (16)$$

where $\eta = \eta(x - C_p t)$ is the depression of the interface and h_1 is the undisturbed thickness of the upper layer (thermocline depth). An appropriate choice of h_1 is the level at which the vertical velocity of a mode one long linear internal wave reaches a maximum. This can be determined from a modal analysis of a density profile, which can be derived from temperature using an appropriate salinity-temperature relationship for the region in question. In practice, the average density profile over a tidal cycle should be used to account for the vertical fluctuations due to the internal tide.

It is assumed that the interface displacement η is well represented by the depth d of the isotherm that undergoes the largest depression and is written $\eta = d_0 - d$, where d_0 is the

depth of the isotherm at which η and thus u are zero. Because only the current fluctuations due to the short-scale solitons are to be considered, while those due to longer-scale internal tide are neglected, d_0 is taken as the envelope of the isotherm depth d . This will ensure that $u = 0$ at the crests between each individual soliton.

The method for estimating the surface velocity field from the temperature field was tested for the case where both measurements were available on August, 12. Because of the short duration of the record the temperature profile could not be averaged over a whole tidal cycle, so it was assumed that the initial conditions before the waves arrived were representative of this average. A value of $h_1 = 25$ m was determined accordingly. The ADCP velocity measurements were averaged over the top three bins within 4 m to 12 m water depth and smoothed. Because the measured near-surface velocity includes the contributions due to the long-term isotherm fluctuations of the internal tide as well as those due to the solitons, the envelope of the velocity profile was subtracted from the observations, so that at the crests of the solitons the local velocity U is always assumed zero. There is good overall agreement between the predicted and observed near-surface soliton current speed (Figure 10), with the predicted currents lying within the ADCP observations. Given this, it seems reasonable to use (16) to estimate the surface currents in the absence of ADCP observations.

In the following calculations, once the currents are retrieved, the surfactant concentration variations follow from (6) and (7), and the contrasts (film, kinematic, and combined) of the energy spectrum at a given wavenumber can be estimated. As an example, model calculations of the IW current velocity and corresponding surfactant concentration variations, film pressure, and the elasticity values over the IW train of August 11, 1994, are presented in Figure 11. Calculated variations of the wavenumber spectrum are shown in Figure 12. In the derivation of the film elasticity variations we have used $\pi_0 = 1.0$ mN/m for the unperturbed surface pressure, which seems to be realistic for not highly contaminated water. As no SAR image was acquired by ERS-1 that day, we based our estimate of IW phase speed on an average of the value

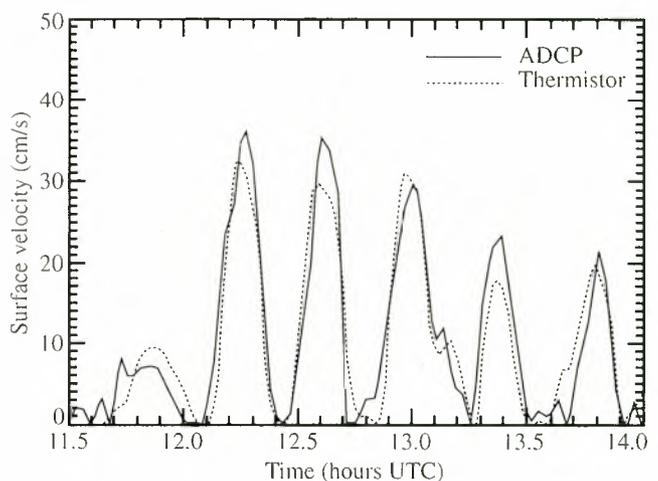


Figure 10. Time series of observed and predicted near-surface soliton current speed for August 12, 1994. Predictions are evaluated from (16) and measurements are evaluated from acoustic doppler current profiler (ADCP).

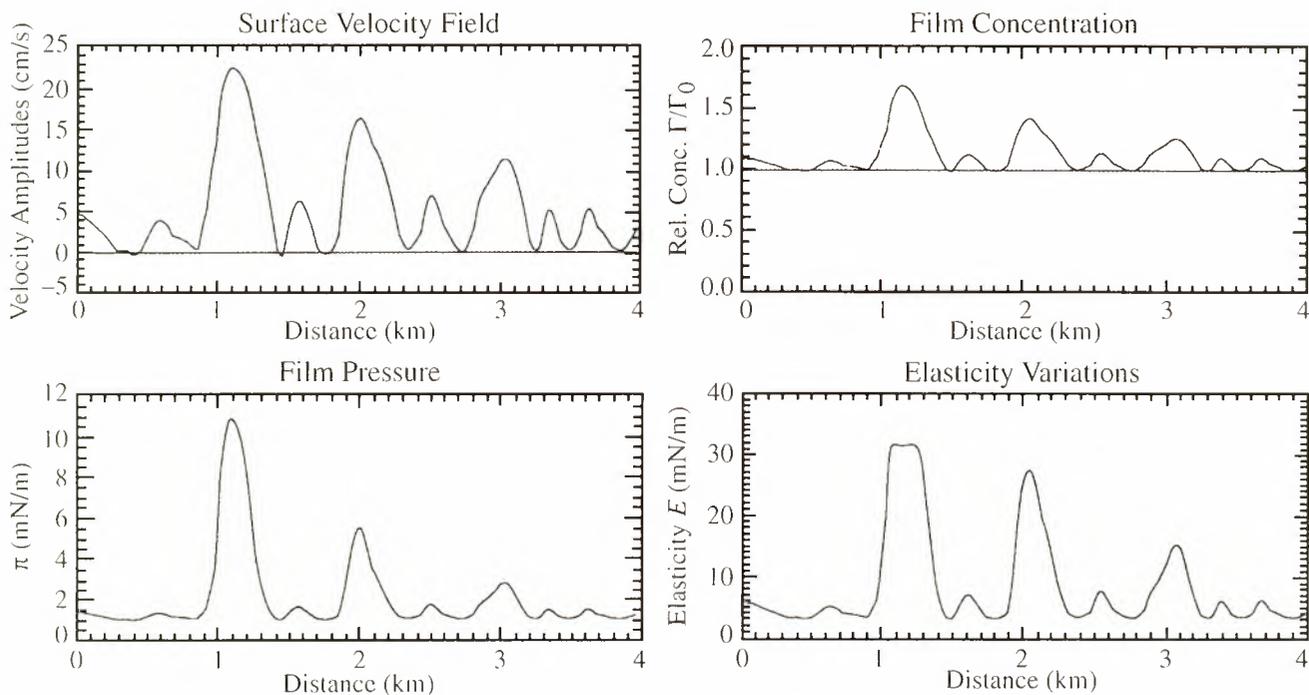


Figure 11. Calculated IW current velocity for thermistor chain measurements of August, 11, 1994, when surface films were sampled from the sea surface. Corresponding surfactant concentration variations, film pressure, and elasticity values are also shown (see also Figure 12).

observed 3 days before and 1 day later, when another wave train was followed along the shelf, and its phase speed was determined. The value of C_p thus estimated is 0.55 m/s. The model calculations of surfactant concentration and surface pressure for the analyzed case indicates ranges of $\Gamma/\Gamma_0 = 1-2$ and $\pi = 1-12$ mN/m. These are within typical values of π previously measured in slicks [Garrett, 1965; Ermakov et al., 1992]. The elasticity values qualitatively correspond to those

obtained experimentally (see section 3.3) and are also consistent with those previously presented by Ermakov et al. [1986] and Bock and Frew [1993]. From Figure 12 it is seen also that the visually recorded location of slicks is in good correspondence with the calculated depressions of the wavenumber spectrum.

5.2. Variations of Wind Wave Spectrum Over IW Trains

In this section the occurrence of different types of spectrum contrasts (double and single sign) is examined by applying the model to observations of IW signatures, in particular to the IW trains imaged by SAR on August 8 and 19, 1994.

5.2.1. Double Sign Spectrum Contrast. For the case of the image from August 8, 1994, the wind velocity estimated from the SAR image using the combination of the CMOD4 and ECMWF models is about 3 m/s from 210°. Retrieved IW current velocity and corresponding combined contrasts at wavelengths of $\lambda = 7$ cm and $\lambda = 21$ cm are given in Figure 13. Comparing Figure 13 with Figure 7 we can see that spatial variability of the calculated relative spectrum and the observed backscatter are in good correspondence. Predicted values of the spectrum variations are stronger than the observed backscatter variations for centimeter-scale waves but correspond well to the observations for decimeter-scale waves. Note that the contrast in the centimeter-scale range is predominantly negative (single negative sign contrast), corresponding to strong wave damping over IW troughs; the film effect for centimeter-waves dominates. For decimeter-scale waves, as can be seen from Figure 13, both wave intensification and damping occur due to a combination of the film and straining effects (a double sign contrast). Our calculations show that the higher the film pressure π_0 is, the stronger is the role of films in the total contrast, and the contrasts also become predominantly negative for decimeter-scale waves.

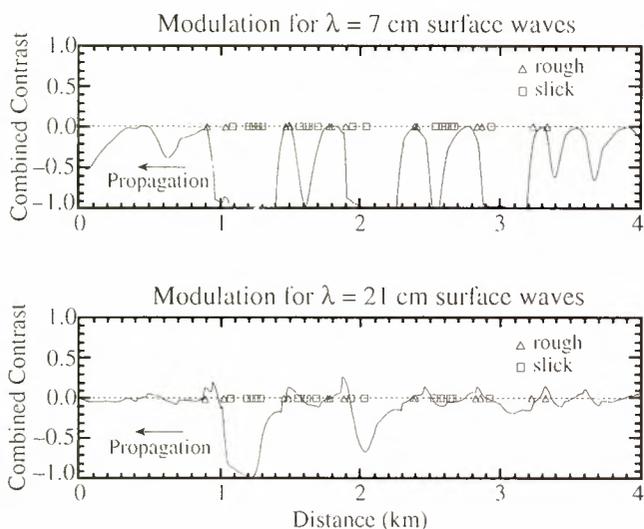


Figure 12. Model results for waveform observed on August 11 which is shown in Figure 6a. Contrasts are presented for surface waves at (a) $\lambda = 7$ cm and (b) $\lambda = 21$ cm. The value of the unperturbed surface pressure used to run the model is $\pi_0 = 1.0$ mN/m. The visual observations of surface roughness are overplotted, and it can be seen that there is overall good agreement with the model predictions (see also Figure 11).

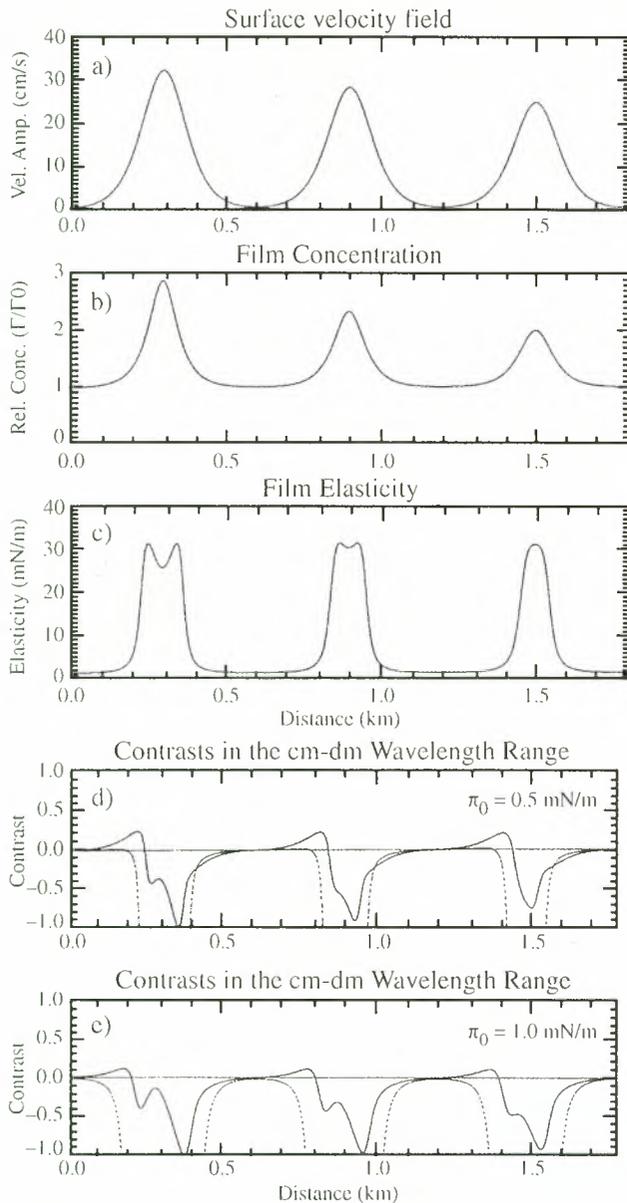


Figure 13. Model simulation for August 8, 1994, based on the in situ measurements. (a) Simulation of the surface velocity field; (b) film concentration variations; (c) correspondent film elasticity; (d) spectrum contrasts for surface waves with 7 cm (dashed line) and 21 cm (solid line) when the unperturbed film pressure is small ($\pi_0 = 0.5$ mN/m); (e) same as Figure 13d but for higher surfactant concentration ($\pi_0 = 1.0$ mN/m). It can be seen that the effect of the increasing film pressure is to reduce the positive part of the decimeter-scale spectrum, transforming double sign into single negative sign. The wind speed used in the simulation is $V = 3$ m/s.

5.2.2. Strong Positive Sign Spectrum Contrast at Very Low Wind Speeds ($V < 2$ m/s). The SAR image of August 19, 1994 (see Figures 2d and 7d-7f), corresponds to the single positive IW SAR signature at very low wind conditions; wind velocity on this day was estimated to be lower than 2 m/s. At such low wind velocities the total growth rate $\chi = \beta - \gamma$ in our model is negative for some wavelengths so that corresponding waves cannot be sustained by wind. Dependencies of χ on wavelength at different wind velocities are plotted in Figure

14 at the film pressure value $\pi_0 = 1.0$ mN/m. Figure 14 shows that at wind speeds $V < 0.5$ m/s, there is no excitation of waves shorter than 50 cm. With the increase of wind speed, the excitation of decimeter waves starts; in particular, a threshold wind velocity for the wave excitation in the wavelength range $10 < \lambda < 50$ cm is about 1 m/s. *Donelan and Pierson* [1987] also claim that the first waves to be amplified and to reach an equilibrium state when the wind is very light are essentially gravity waves with wavelengths about 20 cm. However, it is interesting to note that according to this mechanism the threshold wind velocity is higher for shorter waves. For example, surface waves with a wavelength of 7 cm (which is the Bragg wavelength for ERS-1 SAR) are excited only if $V > 1.5$ m/s. This is due to the effect of surface films, which are more effective in damping capillary-gravity waves in the centimeter-range than gravity waves in the decimeter-range of the spectrum. Although this may seem contradictory to observations by *Kahma and Donelan* [1987], which indicated that the first waves to grow have wavelengths 5-10 cm at an inception wind speed of 0.7 m/s, it follows when the effect of films are taken into account.

For the case of very low wind speeds, contrast values were calculated only for decimeter-scale waves for which wind velocity is over threshold. Figure 15 presents results of modeling of the spectrum variability for decimeter-scale waves ($\lambda = 21$ cm). In the calculations the input velocity field was estimated from the thermistor measurements acquired on August 19 concurrent with the ERS-1 SAR image. The wavelength and amplitudes of the wave train which shows up in the image were retrieved from the original thermistor records. Simulations of surface velocity and film parameters were assumed to be the same as for the case of August 8, shown in Figures 13a, 13b, and 13c. A value of unperturbed film pressure of $\pi_0 = 0.5$ mN/m was assumed for the calculations. Decimeter-scale waves are intensified over IW crests due to the straining effect. One can see that the positive contrasts obtained over IW crests are very strong; this is due to the small values of $\beta_0 - \gamma_0$ in formulas (9) and (10) for the contrasts. Note, however, that our goal for the case of very low wind speeds is simply to make a qualitative comparison of our SAR and visual observations with the model. We stress that at such low wind speeds the model is a poorer approximation than for most of the wind speed range. For the case of near-threshold wind our visual observations were of the form of antislicks on an otherwise calm sea. Similar observations had been made by *Gargett and Hughes* [1972], but no SAR and film measurements were obtained at that time. It is important to stress that we could not discern between the background wave field and the slicks over IW troughs, where $\beta - \gamma$ can become negative because of the increase of film elasticity. For this reason in our model we have set $N \approx N_0$ when $\beta - \gamma < 0$, and consequently the minimum contrast value in this case is zero and not -1, as in the case when the wind speed is considerably higher than threshold. We base our assumption that $N \approx N_0$ (equivalent to a noise spectrum N_{noise}) in the fact that the background image intensity I_0 is equal to the noise equivalent SAR floor (I_{noise}) and is a finite value. At near-threshold winds, when $\beta_0 - \gamma_0 \approx 0$, the IW image profile (I) cannot be smaller than I_{noise} , and that is the reason why we observe single positive contrasts. Following similar arguments, and in order to obtain finite contrasts, we assume the existence of an infinitesimal (but not null) noise wave action

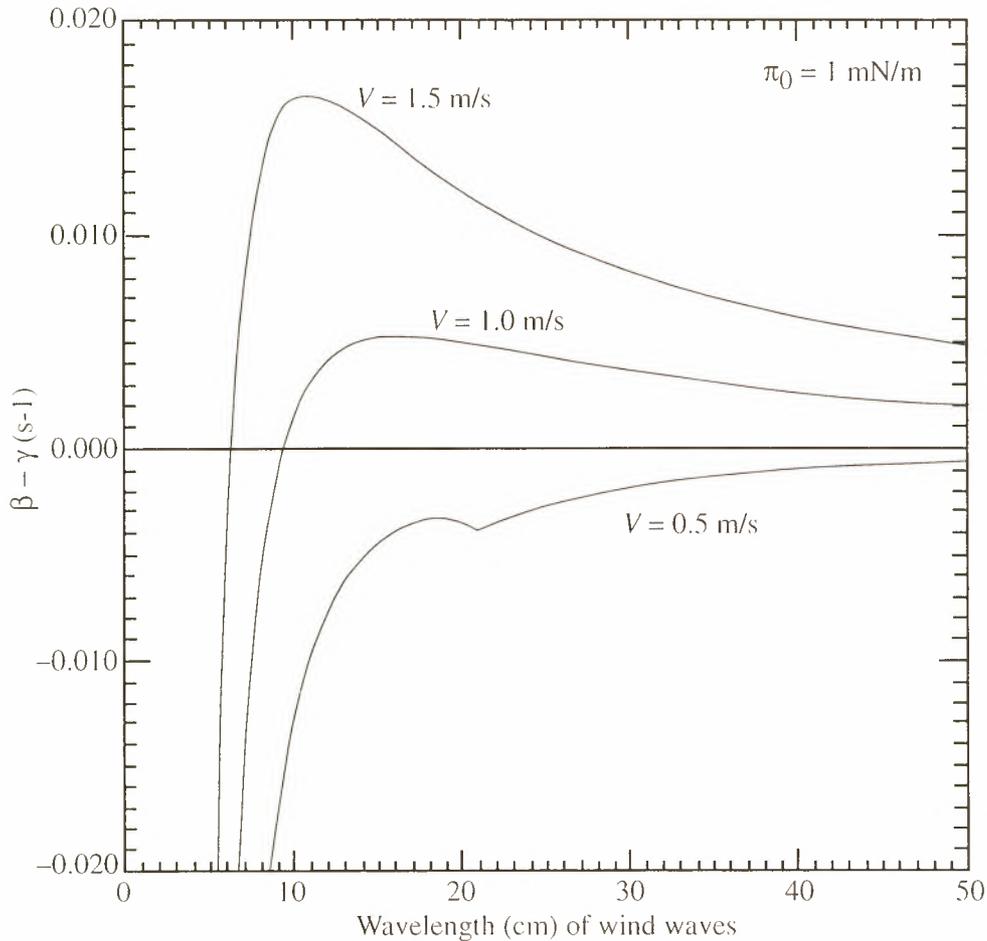


Figure 14. Behavior of the dependence of threshold parameter $\chi = \beta - \gamma$ as a function of surface wavelength for near-threshold wind speeds ($V = 0.5\text{-}1.5$ m/s). It can be seen that Bragg waves corresponding to the ERS SAR begin to be excited only at wind speed $V = 1.5$ m/s, whereas decimeter-scale surface waves are excited at lower wind speeds.

spectrum ($N_0 = N_{\text{noise}}$) of the ambient sea at near-threshold wind speeds. It seems reasonable to assume that such a noise spectrum should not be null since it is difficult to conceive a completely flat sea surface, even when conditions are calm and the wind is too low to be measured.

The model was also run for the observations made on August 12, 1994, when the wind speed recorded on board the ship was $V = 1.2$ m/s and “antislick” bands were seen as the surface manifestations of the IWs in a smooth background.

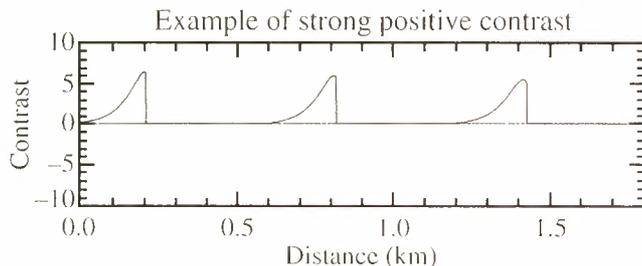


Figure 15. Model result of contrast for surface waves with wavelength of 21 cm corresponding to the in situ observations dated August 19 and shown in Figures 7d-7f. The value of the unperturbed surface pressure used to run the model is $\pi_0 = 1.0$ mN/m. The wind speed used in the calculation is 0.7 m/s. Strong positive contrasts result from the very small value of the unperturbed wave spectrum.

The model results for variations of the decimeter wave spectrum ($\lambda = 28$ cm) presented in Figure 16 qualitatively agree with the visual observations of August 12, when intensified 20-30 cm waves were observed in antislicks over IW crests (see also Figure 6b). This can account for the single positive signatures on SAR images (see Figures 7d, 7e and 16). However, in the low wind speed case, attempts to compare modeled and SAR-observed contrasts quantitatively are not in the scope of this paper.

5.3. Interpretation of SAR Contrasts Over Internal Wave Trains

To interpret SAR signatures of IWs, one should ideally consider a model of radar backscattering. Radar backscattering at low to moderate incidence angles, at which SARs operate, can be considered to result from integration of discrete contributions from surface waves in a wavelength range from the Bragg wavelength λ_b to several λ_b and has to date been most successfully described by a Kirchoff model [see Holliday *et al.*, 1986]. This range in the case of ERS-1 SAR is from 7 cm to about 20 - 30 cm. We thus can conclude that SAR backscatter contrasts result from the spectrum variations in the centimeter-decimeter wavelength range, and we interpret our results in this general sense.

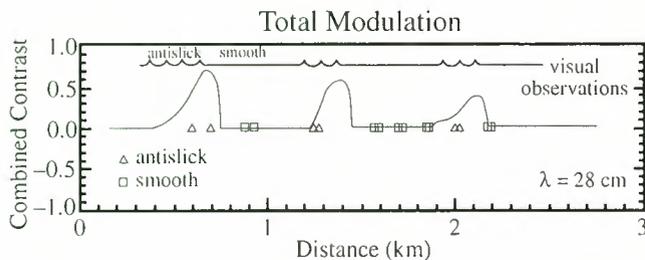


Figure 16. Model simulation for the case of IW train measured on August 12, 1994, at very low wind speeds showing single positive signatures. The model simulation is run for surface waves with 28 cm. The visual observations are reproduced from Figure 6b.

At overthreshold winds ($V > 2$ m/s) centimeter and decimeter waves are excited, reaching an equilibrium state in the background (unperturbed by the short-period IWs), which gives rise to the wave spectrum capable of producing the measured radar backscatter. The resulting backscatter contrast is, in general, double sign (positive/negative) and the contrast value is determined by the relation between negative spectrum variations in the centimeter wavelength range and negative/positive variations in the decimeter wavelength range. In the particular case of azimuthally propagating IWs, positive variations are absent in our model, and the backscatter contrast is expected to be negative, in agreement with the observations.

At very low wind velocities, only slightly exceeding the threshold wind speed for some decimeter-scale waves, there can be positive variations of the wind wave spectrum. The corresponding very high contrasts result from the very low level of the background spectrum N_0 (see definition (8)). Similarly, the positive backscatter SAR contrasts (defined by (1)) obtained at near-threshold wind speeds (see sections 2 and 3) are very strong when compared with the double sign contrasts in the over threshold case. This is because the near-threshold background backscatter intensity I_0 tends to zero at such low wind speeds. For the near-threshold wind speeds the negative spectrum variations are very small (or even nonexistent) compared with the positive part, and the corresponding negative backscatter contrasts are negligible in this case.

The noise level of the SAR is the minimum cross section σ_0 that can be measured. For the ERS-1 Precision Image (PRI) the quoted Noise Equivalent σ_0 ($NE\sigma_0$) is -25 dB, but in practice this value is slightly lower (-28 dB as measured for the case of the image background of August 19, 1994). Below this point any changes in backscatter are not detected by the radar. In the case of over threshold wind speed images the background backscatter level is well above the noise level, and therefore we can observe both positive and negative variations from the mean background level. The background backscatter for the image of August 19 is at the noise level of the SAR, and therefore negative variations of the IW contrast cannot be observed. The single positive sign signature observed at such low wind speeds can be explained by the steep decrease of backscatter cross section as the wind speed falls below the near-threshold.

Given this explanation, it is still necessary to account for why the SAR intensity I_0 falls so sharply at the near-threshold wind speed. For the near-threshold wind speeds some deci-

meter-scale waves are sustained by the wind field, and centimeter-scale Bragg waves may or may not exist. In any case, at such light winds the waves may not be discernable visually as they have very small amplitudes, but it is important to stress that they are present and responsible for the noise wave spectrum (N_{noise}) discussed in the previous section. We suggest that the ambient sea spectrum is too weak to produce measurable backscatter by the SAR. The background backscatter level falls to the $NE\sigma_0$.

The single positive sign IWs are observed as a result of the very strong modulation to which the decimeter-scale waves are exposed. Presumably, this modulation is so strong that the backscatter signal of the strained surface waves over the IW crests rises above the $NE\sigma_0$. The wave field within the antislick bands is characterized by the sharply peaked, long-crested waves, which were observed traveling with crests roughly parallel to the axis of the bands. This highly nonlinear character of the wave field in the antislicks is consistent with the resonant interpretation of the internal-wave/surface-wave interaction, with preferential growth of surface waves for which the energy propagation equals the IW phase speed [see Phillips, 1973]. Gargett and Hughes [1972] observed the same phenomenon and explained it similarly. The "monochromatic" type of waves observed also supports such an interpretation. The IW phase speed was estimated to be $C_p = 0.5$ m/s, for which the resonant surface wavelength unperturbed by the IW flow is ~ 0.6 m, which is about twice the wavelength of the highly nonlinear wavelets observed. Although at present we cannot explain this disagreement, we still consider that the sharply peaked surface waves are connected with the resonant effect. The resonance accounts for the extremely strong modulation that the decimeter-waves encounter in order to produce measurable backscatter.

In the model we considered only surface waves excited by wind, which follow the dispersion relation of linear gravity-capillary waves ("free" surface waves). According to the model, free centimeter-scale waves cannot be excited at very low wind speeds (about 1 m/s), and therefore, if a simple Bragg backscatter radar model is assumed, there should be no radar signal for ERS-1 SAR. However, the improved model of radar scattering developed by Holliday *et al.* [1986] predicts that radar backscatter at low to moderate incidence angles is determined both by a Bragg wave and longer waves (practically up to 2-3 times the Bragg wavelength). Therefore, even in the absence of centimeter-scale Bragg waves, at low wind, IWs can be imaged due to decimeter-scale waves. There is yet another possible mechanism for detection of antislick bands by SAR at near-threshold winds. According to our observations (see section 2) the antislick wavelets were observed as highly nonlinear waves which can be decomposed in several harmonics, so that the 7-cm Bragg waves, for example, could be present as high-order harmonics of the basic 14-cm wave, 21-cm wave, etc. It should be stressed that these Bragg waves would not be "free" waves but bound waves propagating at the same phase speed as the basic harmonics. If we assume this hypothesis, then the strong single positive sign radar signature is to be expected and could occur even for pure Bragg scattering. To test this explanation quantitatively, the hydrodynamic model would need to describe modulation by IWs of both free and bound surface waves; this has not yet been attempted.

5.4. Contrast Dependence on Wind Speed

Since our observations of SAR contrasts show a clear dependence on wind speed, we seek for a similar dependence in the model results. For that we have calculated maximum and minimum contrasts as a function of wind speed for typical parameters of the IWs that are similar to those observed over the Iberian shelf. The parameters chosen were $C_{ph} = 0.5$ m/s, IW wavelength $\lambda_{IW} = 600$ m and maximum IW orbital velocity at the surface $U_0 = 0.35$ m/s, which are based on our SAR observations and in situ measurements (it should be noted here that we define IW wavelength λ_{IW} as the length scale over which a cycle of variations of the surface velocity is observed). Figure 17 presents those results for centimeter ($\lambda = 7$ cm) and decimeter ($\lambda = 21$ cm) waves for the cases of range ($k//K$)-propagating and azimuth ($k \perp K$)-propagating IWs. We have calculated contrasts for two values of unperturbed surface pressure π_0 , representing relatively “clean” (low contamination) ocean waters ($\pi_0 = 0.5$ mN/m) and more contaminated waters ($\pi_0 = 1.5$ mN/m) plotted as triangles and squares, respectively. The behavior of the contrast curves for centimeter surface waves is almost identical for both range- and azimuth-propagating IWs, therefore we omit the plot for azimuth IWs. The dashed lines in Figure 17 represent the kinematic contrast alone, in the absence of films.

Several contrast characteristics can be observed. The first is the single negative sign character of contrasts for centimeter-scale waves at all wind speeds (Figure 17a). This is observed both for surface pressures typical of clean and contaminated waters (triangles and squares). One can see that for low wind speeds in the absence of films the kinematic contrast is not negligible. However, even small amounts of surfactant ($\pi_0 = 0.5$ mN/m) can suppress completely the positive part of the total contrasts in the centimeter-wavelength range.

For decimeter-scale ripples ($\lambda = 21$ cm), double sign contrasts are obtained for clean surfaces ($\pi_0 = 0.5$ mN/m) at low wind speeds (2-4 m/s), but an increase in surfactant concentration reduces strongly the positive part of the contrasts. It is seen from Figure 17b that as surface pressure increases from 0.5 to 1.5 mN/m, the character of the contrasts changes from double to single negative. It should also be noted that even small amounts of surfactants are sufficient to transform symmetric double sign contrasts ($\pi_0 = 0$) into predominantly single negative contrasts ($\pi_0 = 0.5$ mN/m), for which the positive part of contrasts are significantly less intense than their negative pairs. We also note here that the mean SAR contrasts plotted in Figure 5a for range-travelling IWs lie within the contrasts predicted by the model both for centimeter and decimeter waves.

In the case when the IW propagation direction is normal to the radar look direction the backscatter from surface waves propagating in the range direction is not modulated by the straining effect. Figure 17c presents the contrasts in this situation for decimeter waves, and as expected, it is seen that single negative signatures are predicted (we do not present contrasts for centimeter waves, as they are practically the same as in Figure 17a). One should stress that the dependencies of contrasts on wind velocity in Figure 17 were calculated for typical parameters of IWs taken from in situ measurements. Note also that parameters of films used in the model were obtained from empirical pressure-area curves of marine films. It is remarkable how well the results of the

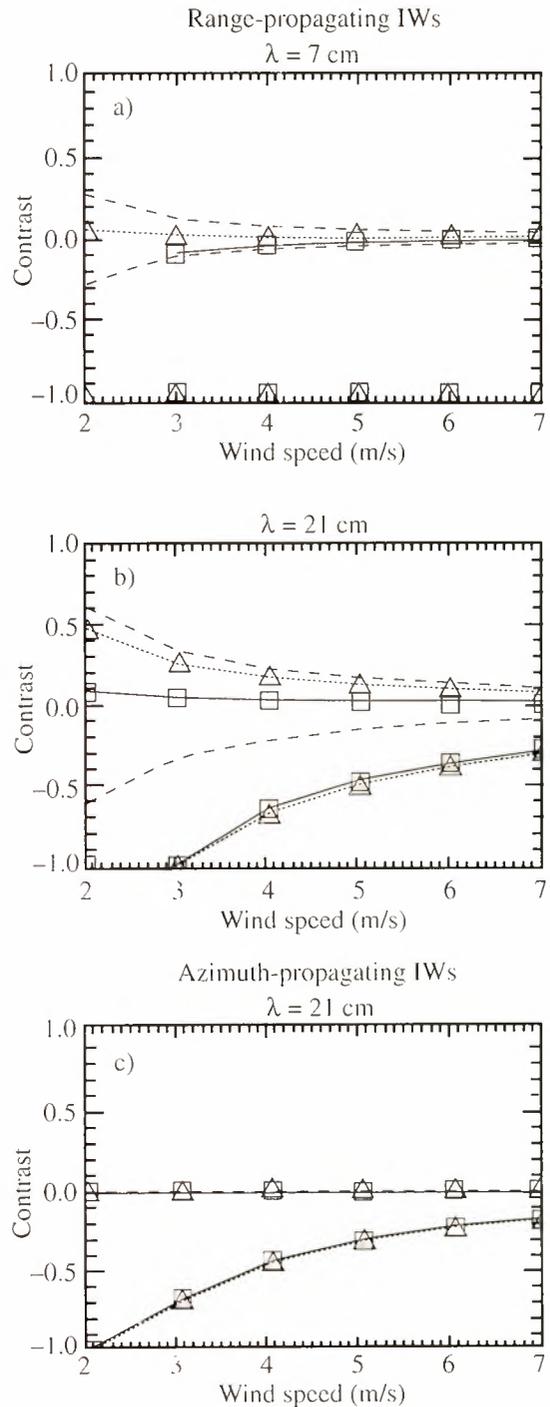


Figure 17. Model predictions of maximum and minimum contrast amplitudes as a function of wind speed for IW parameters typical of the observations in the study area. The cases correspondent of “clean” water surface are represented by the triangles ($\pi_0 = 0.5$ mN/m) and those of “contaminated” water surface by squares ($\pi_0 = 1.5$ mN/m). Pure hydrodynamic modulation in the absence of films ($\pi_0 = 0$) is represented by the dashed lines. (a) Study case of centimeter ($\lambda = 7$ cm) surface waves modulated by IWs propagating in the radar range-direction. (b) Same as in Figure 17a for decimeter surface waves ($\lambda=21$ cm). (c) Contrast modulation of decimeter waves for the case of azimuth- propagating IWs . It can be seen that single negative sign IWs are observed always in the case of azimuth-propagating IWs and for contaminated water in the case of range-propagating waves.

SAR observations presented in Figure 5b correspond to the model contrasts for decimeter waves in Figure 17c. This gives us confidence that the dependencies expressed in Figure 17 can be used to explain observed mean backscatter contrasts.

6. Conclusions

Analysis of an ERS-1 SAR image data set over the western Iberian shelf zone has revealed three types of SAR signatures of short-period IWs: double, single negative and single positive. They correspond respectively to positive/negative, negative, and positive sign variations of radar backscatter. The amplitude of the signatures depends on the wind speed, decreasing as wind speed increases. At low to moderate wind velocities (but higher than 2 m/s), double and single negative signatures are observed for range-propagating IWs; for azimuth-propagating IW the negative signatures prevails. At very low wind velocities (of order 2 m/s and less), strong single positive sign SAR signatures have been observed. This can be considered as a particular case of IW radar signatures when the backscatter background level is comparable with the radar noise level.

Field measurements of IWs simultaneous with ERS-1 SAR images were carried out during the MORENA cruise in August 1994. The three types of radar signatures were observed. Visual observations coincident with the IW measurements were of the form of slick bands in a gently roughened ambient sea when winds were about 2-5 m/s and were of the form of antislick bands when the wind speed was lower than 2 m/s. Not simultaneously with satellite, but corresponding to the two types of visual manifestations of IWs, there was in situ sampling of surfactants at different phases of the IWs. The visual observations showed the occurrence of slick bands over IW troughs, and the antislicks were located over IW crests. Measurements of the damping coefficient of capillary-gravity waves for the sampled films showed that surfactant concentration over IW troughs is higher than over IW crests. Film elasticities retrieved from the damping measurements were found to be higher inside slicks than outside and to agree roughly with previously measured values. One should note, however, that the number of samples was small and the conclusions must therefore be regarded as tentative. Further measurements and improved sampling procedures are needed.

Model calculations of the internal-wave/surface-wave interaction taking into account both the hydrodynamic (kinematic) and film effects were conducted for typical parameters of IWs and films. The hydrodynamic effect predicts positive/negative variations of the surface wave spectrum, and the film effect is responsible mainly for negative variations. The relation between the two effects depends on the surface wavelength, wind velocity, and film parameters. For decimeter-scale waves traveling at small angles to IWs and for low to moderate winds the kinematic and film effects are comparable in magnitude, and their combined effect leads to positive/negative contrasts. The ratio between these positive/negative contrasts depends on film parameters. The film contrast is dominant for centimeter-scale waves and for decimeter-scale waves in the case of azimuth-traveling IWs. At very low winds ($V < 2$ m/s) decimeter-waves can be excited in the ambient sea and are strongly intensified due to the hydrodynamic effect (resonant interaction). In this case we observe

strong positive SAR contrasts in qualitative agreement with the modeled contrast values for short wind waves.

Surfactants have a relevant role in the contrasts for centimeter and decimeter waves. As the values of surface film pressure rise, corresponding to higher concentration of surfactants in the surface microlayer, the double sign (positive/negative) contrast ratios are reduced, and the SAR signatures are expected to turn to single negative. The model can explain the three types of SAR signatures observed in the study case, and their occurrence is determined by variation of two main physical parameters: wind speed and film pressure. Overall, the consistency between observations and the model developed in this paper lead us to the conclusion that the combination of both hydrodynamic and film modulation processes is not only possible but necessary to account for the observed IW features on ERS SAR images at low to moderate wind speeds.

Acknowledgments. This is MORENA publication 16. This work constitutes part of the EU MASTII-supported MORENA project contract MAST2-CT93-0065. J.C.B. da S. gratefully acknowledges the EU Human Capital and Mobility program (ERSBCHBITC941008) for a scholarship to study ocean slicks and the Portuguese JNICT for partial support. Thanks to the Portuguese Meteorological Office (INMG) for providing wind data. We acknowledge the Instituto Geofísico do Porto and its director, J.F.D. Montenegro, for providing the ethanol used during the cruise and Mrs. K. Davies for preparation of figures. The Oceanography Group of the University of Lisbon kindly provided the thermistor chain data records. ERS-1 SAR image data were provided by ESA as part of the AO project AO2.UK123. S.E. gratefully acknowledges the Royal Society; the Winfrith Technology Centre, DERA, United Kingdom, supported his visit to the Southampton Oceanography Centre; and also, the Russian Foundation of Basic Research partly supported the work (projects 96-05-65087 and 97-02-17538).

References

- Alpers, W., Theory of radar imaging of internal waves, *Nature*, 314, 245-247, 1985.
- Amorochio, J., and J.J. DeVries, A new evaluation of the wind stress coefficient over water surface, *J. Geophys. Res.*, 85, 433-442, 1980.
- Apel, J.R., Observations of internal wave surface signatures in ASTP photographs, Apollo-Soyuz test project, *NASA Spec. Publ.*, SP-412, 505-509, 1979.
- Apel, J.R., and F.I. Gonzalez, Nonlinear features of internal waves off Baja California as observed from Seasat imaging radar, *J. Geophys. Res.*, 88, 4459-4466, 1983.
- Baines, P.G., On internal tide generation models, *Deep Sea Res.*, Part A, 29, 307-338, 1982.
- Barger, W.R., and J.C. Means, Clues to the structure of marine organic material from the study of physical properties of surface films, in *Marine and Estuarine Geochemistry*, edited by A.C. Sigleo and A. Hattori, pp. 47-67, A.F. Lewis, New York, 1985.
- Bock, E.J., and N.M. Frew, Static and dynamic response of natural multicomponent oceanic surface films to compression and dilation: Laboratory and field observations, *J. Geophys. Res.*, 98, 14599-14617, 1993.
- da Silva, J.C., D.R.G. Jeans, I.S. Robinson and T.J. Sherwin, The application of near real time ERS-1 SAR data for predicting the location of Internal Waves at Sea, *Int. J. Remote Sens.*, in press, 1997.
- Donelan, M.A., and W.J. Pierson Jr., Radar scattering and equilibrium ranges in wind-generated waves with application to scatterometry, *J. Geophys. Res.*, 92, 4971-5029, 1987.
- Ermakov, S.A., and E.N. Pelinovsky, Variation of the spectrum of wind ripple on coastal waters under the action of internal waves, *Dyn. of Atmos. Oceans*, 8, 95-100, 1984.
- Ermakov, S.A., A.M. Zujkova, A.R. Panchenko, S.G. Salashin, T.G. Talipova and V.I. Titov, Surface film effect on short wind waves, *Dyn. Atmos. Oceans*, 10, 31-50, 1986.

- Ermakov, S.A., S.G. Salashin, and A.R. Panchenko, Film slicks on the sea surface and some mechanisms of their formation, *Dyn. Atmos. Oceans*, 16, 279-304, 1992.
- Ermakov, S.A., J.C.B. da Silva, and I.S. Robinson, Role of surface films in ERS SAR signatures of internal waves on the shelf, 2, Internal tidal waves, *J. Geophys. Res.*, this issue.
- Ewing, G., Slicks, surface films and internal waves, *J. Mar. Res.*, 9, 161-187, 1950.
- Frew, N., and R. Nelson, Isolation of marine microlayer film surfactants for ex situ study of their surface physical and chemical properties, *J. Geophys. Res.*, 97, 5281-5290, 1992a.
- Frew, N., and R. Nelson, Scaling of marine microlayer film surface pressure-area isotherms using chemical attributes, *J. Geophys. Res.*, 97, 5291-5300, 1992b.
- Gargett, A.E., and B.A. Hughes, On the interaction of surface and internal waves, *J. Fluid Mech.* 52, 179-191, 1972.
- Garret, W.D., Collection of slick forming materials from the sea surface, *Limnol. Oceanogr.*, 10, 602-605, 1965.
- Gasparovic, R.F., J.R. Apel, D.R. Thompson, and J.S. Tochko, A comparison of SIR-B synthetic aperture radar data with ocean internal wave measurements, *Science*, 1529-1531, 1986.
- Hasselmann, K., Weak interaction theory of ocean waves, in *Basic Developments in Fluid Dynamics*, vol. 2, edited by M. Holt, pp. 117-182, Academic, San Diego, Calif., 1968.
- Holliday, D., G. St-Cyr, and N.C. Woods, A radar ocean imaging model for small to moderate incidence angles, *Int. J. Remote Sens.*, 7, 1809-1834, 1986.
- Hughes, B.A., The effect of internal waves on surface wind waves, 2, Theoretical analysis, *J. Geophys. Res.*, 83, 455-465, 1978.
- Hughes, B.A. and H.L. Grant, The effect of internal waves on surface wind waves, 1, Experimental measurements, *J. Geophys. Res.*, 83, 443-454, 1978.
- Kahma, K.K., and M.A. Donelan, A laboratory study of the minimum wind speed for wind wave generation, *J. Fluid Mech.*, 192, 339-364, 1987.
- Keller, W.C., and J.W. Wright, Microwave scattering and the straining of wind-generated waves, *Radio Sci.*, 10, 139-147, 1975.
- Laur, H., Derivation of backscattering coefficient σ^0 in ERS-1 SAR PRI products, in *ERS-1 SAR Calibration, ESA Rep. 1*, Eur. Space Agency, Paris, 1992.
- Levich, V.G., *Physicochemical Hydrodynamics*, 591 pp. Prentice-Hall, Englewood Cliffs, N.J., 1962.
- Lucassen-Reynders, E.H., and J. Lucassen, Properties of capillary waves, *Adv. Colloid Interface Sci.*, 2, 347-395, 1969.
- Maxworthy, T., A note on the internal solitary waves produced by tidal flow over a three-dimensional ridge, *J. Geophys. Res.*, 84, 338-346, 1979.
- Onstott, R. and C. Rufenach, Shipboard active and passive microwave measurements of ocean surface slicks off the southern California coast, *J. Geophys. Res.*, 97, 5315-5323, 1992.
- Osborne, A.R., and T.L. Burch, Internal solitons in the Andaman Sea, *Science*, 208, 451-460, 1980.
- Ostrovsky, L.A., and Y.A. Stepanyants, Do internal solitons exist in the ocean?, *Rev. Geophys.*, 27(3), 293-310, 1989.
- Phillips, O.M., On the interactions between internal and surface waves, *Izv. Russ. Acad. Sci. Atmos. Oceanic Phys., Eng. Transl.*, 9, 565-568, 1973.
- Pingree, R.D., G.T. Mardell, and A.L. New, Propagation of internal tides from upper slopes of the Bay of Biscay, *Nature*, 321, 154-158, 1986.
- Scoon, A., I.S. Robinson, and P.J. Meadows, Demonstration of an improved calibration scheme for ERS-1 SAR imagery using a scatterometer wind model, *Int. J. Remote Sens.*, 17, 413-418, 1996.
- Sturdy, G., and W. Fischer, Surface tension of slick patches near kelp beds., *Nature*, 211, 951-952, 1966.
- Thompson, D.R., and R.F. Gasparovic, Intensity modulation in SAR images of internal waves, *Nature*, 320, 345-348, 1986.
- Vesecky, J.F., and R.H. Stewart, The observation of ocean surface radar phenomena using imagery from the Seasat synthetic aperture radar: An assessment., *J. Geophys. Res.*, 87, 3397-3430, 1982.

J. da Silva and I. Robinson, Department of Oceanography, Southampton Oceanography Centre, Waterfront Campus, European Way, Southampton, S014 3ZH, England, U.K. (e-mail: silva@soc.soton.ac.uk; isr@soc.soton.ac.uk)

S. Ermakov and S. Kijashko, Institute of Applied Physics, Russian Academy of Sciences, 46 Ul'yanov Street, 603600, Nizhniy Novgorod, Russia. (e-mail: stas@hydro.appl.sci-nnov.ru)

D. Jeans, GEOS Limited, Hargreaves Road, Groundwell Industrial State, Swindon, SN2 5AZ, England, U.K. (e-mail: jeans@geos.co.uk)

(Received December 30, 1996; revised August 14, 1997; accepted September 25, 1997.)

Role of surface films in ERS SAR signatures of internal waves on the shelf

2. Internal tidal waves

S. A. Ermakov

Institute of Applied Physics, Russian Academy of Sciences, Nizhniy Novgorod, Russia

J. C. B. da Silva and I. S. Robinson

Department of Oceanography, Southampton Oceanography Centre, University of Southampton, Southampton England, United Kingdom

Abstract. Analysis of ERS-1 synthetic aperture radar (SAR) images over the Iberian shelf shows that they contain the signatures of large-scale banded structures with characteristic wavelengths of 10-20 km, parallel to the shelf break. Their consistency with the structure of internal tides leads to the conclusion that they are the surface manifestations of internal tidal waves. The effect of imaging of such large-scale phenomena by SAR operating in L, C and X bands cannot be explained by conventional hydrodynamic modulation models based on the effect of straining of surface wind waves due to the internal tidal currents. An alternative explanation is described by an analysis of modulation of centimeter- and decimeter-scale surface waves caused by the redistribution of surface-active substances by internal tidal currents. Model calculations of variations of the wind wave spectrum across the internal tide profile are presented; an empirical film pressure-area equation and the results of thermistor chain measurements of internal tidal waves in the study area were used in the modeling. The effect of additional modulation of short surface waves due to variations of wind velocity relative to the internal tidal current is also estimated. It is demonstrated that modulation of centimeter-decimeter wind waves in the presence of surfactant films is the dominant factor and can explain the effect of internal tidal wave imaging by SAR.

1. Introduction

Short-period internal waves (IW) having wavelengths of the order of 1 km are known to be imaged by synthetic aperture radars (SARs) as systems of bright and/or dark bands on the image; the periods of which correspond to the spatial scales of the IWs. The commonly accepted imaging mechanism is based on hydrodynamic modulation models [see, e.g., Hughes, 1978; Alpers, 1985] in which the straining of surface waves due to IW orbital currents is shown to cause modulation of decimeter-scale surface waves, which have group velocities close to the IW phase velocity. Another requirement for this mechanism to be effective is that the relaxation length scale of the surface waves should be of the same order as, or longer than, the IW wavelengths. The relaxation length is the distance the waves propagate in the time it takes for a perturbation of the wave energy to relax to the equilibrium spectrum under the influence of wind forcing and dissipation processes. The apparent success of this approach to the interpretation of SAR images has tended to encourage the assumption that image variability at longer length scales should not be associated directly with ocean dynamical processes.

However large-scale periodic bands having 10-20-km length scales characteristic of internal tidal waves can be seen in SAR images over the shelf sea regions. It will be clearly demonstrated in this paper, which examines images over the Iberian shelf off the coast of Portugal, that these are related to the on-shelf propagation of internal tidal waves (ITWs). *Pingree and New* [1995] recently presented evidence for imaging of ITWs in the visible and infrared ranges obtained from satellite observations of the Bay of Biscay, although the imaging mechanism, as they mentioned, was not clear. Surprisingly, such evident phenomena seem to have received very little mention in the SAR literature, except for an allusion by *Fu and Holt* [1982]. The effect cannot be explained by the conventionally accepted IW imaging mechanism connected with hydrodynamic modulation of short wind waves. For wavelengths of order 10-20 km and semidiurnal tidal timescales the straining effect of the ITW surface current becomes negligible for decimeter-scale waves, and the relaxation lengths of short wind waves are much less than the scale of the tidal phenomenon. An alternative suggestion is that the dominant wind waves (10 m or greater), which have a longer relaxation time, might be strain modulated by the internal tidal currents and, in turn, modulate centimeter-decimeter-waves which control the radar backscatter. However, there is no possibility of phase coupling between the group velocities of the surface waves and the ITW phase velocity, and this hypothesis must also be rejected.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02724.
0148-0227/98/97JC-02724\$09.00

To find a satisfactory explanation for the imaging by SAR of ITWs, it is necessary to consider their modulating effect on surface films, which results in spatial variability of the equilibrium amplitude of the short waves which influence radar backscatter. Ermakov and Pelinovsky [1984] and Ermakov *et al.* [1992] have shown how the surface film properties are altered by the redistribution of surfactants in the IW surface current field. In an accompanying paper [da Silva *et al.*, this issue] (hereinafter referred to as paper 1) a model is developed to account for the combined modulation of short wind waves by the hydrodynamic straining and film redistribution effect of short-period ITWs. There it is demonstrated that such an approach is necessary to provide a satisfactory explanation of the variety of different types of SAR signatures of short-period ITWs (i.e., single negative and single positive signatures as well as the double sign phenomenon treated by the hydrodynamic straining approach). In this paper the same approach is applied to the case of larger-scale internal waves of tidal period. It will be shown that the hydrodynamic straining is ineffective as an imaging mechanism, but the film effect can account for sufficiently strong modulation to explain SAR imaging of internal tides.

2. Observations

The area in which the ITWs have been studied is the Iberian Shelf, shown with the bottom topography in Figure 1.

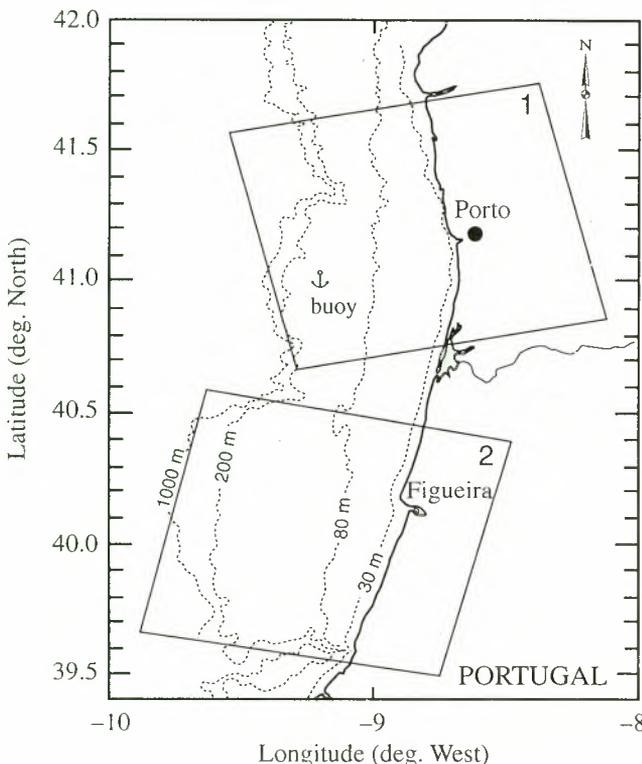


Figure 1. Bottom topography of the Iberian Shelf (studied area near parallel 41°N) where the in situ measurements were acquired in summer 1994 during the Multidisciplinary Oceanographic Research on the Eastern Boundary of the North Atlantic (MORENA 3B) cruise showing evidence of internal wave activity (short and large-period ITWs). Two frames of ERS-1 synthetic aperture radar (SAR) images are shown (labeled images 1 and 2). Location of a moored buoy is marked.

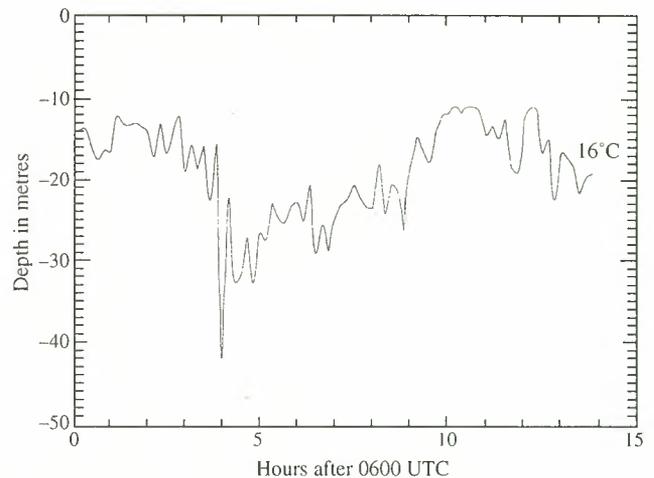


Figure 2. The 16°C isotherm is plotted and is assumed to represent the thermocline displacement during the measurements conducted on August 8, 1994. The isotherm shows a tidal-period of a large scale IW. Phase-locked on the large-period internal tidal wave (ITW) is a packet of several solitary ITWs moving with the same propagation speed as the ITW.

It can be characterized by a shallow water zone and sharp continental break at which internal tidal waves are generated by a barotropic tide. A number of observations of semi-diurnal internal tides have been conducted on the Australian shelf [Holloway, 1994], on the Malin shelf [Sherwin, 1988], in the Bay of Biscay [Pingree and New, 1989] etc. (see also Huthnance [1989], where a comprehensive list of observations of ITWs is given). Until recently, there were no published results about ITWs on the Iberian shelf. In August 1994, as part of the Multidisciplinary Oceanographic Research on the Eastern Boundary of the North Atlantic (MORENA) 3 cruise, measurements of internal tides on the western Iberian shelf were made [see Jeans and Sherwin, 1996]. Moored thermistor chain measurements were conducted from a buoy in August 1994 and allowed us to deduce the thermocline displacement; the buoy location is shown in Figure 1. An example of the buoy measurements on August 8, 1994, is given in Figure 2, which shows ITWs which are supposed to be generated by a barotropic tidal flow over the shelf break [Maxworthy, 1979; Baines, 1982]; a packet of short-period ITWs phase-locked to the ITWs is also seen in Figure 2. A typical sharp thermocline was observed in the test area, the thermocline depth was of order 20-25 m, thermocline displacements of tidal periods are about 4-9 m (D. R. G. Jeans, *personal communication*, 1997), and phase velocities of the ITWs were estimated to be about 0.5 m/s. The estimation of IW phase speed was based on consecutive thermistor chain measurements of thermocline displacements at different locations on the shelf [Jeans and Sherwin, 1996] and is in good agreement with the estimate made from ERS SAR images [see da Silva *et al.*, 1997].

An example of an ERS-1 SAR image of the area is presented in Figure 3a. Groups of bright and dark periodic bands (wavelength about 0.3-0.5 km) spaced about 15-20 km apart are associated with short-period ITWs. It should be noticed, however, that another distinctive feature of this image is the distribution of large-scale dark and bright bands in phase with the groups of short-period bands. This is not an isolated example of such large-scale banded structures. As another ex-

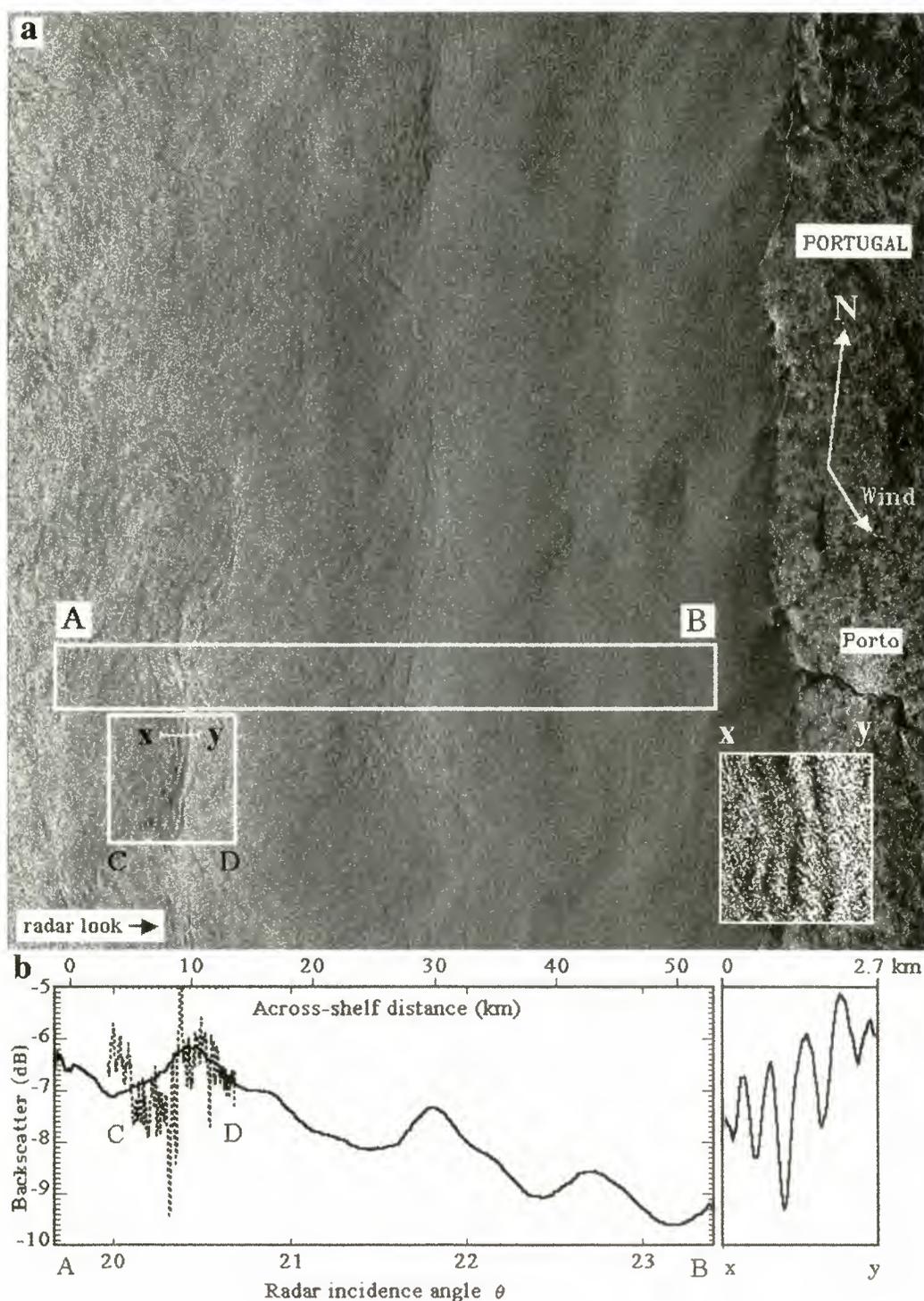


Figure 3. (a) ERS-1 SAR image 1 (orbit: 10481; frame: 819; date: July 17, 1993). The inset box on the bottom right corner is a zoom in the “xy” region of the image showing short-scale IW signatures. (b) “AB” is a smoothed backscatter profile across the shelf zone (along the line “AB”) showing large-scale image contrast variations due to internal tides. Profile “CD” is a fraction of the image data without the low-pass filter applied and shows short-period IWs which are phase-locked with the internal tide. High-resolution image profile “xy” is also presented, showing a packet of IW solitons.

ample, Figure 4a is an image (image 2) acquired 1 year earlier on July 26, 1992. This image was filtered in order to remove the high-frequency signatures of short-period IWs and enhance the large-scale bands. The large-scale periodicity is more clearly demonstrated in Figures 3b and 4b, where backscatter profiles from the images along the lines perpen-

dicular to the bands are shown. The period of these large-scale bands is about 10–20 km. The mean trend of the backscatter in Figures 3b and 4b is due to the normalised radar cross section (NRCS) decrease with the local incidence angle. A fragment of a backscatter profile (“xy”) without the low-pass filter applied is also presented in Figure 3b, showing

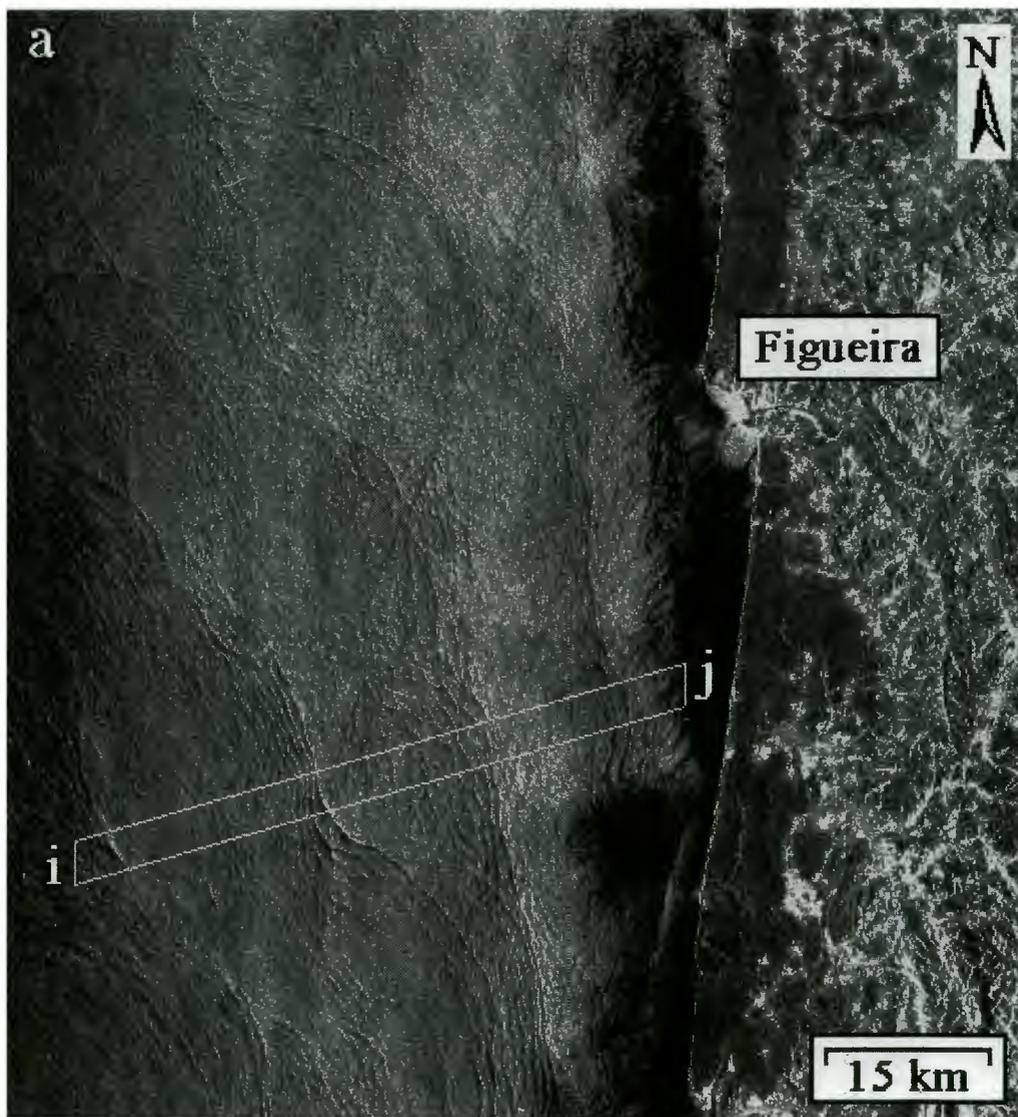


Figure 4. (a) ERS-1 SAR image (orbit: 5378; frame: 2799) labeled 2 in Figure 1, acquired on July 26, 1992, at 1120 UTC during a descending orbit of the satellite. Both short-period and large-period IW signatures are seen. (b) A smoothed backscatter profile along the line "ij" marked in Figure 4a showing the signature of ITWs. Modulation as high as 1 dB is apparent. Wind velocity is 4 m/s from 315°N.

short-period variations associated with a packet of short-period IW solitons in a corresponding phase of the large-scale variations. The large-scale bands are parallel to the shelf break, they have wavelengths typical of internal tidal waves,

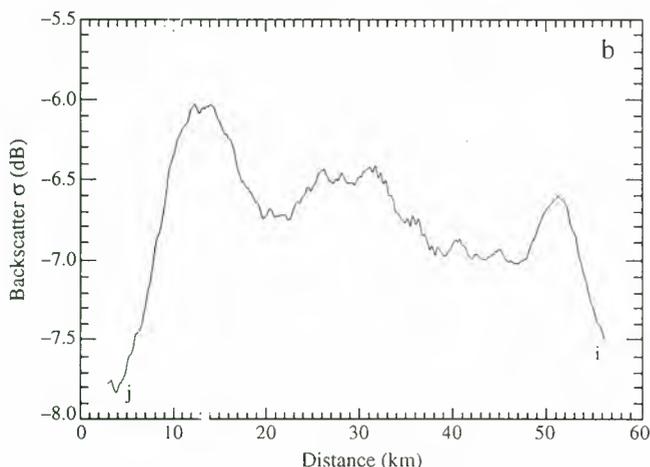


Figure 4. (continued)

and it is reasonable to assume that they are the radar signatures of internal tides. Similar large-scale bands are often seen in SAR images over the shelf, and it is difficult to conceive another origin for these bands which would be topographically locked. We have noticed, however, that at very low wind the large-scale bands are usually not discernible in the images, the effect being masked by manifestations of wind irregularities. Our object in the rest of this paper is to show that contrary to the implications of previous models of surface roughness modulations by ITWs, there is a satisfactory mechanism which can account for the radar imaging of such large-scale ITWs.

3. Model of Modulation

When studying the action of ITWs on short wind waves, we shall follow a model developed in paper 1. Here we reproduce only the general equations of the model. The wind wave spectrum variations due to ITW surface currents can be described by the balance equation for the spectral density of

wave action $N(\mathbf{k}, \mathbf{x}, t)$ [see e.g., Hasselmann, 1968; Keller and Wright, 1975; Willebrand, 1975] and

$$\frac{dN}{dt} = (\beta - \gamma)N - \delta N^2 \quad (1)$$

with the right-hand side in the form suggested by Hughes [1978]. Here $N(\mathbf{k}, \mathbf{x}, t) = \rho\omega(\mathbf{k})S/\mathbf{k}$, where S denotes the surface wave height spectrum, $\omega(\mathbf{k})$ the intrinsic frequency of surface waves, β and γ the wind wave growth rate and the wave damping coefficient, respectively, and δ is a phenomenological coefficient in the last term of (1), describing nonlinear limitation of the wind wave spectrum. Note that in the right-hand side of (1) we have neglected nonlinear wave-wave interaction and wave-breaking terms [see, e.g., Hasselmann, 1968]. Although we assume that these terms should be small compared to wind wave input and film damping, the assumption should be tested in particular cases.

The wave damping coefficient depends on the film elasticity parameter E , and film elasticity can vary because of variations of surfactant concentration Γ in the ITW orbital surface velocity field $U(\mathbf{x}, t)$. The surfactant concentration Γ on the sea surface can be described by the conservation equation in the form

$$\frac{\partial \Gamma}{\partial t} + \text{div}(\Gamma U) = 0, \quad (2)$$

where, for simplicity, processes of surface diffusion and surfactant exchange with subsurface layers are neglected.

Assuming that the internal wave velocity field can be described by a stationary profile in the form

$$U(x, t) = U_m + U_{IW}(x - C_p t),$$

where U_m is the mean current velocity and $U_{IW}(x - C_p t)$ propagates at a fixed velocity C_p , the steady state solution of (2) can be written as

$$\Gamma(z) = \Gamma_0 \left[1 - \frac{U_{IW}(z)}{C_p - U_m} \right]^{-1} \quad z = x - C_p t, \quad (3)$$

where Γ_0 is the unperturbed surfactant concentration when $U_{IW} = 0$. It should be noted that the IW orbital velocities on the sea surface are in counterphase with the thermocline displacements, and according to (2) this implies that there is increased surfactant concentration above wave troughs [see, e.g., Ermakov and Pelinovsky, 1984].

Since the relaxation times of centimeter- and decimeter-scale waves which are of order $(\beta - \gamma)^{-1}$ are small compared with ITW periods, we can neglect the straining effect due to ITWs and consider the wave spectrum to be close to the local equilibrium spectrum. The equilibrium spectrum is found from (1) by setting $(dN/dt) = 0$. Then the relative spectrum variation (contrast) is given by

$$K = N(\mathbf{k}, x, t)/N_0 - 1 \quad (4)$$

and can be written as a sum of the "film" and the "wind" contrasts, which are equal, respectively, to

$$K_{\text{film}} = -\frac{\gamma - \gamma_0}{\beta_0 - \gamma_0} \quad (5)$$

$$K_{\text{wind}} = \frac{\beta - \beta_0}{\beta_0 - \gamma_0} \quad (6)$$

In (4)-(6), N_0 , β_0 , and γ_0 are the unperturbed spectral density of wave action, the coefficient of wind wave growth, and the coefficient of wind wave damping, respectively, nominally in the absence of internal tide flow. The film contrast describes spectrum variations due to variations of wave damping caused by the surfactant redistribution in the ITW current field. The wind contrast corresponds to variations of wind wave growth due to varying relative wind velocity, which is $\mathbf{V} - \mathbf{U}(x, t)$ (here \mathbf{V} is the mean wind velocity). Note that the expression for the film contrast (5) is similar to that obtained in paper 1, where the separation of film and straining effects requires an approximation to be made. The difference from paper 1 is that in this case, where the straining effect is negligibly small, expressions (5) and (6) are much better approximations and are valid even for strong spectrum variations.

Following paper 1, we have used for calculations of the spectrum contrasts the dependence of the damping rate γ on film elasticity E for pure elastic films [see Lucassen-Reynders and Lucassen, 1969; paper 1, equation (4)], and the virial equation relating film pressure and surfactant concentration as deduced from empirical data on marine films [see Frew and Nelson, 1992; Barger and Means, 1985]. We used the virial

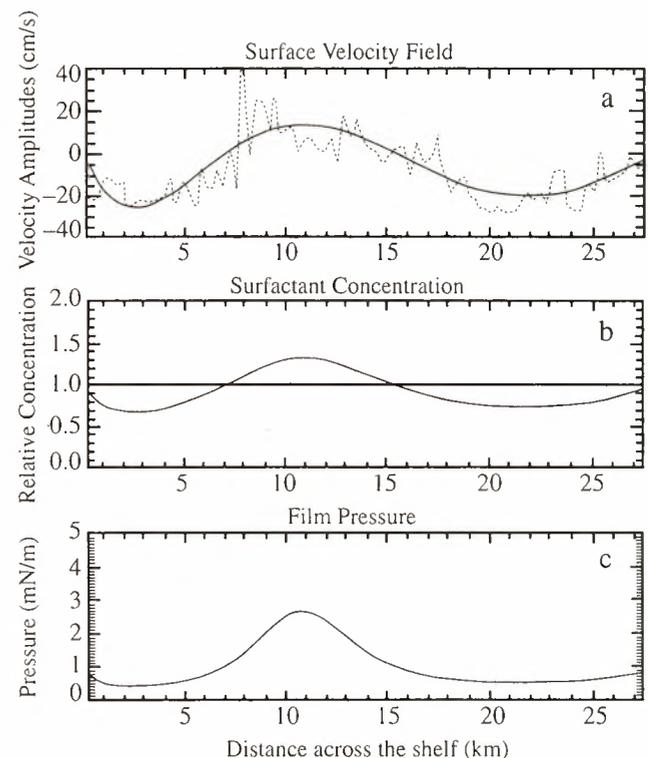


Figure 5. (a) Surface velocity field retrieved from the buoy thermistor measurements of August 8, 1994, and corresponding calculated variations of (b) surfactant concentration and (c) film pressure across the shelf. Solid lines correspond to the smoothed velocity profile, and dotted lines correspond to the unsmoothed profile and describe variations due to both tidal-period and short-period ITWs.

equation with coefficients averaged over those obtained for marine films from coastal productive areas. An expression for the film elasticity parameter deduced from the virial equation is given in paper 1.

4. Model Results

In the absence of in situ measurements coincident in time with SAR observations of the internal tide we sought simply to model the response of the surface film modulation to the velocity field associated with a typical internal tide. Surface currents were deduced from moored thermistor chain measurements of the thermocline displacement conducted on August 8, 1994. The surface currents were estimated using a two-layer model of stratification. The accuracy of this estimation was tested for simultaneous thermistor chain and direct current ADCP measurements of short-period IWs, and the measured and retrieved current velocities were shown to be in good agreement (see paper 1). Figure 5a shows a surface velocity profile retrieved from the thermistor measurements. Estimation of the IW phase velocity enabled us to interpret the time series of thermistor measurements as a function of distance across the shelf. This estimation (see

also paper 1) was based on the arrival time of the leading soliton of the main IW packet at the buoy and its position on a SAR image acquired on August 8, 1994.

The curve fit on the high-frequency velocity profile (Figure 5a) demonstrates the low-frequency (large-scale) internal tide oscillations. Figures 5b and 5c show the calculated profiles of relative surfactant concentration and corresponding variations of film pressure based on an undisturbed film pressure value $\pi_0 = 0.75$ mN/m.

The calculated film and combined contrasts are presented in Figure 6 for surface wavelengths of 7, 14, and 21 cm and for the wind conditions at the time of acquisition of the SAR image in Figure 3a (wind direction, 320°; wind velocity, 5.5 m/s). For comparison, variations of the surface wave spectrum for the unsmoothed velocity profile of Figure 5a were calculated following paper 1. These are shown as dotted lines in Figure 6, clearly demonstrating the significance of the straining (hydrodynamic) effect for short-period IWs and its insignificance for ITWs. For 7 cm waves, Figure 6 demonstrates strong wave damping over ITW troughs due to the film effect which results in a tidal wavelength contrast profile. The film contrast for the geometry of observations on July 17, 1993, dominates over the wind contrast for centimeter-scale

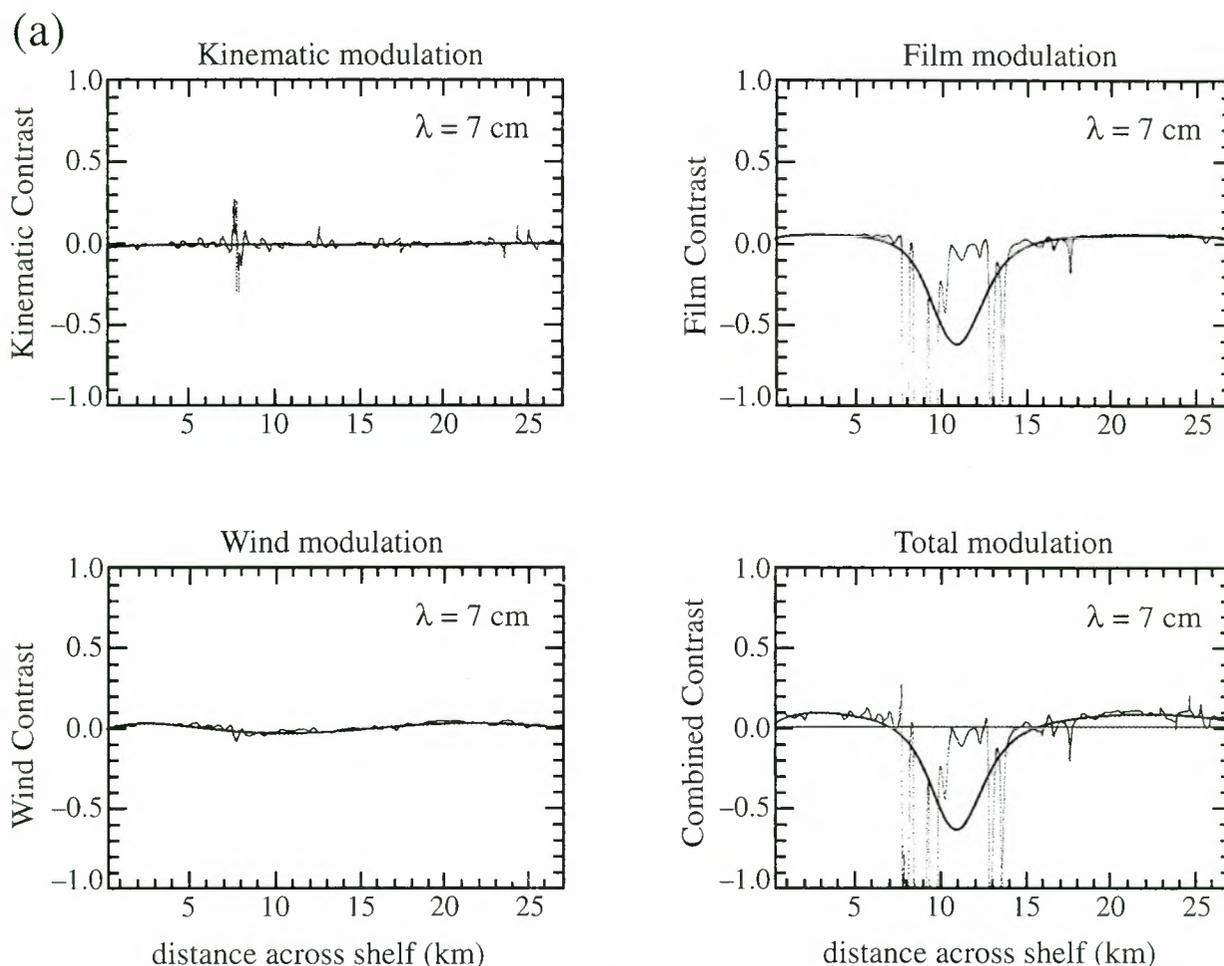


Figure 6. Model variations of the surface wave spectrum calculated separately for the kinematic, film, and wind modulation. The combined effect of the three contrasts is also shown at given surface wavelengths of (a) 7 cm (ERS-1 SAR Bragg wavelength), (b) 14 cm, and (c) 21 cm. The model results are presented for the conditions observed in SAR image shown in Figure 3a (wind velocity 5.5 m/s, direction 320°N). The solid lines correspond to the spectrum variations due to ITW alone, while the dotted lines are for the unsmoothed surface velocity field and incorporate effects from both short-period and internal tidal waves.

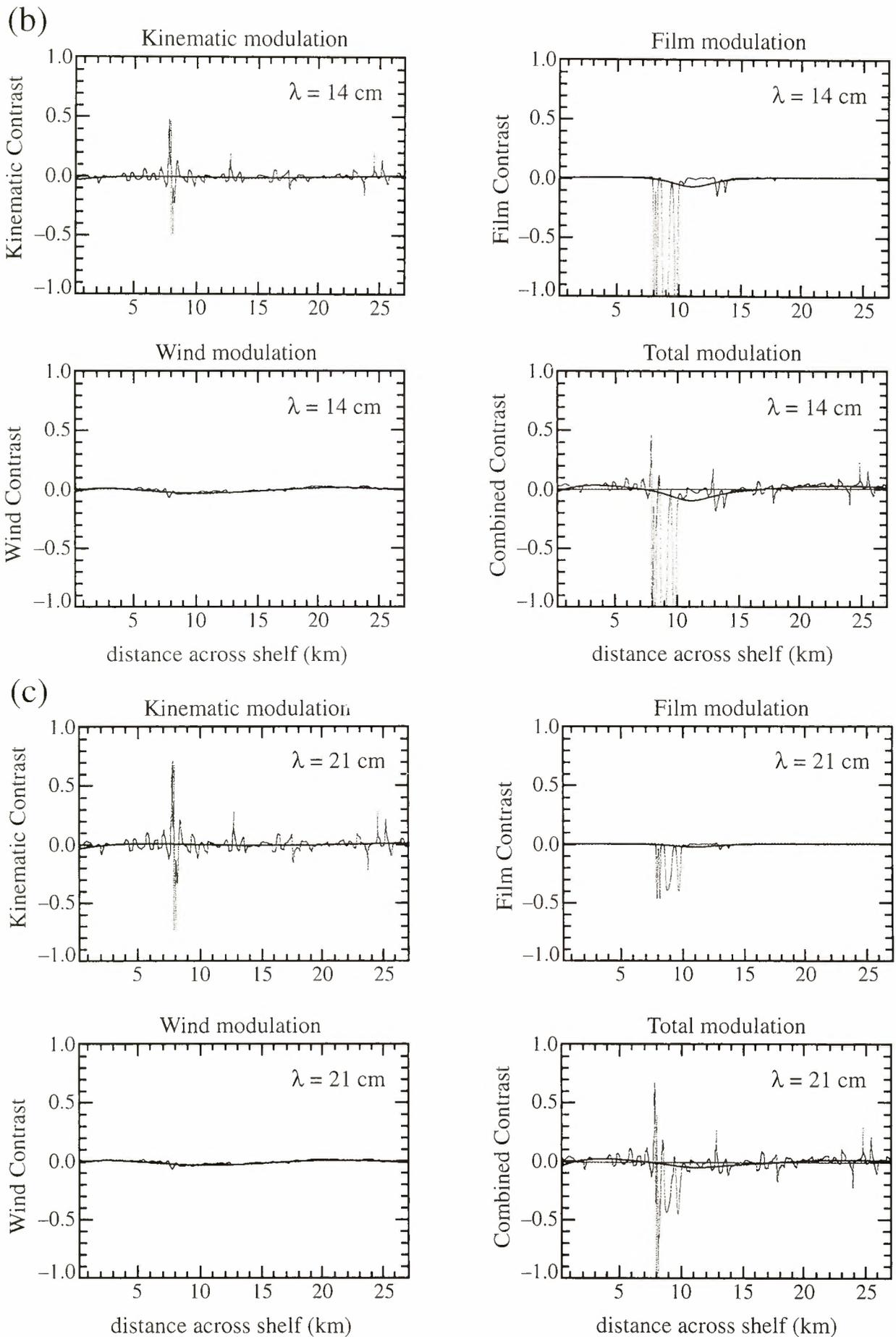


Figure 6. (continued)

waves. With increasing wavelength of wind waves the film contrast decreases and becomes comparable with the wind contrast in the decimeter-wavelength range. The wind conditions for the image acquired on July 26, 1992, are similar to those on July 17, 1993, so that a similar conclusion follows from the contrast modeling.

From the model results presented, it follows that for centimeter-scale waves the film contrast is negative and large; that is centimeter-scale waves are damped strongly over ITW troughs, and over ITW crests the contrast is positive and comparatively weak. Wind modulation is small in the cases when the angle between the wind and ITW propagation directions is quite large, as in the cases analyzed above. But the wind modulation can be important, at least for decimeter-scale waves, if the wind velocity is nearly parallel to the ITW propagation direction. To investigate the wind effect, we calculated contrasts for the cases when wind velocity is directed toward the shore (wind velocity and ITW wave vector are parallel) and when wind velocity is antiparallel to the ITW wave vector. The wind modulation and total modulation due to ITWs for 7-cm-scale surface waves are plotted in Figure 7a for the case when wind velocity is parallel to the ITW wave vector and in Figure 7b when the wind and the ITW wave

vector are antiparallel. When the wind is parallel to the ITW propagation direction, both film and wind modulation are in phase, thus enhancing the combined effect. If the wind and the internal tide propagation directions are opposite to each other, the film and wind effects act in antiphase, and they can partly compensate each other. In this case, internal tide imaging is expected to be less clear than in the previous case. We thus can expect that ITWs should be more readily visible if wind and ITW propagation direction are parallel.

5. Discussion

As is seen from the model results presented, the film modulation can give strong variability of the spectrum of short wind waves. Now we discuss how sensitive the spectrum variations responsible for ITW imaging are to the parameters of the model, particularly to the undisturbed film pressure, the amplitude of ITWs, and the wind velocity and direction. Since the film effect dominates for centimeter-scale waves, we present the sensitivity analysis only for the Bragg 7-cm waves. Figure 8a shows the spectrum film contrast at different pressure values of 1, 0.75, 0.5, and 0.25 mN/m, where the wind velocity and amplitude of ITW are assumed

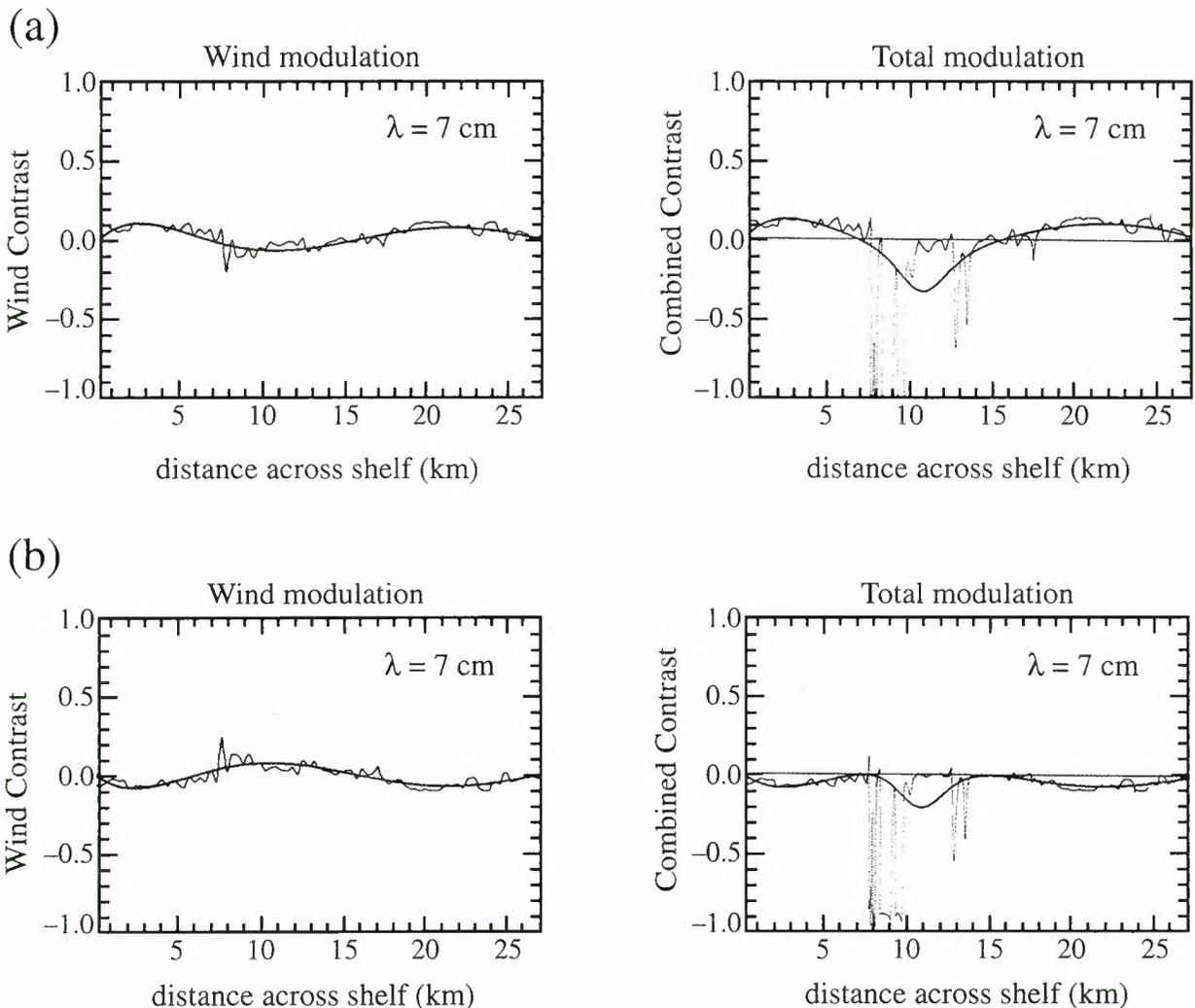


Figure 7. Model wind and combined contrasts when wind velocity and the ITW propagation directions are (a) parallel and (b) antiparallel. Wind velocity is 4 m/s, and unperturbed film pressure is 0.75 mN/m. The surface wavelengths are 7 cm.

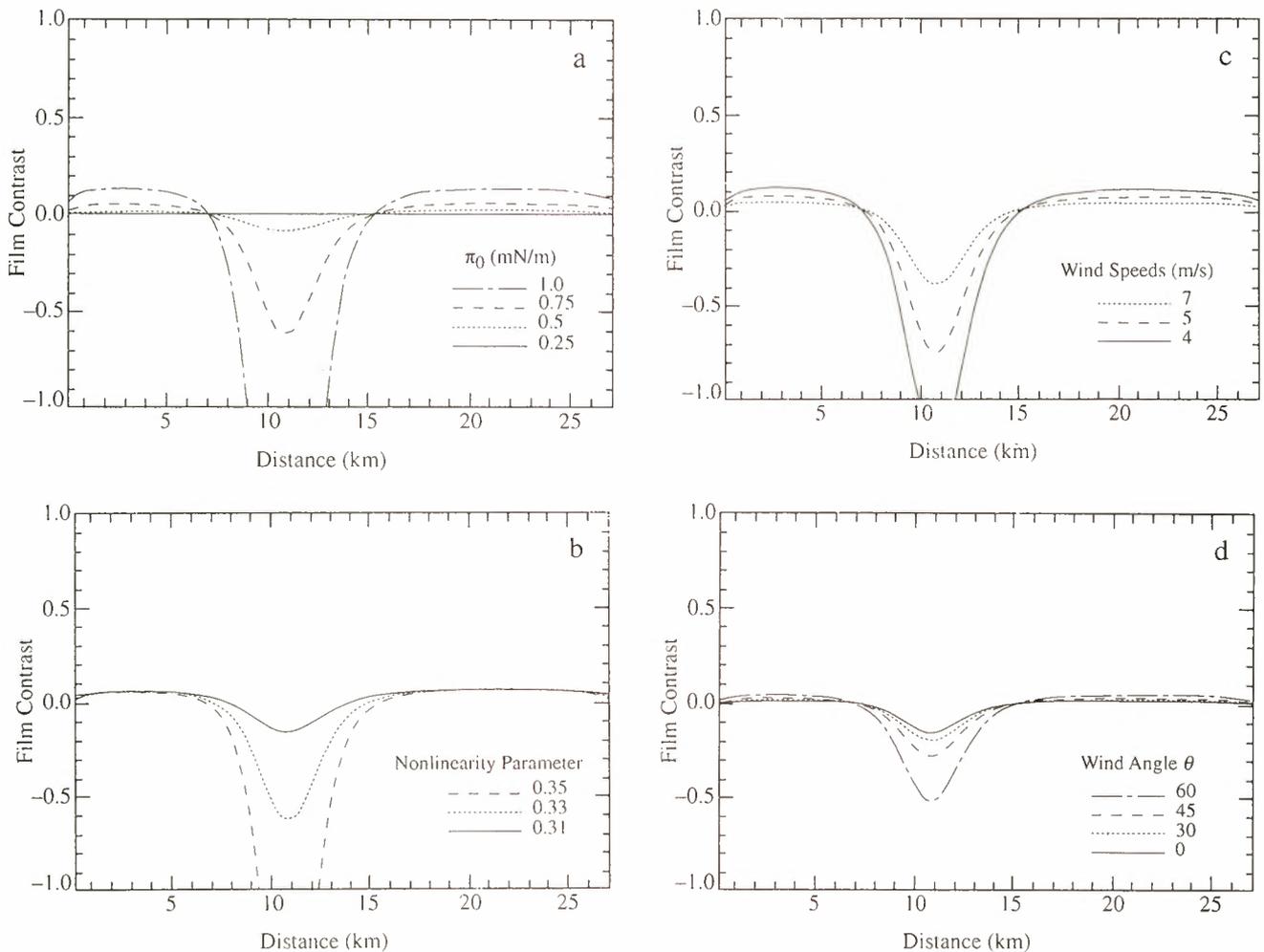


Figure 8. Film contrasts for 7-cm-scale surface waves at different (a) film pressure values, (b) ITW nonlinearity parameters, (c) wind velocities, and (d) wind directions.

to be the same as in Figures 5 and 6. Figure 8b presents film contrasts at different values of the nonlinearity parameter of ITWs, which is defined as a ratio of the thermocline amplitude to the thermocline depth (here the film pressure assumed was 0.75 mN/m, and wind conditions were as in Figure 8a). For the given amplitude of the thermocline elevation, in particular, the variations of the nonlinearity parameter in Figure 8b correspond to variations of the thermocline depth of about 3 m. The film contrast at different wind velocities and wind directions is presented in Figures 8c and 8d (at the nonlinearity parameter of 0.33 and $\pi_0 = 0.75$ mN/m). From the results we conclude that the film contrast is quite sensitive to variations of the model parameters, especially to the nonlinearity parameter and the initial pressure value. This strong sensitivity is not surprising because the two parameters determine variations of the film elasticity over ITWs and, consequently, variations of the wave damping coefficient. The last dramatically depends on the elasticity value, so that even small changes of the elasticity variations can lead to strong differences in the damping coefficient. It should also be noted that all the model results were obtained for particular averaged coefficients in the virial equation and that for a particular film the contrasts can differ from those estimated above.

In the model we neglected relaxation processes in surfactant films. This relaxation can be connected with turbulent mixing and the diffusion of surfactants to subsurface layers and also with horizontal turbulent diffusion. At present the relaxation processes for surfactant films are not well understood. Diffusion relaxation times are small for soluble surfactants, and these are not expected to play an important role in wave damping variations with ITW periods. However, for insoluble surfactants (which are responsible for the strongest wave damping), vertical mixing and diffusion is expected to be weak, at least at low wind speeds. The characteristic time τ of horizontal diffusion can be estimated roughly as L^2/D , where L is the ITW wavelength and D is the eddy viscosity coefficient. For D values $< 10^8$ - 10^9 cm²/s (which is typical for horizontal diffusion in the ocean, see, e.g., *LeBlond and Mysak* [1978]), τ is of the order of tens of hours, i.e., larger than the ITW periods, so that horizontal diffusion can be neglected.

The predicted effect on a radar image of the calculated modulation of the wave spectra depends on the radar model used. For ERS SAR at incidence angle 23° the Bragg wavelength is 7 cm, and if a simple Bragg model is applied, the spectrum contrast at 7 cm should represent the expected image contrast. Although the spectral variations at

wavelength of 7 cm in Figure 6a overestimate the intensity of the SAR image variations, it follows from Figure 8 that realistic parameters of the model can be chosen for which model contrasts at a wavelength of 7 cm fit well to the observed radar intensity variations. We thus conclude that the presented hydrodynamic model incorporating the effect of film damping on short wind waves can give a realistic explanation for the phenomenon of SAR imaging of internal tidal waves.

Additionally, it should be noted that according to field observations [see Mitsuyasu and Honda, 1982; Ermakov et al., 1986; Wei and Wu, 1992], wind stress (friction velocity) decreases over surface slick regions by up to 10-20%. Therefore variations of surfactant concentration due to ITWs could result in additional modulation of wind wave growth of short surface waves with consequent intensification of the expected contrasts. Although this effect could also be significant, we did not take it into account since this effect has not been studied in sufficient detail to parameterize it in the model.

At low to moderate incidence angle range, which is the case for ERS SAR, an improved backscatter model [see Holliday et al., 1986] has been proposed for interpretation of the SAR sea surface imaging. This model suggests that the contribution to the ERS SAR imaging is determined by surface waves from 7 to about 20 cm (up to 2-3 Bragg wavelengths). Since our model-predicted variations at decimeter-scale wavelengths (10-20 cm) are smaller than for centimeter-scale wavelengths, the resulting radar contrasts in the frame of Holliday et al.'s model would be smaller and less sensitive to the film parameters, although the order of magnitude is expected to be the same as estimated for the Bragg model. There is therefore scope for more detailed analysis of SAR imaging of the sea surface using Holliday et al.'s model and incorporating the film effect.

Finally, we note that this explanation of radar imaging of ITWs can also provide the basis for interpreting other large-scale phenomena which may be encountered in radar images when the prevailing wind is no more than 7 m/s. Unlike the hydrodynamic modulation model, the length scale of the phenomenon to be imaged is not restricted to the relaxation length scale but may be tens of kilometers at least. Neither need its timescale be restricted to a few minutes but can be 5-30 hours. Thus surface convergences at fronts, or the surface currents in an upwelling zone, may cause surface film redistribution, giving rise to patterns of backscatter modulation which reveal the underlying dynamical phenomenon in SAR images. In principle, as long as the process is able to maintain a sufficiently strong gradient of surface film properties against the naturally occurring diffusion and dispersion of the film, it is capable of being imaged by SAR. Note, however, that if current gradients are strong, the hydrodynamic modulation cannot be neglected and may lead to wave intensification superimposed on the damping effects of films. In those circumstances both effects are important and must be taken into account to describe the radar images. To determine whether this occurs in practice will require the modeling approach presented here to be adapted for the particular dynamical circumstances.

Acknowledgments. This is MORENA publication 17. We thank Gus Jeans and Toby Sherwin from UCES, Marine Science Laboratories, for processing the thermistor chain data records and for useful discussions; we also thank our colleagues from the University of

Lisbon, including A. Fiúza for providing the thermistor data. S.E. acknowledges the Winfrith Technology Centre, DERA, United Kingdom (project SSDW1/113) and the Russian Foundation of Basic Research (grant 96-05-65087), which partially sponsored this work. J.C. da S. gratefully acknowledges the EU Human Capital and Mobility programme (ERBCHBITC 941008) for a scholarship to study ocean slicks and JNICT for partial support of this work. ERS-1 SAR image data were provided by ESA as part of the AO project AO2.UK123.

References

- Alpers, W., Theory of radar imaging of internal waves., *Nature*, 314, 245-247, 1985.
- Baines, P.G., On internal tide generation models, *Deep Sea Res., Part A*, 29, 307-338, 1982.
- Barger, W.R., and J.C. Means, Clues to the structure of marine organic material from the study of physical properties of surface films, in *Marine and Estuarine Geochemistry*, edited by A.C. Sigleo and A. Hattori, pp. 47-67, A.F. Lewis, New York, 1985.
- da Silva, J.C., D.R.G. Jeans, I.S. Robinson, and T.J. Sherwin, The application of near-real time ERS-1 SAR data for predicting the location of internal waves at sea., *Int. J. Remote Sens.*, in press, 1997.
- da Silva, J.C.B., S.A. Ermakov, I.S. Robinson, D.R.G. Jeans, and S.V. Kijashko, Role of surface films in ERS SAR signatures of internal waves on the shelf, I. Short-period internal waves, *J. Geophys. Res.*, this issue.
- Ermakov, S.A., and E.N. Pelinovsky, Variation of the spectrum of wind ripple on coastal waters under the action of internal waves, *Dyn. Atmos. Oceans*, 8, 95-100, 1984.
- Ermakov, S.A., A.M. Zujkova, A.R. Panchenko, S.G. Salachin, T.G. Talipova, and V.I. Titov, Surface film effects on short wind waves, *Dyn. Atmos. Oceans*, 10, 31-50, 1986.
- Ermakov, S.A., S.G. Salashin, and A.R. Panchenko, Film slicks on the sea surface and some mechanisms of their formation., *Dyn. Atmos. Oceans*, 16, 279-304, 1992.
- Frew, N., and R. Nelson, Scaling of marine microlayer film surface pressure-area isotherms using chemical attributes, *J. Geophys. Res.*, 97, 5291-5300, 1992.
- Fu, L.L., and B. Holt, Seasat views oceans and sea ice with synthetic-aperture radar, *JPL Publ.*, 81-120, 26-41, 1982.
- Hasselmann, K., Weak-interaction theory of ocean waves, in *Basic Developments in Fluid Dynamics*, vol. 2 edited by M. Holt, pp. 117-182, Academic, San Diego, Calif., 1968.
- Holliday, D., G. St-Cyr, and N.C. Woods, A radar ocean imaging model for small to moderate incidence angles., *Int. J. Remote Sens.*, 7, 1809-1834, 1986.
- Holloway, P.E., Observations of internal tide propagation on the Australian North West Shelf, *J. Phys. Oceanogr.*, 24, 1706-1716, 1994.
- Hughes, B.A., The effect of internal waves on surface wind waves, 2. Theoretical analysis, *J. Geophys. Res.*, 83, 455-465, 1978.
- Huthnance, J.M., Internal tides and waves near the continental shelf edge, *Geophys. Fluid Dyn.*, 48, 81-106, 1989.
- Jeans, D.R.G., and T.J. Sherwin, Solitary internal waves on the Iberian shelf, *MORENA Sci. and Tech. Rep. 44, EU MAST program, Univ. C. North Wales*, Nov. 1996.
- Keller, W.C., and J.W. Wright, Microwave scattering and the straining of wind-generated waves, *Radio Sci.*, 10, 139-147, 1975.
- LeBlond, P.H., and L.A. Mysak, *Waves in the Ocean*, 602 pp., Elsevier Sci., New York, 1978.
- Lucassen-Reynders, E.H., and J. Lucassen, Properties of capillary waves, *Adv. Colloid Interface Sci.*, 2, 347-395, 1969.
- Maxworthy, T., A note on the internal solitary waves produced by tidal flow over a three dimensional ridge, *J. Geophys. Res.*, 84, 38-346, 1979.
- Mitsuyasu, H., and T. Honda, Wind induced growth of water waves, *J. Fluid Mech.*, 123, 425-442, 1982.
- Pingree, R.D., and A.L. New, Downward propagation of internal tidal energy into the Bay of Biscay, *Deep Sea Res., Part A*, 36, 735-758, 1989.
- Pingree, R.D., and A.L. New, Structure, seasonal development and sunglint spatial coherence of the internal tide on the Celtic and Armorican shelves and in the Bay of Biscay, *Deep Sea Res., Part 1*, 42, 245-284, 1995.

- Sherwin, T.J., Analysis of an internal tide observed on the Malin shelf, north of Ireland, *J. Phys. Oceanogr.*, 18, 1035-1050, 1988.
- Sherwin, T.J., A.C. Dale, M.E. Inall, and D.R.G. Jeans, Linear and non-linear internal tides around the European Atlantic shelf edge, in *Proceedings of the Sixth International Offshore and Polar Engineering Conference, Los Angeles*, pp. 131-137, Int. Society of Offshore and Polar Engineers, Golden, Colorado, U.S.A., 1996.
- Wei, Y., and J. Wu, In situ measurements of surface tension, wave damping, and wind properties modified by natural films, *J. Geophys. Res.*, 97, 5307-5313, 1992.
- Willebrand, J., Energy transport in a non-linear and inhomogeneous random gravity wave field, *J. Fluid Mech.*, 70, 113-126, 1975.
-
- J. C. B. da Silva and I. S. Robinson, Department of Oceanography, Southampton Oceanography Centre, University of Southampton, European Way, Southampton, SO14 3ZH, England, U.K. (e-mail: silva@soc.soton.ac.uk; isr@soc.soton.ac.uk)
- S. A. Ermakov, Institute of Applied Physics, Russian Academy of Sciences, 46 Ul'yanov St., 603600 Nizhniy Novgorod, Russia. (e-mail: stas@hydro.appl.sci-nnov.ru)

(Received January 29, 1997; revised September 15, 1997; accepted September 24, 1997.)

ERS-1/2 orbit improvement using TOPEX/POSEIDON: The 2 cm challenge

P.-Y. Le Traon and F. Ogor

Collecte, Localisation, Satellites Space Oceanography Division, Ramonville Saint-Agne, France

Abstract. The ERS orbit error reduction method using TOPEX/POSEIDON (T/P) data as a reference [Le Traon *et al.*, 1995a] was applied to ERS-1 cycles from phases C, E, F, and G and to the first 16 cycles of the ERS-2 mission (phase A). T/P M-GDR (geophysical data record) (version C) and ERS-1/2 ocean product (OPR) data were used. ERS-1/2 orbits are the D-PAF (processing and archiving facility) orbits and, when necessary, ERS-1/2 altimetric corrections were updated to make the T/P and ERS-1/2 corrections homogeneous. The adjustment method has been refined, and formal error on the estimation is now calculated. The ERS-1/2 orbit error estimation is thus estimated to be precise to within about 2 cm root-mean-square (rms). E-E crossover differences are reduced from 12 to 17 cm to only 6.5 cm rms for all processed cycles. Similarly, the T/P-E crossover differences are reduced from 11 to 14 cm to only 7 cm rms. The adjusted D-PAF orbit error varies between 6 and 12 cm rms. The adjustment has also been performed for the Joint Gravity Model 3 (JGM 3) orbits of ERS-1 phases C, E, and F. The rms difference between the corrected orbits for the D-PAF and JGM 3 orbits is only about 1 cm rms, while it is about 11 cm before T/P orbit error correction. This shows that the adjustment is almost insensitive to the initial ERS-1 orbit used. It also confirms the 2 cm precision of the method. We also do repeat-track analysis on the 35 day repeat cycles of ERS-1 phase C. The mean difference in sea level variance before and after orbit error correction is 34 cm^2 (D-PAF orbit) and 17 cm^2 (JGM 3 orbit). The corrected ERS-1 and T/P sea level variabilities, however, are in excellent agreement. The study thus shows that ERS-1/2 orbit error must be corrected before analyzing large-scale oceanic signals and combining ERS-1/2 with T/P data. The proposed method provides a very effective correction and thus significantly enhances the quality of ERS-1/2 data. Corresponding data sets will be distributed to the scientific community by Archiving, Validation, and Interpretation of Satellite Oceanographic Data (AVISO).

1. Introduction

Since the launch of TOPEX/POSEIDON (T/P) in August 1992, several altimetric satellites (T/P, ERS-1, and ERS-2) have been flying simultaneously. This is likely to recur in the future with the Geosat follow-on, Envisat, and Jason missions. While T/P, with its unprecedented accuracy, has provided a new picture of the ocean, it cannot observe the full spectrum of the sea level and oceanic circulation. At least one more mission is needed, in particular, to resolve the mesoscale oceanic circulation. The ERS satellites are thus an excellent complement for T/P sampling. They provide a good resolution of the mesoscale oceanic circulation [e.g., Hernandez *et al.*, 1995] while T/P provides the large-scale oceanic circulation. They also provide coverage at high latitudes.

Merging the different altimetric data sets will provide better mappings of sea level variation and the geoid (or more exactly the mean sea surface) [e.g., Rapp *et al.*, 1994; P. Mazzega *et al.*, Maps of the mean sea surface and corresponding gravity anomalies from ERS-1 geodetic mission, submitted to the *Journal of Geophysical Research*, 1996]. Merging multisatellite data sets requires homogeneous, intercalibrated data sets. This means first that the same altimetric corrections should be used

(same tidal models, same meteorological models for inverse barometer and dry tropospheric corrections, etc.) and that orbits should generally be calculated with the same geopotential model and with consistent reference systems. To obtain intercalibrated data sets, we recommend using the most precise altimetric missions (T/P, Jason) as a reference for the other altimetric missions. This method has been shown to provide very good results with ERS-1 [Le Traon *et al.*, 1995a, b; Carnochan *et al.*, 1995]. Although ERS-1/2 orbit accuracy has been significantly improved since these studies, T/P is still much more precise. The ERS-1/2 orbits are accurate to within about 7–10 cm [Scharroo *et al.*, 1994; Massmann *et al.*, 1997] while T/P accuracy is 2 cm [e.g., Fu *et al.*, 1994; Tapley *et al.*, 1996]. In addition, contrary to T/P, the altimetric data themselves are generally used in ERS-1/2 orbit determination.

The purpose of the paper is to demonstrate that T/P can dramatically improve ERS-1/2 orbit accuracy and that the technique proposed by Le Traon *et al.* [1995a] is very effective at providing homogeneous and intercalibrated altimetric data sets. The paper is organized as follows. The ERS orbit error reduction method is briefly described in section 2, and the accuracy of the fit is analyzed in section 3. The method is then applied to the ERS-1 35 day repeat cycle mission (phase C), the ERS-1 geodetic mission (phases E and F), the second ERS-1 35 day repeat cycle mission (phase G), and the first 16 cycles of the ERS-2 35 day repeat cycle mission (phase A).

Copyright 1998 by the American Geophysical Union.

Paper number 97JC01917.
0148-0227/98/97JC-01917\$09.00

Results are discussed in section 4. Analysis of sea level variability is performed in section 5 to demonstrate the improvement in ERS-1/2 data accuracy. The main conclusions and perspectives are given in section 6.

2. ERS Orbit Error Reduction Method

The ERS orbit error reduction method is described in detail by *Le Traon et al.* [1995a] and is only summarized here. The method is based on global minimization of T/P-E and E-E crossover differences. T/P-E crossover differences give an estimate of the ERS orbit almost directly, leading to a "geometric" estimation of orbit error. Smoothing cubic spline functions [Hayes, 1974] are then used to obtain a continuous estimation of orbit error over time. E-E crossovers further constrain the solution; this is particularly useful at high latitudes where there are no T/P-E crossovers.

The cubic spline representation of the ERS orbit error $E(t)$ is given by

$$E(t) = \sum_{k=1}^{K+4} c_k N_k(t) \quad (1)$$

$N_k(t)$ is the so-called normalized third-degree B spline defined on knots $l_k, l_{k+1}, l_{k+2}, l_{k+3},$ and l_{k+4} . The ERS orbit error reduction therefore consists of determining the $K + 4$ coefficients c_k which minimize the function $F(c_1, c_2, \dots, c_K)$

$$F = \sum_{i=1}^N w_i^2 [dc(t_i) - E(t_i)]^2 + \sum_{j=1}^M w_j^2 [E(t_{jd}) - E(t_{jd}) - d_j]^2 \quad (2)$$

where $dc(t_i)$ is the dual crossover difference (T/P-ERS) at crossover i , which provides a direct estimate of ERS orbit error $E(t_i)$; t_{ja}, t_{jd} , and d_j are the times of the ascending and descending arcs and the ERS-ERS crossover differences at crossover j , respectively; N is the number of T/P-E dual crossovers; M is the number of E-E crossovers; and w_i are the weights for each observation.

To avoid removing part of the oceanic signal in the adjustment, the method applied by *Le Traon et al.* [1995a, b] only used T/P-E and E-E crossovers with time differences below 5 days. The method has now been somewhat refined. For T/P-E crossovers the T/P sea surface height estimation is linearly interpolated from the two T/P cycles surrounding the ERS arc (the two T/P cycles are separated by 10 days, the T/P cycle duration). This provides an estimate of the sea surface height at the time of the ERS arc and minimizes the oceanic signal contribution in the T/P-E crossover differences. E-E crossovers with time differences of up to 10 days are now also used but with weights w_i depending linearly on time differences. Weights decrease by a factor of 2 from 0 to 10 days. This was derived from an analysis of E-E crossover differences since weights should be inversely proportional to the root-mean-square (rms) crossover differences. The weights for T/P-E crossovers correspond to the weights of E-E crossovers with no time differences. At latitudes over 66° , where there are no T/P-E crossovers, the E-E crossover weights are multiplied by 2 to better constrain the solution, as explained by *Le Traon et al.* [1995a]; see also the discussion in section 3. Note, however, that there is no downweighting for the larger number of crossovers at high latitudes. T/P-E and E-E crossovers in shallow regions (<200 m) are not used in the minimization.

The strategy for positioning the knots was also slightly modified. Two knots are initially placed at the first and last T/P-E crossovers on a given ERS arc. When the two knots are more than 10,000 km apart and there are more than 20 T/P-E crossovers on the arc, an extra knot is placed at the middle T/P-E crossover on the arc. If the two knots are less than 5000 km apart or if there are fewer than 10 T/P-E crossovers on the arc, the second knot is removed. This typically corresponds to 2500 knots in all for about 50,000 dual crossovers and 20,000 single ERS-1 crossovers for a full 35 day cycle.

3. Accuracy of ERS Orbit Error Estimation

The adjustment of ERS data using T/P as a reference was previously shown to provide a very good estimate of ERS orbit error. This was assessed through crossover and repeat-track analysis before and after orbit error correction [*Le Traon et al.*, 1995a, b]. For a more quantitative assessment the formal error on the estimation is now calculated. The formal error is derived from standard least squares theory, but the calculation is complicated by the spline representation. The ERS orbit error estimation consists of minimizing the function F . F can be written as $(\mathbf{AX}-\mathbf{B})(\mathbf{AX}-\mathbf{B})^T$, where \mathbf{X} is the vector of unknowns, i.e., the spline coefficients c_k . The standard least squares solution is $\mathbf{X} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{B}$, and the covariance matrix of \mathbf{X} , $\text{Cov}(\mathbf{X})$ is given by $\text{Cov}(\mathbf{X}) = (\mathbf{A}^T\mathbf{A})^{-1}\sigma^2$, where σ^2 is the least squares residual (F) after minimization.

To obtain the error on the orbit estimation $E(t)$ from the error on \mathbf{X} , the nondiagonal terms of $\text{Cov}(\mathbf{X})$ have to be taken into account

$$\langle E(t)^2 \rangle = \sum_{k=1}^{K+4} \sum_{k'=1}^{K+4} \langle c_k c_{k'} \rangle N_k(t) N_{k'}(t) \quad (3)$$

This implies calculating the inverse of the normal matrix $\mathbf{A}^T\mathbf{A}$ and estimating the variance of the error on the orbit error estimation $\langle E(t)^2 \rangle$ for any given t , given (3). This represents the fitting error only, which depends on the number and position of spline knots and on the least squares residual. It will not include the error due to orbit or large-scale errors unresolved by the spline representation.

The error estimate was used to refine the strategy for choosing the position of spline knots (see section 3). Plate 1a shows the estimated formal error when only T/P-E crossovers are used in the adjustment. The error is below 3 cm for latitudes below 66° and can reach 5 cm at high latitudes. The Mediterranean is also not very well constrained because of its geometry. This confirms, nevertheless, the very good fit obtained with the T/P-E crossovers. When the E-E crossovers are used, the error decreases to less than 2 cm below 66° and to about 3 cm at higher latitudes and in the Mediterranean (Plate 1b). With the overweighting of E-E crossovers at high latitudes the error is much more uniform, but given the overweighting, the result is not really representative of the true accuracy.

4. Application

4.1. Data

The method was applied to (1) ERS-1 cycles 6–18 of the first 35 day repeat cycle mission (October 1992 to December 1993) (ERS-1 phase C), (2) the two ERS-1 168 day cycles (geodetic mission) (ERS-1 phases E and F) split into eight subcycles of

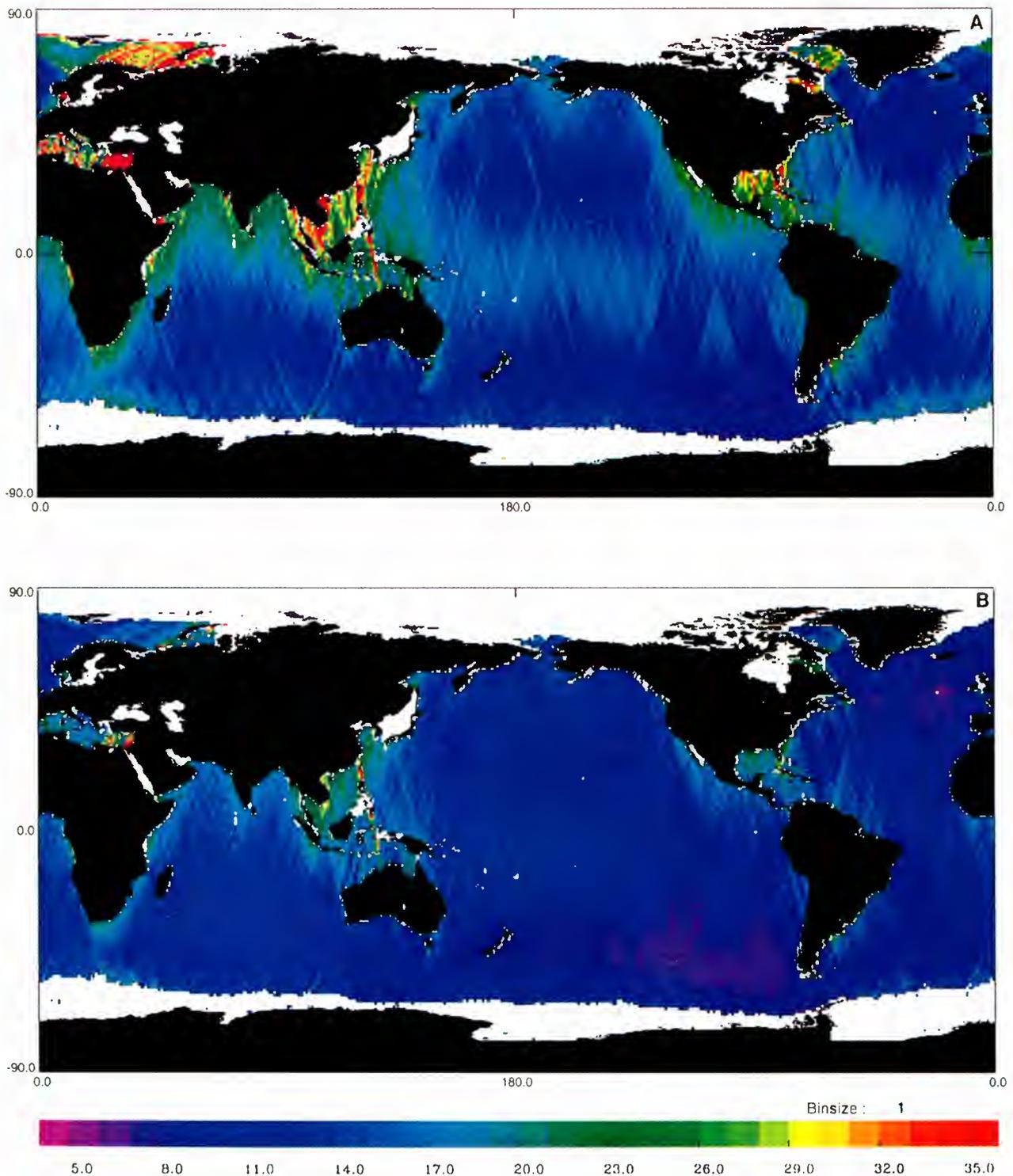


Plate 1. Formal error on ERS orbit error estimation (in centimeters) when (a) only TOPEX/POSEIDON (T/P)-E crossovers are used in the minimization and (b) when T/P-E and E-E crossovers are used in the minimization.

37 days and two subcycles of 20 days (April 1994 to March 1995). (3) ERS-1 cycles 1–13 of the second 35 day repeat cycle mission (March 1995 to June 1996) (ERS-1 phase G), and (4) ERS-2 cycles 1–16 (35 day repeat cycle mission) (May 1995 to November 1996) (ERS-2 phase A). T/P merged geophysical data records (GDR) distributed by Archiving, Validation, and Interpretation of Satellite Oceanographic Data (AVISO) [AVISO, 1996a] and ERS-1/2 ocean products (OPRs) distrib-

uted by Centre ERS d'Archivage et de Traitement (CERSAT) were used [CERSAT, 1994, 1996]. T/P data are the reprocessed merged GDRs (GDR-M, version C) [AVISO, 1996a]. They include the Center for Space Research (CSR) 3.0 tidal model [Eanes *et al.*, 1995], the National Aeronautics and Space Administration (NASA) and Centre National d'Etudes Spatiales (CNES) Joint Gravity Model 3 (JGM 3) orbits [Tapley *et al.*, 1996; Barotto and Berthias, 1998], and a series of minor cor-

Table 1. TOPEX/POSEIDON and ERS-1/2 Data Sets Used for the Study

	TOPEX/POSEIDON	ERS
GDRs	GDR-M products, version C [AVISO, 1996a]	OPR products, version 3 for ERS-1 phases C, E, and F and version 6.2 for ERS-1 phase G and ERS-2 phase A [CERSAT, 1994, 1996]
Orbit	CNES JGM 3/ELFE orbits	D-PAF precise orbit referenced to TOPEX ellipsoid
Geophysical corrections		
Dry troposphere and inverse barometer effect	from ECMWF	from ECMWF, this is an upgrade for the ERS-1 version 3 products which used the ARPEGE model
Wet troposphere	from TMR radiometer	from ATSR-M radiometer ^a
Ionosphere	from dual-frequency altimeter range measurements for TOPEX data ^b , from DORIS for POSEIDON data	BENT model
Sea state bias	BM4 [Gaspar <i>et al.</i> , 1994]	version 3, -5.5% of SWH [Gaspar and Ogor, 1994]; version 6.2, BM3 [Gaspar and Ogor, 1996] (update of the OPR)
Tides		
Ocean tide and loading tide	CSR3.0 [Eanes and Bettadpur, 1995]	CSR3.0 (update of the OPR)
Solid Earth tide	Cartwright and Tayler [1971]	Cartwright and Tayler [1971]
Pole tide	applied	not applied

^aCorrected as indicated on the CLS ERS quality assessments [CLS, 1992, 1996a, b] and extrapolated near the coasts.

^bSmoothed using a Lanczos filter (300 km).

TMR is TOPEX microwave radiometer, ATSR-M is along-track scanning radiometer-microwave, and DORIS is Doppler orbitography and radiopositioning integrated by satellite.

rections to the initial GDRs (e.g., sigma-0 calibration, polar tide correction). In addition, the TOPEX instrumental drift and bias were also corrected for. For ERS-1/2, D-PAF (processing and archiving facility) orbits were used [Massmann *et al.*, 1997]. When necessary, ERS-1/2 altimetric corrections were updated to make the T/P and ERS-1/2 corrections homogeneous. The CSR3.0 tidal model was thus used. The dry tropospheric and inverse barometer corrections, derived from the French meteorological model ARPEGE (until the end of the ERS-1 geodetic mission), were also replaced by corrections derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) model as for T/P. The corrections applied to T/P and ERS-1/2 data are summarized in Table 1. ERS-1/2 altimeter biases, ultra stable oscillator (USO) drift, and single-point target response (SPTR) corrections were not applied since they can be fully corrected for by the orbit error reduction method. The same holds for the pole tide, which is not present in ERS-1/2 OPR. Finally, time tag biases of -1.3 and -1.1 ms were applied to ERS-1 and ERS-2 range measurements, respectively. The biases were derived from crossover analysis.

Given the very high constraint of the fit due to the very low T/P orbit error, the method should be almost insensitive to the initial orbit used. This was checked by also performing the adjustment for the Delft JGM 3 precise orbits [Scharroo *et al.*, 1994] for ERS-1 phases C, E, and F and comparing with the results obtained from D-PAF orbits.

4.2. Results

Figures 1 and 2 show the statistics of T/P-E and E-E crossovers for all processed cycles for the D-PAF and JGM 3 ERS-1 orbits (phases C, E, and F), for the D-PAF ERS-1 orbits (phase G), and for the D-PAF ERS-2 orbits (phase A). Before orbit error adjustment the E-E crossovers for time lags below 10 days are between 15 and 18 cm (D-PAF orbits for ERS-1 phases C, E, and F), 10 and 14 cm (JGM 3 orbits for ERS-1 phases C, E, and F; D-PAF orbits for ERS-1 phase G and ERS-2 phase A). Similarly, the T/P-E crossover differences for time lags below 10 days are between 13 and 15 cm (D-PAF

orbits for ERS-1 phases C, E, and F) or between 10 and 12 cm (JGM 3 orbits 44 for ERS-1 phases C, E, and F; D-PAF orbits for ERS-1 phase G and ERS-2 phase A). Note that D-PAF orbit accuracy is enhanced for ERS-1 phase G and ERS-2 phase A and is about the same as JGM 3 orbits for ERS-1 phases C, E, and F. This mainly results from an update of the geopotential model (from precise gravity models PGM035 to PGM055) in the D-PAF orbit calculation [Massmann *et al.*, 1997]. Note also that starting from ERS-2 cycle 7, only, satellite laser ranging (SLR) data and Precise Range and Range Rate Experiment (PRARE) data are used as tracking data in the ERS-2 orbit calculation. For the other ERS-1/2 cycles both D-PAF and JGM 3 orbits used SLR and radar altimeter (RA) crossovers as tracking data.

The crossover statistics after adjustment for the two orbits (D-PAF and JGM 3) are almost the same. E-E and T/P-E rms crossover differences are about 6.5 and 7 cm, respectively. They are very similar to T/P-T/P rms crossover differences, which (with the same editing) are about 7 cm rms. T/P-E crossover differences are slightly larger than E-E crossover differences, probably because of the influence of T/P orbit error. Crossover statistics are slightly better with the JGM 3 orbit. The difference in variance is typically 1-2 cm². This suggests that the difference in geopotential models between the D-PAF and JGM 3 orbits introduced an additional small high-frequency signal in dual T/P-ERS crossover differences.

The adjusted D-PAF orbit error (i.e., $E(t)$) for ERS-1 phases C, E, and F varies between 9 and 12 cm rms for all processed cycles. The JGM 3 orbit for the same ERS-1 phases is more accurate; its estimated error is between 6 and 9 cm for all processed cycles. D-PAF orbit error for ERS-1 phase G and ERS-2 phase A varies between 6.5 and 9.5 cm, about the same as JGM 3. The rms (SLR plus PRARE) and (SLR plus RA) ERS-2 D-PAF orbit errors are similar. This agrees with results found by Massmann *et al.* [1997].

These estimates provide an independent measure of the accuracy of these ERS-1/2 orbits. They show that good accuracy (7-8 cm) is achieved. Still, these orbit errors are too large

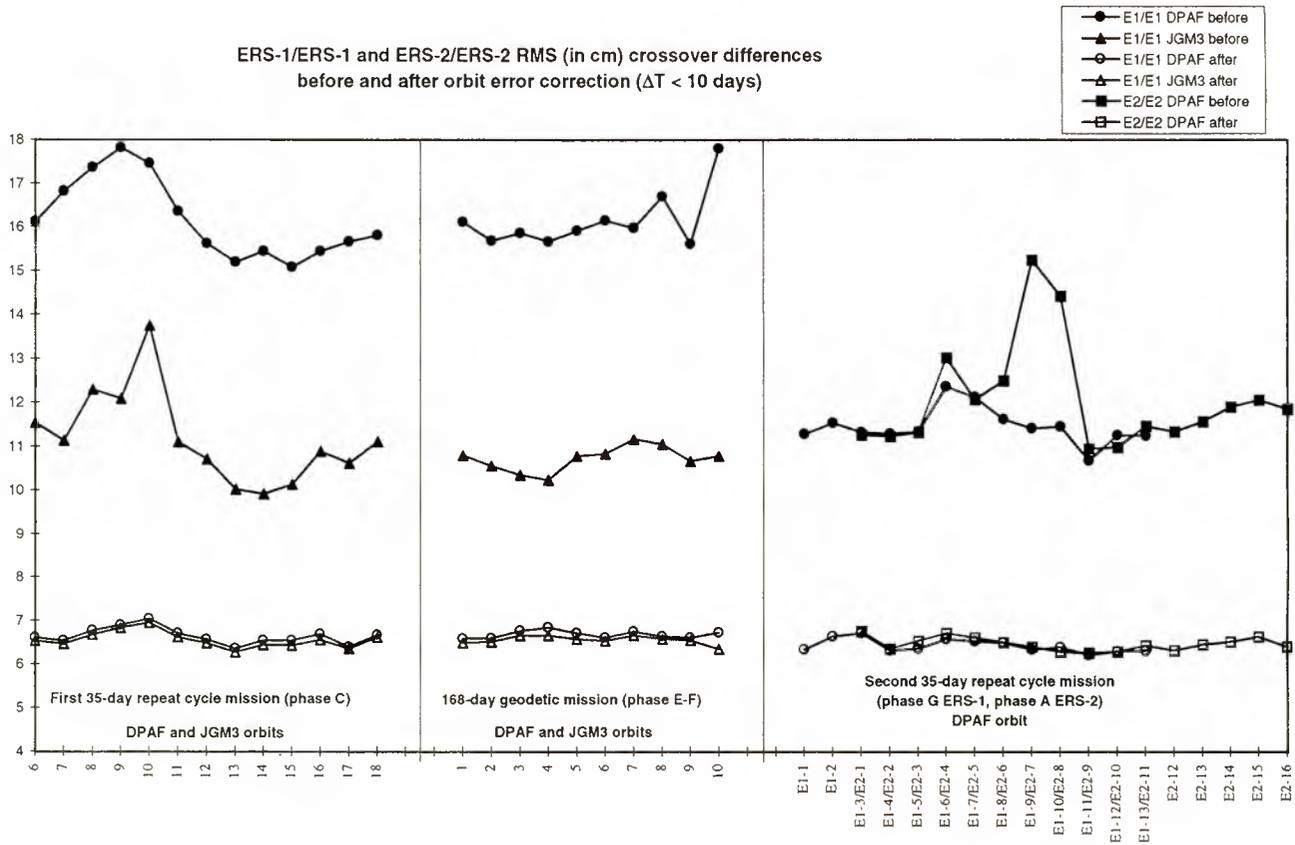


Figure 1. E-E crossover statistics for D-PAF and Joint Gravity Model 3(JGM 3) orbits for ERS-1 cycles 6–18 (phase C, 35 day repeat period), ERS-1 subcycles 1–10 (phases E and F, 168 day repeat period), ERS-1 cycles 1–13 (phase G), and ERS-2 cycles 1–16 (phase A) before and after orbit correction. Units are centimeters.

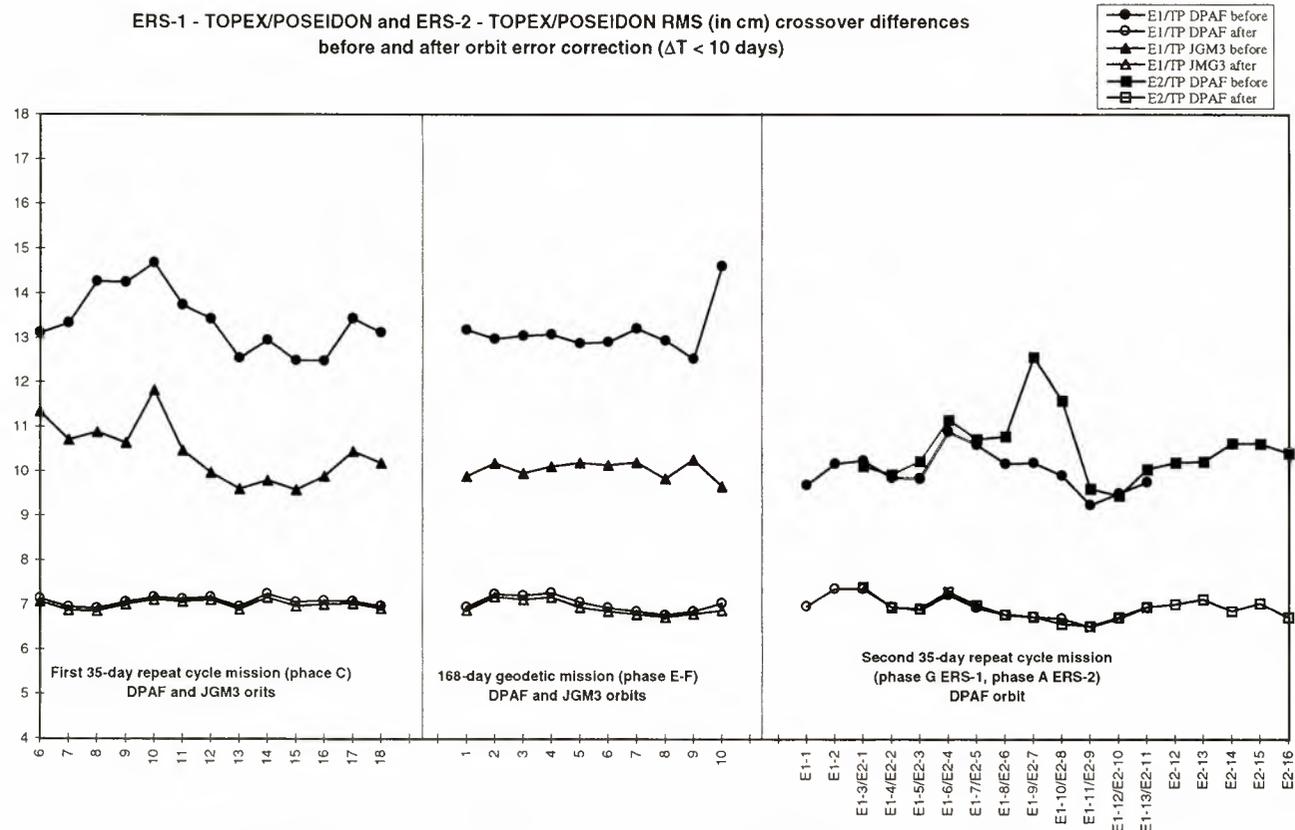


Figure 2. Same as Figure 1 but for T/P-E crossovers. Units are centimeters.

for analyzing the large-scale oceanic signals. The nondynamical adjustment onto T/P is needed to reduce them to a level comparable to T/P orbit error.

4.3. Comparison of D-PAF and JGM 3 Orbits Before and After Orbit Error Correction (ERS-1 Phases C, E, and F)

The comparison of sea surface heights obtained with the D-PAF and JGM 3 orbits (i.e., the difference in orbits) for ERS-1 cycles 6–18 gives a mean rms difference of 11 cm. About half of the difference is constant over time and corresponds to the difference in geographically correlated orbit errors between the D-PAF and JGM 3 orbits (Plate 2a). The variable part of the difference has an rms difference of 8 cm rms (Plate 3a). After orbit error correction the rms difference is only 1.2 cm. About half of the difference between the two corrected solutions is constant over time and is mainly concentrated in equatorial regions (Plate 2b). This probably corresponds to the difference in high-frequency geographically correlated D-PAF and JGM 3 orbit errors which cannot be removed with the spline representation used. This is exemplified on Plate 4, which shows the difference between the two orbits along a given track and the difference between the two orbit error corrections. While the adjustment removes all the long wavelength differences between the two orbits, it cannot remove high-frequency signals which have, in this example, wavelengths shorter than 5000 km. To remove the high-frequency geographically correlated orbit, far more spline knots would be needed. This actually could be achieved by adjusting ERS-1/2 mean profile onto T/P mean profile. In this case the noise at crossovers would be much smaller (because of averaging), and more knots could be added without increasing the adjustment error.

The variable part of the difference (Plate 3b) has a mean rms of 0.7 cm only. The rms is less than 0.5 cm in most areas and reaches 2 cm at high latitudes and in the eastern Mediterranean. It is consistent with the formal error estimates shown in Plate 1b. Note that sea level anomaly is only affected by the variable part of the orbit error. The corrected sea level anomaly for D-PAF and JGM 3 will thus be identical at the centimeter level.

This comparison thus shows that the method is almost insensitive to the initial orbit used because of the very high constraint of the fit. It also confirms the 2 cm precision of the adjustment.

4.4. Influence of T/P Orbit Error

Given the very low T/P orbit error, the estimation of the ERS-1/2 orbit error should not be sensitive to it. T/P JGM 3 orbit error is about 2 cm, and this is not the main source of error in T/P-E crossover differences. Tests were performed using the two available T/P orbits (NASA JGM 3 orbits and CNES JGM 3 estimation by empirical smoothing and filtering (ELFE) orbits). They agree to within about 1.5 cm rms. The influence of these orbits on the ERS orbit error correction is, however, always below 0.5 cm. This means that T/P orbit error is partly filtered out by our method. Of course, the very low frequency part (e.g., frequencies lower than 10 days^{-1}) of the T/P orbit error will directly impact the ERS orbit error estimation and will induce biases on the ERS orbit error. However, the low-frequency part of the JGM 3 T/P orbit error is small [Marshall *et al.*, 1995] and certainly much smaller than the corresponding ERS-1/2 orbit errors.

5. Sea Level Variability

Repeat-track analysis was performed for ERS-1 cycles 6–18 both for D-PAF and JGM 3 orbits. The rms sea level variabilities before orbit error correction for D-PAF and JGM 3 orbits and after orbit error correction for JGM 3 are shown on Plates 5a, 5b, and 6a, respectively. The result with the corrected D-PAF orbit (not shown) is, as expected, almost identical to Plate 6a. Signals in Plates 5a and to a lesser extent 5b are still dominated by orbit error, especially in low-energy areas. The mean difference in sea level variance before and after orbit error correction is 34 cm^2 (D-PAF orbit) and 17 cm^2 (JGM 3 orbit). The rms sea level variability computed from T/P data (Plate 6b) over the same period is in excellent agreement with the ERS-1 corrected sea level variability (Plate 6a). The difference in noise between the T/P and ERS-1 missions explains the slightly larger signals given by ERS-1 in low eddy energy. In high eddy energy (e.g., Gulf Stream) the signals given by ERS-1 are actually smaller. The difference in variance (not shown) is about 40 cm^2 , i.e., about 5% of the signal variance. This is because the mean, which is calculated with 3.5 times fewer cycles than T/P, has absorbed a small part of the oceanic signal. As a result, the mean global difference in variance between the ERS-1 and T/P maps is slightly negative (-2 cm^2).

These results show that the ERS-1 data have to be corrected for orbit error to map the oceanic signal. This is shown in Plates 7a and 7b, the mid-November 1992 maps of sea level anomaly obtained with the JGM 3 orbit and the corrected JGM 3 orbit. The maps were obtained with a global suboptimal space-time objective analysis method. The enhancement of ERS-1 data accuracy using T/P as a reference is plain to see. A similar map was obtained using T/P data. The map (Plate 7c) shows that the large-scale features are in excellent agreement with the ERS-1 map. In particular, the large-scale steric effects related to the heating/cooling of surface waters, which are the most difficult signal to preserve in conventional orbit error removal schemes [e.g., *Tai*, 1991], are completely preserved. The differences at smaller scales mainly represent the different space and time sampling of the two satellites. This shows that the T/P and ERS-1/2 data sets are much more consistent after orbit error correction and that they can be merged in the analyses.

6. Conclusion

Formal error estimates and crossover and repeat-track analyses demonstrate that ERS-1/2 orbit error can be estimated to within about 2 cm rms using T/P as a reference. This is comparable to T/P orbit accuracy. The adjustment will also remove any long wavelength error in ERS-1/2 data (e.g., altimeter bias and drift, USO drift, large-scale error in ionospheric correction) which is necessary to obtain consistent T/P and ERS-1/2 data sets. It is also shown that ERS-1/2 orbit error must be corrected before analyzing large-scale oceanic signals and combining ERS-1/2 with T/P data.

The method will be applied on an operational basis to ERS-2 data when they are produced by CERSAT. Corresponding data sets (corrected sea surface heights and sea level anomalies files) will be distributed to the scientific community by AVISO [AVISO, 1996b, 1997]. These consistent, homogeneous T/P, ERS-1, and ERS-2 data sets will then be merged using the mapping method described by *Le Traon et al.* [1998] to analyze the sea level variability with a high resolution and a high accuracy.

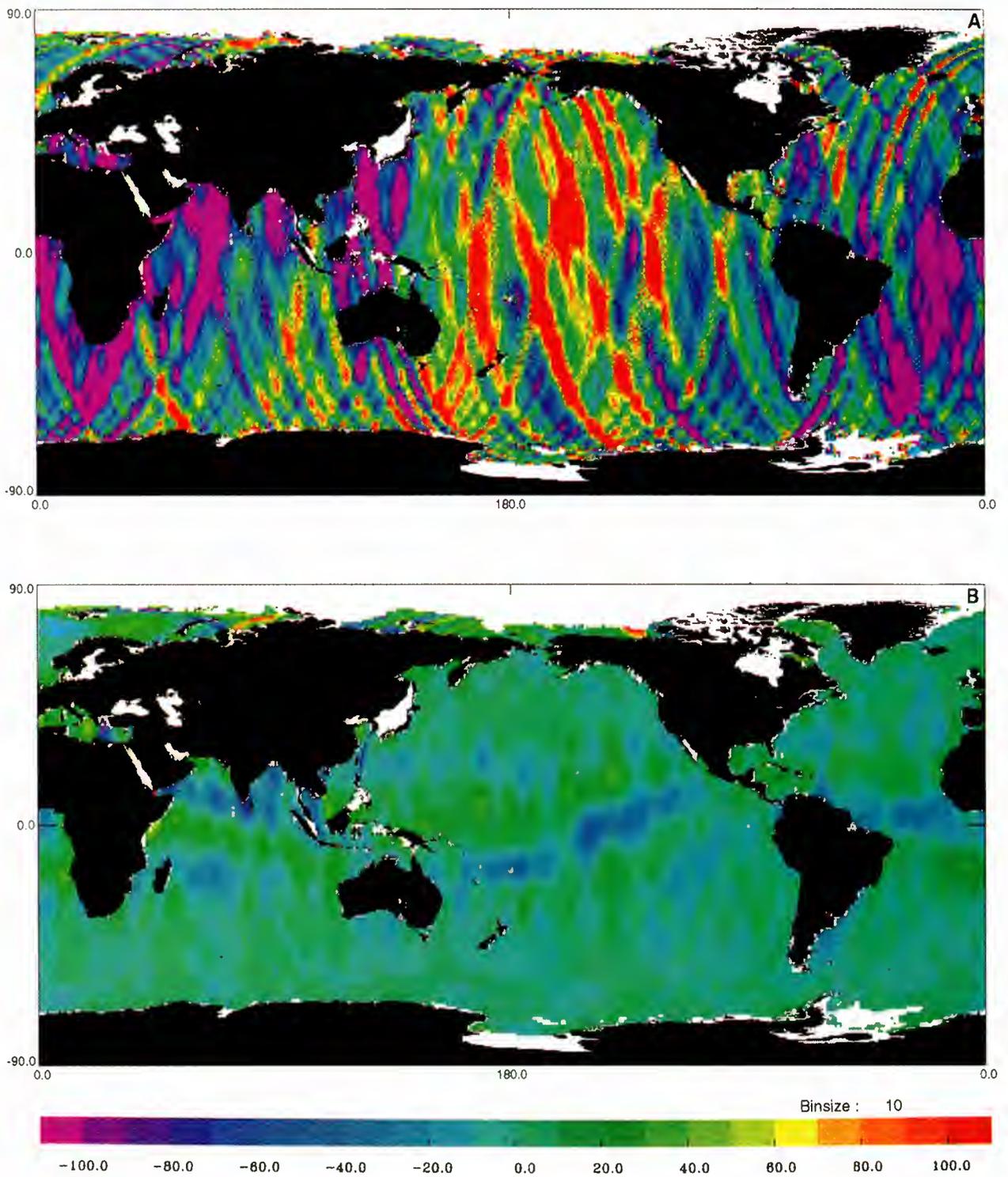


Plate 2. Mean difference (ERS-1 cycles 6–18) in corrected sea surface heights for D-PAF and JGM 3 orbits (a) before and (b) after orbit error reduction.

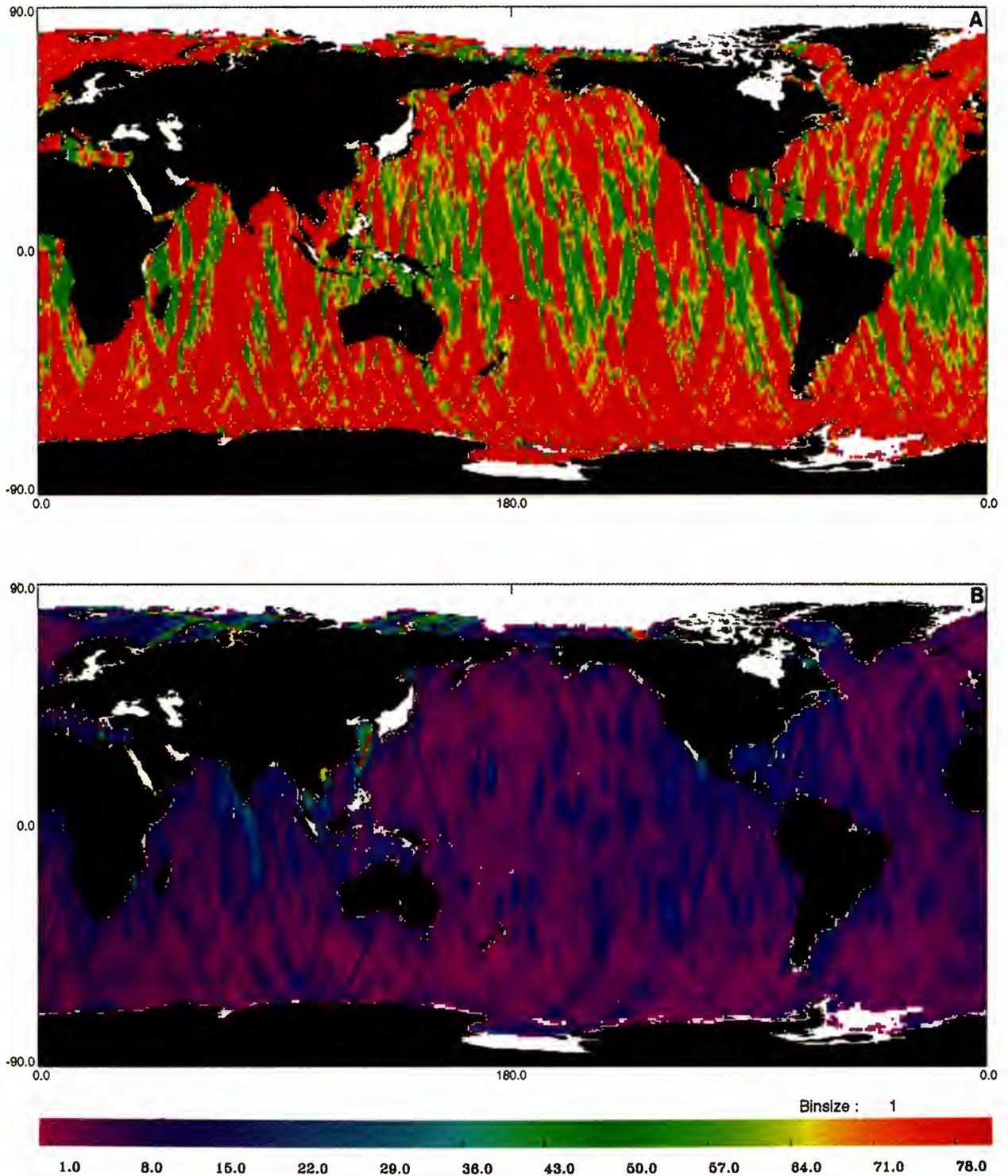
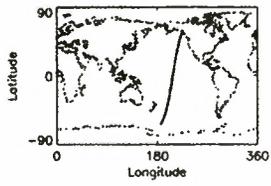


Plate 3. Standard deviation of the difference (ERS-1 cycles 6–18) in corrected sea surface heights between D-PAF and JGM 3 (a) before and (b) after orbit error reduction.



ERS-1 Cycle 006, pass 600

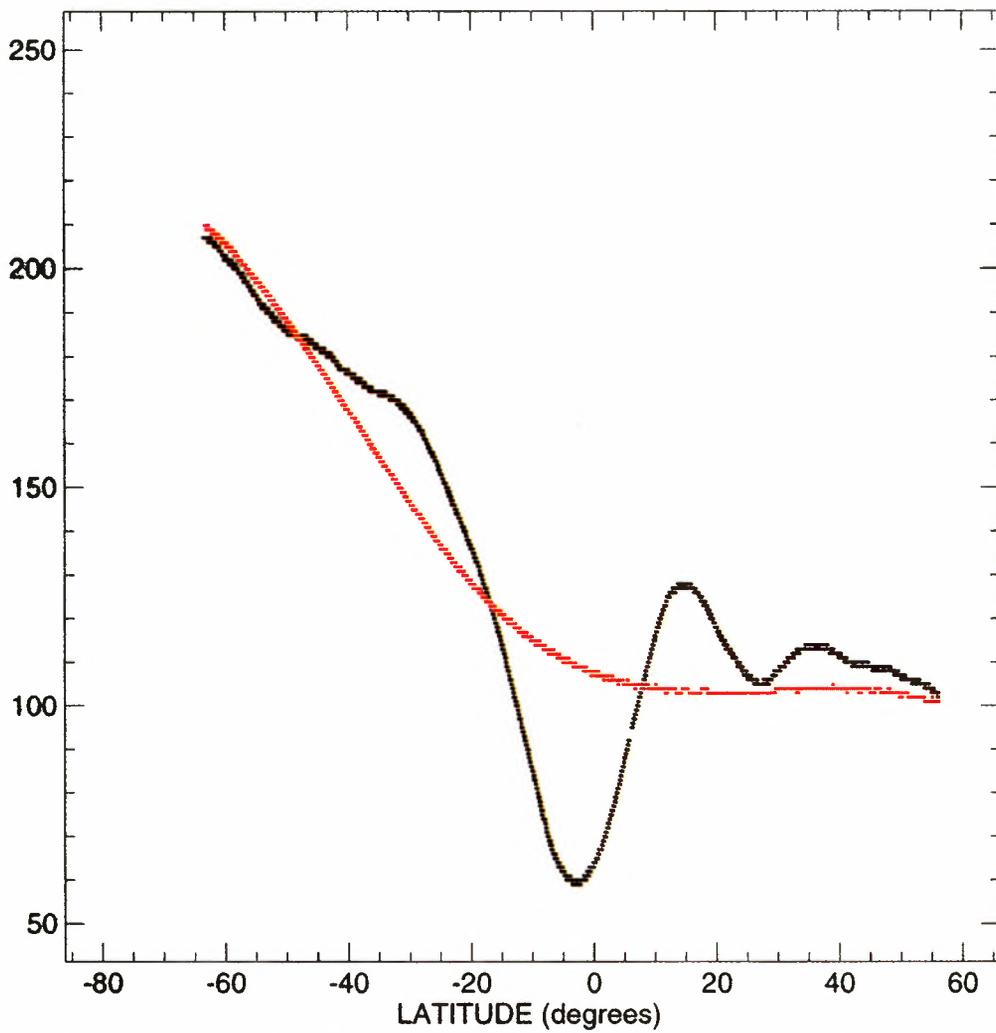


Plate 4. Difference between D-PAF and JGM 3 orbits along ERS-1 track 600 and difference between the D-PAF and JGM 3 orbit error corrections. Units are millimeters.

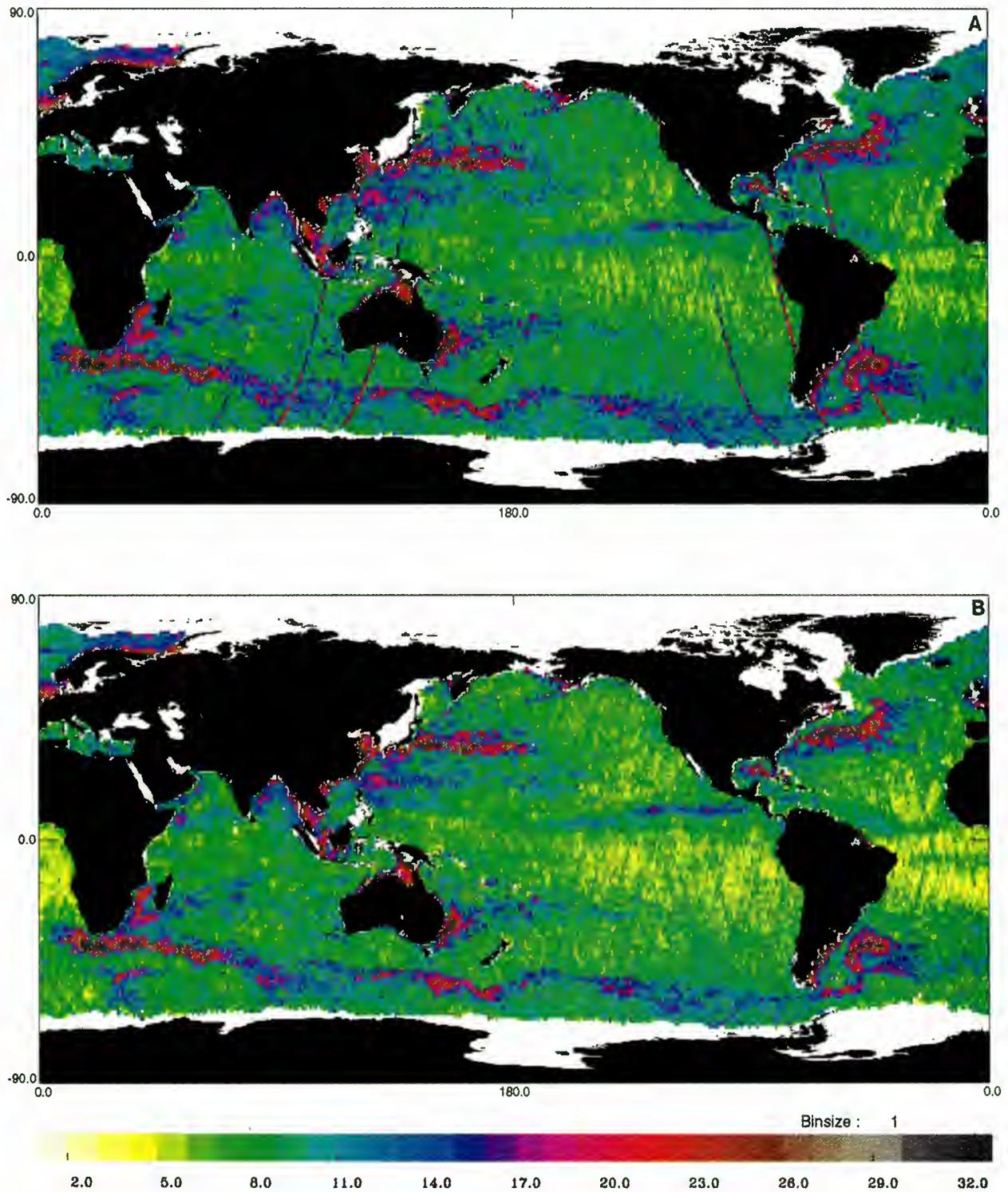


Plate 5. The root-mean-square (rms) sea level variability for ERS-1 cycles 6-18 before orbit error correction for (a) D-PAF and (b) JGM 3 orbits. In centimeters.

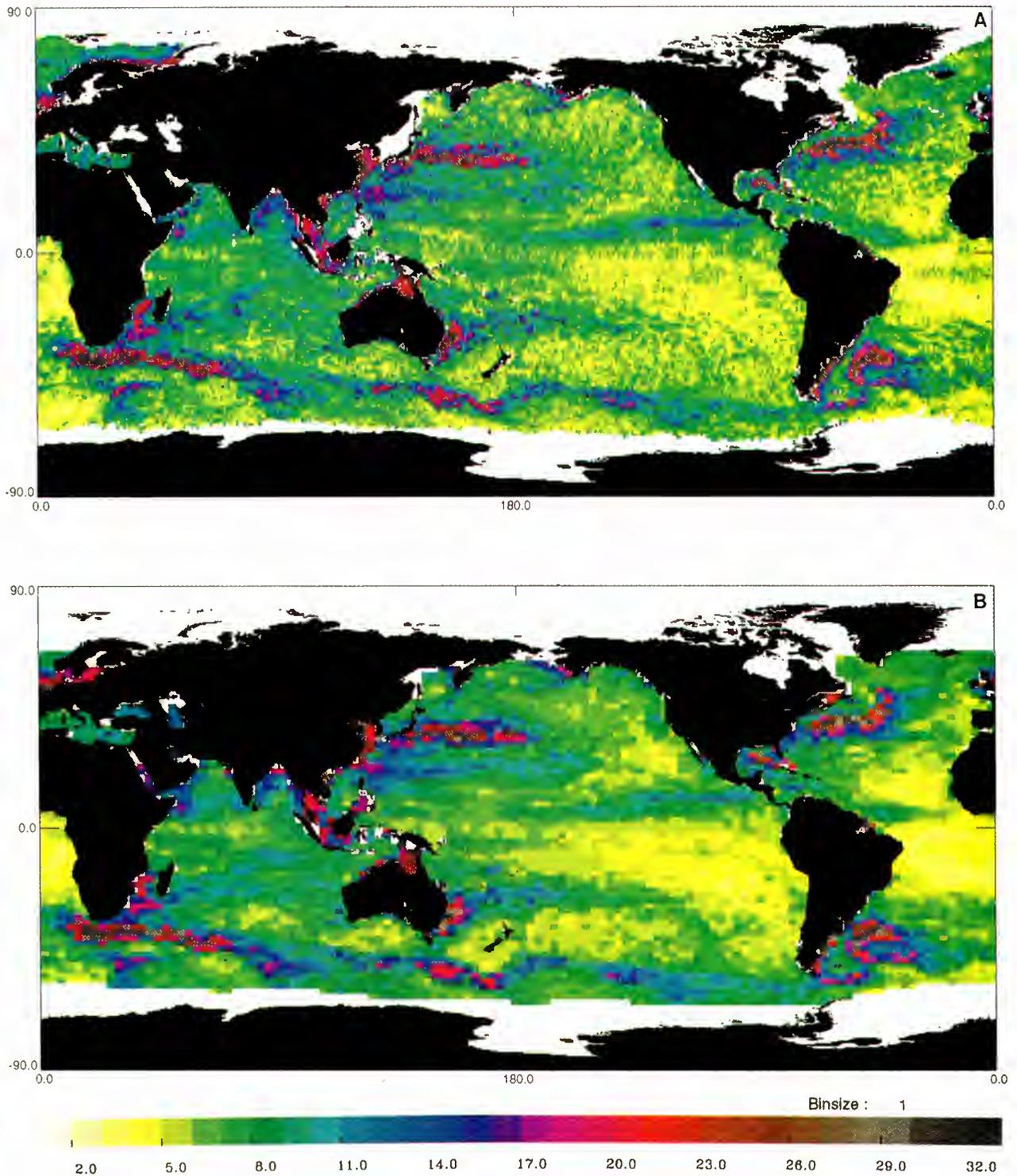


Plate 6. (a) The rms sea level variability for ERS-1 cycles 6-18 after orbit error correction. (b) The rms sea level variability for TOPEX/POSEIDON over the period corresponding to ERS-1 cycles 6-18 (October 1992 to December 1993). In centimeters.

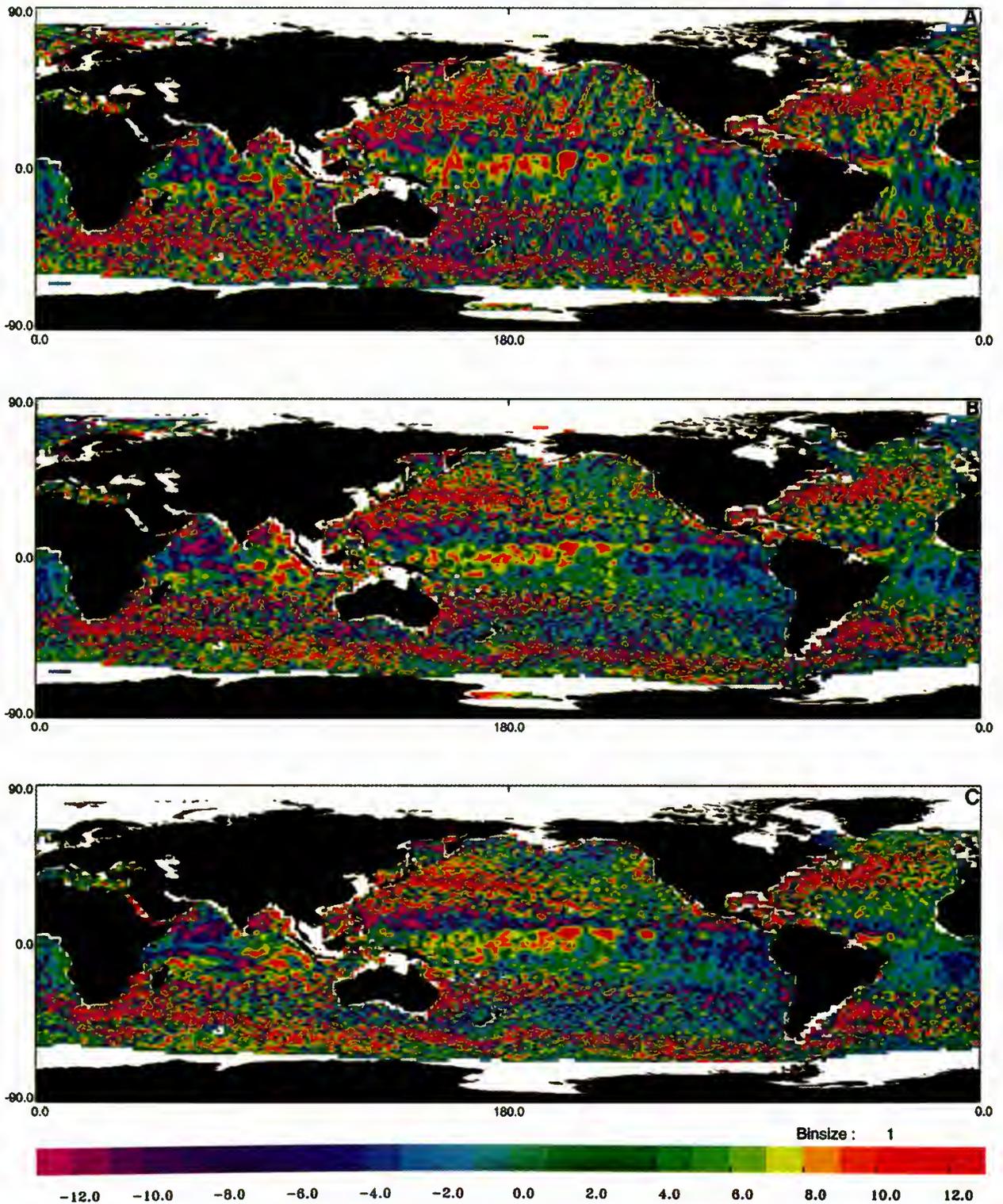


Plate 7. Map of ERS-1 sea level anomaly obtained using (a) the JGM 3 orbit and (b) the corrected JGM 3 orbit for mid-November 1992. (c) Map of T/P sea level anomaly for mid-November 1992. In centimeters.

Acknowledgments. JGM 3 ERS-1 orbits were kindly provided to us by Remko Scharroo from Delft University. The study was supported by the Midi-Pyrénées region, the Environment and Climate AGORA project of the European Union and the French Space Agency (CNES). This work is part of the ESA proposal AO2.F105 "Joint analysis of ERS-1, ERS-2 and TOPEX/POSEIDON altimeter data for oceanic circulation studies."

References

- Archiving, Validation, and Interpretation of Satellite Oceanographic Data, AVISO handbook for merged TOPEX/POSEIDON products, 3rd ed., *Handb. AVI-NT-02-101-CN*, Toulouse, France, 1996a.
- Archiving, Validation, and Interpretation of Satellite Oceanographic Data, AVISO user handbook corrected sea surface height (CORSSH), 2nd ed., *Handb. AVI-NT-011-311-CN*, Toulouse, France, 1996b.
- Archiving, Validation, and Interpretation of Satellite Oceanographic Data, AVISO user handbook: Sea level anomalies (SLA), 2nd ed., *Handb. AVI-NT-011-312-CN*, Toulouse, France, 1997.
- Barotto, B., and J. P. Berthias, First results of reduced dynamics with DORIS on TOPEX/POSEIDON and SPOT, *J. Guidance Control*, in press, 1998.
- Carnochan, S., P. Moore, and S. Ehlers, ERS-1 radial positioning refinement by dual crossover analysis with TOPEX/POSEIDON, *Adv. Space Res.*, **16**, 119–122, 1995.
- Cartwright, D. E., and R. J. Tayler, New computations of the tide-generating potential, *Geophys. J. R. Astron. Soc.*, **33**, 253–264, 1971.
- Centre ERS d'Archivage et de Traitement, ERS-1 altimeter products user manual, *Handb. C1-EX-MUT-A21-01-CN*, Brest, France, 1994.
- Centre ERS d'Archivage et de Traitement, Altimeter and microwave radiometer ERS products user manual, *Handb. C2-MUT-A-01-IF*, Brest, France, 1996.
- Collecte, Localisation, Satellites, Quality assessment of CERSAT altimeter products, *Rep. CLS.OC/NT/93.005*, Agne, France, 1992.
- Collecte, Localisation, Satellites, Validation of ERS-1/2 OPR cycles (ERS-1), *Rep. CLS.OC/NT/96.010*, Agne, France, 1996a.
- Collecte, Localisation, Satellites, Validation of ERS-1/2 OPR cycles (ERS-2), *Rep. CLS.OC/NT/96.011*, Agne, France, 1996b.
- Eanes, R. J., and S. V. Bettadpur, The CSR 3.0 Global Ocean Tide Model, *Rep. CSR-TM-95-06*, Cent. for Space Res., Univ. of Texas, Austin, 1995.
- Fu, L.-L., E. J. Christensen, C. A. Yamarone Jr., M. Lefebvre, Y. Menard, M. Dorrer, and P. Escudier, TOPEX/POSEIDON mission overview, *J. Geophys. Res.*, **99**, 24,369–24,381, 1994.
- Gaspar, P., and F. Ogor, Estimation and analysis of the sea state bias of the ERS-1 altimeter, technical report of IFREMER contract 94/2.426016/C, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, 1994.
- Gaspar, P., and F. Ogor, Estimation and analysis of the sea state bias of the new ERS-1 and ERS-2 altimetric data (OPR version 6), *Tech. Rep. CLS/DOS/NT/96.041*, Collect., Localisation, Satell., Agne, France, 1996.
- Gaspar, P., F. Ogor, P. Y. Le Traon, and O. Z. Zanife, Estimating the sea state bias of the TOPEX and POSEIDON altimeters from cross-over differences *J. Geophys. Res.*, **99**, 24,981–24,994, 1994.
- Hayes, J. G., Numerical methods for curve and surface fitting, *Bull. Inst. Math. Appl.*, **10**, 144–152, 1974.
- Hernandez, F., P. Y. Le Traon, and R. Morrow, Mapping mesoscale variability of the Azores Current using TOPEX/POSEIDON and ERS-1 altimetry, together with hydrographic and Lagrangian measurements, *J. Geophys. Res.*, **100**, 24,995–25,006, 1995.
- Le Traon, P. Y., P. Gaspar, F. Bouyssel, and H. Makhmaraa, Use of TOPEX/POSEIDON data to enhance ERS-1 data, *J. Atmos. Oceanic Technol.*, **12**, 161–170, 1995a.
- Le Traon, P. Y., P. Gaspar, F. Ogor, and J. Dorandeu, Satellites work in tandem to improve accuracy of data, *Eos Trans. AGU*, **76**(39), 385, 389, 1995b.
- Le Traon, P. Y., F. Nadal, and N. Ducet, An improved mapping method of multiple satellite altimeter data, *J. Atmos. Oceanic Technol.*, in press, 1998.
- Marshall, J. A., N. P. Zelensky, S. M. Klosko, D. S. Chinn, S. B. Luthcke, K. E. Rachlin, and R. G. Williamson, The temporal and spatial characteristics of TOPEX/POSEIDON radial orbit error, *J. Geophys. Res.*, **100**, 25,331–25,352, 1995.
- Massmann, F. H., K. H. Neumayer, J. C. Raimondo, K. Enninghorst, and H. Li, Quality of the D-PAF orbit before and after the inclusion of PRARE data, paper presented at 3rd ERS Symposium, Eur. Space Agency, Florence, Italy, March 17–21, 1997.
- Rapp, R. H., Y. Yi, and Y. M. Wang, Mean sea surface and geoid gradient comparisons with TOPEX altimeter data, *J. Geophys. Res.*, **99**, 24,657–24,667, 1994.
- Scharroo, R., K. F. Wakker, and G. J. Mets, The orbit determination accuracy of the ERS-1 mission, in *Proceedings of the Second ERS-1 Symposium—Space at the Service of our Environment, Hamburg, Germany, 11–14 October 1993*, Eur. Space Agency Spec. Publ., *ESA SP-361*, 735–740, 1994.
- Tai, C. K., How to observe the gyre to global-scale variability in satellite altimetry: Signal attenuation by orbit error removal, *J. Atmos. Oceanic Technol.*, **8**, 271–288, 1991.
- Tapley, B. D., et al., The Joint Gravity Model 3, *J. Geophys. Res.*, **101**, 28,029–28,049, 1996.

P.-Y. Le Traon and F. Ogor, CLS Space Oceanography Division, 8-10 rue Hermes, Parc Technologique du Canal, 31526 Ramonville Saint-Agne, France. (e-mail: letraon@cls.cnes.fr; ogor@cls.cnes.fr)

(Received November 22, 1996; revised June 12, 1997; accepted July 3, 1997.)

Complex empirical orthogonal functions analysis of ERS-1 and TOPEX/POSEIDON combined altimetric data in the region of the Algerian current

Catherine Bouzinac

Instituto de Ciencias del Mar, Consejo Superior de Investigaciones Científicas, Barcelona, Spain

Jorge Vazquez

Jet Propulsion Laboratory, California Institute of Technology, Pasadena

Jordi Font

Instituto de Ciencias del Mar, Consejo Superior de Investigaciones Científicas, Barcelona, Spain

Abstract. Maps of sea level anomalies (SLA) relative to the 1993 annual mean sea level combine the data from the two altimetric missions, ERS-1 and TOPEX/POSEIDON, during the overlap period (October 1992 to December 1993). These regular maps in space and time of residual sea level every 10 days on a 0.2° regular grid are used in the region of the Algerian current where the mesoscale eddies are of primary importance to the circulation of all the Mediterranean water masses. They are first compared with ERS-1 along-track scanning radiometer sea surface temperature images to get information on two anticyclonic eddies produced by instabilities of the Algerian current and visible in both infrared and altimetric data sets. Then, an analysis of complex empirical orthogonal functions (CEOFs) is performed on the SLA data set to see the correlation of the different dynamic features of the observed variability. The CEOF analysis is applied to the complex time series formed from the original SLA time series and their Hilbert transforms to separate the variability into spatially coherent modes. The spatially correlated signal in the study area ($0\text{--}15^\circ\text{E}$ and $35^\circ\text{--}40^\circ\text{N}$) was found to be dominated by the first two CEOFs. These first two modes explain nearly 85% of the variability, with 80% of the total variance for the first one and 5% of the total variance for the second one. The temporal phase of the first mode indicates that a constant frequency of one cycle per year is clearly dominant, corresponding to the seasonal signal. The strongest amplitude is obtained in the southern part of the channel of Sardinia and south of the Strait of Sicily. The temporal amplitude and the temporal phase of the second mode show a periodicity of about 6 months which appears to be associated with the variability of the Algerian current as the phase isolines are parallel to the mean current path along the Algerian coast. The strongest amplitude of the second mode is located near the African coast at $\sim 4^\circ\text{E}$ and 8°E . These two points of high variability could correspond to eddy detachments from the main current.

1. Introduction

The Atlantic water enters the western Mediterranean Sea by the Strait of Gibraltar in the surface layer. In the Alboran basin it mixes with saltier Mediterranean waters and reaches the Greenwich meridian with a salinity >36.5 psu. This eastward flow of Modified Atlantic Water (MAW) then forms a well-defined jet along the Algerian coast called the Algerian current which is driven by the density difference between the Atlantic and the Mediterranean surface waters [Millot, 1985]. The Algerian current is hence a coastal current which flows eastward along the North African coast and drives the MAW from the Alboran basin to the channel of Sardinia and the Strait of Sicily. Meanwhile, part of the MAW spreads into the Algerian basin by instabilities of the current which begin to meander at

1° or 2°E (Figure 1) [Millot, 1987]. The meanders can then generate cyclonic and anticyclonic eddies. However, the cyclonic ones disappear rapidly, while the anticyclonic ones can develop as they propagate eastward with the current and may detach from the main current [Millot, 1994]. Indeed, large anticyclonic eddies with diameters of 200 km have been observed in the Algerian basin during several months [Taupier-Letage and Millot, 1988]. These huge eddies, which can occupy an extensive part of the basin as far as Sardinia, could be old stages of eddies generated by the instabilities of the Algerian current which may have drifted after being detached from the coast. During their residence they are able to deviate seaward a significant portion of the Algerian current [Benzohra and Millot, 1995a]. Perkins and Pistek [1990] described an eastward flow of MAW in the northern part of the channel of Sardinia and a westward flow in the southern part during June 1986. This situation seemed to be due to the presence of an anticyclonic eddy impeding the usual MAW eastward circulation

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02909.
0148-0227/98/97JC-02909\$09.00

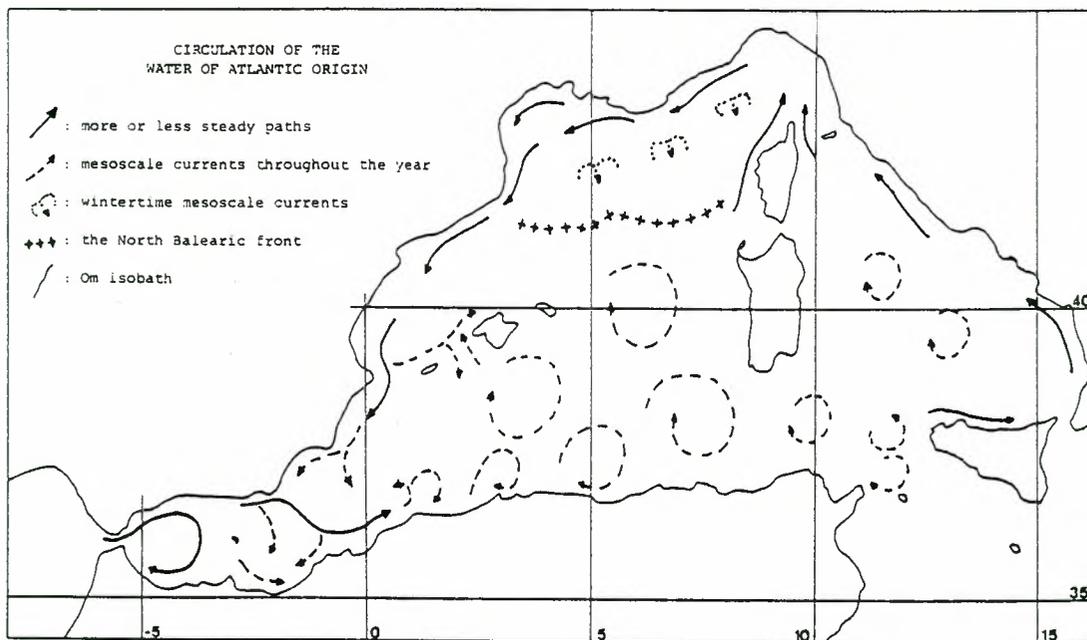


Figure 1. Modified Atlantic Water (MAW) circulation in the western Mediterranean Sea, from Millot [1987].

near the Tunisian coast. Recently, *Fuda et al.* [1997] have performed expendable bathy thermograph (XBT) sections between Tunisia and France. Their results show that at $\sim 8^\circ\text{E}$ near the channel of Sardinia some eddies detach quite periodically from the coast and drift northward and westward after reaching Sardinia, following the 2500 m isobath. This is supported by current-meter measurements in the western part of the channel during the Programme de Recherche Internationale en Mediterranee (PRIMO)-1 experiment (1993–1994), which revealed a 3 month long anticyclonic perturbation propagating northward at a depth of 200 m with a strong influence on the bottom circulation [Bouzinac *et al.*, 1997].

The Algerian eddies are hence of primary importance to the circulation of all the water masses in the western Mediterranean Sea [Millot *et al.*, 1990] because of their deep extension [Benzohra and Millot, 1995b]. In the surface layer the MAW circulation is affected by this mesoscale activity, and its temporal and spatial variability is thus very large. Its seasonal variability has also been observed in the Strait of Sicily by *Manzella et al.* [1990], who reported a summer flux of about one-half that of the winter. However, the variability scales of this surface circulation are still to be established.

Variability of the Mediterranean sea level has been studied using the data from TOPEX/POSEIDON (T/P) altimeters [Larnicol *et al.*, 1995], but the resolution of the gridding ground tracks of this satellite is too coarse (the distance between two tracks is 2.8° which is equivalent to about 250 km at the 38°N latitude) to observe the circulation in small basins such as the Mediterranean one where the mesoscale signals, of the order of 100 km, can be strong but not very large. The Algerian eddies have been clearly detected by T/P, but the space and time sampling of T/P is not dense enough to follow their evolution. The variability of the Alboran basin has been observed with the data of the European Remote Sensing (ERS-1) satellite altimeter [Vazquez *et al.*, 1996], but the resolution in time is low because of a 35 day satellite repetition cycle and does not allow a good estimation of the temporal variability of

the anticyclonic gyres usually observed in this basin. It was therefore necessary to combine data from ERS-1 (with a repetition cycle of 35 days and a distance between two tracks of ~ 60 km) and T/P (with a repetition cycle of 10 days and a distance between two tracks of ~ 250 km) altimeters to study the variability of the Mediterranean sea level with adequate temporal and spatial resolutions.

The main goal of this study is to estimate the temporal and spatial scales of the mesoscale activity of the Algerian current from a synoptic point of view provided by the satellite altimetry. The frequency of the formation of the anticyclonic anomalies and their direction of propagation are investigated by means of a complex empirical orthogonal function (CEOF) analysis. The data set is described in section 2. Two sea surface temperature (SST) images of anticyclonic eddies are compared with the altimetric data in section 3. Basic statistics are performed on sea level anomaly (SLA) maps in section 4. The CEOF analysis and its results are presented in section 5. Conclusions are drawn in section 6.

2. Data Set

The T/P and ERS-1 altimeters provide complementary space and time sampling of the oceanic circulation, but they do not obtain data with the same accuracy. The accurate determination of the satellite geocentric position, and especially its radial component, is a basic requirement for accurate measurement of the sea level. The radial component of the T/P orbit is determined to within 5 cm, while ERS-1 orbits are only determined to within 15 cm. However, since T/P and ERS-1 are flying simultaneously, the more precise T/P data have been used to correct the ERS-1 orbit error [Le Traon *et al.*, 1995].

After applying the usual corrections including the inverse barometer effect the data of the first overlap period of the two altimetric missions, which begins on October 1992 and ends on December 1993, have been merged by *Ayoub et al.* [1996]. The lack of knowledge of the geoid impedes the use of the absolute

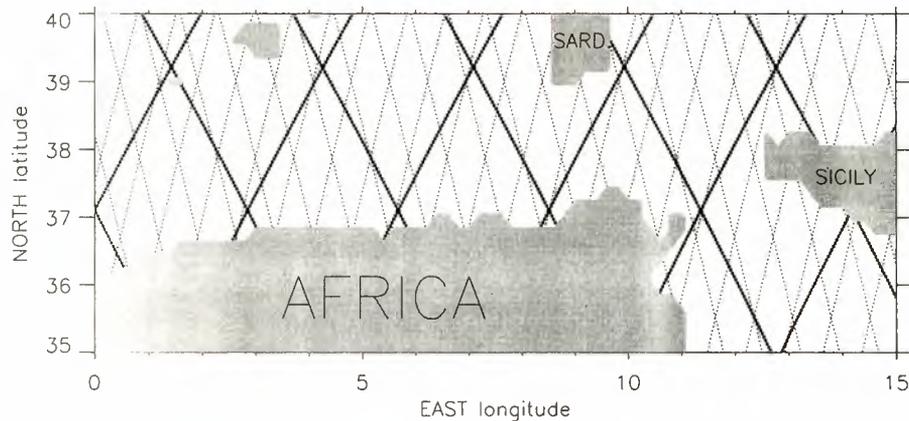


Figure 2. ERS-1 (dotted lines) and TOPEX/POSEIDON (bold lines) tracks in the study area.

heights of the sea level; hence only SLA are considered. They have calculated SLA relative to the annual mean sea level and not relative to the 15 months of the complete data set since the seasonal signal is very important in the Mediterranean Sea [Larnicol *et al.*, 1995]. Then, SLA maps have been created every 10 days with a spatial resolution of 0.2° using an objective analysis (spatial radius of 200 km and temporal radius of 20 days) with the correlation function of *Le Traon and Hernandez* [1992], the a priori errors being 10% and 15% for T/P and ERS-1, respectively.

For the present study the data from 0° to 15°E in longitude and from 35° to 40°N in latitude have been extracted to focus on the MAW flow between the Alboran basin and the eastern Mediterranean Sea. The ERS-1 and T/P tracks in the area are shown in Figure 2. The resulting 44 regular maps in time and space and their associated error maps also provide an excellent database to make comparisons with SST images from the along-track scanning radiometer (ATSR) of ERS-1.

The structures observed from the SLA maps are more or less circular forms of positive or negative anomalies that evolve in time and space with horizontal scales from 50 to 250 km in diameter. To allow a direct comparison with ATSR images by superimposition of the SST on the SLA, vector maps of surface turbulent geostrophic velocities or geostrophic velocity anomalies (GVA) have been calculated, computing the SLA gradient in both directions (east and north) at each grid point. The resulting vector ring structures are cyclonic for a local SLA minimum or anticyclonic for a local SLA maximum, as a consequence of the geostrophic balance. Even if the majority of the resulting GVA structures look like surface vortex circulation, they do not correspond necessarily to real closed cells or eddies. Here such ring structure is called an eddy only if its existence is confirmed by SST; otherwise, the term anticyclonic or cyclonic anomaly is used. However, in the regions where anticyclonic eddies have often been observed with infrared pictures the circular shapes of positive SLA probably correspond to such eddies.

3. Comparisons With ERS-1/ATSR/SST Images

Some Algerian eddy characteristics are pointed out here from the comparison between GVA maps and ATSR/SST instantaneous images, where the clouds are white and the land is black. Plate 1 presents an SST image taken on October 28, 1992, at 2200 UT with the GVA of October 27, 1992, super-

imposed. An anticyclonic eddy of cooler surface water is visible just between two masses of clouds, centered at $\sim 4.5^\circ\text{E}$ and 37.5°N with a diameter of ~ 100 km. At this location the altimetric map does not display a well-defined circular positive anomaly but a rather oval one with a northeast extension overlying the cooler water of the picture. The resulting GVA eddy is centered at 4°E and 37°N . The offset between the SST and the GVA may be due to the SLA mapping interpolation and to the motion of the eddy. The hypothesis that this eddy began as a meander of the current a few days before October 27, then developed, and moved eastward a few days after could explain the resulting oval (and not circular) form of the altimetric signal. The strongest GVA is 32 cm s^{-1} at 3.8°E , 37.6°N for the western edge and 38 cm s^{-1} at 4.8°E , 37.4°N for the eastern edge. Both maxima are located in the coolest vein of the eddy. These velocities are of the same order of magnitude as the ones obtained by *Benzohra and Millot* [1995a] from hydrography, who reported 30 cm s^{-1} eastward in the northern part of a smaller coastal anticyclonic eddy of about 50 km in diameter at 4°E . The evolution of this anticyclonic anomaly can be followed in consecutive maps of SLA. Even if this anomaly really corresponds to an eddy at the end of October as shown by the SST image, this cannot be asserted during the following months; that is why the word “anomaly” is employed here instead of “eddy.” First, it moves eastward along the coast until the end of December 1992. Then, from January to February 1993, it deviates progressively toward the center of the basin at $\sim 6^\circ\text{E}$ and disappears. Unfortunately, no ATSR image free of clouds has been found between November 1992 and February 1993 to corroborate that this anomaly still corresponds to an anticyclonic eddy.

Plate 2 presents an SST image taken on August 7, 1993, at 2200 UT. Although the SST gradients are not very clear, a circular shape of cooler surface waters is detected near the Algerian coast at $\sim 2^\circ\text{--}3^\circ\text{E}$, 37°N . The nearest GVA map in time (August 3, 1993) is overlaid on this SST picture. The velocities reveal an anticyclonic eddy with a diameter of more than 100 km. The center of this eddy is 8 cm higher than the edge. The small shift between the eddy described by the GVA and the warmer center visible on the SST picture, along with the 4 days separating the GVA and the SST data, suggest an eastward propagation of the eddy. The strongest velocity anomaly of 28.5 cm s^{-1} is located at 2°E , 37.4°N in the north-western part of the eddy. Ten days later (August 13, 1993; map

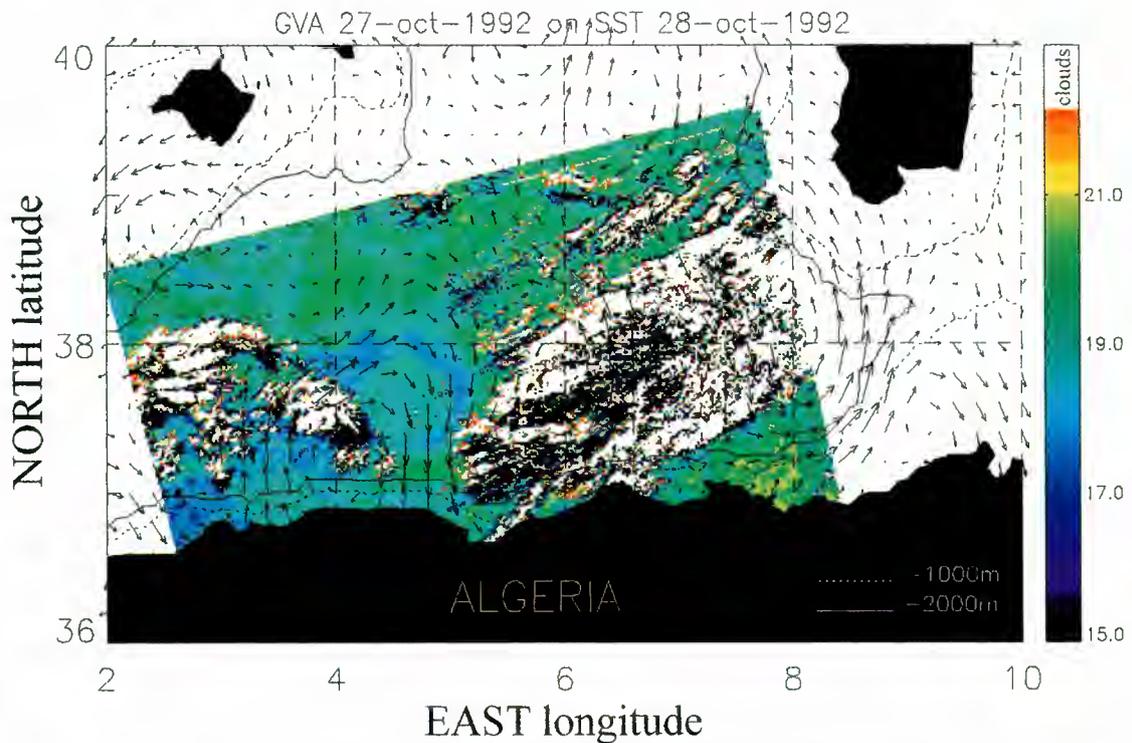


Plate 1. Sea surface temperatures (SST) of October 28, 1992, with geostrophic velocity anomalies (GVA) of October 27, 1992.

not shown), the GVA maximum in this anticyclonic structure is 34 cm s^{-1} at 2.6°E , 37.4°N in the northeastern part of the eddy. In 20 days the eddy center has drifted $\sim 40 \text{ km}$ eastward at 37°N . This gives a propagation velocity of $\sim 2 \text{ km d}^{-1}$ which is

slightly weaker than the 13 km d^{-1} observed by *Taupier-Letage and Millot* [1988] during July 1984 in the same area and the $3\text{--}5 \text{ km d}^{-1}$ estimated by *Millot et al.* [1990] from infrared imagery and current measurements.

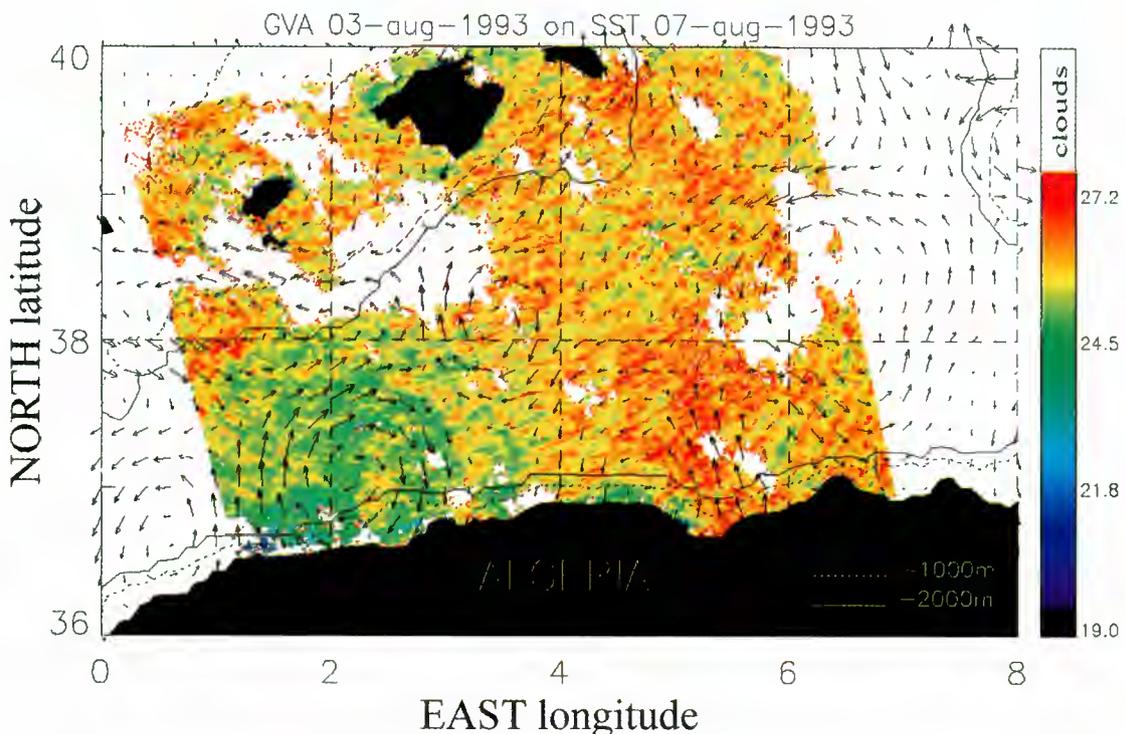


Plate 2. SST of August 7, 1993, with GVA of August 3, 1993.

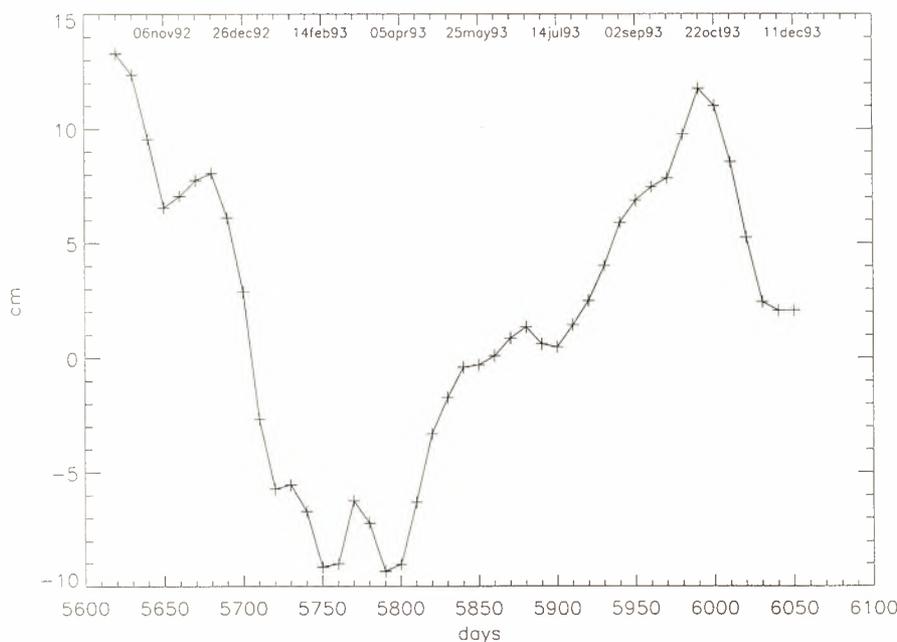


Figure 3. Temporal evolution of the spatial mean sea level anomaly (SLA) from October 7, 1992, to December 11, 1993.

4. Basic Statistics

The temporal evolution of the spatial mean SLA (Figure 3) shows the annual seasonal cycle of variability with a maximum value of +10 cm in the fall (October 1992 and 1993) and a minimum value of -8 cm between the winter and spring (February, March, and April 1993). This curve is very similar to the one obtained by *Lamicol et al.* [1995] for all the western Mediterranean Sea from T/P data only. This means that the study area has a seasonal response quite identical to the mean response of the entire western basin. This temporal variation of the mean sea level is also in good agreement with the description of the seasonal variability of the Atlantic water inflow at the Strait of Gibraltar made by *Ovchinnikov* [1974]. From in situ measurements and from salt conservation he deduced that the sea level continuously increases from February to November and rapidly decreases between December and February.

The map of the seasonal mean SLA for the fall of 1992 (Figure 4a) shows various maxima along the African coast at 1° - 2° E, 4° - 5° E, 7° - 8° E, and 9° - 10° E with anomalies higher than 12 cm including a mean anomaly of ~ 8 cm for this season. In the spring of 1993 (Figure 4c), two similar maxima are observed at 4° E and 6° - 7° E with anomalies higher than 0 cm, while the mean anomaly is ~ -2 cm during this season whereas in the winter and summer of 1993 (Figures 4b and 4d), no such maxima can be seen along the African coast. This indicates that the Algerian current is subject to an important seasonal variability with strong positive level anomalies in the fall and spring that could favor the formation of anticyclonic eddies.

The map of standard deviation for the annual period from October 7, 1992, to October 2, 1993, is shown in Figure 5. A large part of the standard deviation is due to the annual cycle of the sea level. This explains the relatively high background values (5-6 cm) found in the quietest areas. In order to compare, according to *Zlotnicki et al.* [1989], most of the world ocean has variability below 8 cm rms, with uniform variability below 4 cm rms in some equatorial areas. Standard deviations

higher than 10 cm are visible between 4° E and 6° E and between 7.5° E and 9° E near the African coast, revealing two intense variability points in the Algerian current.

There are obviously two distinct reasons for the altimetric sea level variability to be highest in the neighborhood of the Algerian current. The first reason is that the axis of the current does not stay in a fixed position but shifts by ~ 50 km, mainly where the meanders develop. Thus, at a fixed position the sea level rises and falls as the current axis moves on or off that position. This is also the reason why the current cannot be seen all along the Algerian coast since its position does not shift enough and the lack of a geoid model allows one to measure only sea level variability, not absolute departure of the sea level from the geoid. The second reason for the increased variability approaching the Algerian current is the passage of eddies shed by the instabilities of the current, as in the case of the Gulf Stream [*Zlotnicki et al.*, 1989].

5. CEOF Analysis

5.1. Methodology

Oceanic phenomena are usually the combination of various parameters difficult to dissociate interacting at many different timescales and space scales. The decomposition into empirical orthogonal functions (EOFs) is advantageous because it is built on modes calculated from the data themselves, hence proper to the data set and not imposed. If the correlation between the data and the base functions is low, a high number of base functions will be necessary to describe the phenomena. That is why it is better to choose a set of orthogonal functions where the few first ones are sufficient. Hence the empirical eigen functions are useful in analyzing large data sets. Their principal virtues are that they provide the most efficient method of compressing data (they reduce the amount of data by keeping only the most significant phenomena), and they may be regarded as uncorrelated modes of variability of the

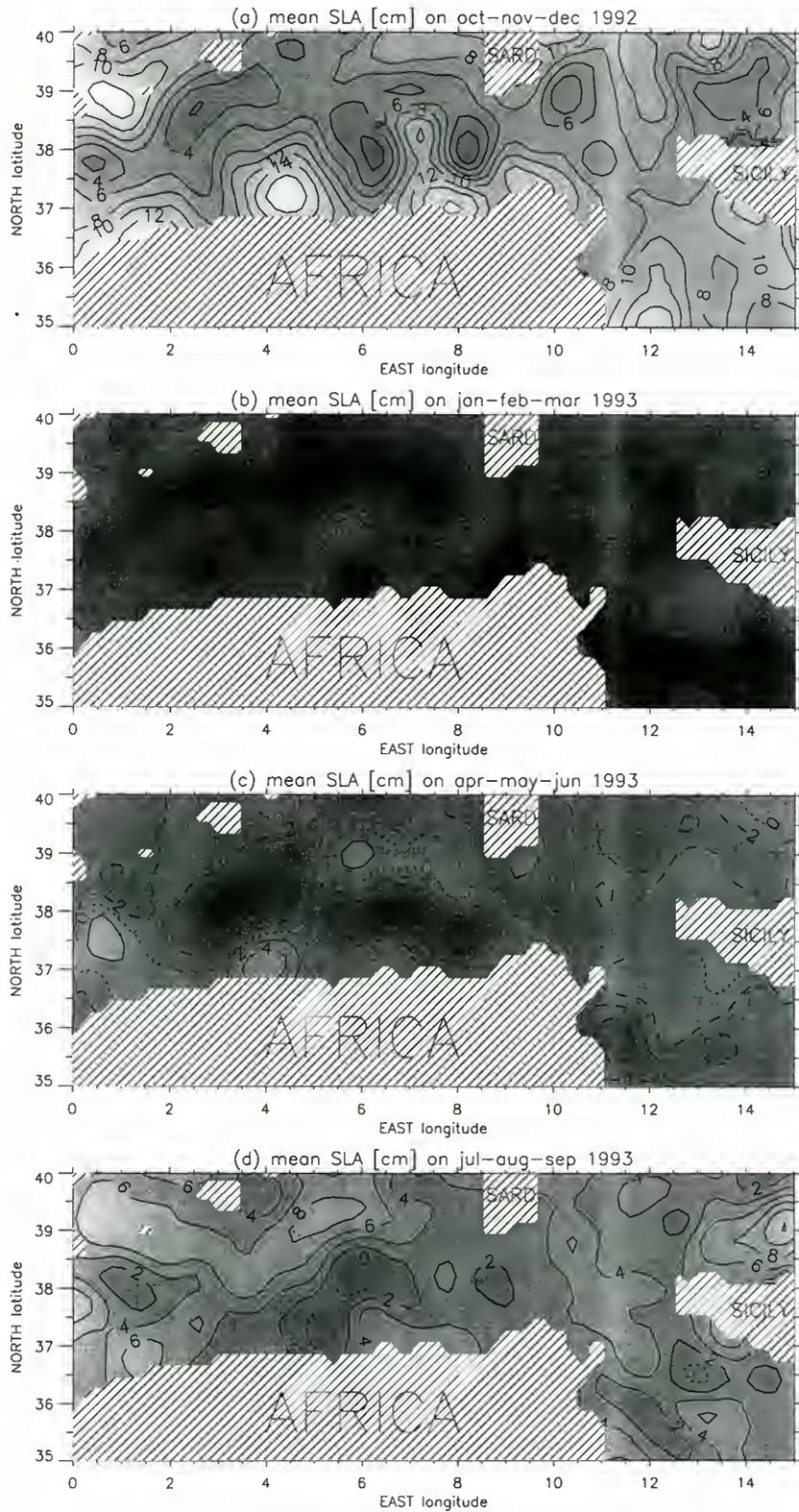


Figure 4. Seasonal mean SLA: (a) fall (October, November, and December 1992), (b) winter (January, February, and March 1993), (c) spring (April, May, and June 1993), (d) summer (July, August, and September 1993). The contour interval is 2 cm. Dashed, dotted, and solid lines correspond to negative, null, and positive values, respectively.

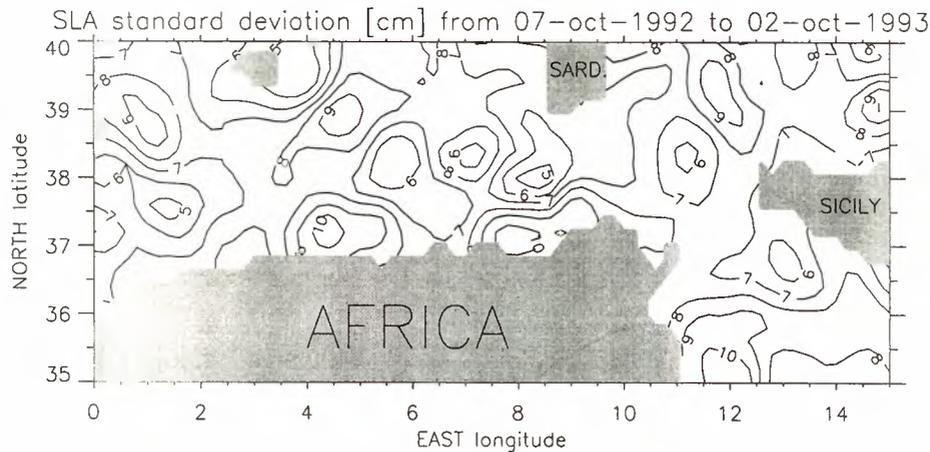


Figure 5. Map of SLA standard deviation for the annual period from October 7, 1992, to October 2, 1993. The contour interval is 1 cm.

field [Davis, 1976]. Unlike real empirical orthogonal functions where the modes represent standing wave patterns, complex EOFs can resolve propagating waves [Horel, 1984]. The decision to apply CEOFs instead of real EOFs is appropriate because of the wave propagation associated with the instabilities of the Algerian current. This method of analysis has given good results in the Gulf Stream [Vazquez, 1993] and on the surface circulation variability in the Alboran basin [Vazquez et al., 1996]. Complex time series $H(x, y, t)$ are formed from the original time series and their Hilbert transforms. The real part of $H(x, y, t)$ is simply the value of the residual sea level at grid point (x, y) and time t , while the imaginary part is the Hilbert transform of the original data field where the amplitude remains unchanged but where the phase is shifted by $\pi/2$. In practice, an estimate of the Hilbert transform is obtained directly from the Fourier coefficients [Barnett, 1983] derived from fast Fourier transform (FFT) routines. This approach is subject to the problems that Fourier analysis normally presents (aliasing, end effects, etc). Hilbert transforms obtained from FFT routines are more strongly influenced by end effects than are the Fourier transforms themselves [Horel, 1984]. Hence each end (roughly 10% of the total time series) is cosinetapered to eliminate this problem.

No assumption is made about phase propagation in either the x or y direction. The covariance statistics are applied in the spatial domain to extract the correlated orthogonal signal. Mathematically, the rest of the procedure is similar to real EOF analysis except that the statistics are performed on the complex time series $H(x, y, t)$. For detailed examples on the use of EOFs, see Stidd [1967] and Davis [1976]. Since the covariance matrix is Hermitian, it possesses real eigenvalues which give the variance associated with each mode. In the CEOF procedure the complex time series $H(x, y, t)$ is decomposed in terms of a set of orthogonal functions $F(x, y)$.

$$H(x, y, t) = \sum_{m=1}^N w_m(t) F_m^*(x, y)$$

where the asterisk is the complex conjugation, N is the total number of grid points in the sea ($N = 1306$) and $w(t)$ is the amplitude function. Thus both the eigenvectors $F_m(x, y)$ of the spatial covariance matrix and the temporal amplitudes

$w_m(t)$ are complex and have an amplitude and a phase component. For each dominant mode m are plotted

The spatial amplitude

$$A_m(x, y) = \sqrt{N F_m(x, y) F_m^*(x, y)}$$

The spatial phase

$$\varphi_m(x, y) = \arctan \left\{ \frac{\text{Im} [F_m(x, y)]}{\text{Re} [F_m(x, y)]} \right\}$$

The temporal amplitude

$$w_m(t) = \sqrt{\frac{\sum_{x,y}^N H(x, y, t) F_m(x, y) \sum_{x,y}^N H^*(x, y, t) F_m^*(x, y)}{N}}$$

The temporal phase

$$\varphi_m(t) = \arctan \left\{ \frac{\text{Im} \left[\sum_{x,y}^N H(x, y, t) F_m(x, y) \right]}{\text{Re} \left[\sum_{x,y}^N H(x, y, t) F_m(x, y) \right]} \right\}$$

For plotting purposes the temporal (spatial) amplitude is divided (multiplied) by the square root of N .

5.2. Results

The spatially correlated signal in the study area was found to be dominated by the first two CEOFs. These two modes explain nearly 85% of the variability, with 80% of the total variance contained in the first one and 5% of the total variance in the second one (Table 1). It is readily seen that higher modes add little variance as the modal structure becomes quickly degenerate and are not significant. The question is whether the second mode, representing only 5% of the variability, is significant or not. A Monte-Carlo simulation analysis [Overland and Preisendorfer, 1982] was applied by running 100 simulations of the CEOF analysis on Gaussian random numbers with a standard deviation of 5. None of the resulting random modes explain $>6\%$ (Figure 6). This means that the

Table 1. Contributions (c(m)) of Each Complex Empirical Orthogonal Function (CEOF) Mode (m) in Percentage of the Total Variance

m	c(m)
1	79.92
2	5.00
3	2.89
4	2.27
5	1.89
6	1.38
7	1.21
8	0.96
9	0.74

second mode of the present analysis could be considered a priori as noise and not significant. However, its temporal and spatial amplitudes do not display any random behavior, as displayed by the modes of the Monte-Carlo simulation. In order to verify that the CEOF analysis is able to separate a propagating wave and a noise background of the same amplitude the analysis was applied on various simulated fields, each one composed of the sum of three terms: a propagating wave of high amplitude (10 cm), a propagating wave of small amplitude (between 2 and 5 cm), and a Gaussian random noise of the same amplitude. The two propagating waves have different periods and wavelengths. The percent variance associated with each mode for each simulation and their respective amplitudes are displayed in Table 2. In every case the three features are spatially and temporally well resolved by the three lowest modes of the CEOF analysis, even in simulations 1 and 2 where the second modes represent $<6\%$ of the total variance. Since the rms error associated with the altimetric data set is ~ 3 cm [Ayoub *et al.*, 1997], the simulation closest to the real data is simulation 3 which shows that the second wave is completely separated from the noise. Even in the last case (simulation 5 in Table 2) where the noise rms is greater than the amplitude of the second wave, the period and the wavelength of the second CEOF are in agreement with the period and the wavelength of the second input wave, and the variability explained by the second CEOF is higher than the one explained by the third CEOF. Thus one can assert that the second mode of the present analysis is the representation of a weak but coherent phenomenon.

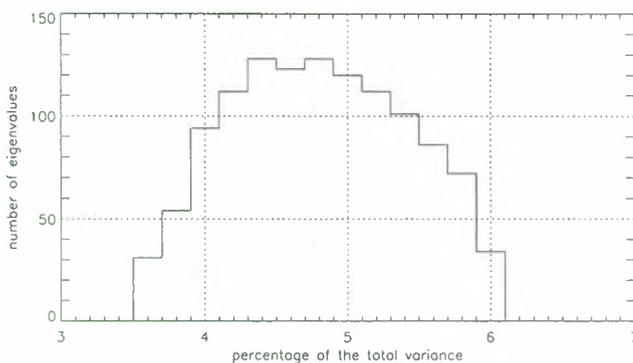


Figure 6. Histogram of all eigenvalues representing $>3\%$ of the total variance for 100 simulation runs of CEOF analysis using a Gaussian random number generator with standard deviation of 5.

Table 2. Amplitudes of the Input Propagating Waves (1 and 2) and Amplitude of the Input Noise (3) for Each Simulation and Percentages of the Total Variance Explained by Each Resulting CEOF

Simulation	Amplitude 1, cm	Percent 1	Amplitude 2, cm	Percent 2	Amplitude 3, cm	Percent 3
1	10	89.05	2	3.87	2	0.82
2	10	84.25	2.5	5.67	2.5	0.88
3	10	79.07	3	7.61	3	0.98
4	10	58.58	5	15.31	5	1.67
5	10	47.21	5	13.07	7	2.58

5.2.1. The first mode (CEOF 1). The temporal phase of the first mode (Figure 7b) indicates a constant frequency of one cycle per year corresponding to the seasonal variability. The strongest signal is observed in the southern part of the channel of Sardinia, south of the Strait of Sicily near the Tunisian coast, at the entrance to the Tyrrhenian basin, and southeast of Mallorca (Figure 8a). The spatial phase (Figure 8b) is homogeneous, indicating a stationary seasonal oscillation in all the study area. The normalized temporal amplitude (Figure 7a) does not show a clear annual cycle, but its multiplication by the cosine of the temporal phase gives back the seasonal curve of Figure 3.

According to Larnicol *et al.* [1995], who analyzed 2 years of T/P data in the whole Mediterranean Sea and characterized changes in the mean level, roughly half of the seasonal signal

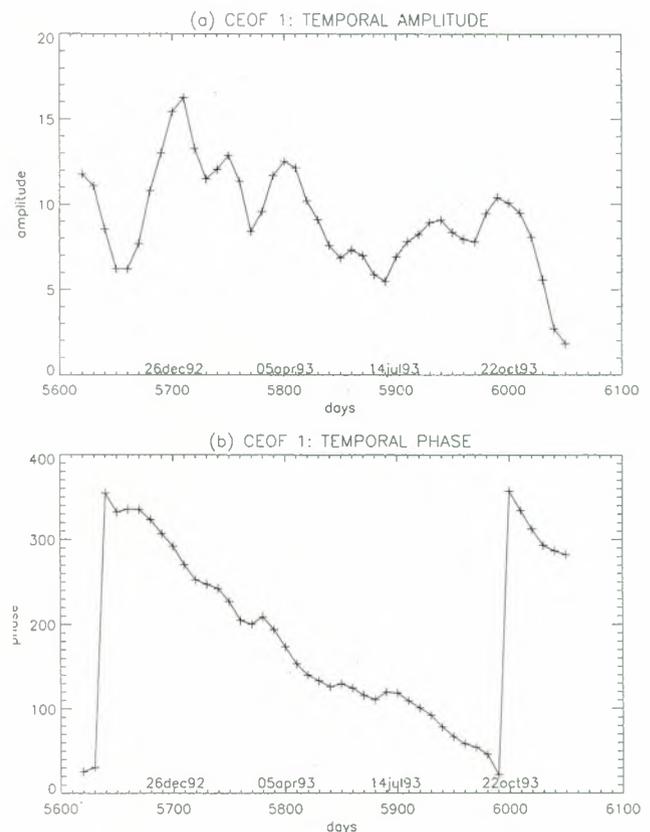


Figure 7. (a) Temporal normalized amplitude and (b) temporal phase in degrees of complex empirical orthogonal function (CEOF) 1.

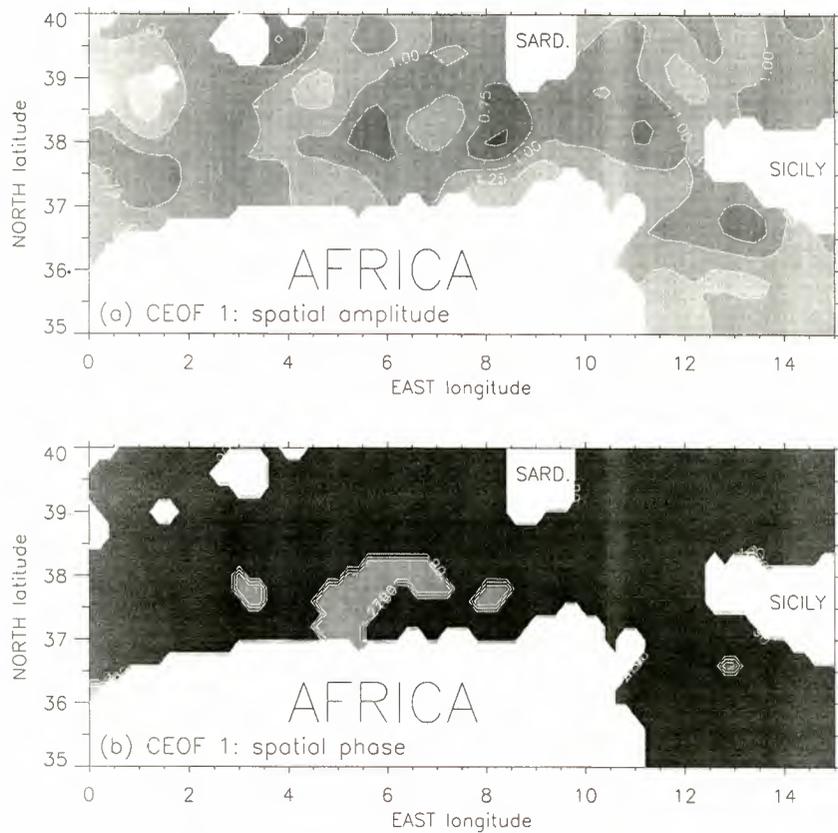


Figure 8. (a) Spatial normalized amplitude with a contour interval of 0.25 and (b) spatial phase with a contour interval of 90° of CEOF 1.

is related to the steric effect of contraction and dilatation of the surface waters due to the heat exchanges at the ocean-atmosphere interface. The other half could be due to a slight seasonal imbalance between inflow and outflow at the Straits of Gibraltar and Sicily and between evaporation and precipitation. These contributions to the seasonal sea level variations are still an important subject of investigation on the Mediterranean Sea [Manzella and La Violette, 1990; Millot et al., 1992] which will not be discussed here.

5.2.2. The second mode (CEOF 2).

The temporal normalized amplitude (Figure 9a) and the temporal phase (Figure 9b) of the second mode show a periodicity of ~ 6 months which appears to be associated with the variability of the Algerian current as the phase isolines are parallel to the mean path along the Algerian coast up to the channel of Sardinia (Figure 10b). The strongest amplitudes of the second mode are located near the African coast between 4°E and 6°E and north of Sicily (Figure 10a). This map is very similar to the map of SLA standard deviation (Figure 5). This indicates that the second CEOF represents the mesoscale activity. The phase isolines all along the coast (Figure 10b) show a deviation offshore at 4°E that corresponds to a maximum in amplitude. From this point the spatial phase decreases north-eastward. In the western part of the channel of Sardinia ($8^\circ\text{--}9^\circ\text{E}$) the amplitude presents a secondary maximum, and the spatial phase decreases from south to north.

The real spatial amplitudes (centimeters) of the second mode bimonthly average (Figure 11) present the same situation in October–November 1992 and April–May 1993 as a consequence of the 6 month periodicity. The highest values are

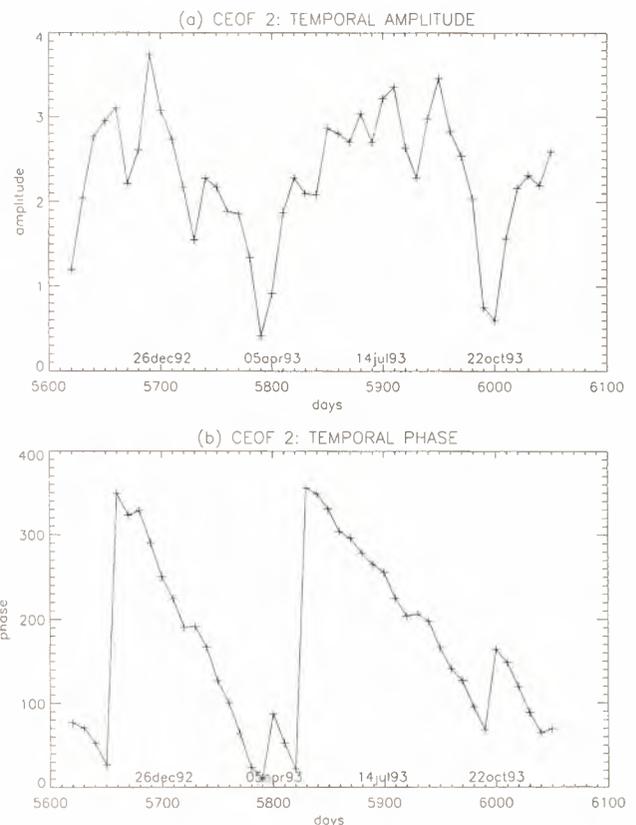


Figure 9. (a) Temporal normalized amplitude and (b) temporal phase in degrees of CEOF 2.

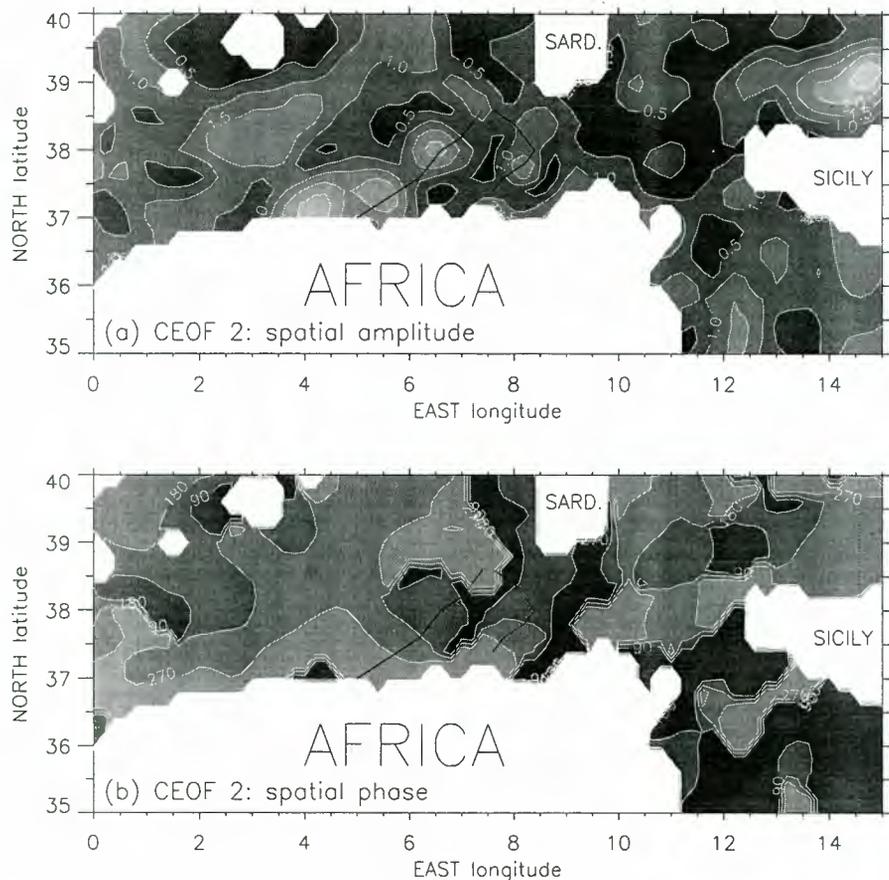


Figure 10. (a) Spatial normalized amplitude with a contour interval of 0.5 and (b) spatial phase with a contour interval of 90° of CEOF 2. The two black lines are the sections used for the time-latitude diagrams plotted in Figure 12.

located at 4°E near the coast and at $8^\circ\text{--}9^\circ\text{E}$ in the southern part of the channel. In December–January 1993 the former has progressed eastward, while the latter has moved northward across the channel. In February–March 1993 both positive anomalies have weakened and moved to more northerly positions. This cycle repeats from April to May 1993 with an evolution slightly longer than 6 months as the amplitudes are still significant in August–September 1993 but vanish in October–November 1993 (this last map is not shown). Looking at the amplitude maps of this mode every 10 days like a movie, both areas of maximum variability evoke wave propagation eastward and northward, since positive and negative anomalies alternate, with wavelengths of 140 ± 30 km and phase velocities of 1 or 2 km d^{-1} . The propagation from the two high-variability points is also visible on the time-latitude diagrams shown in Figures 12a and 12b. Figures 12a and 12b display the time evolution of the second-mode amplitude along the two sections drawn on Figure 10. In Figure 12a the propagation is northeastward between the Algerian coast and Sardinia. Pulsation of the propagating anomaly amplitude is observed with a period of ~ 60 days. In Figure 12b the propagation follows approximately the deep isobaths (~ 2000 m) with the same period of pulse. Both wave propagations end in the same area, southwest of Sardinia.

After comparison with the map of satellite tracks (Figure 2) it can be stated that the two areas of maximum variability roughly coincide with the two T/P track crossover points at

37°N and could be due to an artifact of the mapping routine. However, the other T/P crossover points do not present any variability maxima. One can also remark that the northward propagation near the channel of Sardinia is nearly parallel to ERS-1 descending tracks for one half and parallel to ERS-1 ascending tracks for the other half and could be due to temporal aliasing between measurements of adjacent or crossing tracks. However, the northeastward propagation from the Algerian coast (5°E) is not parallel to any tracks. In more general terms, no coherent relationship has been found between the CEOF results and the sampling of the original data or the mapping analysis used to create the SLA data set. Thus the CEOF patterns do not seem to be due to any artifact of the procedure.

6. Discussion and Conclusion

A CEOF analysis performed on 15 months of combined altimetric SLA from T/P and ERS-1 shows that the variability in the region of the Algerian current is dominated by two modes which are the annual cycle and an almost semiannual cycle. Because of their respective periodicity, the second CEOF could be a harmonic of the first CEOF, but the similarity between the map of SLA standard deviation (Figure 5) and the amplitude map of the second mode (Figure 10a) supports the fact that the two modes are associated with two very

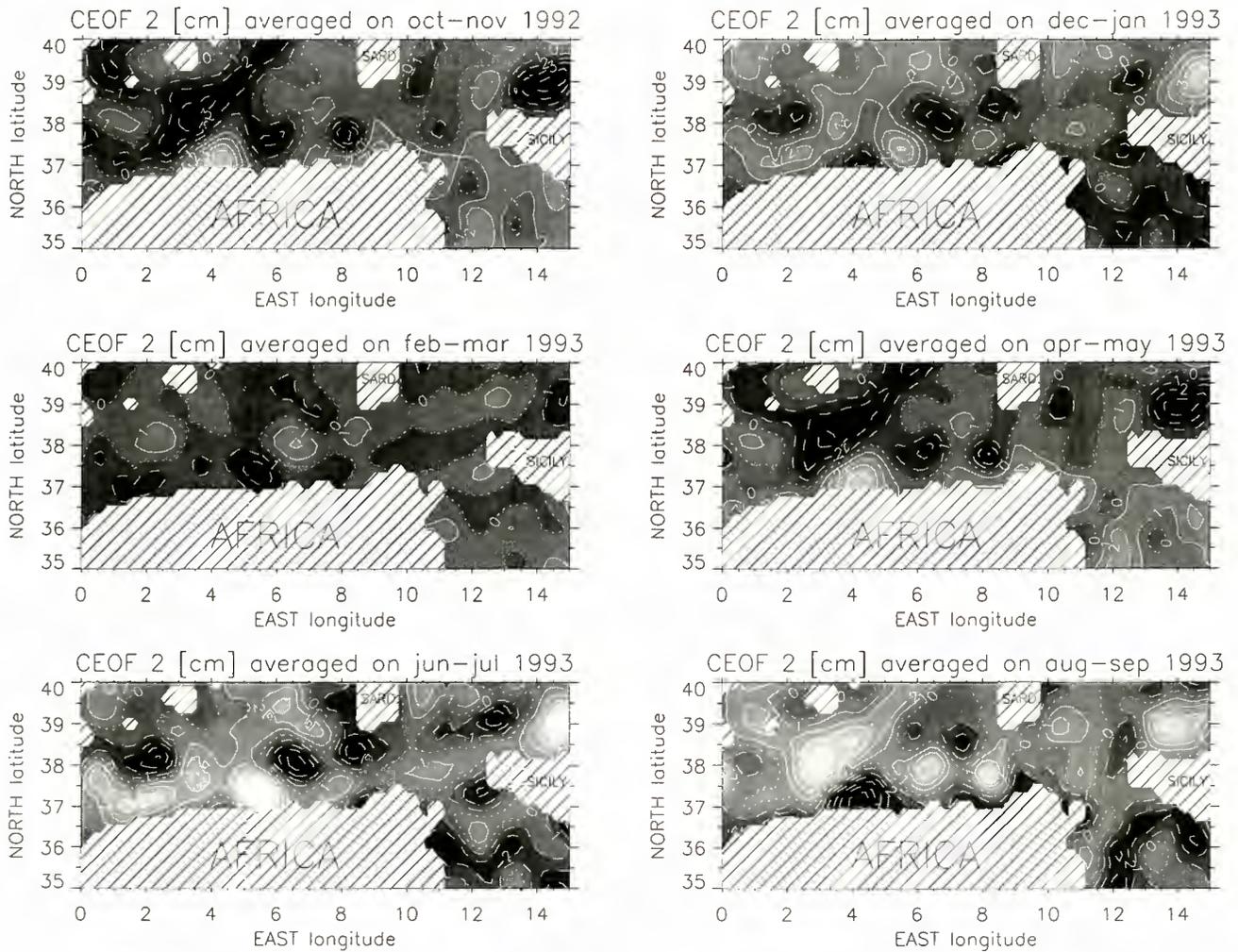


Figure 11. Bimonthly averaged amplitudes of CEOF 2. The contour interval is 1 cm. Dashed, dotted, and solid lines correspond to negative, null, and positive values, respectively.

different physical phenomena which are the seasonal steric effect and the mesoscale variability.

The path of the Algerian current in the southern channel of Sardinia could be strongly affected by the seasonal variability as the first CEOF amplitude is higher in this area with the maximum seasonal mean value obtained in the fall of 1992 and fall of 1993. This can be related to the seasonal variability of MAW flux observed in the Strait of Sicily by the fact that these positive SLA (visible in Figure 4a) could be linked to a stronger MAW circulation around the northeast cape of Tunisia and a direct path toward the Strait of Sicily. This is in agreement with the recent observations of *Astraldi et al.* [1996], who report a wide MAW flow, as broad as the strait, in the winter and a narrow vein close to Tunisia during the summer and fall. However, the in situ data in the Strait of Sicily are sparse, especially along the Tunisian coasts, and according to *Manzella et al.* [1990], the maximum MAW flow occurs during the winter and not during the fall. Even if the seasonal variation of the mean sea level is similar to the one described by *Ovchinnikov* [1974], the 15 month time series of altimetric data used in this study is too short to get any solid conclusion on the annual seasonal variability locally represented by the first CEOF pattern.

The second CEOF reveals two points of higher variability along the Algerian current path. The decreasing phase from

south to north across the western opening of the channel of Sardinia could correspond to the propagation of eddies described by *Fuda et al.* [1997] or by *Bouzinac et al.* [1997], who report an anticyclonic eddy evolving between June and August 1994 in the same area. This is in agreement with the hypothesis that some anticyclonic eddies could detach from the African coast at the entrance of the channel, following the deep bathymetry, because of their vertical extension, as events of ~6 months between the formation and the complete disappearance of the eddy. In the same way such detachments could occur between 4° and 6°E if one assumes that positive anomalies correspond to real anticyclonic eddies. The amplitude and phase isolines imply a northeast propagation of the eddies. This is supported by the anticyclonic eddy observed at the end of October 1992 (Plate 1), which is part of the second CEOF features (Figure 11a).

The pulsation of ~2 months observed in the amplitude of the second CEOF may be related to the growth rate of the baroclinic instabilities in the Algerian current which have been modeled by *Beckers and Nihoul* [1992]. Indeed, they obtained a development time of a few weeks for the perturbations to affect the whole current. However, the fact that this pulsation is visible in all the basin cannot be explained.

The 6 month periodicity of the mesoscale activity does not

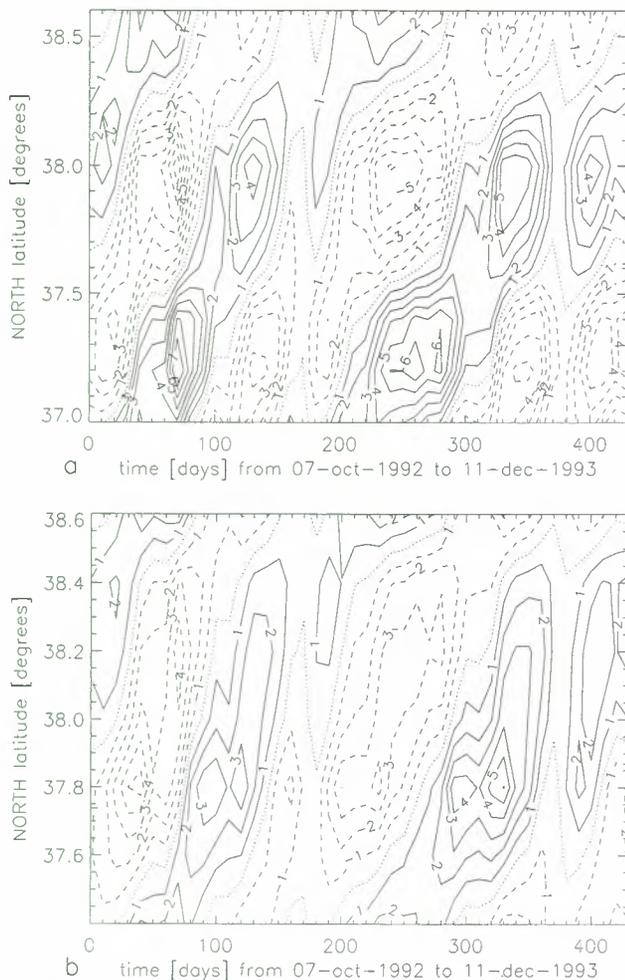


Figure 12. Time-latitude diagrams of the CEOF 2 amplitude (centimeters) (a) between 5°E, 37°N and 7.4°E, 38.6°N and (b) between 7.6°E, 37.4°N and 7.6°E, 38.6°N. These sections are shown in Figure 10.

appear in the previous studies of the Algerian current from in situ and infrared observations. Anticyclonic eddies have been detected in the spring, summer, and fall [Taupier-Letage and Millot, 1988], and the largest ones have lasted more than 6 months [Millot et al., 1990]. This periodicity could also be a specific feature of the year 1993. It is therefore impossible to confirm this periodicity revealed by the second mode of the CEOF analysis whereas the propagation path and the propagation speed of 1 or 2 km d⁻¹ fit quite well with in situ observations near the channel of Sardinia. In the future it would be necessary to monitor these eddies on a longer time period with the same almost continuous synoptic maps to ascertain whether the mesoscale variability is subject to a semi-annual cycle. This will be possible using the combined altimetric data of T/P and ERS-2.

Acknowledgments. This research has been undertaken in the framework of the Mediterranean Targeted Project (MTP) phase II-MATER (contribution MTP-II-MATER/004). We acknowledge support from the European Commission's Marine Science and Technology Programme (MAST III) under contract MAS3-CT96-0051. C. Bouzinac benefited from financial support from Ministerio de Asuntos Exteriores in the frame of the Franco-Spanish exchange. We are grateful to N. Ayoub (CLS, Toulouse) who provided the altimetric SLA maps and to the Centre de Supercomputació de Catalunya

(CESCA) whose CRAY YMP was used for CEOF computation. We warmly thank C. Millot, F. Martel, and E. García for their helpful comments. This study is a contribution to ESA ERS Announcement of Opportunity project ALGERS (E102/0). The Spanish National Program on Marine Sciences and Technology (CICYT) gave economic support to the research through project MAR95-1861.

References

- Astraldi, M., G. P. Gasparini, S. Sparnocchia, M. Moretti and E. Sansone, The characteristics of the water masses and the water transports in the Sicily channel at long time scales, edited by F. Briand, *Bull. Inst. Océanogr.*, 17, 95–115, 1996.
- Ayoub, N., P. Y. Le Traon, and P. De Mey, A description of the Mediterranean surface variable circulation from combined ERS-1 and TOPEX/POSEIDON altimetric data, *J. Mar. Syst.*, in press, 1997.
- Barnett, T. P., Interaction of the monsoon and Pacific trade wind system at interannual time scales, The equatorial zone, *Mon. Weather Rev.*, 111, 756–773, 1983.
- Beckers, J. M., and C. J. Nihoul, Model of the Algerian current's instability, *J. Mar. Syst.*, 3, 441–451, 1992.
- Benzohra, M., and C. Millot, Characteristics and circulation of the surface and intermediate water masses off Algeria, *Deep Sea Res.*, 42, 1803–1830, 1995a.
- Benzohra, M., and C. Millot, Hydrodynamics of an open Algerian eddy, *Deep Sea Res.*, 42, 1831–1847, 1995b.
- Bouzinac, C., C. Millot, and J. Font, Hydrology and currents observed in the channel of Sardinia during the PRIMO-1 experiment from November 1993 to October 1994, *J. Mar. Syst.*, in press, 1997.
- Davis, R. E., Predictability of sea surface temperature and sea level pressure anomalies over the north Pacific ocean, *J. Phys. Oceanogr.*, 6, 249–266, 1976.
- Fuda, J. L., C. Millot, I. Taupier Letage, and X. Bocognano, XBT sections across the western Mediterranean sea, *J. Mar. Syst.*, in press, 1997.
- Horel, J. D., Complex principal component analysis: Theory and examples, *J. Clim. Appl. Meteorol.*, 23, 1660–1673, 1984.
- Larnicol, G., P. Y. Le Traon, and P. De Mey, Mean sea level and variable surface circulation of the Mediterranean Sea from 2 years of TP altimetry, *J. Geophys. Res.*, 100, 163–177, 1995.
- Le Traon, P. Y., and F. Hernandez, Mapping of the oceanic mesoscale circulation; validation of satellite altimetry using surface drifters, *J. Atmos. Oceanic Technol.*, 9, 687–698, 1992.
- Le Traon, P. Y., P. Gaspar, F. Bouyssel, and H. Makhmara, Using TOPEX/POSEIDON data to enhance ERS-1 data, *J. Atmos. Oceanic Technol.*, 12, 161–170, 1995.
- Manzella, G.M.R., and P. E. La Violette, The seasonal variation of water mass content in the western Mediterranean and its relationship with the inflows through the straits of Gibraltar and Sicily, *J. Geophys. Res.*, 95, 1623–1626, 1990.
- Manzella, G.M.R., T. S. Hopkins, P. J. Minnett, and E. Nacini, Atlantic water in the strait of Sicily, *J. Geophys. Res.*, 95, 1569–1575, 1990.
- Millot, C., Some features of the Algerian current, *J. Geophys. Res.*, 90, 7169–7176, 1985.
- Millot, C., Circulation in the western Mediterranean Sea, *Oceanol. Acta*, 10, 143–149, 1987.
- Millot, C., Models and data: A synergetic approach in the western Mediterranean sea, in *Ocean Processes in Climate Dynamics: Global and Mediterranean Examples*, pp. 407–425, Kluwer Acad., Norwell, Mass., 1994.
- Millot, C., I. Taupier Letage, and M. Benzohra, The Algerian eddies, *Earth Sci. Rev.*, 27, 203–219, 1990.
- Millot, C., G.M.R. Manzella, and P. E. La Violette, Correction to “The seasonal variation of water mass content in the western Mediterranean and its relationship with the inflows through the straits of Gibraltar and Sicily” by G. M. R. Manzella and P. E. La Violette, *J. Geophys. Res.*, 97, 17,961–17,962, 1992.
- Ovchinnikov, I. M., On the water balance of the Mediterranean Sea, *Oceanology*, 14, 198–202, 1974.
- Overland, J. E., and R. W. Preisendorfer, A significance test for principal components applied to a cyclone climatology, *Mon. Weather Rev.*, 110, 1–4, 1982.
- Perkins, H., and P. Pistek, Circulation in the Algerian basin during June 1986, *J. Geophys. Res.*, 95, 1577–1585, 1990.
- Stidd, C. K., The use of eigenvectors for climatic estimates, *J. Appl. Meteorol.*, 6, 255–264, 1967.

- Taupier-Letage, I., and C. Millot, Surface circulation in the Algerian basin during 1984, *Oceanol. Acta*, 9, 79–85, 1988.
- Vazquez, J., Observations on the long period variability of the Gulf Stream downstream of Cape Hatteras, *J. Geophys. Res.*, 98, 20,133–20,147, 1993.
- Vazquez, J., J. Font, and J. J. Martinez Benjamin, Observations on the circulation in the Alboran Sea using ERS-1 altimetry and sea surface temperature data, *J. Phys. Oceanogr.*, 26, 1426–1439, 1996.
- Zlotnicki, V., L. L. Fu, and W. Patzert, Seasonal variability in global sea level observed with GeoSat altimetry, *J. Geophys. Res.*, 94, 17,959–17,969, 1989.
- C. Bouzinac and J. Font, Instituto de Ciencias del Mar, Consejo Superior de Investigaciones Científicas, Paseo Juan de Borbon, 08039 Barcelona, Spain. (e-mail: cath@icm.csic.es; jfont@icm.csic.es)
- J. Vazquez, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109-8099. (e-mail: jv@pacific.jpl.nasa.gov)

(Received December 19, 1996; revised October 3, 1997; accepted October 10, 1997.)

Correlations between altimetric sea surface height and radiometric sea surface temperature in the South Atlantic

Matthew S. Jones¹, Myles Allen^{2, 3}, Trevor Guymer¹ and Mark Saunders⁴

Abstract. In the last decade, satellite altimetric measurements of sea surface height (SSH) and infrared radiometric measurements of sea surface temperature (SST) have provided a wealth of information about ocean circulation and atmosphere-ocean interactions. SSH is a depth-integrated quantity dependent upon the temperature and salinity structure of the water column and on the depth independent barotropic contribution. SST from infrared radiometers is a surface parameter representing the temperature of the top few microns of the ocean surface. Hence any relationship between SST and SSH provides dynamical information about the coupling between the ocean surface and subsurface. It also offers a promise of new techniques such as interpolating SSH data using SST and of improved calculations of eddy kinetic energy. We use SST data from the along-track scanning radiometer on ERS-1 and SSH data from the TOPEX/POSEIDON instrument to examine the relationship between SST and SSH anomalies within the South Atlantic region for 1993 and 1994. We find that positive (≈ 0.2 - 0.6) spatial cross correlations between SST and SSH anomalies at zero lag are present throughout the region at large scales (wavelengths > 1000 km). Small-scale correlations, however, are high (≈ 0.7) only in areas associated with fronts and mesoscale variability. These small-scale correlations are seasonal, being strongest in winter and weakest in summer. We discuss the application of these correlations to various techniques requiring the synergistic use of SSH and SST data.

1. Introduction

Satellite remote sensing has established itself as an essential means of acquiring global information about our oceans. Satellites now measure sea surface temperature (SST), sea surface height (SSH), wind velocity, significant wave height, wave period, and ocean color. However, many of these parameters are traditionally still studied in isolation of the others. The future challenge must be to start combining multi-instrument satellite measurements to obtain extra insight on ocean dynamics and circulation. An initial attempt at combining two satellite-measured parameters, SST and SSH, is presented here.

SST and SSH are arguably the remote-sensed parameters most fundamental to ocean circulation and climate. SST is used to identify ocean fronts and eddies [e.g., Olson *et al.*, 1988] and is one parameter controlling the transfer of heat and moisture between the ocean and atmosphere. Furthermore, SST is a dominant factor controlling the genesis of hurricanes [Saunders and Harris, 1997] and is an indicator of climatic changes such as the El Niño-Southern Oscillation and the North Atlantic Oscillation. SSH is dependent on the full depth density structure of the ocean, and knowledge of SSH relative

to the geoid provides the surface geostrophic circulation. Moreover, SSH is of use in observing sea level change and its associated consequences [e.g., Minster *et al.*, 1995].

The accurate measurement of SST and SSH from space has considerably improved over the last few years. The TOPEX/POSEIDON (T/P) altimeter can measure SSH relative to a reference ellipsoid to an absolute accuracy of 5 cm [Fu *et al.*, 1994]. This is an order of magnitude improvement over previous altimeters and is due mainly to T/P's improved orbit determination. The accuracy of infrared measurements of SST has also advanced recently because of the ERS-1 mounted along-track scanning radiometer (ATSR). In most areas of the ocean the overall relative accuracy of a point measurement of SST from ATSR is around 0.3 K [Mutlow *et al.*, 1994; Forrester and Challenor, 1995; Harris and Saunders, 1996]. This is a factor of 2 improvement on previous instruments.

In this study we take advantage of data from the pioneering T/P altimeter and ATSR instruments to investigate the relationship between SST and SSH within the South Atlantic Ocean. This area contains a variety of different dynamical regimes, such as the highly energetic Brazil-Falkland [Gordon, 1989] and Agulhas [Lutjeharms and van Ballegooyen, 1988] regions, the quiescent central South Atlantic, and the ocean-linking Antarctic Circumpolar Current [Johnson, 1989]. For an excellent overview of the upper level geostrophic circulation in the South Atlantic the reader should consult Peterson and Stramma [1991]. The structure of this present paper is as follows. A description of the previous studies relating SST and SSH is given in section 2, and a brief introduction of the theory behind possible relationships between SST and SSH is given in section 3. Observational results of the relationship between T/P SSH and ATSR SST data are presented in section 4, followed by a discussion of the applications of these results in section 5. Conclusions are then drawn in section 6.

¹James Rennell Division, Southampton Oceanography Centre, Southampton, England, United Kingdom.

²Space Science Department, Rutherford Appleton Laboratory, Chilton, Didcot, England, United Kingdom.

³Also at Department of Physics, University of Oxford, Oxford, England, United Kingdom.

⁴Department of Space and Climate Physics, University College, London, England, United Kingdom.

2. Previous Studies

Previous attempts to relate SST and SSH are scarce. *Vastano and Reid* [1985] apply an interactive feature-tracking technique [*Vastano and Borders*, 1984] to two AVHRR brightness temperature images in order to obtain surface current velocities. From these velocity vectors a stream function is obtained, which they relate to SSH by assuming geostrophy and integrating the geostrophic relationship. A problem with this ambitious approach is that feature-tracking techniques do not always yield an accurate geostrophic velocity field; Ekman and inertial currents may advect the feature being tracked; and if an eddy is perfectly circular, it is impossible to obtain the rotational component of the eddy, even though this may be larger than the component owing to advection. Errors in the estimation of the geostrophic velocity will lead to errors in the resultant SSH field.

Several studies comment on similar patterns in both infrared imagery and satellite altimeter data. *Scott and McDowall* [1990] conduct a thermistor chain survey along a section of Geosat track across the Iceland-Faeroes Frontal Zone. They use several AVHRR brightness temperature images to provide the context for the thermistor and the Geosat data. Several cold cross-frontal jets are observed in both the AVHRR SST and the Geosat measurements of SSH anomalies (Δ SSH), with cold SST corresponding to a lowering in Δ SSH. *Vazquez et al.* [1990] analyze 2 years of Geosat data in the Gulf Stream region, together with the position of the north wall of the Gulf Stream as measured by AVHRR. They find that the Gulf Stream front is visible in several maps of Δ SSH (because of the front meandering from its mean position and therefore appearing in maps of Δ SSH). They find a root-mean-square difference between the position of the Gulf Stream from Δ SSH and from SST of 92 km between 75° and 50°W.

Halliwel et al. [1991] use Geosat and AVHRR data to show the presence of westward propagating features in the subtropical convergence zone region of the Sargasso Sea. These features have wavelengths of 800 km and periods of 200 days and propagate westward with SST anomalies (Δ SST) leading Δ SSH by ~1 month and 100 km. *Van Woert and Price* [1993], in a study of Rossby waves near the Hawaiian Islands, observe cusp-shaped features in AVHRR SST fields which match up well with gridded Δ SSH data from Geosat. By estimating the propagation of these features in the Δ SSH data they deduce that these cusp-shaped patterns are the surface manifestations of baroclinic Rossby waves. *Van Woert and Price* [1993] also compare Δ SSH data along four Geosat tracks with AVHRR SST data interpolated onto the Geosat tracks. They find that in April 1988 for the region 16°-26°N, water cooler than 24°C corresponds to negative Δ SSH and warmer water corresponds to positive Δ SSH. The exception to this is an area southwest of the Hawaiian Islands, where there is no relationship between SST and Δ SSH. They attribute this to the weaker winds in this area allowing surface heating to obscure any underlying horizontal structure, whereas elsewhere in the region the winds are stronger and this decoupling mechanism does not occur. *Cipollini et al.* [1997] observe coincident westward propagating features in zonal SST gradients from ATSR and zonal SSH gradients from T/P in the Azores current region (34°N). A spectral decomposition of longitude-time plots reveals three similar peaks in both the SST and the SSH spectra with wavelengths of 520, 380, and 390 km and periods of 200, 240, and 490 days, respectively. There is tentative

evidence to suggest that these three peaks may be related to the first three baroclinic modes of Rossby wave propagation.

Knudsen et al. [1996] take a different approach to relating SST and SSH by using T/P altimeter data and ATSR SST data. Rather than studying spatial variations in SST and SSH, they average both SSH and SST data into 2° (latitude) by 5° (longitude) bins and compare time series of SST and SSH at each location. By averaging the data into such large bins they effectively reduce any mesoscale signal and emphasize the large-scale seasonal signal in both Δ SST and Δ SSH owing to the seasonal change in heat and freshwater fluxes [*Gill and Niiler*, 1973]. In most regions they find similarities between the amplitudes and phases of the annual cycles of SST and SSH, with largest amplitudes in the northern hemisphere being off the eastern coasts of continents and in the southern hemisphere being in a latitude band 25°-50°S.

In summary, direct observations of any relationship between SST and SSH are limited to a handful of studies. Before using model or in situ data to attempt to establish a relationship between SST and SSH the approach taken here is to see whether or not the available remotely sensed data suggests a relationship. This will then guide future research in this area. The relationship between SST and SSH is quantified in the South Atlantic area for 1993 and 1994 in section 4. First, however, the reason why a relationship between SST and SSH might be expected is described.

3. Theory

Several mechanisms exist which could result in a relationship between SST and SSH. One mechanism is that the surface temperature structure is correlated with the temperature structure at depth. This would result in a surface temperature structure that is correlated with the SSH structure (ignoring salinity and barotropic effects) independent of the velocity structure at depth. Another mechanism is that in the presence of a meridional surface temperature gradient (cold north, warm south for example) a zonal slope in SSH would cause a meridional geostrophic surface velocity that would either advect warm water north, or cold water south (in the case of this example). In this instance the SST would be correlated with the gradient in SSH rather than with the SSH itself. A third (and more tentative) mechanism that may result in a relationship between SST and SSH is ocean-atmosphere coupling. For example, a SST anomaly may cause a horizontal temperature gradient in the lower atmosphere. This will cause a change in the surface wind field which, in turn, could force anomalies in the surface height field.

The observational study described in section 4 shows that the correlation between Δ SST and Δ SSH is much stronger than the correlation between Δ SST and the gradient of the Δ SSH. This suggests that advection of a temperature gradient is a weaker mechanism than the Δ SST structure correlating with the temperature structure at depth. The manner in which a temperature change at depth is related to the Δ SSH response is developed in the rest of this section.

Let η_s represent variations which are caused by changes in the density structure of the water column (steric variations)

$$\dot{\eta}_s = -\frac{1}{\rho} \int_{-H}^0 \dot{\rho} dz \quad (1)$$

Here $z = -H$ is the depth of the water, $z = 0$ is the mean sea surface, $\bar{\rho}$ is the time-averaged mean density of the water column, and ρ' is the density variation from the time-averaged mean density at a particular depth. The way in which ρ' varies depends upon the mechanism causing the change. Two mechanisms are important: (1) changes in the upper several hundred meters caused by variations in ocean-atmosphere heat and water fluxes and (2) changes caused by advection. *Gill and Niiler* [1973] show that on large spatial scales (~1000 km) the former is the dominant mechanism causing changes in SSH; seasonal changes in these fluxes can cause SSH variations of several centimeters. On smaller spatial scales (~100 km) the effects of advection are important. A mesoscale eddy advected by the mean flow can cause changes in the density of the water column that result in changes in SSH of several tens of centimeters [e.g., *Gordon and Haxby*, 1990; *Smythe-Wright et al.*, 1996], an order of magnitude larger than SSH variations caused by ocean-atmosphere heat fluxes. The meandering of an ocean front can also cause such changes. Indeed, *Vazquez et al.* [1990] study the SSH variability in the Gulf Stream area and conclude that the majority of the seasonal variability can be explained by a meandering front. The reason that changes in SSH caused by advection are larger than those caused by ocean-atmosphere heat fluxes is that advection changes can occur over a large depth range (~1000 m), depending on the structure of the feature being advected. SSH changes induced by surface fluxes, however, are limited to the depth of the mixed layer of the ocean (~100 m).

Irrespective of the mechanism causing the density change, variations in SSH will only be related to changes in SST if the variations in density are related to changes in SST. Changes in density at a particular pressure ($\delta p = 0$) are related to variations in temperature and salinity by

$$\delta\rho = \frac{\partial\rho}{\partial T}\delta T + \frac{\partial\rho}{\partial S}\delta S \tag{2}$$

Hence a change in density will only correlate with temperature variations if the change in density caused by salinity is either small or if this change correlates with the change in density caused by temperature. Assuming that temperature and salinity changes from mean values are sufficiently small ($\delta T = T'$, $\delta S = S'$, hence $\delta\rho = \rho'$), then (2) can be substituted into (1), which gives

$$\eta'_s = -\frac{1}{\rho} \int_{-H}^0 \left(\frac{\partial\rho}{\partial T} T' + \frac{\partial\rho}{\partial S} S' \right) dz \tag{3}$$

or

$$\eta'_s = \int_{-H}^0 (\alpha T' - \beta S') dz \tag{4}$$

where

$$\alpha = -\frac{1}{\rho} \frac{\partial\rho}{\partial T}, \quad \beta = \frac{1}{\rho} \frac{\partial\rho}{\partial S} \tag{5}$$

are the thermal expansion coefficient and the expansion coefficient for salinity, respectively.

Thus, for a relationship to exist between SST and SSH, three criteria must be satisfied: (1) the variation in Δ SSH is steric, rather than due to either barotropic motions or the effects of

tides or atmospheric pressure (the latter two effects have been eliminated from our analysis, as described in the next section); (2) a relationship exists between $\alpha T'$ and $\beta S'$ (or $\beta S'$ is small); and (3) surface temperature variations are related to temperature variations at depth.

4. Observation

4.1. Instrument and Data Description

The U.S.-French T/P altimeter was launched in August 1992 into an orbit with an inclination of 66° and a repeat period of 9.9156 days [*Fu et al.*, 1994]. The global accuracy of the determination of the SSH relative to some fixed reference is ≈ 5 cm in conditions of 2 m significant waveheight [*Fu et al.*, 1994]. This order-of-magnitude improvement on previous altimetric missions is achieved mainly by a more accurate orbit determination but also by direct measurement of both the ionospheric and wet tropospheric corrections.

The altimetry data used for this study are T/P SSH data spanning 1993 and 1994 (T/P cycles 11-84 inclusive). T/P data are provided as merged geophysical data records on CD-ROM by Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO). The standard set of geophysical corrections are applied to the data. In particular, the Center for Space Research (CSR) 3.0 tidal model [*Eanes and Bettadpur*, 1995] is used to remove the effect of tides, an inverse barometer correction is applied, and the four parameter sea state bias correction by *Gaspar et al.* [1994] is used. Data are rejected if the rejection criteria specified by AVISO [1996] are exceeded. The data are collocated to the subsatellite points defined by T/P cycle 18 (this being close to the nominal ground track) using a perpendicular bisector approach [*Cheney et al.*, 1983]. The Ohio State University (OSU91) mean sea surface model is used to reduce errors caused by the combined effect of across-track variations in the altimeter sampling and across-track mean sea surface gradients [*Brenner et al.*, 1990].

The ATSR instrument was designed and built by a consortium led by the U.K. Rutherford Appleton Laboratory to meet the need for more accurate SST measurements from space. The precise calibration (0.1 K [*Mason et al.*, 1996]) of this instrument results in a more stable measurement of SST than the AVHRR sensors provide (0.55 K [*Weinreb et al.*, 1990]), whilst a novel viewing geometry allows two near simultaneous ($\Delta t = 150$ s) measurements of the same patch of ocean through different atmospheric path lengths. This should result in a more accurate atmospheric correction than that achieved by the AVHRR instruments [*Zavody et al.*, 1995]. Validation studies have shown that in most areas of the ocean the overall relative accuracy of a point measurement of SST from ATSR is around 0.3-0.4 K [*Mulrow et al.*, 1994; *Barton et al.*, 1995; *Forrester and Challenor*, 1995; *Harris and Saunders*, 1996], which is an improvement on the 0.6 K accuracy of AVHRR [*McClain et al.*, 1985].

The SST data used are the Synthesis of ATSR Data Into Surface Temperature (SADIST) version 500 spatially averaged 0.5° SST data (ASST) from ATSR over the same time period as the T/P data. Studies by *Jones et al.* [1996a, b] have shown that a problem exists in the nighttime ASST data in specific geographical locations associated with marine stratiform cloud formation. Within the South Atlantic region the affected area lies between 30° and 50°S and extends to the equator along the west coast of Africa. This contamination manifests as an

increased nighttime SST variability and a decreased nighttime mean SST in such regions and is caused by inadequate cloud screening of the data (partly caused by the failure of the 3.7 μm channel on May 26, 1992). A filtering scheme to reduce the cloud contamination is proposed and tested by Jones *et al.* [1996a, b]. This scheme involves fitting an annual and semiannual cycle to the daytime data and rejecting data that are further than three standard deviations of the daytime residuals away from the model. This results in the rejection of 4.7% of the ASST data (most of which is nighttime data). The data used in this study are filtered with this scheme.

4.2. SST Overlaid on Dynamic Topography

Initially, patterns of absolute SST and SSH are compared. To obtain an estimate of absolute SSH, the OSU91A-JGM3 (Joint Gravity Model) hybrid geoid is used. This geoid is subtracted from the time-averaged mean SSH at each subsatellite location. The resulting dynamic topography contains both the ocean signal and the geoid error due to our lack of knowledge of the geoid on small scales (<2000 km, e.g., Nerem *et al.* [1994]). To remove the small-scale geoid errors, the dynamic topography map is smoothed and interpolated to a 0.5° grid using a Gaussian filter with a full width at half maximum (FWHM) of 800 km. The SSH anomaly

(from the T/P 1993-1994 mean SSH) for January 1993 is then interpolated onto a 0.5° grid using a Gaussian interpolator with a FWHM of 180 km. The smoothed mean dynamic topography is then added to the height anomaly map to obtain an estimate of the absolute dynamic topography. This estimate will be accurate in regions where the small spatial scale, temporal mean flow is weak.

The dynamic topography that the above method yields for January 1993 is converted into geostrophic velocity (by first differencing the gridded dynamic topography field and using the geostrophic equation) and overlaid on the ATSR SST averaged within the same month. An example from the Agulhas region is shown in Plate 1. It is clear that in some areas a striking similarity exists between the SST contours and the T/P flow field. For example, the large meanders (wavelength ~ 600 km) in the Agulhas Return Current at 37°S , $30\text{--}40^\circ\text{E}$ are present in both the flow field and the SST field; regions of high dynamic topography correspond to regions of high temperature and vice versa. At 41°S , 18°E , however, there is an anticyclonic eddy present in the altimetric flow field which is not clearly visible in the SST. Similar patterns in these two fields can be observed in many energetic regions in this way, such as the Brazil-Falklands confluence and the northern area of the Drake Passage (not shown).

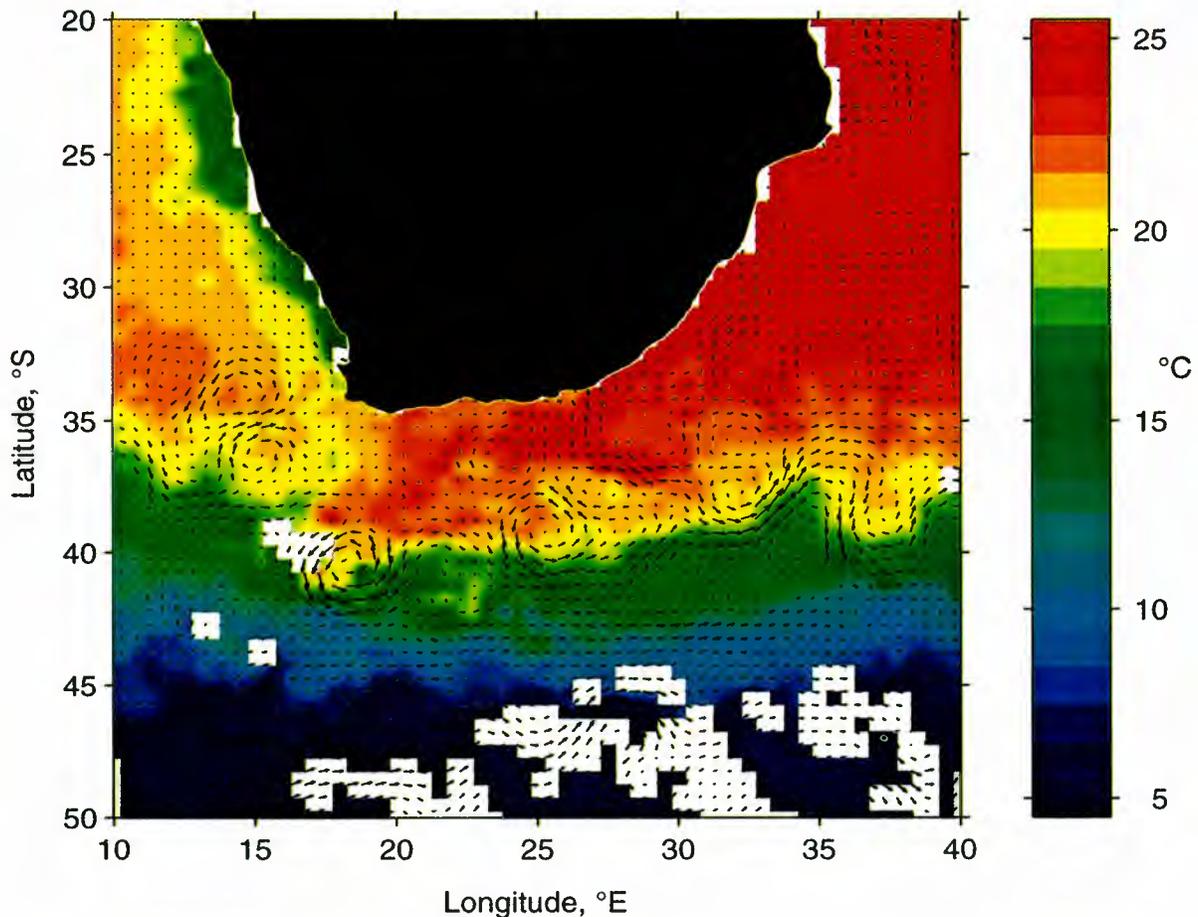


Plate 1. TOPEX/POSEIDON (T/P) "absolute" (see text for details) geostrophic flow vectors overlaid on the along-track scanning radiometer (ATSR) mean sea surface temperature (SST) for January 1993. White areas denote missing SST data because of clouds. Note the similarity between the geostrophic flow and the isotherms at 35°E , 38°S .

Given that these two data sources are completely independent, that the altimetry is interpolated, and that the method of estimating absolute dynamic topography is error prone (because of the omission of the small-scale mean flow), it is surprising to see such marked similarity between the two fields in certain areas. Although this is proof that there certainly is mutual information in both data sources, establishing a quantitative estimate of the relationship between SST and SSH from this type of comparison is difficult because of the large errors (due to interpolation and inadequate knowledge of the geoid on small scales) involved. Thus, to quantify the relationship between SST and SSH, a different approach is taken, which reduces the errors both in interpolating the altimetry and in calculating absolute dynamic topography. This is described in the following section.

4.3. Comparisons of ΔSST With ΔSSH

To narrow the uncertainties, a more quantitative attempt is made to relate SST and SSH by using SSH anomalies (from the 1993-1994 mean SSH) along the T/P ground tracks. ATSR SST anomalies (from the 1993-1994 mean SST) are then collocated onto the T/P ground tracks by using matchup criteria between the SST and SSH data of ±2 days and ±60 km. The criterion of 60 km was selected by considering that the spatial resolution of the ASST data set is 55 km at the equator. This immediately sets a minimum spatial matchup criterion of 32 km (the distance of the corner of an ASST cell from its center), since if the criterion is less than this, there will be locations where it is not possible to collocate the SST even with full coverage. It is also desirable that each collocation point has the potential to obtain information from more than one ASST cell. This is to reduce the effects of noise and increase the amount of matchups available. To achieve this, a criterion of >55 km is necessary, and 60 km was chosen. The criterion of 2 days was chosen after experimenting with various criteria and taking a compromise between a small time criterion allowing more precise but fewer matchups and a large time criterion allowing a larger quantity of matchups.

An estimate of the SST at the collocation point (T_0) is then obtained from the SSTs within the matchup criteria by using the following equation

$$T_0 = \frac{\sum_{i=1}^N w_i T_i}{\sum_{i=1}^N w_i} \tag{6}$$

where T_i are the SSTs within the matchup criteria, N is the number of SSTs within the matchup criteria, and the weights w_i are given by

$$w_i = e^{-\left(\frac{(x_i - x_0)^2}{\sigma_d^2} + \frac{(t_i - t_0)^2}{\sigma_t^2} \right)} \tag{7}$$

where $(x_i - x_0)$ are the distance of the SSTs from the collocation point, $(t_i - t_0)$ are the time between the SST measurements and the altimeter SSH measurement at the collocation point, and σ_d and σ_t are 30 km and 1.2 days, respectively.

The purpose of such a weighting function is to attach more weight to data nearer in time and space to the collocation point and less weight to data near the edge of the matchup criteria.

The results presented here are largely insensitive to the choice of parameters. To ensure this, the results described below were also calculated with $w_i = 1$ (i.e., just a simple box average). It was found that the correlations between ΔSST and ΔSSH are only marginally smaller ($\approx 0.01-0.05$) using the simple average compared to using the Gaussian weighting.

The collocation of the ΔSST onto the T/P ground tracks does not present the same problems as interpolation of altimeter data onto a regular grid because the sampling of the ASST data is regular and well posed for this purpose. The alternative method of interpolating the altimeter data onto a 0.5° grid and comparing this with the ASST data would incur large interpolation errors in the middle of the track diamonds [e.g., Greenblade *et al.*, 1997] and has therefore been avoided. These collocated data are analyzed over the entire South Atlantic region for 1993 and 1994 to identify the regions and seasons when correlations between ΔSST and ΔSSH exist.

Three profiles of collocated ΔSST and ΔSSH along T/P ground tracks are shown in Figures 1a-1c for pass 174 (Figure 1d) and cycles 11, 30, and 49 (this corresponds to January 7, 1993; July 14, 1993; and January 18, 1994, respectively). In certain places, there are close similarities between ΔSST and ΔSSH, and it is clear that the shapes of ΔSST and ΔSSH profiles are related. For example, in Figure 1a, a similarity in the region near 37°S is evident. In Figure 1b the structure of the SST and SSH profiles between 38° and 45°S is similar. In Figure 1c the SST and SSH structure at around 40° and 35°S appears related. It is also clear, however, that the relationship is far from simple. Indeed, it is impossible to deduce from these few profiles any statistically significant relationship between ΔSST and ΔSSH.

As a preliminary test to decide upon the parameters to cross correlate, data in the Agulhas region (0-40°E, 10-50°S) are considered. Zero lag cross-correlation coefficients (r) are calculated for each T/P cycle using (8) below

$$r = \frac{\sum_{i=1}^N (\eta'_i - \bar{\eta})(s'_i - \bar{s})}{N\sigma_\eta\sigma_s} \tag{8}$$

where N is the number of collocated data points in each cycle, s denotes SST, η denotes SSH, σ is the standard deviation, a dash denotes anomalies from the time mean at a particular location, and an overbar denotes the mean over that particular T/P cycle.

The results of these calculations for 1993 and 1994 (T/P cycles 11-84) are shown in Figure 2. The cross correlations between ΔSST and ΔSSH, between ΔSST gradients and ΔSSH gradients, between ΔSSH and ΔSST gradients, and between ΔSSH gradients and ΔSST are shown. It is clear from the temporal consistency of the time series shown in Figure 2 that the correlations between ΔSST and ΔSSH and between the gradients of ΔSST and ΔSSH are all consistently different from zero, whereas correlations between the gradients of one parameter and the actual values of the other parameter are much weaker. The correlations between gradients of ΔSST and gradients of ΔSSH are nearly always slightly less than the correlations between ΔSST and ΔSSH. This can be attributed to two factors: (1) removing any large-scale correlation between ΔSST and ΔSSH, since taking the gradient high-pass filters the data and (2) increasing the noise in the data due to the calculation of gradients and therefore reducing the

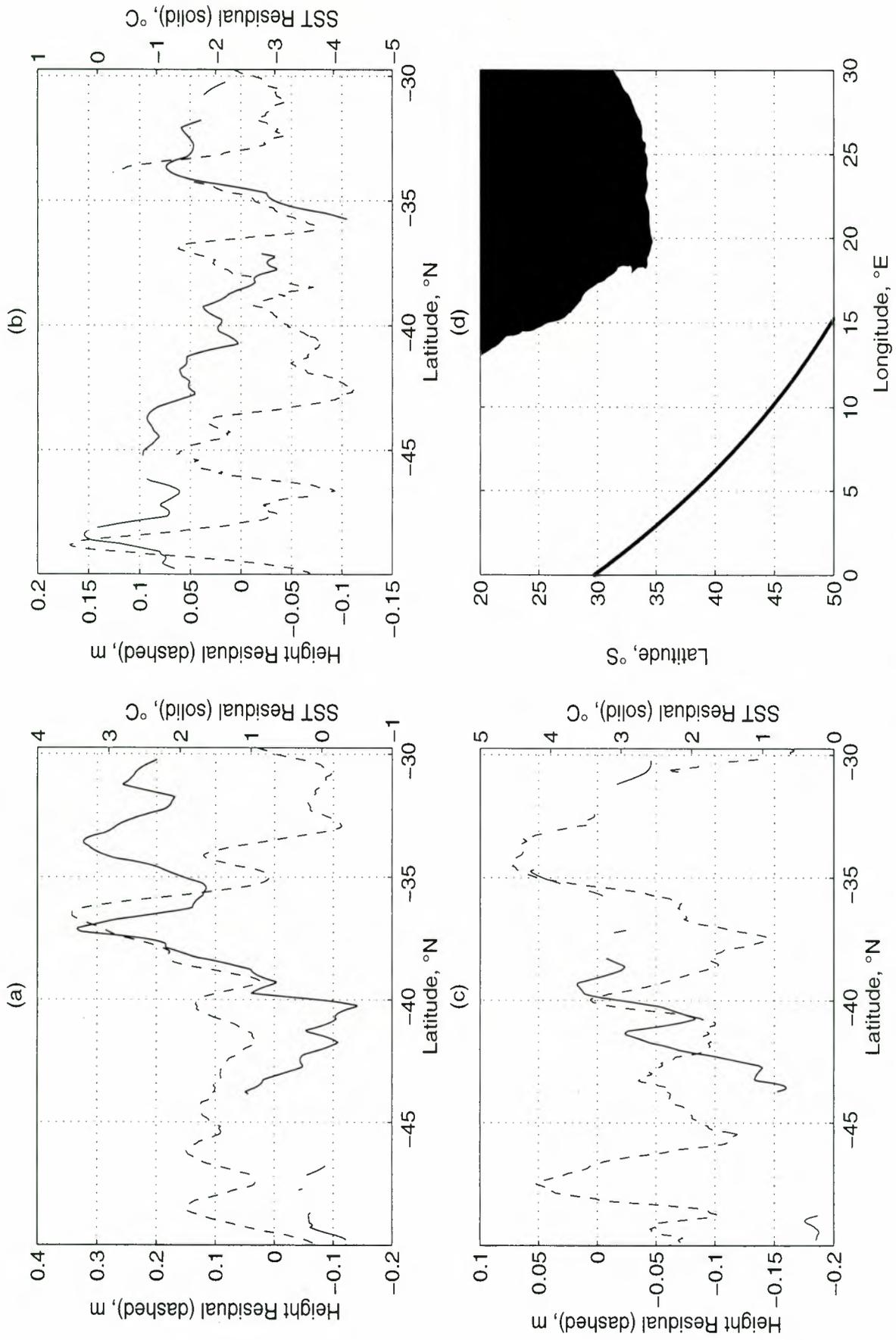


Figure 1. Examples of collocated ATSR Δ SST and T/P Δ SSH for T/P pass 174. (a) Cycle 11, January 7, 1993; (b) cycle 30, July 14, 1993; and (c) cycle 49, January 18, 1994. (d) The location of T/P pass 174. In many cases the structure of the Δ SST is similar to that of Δ SSH, although the relationship is far from simple.

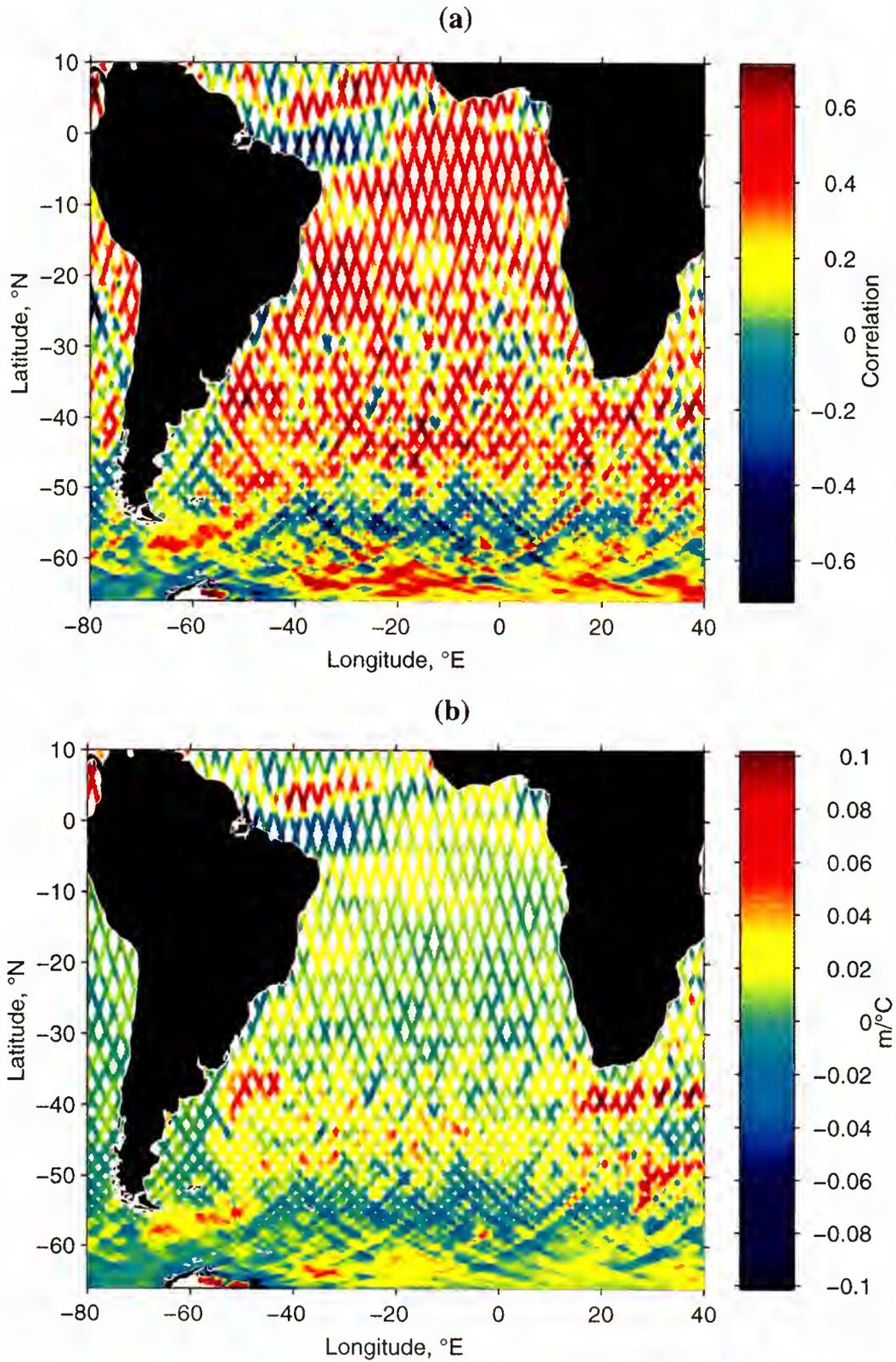


Plate 2. (a) The zero-lag cross correlation between Δ SST and Δ SSH calculated with successive 240 km along-track segments and (b) the gradient $\partial(\Delta$ SSH)/ $\partial(\Delta$ SST) calculated for the years 1993 and 1994. These results are dominated by the large-scale seasonal response of the ocean to heat and freshwater fluxes.

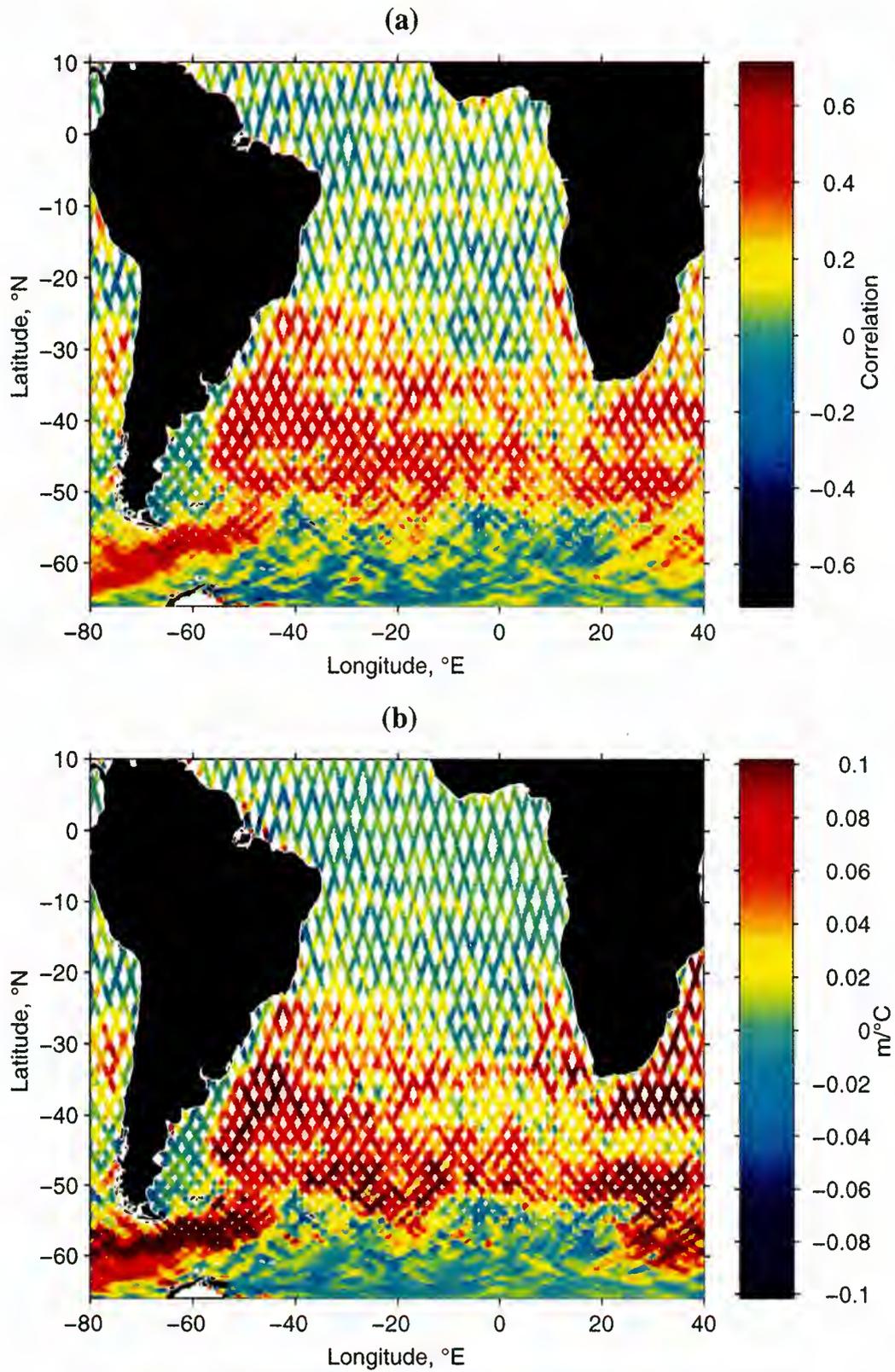


Plate 3. (a) The zero-lag cross correlation between ΔSST and ΔSSH calculated with successive 240 km along-track segments and (b) the gradient $\partial(\Delta\text{SSH})/\partial(\Delta\text{SST})$ calculated for the years 1993 and 1994 from data filtered to remove wavelengths >400 km. The correlations are large in regions associated with fronts and mesoscale variability.

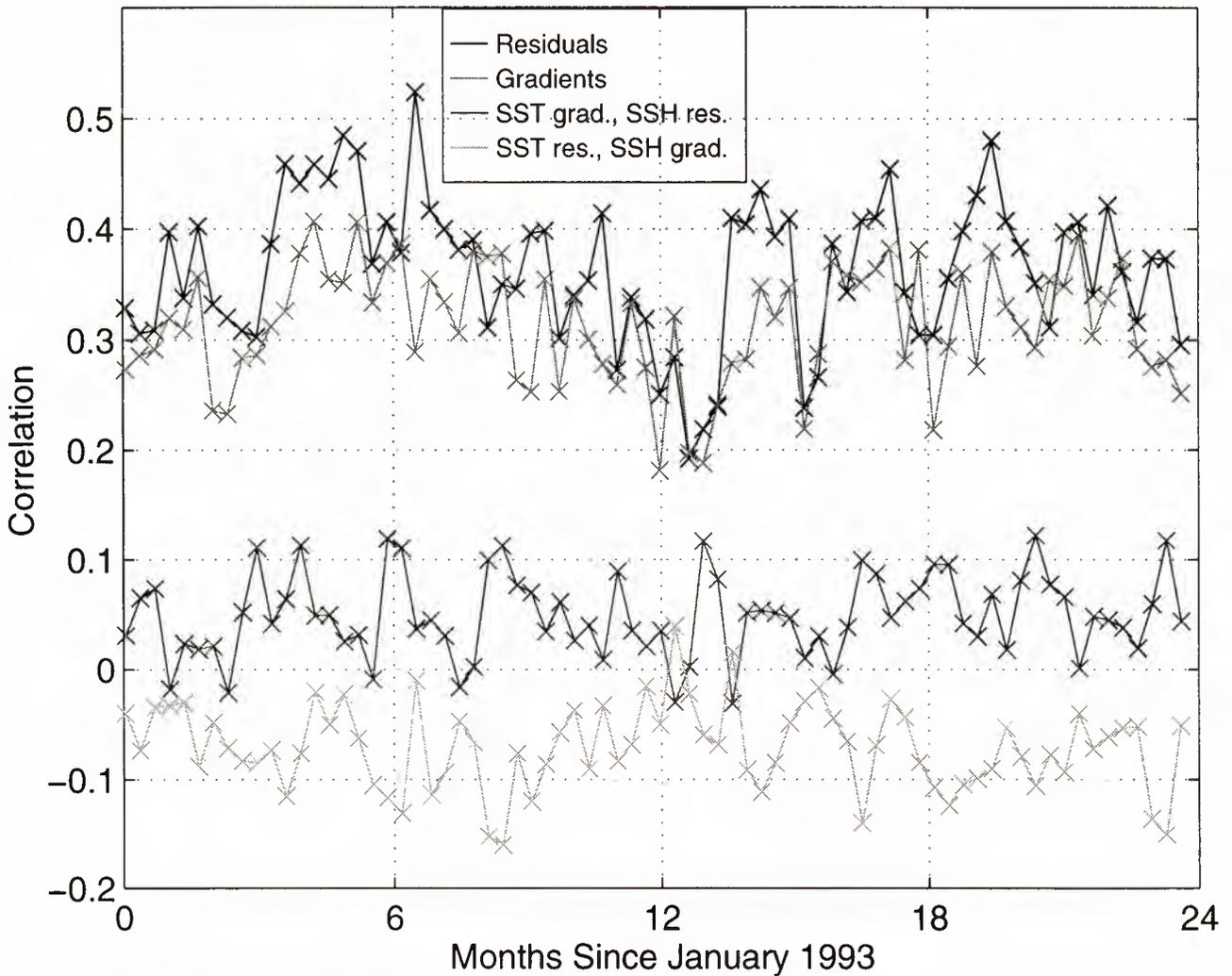


Figure 2. Time series showing the correlation between Δ SST and Δ SSH, Δ SST gradients and Δ SSH gradients, Δ SST gradients and Δ SSH, and Δ SST and Δ SSH gradients. The correlations are calculated for each cycle over the Agulhas region.

correlations. It is likely that both of these contribute to the slight lowering of the gradient correlations. The magnitude of the correlations, although significantly different from zero (see discussion in section 4.4), is fairly small. Typical values of about 0.35 imply that SST variability can only explain about 13% of the SSH variability. Even the maximum value of 0.53 means that only 28% of the variability of one field can be explained by the other. One of the reasons why the correlations are so low is that these are correlations over the entire Agulhas region. If the relationship between SST and SSH is strong in some areas, but weak in others, or $\partial(\Delta$ SSH)/ $\partial(\Delta$ SST) varies throughout the region, the net result will be an intermediate level of correlation. The geographical variation of the correlations is explored in the next section.

From this preliminary investigation in the Agulhas region, it is clear that (1) a relationship does exist between Δ SST and Δ SSH, (2) the relationship is strongest between Δ SST and Δ SSH rather than between the gradients of these parameters or any combination thereof and (3) a seasonality exists in the strength of the correlations. This study is extended in the following sections using data for the whole of the South Atlantic region and studying in more detail the geographical and seasonal variation of the correlations between Δ SST and Δ SSH.

4.4. Geographical Dependence of the Δ SST Δ SSH Relationship

The results of the previous section are interesting, but to investigate fully any relationship between Δ SST and Δ SSH, it is important to quantify the geographical dependence. One method of doing this is to calculate correlations between Δ SST and Δ SSH from time series at fixed geographical locations (the subsatellite points that the Δ SST have been collocated to). An estimate of the use of such a calculation can be made by examining the significance of correlations calculated in such a way. Assuming that Δ SST and Δ SSH are binormally distributed, it can be shown that

$$t = r \sqrt{\frac{N-2}{1-r^2}} \tag{9}$$

where t is distributed as Student's t statistic with $N - 2$ degrees of freedom, r is the correlation given by (8), and N is the number of independent samples.

An estimate of the number of independent samples in each time series is made by dividing the length of the time series by the integral timescale. For midlatitude regions a typical timescale is ≈ 25 days [Stammer, 1997]. For a 2 year time series this gives $730/25 = 29.2$ independent samples.

Substituting for $N = 29.2$ and $t = 2.05$ (95% confidence limits for a Student's t distribution with 27.2 degrees of freedom) gives $r = 0.37$. This shows that correlations smaller than ≈ 0.4 are not significantly different from zero. It was shown in section 4.3 that, on a regional average, correlations between Δ SST and Δ SSH are around 0.4. Hence it is desirable to increase the number of independent samples to reduce the level at which correlations become distinguishable from zero. This can be accomplished by using an along-track window and correlating Δ SST and Δ SSH over a segment of ground track rather than calculating correlations at single points. The number of independent samples within a spatial window can be approximated by the window length divided by the integral length scale. A typical value for an ocean midlatitude integral length scale is 100 km [Stammer, 1997]. Choosing a window length of 240 km (39 along-track data points) gives $240/100 \times 730/25 = 70$ independent samples. It follows (from (9)) that correlations >0.23 are significantly different from zero at the 95% level. The high along-track spatial resolution can be maintained by overlapping successive windows, although, of course, this has a smoothing effect. The means and standard deviations used to calculate the correlations are computed from the ensemble of data within each 240 km window for all 74 T/P cycles. A 240 km window was chosen after experimentation with several different window lengths and is thought to be a reasonable compromise between increasing the number of independent samples and maintaining a useful spatial resolution.

We applied this technique to all the collocated data within the South Atlantic region for 1993 and 1994. The resulting correlation map for the South Atlantic is shown in Plate 2a. This plate shows that positive correlations between 0.2 and 0.6 dominate in most locations. Two exceptions, where negative correlations occur, are (1) the region centered on the equator and stretching from the Amazon Delta to Liberia and (2) the region across most of the South Atlantic between 50° and 58°S. Caution must be used when interpreting correlations at high latitudes since the seasonal variation in cloud cover means that matchups between Δ SST and Δ SSH are biased toward summer months. Although care was taken to reduce remnant cloud contamination in Δ SST data by using the filtering scheme of Jones *et al.* [1996a, b], there is still a possibility that some cloud contamination is still present. Marine stratiform cloud is known to occur in the region 40°–60°S in the austral summer [Klein and Hartmann, 1993]. Cloud contamination usually acts to reduce SST. Hence cloud contamination in summer months would cause a negative correlation between Δ SST and Δ SSH (assuming that Δ SSH is positive in summer). This, coupled with a reduction in the variability at latitudes higher than 50°S [Knudsen *et al.*, 1996], could lead to the observed pattern of negative correlation. The equatorial region of negative correlation is more interesting. Cloud contamination is not thought to be a problem in this area, and although the cloud cover varies (because of the movement of the intertropical convergence zone), data gaps are not nearly so frequent as at high latitudes. A possible cause for this region of negative correlation is that the variation in the net freshwater flux is dominating the steric response of the ocean here. This situation may occur when the net flux of freshwater into the ocean is out of phase with the net heat flux into the ocean and dominates in terms of the changes in density that are caused. A study into whether this is the case is possible given heat and freshwater fluxes together

with knowledge of mixed-layer depth. This, however, will be the subject of future research.

The region of positive correlation is fairly homogeneous with values ranging mostly from 0.2 to 0.6 for latitudes between 0° and 50°S. Several areas with correlations as high as 0.7 are evident in regions of high mesoscale variability (e.g., the Brazil-Falklands confluence at 36°S, 50°W and the Agulhas Return Current region (40°S, 25°E)). Interpreting this map is difficult because, as described in section 3, correlations may be due to the seasonally varying heat fluxes (which is thought to dominate on large scales (wavelengths >1000 km)) or to horizontal advection (which is thought to dominate on small scales (wavelengths <1000 km)). The map of $\partial(\Delta$ SSH)/ $\partial(\Delta$ SST) is shown in Plate 2b. This is calculated by fitting a straight line to the ensemble of data within each 240 km window using ordinary least squares with Δ SST as the independent variable. It is clear that throughout most of the region the gradient between Δ SSH and Δ SST is ≈ 0.01 – 0.02 m/°C. This is typically the magnitude that is associated with the ratio of the amplitudes of the large-scale seasonal variations in Δ SST and Δ SSH due to the seasonally varying heat fluxes [Knudsen *et al.*, 1996]. In regions associated with mesoscale variability (e.g., Agulhas Retroflexion region and the northern part of the Drake Passage), gradients are as high as 0.1 m/°C. This is more typical of the signal caused by the horizontal advection of ocean fronts and eddies [e.g., Gordon and Haxby, 1990]. However, a more rigorous interpretation of the correlations cannot be made unless a scale analysis is performed. This is described in the next section.

4.5. Spatial Scale of the Δ SST Δ SSH Correlations

The correlation map described in the previous section included contributions from the whole range of wavelengths from ≈ 100 km upward (wavelengths below ≈ 100 km are not resolvable because of the 0.5° resolution of the ASST data). To gain some insight into the physical mechanism causing the correlations, we filter the data into three parts. The data are (1) high-pass filtered by removing an along-track running mean with a 240 km (39 point) window, (2) band-pass filtered by smoothing with a 240 km running mean and then subtracting a 600 km running mean, and (3) low-pass filtered by smoothing using a 600 km running mean. The transfer functions of these filters were calculated, and the demarcations between the filters were evaluated by choosing the wavelengths at the half-power points of the transfer functions. This results in the data, filtered as above, containing wavelengths of (1) 100–400 km, (2) 400–960 km, and (3) 960 km and greater, respectively. The same methodology as in section 4.4 but using filtered data produces scale dependent correlation maps.

The correlation map for the large scales is virtually identical to the correlation map obtained with no filtering (Plate 2a) and is therefore not shown. The gradient map is also similar, although in the regions of high mesoscale variability the gradients are reduced. Instead of reaching 0.1 m/°C in these regions (as shown in Plate 2b), the gradients maximize at around 0.04 m/°C. The large-scale correlations are dominated by the seasonal response of the ocean to the seasonally varying heat and freshwater fluxes [Gill and Niiler, 1973]. An analysis of the similarities between Δ SST and Δ SSH on large scales is described by Knudsen *et al.* [1996]. For many purposes (see section 5) a large-scale correlation between

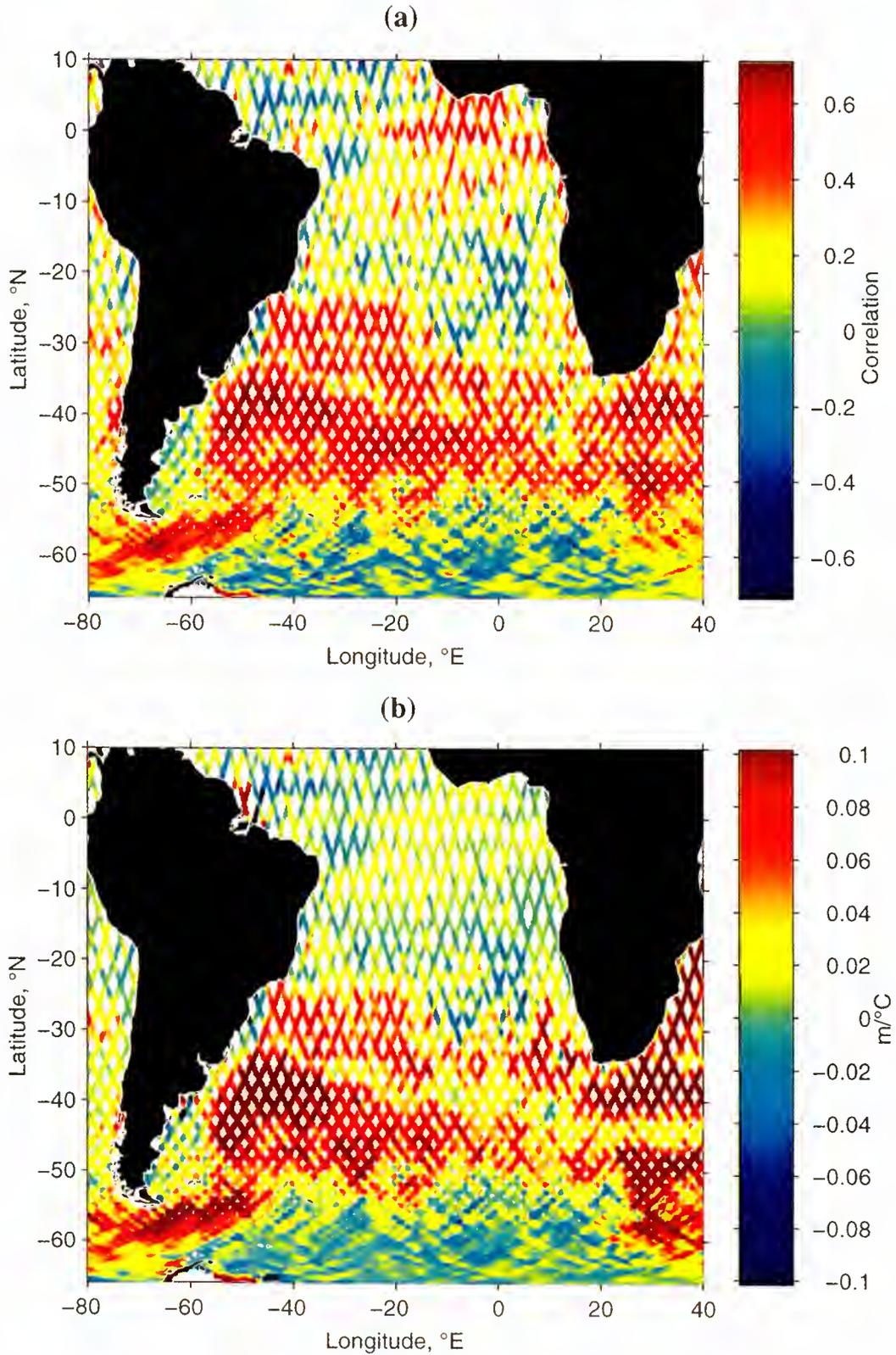


Plate 4. (a) The zero-lag cross correlation between ΔSST and ΔSSH calculated with successive 240 km along-track segments and (b) the gradient $\partial(\Delta\text{SSH})/\partial(\Delta\text{SST})$ calculated for the years 1993 and 1994 from data filtered to remove wavelengths <400 km and >1000 km. A similar pattern is seen to that in Plate 3. The difference is that the correlations are ≈ 0.1 larger, possibly owing to the requirement for more stringent collocation criteria on smaller scales.

Δ SST and Δ SSH is not useful. The focus of this paper is on correlations at small and intermediate wavelengths (which means that the large-scale seasonal response of the ocean is excluded). This is described in the remainder of this section.

Correlations caused by variability on wavelengths smaller than 400 km are shown in Plate 3a. It is clear that the correlations between Δ SST and Δ SSH are regional in this case, with high correlations (~ 0.6 - 0.7) in regions associated with mesoscale variability and strong fronts. For example, the region of high correlation stretching through the northern part of the Drake Passage is in the vicinity of the Subantarctic and Polar Fronts, as is the broad swath of high correlation extending across the latitude band 40° - 50° S. The region of high correlation just southeast of the tip of South Africa is associated with the Agulhas Return Current, and the region to the west of South Africa is associated with the Benguela Current. An interesting area of high correlation extends from 30° S at the east coast of South America to 35° S at 0° E. The reason for this is unclear; however, it may be due to eddies generated in the Agulhas region and propagating across the South Atlantic. The structure of the correlations in Plate 3a is quite different to the unfiltered correlation map (Plate 2a), where the correlations are fairly homogeneous throughout most of the region. The map of $\partial(\Delta$ SSH)/ $\partial(\Delta$ SST) for the small scales (Plate 3b) is also quite different from the equivalent map for all scales (Plate 2b). The gradients in the regions of high correlation (and high mesoscale variability) are now much higher at around 0.1 - 0.2 $m^\circ C$, indicating that different physical processes are responsible for the correlations in these regions.

The correlation map for intermediate scales (Plate 4a) is very similar to the equivalent map for small scales (Plate 3a). The major difference is that the correlations in most regions are slightly higher (typically 0.7 in regions of high mesoscale variability rather than 0.6). This reduction in correlation at smaller scales may purely be due to the matchup criteria used to collocate the data. For example, assuming a horizontal advection speed of 5 cm/s [e.g., Gordon and Haxby, 1990], a feature can move 9 km in 2 days (the matchup criteria). Hence, if it had a wavelength of less than $4 \times 9 = 36$ km, it would be completely decorrelated. Obviously, the longer the wavelength of the feature, the less the reduction in correlation. The collocation therefore places a limit on the extent to which the actual small-scale dependence of the relationship between SST and SSH can be established. One region where the correlations at intermediate scales are less than those at small scales is the region associated with the Benguela Current (and with upwelling) to the west of the South African coast. The gradients associated with the correlations at intermediate scales are shown in Plate 4b. The gradients are similar but slightly larger than those shown for the smaller scales (Plate 3b) with gradients in regions of high correlation reaching 0.15 - 0.2 $m^\circ C$.

The implication of Plates 3 and 4 is that there is indeed a relationship between Δ SST and Δ SSH in specific geographical regions associated with mesoscale variability. The square of the correlation values is the percentage of the variance in one parameter that can be explained by the other parameter in a linear relationship. Values in some regions in Plates 3 and 4 are as high as 50%. The fact that these correlations are present at small and intermediate wavelengths imply that they are caused by large eddies, meanders, or Rossby waves [Cipollini

et al., 1997] rather than by the very large-scale seasonal response of the ocean to varying heat and water fluxes. Distinguishing between mechanisms operating on different scales is critical if any relationship between Δ SST and Δ SSH is to be used to interpolate altimeter data, provide improved eddy statistics, or to validate eddy-resolving models. If the relationship between Δ SST and Δ SSH was purely due to the large-scale seasonal response, there would not be a relationship on small enough spatial scales (~ 100 km) to be useful for the above mentioned purposes.

4.6. Seasonal Cycle in the Δ SST Δ SSH Correlations

Figure 3 shows a time series of the spatial mean cross correlations over the entire South Atlantic region for intermediate wavelengths. It is clear that an annual signal exists in these correlations. Therefore a model of a sinusoid with annual period plus trend is fitted to the data using conventional least squares. The peak of this cycle is -0.38 in mid-August, the middle of the austral winter, compared to a trough of -0.29 in midsummer. Hence the correlations in the winter are $\sim 30\%$ stronger than in the summer. There are two possible explanations for this seasonal cycle. First, it is possible that development of a diurnal thermocline might be stronger in summer, and hence the correlations are lower because of this effect decoupling the surface from the subsurface. Second, the development of a seasonal thermocline in midlatitudes may be responsible for the poor correlations in summer. This is the effect that Legeckis [1978] observed, where summer heating creates a horizontally homogeneous warm pool of water that masks the underlying frontal structures. Legeckis observed that between latitudes of 25° and 35° in both hemispheres, frontal structures are visible in winter but not in summer, whereas for higher latitudes frontal structures are visible throughout the year.

To investigate whether or not the diurnal thermocline effect is causing a seasonal cycle in the correlations, the above analysis was undertaken for day and night SST data separately. The results (not shown) indicate, surprisingly, that there is virtually no difference between the correlations according to whether day or night SST data are used. The diurnal thermocline effect, in some circumstances, is known to obscure the underlying thermal structure [e.g., Stramma *et al.*, 1986]. The fact that no difference in the day and night correlations occurs suggests that although the diurnal thermocline effect may occur occasionally, it is not the dominant effect in terms of obscuring underlying structure.

To examine the geographical variation of the seasonal differences, the small and intermediate scale correlation maps are calculated for January, February and March (JFM) and for July, August and September (JAS). These maps (not shown) are not as clear as maps for the entire two year period due to the reduction in the number of independent samples. However, they suggest that the winter correlations are stronger than the summer correlations mainly in the region of high correlation extending across the South Atlantic at ~ 30 - 35° S (e.g. Plate 4a) and in the region associated with the Benguela Current at $\sim 25^\circ$ S, 10° E (Plate 3a). These regions are both at latitudes close to those most affected by the seasonal thermocline [Legeckis, 1978] which is tentative evidence that this is the main cause of the differences in correlation between summer and winter.

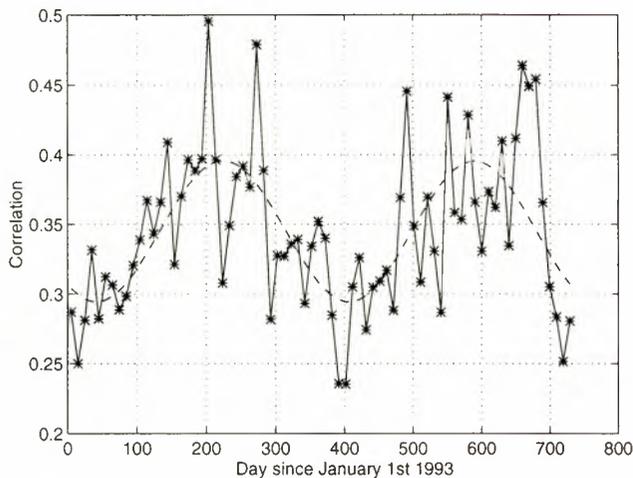


Figure 3. The seasonal cycle in the strength of the correlations between Δ SST and Δ SSH over the whole South Atlantic region. An annual plus drift model is fitted to the data (see text for details). The correlations minimize in mid-February and maximize in mid-August. A possible mechanism for this is the development of the seasonal thermocline during summer and the obscuration of the underlying density structure that this may cause.

5. Application

In most forms of remote sensing of the ocean a major difficulty is obtaining information about the subsurface ocean from purely surface parameters. SST, as measured by infrared radiometers, is the temperature of the top few microns of the sea surface, and thus any relationship between SST and SSH is evidence that the temperature of the top few microns of the ocean surface contains information about the entire water column. In particular, the SST structure must be correlated with the temperature structure to some depth before it will correspond to a similar SSH structure. Establishing the regions and seasons where there is any relationship between SST and SSH will show when and where to expect the SST structure as observed from infrared radiometers to represent more than just the surface layer. Furthermore, by examining the temperature-salinity structure of the water column (from ocean model data, or wherever possible from in situ data) in regions where a relationship holds, it may be possible to infer dynamical information. Examples of such information are the baroclinicity of the variability or the depth to which surface temperature structure is manifested in subsurface temperature structure.

As global ocean models become more sophisticated and are able to represent the state of the ocean more accurately [Semtner, 1995], it becomes more important to test these models to assess their performance. Remote sensing provides the only method for global validation of models at the high spatial resolution that these models require. Traditionally, altimeter SSH data are used, and statistics such as SSH variability and eddy kinetic energy (EKE) are compared to the corresponding model statistics [Stammer *et al.*, 1996]. A more stringent test of a model, however, would be to compare the relationship between SST and SSH as observed from satellites with the corresponding SST-SSH relationship from the model. To achieve a realistic coherency between SST and SSH, the model must correctly represent both the near-surface thermodynamics and the deeper structure of the water column.

This is a challenging task but is especially important for coupled ocean-atmosphere models where SST is a central factor.

Establishing a relationship between SST and SSH would open up the way for new techniques based on any such relationship. One example is the interpolation of altimeter data. For an altimeter with the sampling characteristics of T/P the results of Greenslade *et al.* [1997] suggest that it is impossible to consistently resolve scales with wavelengths shorter than about 400 km. However, if there is any relationship between SST and SSH, it may be possible to use SST data within an optimal interpolation scheme for SSH, therefore providing pseudo-SSH fields of much higher resolution than that given by altimetry alone. This would be useful for tracking small wavelength features such as mesoscale eddies, which are difficult to track as they disappear between altimeter tracks. It would also be helpful for assimilating into ocean models where a high-resolution gridded SSH field would be much easier to use than along-track altimeter data.

A second example where a relationship between SST and SSH would be of use is in the calculation of eddy kinetic energy (EKE) (the variance of the geostrophic velocity anomalies). To calculate EKE from altimetry, both horizontal components of the velocity field are required. Hence either EKE calculations are limited to crossover points or isotropy is assumed and EKE is calculated at the full along-track resolution of the altimeter. However, the assumption of isotropy is known to be poor in certain regions [Morrow *et al.*, 1994], especially in the Southern Ocean where the bottom topography exerts a large influence on the eddy field. Hence EKE maps calculated at the full resolution of the altimeter will be in error. SST fields are fully two dimensional, however, and if there is any relationship between SST and SSH, it may be possible to use the SST data to remove the assumption of isotropy and calculate more accurate high-resolution EKE fields.

6. Conclusions

Following the introduction, this paper commences with a literature review of previous work relating SST and SSH. This is followed by a brief outline of the reasons why a relationship could exist between these two parameters. An observational study from T/P SSH and ATSR SST shows that a relationship does indeed exist. The main conclusions of this study are described below.

It is first shown that if the T/P geostrophic flow vectors are overlaid on the SST for January 1993, a remarkable agreement is present in some locations. Meanders in the geostrophic flow field in the Agulhas Return Current closely follow isotherms. This qualitative agreement is investigated using collocated SST and SSH anomalies.

A pilot study in the Agulhas region shows that the spatial correlation between height and temperature residuals is slightly stronger than the correlation between along-track gradients of height residuals and along-track gradients of temperature residuals. Furthermore, correlations between temperature residuals and along-track gradients of height residuals are much weaker (≈ 0.05 compared to ≈ 0.4).

The geographical variation in the correlations between SST and SSH is investigated by computing zero-lag cross-correlation coefficients within overlapping successive 240 km

along-track segments for 1993 and 1994. The resulting correlation map (Plate 2a) shows correlations ranging from 0.2-0.6 throughout most of the South Atlantic. This pattern is dominated by the large-scale correlation between Δ SST and Δ SSH due to seasonally varying heat and freshwater fluxes. When the large-scale variability (wavelengths >1000 km) is removed from the data, the correlation pattern is quite different. Regions of high correlation (≈ 0.7) are now evident in areas generally associated with ocean fronts and mesoscale variability. Such regions include the Polar Front, the Sub-Antarctic Front, the Agulhas Return Current, the Benguela Current, and the Brazil-Falkland Confluence. A relationship between SST and SSH at these small and intermediate spatial scales will allow techniques such as interpolation of altimeter data and improvement of eddy statistics to be developed.

The strength of the small- and intermediate-scale correlation (wavelengths <1000 km) between Δ SST and Δ SSH is shown to vary with season, with correlations in mid-August of 0.38 (austral winter) being $\sim 30\%$ higher than correlations in mid-February of 0.29 (austral summer). The cause of this seasonality is investigated, first by studying the correlations using day and night SST data separately. It is found that using day rather than night SST data, and vice versa, makes virtually no difference to the strength of the correlations. This rules out the diurnal thermocline as a mechanism for the seasonal variation in correlations. The geographical pattern of the correlation map for summer and winter reveals that the main difference between the correlations is at latitudes of $\sim 30^\circ$ S. This suggests that the seasonal thermocline and the associated horizontally homogeneous pool of warm water, caused by summer heating, is obscuring the underlying temperature structure and causes lower correlations [Legeckis, 1978]. Extending the results of this study to gain more information about the dynamical processes occurring as well as developing new techniques based on the correlations between SST and SSH are exciting subjects for future research.

Acknowledgments. The authors are grateful to ESA, Chris Mutlow, Albin Závody, and the ATSR Science Team at RAL for supplying the ATSR data. Thanks are also due to AVISO for provision of the TOPEX/POSEIDON data and to Helen Snaith for help with the altimeter data processing. Zöe Jones is thanked for her comments on the manuscript. Matthew Jones is funded by grant GR3/10796 from the U.K. Natural Environment Research Council.

References

- AVISO, User Handbook: Merged TOPEX/POSEIDON Products, *Rep. AVI-NT-02-101-CN, Edition 3.0*, AVISO Altimetry, CNES, 1996.
- Barton, I. J., A. J. Prata, and R. P. Cechet, Validation of the ATSR in Australian waters, *J. Atmos. Oceanic Technol.*, **12**, 290-300, 1995.
- Brenner, A. C., C. J. Koblinsky, and B. D. Beckley, A preliminary estimate of geoid-induced variations in repeat orbit satellite altimeter observations, *J. Geophys. Res.*, **95**, 3033-3040, 1990.
- Cheney, R. E., J. G. Marsh, and B. D. Beckley, Global mesoscale variability from collinear tracks of Seasat altimeter data, *J. Geophys. Res.*, **88**, 4343-4354, 1983.
- Cipollini, P., D. Cromwell, M. S. Jones, G. D. Quartly, and P. Challenor, Concurrent altimeter and infrared observations of Rossby wave propagation near 34° N in the northeast Atlantic, *Geophys. Res. Lett.*, **24**, 889-892, 1997.
- Eanes, R. J., and S. V. Bettadpur, The CSR 3.0 global ocean tide model: Diurnal and semi-diurnal ocean tides from TOPEX/POSEIDON altimetry, *Rep. CSR-TM-95-06*, Cent. for Space Res., Univ. of Texas, Austin, 1995.
- Forrester, T. N., and P. G. Challenor, Validation of ATSR sea surface temperatures in the Faeroes region, *Int. J. Remote Sens.*, **16**, 2741-2753, 1995.
- Fu, L.-L., E. J. Christensen, C. A. Yamarone, M. Lefebvre, Y. Menard, M. Dorrer, and P. Escudier, TOPEX/POSEIDON mission overview, *J. Geophys. Res.*, **99**, 24,369-24,381, 1994.
- Gaspar, P., F. Ogor, P.-Y. Le Traon, and O.-Z. Zanife, Estimating the sea state bias of the TOPEX and POSEIDON altimeters from crossover differences, *J. Geophys. Res.*, **99**, 24,981-24,994, 1994.
- Gill, A. E., and P. P. Niiler, The theory of the seasonal variability in the ocean, *Deep Sea Res. Oceanogr. Abstr.*, **20**, 141-177, 1973.
- Gordon, A. L., Brazil-Malvinas Confluence: 1984, *Deep Sea Res., Part A*, **36**, 359-384, 1989.
- Gordon, A. L., and W. F. Haxby, Agulhas eddies invade the South Atlantic: Evidence from Geosat altimeter and shipboard conductivity-temperature-depth survey, *J. Geophys. Res.*, **95**, 3117-3125, 1990.
- Greenslade, D. J. M., D. B. Chelton, and M. G. Schlax, The midlatitude resolution capability of sea level fields constructed from single and multiple satellite altimeter datasets, *J. Atmos. Oceanic Technol.*, **14**, 849-870, 1997.
- Halliwell, G. R., Y. J. Ro, and P. Cornillon, Westward-propagating SST anomalies and baroclinic eddies in the Sargasso Sea, *J. Phys. Oceanogr.*, **21**, 1664-1680, 1991.
- Harris, A. R., and M. A. Saunders, Global validation of the along-track scanning radiometer against drifting buoys, *J. Geophys. Res.*, **101**, 12,127-12,140, 1996.
- Johnson, M. A., Southern Ocean surface characteristics from FGGE buoys, *J. Phys. Oceanogr.*, **19**, 696-705, 1989.
- Jones, M. S., M. A. Saunders, and T. H. Guymer, Global remnant cloud contamination in the along-track scanning radiometer data: Source and removal, *J. Geophys. Res.*, **101**, 12,141-12,147, 1996a.
- Jones, M. S., M. A. Saunders, and T. H. Guymer, Reducing cloud contamination in ATSR averaged sea surface temperature data, *J. Atmos. Oceanic Technol.*, **13**, 492-506, 1996b.
- Klein, S.A. and D.L. Hartmann, The seasonal cycle of low stratiform clouds, *J. Clim.*, **6**, 1587-1606, 1993.
- Knudsen, P., O. B. Andersen, and T. Knudsen, ATSR sea surface temperature data in a global analysis with TOPEX/POSEIDON altimetry, *Geophys. Res. Lett.*, **23**, 821-824, 1996.
- Legeckis, R., A survey of worldwide sea surface temperature fronts detected by environmental satellites, *J. Geophys. Res.*, **83**, 4501-4522, 1978.
- Lutjeharms, J. R., and R. C. van Ballegooyen, The retroflexion of the Agulhas Current, *J. Phys. Oceanogr.*, **18**, 1570-1583, 1988.
- Mason, I. M., P. H. Sheather, J. A. Bowles, and G. Davies, Black body calibration sources of high accuracy for a space-borne infra-red instrument, the along track scanning radiometer, *Appl. Opt.*, **35**, 629-639, 1996.
- McClain, E. P., W. G. Pichel, and C. C. Walton, Comparative performance of AVHRR-based multichannel sea surface temperatures, *J. Geophys. Res.*, **90**, 11,587-11,601, 1985.
- Minster, J. F., C. Brossier, and P. Rogel, Variation of the mean sea level from TOPEX/POSEIDON data, *J. Geophys. Res.*, **100**, 25,153-25,161, 1995.
- Morrow, R., R. Coleman, J. Church, and D. Chelton, Surface eddy momentum flux and velocity variances in the Southern Ocean from Geosat altimetry, *J. Phys. Oceanogr.*, **24**, 2050-2071, 1994.
- Mutlow, C. T., A. M. Závody, I. J. Barton, and D. T. Llewellyn-Jones, Sea surface temperature measurements by the along-track scanning radiometer on the ERS-1 satellite: Early results, *J. Geophys. Res.*, **99**, 22,575-22,588, 1994.
- Nerem, R. S., et al., Gravity model development for TOPEX/POSEIDON: Joint Gravity Models 1 and 2, *J. Geophys. Res.*, **99**, 24,421-24,447, 1994.
- Olson, D. B., G. P. Podesta, R. H. Evans, and O. B. Brown, Temporal variations in the separation of Brazil and Malvinas Currents, *Deep Sea Res., Part A*, **35**, 1971-1990, 1988.
- Peterson, R. G., and L. Stramma, Upper-level circulation in the South Atlantic Ocean, *Prog. Oceanogr.*, **26**, 1-73, 1991.
- Saunders, M. A., and A. R. Harris, Statistical evidence links exceptional 1995 Atlantic hurricane activity to record sea warming, *Geophys. Res. Lett.*, **24**, 1255-1258, 1997.
- Scott, J. C., and A. L. McDowall, Cross-frontal cold jets near Iceland: In-water, satellite infrared, and Geosat altimeter data, *J. Geophys. Res.*, **95**, 18,005-18,014, 1990.
- Semtner, A. J., Modeling ocean circulation, *Science*, **269**, 1379-1385, 1995.
- Smythe-Wright, D., A. L. Gordon, P. Chapman, and M. S. Jones, CFC-113 shows Brazil Eddy crossing the South Atlantic to the Agulhas Retroflexion region, *J. Geophys. Res.*, **101**, 885-895, 1996.

- Stammer, D., Global characteristics of ocean variability from regional TOPEX/POSEIDON altimeter measurements, *J. Phys. Oceanogr.*, 27, 1743-1769, 1997.
- Stammer, D., R. Tokmakian, A. Semtner, and C. Wunsch, How well does a $1/4^\circ$ global circulation model simulate large-scale oceanic observations?, *J. Geophys. Res.*, 101, 25,779-25,811, 1996.
- Stramma, L., P. Cornillon, R. A. Weller, J. F. Price, and M. G. Briscoe, Large diurnal sea surface temperature variability: Satellite and in situ measurements, *J. Phys. Oceanogr.*, 16, 827-837, 1986.
- Van Woert, M. L., and J. M. Price, Geosat and advanced very high resolution radiometer observations of oceanic planetary waves adjacent to the Hawaiian Islands, *J. Geophys. Res.*, 98, 14,619-14,631, 1993.
- Vastano, A. C., and S. E. Borders, Sea surface motion over an anticyclonic eddy on the Oyashio Front, *Remote Sens. Environ.*, 16, 87-90, 1984.
- Vastano, A. C., and R. O. Reid, Sea surface topography estimation with infrared satellite imagery, *J. Atmos. Oceanic Technol.*, 2, 393-400, 1985.
- Vazquez, J., V. Zlotnicki, and L.-L. Fu, Sea level variabilities in the Gulf Stream between Cape Hatteras and 50°W : A Geosat study, *J. Geophys. Res.*, 95, 17,957-17,964, 1990.
- Weinreb, M. P., G. Hamilton, S. Brown, and R. J. Koczor, Nonlinearity corrections in calibration of advanced very high resolution radiometer infrared channels, *J. Geophys. Res.*, 95, 7381-7388, 1990.
- Závody, A. M., C. T. Mutlow, and D. T. Llewellyn-Jones, A radiative transfer model for sea surface temperature retrieval for the along-track scanning radiometer, *J. Geophys. Res.*, 100, 937-952, 1995.

M. Allen, Space Science Department, Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, England, UK. (e-mail: allen@wobble.ag.rl.ac.uk)

T. Guymer and M. S. Jones, James Rennell Division, Southampton Oceanography Centre, European Way, Southampton, SO14 3ZH, England, UK. (e-mail: thg@soc.soton.ac.uk)

M. Saunders, Department of Space and Climate Physics, Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Dorking, Surrey RH5 6NT, England, UK. (e-mail: mas@mssl.ucl.ac.uk)

(Received December 23, 1996; revised June 17, 1997; accepted July 30, 1997.)

Fully reprocessed ERS-1 altimeter data from 1992 to 1995: Feasibility of the detection of long term sea level change

M. Anzenhofer and T. Gruber

GeoForschungsZentrum Potsdam, Division 1, German Processing and Archiving Facility Oberpfaffenhofen
Germany

Abstract. Global mean sea level observations are necessary to answer the urgent questions about climate changes and their impact on socio-economy. At GeoForschungsZentrum/German Processing and Archiving Facility ERS altimeter data is used to systematically generate geophysical products such as sea surface topography, high-resolution geoid and short- and long-period sea surface height models. On the basis of this experience, fully reprocessed ERS-1 altimeter data is used to generate a time series of monthly sea surface height models from April 1992 to April 1995. The reprocessing consists of improved satellite ephemerides, merging of Grenoble tidal model, and application of range corrections due to timing errors. With the new data set the TOPEX/POSEIDON prelaunch accuracy requirements are fulfilled. The 3-year time series is taken to estimate the rate of change of global mean sea level. A careful treatment of seasonal effects is considered. A masking of continents, sea ice, and suspect sea surface heights is chosen that is common for all sea surface height models. The obtained rate of change is compared to external results from tide gauge records and TOPEX/POSEIDON data. The relation of sea level changes and sea surface temperature variations is examined by means of global monthly sea surface temperature maps. Both global wind speed and wave height maps are investigated and correlated with sea surface heights and sea surface temperatures in order to find other indicators of climate variations. The obtained rate of changes of the various global maps is compared to an atmospheric CO₂ anomaly record, which is highly correlated to El Niño events. The relatively short period of 3 years, however, does not allow definite conclusions with respect to possible long-term climate changes.

1. Introduction

Urgent questions about possible climate changes may be answered by means of long-period and secular rates of change of the sea level. Trends in the sea level are considered as possible indicators of a global temperature rise caused by the increase of greenhouse gases. For the next century a rise between 1° C and 4° C can be expected. If this happens, then the sea level will rise 30-50 cm [Church *et al.*, 1991; Houghton and Woodwell, 1989], caused by the melting of glaciers, polar ice caps, and thermal expansion of the oceans [Church *et al.*, 1991; Meier, 1984]. However, results presented further below will show that the sea level rise has regional tendencies; that is, there are areas, that are much more affected than others. Within such a future scenario, the socioeconomic consequences are virtually unpredictable [Broecker, 1996].

Before the advent of satellite altimetry the sea level could only be observed through tide gauges. The latter has three major disadvantages. (1) As tide gauges are located on the shores of continent, and islands exclusively, their global distribution is necessarily uneven (see Plate 8). (2) Tide

gauge measurements reflect relative motions between the ground the instrument stands on and the sea surface, which, for example, could be shifted systematically due to postglacial rebound or tectonic uplift toward other stations. (3) A modification of the ocean circulation pattern may also influence the local sea level and hence tide gauge measurements.

With altimeter data it was the first time that the sea surface could be monitored in a continuous and repeated manner in a unique reference datum. However, because of unknown altimeter drifts and systematics of Seasat, Geos-3, and Geosat [Allan, 1983; Bonavito *et al.*, 1975; Cheney *et al.*, 1991; Horai, 1982] and missing overlapping data with other altimeter missions, only data of the European satellites ERS-1 and ERS-2 and the U.S./French satellite TOPEX/POSEIDON can be used to measure the sea level change with a sufficient accuracy. This means that the sea level change can be investigated from 1991 onward. It should, however, not be neglected that recent efforts have demonstrated that the use of tide gauges to link Geosat to TOPEX/POSEIDON and ERS-1 has proven useful and sea level estimates have been obtained. An uncertainty with tide gauges, however, remains and the estimates are difficult to prove. The problems of linking tide gauges and altimeter data, as described above, should not degrade the usefulness and importance of tide gauge measurements. They play an important role in calibrations and help to link ocean and

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02566.
0148-0227/98/97JC-02566\$09.00

land. The anticipated secular changes of the sea surface are so small that very careful data preparation and calibration are needed. Earlier investigations [Nerem, 1995] and the ERS-1 data upgrade as described below demonstrate that there are further and maybe unknown elements of uncertainty [Kopytoff, 1996].

The uncertainties of the data and the missing possibility of validating the sea level results (tide gauges are not suitable for open ocean) led to following consideration: If the sea level (or the sea surface) is an indicator of climate change, then other parameters of the global system "ocean-atmosphere" must also show variations within the same time period. For this reason, sea surface temperatures, wind speeds, and wave heights were analyzed in the same manner as sea surface heights. The changes obtained then were correlated and interpreted within the framework of the ocean-atmosphere system.

The paper begins with the data and data preparation description. Because of the enormous work package and its importance this part is described in detail. The second part is dedicated to the methodology of the sea level study. The third part shows the obtained results. In the fourth part a quality assessment of the results is performed. Finally, the obtained results are discussed in the context of an investigation of a long time series of sea surface temperatures.

2. Data

The following sections provide a detailed summary of all data used for the sea level study. Starting from the operational ERS-1 and TOPEX altimeter data, all additional data upgrades and correction models are described in detail. Some additional analyses for specific upgraded parameters are given in the subsequent chapters. Furthermore, sea surface temperatures, wind speed, and tide gauge data sets are described.

2.1. Altimeter Data

Precise altimeter products from the European Space Agency's ERS-1 mission and the U.S./French TOPEX/POSEIDON mission are used in this study as a baseline. For each mission, 3 years of data are selected, whose temporal separation is as small as possible and which take into account the different mission scenarios. In the following a detailed description for both data sets is given.

2.1.1. ERS-1 altimeter data. In July 1991 the European Space Agency launched the first of the two remote sensing satellites for studying the Earth's environment. The ERS-1 satellite (European Remote Sensing Satellite) carries a number of microwave instruments, which are designed to provide weather independent, repeated observations and images. Mission goals are, beside others, the improvement of the understanding of ocean-atmosphere interactions and the knowledge of ocean circulation and the transfer of energy. Both goals are linked very closely to a possible global climate change. Since its launch, ERS-1 has gone through different repeat cycles, which were optimized for particular mission objectives [European Space Agency, 1993]. Table 1 provides an overview of the different mission scenarios. During the commissioning phase, on-board instruments were calibrated with frequent flights over the calibration sites. Two ice phases, separated by 2 years, were flown to satisfy the ice community. The multidisciplinary phases were flown to satisfy as many applications as possible, with particular emphasis to synthetic aperture radar (SAR) ice and land mapping. Finally, two geodetic cycles, the second one shifted by half a track distance from the first one, were flown for the geodetic community. The main objective of these phases was the mapping of the geoid with an extremely dense spatial resolution. In April 1995 the second quasi-identical satellite ERS-2 was launched successfully. As an additional instrument, the Global Ozone Monitoring Experiment (GOME) is included. Also on ERS-2 the precise range and range rate equipment (PRARE) microwave tracking system, which failed on ERS-1, is working successfully and provides independent orbit information for altimetry [Flechtner et al., 1997]. After the launch of ERS-2 both ERS satellites flew for 1 year the 35-day repeat cycle in-line; that is, one satellite repeated the ground tracks of the other 1 with a temporal separation of one day. This so-called tandem scenario was especially designed for SAR interferometry [Duchossois and Martin, 1995]. On June 3 in 1996, ERS-1 went into hibernation, which means that the satellite and its instruments are monitored but that all instruments are switched into a stand-by mode, except for some SAR interferometric campaigns. ERS-2 will remain for its complete lifetime in the multidisciplinary 35-day repeat cycle. For both satellites, operational altimeter products such as waveforms, ocean products and sea surface height models are generated by the different Processing and Archiving Facilities (PAF) within the ERS ground segment. Geophysical data records (GDR), the so-called ocean product

Table 1. Overview of ERS-1 Mission Phases

Phase	Period	Repeat Cycle, days	Track Distance, km	
			Equator	$\phi=60^\circ$
Commissioning	July 29, 1991 - Dec. 10, 1991	3	930	460
Ice (1)	Dec. 20, 1991 - March 30, 1992	3	930	460
Multidisciplinary (1)	April 14, 1992 - Dec. 20, 1993	35	80	40
Ice (2)	Dec. 21, 1993 - April 10, 1994	3	930	460
Geodetic (1)	April 10, 1994 - Sept. 27, 1994	168	17	8
Geodetic (2)	Sept. 28, 1994 - March 20, 1995	168	17	8
Multidisciplinary (2)	March 21, 1995 - June 2, 1996	35	80	40
Hibernation	since June 3, 1996	35	80	40

records (OPR) for ERS, are operationally generated by the German (D-PAF) and the French (F-PAF) Processing and Archiving Facilities. While at D-PAF, the Quick-Look Ocean Products (QLOPR) are generated and distributed within 2 weeks after acquisition [Gruber *et al.*, 1993]; F-PAF is responsible for the precise and final ocean products (OPR). The data is generated with a delay of some months. It is the baseline for the sea level study described in this paper. At F-PAF, raw data, which are recorded on-board the satellite and dumped after each pass to the ground station, are transformed to physical quantities, which are then time tagged, located, and corrected for the instrumental effects. For this purpose a complete retracking of the altimeter signal is performed. In a second step, the precise orbit information and the environmental and geophysical corrections for each 1-Hz measurement are included before the products are distributed on CD-ROM [Centre ERS d'Archivage et de traitement, 1995]. For the sea level study, consistently processed ERS-1 data for the period between launch and August 1995 were available. Because of the limited spatial resolution of the 3-day repeat cycle period during the first 9 month of the ERS-1 mission, however, only data from April 15, 1992, until

March 20, 1995, were used. The other available data for the second multidisciplinary phase were not used because the data that was available at the time of the processing did not cover a complete year. The annual cycle causes periodic sea surface changes, whose power is much higher than that of possible secular sea level change. Thus incomplete years of altimeter data cause systematic artifacts in the analysis result. Furthermore, OPR data starting with the second multidisciplinary phase have been processed with a new software version. This could cause systematics in the range measurement and in other parameters, such as the microwave radiometer measurements, which are used for the wet tropospheric correction. In Table 2 a summary of all relevant parameters of the ERS-1 ocean product records used for this study is given. The analysis time period includes 3 months of 3-day repeat cycles, the so-called second ice phase. The implicit coarse spatial resolution of that period was overcome by a special gridding algorithm, which is described in section 3.1.

2.1.2. TOPEX altimeter data. In August 1992 the United States National Aeronautics and Space Administration (NASA) and the French Centre National d'Etudes Spatiale

Table 2. ERS-1 OPR Content Overview

Parameter	Description
Time, μ s	UTC time elapsed since January 1, 1990, 0 h of the 1-Hz altimeter measurement
Latitude, μ deg	geodetic latitude with respect to WGS84 ellipsoid calculated from orbit
Longitude, μ deg	geodetic longitude with respect to WGS84 ellipsoid calculated from orbit
Altimeter range, μ mm	1-Hz altimeter range corrected for the following instrumental effects: look-up table correction, effect of on-board filter on waveforms and retracking Doppler correction from altimeter range derivatives internal calibration correction from measurement of altimeter single-point target response ultrastable oscillator drift center of gravity correction, distance antenna to center of gravity
Range standard deviation, mm	standard deviation of 1-Hz mean from 20 Hz range measurements
Dry troposphere, mm	altimeter path delay correction due to dry troposphere [Saastamoinen, 1972]; atmospheric pressure is calculated by linear interpolation in time and space in global half-degree 6-hour pressure fields based on ECMWF model outputs
Wet troposphere, mm	altimeter path delay correction due to wet troposphere, two corrections are provided: correction computed from the microwave radiometer brightness temperatures and the altimeter's wind speed correction computed from ECMWF global half-degree model temperature and humidity fields; linear interpolation in time and space for each measurements is performed
Ionosphere, μ mm	ionospheric path delay correction, calculated from total electron content computed with the Bent model [Llewellyn and Bent, 1973]; the model input consists of predictions for the 12-month running average of the monthly sunspot numbers
Ocean tide, μ mm	elastic ocean tide correction calculated from 13 tide waves, first eight waves (M2 to Q1) calculated by Schwiderski [1980], and last five waves (L2 to μ 2) by Le Provost <i>et al.</i> [1991]; in the Mediterranean Sea, tides are calculated for the first eight waves by Cancell <i>et al.</i> [1993]
Tidal loading, μ mm	tidal loading correction calculated from the eight main tide waves (M2 to Q1) based on calculations of Francis and Mazzega [1990]
Solid Earth tides, mm	solid Earth tide correction taking into account the contributions of Sun and Moon and Wahr's [1995] radial correction but not the permanent deformation term
Satellite altitude, μ mm	satellite height above WGS84 reference ellipsoid, calculated by D-PAF based on PGM035 gravity field, which is an updated GRIM4-S1 satellite-only model [Schwintzer <i>et al.</i> , 1991]
Sea state bias, mm	sea state bias altimeter range correction (corresponds to electromagnetic bias)
Significant wave height, cm	1-Hz significant wave height corrected for instrumental effects
Wind speed, cm s^{-1}	1-Hz wind speed intensity calculated using a table, which relates the backscatter coefficient to the wind speed [Witter and Chelton, 1991]

* Parameter upgraded for sea level analysis (see section 2.2).

** The overview is from CERSAT [1995].

(CNES) launched the TOPEX/POSEIDON altimeter satellite. In contrast to the ERS missions, this satellite carries, except of the tracking instruments (Doppler orbitography and radiopositioning integrated by satellite (DORIS), Global Positioning System (GPS), laser retroreflector), only two altimeters and the microwave radiometer. The primary goal of the mission is the improvement of the knowledge of the global ocean circulation and, consequently, of the understanding of the ocean's role in global climate change [Fu *et al.*, 1994]. The main instrument is the TOPEX dual-frequency microwave altimeter, which can measure the ionospheric electron content and therefore is able to minimize range errors caused by this phenomenon. As an experimental altimeter the POSEIDON solid-state one-frequency instrument is on-board. It is meant to demonstrate new technologies for low-power and low-weight altimeter instruments to be used for future Earth observation missions. As both altimeters share the same antenna, only one of them can work at a given time. This means that approximately for every 10 cycles, 1 cycle is operated with the POSEIDON altimeter. The orbital configuration (an inclination of 66° , an altitude of about 1336 km, and a repeat cycle of 10 days (equatorial track distance of 316 km)) was mainly chosen in order to be convenient for oceanographic applications. Because of the higher altitude and additional tracking

systems the spacecraft positions can be determined better than for ERS. Merged TOPEX and POSEIDON geophysical data records are operationally generated and distributed with a delay of ~ 3 months. NASA is processing the TOPEX altimeter ranges and the GPS orbits, while CNES is processing the POSEIDON altimeter ranges and the DORIS orbits. For this study, only TOPEX altimeter data between September 25, 1992, which marks the begin of operations, and September 30, 1995, which completes the first 3 years of operations, from the Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO) data center in Toulouse were used. This period is as close as possible to the ERS-1 data period used for the study. Because of the constraints for ERS-1 and TOPEX data availability, no common 3-year period is available. All TOPEX data for these 3 years are processed with the same software and models, except the gravity model used for orbit determination. Until cycle 92 (March 24, 1995) the Joint Gravity Model (JGM)- 2 [Nerem *et al.*, 1994] was used. Since cycle 93 all orbits are based on the JGM- 3 gravity model [Tapley *et al.*, 1996], showing significant improvements with respect to the previous solution. Small inconsistencies in the product series may be a side effect of this transition. Table 3 summarizes the relevant parameters of Topex GDR used for the investigations.

Table 3. TOPEX GDR-M Content Overview

Parameter	Description
Time, μs	UTC time elapsed since January 1, 1958, 0 h of the 1-Hz altimeter measurement
Latitude, $^\circ \mu\text{deg}$	geodetic latitude with respect to TOPEX ellipsoid calculated from CNES orbit
Longitude, $^\circ \mu\text{deg}$	geodetic longitude with respect to TOPEX ellipsoid calculated from CNES orbit
Altimeter range, $^\circ \text{mm}$	1-Hz altimeter range corrected for following instrumental effects: pointing angle, sea state errors, satellite center of gravity motion, Doppler shift, oscillator drift, and others
Range standard deviation, mm	standard deviation of 1-Hz mean from 10-Hz range measurements
Dry troposphere, mm	altimeter path delay correction due to dry troposphere [Saastamoinen, 1972]; atmospheric pressure is calculated by linear interpolation in time and space in global half-degree 6-hour pressure fields based on ECMWF model outputs
Wet troposphere, mm	altimeter path delay correction due to wet troposphere; two corrections are provided: correction computed from the three frequency microwave radiometer brightness temperatures correction computed from French Meteorological Office models for the water vapor content
Ionosphere, $^\circ \text{mm}$	ionospheric path delay correction, calculated from two frequency altimeter measurements
Ocean tide, $^\circ \text{mm}$	elastic ocean tide correction calculated from 13 tide waves; first 8 waves (M2 to Q1) calculated by Schwiderski [1980], and last 5 waves (L2 to $\mu 2$) by Le Provost <i>et al.</i> [1990]. In the Mediterranean sea tides are calculated for the first 8 waves by Canceill <i>et al.</i> [1993] Geosat derived Cartwright and Ray [1990] model
Tidal loading, $^\circ \text{mm}$	tidal-loading correction calculated from both models of elastic ocean tides, using eight main diurnal and semidiurnal constituents
Solid Earth tides, mm	solid Earth tide correction [Cartwright and Taylor, 1971] including pole tide [Wahr, 1985]
Satellite altitude, $^\circ \text{mm}$	satellite height above reference ellipsoid, calculated by NASA based on JGM-2/3 gravity field models [Nerem <i>et al.</i> , 1994; Tapley <i>et al.</i> , 1996]
Electromagnetic bias, mm	electromagnetic bias altimeter range correction [Fu and Glazman, 1991]
Significant wave height, cm	1-Hz significant wave height corrected for instrumental effects
Wind speed, dm s^{-1}	1-Hz wind speed intensity calculated using a table, which relates the backscatter coefficient to the wind speed [Witter and Chelton, 1991]

* Parameter upgraded for sea level analysis (see section 2.2).

** The overview is from AVISO [1992].

2.2. Altimeter Data Upgrade

The original data sets provided by AVISO for TOPEX and F-PAF for ERS-1 were further upgraded to make them as consistent as possible and to apply additional corrections, which are not included in both data sets.

2.2.1. ERS-1 time tagging. For ERS-1 OPR data it is known from analysis of various groups that the data contains a time tagging error. To estimate the time bias, the method described by *Marsh and Williamson* [1982] was used, which relates the crossover differences and the range derivative differences between ascending and descending arcs with the time bias. Another method to estimate the timing error is to use altimeter data (crossovers or ranges) in the orbit determination. There the altimeter times are fixed to the very precise timing of the laser tracking data. After an extensive analysis of both methods over the 3-year analysis period a time bias of +1.5 ms in the OPR altimeter data was identified and added to the original times. In order to apply the time bias correctly, all orbits had to be interpolated again for the 3-year ERS-1 altimeter data. It should, however, be noted that the naming convention "time bias" is a little misleading. The so-called time bias contains also error sources that are absorbed in the computational procedures.

2.2.2. TOPEX time tagging. For TOPEX, no additional timing correction had to be applied. The original time parameters were used for the study.

2.2.3. ERS-1 orbit and orbit accuracy. The original OPR data set is based on D-PAF orbits revision 1. This means that all orbits are based on an upgraded GRIM4-S3 (PGM035) gravity field [*Schwintzer et al.*, 1993]. The average radial accuracy of these orbits is estimated to be some 15 cm. This accuracy is not sufficient for the sea level study, where a signal in the millimeter level is expected. For this reason the GFZ/D-PAF orbit determination group reprocessed all orbits for the period of interest with a more recent gravity model PGM055 and improved modeling. The PGM055 gravity field is an upgraded GRIM4-S4 gravity field, which contains additional ERS-1 laser and crossover observations for the 35-day and 168-day cycles [*Schwintzer et al.*, 1997]. The PGM055 model is given by a solution in spherical harmonics up to degree and order of 60 and some resonant terms with maximal degree of 69. It corresponds to an equatorial spatial resolution of 330 km at the Earth's surface. This dimension was chosen according to maximal orbit perturbations following from the orbit characteristics of the satellites which have contributions to the data processing. ERS-1 is the only satellite whose information was used to determine the resonant terms of degrees over 60. The gravitational harmonic coefficients have been estimated by a rigorous least squares adjustment simultaneously with parameters of ocean tidal terms, tracking station coordinates, and rates of motion. Therefore the gravity model is associated with a consistent ocean tide model and a terrestrial reference frame spanned by over 300 optical, laser, and Doppler tracking stations. Gravity field recovery from satellite tracking data is based on satellite orbit perturbation theory. Observations given by the tracking data and crossovers are used to reconstitute a satellite orbit. For any observation, partial derivatives are computed with respect to initial orbital elements, which are the parameters of the gravity model and ocean tides, station coordinates, rates of motion, drag and radiation pressure, and empirical nuisance

parameters. For ERS-1 data processing, an empirical acceleration once per revolution was taken into account. To minimize the residuals between observed and computed quantities, the system of normal equations is generated for any reconstituted orbit. The normal equation systems are accumulated according to a relative weighting scheme, which depends on the satellite and the data type. Constraints and stabilizing equations derived from Kaula's rule complete the accumulated system. The resulting system is solved for the parameters by inversion including the computation of the variance-covariance matrix and the variance of the unit of weights according to the Gauss-Markoff model. For the PGM055 model, data of 34 satellites have been processed. The orbit altitudes of the satellites used range from about 800 to 20,000 km. Among the 34 satellites, 15 have been tracked by cameras, 17 by laser, and 9 by microwave instrumentation (Doppler, GPS); some of these have been tracked by mixed optical/laser or laser/microwave systems at the same time. Some remarks to the ERS-1 tracking data used for the GRIM4-S4 and PGM055 models are of special interest. One 35-day cycle of 1993 with ERS-1 laser tracking and crossover data and, additionally, with TOPEX laser tracking data, single TOPEX and double ERS-1/TOPEX crossovers were processed for both models. Two more 35-day cycles of 1993 and the first 168-day cycle of 1994 with ERS-1 laser tracking and crossover data were processed for the PGM055. All together, a total sum of more than 2.7 million satellite tracking observations were compiled for the model. The total number of observations is composed of 3% optical data, which were acquired in the 1960s from 15 satellites; of 25% laser tracking data to 21 satellites starting in the 1970 s; of 70% microwave tracking data, mainly from TOPEX, Geosat and Nova-3; and finally only of 2% altimeter crossover observations. Of great importance for the quality assessment of a gravity field model is the complete calibrated variance-covariance matrix. To get realistic quality parameters, a single calibration factor of 5.5 has been estimated for the PGM055 model by comparisons to external data sets. Plates 1a and 1b show the predicted geographically correlated (mean) and anticorrelated (variable) radial orbit errors for ERS-1/2 orbits, computed by error propagation of the full variance-covariance matrix of the PGM055 solution. While the variable, or anticorrelated part, could be estimated to a large extent from single-mission altimeter crossover observations, the mean or correlated part could not be determined from this data type. It would be possible to determine it from double-mission crossovers if perfect knowledge of the orbit and the altimeter ranges can be assumed. But there are problems. For example take TOPEX data, where the orbits also contain a geographically correlated orbit error, which is indeed due to its orbit characteristics approximately by a factor of 3 smaller than for ERS. Furthermore, the altimeter ranges are not perfect due to calibration (e.g., oscillator drift) and other problems. This situation is insufficient for a precise determination of the geographically correlated error. Therefore the error propagation by the calibrated variance-covariance matrix is a welcome instrument to determine the geographically correlated orbit error. The sum of both radial orbit error components reflects the part of the orbital signal, which may go into the recovered sea surface heights from ERS-1/2 altimetric measurements. This error propagates completely into the estimation of the permanent part of the sea surface

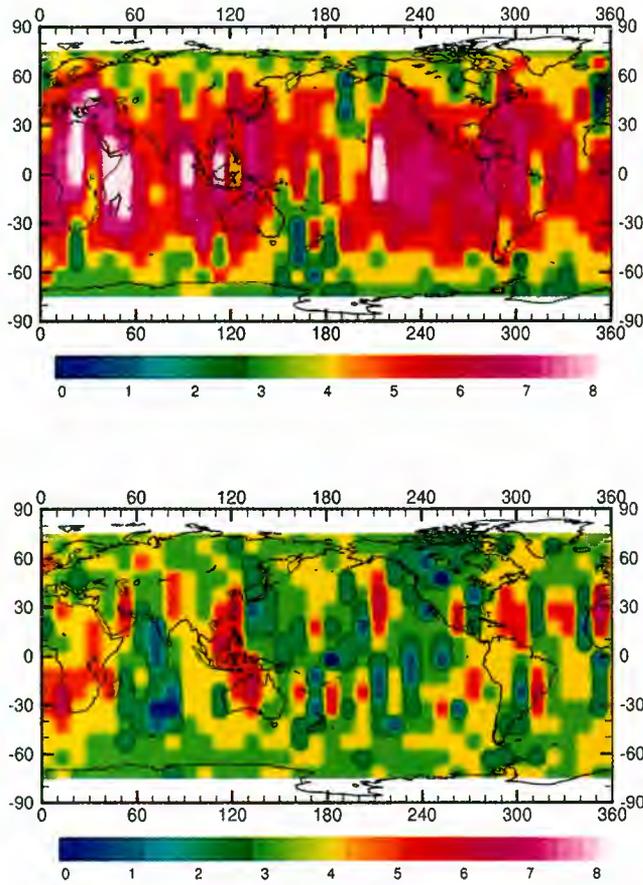


Plate 1. (1) ERS-1/2 mean radial orbit error (centimeters), (b) ERS-1/2 variable radial orbit error (centimeters).

topography, which represents the deviation of the mean sea surface from the geoid and, consequently, is the fundamental parameter for determination of the water mass transport across the oceans. The global geoid error for different models with respect to the degree of the spherical harmonic series is shown in Plate 2a. When interpreting this figure one has to have in mind that for each of these models an individual calibration factor for the variance-covariances was determined by the originator of the model. Therefore it is difficult to compare different models. Here two different types of models are shown. While satellite-only solutions like PGM055 and Earth Gravity Model 1996 (EGM96S) [Lemoine et al., 1997] are suffering from the limited availability of spacecrafts designed for gravity field determination and are showing a strong increasing error when increasing the degree of the series, combined models like EGM96 [Lemoine et al., 1997] and TEG3 are providing much more optimistic error estimates. But, for combined models terrestrial data (gravimetry, altimetry) were used, which could cause aliasing effects when using such models for oceanographic purposes. Furthermore, because of unknown datum of terrestrial data, the very long wavelengths could be influenced. To overcome such problems, only dedicated gravity field missions like GFZ-1, Challenging Microsatellite Payload (CHAMP) [Reigber et al., 1996], and others can help. The long wavelengths (up to degree and order of 70 or even more) should be completely determined from satellite observations, which are the only data type providing global and consistent distributed data sets. For

testing the long wavelengths of a gravity field, orbital fits for different satellites usually provide the most valuable information. Because PGM055 is currently used as the base for many ERS standard products, orbital tests are mainly focused on ERS-1/2. Orbital fits for the ERS satellites are a good parameter for the overall quality assessment because of their relatively low orbit, they are much more sensitive to gravity than, for example, the TOPEX satellite. In summary, using PGM055 for orbit determination, a mean fit of laser observations of about 7 cm and a crossover fit of about 8 cm for each arc are reached. Plate 2b compiles monthly crossover statistics (all possible crossovers within 1 month) over 3 years for the reprocessed precise orbits (revision 3) based on PGM055 and the completely independent orbits from Delft University of Technology (DUT), on the basis of the JGM-3 gravity field model. All together, both orbits show very similar behavior over the 3 years, with a mean crossover root mean square (RMS) of about 14 cm. While the DUT orbits show some smaller values for the 168-day cycle, the PGM055 orbits fit better for the 35-day repeat cycles. More interesting is the mean of crossover differences because it shows systematic differences between ascending and descending arcs. The PGM055 orbits are varying, except for two outliers, by some millimeters around the zero value,

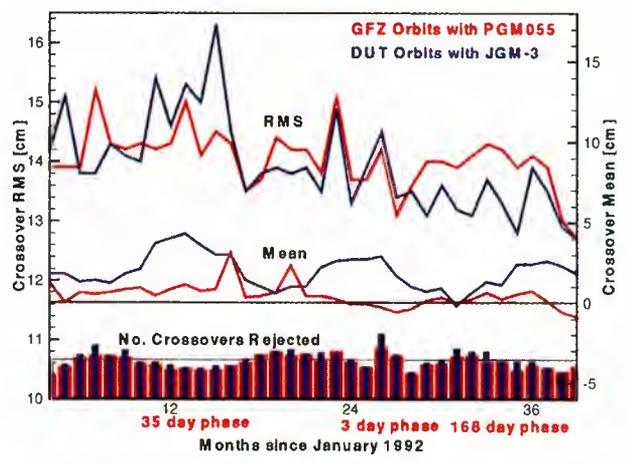
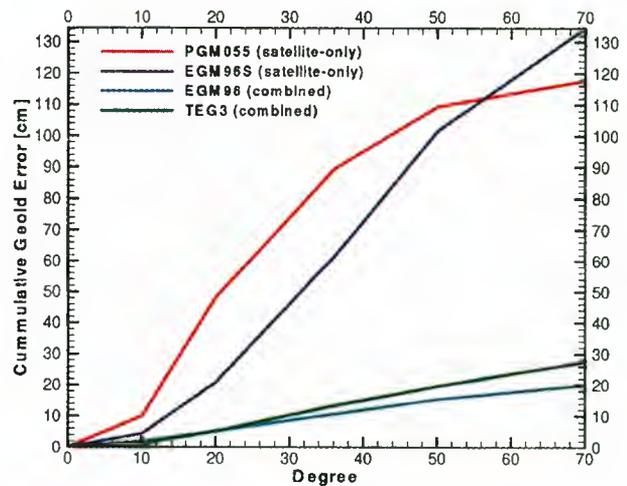


Plate 2. (a) Cumulative geoid error with respect to degree, (b) ERS-1 crossover statistic for different orbits.

while the DUT orbits show significant positive values of up to 4 cm. Obviously, there are some systematic effects in the DUT JGM-3 orbits, which cannot be completely addressed to the geographically anticorrelated orbit error, which is assumed to be smaller for the JGM-3 combined gravity model [Tapley *et al.*, 1996]. For the sea level investigation the obtained accuracy of the reprocessed ERS-1 orbits was sufficient. Thus no additional orbit error adjustment was applied. All orbits were merged into the OPR altimeter data set. In each record the satellite height, the geodetic latitude, and the geodetic longitude based on the World Geodetic System (WGS84) reference ellipsoid were computed and included into the record.

2.2.4. TOPEX orbit. The original AVISO products refer to a special TOPEX reference ellipsoid. To be consistent with ERS-1, satellite heights, geodetic latitude and geodetic longitude for each record were transformed into the WGS84 reference system. Because of the high orbit accuracy no further orbit error analysis is necessary for the TOPEX data.

2.2.5. ERS-1 altimeter range. Corresponding to the OPR documentation, altimeter ranges are corrected for all instrumental effects. During the cross-calibration of the ERS-2 altimeter against the ERS-1 altimeter [Anzenhofer *et al.*, 1996b] it was apparent that two additional corrections have to be applied to the altimeter ranges [Benveniste, 1996]. These corrections are necessary to compensate for two effects: first, the drift of the onboard ultrastable oscillator (USO drift), which causes an increasing range error and second, range jumps, which are produced by changes in the clock asymmetry caused by low temperatures during switch-off of the instrument after instrument anomalies. The latter error can be quantified from the single-pulse target response (SPTR) internal calibration data to correct the measured range. This so-called SPTR correction, which reaches values up to ± 2 cm, is operationally computed by the European Space Research Institute (ESRIN) and was applied to the 3 years of ERS-1 data. A critical element of the radar altimeter system for performing sea level studies is the USO drift, which causes systematic shifts in the final result. For the ERS satellites the USO frequency is measured once per week generally over the Kiruna ground station. The frequency is then directly recoverable on-ground from the real-time bit-synchronizer on the ground station and

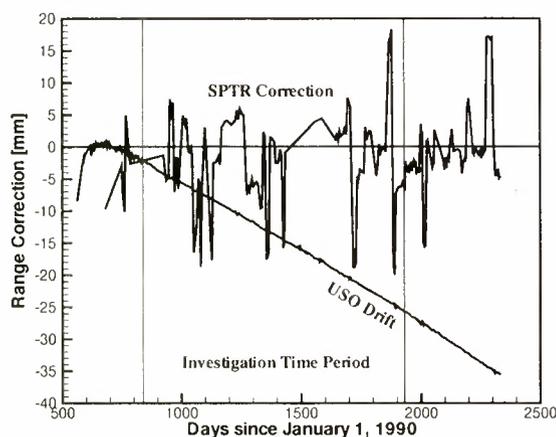


Figure 1. ERS-1 SPTR and USO drift altimeter range corrections to be applied to original (OPR) ranges.

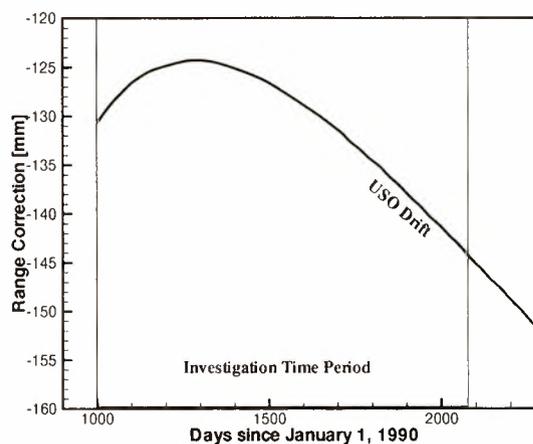


Figure 2. TOPEX USO drift range correction to be applied to Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO) TOPEX altimeter ranges.

comparable with an atomic frequency standard. This correction also is operationally estimated by ESRIN and is applied to the ERS-1 data. Both additional corrections are shown in Figure 1, where the curvatures display the SPTR and USO drift corrections for the complete ERS-1 mission. Both terms have to be subtracted from the original ranges.

2.2.6. TOPEX altimeter range. In July 1996, NASA headquarters issued a press release [Isbell and Hardin, 1996], which announced that the TOPEX sea level measurements have to be revised because of an altimeter data processing error found by O.-Z. Zanife, P. Escudier, and P. Vincent, which incorrectly adjusted for the natural drift of the onboard oscillator that controls the satellite clock. At NASA the Wallops Flight Facility [Hancock and Hayne, 1996] produced a new correction table taking into account this oscillator drift and, additionally, the constant offset between the TOPEX and the POSEIDON altimeters of 13 cm. Figure 2 shows the effect of the sum of both corrections for the TOPEX ranges, which have to be subtracted from the original ranges. This was done for the 3 years of TOPEX data used for the sea level study.

2.2.7. ERS-1 and TOPEX ionosphere. As described earlier, the ERS-1 altimeter is a single-frequency instrument, while the TOPEX instrument can derive the ionospheric path delay from its dual-frequency range measurements. To harmonize both data sets as much as possible, an additional ionospheric correction, computed from the International Reference Ionosphere Model from 1995 (IRI95) (D. Bilitza, personal communication, 1996) was included in both data sets. The International Reference Ionosphere is an international project sponsored by the Committee on Space Research (COSPAR) and the International Union of Radio Science (URSI). IRI95 is an empirical standard model of the ionosphere based on all available data sources, which mainly are the worldwide network of ionosondes, the ionospheric backscatter radars, the ISIS and Alouette topside sounders, and in-situ instruments on several satellites and rockets. IRI95 describes, beside other quantities, the electron density in the altitude range from about 50 to about 2000 km. The model inputs are the solar and ionospheric indices (IG12, Rz12) and the International Radio Consultative Committee (CCIR) and URSI coefficient files for the global

representation of F2 peak height and critical frequency. The ionospheric path delay correction is then calculated for each altimeter measurement from the total electron content, which is computed by integration of the electron density at different levels up to the satellite height. Using the IRI95 model instead of the 1990 solution (IRI90), better results have been reached in terms of crossover difference statistics and other tests. Additional analysis on the influence of this correction on the sea level study and comparisons to the TOPEX measurements and other models are given later.

2.2.8. ERS-1 and TOPEX ocean tide and loading. For elastic ocean tides and loading effects the recent FES95.1 model [Le Provost *et al.*, 1998] was additionally included to the standard data sets. This model is based on the purely hydrodynamic solution FES94.1, produced on the basis of the finite element model [Le Provost *et al.*, 1991]. This altimetry independent solution was produced to provide a completely independent tide model to the altimeter user community. Comparison of FES94.1 to empirical TOPEX/POSEIDON solutions showed large-scale errors in the hydrodynamical model of the order of up to 6 cm in the M2 wave. Therefore the FES94.1 model was updated by assimilating into the hydrodynamic model the earlier empirical TOPEX/POSEIDON CSR2.0 tidal solution [Shum *et al.*, 1997] from the University of Texas in Austin. The FES95.1 model represents a major improvement with respect to the models used in the official data sets. Also in the meantime the FES95.2 model from the Grenoble group is available, which further improves the ocean tide corrections for the sea surface height measurements. Because of the huge amount of data it was not possible to also include this new model into the complete data set. Therefore, in this study, the earlier FES95.1 model was used. Wave Heights and wind speeds are extracted from the OPR data [CERSAT, 1995].

2.3. Sea Surface Temperatures From the National Meteorological Center (NMC)

During the investigation time period, no ERS-1 along-track scanning radiometer (ATSR) data were available, although several requests have been initiated. Hence the very popular sea surface temperatures from NMC were extracted from the data server. In 1985 the World Meteorological Organization and the U.S. National Weather Service established a data center for global sea surface temperatures at the U.S. National Meteorological Center in Washington, D.C.. At the NMC, 3 different types of sea surface temperature analyses are performed: the processing of in-situ measurements (ships and buoys), the evaluation of advanced very high resolution radiometer (AVHRR) data of the National Oceanic and Atmospheric Administration (NOAA) satellites and the generation of mixed sea surface temperature models [Reynolds, 1987]. Both the in-situ measurements and satellite data run through a preprocessing chain in order to get validated sea surface temperatures. The basis of the mixed models is in-situ data. The satellite measurements then are used to complement areas with sparse data distribution. By means of a regression analysis the transformation of the satellite's skin to bulk temperatures is performed. By an optimum interpolation technique, weekly sea surface temperature grids are interpolated. The spatial resolution of the global grids is $1^{\circ} \times 1^{\circ}$ [Reynolds, 1987]. Monthly grid models are also available. Both weekly and

monthly sea surface temperature grids can be obtained through file transfer.

2.4. Wind Speeds From ERS-1 Scatterometer

Apart from wind speeds from ERS-1 altimetry, global wind speed grids derived from the ERS-1 scatterometer are received from Institut Français de Recherche et d'Exploitation de la Mer (IFREMER), France [Quilfen and Cavanie, 1991]. The scatterometer backscatter principle is Bragg dominated and increases with wind speed and is very different from radar altimeter backscatter.

The experiences with the Seasat wind scatterometer led to a special design of the corresponding instrument on ERS-1. Three antennas transmit radar pulses that are directed 45° forward, 45° backward, and sideways. The differences of the radar pulses and wind directions lead to different radar backscatter coefficients. A combination of the three different radar backscatter coefficients in turn can be used to estimate the wind field, which consists of wind speed and direction. For the processing of the ERS-1 scatterometer data to the wind field, an empirical model was implemented. This model sometimes leads to ambiguities of the wind direction, but these ambiguities can be identified and removed by comparisons to model wind fields [Quilfen and Cavanie, 1991].

2.5. Atmospheric Carbon Dioxide

A monthly time series of atmospheric carbon dioxide was provided by NOAA/Climate Monitoring and Diagnostics Laboratory (CMDL). The time period is from 1958 to 1996. The data is observed only at Mauna Loa, Hawaii [Keeling *et al.*, 1989; Thoning *et al.*, 1989]. The record is believed to be representative of the average CO_2 concentration throughout the atmosphere [Meyers and O'Brien, 1995].

2.6. Tide Gauges

The tide gauges used for the sea level study are obtained from files of the Permanent Service for Mean Sea Level (PSMSL) Public Access Directory. The PSMSL data set comprises monthly and annual means of sea level [Pugh, 1987] measured at tide gauge stations.

3. Strategy for Sea Level Investigation

All sea level investigations are performed by comparing monthly grid models to a long-term mean. The transition to relative measures make the representation of small values very easy and leads to numeric stable results (for example, the scale of absolute to relative sea surface heights is $100 / 0.2$ m). For the long-term mean a averaging of the 36-month grid models is performed.

3.1. Generation of Monthly Grids

Monthly solved grid models of sea surface heights, sea surface temperatures, wind speeds, and wave heights are the basis of this sea level investigation. This means that the data of 1 month are equally weighted and interpolated to one grid model. The spatial resolution is always $1^{\circ} \times 1^{\circ}$.

3.1.1. Sea surface height models. At the German Processing and Archiving Center for ERS (D-PAF) [Gruber

Table 4. Altimeter Corrections

Altimeter Correction	ERS-1 Sea Surface Heights	TOPEX Sea Surface Heights
Satellite height	GFZ/D-PAF orbits based on PGM055 gravity model	NASA orbits based on JGM-2 and JGM-3 gravity models
EM bias	F-PAF	<i>Fu and Glazman</i> [1991]
Inverse barometer effect	global rule: -9.948 (P - 1013.3)	global rule: -9.948 (P - 1013.3)
Ocean tides	FES95.1	FES95.1
Ocean loading	based on FES95.1	based on FES95.1
Earth tides	<i>Wahr</i> [1985]	<i>Cartwright and Taylor</i> [1971]
Ionosphere	<i>Bent et al.</i> [1975] (and IRI95)	TOPEX dual-frequency (and IRI95)
Dry troposphere	French Meteorological Office	ECMWF
Wet troposphere	radiometer	radiometer
Additional correction	time bias of 1.5 ms SPTR correction oscillator drift	oscillator drift

et al., 1993a] an operational processing chain has been implemented for the generation of sea surface height models [Anzenhofer *et al.*, 1992, 1994; Spöcker *et al.*, 1992]. The installed processing chain is capable of reading and processing different mission data. After passing through a preprocessing chain all altimeter data are compiled into a mission independent database which, in turn, is used to generate a catalogue of crossover events and crossover differences. The altimeter data are separated in half revolution arcs (ascending/descending). For the identification of these arcs an order file is generated, which holds the times and the longitudes of the equator crossing. This file handling enables an easy data extraction for any time period and region. The corrections as compiled in Table 4 then are applied to the altimeter data. In a following step the unevenly spaced altimeter data are interpolated to a geographical grid. The interpolation is performed locally at each grid node by means of local planes. For each grid node a functional model is defined by setting up n observation equations (n is the number of altimeter measurements for one local plane)

$$\begin{aligned} \hat{v}_i &= h_{grid\ node} - h_i \\ &= h_{grid\ node} - (\hat{a} + \hat{b} \Delta\lambda_i + \hat{c} \Delta\phi_i) \end{aligned} \quad (1)$$

The coefficients a , b , c define the local plane, which are estimated by a rigorous least squares process

$$\hat{x} = [D^T P D]^{-1} [D^T P L] \quad (2)$$

The vector \hat{x} holds the unknown (a , b , c). P defines the weighting matrix of each incoming measurement. L is the observation matrix holding the measurements h_i . D is the so-called design matrix as follows

$$D_{n,3} = \begin{pmatrix} 1 & \Delta\lambda_1 & \Delta\phi_1 \\ 1 & \Delta\lambda_2 & \Delta\phi_2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & \Delta\lambda_n & \Delta\phi_n \end{pmatrix} \quad (3)$$

The standard deviations of the unknown plane coefficients are given by:

$$K_x = \sigma_0^2 [D^T P D]^{-1} \quad \sigma_0^2 = \frac{\hat{v}^T P \hat{v}}{n - 3} \quad (4)$$

The altimeter data is scanned, for each observation the surrounding nodes are identified, and the corresponding normal equation system is accumulated. An influence circle around each observation restricts the number of corresponding grid nodes. The size of the circle depends on the orbit characteristics taking into account the density of the satellite track pattern. For the sea level investigation a smoothed sea surface is desired that minimizes residual radial errors in the data. Therefore a larger ($2^\circ 30'$) radius than the radius for high-resolution sea surface models like MSS95A [Anzenhofer *et al.*, 1996a] is chosen. The chosen weighting scheme preserves the high along-track gradients and smooths cross-track grid nodes to avoid trackiness in the sea surface. As mentioned above, altimeter data of the second ice phase with track separation of 930 km at the equator (3-day repeat cycles) are compared with data of the multidisciplinary phases with 80 km track separation at the equator (35-day repeat cycles). In order to improve the coarse spatial resolution of the 3-day repeat cycles, monthly grids were generated in a first run with defined grid nodes that were very close to the altimeter tracks. In a second run, data from TOPEX covering the same time period were fitted to the base models by removing large-scale and mesoscale systematics between ERS-1 and TOPEX, which were achieved by a fifth-degree polynomial representation of the differences at track transitions. This procedure preserved the short-scale oceanic features of TOPEX altimeter data, which improved the spatial resolution of the ERS-1 model. A systematic shift due to the introduction of TOPEX data, however, was avoided by the algorithm. Each sea surface height model obtained had to undergo a quality control with shading reliefs, comparison to external models, and outlier detection [Anzenhofer and Gruber, 1995].

3.1.2. Sea surface temperature models. Monthly gridded and $1^\circ \times 1^\circ$ spatially solved models of the sea surface temperature are provided by NMC, Washington D.C.. Thus no extra interpolation had to be performed [Reynolds, 1987; Reynolds and Smith, 1993].

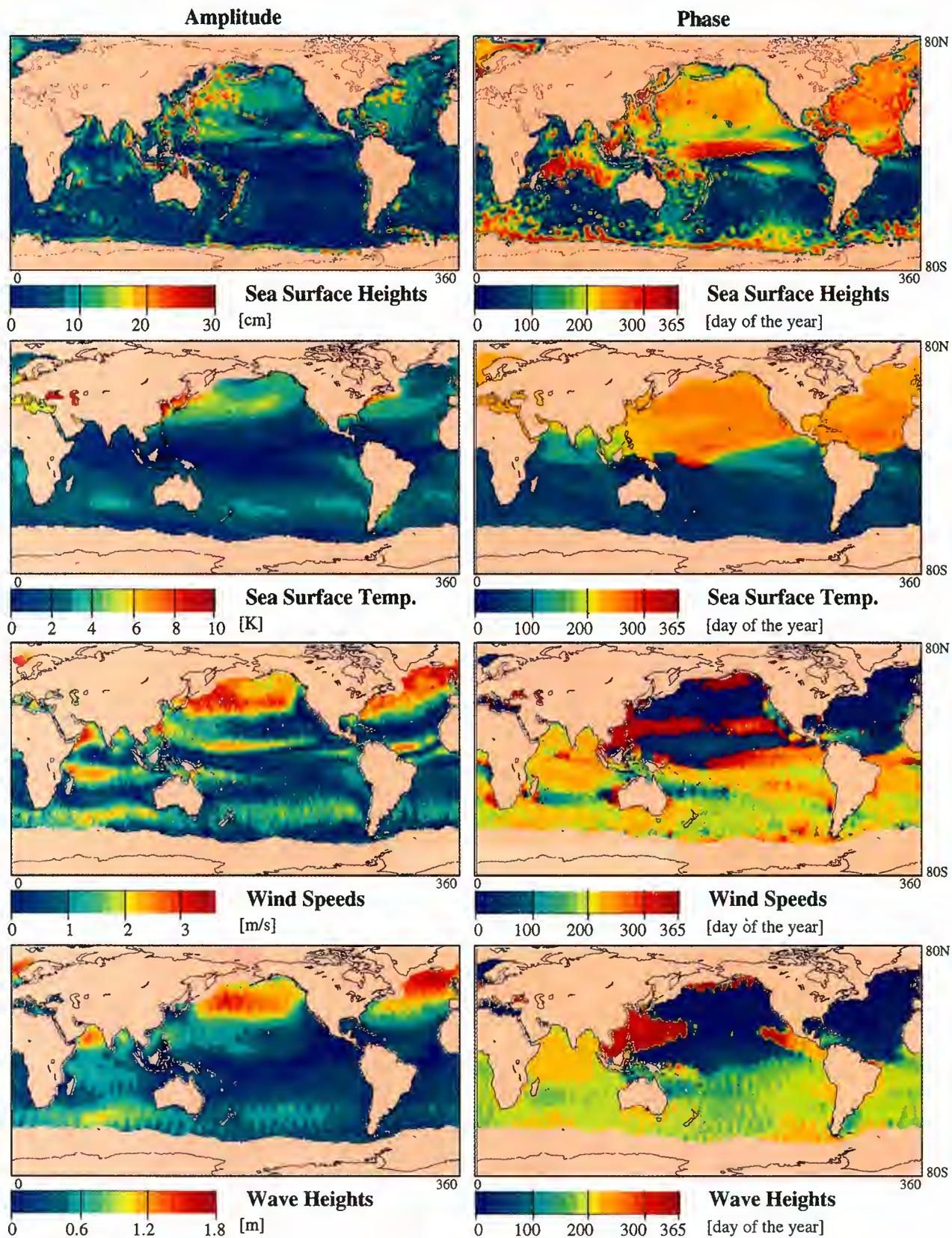


Plate 3. Annual oscillation of sea surface heights, sea surface temperatures, wind speed, and wave heights.

3.1.3. Wind speed models. For the wind speed derived from ERS-1 scatterometer, no grid interpolation was necessary. Monthly solved $1^{\circ} \times 1^{\circ}$ grids were provided by IFREMER [IFREMER, 1996].

3.1.4. Other grid models. For wind speed, wave height and range corrections (dry and wet troposphere, ionosphere, inverse barometer effect, electromagnetic (EM) -bias) from ERS-1 altimetry grid interpolations were necessary because measurements were available as 1 s along-track data. The algorithm chosen is very similar to the local planes as described above. Instead of local planes (coefficients a , b , c), a weighted mean is estimated (only coefficient a).

3.2. Methodology

The investigation of the sea level is embedded into the comparison with other parameters in the global system ocean-atmosphere, like sea surface temperature, wind speed, and wave height. The data come from different instruments and sources. Further on for the generation of grid models, different editing and quality criteria are used. This implies grid models with undefined nodes that are not common to other grid models. Investigations in the global system ocean-atmosphere have demonstrated [Fu and Cheney, 1995] that there are significant seasonal differences between the northern and southern hemisphere. This means that if grid models with different numbers of defined and undefined grid nodes are intercompared, then systematic shifts may happen. In order to avoid this error source a common masking for all grid models was performed. Therefore a grid model was generated that was successively filled with all undefined grid nodes found in the considered monthly grids. Then all the monthly grids were combined with the mask as shown in Table 5. This processing scheme produces consistent grid models and thus time series which can be intercompared.

For each parameter (sea surface height, sea surface temperature, wind speed, wave height, and range corrections), two analyses have been performed to extract variations within the investigation time period: (1) global rate of change and (2) local rates of change.

3.2.1. Global rate of change. The first step of this method is to subtract the monthly grid models from the long-term mean (3-year model). Then the differences in the grid nodes are averaged to one mean, taking into account that the area covered by each grid node decreases toward polar regions, which was done by a cosine (latitude) weighting. The obtained value now represents the mean deviation from a long-term state. The successive processing of all monthly grids leads to a time series which in turn is used to estimate the rate of change. This is done by a regression analysis. The

regression coefficient (and its standard deviation) defines the global rate of change for the above-mentioned parameters, like sea surface height, sea surface temperature, wind speed, wave height, and altimeter corrections.

3.2.2. Local rate of change. With the availability of the global rate of change value the question is raised whether the trend is evenly distributed over the oceans or if there are regional differences. Therefore a regression analysis was performed for each grid node of the monthly models. The estimated regression coefficients (and standard deviations) or local rates of change were then visualized by raster plots.

4. Results

The results of the sea level investigation are mainly focused on ERS-1 altimetry. The obtained results are embedded in the global system ocean-atmosphere. After the discussion of the results an internal quality control is performed that shows the influence of altimeter corrections on the sea level. By means of TOPEX altimeter data and tide gauges an external quality control is done.

4.1. Annual Oscillation

Because of the annual climate variations on Earth, the sea level exhibits a strong signal with a 1-year period [Fu and Cheney, 1995]. On the basis of monthly grid models of sea surface heights, sea surface temperatures, wind speeds, and wave heights, a harmonic analysis was performed. The results (amplitude and phase lag) can be seen in Plate 3.

4.1.1. Sea surface heights. In Plate 3 (left) the amplitude of the 1-year period is displayed. The scale of the amplitudes is between 0 and 30 cm. Plate 3 (right) picture shows the phase lag of the amplitude, which points to the day of the year when the amplitude has its maximum. Looking at the amplitudes some features can be identified.

1. The magnitudes of the amplitudes of the 1-year signal is different between the northern and southern hemisphere. The ratio is at least 2:1 [Stammer and Wunsch, 1994; Fu and Cheney, 1995].

2. The large amplitudes are in the surrounding areas of the western boundary currents like the Gulf Stream in the North Atlantic and the Kuroshio in the North Pacific.

3. The overall distribution of the amplitudes of the 1-year period in the vicinity of the western boundary currents is likely to be caused to a large extent by the thermal expansion due to their immense heat transport from the tropics to higher latitudes.

The phase lag picture confirms the differences between the northern and southern hemisphere. The maximal amplitudes on the northern hemisphere are between August and September, while on the southern hemisphere the maximum is reached during January and February. The separation line between the climate processes lies on the Intertropical Convergence Zone, which appears to be very sharp in the picture. The Intertropical Convergence Zone does not coincide with the equator line; for example, in the Atlantic, the zone is a few degrees north of the equator [Houghton, 1991].

4.1.2. Sea surface temperatures. The amplitudes of the annual oscillation of the sea surface temperatures exhibit the seasonal varying heat transport of the western boundary currents Gulf Stream and Kuroshio with ranges up to 6 K. A

Table 5. Masking of Monthly Grid Models

Mask	Grid	
	Monthly	Output
Defined	defined	defined
Undefined	defined	undefined
Defined	undefined	not possible
Undefined	undefined	undefined

high correlation to the corresponding sea surface pattern is obvious [Knudsen *et al.*, 1996]. The disproportion between amplitudes of the northern and southern hemisphere is also clearly visible. The phase lags of the sea surface temperatures correspond very well with the sea surface height phases. The separation line between the northern and southern hemisphere does not coincide with the equator. It is concentrated around the equatorial convergence zones.

4.1.3. Wind speeds. The amplitudes of the wind speeds mark the seasonal varying wind zones of the world, which are concentrated around the 40° latitude band (roaring forties). Only small annual oscillations can be identified in the trade zones. The phase lags show the anticorrelation between the atmospheric quantity wind speed and the oceanic quantities, sea surface heights and sea surface temperatures.

4.1.4. Wave heights. The amplitudes and phase lags of the wave heights demonstrate the relation between wind and waves (the wind system induces waves). Thus an anticorrelation to sea surface heights and temperatures is visible too.

4.2. Global Rate of Change

As described in the preceding sections the global rates of change of the sea surface, sea surface temperatures, wind speeds, and wave heights are estimated by comparisons to a long-term mean. Because of the analysis time period of 3 years, 36 monthly values per time series (time, deviation to the long-term mean) are available. To estimate the global trend, a regression analysis is performed. By means of a Fourier transform the amplitude spectrum of each time series is computed in order to show the significant frequencies of the parameters. The results are displayed in Figure 3, which is separated into two columns. Figure 3 (left) shows the deviations of each parameter to its long-term mean. The abscissas mark the times in years. The plotted line in each picture is the result of the regression analysis. Figure 3 (right) compiles the results of the frequency investigations. The ordinates mark the amplitudes. The abscissas point to the corresponding periods. The results are presented below.

4.2.1. Sea surface heights. The time series of sea surface heights contains a significant 1 year period. A half-year period is not evident. Positive deviations to the long-term mean are in August/September, and negative deviations are shifted by a half year. In the time series, three striking annual oscillations can be identified, including a positive trend during the investigation time period. The annual oscillation is induced by the inverse barometer effect (see section 4.4). The thin line in Figure 3 shows the deviations of the sea surface heights with respect to the mean, but without the annual oscillation. The filtering was done by a Fourier transformation. The regression coefficient (or global rate of change) is 2.2 ± 1.6 mmyr⁻¹. This means that the sea level rises globally. The standard deviation of the regression coefficient marks a significant estimation (1σ criterion). The standard deviation, however, is an artifact of the inverse barometer correction and is much less for the filtered curve (± 0.1 mm). However, it should be mentioned that this value does not contain the errors of altimeter data and methodology. It is very difficult to give an absolute estimate of the sea level change error bar. Let us assume that the inputs of the sea level investigation, the sea surface height models, are error-free (the real rms of the single-grid nodes

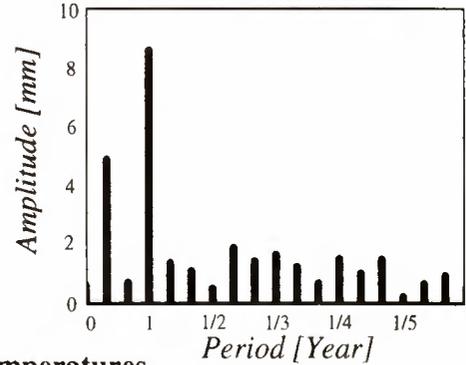
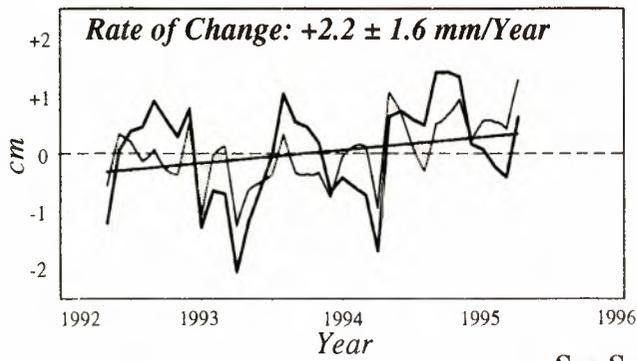
range between a few millimeters to centimeters and have a mean of about 2 cm). The differences between the monthly maps and the mean result in a root mean square of the variance of about ± 5 cm. The standard deviation of the mean of the differences is about ± 0.3 mm ($\sigma_{\text{mean}} = (\text{square root of variance})/(\text{square root of (number of defined grid nodes)})$). An error propagation of the regression coefficient standard deviation and standard deviations of the mean leads to a slight increase of the regression coefficient standard deviation. This means that, assuming error-free input data, the sea level change can be estimated significantly by the methodology, especially when the seasonal signal is filtered out. This does, however, not include errors and, more problematic, systematic shifts from corrections and models adopted to the altimeter data (see section 4.4).

4.2.2. Sea surface temperatures. A dominant annual oscillation can also be identified in the sea surface temperature time series. However, a significant half-year period is evident. The time series of sea surface heights and sea surface temperatures correspond well if the filtered curve of the sea surface height deviations is considered (Figure 3, thin line). Again, this shows the high correlation between sea surface heights and temperatures [Knudsen *et al.*, 1996]. The estimated global rate of change is 0.048 ± 0.032 Kyr⁻¹. This means that within the 3 years there exists a global heating of the oceans. In order to get a rough estimation for the question, if the temperature rise of 0.048 Kyr⁻¹ is responsible for the sea level rise of 2.2 mmyr⁻¹, the annual oscillations of sea surface heights and sea surface temperatures are compared (see section 4.1). In the Gulf stream region a sea surface height amplitude of 20 cm corresponds to a sea surface temperature of 5 K. If this would be the global relationship, then a temperature rise of 0.048 Kyr⁻¹ would lead to a sea level rise of 1.9 mmyr⁻¹, which is very close to the obtained 2.2 mmyr⁻¹. This rough estimate shows that a heating of the oceans could be responsible for the obtained sea level rise.

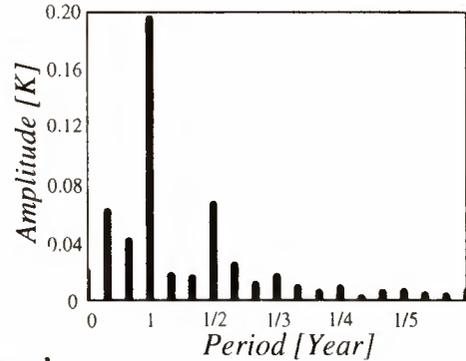
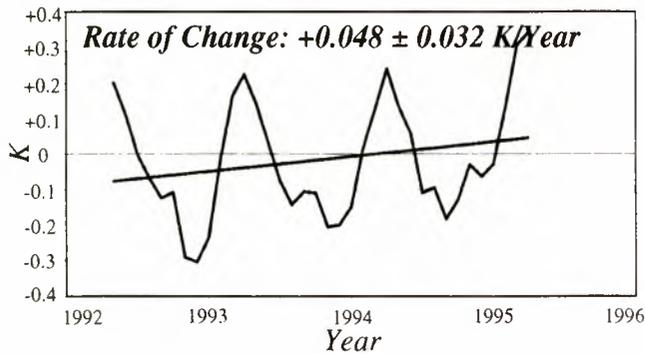
4.2.3. Wind speeds (scatterometer). In contrast to both the parameters described above, the wind speeds contain a perceptible half-year period. For the northern hemisphere maximal wind speeds are found in January, and they are found in July for the southern hemisphere. The lowest wind speeds occur around the dates of the change of the seasons (March 21 and September 23). The intercomparisons between the time series of sea surface temperatures and wind speeds show a significant anticorrelation. In contrast to sea surface heights and temperatures, the global trend is negative, -44 ± 32 mms⁻¹yr⁻¹. This means that the wind speeds have globally decreased between 1992 and 1995.

4.2.4. Wave heights. Wave heights must be highly correlated to wind speeds. The intercomparison of both time series confirm the statement that high wind speed deviations coincide with corresponding wave heights and vice versa. The amplitude spectrum also shows a half-year period, but it is not so dominant. There is, however, a remarkable difference between the time series of wind speed and wave heights; the global rate of change of the wave heights is positive, 6 ± 13 mmyr⁻¹, which implies a global rise of the wave heights. The negative wind speed trend should be accompanied by a decrease of the wave heights. It seems that there exists a small discrepancy between wind speed and wave heights, which must be clarified (see section 4.3).

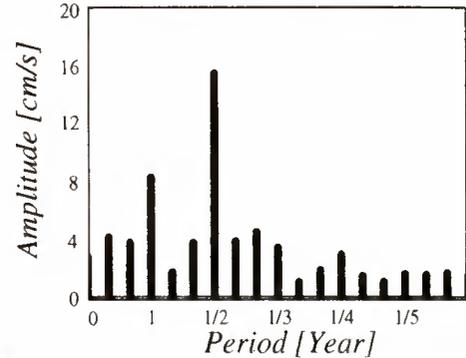
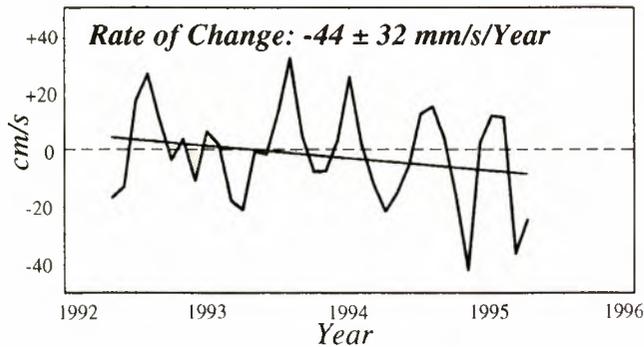
Sea Surface Heights



Sea Surface Temperatures



Wind Speed



Wave Heights

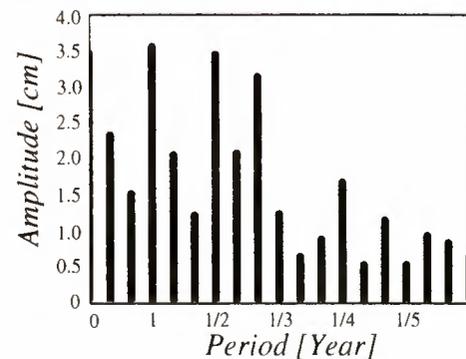
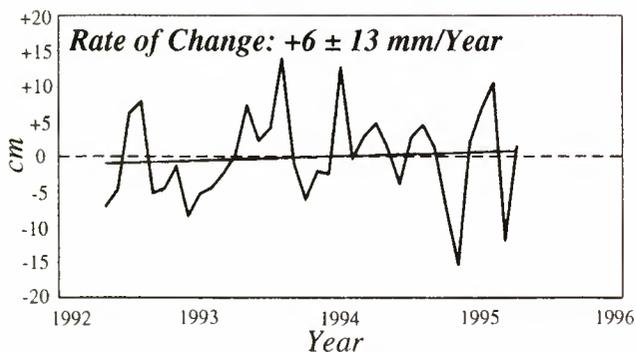


Figure 3. Global rates of change and corresponding amplitude spectra.

Looking at the global trends, it can be stated that there are global changes within the investigation time period from 1992 to 1995. But that is only true for global means of the quantities involved. The regional variations must now be investigated, which are shown in the following section.

4.3. Local Rates of Change

The global rates of change raised the question about the regional structures of long-term changes. The results are compiled in Plate 4.

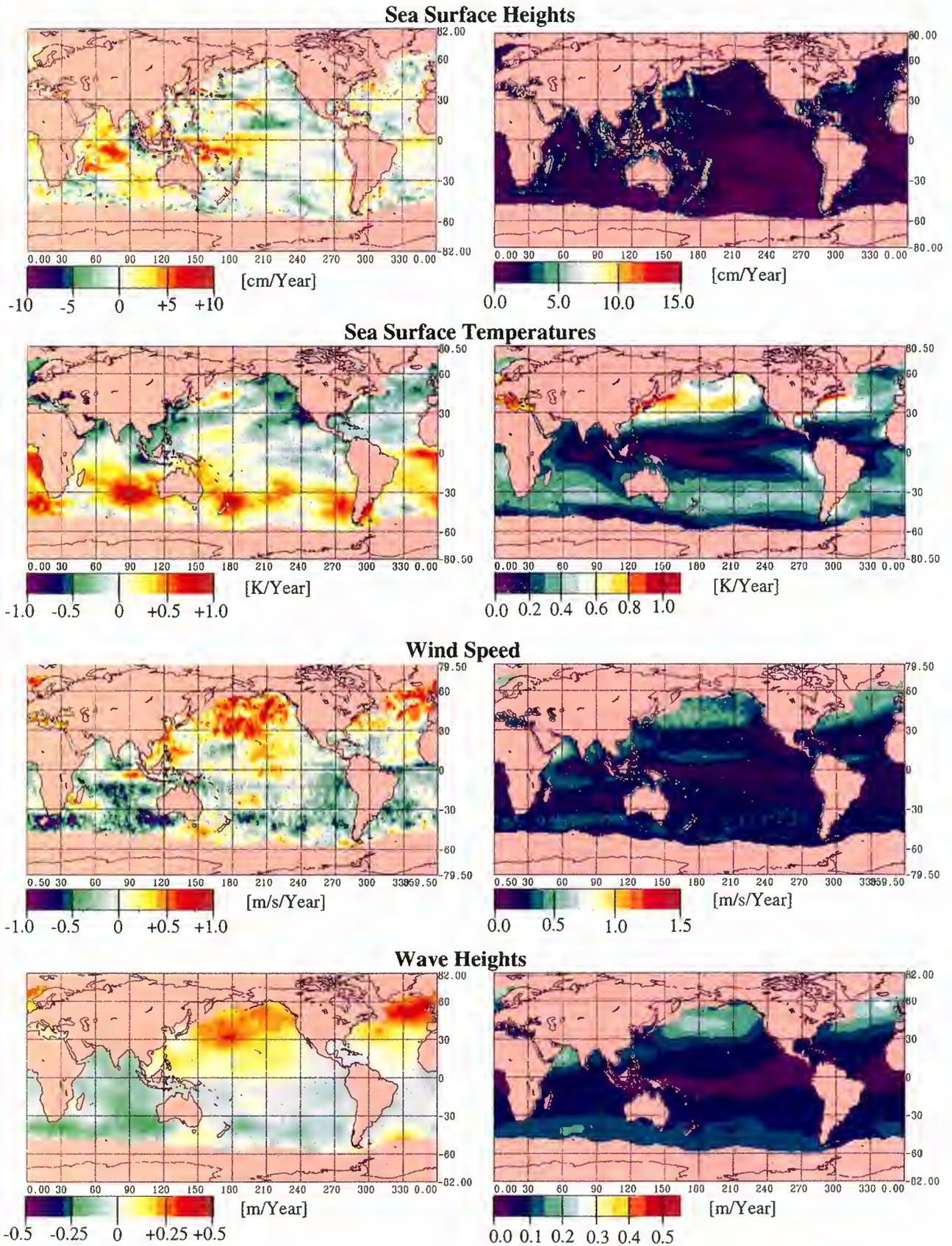


Plate 4. Local rates of change and corresponding standard deviations.

On Plate 4 (left) the local rates of change of sea surface heights (cm yr^{-1}), sea surface temperatures (K yr^{-1}), wind speeds ($\text{ms}^{-1} \text{yr}^{-1}$) and wave heights (m yr^{-1}) are displayed. Plate 4 (right) shows the corresponding standard deviations. The results are presented below:

4.3.1. Sea surface heights. The structures of the local rates of change of the sea surface clearly demonstrate that it is absolutely necessary to view not only global trends; the sea level trends are highly variable. Large regions with positive local rates of change (Plate 4, colors yellow to red), e.g., the tropics or within the western boundary currents, can be identified, but there are also regions with negative trends (colors blue to green), such as the western Pacific or in higher latitudes. The corresponding standard deviation plot marks significant estimates of the local trends of the sea level. Only in the highly variable western boundary current regions can higher standard deviations be recognized.

4.3.2. Sea surface temperatures. When viewing the local trends of the sea surface temperature a striking difference between the northern and southern hemisphere can be noticed. In the southern hemisphere an overall heating is evident, while in the northern hemisphere a cooling can be observed. There are, however, also regions in the northern hemisphere where heating is evident (Kuroshio, Gulf Stream). If the sea level change is dominated by the thermal expansion of the oceans, then a high correlation to the sea surface temperatures must be observable. This is true for the western boundary currents and for the equator region. But the striking difference of the sea surface temperature between the northern and southern hemisphere cannot be clarified by this relation. It is therefore a not too far-fetched question if errors in the data caused the systematics. Another indicator will be in how far the local trends of wind speeds and wave heights correspond to sea surface temperatures. If the apparent difference between the northern and southern hemisphere cannot be reproduced, then an error in the sea surface temperature data must be supposed. The corresponding standard deviations of the local rates of change do not indicate data errors. Only the highly variable regions of the oceans have enlarged standard deviations. The overall relation between the local trends and their standard deviation is better than the 1σ criterion.

4.3.3. Wind speeds (scatterometer). The local trends of wind speeds are highly (anti)correlated to sea surface temperatures. There is also a striking disproportion between the northern and southern hemisphere with an increase of wind speed in the northern hemisphere, and a decrease in the southern hemisphere respectively. In combination with sea surface temperatures this means that between 1992 and 1995 a climate variation had happened, which looks like it belongs to a period that is > 3 years. Maybe it is part of an interannual variation or it belongs to the detected decadal period [Latif *et al.*, 1995]. The corresponding standard deviations of the local trends reflect the zones of variable wind speed, e.g., around 40°N . However, the standard deviations point out a significant estimation of the local rates of changes.

4.3.4. Wave heights. The local trends of the wave heights are highly correlated to corresponding wind speed rates. In the very regions where the wind speed trends are high, the same is true for the wave height trends and vice versa. This is another indicator for the fact that within the 3

year investigation time period a long-term variation of the climate is evident. The standard deviations of the wave height trends also point out a significant estimation.

The local rates of change demonstrate that for each parameter, there are significant regional structures. If only global trends are considered, then these systematics remain hidden.

4.4. Quality of Sea Level Changes

The obtained values for the global and local rates of change for the sea level, sea surface temperatures, wind speeds and wave heights raise the question about the quality of the results. Because of the fact that ERS-1 altimeter data are used for the sea level study and because of the large amount of applied corrections and models for the altimeter data upgrade, an intense analysis of the altimeter data was performed and is described in the following paragraphs. This analysis constitutes an internal quality control. An external quality assessment is attempted with the analysis of TOPEX data and tide gauges. But first, the problem is discussed that the wind speed is only one among the parameters involved whose global change exhibits a negative trend.

4.4.1. Wind speed discrepancy. The identified discrepancy (negative trend) of the wind speeds within the analysis time period is investigated by (nadir) wind speeds of ERS-1 altimeter data and Comprehensive Ocean-Atmosphere Data Set (COADS) wind speeds [Lindau, 1995]. The data is processed as described in section 3. For the intercomparison of ERS-1 scatterometer, altimeter and COADS wind speeds one monthly grid is analyzed in addition. The results are compiled in Plate 5.

The global rates of change are presented Plate 5 (left). Plate 5 (top left) shows the already displayed result of ERS-1 scatterometer data with its negative trend of $-44 \pm 32 \text{ mms}^{-1} \text{yr}^{-1}$. Plate 5 (middle left) displays the result of the ERS-1 altimeter but with a positive trend of $54 \pm 33 \text{ mms}^{-1} \text{yr}^{-1}$. Comparing Plate 5 (top left) and Plate 5 (middle left) it can be seen that between both data there exists a difference of $90 \text{ mms}^{-1} \text{yr}^{-1}$. But the structures of the time series are very similar. Both data should lead to the same results, which means that one data set is not fully calibrated. It is not clear whether the error can be attributed to the ERS-1 scatterometer or the altimeter. As the wave heights slightly increased during 1992 to 1995, a wind speed increase is more likely. This would mean that the scatterometer data contains the error. However, just like wind speeds, wave heights are extracted from altimeter signal processing. If there is a processing error, then both data types can be systematically affected in the same direction. If this is true, then the scatterometer data could be closer to reality. Another indicator could confirm the last statement that sea surface temperatures and wind speed are highly anticorrelated (see section 4.1). Therefore the positive trend of the sea surface temperatures should lead to a negative trend of the wind speeds, which makes the scatterometer result more likely. Besides, it seems that the scatterometer data passed through a more extensive quality control [Bentamy *et al.*, 1997] than the altimeter wind speeds. From these contrary facts it is still not clear what data (scatterometer or altimeter wind speed) is correct. Hence an independent wind speed source, the so-called COADS wind speeds [Lindau, 1995], are taken as a

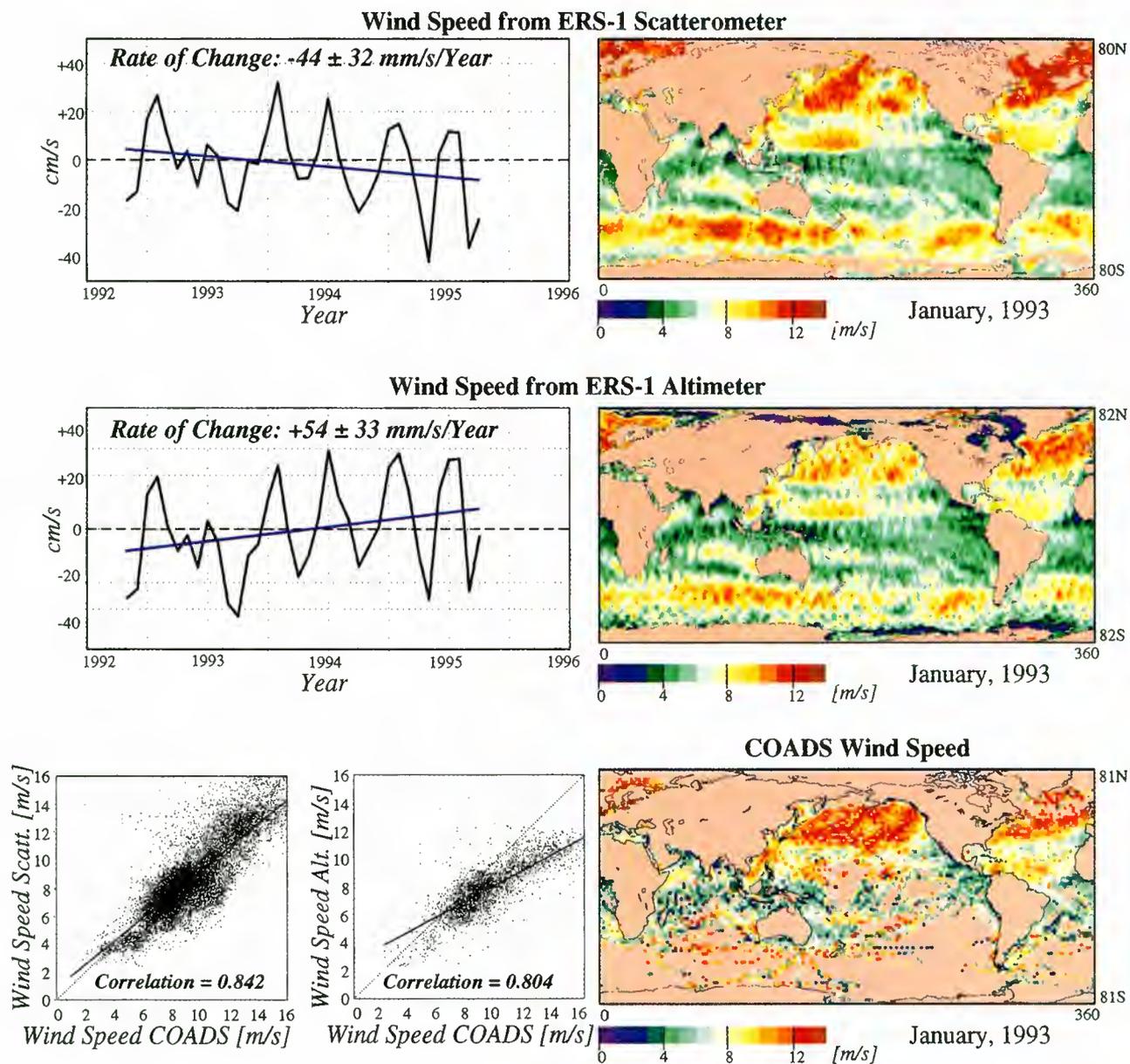


Plate 5. Wind speed analysis.

reference. The result in Plate 5 shows that the wind speed amplitudes match the scatterometer result much better than the altimeter result. It seems that the altimeter wind speeds are underestimated. The scatterplots of wind speeds in the Atlantic region and the corresponding correlation coefficients confirm that the scatterometer result is more likely, which means a global wind speed decrease of $-44 \text{ mms}^{-1}\text{yr}^{-1}$ in the analysis time period.

In order to get a feel for the differences of both wind speeds, monthly grid models are displayed on Plate 5 (right). For this purpose one arbitrary month was chosen (January 1993). The comparison of Plate 5 (top right) and Plate 5 (bottom right) demonstrates that the scatterometer data matches the COADS winds much better than the altimeter data, which makes the scatterometer result more reliable.

The accuracy of the sea level result directly depends on the adopted geophysical corrections. The influence of these corrections on the sea level result is described in following

paragraphs. The altimeter range corrections are processed to global and local rates of change as shown in section 3. The results are compiled in Plate 6. Before going into detail a cross correlation between ERS-1 and TOPEX altimeter range corrections must be described. Within an arbitrary time period of 15 days (January 1-15, 1993) all possible crossovers between ERS-1 and TOPEX altimeter data (ascending-ascending, descending-descending, ascending-descending, descending-ascending) have been computed. The maximum time difference at the crossover point was restricted to 6 hours in order to avoid large atmospheric and oceanic variability. A smaller time difference would have been preferable, but this would have caused unevenly distributed crossover events (only in higher latitudes). Table 6 compiles the results.

The crossover result demonstrates that the corrections match each other very well in the mean. The larger value for the ionospheric correction and the higher rms values (square

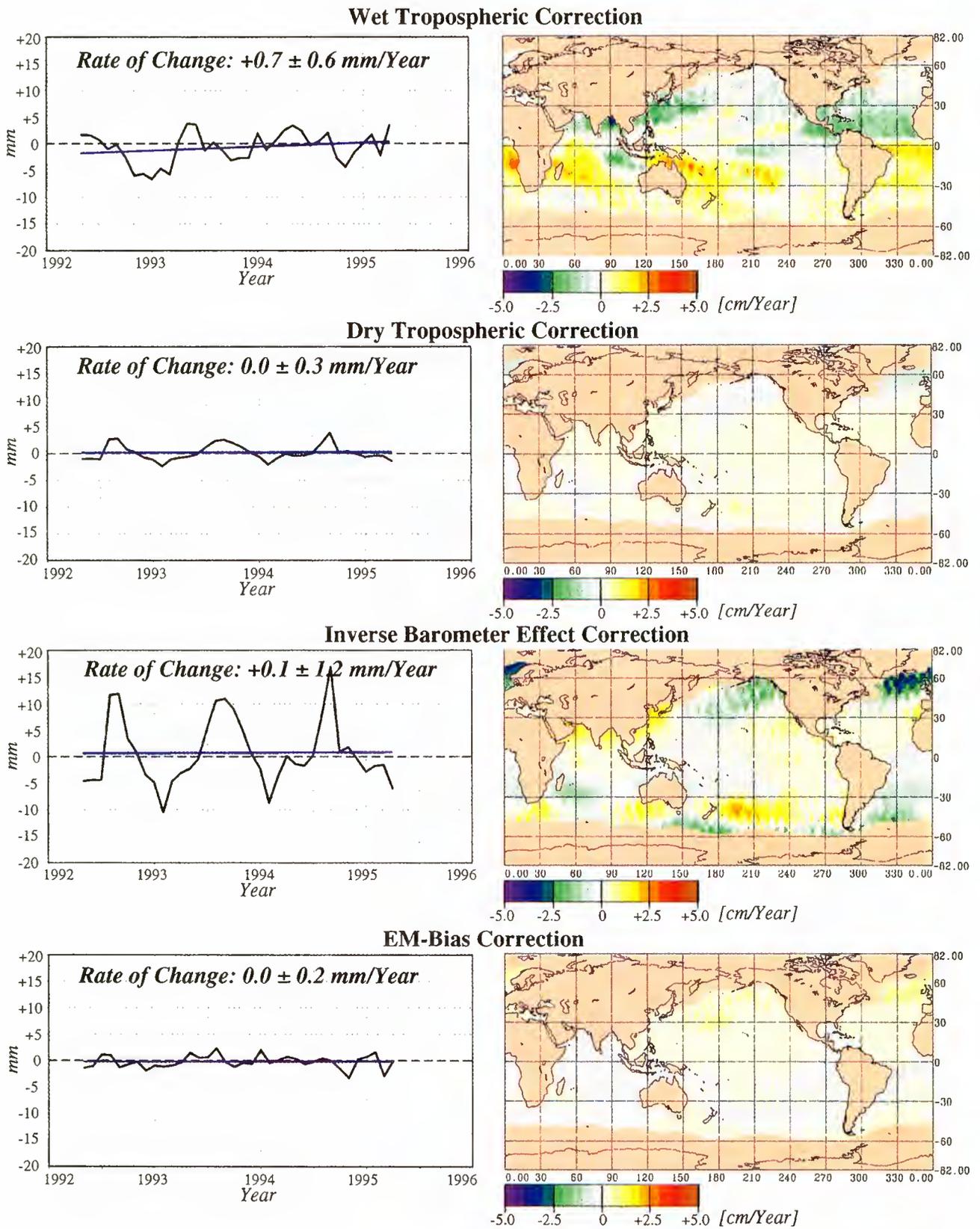


Plate 6. Influence of ERS-1 altimeter corrections on sea level estimate.

Table 6. Crossover Results Between ERS-1 and TOPEX

Correction	Mean, mm	Root-mean-square, mm
Wet troposphere	-1	25
Dry troposphere	0	6
Ionosphere	9	16
EM bias	-1	19
Inverse barometer effect	-2	42

root of the variance) is caused by the 6-hour time difference at the crossover event, for example, the ionosphere has a striking one day oscillation.

4.4.2. Wet tropospheric range correction. The wet tropospheric corrections are obtained by the onboard radiometer. A glance at Plate 6 shows that the wet tropospheric correction has global deviations from the mean of about 3 mm. A dominant annual oscillation caused by the incoming solar radiation is evident. The global rate of change amounts to 0.7 mmyr^{-1} for the sea level. The corresponding local rates of change again show the disproportion between the northern and southern hemisphere, which can be explained by the changes of the sea surface temperatures (see Plate 4). Exactly in those regions where the ocean surface is warmed, an increase of humidity, which is the controlling factor of the wet tropospheric correction, can be recognized and vice versa. In retrospect the wet tropospheric changes confirm the results of the local rates of change of sea surface temperatures, wind speeds, and wave heights as displayed in Plate 4. It seems that these parameters are affected by a long-term variation in the global system ocean-atmosphere that can possibly be attributed to a decadal variability [Latif *et al.*, 1995]. The question, whether or not the wet tropospheric correction causes a systematic error for the estimation of the sea level, depends on the accuracy and long-term stability of the ERS-1 radiometer. However, the structures of the local rates of change and the small deviations to the corresponding TOPEX correction (table 6) show that the influence of the wet tropospheric correction should be negligible.

4.4.3. Dry tropospheric range correction. The dry tropospheric correction is provided by the French Meteorological Office [CERSAT, 1995]. The global and local rates of change of the dry tropospheric correction are zero (0.0 mmyr^{-1}). The crossover statistic with TOPEX data confirms that the dry tropospheric correction is very accurate and without any influence on the sea level estimate.

4.4.4. Inverse barometer effect correction. The inverse barometer effect is a critical measure for the sea level estimation. It accounts for the pressure that the atmosphere exerts upon the sea surface. Its value is about 1 cm for 1-mbar pressure change. There is a global application of the rule, although analyses have shown that spatial and temporal deviations exist [Fu and Pihos, 1994; Trupin and Wahr, 1990]. The controlling factor of the inverse barometer effect is the atmospheric pressure. Mostly, this quantity is obtained from model runs, which in turn are mainly influenced by pressure measurements on continental surfaces. A look at the global rate of change (0.1 mmyr^{-1}) points to a vanishing global effect on the sea level. However, the time series of deviations to the long-term mean shows that the seasonal

variations are pretty large and lead to the seasonal cycle of the sea surface height deviations as mentioned above (see section 4.2). The local rates of change indicate only a small influence on the sea level estimate because the effects emerge perceptibly only in higher latitudes.

It is an ongoing controversy in the scientific community whether it is a reasonable thing to apply the inverse barometer correction to the sea surface or not. A desirable effect is the reduction of ocean variability when the correction is applied. A negative effect is the uncertainty of data quality and temporal deviations.

For the sea level study at GFZ/D-PAF, sea surface height models with and without the inverse barometer effect applied were generated. If the corrections were not applied, then the following effects happened to the sea level: (1) the sea level will rise 3.0 mmyr^{-1} (instead of 2.2 mmyr^{-1}), (2) the structures of the local rates of change differ only by a constant shift and (3) the time series of the deviations from the long-term mean is more noisy.

The sea level analysis was performed with sea surface heights, which were corrected for the inverse barometer effect. This implies possible errors of the atmospheric pressure data, but the ocean variability is minimized by the correction.

4.4.5. EM-bias range correction. The global and local rates of change of the EM-bias correction (0.0 mmyr^{-1}) indicate that there is not any influence on the sea level estimate. The pattern of the local rates of change correspond very well to the corresponding wave heights, which is expected because of its relationship.

The following paragraphs discuss the influence of the ionospheric correction and tides. Both are extracted from models as direct measurements are not available.

4.4.6. Ionospheric correction. The ionospheric correction for ERS-1 altimeter data is derived from the Bent model [Bent *et al.*, 1972]. As described in section 2, the correction was also available from the IRI95 model. The necessity to use model data is a significant disadvantage of ERS-1 altimetry in contrast to TOPEX/POSEIDON, where the correction can be obtained from the dual-frequency altimeter instrument. In order to investigate the influence of the ionospheric correction for the sea level estimate, global and local rates of change are generated as described in section 3. The global trend of $-9.1 \pm 1.8 \text{ mmyr}^{-1}$ clearly demonstrates that this is a very critical quantity (for comparison, global rate of change of wet tropospheric correction: 0.7 mmyr^{-1}). The local rates of change and thus the effect on the sea level look like a constant drift with maximum values in the vicinity of the geomagnetic equator. The sea level trend of 2.2 mmyr^{-1} was obtained by using the Bent model. The sea level study was also performed using the IRI95 model. With the IRI95 model the global rate of change for the sea surface is $-0.9 \pm 2.3 \text{ mmyr}^{-1}$; in contrast to the Bent model, we are looking at a decrease rather than an increase. These trends with an opposite sign show that long-term systematics between the ionospheric models exist.

In order to get an idea about the quality of the ionospheric models, two tests have been performed. The principle of the quality assessment is to assume that the ionospheric correction derived from TOPEX dual-frequency measurements are the ones that are true. For the quality assessment three different ionospheric time series were produced from TOPEX dual-frequency measurements, the

Bent [Bent *et al.*, 1972] model, and the IRI95 model. The time series consist of global and monthly solved ionospheric grid models from October 1992 until March 1995. First, the grid models from TOPEX dual-frequency measurements are subtracted from the corresponding Bent grid models. Second, the differences in the grid nodes are simply averaged to one mean. The successive processing of all monthly grids leads to a time series, which in turn is used to estimate the mean and a trend value. The same procedure is performed by using the IRI95 grid models. The result is compiled in Plate 7. The plate indicates that the Bent model reproduces the ionospheric conditions, as measured by the TOPEX dual-frequency altimeter, very well. Only a small systematic of -0.8 mmyr^{-1} is apparent. The comparison of IRI95 and TOPEX grids, however, indicates a large trend of -3.2 mmyr^{-1} . The curvature itself can be correlated to the solar cycle (solar sunspot numbers are the controlling factors of the ionospheric models) with large differences within high solar activity (1992 and 1993) and minor differences in 1995.

This results supports the use of the Bent model for the sea level investigation. However, it shows that the ionosphere is the most critical measure of the whole study, which makes the obtained sea level rise questionable. In order to further improve the ERS altimeter data, one can think of merging GPS-derived daily ionospheric models into the data. This was tested for the first 3 month of 1995 with the so-called Center for Orbit Determination in Europe (CODE) global ionosphere maps of the University of Berne [Schaer *et al.*, 1996], which were merged into the TOPEX data. As described above, monthly grids were generated and compared to the corresponding TOPEX dual-frequency ionospheric grids. The result is tabulated in Table 7. When looking at Table 7, it is still not clear which model best matches the TOPEX dual-frequency ionospheric corrections. Further analyses with longer time series are necessary to answer the open questions.

4.4.7. Ocean tides. For the ocean tides correction the hydrodynamic model FES95.1 was used [Le Provost *et al.*, 1991]. The model only causes periodic variations of the sea surface, which do not affect the sea level estimate. A problem arises with the 18.6-year lunar nodal tide. Investigations have shown that this tide affects the sea level by 0.4 mmyr^{-1} within the investigation time period [Nerem, 1995].

The analysis of the ocean tides closes the internal quality assessment of the sea level estimate. It has been demonstrated that the altimeter corrections affect the sea level estimate globally and locally. Especially, the ionospheric correction is due to its extraction from a model, the main source of uncertainty. Furthermore, the long-term behavior of all the corrections is unknown, and the errors bars of the corrections often exceed the magnitudes of the estimates themselves.

4.4.8. Comparison between ERS-1, TOPEX, and tide gauges data. In order to find clues about the quality of the ERS-1 sea level result two comparisons have been performed: (1) analysis of sea level change from TOPEX data and (2) analysis of local sea level changes from tide gauges.

The TOPEX data ran through the same processing chain as ERS-1 data to get estimates of the global and local rates of change of the sea surface. For the tide gauges only local rates of change have been estimated because of the uneven

Table 7. Comparison of Ionospheric Corrections

Period	GIM-T/P, mm	Bent-T/P, mm	IR195-T/P, mm
January 1995	8.0	4.4	-0.1
February 1995	6.2	8.6	6.2
March 1995	4.4	5.5	2.5

* T/P means TOPEX/POSEIDON ionospheric data.

distribution of the tide stations. The results are compiled in Plate 8.

The time series of the deviations from the long-term mean shows for TOPEX data a global rate of change of $2.3 \pm 1.9 \text{ mmyr}^{-1}$. It should be noted that the time series consists only of 35 monthly grids. During the investigations it became apparent that the data of October 1992 were contaminated by errors. Thus the data was excluded from the investigations. The global rate of the sea level conforms very well with the one obtained from ERS-1 data. The local rates of change are also very similar to ERS-1 trends shown before. The locations of the sea level rise and fall are found at the same geographical positions. Merely the amplitudes of the sea level rise (fall) differ slightly, e.g., in the Indian ocean. Because of the transition from JGM-2 to JGM-3 for the orbit computations the TOPEX-derived result, however, must be contaminated by the differences between both gravity fields.

The rates of change derived from tide gauges are well correlated with the corresponding estimates from ERS-1 data for their locations. The scatterplot of sea level trends from tide gauges and sea surface heights exhibits a correlation coefficient of 0.49. On one hand, this value is caused by the fact that the sea surface near coastlines can only be determined with reduced accuracy because of mismodeling of altimeter corrections, like ocean tides. On the other hand, tide gauge measurements can be affected by tectonic uplift effects. The gridded tide gauge trends far from the continents (like in the Pacific), however, correlate well with the corresponding altimeter trends (see section 4.3). As tide gauges do only give isolated local sea level estimates, a comparison on a global scale is not possible.

Taking everything into account as presented so far, it seems that within the investigation time period there is a global sea level rise of about 2 mmyr^{-1} , which coincides well with former estimates [Fu and Cheney, 1995; Trupin and Wahr, 1990].

5. Discussion

In the meantime several sea level investigations from different groups exist [Nerem, 1995; Shum *et al.*, 1997]. Together with the analysis at hand, the following different statements can be drawn: (1) the sea level trend is very small, which means that a very careful data handling and preparation is necessary [Isbell and Hardin, 1996; Kopytoff, 1996]; (2) input data must be well calibrated or correction tables must be available [Benveniste, 1996a; Hancock and Hayne, 1996]; (3) there is a necessity to continuously monitor instruments and corrections in terms of drift, not only for a distinct time, but throughout the whole mission duration in order to detect non-geophysical drifts; (4)

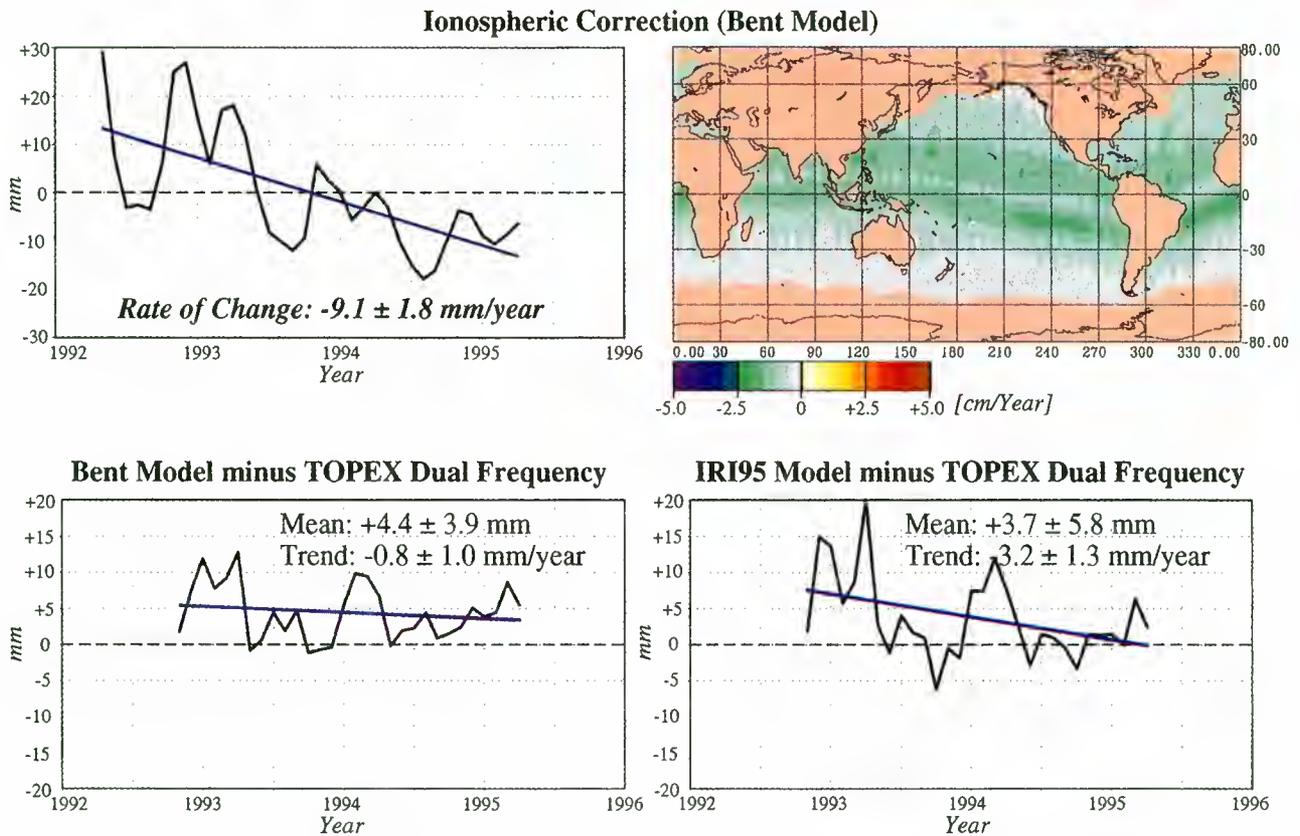


Plate 7. Comparison of ionospheric corrections for TOPEX.

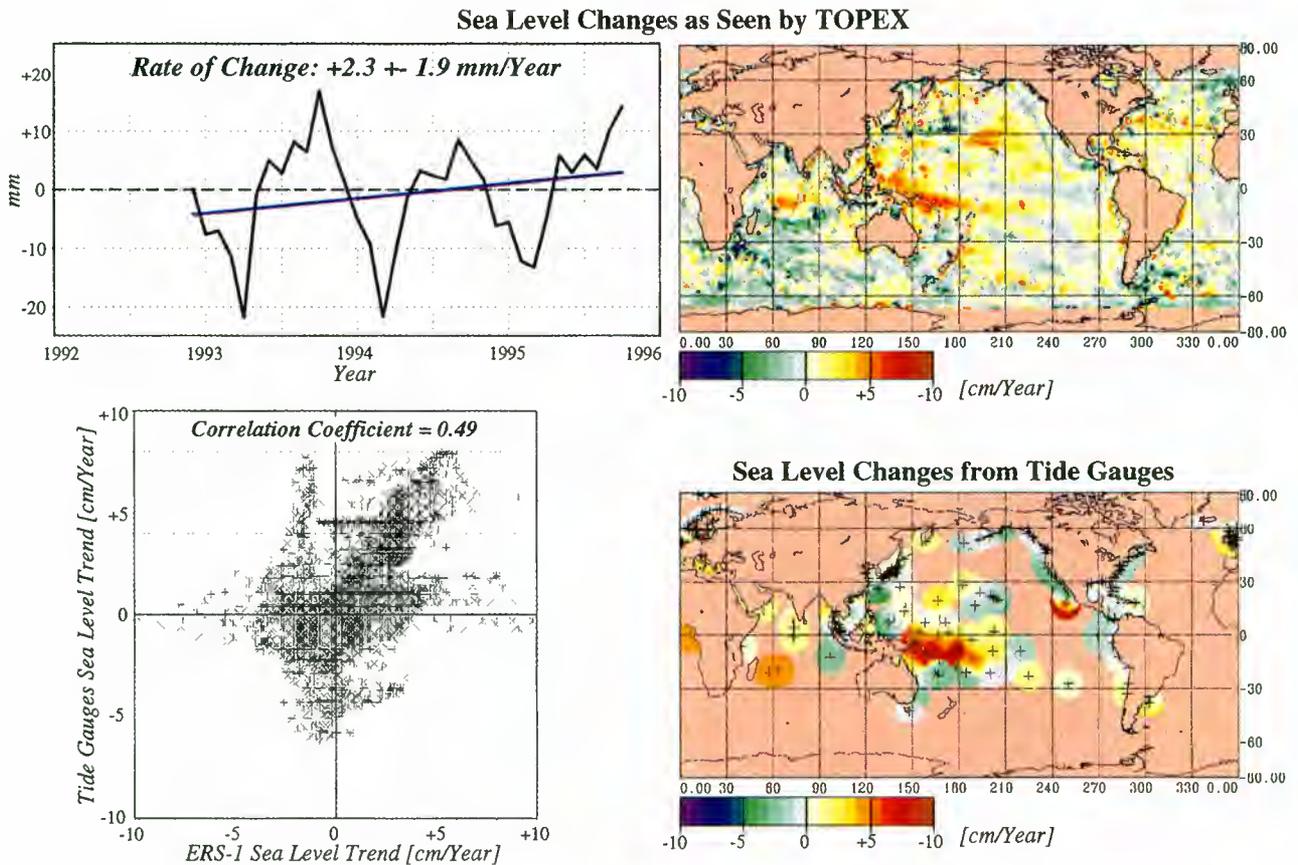


Plate 8. Sea level changes from TOPEX altimetry and tide gauges.

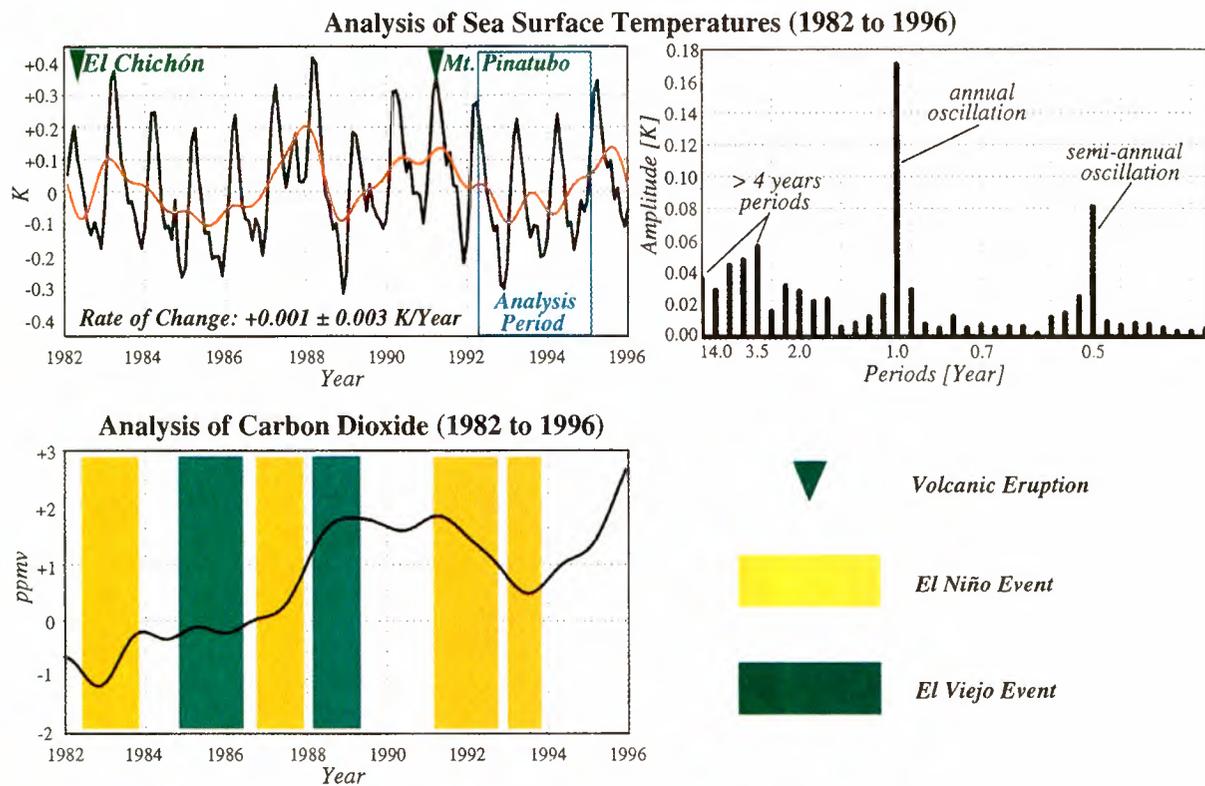


Plate 9. Analysis of sea surface temperatures and atmospheric carbon dioxide from 1982 until 1996.

altimeter range corrections like for example, the wet tropospheric corrections must have the same accuracy level as the altimeter range; (5) altimeter range corrections like the ionospheric correction must be derived from actual measurements and not from models; and because (6) it is very difficult to verify obtained results the analysis must be embedded in processes of the global system ocean-atmosphere.

The points mentioned above refer to input data and their quality. Another aspect, which must be discussed, is the question about the time relation of the obtained 2.2 mmyr^{-1} sea level rise.

An attempt is made to show the problem through sea surface temperatures and carbon dioxide data of a 14-year time series (1982-1995). Plate 9 compiles the results.

5.1. Sea Surface Temperature Analysis

As described in section 3, the 14 years of monthly sea surface temperature grids are processed to global and local rates of change. The deviations of the monthly grids from the long-term mean can be seen in Plate 9 (top) (black curve). Right next to the time series the corresponding amplitude spectrum is plotted. Three individual peaks can be identified. The annual oscillation is the dominant period in the spectrum followed by the half-year period. However, it seems that the time series contains some long period effects (4-year period and longer). This means that within the 14 years, there are longer timescales with global ocean heating and cooling as well. In order to highlight those periods a low-pass filter was applied to the time series that removed all periods ≤ 1 year (Plate 9 (top), orange curve). A light blue box marks the

investigation period. It can be noted that the investigation time period lies within a warming period of the oceans. Keeping in mind the high correlation between sea surface heights and sea surface temperatures, it is clear that the global sea level rise is partly caused by the thermal expansions of the oceans. In Plate 9 (top) the two triangles mark events that influenced the heat budget of the oceans, the eruptions of El Chichón and Mount Pinatubo. Both volcanic eruptions catapulted large amount of ashes to the atmosphere, so that the incoming solar radiation was weakened. This led to a global cooling of the oceans, which was -0.1 K for El Chichón and -0.15 K for Mount Pinatubo within 1 year [Reynolds, 1993; Stendel and Bengtsson, 1996]. The following rewarming periods lasted about 5 years. The investigation time period from 1992 to 1996 lies within the warming period caused by Mount Pinatubo.

5.2. Carbon Dioxide Analysis

There are two more oceanic phenomena that influence the sea surface temperatures, El Niño and El Viejo. El Niño is defined as a large downwelling Kelvin wave propagating eastward along the equatorial Pacific [Busalacchi and O'Brien, 1981]. Two methods are used to fix such an event: (1) the sea surface temperatures in the tropical Pacific are at least 0.5 K higher than in the six preceding months or (2) the sea level at Galapagos is at least 2 cm higher than in the six preceding months.

An El Niño event changes the vertical structure of the ocean and its CO_2 cycle. Usually the tropical Pacific is a huge source of atmospheric CO_2 because of its equatorial upwelling and outgassing of CO_2 rich deep water. However,

the downwelling Kelvin wave of El Niño prevents deep water from reaching the surface [Feely *et al.*, 1987]. Thus CO₂ is not released, and the absence of mixing of warm upper layer water with cold deep water causes the positive sea surface temperature anomaly.

In Plate 9 (bottom) the low pass filtered time series of atmospheric CO₂ is displayed (only periods greater equal one year). With green and yellow bars the El Niño and El Viejo events are marked. The plot shows that an El Viejo coincides with a local maximum of CO₂ and El Niño with a local minimum respectively [Keeling *et al.*, 1989]. The correlations, however, are not always obvious, and the magnitudes of the events are not always proportional to observed CO₂ concentrations [Meyers and O'Brien, 1995]. The time series of sea surface temperatures and atmospheric CO₂ span the same time period. Thus an intercomparison can be performed. In Plate 9 (bottom) the abovementioned coherence (El Niño, sea surface temperature increase and CO₂ decrease; El Viejo, sea surface temperature decrease and CO₂ increase) can be seen.

The analysis of atmospheric CO₂ was added to the sea level investigation to show that there are various phenomena, which in turn can change sea surface temperatures and thus the sea surface. The analysis of both CO₂ and sea surface temperatures have demonstrated that long-term events exist in the global system ocean-atmosphere. Because of the fact that the sea level is also one controlling factor in this system, it can be concluded that both sea level rise and fall must have happened several times in the past.

6. Conclusion

The comparison of consistent time series from ERS-1 and TOPEX altimeter data yields a global sea level rise of about 2 mmyr⁻¹ within April 1992 and March 1995. For the same time period wind speeds, wave heights and sea surface temperatures were investigated. It was found that there are positive trends for the sea surface temperatures and wave heights. For the wind speeds from altimetry a positive trend was estimated, while for the wind speeds from ERS-1 scatterometer a negative trend was derived. It can be concluded that one of the data sets was insufficiently calibrated. Besides global trends, local rates of change were also estimated. A sea level rise was detected in the tropics and in the Indian Ocean and a sea level fall in higher latitudes, respectively. The sea level study was embedded in the global system ocean-atmosphere, which was represented by wind speeds, wave heights, sea surface temperatures, and atmospheric CO₂. The local rates of change of wind speeds, wave heights, and sea surface temperatures showed a disproportion between variations on the northern and southern hemisphere. Correlation to the sea level change could also be identified. A high coherence between those parameters of the global system ocean-atmosphere is evident and could be demonstrated. Hence it can be supposed that the variations can be attributed to periods longer than the investigated 3 years [Hastenrath, 1984; Houghton, 1991].

The small amounts of global and local rates of change raise the question about their significance. Therefore the altimeter corrections were individually processed and analyzed. It was found that deficiencies with respect to their quality and long-term behavior exist. Especially, the ERS-1 ionospheric correction, which is derived from a model, is

very critical. Depending on the model adopted for the correction, either a sea level fall or a sea level rise can be produced. Investigations, however, have shown that the Bent model conforms much better with the ionospheric correction derived from TOPEX dual-frequency measurements than the IRI95 model. Thus a sea level rise is more likely. Because of the uncertainties of the ionospheric models it is absolutely necessary that a dual-frequency altimeter is an onboard feature of future missions. Furthermore, an attempt should be initiated to improve the ERS altimeter data by including the DORIS-derived ionospheric model. This could be achieved by analyzing the DORIS bifrequency Doppler receiver onboard of SPOT-2 and SPOT-3, which were orbiting at the same time as ERS-1, on analogous sun-synchronous orbits with nearly equivalent local hours for the ascending/descending nodes.

By means of TOPEX altimeter data and tide gauges an external quality assessment was attempted. The global and local rates of change between ERS-1 and TOPEX coincided to a high extent. Between ERS-1 and tide gauges an impressive level of local coherence was shown to exist. Furthermore, the obtained results were embedded in the discussion of 14 years of sea surface temperature and carbon dioxide. It was shown that the investigation time period lies within a global heating phase.

If all the results are summarized, then the 2.2 mmyr⁻¹ sea level rise is caused by a combination of thermal expansion, El Niño, and interannual variability. At this moment the effects cannot be separated. The question as to whether the melting of continental ice or polar caps are (partly) responsible for a sea level rise cannot be answered. Recent results on the net amount of ice concentration of Antarctica indicate an increase of the ice masses [Heygster, 1996]. The growth of polar ice, however, leads to a sea level fall, which is in contradiction to the sea level rise obtained by the study.

Open questions lead to two important requirements for the future.

1. The investigation period of 3 years is much too short for long-term predictions of the sea level or the climate system. Therefore the observation and analysis of oceanic and atmospheric data must be seen as a continuous process. Without consistent time series a description of present and future conditions of the "planet Earth" is not possible.
2. Accurate and well-calibrated input data and their corrections are required. Therefore also different data sources with a temporal overlapping are needed. The calibration of data must be a continuous task.

As a final note, it can be stated that the data sets for ERS-1, ERS-2, and TOPEX/POSEIDON are still improving, so that a number of uncertainties like the ionospheric correction could be overcome in the next years. This will help to iterate and consolidate the sea level study with improved data and longer time series.

Acknowledgments. A number of data was provided by different institutes. We thank the F-PAF for the operational provision of ERS OPR2 altimeter data. The IRI95 model was kindly provided by D. Bilitza. Thanks belong to O.B. Andersen and R. Ray for the preparation of the FES95.1 tidal loading modules. The range correction tables for ERS-1 and TOPEX were obtained from ESRIN/ESA and GSFC Wallops Flight Facility. The recent result from the ice concentration of Antarctica was delivered by the University of Bremen, Germany. The atmospheric carbon dioxide time series was kindly provided by Pieter Tans, NOAA/CMDL. We

received ERS-1 scatterometer data from IFREMER, France. On a regular basis AVISO provided TOPEX/POSEIDON data. Thanks belong to Richard Reynolds from NMC, Washington, D.C., for the preparation of sea surface temperature models and his helpful comments. Special thanks belong to Hans Neumayer for the review of the paper.

References

- Allan, T.D., Satellite Microwave Remote Sensing, *Ellis Horwood Ser. Mar. Sci.*, Ellis Horwood, Chichester, England, 1983.
- Anzenhofer, M., and T. Gruber, MSS93A: A new stationary sea surface combining one year upgraded ERS-1 fast delivery data and 1987 Geosat altimeter data, *Bull. Géod.*, 69, 157-163, 1995.
- Anzenhofer, M., W. Bosch, T. Gruber, and T. Spöcker, D-PAF quick-look sea surface monitoring with ERS-1 fast delivery altimeter data, in *International Association of Geodesy Symposia, Geodesy and Physics of the Earth*, vol. 112, pp. 87-91, Springer-Verlag, Berlin, 1992.
- Anzenhofer, M., T. Gruber, and M. Rentsch, D-PAF sea surface models from different data sources, in *Proceedings of the Second ERS-1 Symposium: Space at the Service of our Environment*, Eur. Space Agency Spec. Publ., ESA SP-361, 783-787, 1994.
- Anzenhofer, M., T. Gruber, and M. Rentsch, Global high resolution mean sea surface based on ERS-1 35- and 168- day cycles and TOPEX data, in *Proceedings of International Association of Geodesy Symposia, Global Gravity Field and Its Temporal Variations*, vol. 116, Springer-Verlag, New York, 1996a.
- Anzenhofer, M., C. Rajasenan, T. Gruber, and F.-H. Massmann, ERS-2 RA calibration using OPR2, *final tech. rep.*, GeoForschungsZentrum, Potsdam, Germany, 1996b.
- Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO), AVISO user handbook: Merged TOPEX/POSEIDON products, *Handb. AVI-NT-02-101-CN*, ed. 2.1, Toulouse, France, 1992.
- Bent, R. B., S.K. Llewellyn, and P.E. Schmid, Ionospheric refraction corrections in satellite tracking, *Space Res. XII*, pp. 1186-1194, Akademie-Verlag, Berlin, 1972.
- Bentamy, A., N. Grima, Y. Quilfen, V. Harscoat, C. Maroni, and S. Pouliquen, *An Atlas of Surface Wind From ERS-1 Scatterometer Measurements*, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest, France, 1997.
- Benveniste, J., Minutes of the ERS-2 radar altimeter and microwave radiometer commissioning working group (#9), ESA-ESRIN, *Minutes rs/pml/jb196.009*, Eur. Space Res. Inst., Frascati, Italy, January 30-31, 1996a.
- Benveniste, J., Minutes of the ERS-2 radar altimeter and microwave radiometer commissioning working group (#10), ESA-ESRIN, *Minutes rs/pml/jb196.026*, Eur. Space Res. Inst., Frascati, Italy, April 25-26, 1996b.
- Bilitza, D., International reference ionosphere: Recent developments, *Radio Sci.*, 21(3), 343-346, 1986.
- Bonavito, N.L., R.A. Gordon, and J.G. Marsh, On the nature of the radial and cross track errors for artificial earth satellites, report, Goddard Space Flight Cent., Greenbelt, Md., 1975.
- Broecker, W.S., Plötzliche Klimawechsel, *Spektrum der Wiss.*, *Januar*, 86-92, 1996.
- Busalacchi, A.J., and J.J. O'Brien, Interannual variability of the equatorial Pacific in the 1960s, *J. Geophys. Res.*, 86, 10901, 1981.
- Cartwright, D.E., and R.D. Ray, Oceanic tides from Geosat altimetry, *J. Geophys. Res.*, 95, 3069-3090, 1990.
- Cartwright, D.E., and R.J. Taylor, New computations of the tide-generating potential, *Geophys. J. R. Astron. Soc.*, 23, 45-74, 1971.
- Centre ERS d'Archivage et de Traitement (CERSAT), Altimeter and microwave radiometer ERS products, user manual, version 1.2, Plouzané, France, 1995.
- Cheney, R.E., N.S. Doyle, B.C. Douglas, R.W. Agreen, L. Miller, E.L. Timmerman, and D.C. McAdoo, The complete Geosat altimeter GDR handbook, *NOAA Manual NOS NGS 7*, Nat. Ocean Serv., Rockville, Md., 1991.
- Church, J.A., J.S. Godfrey, D.R. Jackett, and T.J. McDougall, A model of sea level rise caused by ocean thermal expansion, *J. Clim.*, 4(4), 438-456, 1991.
- Duchossois, G., and P. Martin, ERS-1 and ERS-2 tandem operations, *ESA Bull.* 83, 54-60, 1995.
- European Space Agency (ESA), ERS-1 system, *Eur. Space Agency Rep. SP-1146*, ESA Publ. Div., Noordwijk, Netherlands, 1992.
- European Space Agency (ESA), ERS user handbook, *Eur. Space Agency Rep. SP-1148*, ESA Publ. Div., Noordwijk, Netherlands, 1993.
- Feely, R.A., R.H. Gammon, B.A. Taft, P.E. Pullen, L.S. Waterman, J.F. Conway, J.F. Gendron, and D.P. Wisegarver, Distribution of chemical tracers in the eastern equatorial Pacific during and after the 1982-1983 El Niño/Southern Oscillation event, *J. Geophys. Res.*, 92, 6545-6558, 1987.
- Flechtner, F., S. Bedrich, and F.-H. Massmann, PRARE/ERS-2: System status and results, *CSTG Bull.* 13, 67-71, 1997.
- Francis, O., and P. Mazzeo, Global charts of ocean tide loading effects, *J. Geophys. Res.*, 95, 11411-11424, 1990.
- Fu, L.-L., and R.E. Cheney, Application of satellite altimetry to ocean circulation studies: 1987-1994, *U.S. Nat. Rep. Int. Union Geod. Geophys.*, 1991-1994, *Rev. Geophys.*, 33, 213-223, 1995.
- Fu, L.-L., and R. Glazman, The effect of wave age on the sea state bias in radar altimetry measurements, *J. Geophys. Res.*, 96, 8829-8834, 1991.
- Fu, L.-L., and G. Pihos, Determining the response of sea level to atmospheric pressure forcing using TOPEX/POSEIDON data, *J. Geophys. Res.*, 99, 24633-24642, 1994.
- Fu, L.-L., E.J. Christensen, C.A. Yamarone, M. Lefebvre, Y. Ménard, M. Dorrer, and P. Escudier, TOPEX/POSEIDON mission overview, *J. Geophys. Res.*, 99, 24369-24381, 1994.
- Groetzer, A., M. Latif, and T.P. Barnett, A decadal climate cycle in the north atlantic ocean as simulated by the echo coupled GCM, *Rep. 208*, Max-Planck-Inst. für Meteorol., Hamburg, Germany, 1996.
- Gruber, T., F.-H. Massmann, and C. Reigber, ERS-1 D-PAF Global Products Manual, GeoForschungsZentrum, Potsdam, Germany, 1993a.
- Gruber, T., M. Anzenhofer, M. Rentsch, Quick-look ocean products generation at D-PAF, in *Proceedings of the Second ERS-1 Symposium: Space at the Service of Our Environment*, edited by B. Kaldeich, Eur. Space Agency, Spec. Publ., SP-361, 1191-1196, 1993b.
- Hancock, III D.W., and G.S. Hayne, Error in TOPEX oscillator drift correction, report NASA/GSFC Wallops Flight Facil., Obs. Sci. Branch, Wallops Island, Va., July 10, 1996.
- Hancock, D.W., F.G. Forsythe, and J. Lorell, Seasat altimeter sensor file algorithms, *IEEE J. Oceanic Eng.*, OE-5, 93-99, 1980.
- Hastenrath, S., Interannual variability and annual cycle: Mechanisms of circulation and climate in the tropical atlantic sector, *Mon. Weather Rev.*, 112, 1097-1107, 1984.
- Heygster, G., PELICON: Project for estimation of long-term variability in ice concentration, final rep.: *EC Contract EV5V-CT93-0268 (DG 12 DTEE)*, Univ. of Bremen, Germany 1996.
- Horai, K.-I., Adjustment, interpolation and smoothing of GEOS- 3 altimeter data, *J. Geophys. Res.*, 87, 8693-8707, 1982.
- Houghton, R.A., and G.M. Woodwell, Global climatic change, *Sci. Am.*, 260 (4), 18-26, 1989.
- Houghton, R.W., The relationship of sea surface temperature to thermocline depth at annual and interannual timescales in the tropical atlantic ocean, *J. Geophys. Res.*, 96, 15173-15185, 1991.
- Institut Français de Recherche et d'Exploitation de la Mer (IFREMER), Scatterometer mean wind field products, user manual, *C2-MUT-W-02-IF*, version 1.0, Inst. Fr. de Rech. pour l'Exploit. de la Mer, Brest,, France, 1996.
- Isbell, D., and M. Hardin, TOPEX/POSEIDON sea-level measurements to be revised, *NASA Press Release 96-148*, July 25, 1996.
- Keeling, C.D., R.B. Bacastow, A.F. Carter, S.C. Piper, T.P. Whorf, M. Heimann, W.G. Mook, and H. Roeloffzen, A three-dimensional model of atmospheric CO₂ transport based on

- observed winds: Analysis of observational data, in *Aspects of Climate Variability in the Pacific and the Western Americas*, *Geophys. Monogr. Ser.*, vol. 55, edited by D.H. Peterson, pp. 165-236, AGU, Washington D.C., 1989.
- Knudsen, P., O.B. Andersen, and T. Knudsen, ATSR sea surface temperature data in a global analysis with TOPEX/POSEIDON altimetry, *Geophys. Res. Lett.*, 23 (8), 821-824, 1996.
- Kopytoff, V.G., Error inflated estimate of rising sea levels, researchers report, *New York Times*, July 30, 1996.
- Latif, M., R. Kleeman, and C. Eckert, Greenhouse warming, decadal variability, or El Niño? An attempt to understand the anomalous 1990s, *Rep. 175*, Max-Planck-Inst. für Meteorol., Hamburg, Germany, 1995.
- Lemoine, F.G., et al., The development of NASA-GSFC/DMA geopotential model, in *Proceedings of IAG Symposium: Gravity, Geoid and Marine Geodesy*, vol. 117, edited by K.-P. Schwarz, pp. 461-469, Springer-Verlag, New York, 1997.
- Le Provost, C., F. Lyard, and J.M. Molines, Improving ocean tide predictions by using additional semidiurnal constituents from spline interpolation in the frequency domain, *Geophys. Res. Lett.*, 18, 845-848, 1991.
- Le Provost, C., F. Lyard, J.M. Molines, M.L. Genco, F. Rabilloud, A hydrodynamic ocean tide model improved by assimilating a satellite altimeter derived data set, *J. Geophys. Res.*, in press, 1998.
- Lindau, R., Time dependent calibration of marine Beaufort estimates using individual pressure differences, in *Proceedings of the International COADS Winds Workshop*, editor H.F. Diaz, rep. 265, Berichte aus dem Inst. für Meereskunde an der Christian-Albrechts-Univ. Kiel, Germany, 1995a.
- Lindau, R., Klima des Atlantiks, *Ann. Meteorol.*, 82-83, 1995b.
- Llewellyn, S.K., and R.B. Bent, Documentation and description of the Bent ionospheric model, *Rep. AFCRL-TR-73-0657*, 1973.
- Marsh, J.G., and R.G. Williamson, Seasat altimeter timing bias estimation, *J. Geophys. Res.*, 87, 3232-3238, 1982.
- Meier, M.F., The contribution of small glaciers to global sea level rise, *Science*, 226, 1418-1421, 1984.
- Meyers, S.D., and J.J. O'Brien, Pacific ocean influences atmospheric carbon dioxide, *Eos Trans. AGU*, 76 (52), 533, 1995.
- Nerem, R.S., et al., Gravity model development for TOPEX/POSEIDON: Joint Gravity Models 1 and 2, *J. Geophys. Res.*, 99, 24421-24447, 1994.
- Nerem, R.S., Measuring global mean sea level variations using TOPEX/POSEIDON altimeter data, *J. Geophys. Res.*, 100, 25135-25151, 1995.
- Pugh, D.T., The global sea-level observing system, *Hydrogr. J.*, 45, 5-8, 1987.
- Quilfen, and Y., A. Cavanie, A high precision wind algorithm for the ERS-1 scatterometer and its validation, in *IGARSS '91 Remote Sensing: Global Monitoring for Earth Management*, IEEE Publ. 91CH2971-0, Inst. of Elec. and Electrom. Eng., New York, 1991.
- Reigber, Ch. et al., CHAMP Phase B, executive summary, *Scie. Tech. Rep. STR96/13*, GeoForschungsZentrum Potsdam, Germany, 1996.
- Reynolds, R.W., A real-time global sea surface temperature analysis, *J. Clim.*, 1, 75-86, 1987.
- Reynolds, R.W. Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures, *J. Clim.*, 6, 768-774, 1993.
- Reynolds, R.W. and D.C. Marsico, and T.M. Smith, An improved real-time global sea surface temperature analyses, *J. Clim.*, 6, 114-119, 1993.
- Saastamoinen, J., Atmospheric correction for the troposphere and stratosphere in radio ranging of satellites, in *The Use of Artificial Satellites for Geodesy*, *Geophys. Monogr. Ser.*, vol. 15, edited by S.W. Henriksen, A. Mancini, and B.H. Chovitz, pp. 247-251, AGU, Washington D. C., 1972.
- Schaer, S, M. Rothacher, and G. Beutler, Global ionosphere maps based on GPS carrier phase data from the IGS network produced by the CODE analysis center, *Eos Trans. AGU Spring Meeting*, 77 (12), 71, 1996.
- Schwiderski, E.W., On charting global tides, *Rev. Geophys.*, 18 (1), 243-268, 1980.
- Schwintzer, P., et al., A new earth gravity field model in support of ERS-1 and Spot-2: GRIM4-S1/C1, final rep. Ger. Space Agency (DARA), München, Germany, 1991.
- Schwintzer, P., et al., Improvement of GRIM4 earth gravity models using Geosat altimeter and Spot-2 and ERS-1 tracking data, in *Proceedings of IAG Symposium: Geodesy and Physics of the Earth*, vol. 112, edited by H. Montag and Ch. Reigber, pp. 75-78, Springer-Verlag, New York, 1993.
- Schwintzer, P., et al., Long-wavelength global gravity field models: GRIM4-S4, GRIM4-C4, *J. Geod.*, 7, 189-208, 1997.
- Shum, C.K., et al., Accuracy assessment of recent ocean tide models, *J. Geophys. Res.*, 102, 25173-25194, 1997.
- Spöcker, N., M. Anzenhofer, W. Bosch, T. Gruber, Quick-look monitoring of the sea surface with ERS-1 fast delivery altimeter data, in *Proceedings of the First ERS-1 Symposium*, vol. 1, pp. 469-472, Eur. Space Agency Publ. Div., Noordwijk, Netherlands, 1992.
- Stammer, D., and C. Wunsch, Preliminary assessment of the accuracy and precision of TOPEX/POSEIDON altimeter data with respect to large-scale ocean circulation, *J. Geophys. Res.*, 99, 24584-24604, 1994.
- Stendel, M., and L. Bengtsson, Monitoring the temperature of the troposphere by means of a general circulation model, *Rep. 186*, Max-Planck-Inst. für Meteorol., Hamburg, Germany, 1996.
- Tapley, B.D., et al., The Joint Gravity Model 3, *J. Geophys. Res.*, 101, 28029-28049, 1996.
- Thoning, K.W., P.P. Tans, and W.D. Komhyr, Atmospheric carbon dioxide at Mauna Loa Observatory, 2, Analysis of the NOAA GMCC data, 1974-1985, *J. Geophys. Res.*, 94, 8549-8565, 1989.
- Trupin, A., and J. Wahr, Spectroscopic analysis of global tide gauge sea level data, *Geophys. J. Int.*, 100, 441-453, 1990.
- Wahr, J.W., Deformation of the earth induced by polar motion, *J. Geophys. Res.*, Vol. 90, 9363-9368, 1985.
- Witter, D.L., and D.B. Chelton, A Geosat altimeter wind speed algorithm and a method for altimeter wind speed algorithm development, *J. Geophys. Res.*, 96, 8853-8860, 1993.

M. Anzenhofer and T. Gruber, GFZ/D-PAF, Münchnerstr. 20, Postfach 1116, D-82230 Oberpfaffenhofen, Germany. (e-mail: anzenhof@dfd.dlr.de)

(Received January 6, 1997; revised August 22, 1997; accepted September 8, 1997.)

Precise orbit determination and gravity field improvement for the ERS satellites

Remko Scharroo and Pieter Visser

Delft Institute for Earth-Oriented Space Research, Delft University of Technology, Delft, Netherlands

Abstract. The radial orbit error has long been the major error source in ERS-1 altimetry, crippled by having only satellite laser ranging for precise tracking and relying on insufficiently accurate general-purpose gravity field models. Altimeter crossovers are used very effectively as additional tracking data to laser ranging. The ERS Tandem Mission even provides the unique possibility to simultaneously determine orbits of two similar satellites flying the same orbit. Altimeter crossovers between the two satellites then link the two orbits into a common reference frame. Tailoring of the Joint Gravity Model 3 (JGM 3) is another step to reduce orbit errors. This technique is aimed at the reduction of the geographically anticorrelated orbit error (observed in the crossover height differences) through the adjustment of selected gravity field parameters. The resulting Delft Gravity Model (DGM)-E04 has reduced this part of the orbit error by a factor of 2, performs even better with respect to the ESA-provided orbits, and also outperforms the recent Earth Geopotential Model EGM96 in this respect. ERS-1 and ERS-2 orbits for the entire Tandem Mission are computed and studied in detail, and orbit errors due to the gravity field and nonconservative forces are identified. Analyses systematically show that the orbits computed with JGM 3 have a radial root-mean-square orbit accuracy of 7 cm, with DGM-E04 5 cm.

1. Introduction

The Tandem Mission of the two European Remote Sensing Satellites ERS-1 and ERS-2 started soon after launch of ERS-2, April 21, 1995. Since then, both satellites are flying in the same near-circular orbit with an average altitude of 800 km and an inclination of 98.5° . The orbits are Sun-synchronous and are phased such that the satellites cross the equator at 2230 local solar time (LST) in northern direction and 1030 LST in southern direction. Generally speaking, the satellites observe the Earth at nightfall on ascending tracks and at dawn on descending tracks. After making 501 revolutions in exactly 35 days the satellites repeat their ground tracks. Until today, even though ERS-1 is no longer providing data, ERS-2 is following it at a 32-min lag in the same orbit plane, which causes ERS-1 to precede ERS-2 by 24 hours along the same ground track.

Satellite tracking is the only means to tie the (at itself relative) radar altimeter range measurement into a well-defined global reference frame. Precise orbit determination (POD), on the basis of these tracking data and various dynamical models, ensures continuity of this link also at locations where there is no immediate tracking. The accuracy at which the absolute sea level, land, or ice elevation is inferred by differencing the orbital altitude and the altimeter range measurement is always limited by the accuracy of the orbit computation. The radial orbit error has been one of the larger errors in recovering the sea surface height from ERS altimeter measurements and, because of their long-wavelength character, disqualified ERS altimeter data from application in global ocean circulation studies, ocean tide modeling, and monitoring of seasonal and secular (climate-related) change. The 10-cm overall error

budget required by the oceanographic community [Tapley, 1992] just could not be met. Recently, SAR interferometry, with its applications of mapping land surface elevation and elevation changes, is now also demanding ever increasing accuracy of the orbit computation in cross-track direction.

The bulky satellites ERS-1 and ERS-2 were never designed for high-accuracy orbit determination, and the loss of the Precise Range and Range-Rate Equipment (PRARE) tracking system left ERS-1 even more poorly equipped for orbit determination. Yet, subdecimetric orbit accuracy is not of academic interest only. The ERS altimetric system has performed well above expectations and is unique because of its multidisciplinary character, sampling not only ocean but also land and ice surfaces, in combination with the suite of instruments on board, providing, e.g., simultaneous measurements of wet tropospheric content and surface temperature. Undoubtedly, ERS will always lag behind on the 2-cm root-mean-square orbit accuracy of the TOPEX/POSEIDON altimeter mission [Marshall *et al.*, 1995], so only when the precise orbit determination is stretched to its very limits, ERS altimetry will be regarded a reliable source of information. Only then will ERS be able to demonstrate its additive value in ocean research and its unique capabilities in, e.g., monitoring of the ice sheet mass balance.

In section 2 we will discuss the numerous advances made in ERS operational and precise orbit determination over the years up to the current state-of-the-art modeling. An important step in this is the development of the ERS-tailored Delft Gravity Model (DGM)-E04 (section 3). In section 4, DGM-E04 demonstrates that it constitutes a remarkable improvement on the ERS orbits but also has a more general applicability. Section 5 examines the effect of nonconservative forces acting on the satellite. Section 6 combines all results and attempts to give a realistic error budget for the radial accuracy of the various orbit solutions discussed in this paper and demonstrates that, with the selection of the right orbit solution, the overall error budget for the ERS altimetric system is well within the origi-

Copyright 1998 by the American Geophysical Union.

Paper number 97JC03179.
0148-0227/98/97JC-03179\$09.00

nal goals set for TOPEX/POSEIDON. Finally, we will discuss the achievements and limitations in section 7.

2. Operational and Precise Orbit Determination

In December 1991 the Delft University of Technology (DUT) produced the first operationally determined orbit for ERS-1 and has been doing so ever since at the regular pace of two orbital arcs per week. With these orbits and additional geophysical models the United States' National Oceanic and Atmospheric Administration (NOAA) has been upgrading the European Space Agency's (ESA) ERS fast-delivery altimeter product to interim geophysical data records (IGDRs).

Major advances have been made during ERS' lifetime in the accurate restitution of the orbit. Starting at radial orbit errors of around 140 cm in 1991, the best available orbits are now believed to be accurate up to about 5 cm. This could only be obtained by adopting improved models for the gravity field, the satellite surface forces, and adding altimeter crossover data to the POD as long as the PRARE tracking data were not generally available.

2.1. History

At the time of the launch of ERS-1 the NASA Goddard Space Flight Center (GSFC) GEM-T2 gravity model [Marsh *et al.*, 1989] was the most commonly used model. Because this model lacked sufficient high-inclination orbit information in its development, the resulting ERS-1 orbits started at about 140-cm radial accuracy. After GSFC included tracking data of SPOT-2, which runs in an orbit very similar to ERS, orbit errors reduced drastically to ~ 30 cm [Scharroo *et al.*, 1993a]. ERS-1 tracking data were first included in the first Joint Gravity Models for the TOPEX/POSEIDON mission,

JGM 1 and JGM 2 (NASA/University of Texas at Austin) [Nerem *et al.*, 1994], and in the GeoForschungsZentrum (GFZ) PGM035 gravity model used to produce the official ESA orbits from the German Processing and Archiving Facility (D-PAF). The introduction of these models in the ERS-1 POD reduced the orbit error to a level of 15–20 cm [Aksnes *et al.*, 1994; Massmann *et al.*, 1994; Scharroo *et al.*, 1993b, 1994], at which the surface force modeling errors (drag and solar radiation) started to dominate [Le Traon *et al.*, 1995]. A new model for the satellite geometry was developed and successfully implemented, but incorporating altimeter range measurements as an additional tracking type scored little success in reducing the remaining orbit errors [Scharroo *et al.*, 1993a, b; Visser, 1993; Shum *et al.*, 1994]. Nondynamic orbit improvement using TOPEX/POSEIDON as a reference, with all its associated limitations, seemed the only way to provide subdecimetric orbits for ERS-1 [Smith and Visser, 1995; Le Traon *et al.*, 1995; Le Traon and Ogor, this issue].

2.2. Second-Generation Precise Orbits

Simultaneous with the launch of ERS-2 and the start of the ERS Tandem Mission (May 1995), DUT introduced JGM 3 [Tapley *et al.*, 1996] in the operational orbit determination. This brought the radial orbit error down to ~ 10 cm, but satellite laser ranging (SLR) to ERS-1 and ERS-2 simply remained insufficiently abundant and lacked a regular global distribution required to provide subdecimetric accuracy everywhere on the globe. With most of the tracking stations located in Europe and North America (Figure 1), orbit accuracy remained poor, especially at southern latitudes. Inclusion of additional tracking data was imperative and found in the form of altimeter crossover height differences, nailing down the orbits everywhere over the oceans and paving the way to introduce additional parameters in the POD to absorb remaining unmodeled or inadequately modeled forces.

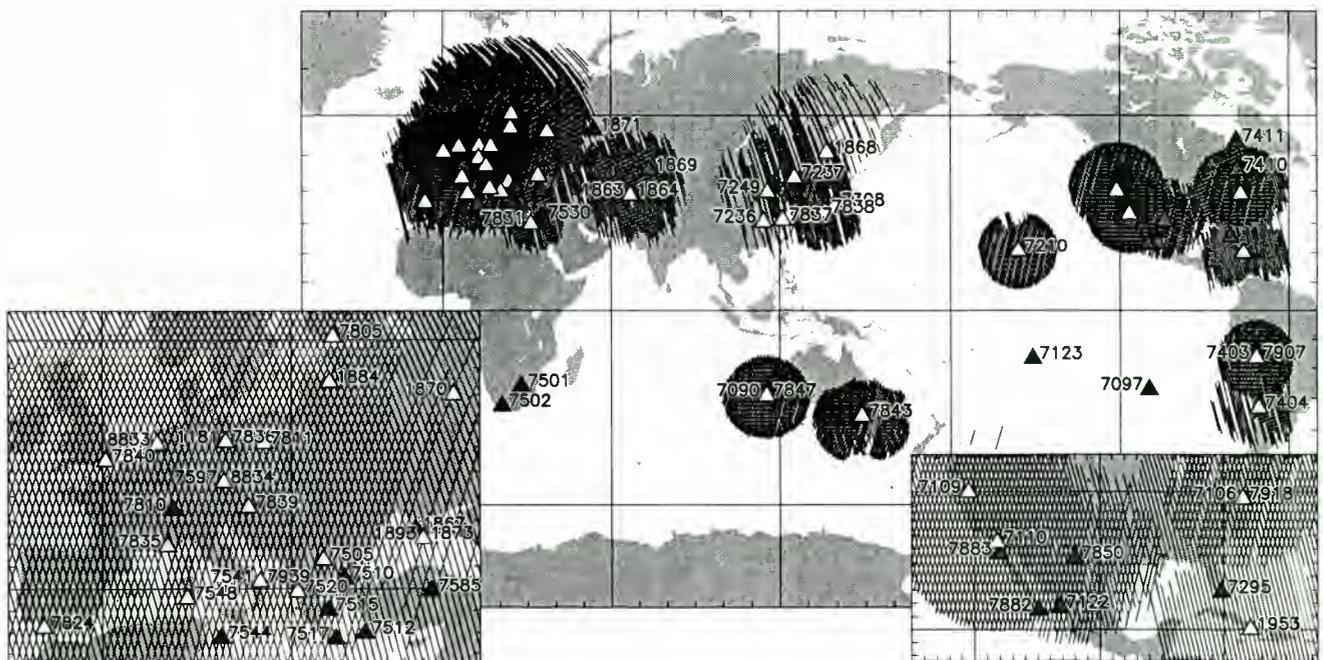


Figure 1. Satellite laser ranging coverage for ERS-1 and ERS-2 during the Tandem Mission. Locations of SLR stations that have tracked ERS-1 or ERS-2 during the Tandem Mission are indicated by open triangles and site numbers.

In spring 1996, DUT produced a full set of second-generation precise ERS-1 and ERS-2 orbits for the period April 1992 until August 1995, on the basis of the JGM 3 gravity model and SLR and altimeter crossover (XO) tracking data. Of these, the ERS-1 orbits for mission phases C, D, E, and F (multidisciplinary phase, second ice phase, and the two geodetic phases) were made available at <http://deos.lr.tudelft.nl/ers/precorbs/>. The radial orbit accuracy of about 7 cm [Scharroo *et al.*, 1996a, b] was a significant improvement over the 10-cm accuracy of the GFZ/D-PAF orbits featured on the official ESA altimeter products. This was mainly due to the superiority of JGM 3 over the GFZ PGM035 gravity model used in those days to compute these orbits.

2.3. Third-Generation Precise Orbits

A natural step was to develop a gravity model, tailor-made for ERS orbit determination, starting from JGM 3, and so to reduce the gravity-induced orbit error. The resulting Delft Gravity Model DGM-E04, discussed in section 3 now forms the basis for the third-generation precise ERS-1 and ERS-2 orbits, currently all accessible at <http://deos.lr.tudelft.nl/ers/precorbs/>. In this paper (section 4 and further) we will limit ourselves to the POD for the ERS Tandem Mission, exploiting the unique situation of having the two altimeter missions flying the same orbit. During this period, ERS orbits are computed simultaneously, using ERS-1/2 dual-satellite crossovers to link the orbits tightly in a common reference frame and to have one orbit benefit from the other when SLR tracking is sparse. At the same time, altimeter height differences along collinear tracks (section 5.2) and ERS/TOPEX dual-satellite crossovers (section 4.4) remain as independent indicators for the radial orbit error. This situation is significantly different from the ERS orbit determination based on ERS/TOPEX dual-satellite crossovers, where TOPEX orbits are merely used as a reference and are not simultaneously adjusted.

2.4. Applied Models and Constants

The procedures and models used for the POD are based on the most up-to-date knowledge of gravity and nonconservative force modeling. In this study, three different gravity models, JGM 3, EGM96 (NASA/National Imagery and Mapping Agency, truncated to degree and order 70) [Lemoine *et al.*, 1997], and DGM-E04 (DUT), have been used to describe the gravitational field of the Earth (including tides), and results are intercompared. The satellite surface forces are accurately modeled by means of satellite-specific macromodels consisting of 10 panels. Additional unmodeled forces are parameterized through so-called empirical accelerations. Because of the inclusion of SLR and single- and dual-satellite crossovers, sufficient data are available to estimate these additional parameters.

Other estimated parameters are the station location for some (mainly mobile) SLR stations and time tag and range biases for stations that are notorious for producing SLR measurements with significant offsets. Since the datation of the altimeter data may not be the same as the SLR standard (UTC), it is important to estimate a time tag bias on the altimeter data. If not, the altimeter ranges appear corrupted by a signal with a frequency of two cycles per revolution (cpr) introduced by shifting the altimeter ranges forward or backward on a flattened Earth; a signal that may well alias into the orbit [Schutz *et al.*, 1982].

Especially relevant to the orbit determination are the location of the satellite center of mass, and the laser retroreflector and altimeter reference points, listed in Table 1. A concise description of the modeling is given in Table 2.

Table 1. Coordinates of Some ERS Reference Points in a Body Fixed Frame

Reference Point	X_s	Y_s	Z_s
Altimeter	-3786.4	570.0	-840.4
Laser retroreflector	-2850.4	-700.0	-995.0
ERS-1 nominal CM	-1827.0	11.8	11.9
ERS-2 nominal CM	-1853.0	-9.0	-3.0

Coordinates are in millimeters. CM is Center of Mass.

2.5. Tracking Data

SLR 15-s normal points are collected from the Eurolas Data Center (EDC) and Crustal Dynamics Data Information System (CDDIS). All ranges are corrected for a distance of 4.3 cm between the effective sphere of reflection and the laser retroreflector reference point. The data weight is the root-sum-square (rss) of a system-dependent noise value (ranging from 1 to 20 cm) and an estimate of the overall solution error (5 cm).

All ERS altimeter data are retrieved from the ESA altimeter ocean products (OPR). Unfortunately, for the actual orbit determination we had to harmonize several different versions of this product. Version 3 OPRs were available prior to the Tandem Mission [Centre ERS d'Archivage et de Traitement (CERSAT), 1994], followed by a number of intermediate versions. The verification of the orbits of the Tandem Mission, however, is all based on the current Version 6 data [CERSAT, 1996].

All 1-Hz altimetric sea heights are screened and corrected for the following geophysical and instrument corrections: (1) range corrections for bias jumps and oscillator drift (ESA, public data, 1996); (2) GFZ/D-PAF precise orbits based on the GFZ PGM055 gravity model [Gruber *et al.*, 1997]; (3) meteorological dry tropospheric (ECMWF) and ionospheric corrections (Bent); (4) wet tropospheric correction from the ERS Microwave Radiometer; (5) solid Earth and pole tides; (6) ocean tides (Grenoble FES95.2.1) and tidal loading; (7) sea-state bias; OPR version 3, 5.5% of Significant Wave Height (SWH); [Gaspar and Ogor, 1994]; OPR version 6, BM3 model [Gaspar and Ogor, 1996]; (8) 100% inverse barometer correction; and (9) mean sea surface (Ohio State University (OSU) MSS95). Forthwith, the 1-Hz relative sea heights are converted to altimeter crossover height differences (XD).

First, the location of the crossing point of all ascending and descending passes is computed. Combinations of passes with extreme shallow angles, passes with a time interval larger than 17.5 days, and XOs with too few surrounding 1-Hz measurements are rejected. (Theoretically, 17.5 days is the largest possible time interval between consecutive ascending and descending passes in a XO is half the repeat period.) Relative sea heights are filtered and interpolated with a quadratic polynomial at the location of the XO.

The two relative sea heights in the XOs are converted back to range measurements by subtracting them from the orbital altitude. The measurement weight varies with the geographical position and is based on a combination of a posteriori orbit error and sea surface variability determined from an earlier orbit solution (values ranging from 4 to 40 cm). Because with XDs we are only concerned with a height difference, the usage and choice of the reference mean sea surface is irrelevant but facilitates the data screening.

3. Development of an ERS-Tailored Gravity Model

The perception that much of the radial orbit error in second-generation orbits (section 2.2) was still caused by deficiencies of

Table 2. Summary of the Dynamical and Measurement Models Used for the ERS Third-Generation Precise Orbit Determination.

Item	Description
<i>Measurement Models and Constants</i>	
SLR observations	global quick-look SLR data (1 per 15 s normal points) retrieved from EDC and CDDIS; corrected for 4.3-cm distance to LRR reference point; 10° elevation cutoff and editing of spurious measurements; weight is rss combination of overall model accuracy (5 cm) and system noise level (1–20 cm)
Station coordinates	LSC(DUT)95L02 LAGEOS I/II solution (September 1983 – December 1993, epoch January 1, 1988), advanced to epoch by three-dimensional motions incorporated in the coordinate solution
Tidal displacement SLR geometric correction	Love model, including frequency dependent and permanent tides ($h_2 = 0.609$, $l_2 = 0.0852$); pole tide offset of LRR optical center with respect to LRR reference point (4.3 cm) and to the spacecraft nominal center of mass (see Table 1)
Crossover observations	global ERS radar altimeter data, retrieved from the ESA ERS OPR products, plus additional geophysical corrections (see section 2.5); converted to single- and dual-satellite XOs; weight is based on local rms XDs coming from an earlier solution
Speed of light	$c = 299792.458 \text{ km s}^{-1}$
<i>Force Model</i>	
Gravity model	Delft Gravity Model DGM-E04, complete to degree and order 70, including secular C_{21} and S_{21} and dynamic polar motion; $GM = 398600.4415 \text{ km}^3 \text{ s}^{-2}$, $a_e = 6378.1363 \text{ km}$, $1/f = 298.2564$
Tidal gravity	Wahr solid Earth tides; background ocean tides: JGM 3*
Third body attraction	Sun, Moon, Mercury, Venus, Mars, Jupiter, Saturn, Neptune, according to JPL DE200 ephemeris
Atmospheric drag	French Density Temperature Model (DTM) [Barlier et al., 1977] with daily $F_{10.7}$ and 3-hourly K_p values
Radiation	solar radiation including umbra, penumbra, and occultation by Moon; Earth albedo
Orbit maneuvers	a priori information according to ESOC predictions; adjusted during POD
<i>Satellite Model</i>	
Mass	ERS-1, 2377.13 kg; ERS-2, 2502.00 kg
Cross-sections	satellite-specific macromodels, each consisting of eight fixed and two rotating panels (see section 5.1)
<i>Reference Frame</i>	
Polar motion	Earth orientation and length of day from IERS EOP 90 C 04 solution
Coordinate system	J2000; precision IAU 1976 (Lieske model); nutation IAU 1980 (Wahr model)
<i>Estimated Parameters Per Orbital Arc</i>	
State vector	position and velocity at epoch over 5.5-day orbital arc at 3.5-day intervals
Nonconservative forces	6-hourly drag coefficients; 22-hourly 1-cpr along-track and cross-track accelerations; orbit maneuvers (three-dimensional accelerations)
Measurement offsets	coordinates of some (mobile) stations; range and timing bias for some SLR systems; timing bias for both altimeters; relative range bias between ERS altimeters

*In fact, this model was not adjusted since JGM 1 [Nerem et al., 1994; Tapley et al., 1996].

JGM 3, led to the attempt to develop a model specifically tailored to ERS orbit determination. This is done by first isolating the gravity-induced orbit error observed in XOs and then adjusting a well-chosen set of gravity field coefficients such that the error is minimized.

3.1. Linear Perturbation Theory

Gravity-induced orbit errors are, because of their origin, geographically correlated; that is, they are repetitive along the same ground track, repeat cycle after repeat cycle. This means that along two such collinear ground tracks gravity-induced orbit errors are identical and cancel when differencing the altimetric sea surface profiles along these tracks. This is not the case for two crossing tracks; the different “history” of gravity sensed along the ascending and descending passes causes the orbit error to be essentially different along each pass.

Figure 2 shows how these two gravity-induced radial orbit errors (δr^{asc} and δr^{des}) can alternatively be decomposed into a geographically fully correlated and anticorrelated component (δr^c and δr^s). The first component is identical on both passes; the second is of equal magnitude on both passes but of opposite sign. In literature [e.g., Tapley and Rosborough, 1985; Rosborough, 1986] these components are often given the confusing indications “mean error” and “variability error.” It should be stressed that there is nothing

variable about the component δr^s , which is just as time invariant as the “mean” component δr^c , with the only difference that δr^s has an opposite sign on the two passes. Thus we write for the gravity-induced radial orbit error in a crossover point

$$\delta r = \delta r^c \pm \delta r^s \quad (1)$$

where the plus sign is used for the ascending pass and the minus sign for the descending pass.

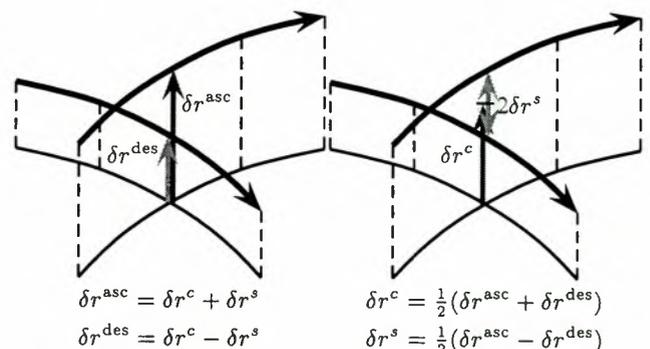


Figure 2. Geographically correlated orbit errors along two crossing passes.

Linear Perturbation Theory (LPT) [Kaula, 1966] describes the three orthogonal components of the orbit error as a linear combination of terms due to (commission or omission) errors in the gravity model coefficients. Each coefficient produces errors at various frequencies, depending on their harmonic degree l and order m . It can be easily shown [e.g., Schrama, 1989] that on the near-circular, Sun-synchronous, and "frozen" ERS orbit, the dominant errors have frequencies

$$\dot{\psi}_{lmp} = (l - 2p) \text{ cpr} - m \text{ cpd} \quad (2)$$

where p ranges from 0 to l and cpd is cycles per day. Some combinations of $l - 2p$ and m cause $\dot{\psi}_{lmp}$ to be close to the resonance frequencies 0 and 1 cpr. For ERS' 35-day repeat this happens for coefficients of order 43 because 43 cpd is very close to 3 cpr.

Rosborough [1986] focuses on the spatial distribution of the radial orbit error and conveniently formulates

$$\delta r = \sum_{l=1}^{\infty} \sum_{m=0}^l \delta r_{lm}^c \pm \delta r_{lm}^s \quad (3)$$

with

$$\delta r_{lm}^c = \bar{Q}_{lm}^c (\Delta \bar{C}_{lm} \cos m\bar{\lambda} + \Delta \bar{S}_{lm} \sin m\bar{\lambda}) \quad (4)$$

$$\delta r_{lm}^s = \bar{Q}_{lm}^s (\Delta \bar{C}_{lm} \sin m\bar{\lambda} - \Delta \bar{S}_{lm} \cos m\bar{\lambda}) \quad (5)$$

where \bar{Q}_{lm}^s and \bar{Q}_{lm}^c are both functions of latitude only and further depend on the orbit mean semimajor axis, eccentricity, and inclination.

This yields the following: (1) when computing XDs, x , the fully correlated part cancels, but the anticorrelated part is observed at double efficiency since

$$\delta x = \delta r^{\text{asc}} - \delta r^{\text{des}} = 2\delta r^s \quad (6)$$

(2) the zonal coefficients do not contribute to XDs since $\sin m\bar{\lambda} = 0$ and $\bar{S}_{lm} = 0$ for $m = 0$; (3) nonzonal coefficients of the gravity field do contribute to XDs, each with a distinct global pattern; and (4) vice versa, when we can isolate and observe the effect of gravity model deficiencies in XDs, the nonzonal coefficients can be tuned to reduce the observed effect.

The LPT has been used before to cancel gravity-induced radial orbit errors but never in such a way that it actually provided a well-tuned gravity model. Engelis [1987, 1988] and Visser [1992] use LPT to simultaneously improve dynamic topography and geoid from Seasat and Geosat altimeter data. The link between geoid and orbit errors provide the means to partially separate geoid and dynamic topography. Visser [1995] extends the technique by including SLR and single- and dual-satellite XDs. Because the gravity-induced orbit error was not isolated from other orbit errors, each of the additional estimated parameters in orbit determination also had to be readjusted. Novel in the approach proposed here is the use of crossover height differences instead of altimetric sea heights (thus canceling geoid and dynamic topography errors) and the use of the two-step approach: first isolating the gravity-induced signal and then adjusting the gravity field coefficients.

3.2. Isolating the Gravity-Induced Orbit Errors

The geographically anticorrelated orbit error is but one of the many contributions to XDs. The assumption we make is that all contributions, except the constant gravitational, are time variant and average out to zero over a sufficiently long period, like a year

or more, or are not observed at the maximum time interval of 17.5 days in our set of XDs.

The procedure of generating local average XDs can easily be performed by gridding the data on a regular latitude-longitude grid. However, this will remove some signal at high latitudes where, because of the closure of XO locations, the spatial scales are much shorter than around the equator.

A more sophisticated approach takes advantage of the fact that XOs are already located on a more or less regular "grid" which densifies toward the poles. XOs can be stacked together at the "grid points." Following a 3.5- σ editing in each stack, the average of the XDs in each grid point (\bar{x} and σ_x) represents twice the geographically anticorrelated orbit error ($2\delta r^s$), while the variance resembles $\sqrt{2}$ times the rss of the nongravitational orbit error, sea surface variability, and altimeter correction errors. An additional 3.5- σ editing is performed over all σ_x values to root out areas of extreme variability.

We ignore the fact that the XO locations actually vary from repeat cycle to repeat cycle within a radius of about 1 km because of drifting of the ground track around the nominal position by the same amount. This is allowed since the signal we are isolating has much longer wavelengths than 1 km. Also, we are not hampered by cross-track geoid slopes since we have constructed the XDs at their true locations and not nominal locations, and if we had, this would have been tackled by taking the OSU MSS95 model as reference.

This technique was used for a total of about 1,000,000 ERS-1 and ERS-2 single-mission XOs for the period of April 1992 until August 1995, created from the OPR altimeter products as described in section 2.5. All XDs were first adjusted by replacing the GFZ orbits by DUT second-generation JGM 3 orbits, simultaneously accounting for an estimated time tag bias of -1.3 and -1.1 ms for ERS-1 and ERS-2, respectively. The average and the variance of the XDs at the remaining 45,000 locations are sketched in Plate 1. The variance clearly resembles what we know of sea surface variability (high in Western Boundary Currents) and areas with large ocean tide modeling errors (e.g., East China Sea, Indonesian Archipelago). Clearly noticeable are a few latitude bands with remarkable low local variance; these pertain to those XOs with a time lag of about 12 hours, during which sea surface variability is minimal. Note also the significance of the averages compared to the variances: in areas with little or no variance (central Pacific) the average may be a couple of times larger than the variance. This indicates that the isolation of the gravity-induced contribution to XD functions quite well and that the gravity-induced orbit error is sizeable compared to nongravitational errors.

3.3. Tailoring the Gravity Model

Tailoring a gravity model involves the tuning of its coefficients such that residuals of observations from a single satellite are reduced. Usually, only a subset of coefficients is adjusted because others may not affect the residuals; like, in our case, we cannot observe errors in the zonal coefficients in the XDs. We limit ourselves to those combinations of degree and order (l, m) that, according to LPT, produce a global root-mean-square (rms) XD of 2 mm or more, assuming a 1σ error in either JGM 3 coefficient \bar{C}_{lm} or \bar{S}_{lm} . This leads to a set of 550 pairs of \bar{C}_{lm} and \bar{S}_{lm} coefficients to be adjusted while other gravity coefficient errors are deemed to be too poorly observed in the XDs for any adjustment to be realistic.

Solving the gravity coefficient adjustments $\Delta \bar{C}_{lm}$ and $\Delta \bar{S}_{lm}$ from (5) and (6) leads to a set of linear equations with 45,000 observations and 1100 unknowns. Because the number of observations far exceeds the number of unknowns, we solve the unknowns

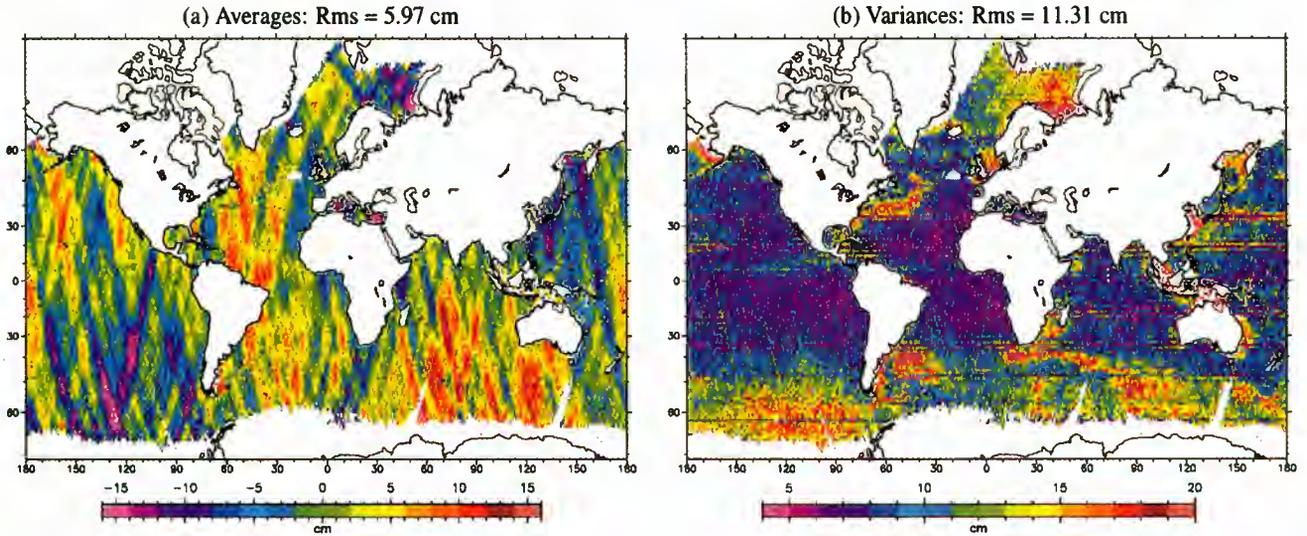


Plate 1. (a) Averages and (b) variances of stacks of ERS-1 and ERS-2 single-mission altimeter crossover height differences for the period of April 1992 till August 1995. Gravity model JGM 3. Apparent altimeter time tag bias is applied.

in a Bayesian linear least squares fashion

$$(A^T W A + f N) \mathbf{c} = A^T W \mathbf{x} \tag{7}$$

where \mathbf{c} is the vector of unknowns ($\Delta \bar{C}_{lm}, \Delta \bar{S}_{lm}$); \mathbf{x} is the vector of average XDs ($\bar{x} \approx \delta x$); A is the matrix of partials ($2\bar{Q}_{lm}^s \sin m\lambda, -2\bar{Q}_{lm}^s \cos m\lambda$); W is a diagonal matrix of observation weights, given by $n/(\sigma_x)^2$, where n is the number and σ_x is the variance of XDs in each stack; and fN is the JGM 3 normal matrix N (= inverse of the error variance-covariance matrix, kindly provided by John Ries, CSR) multiplied by a weighting factor f .

Because we only have observations over oceans and only up to a certain latitude, we need to constrain the solution elsewhere. This is done by adding the relevant part of the JGM 3 normal matrix to the normal equations. The factor f further determines how tight

the solution is to be constrained. The optimal value for f (=10) was found experimentally, weighing between an almost total reduction of the geographically anticorrelated orbit error but unrealistically large gravity field adjustments (small f) and small adjustments but insignificant reduction of the orbit error (large f).

Plate 2 shows the adjustments to the 1100 \bar{C}_{lm} and \bar{S}_{lm} coefficients of JGM 3 divided by their formal standard deviations (Δ/σ). The rms of all values Δ/σ is 0.52, which means that the new solution falls well within the error budget of JGM 3. Largest adjustments are to some coefficients around orders 16, 33, and 41 but are still within acceptable limits. Thus we have generated the ERS-tailored Delft Gravity Model DGM-E04.

With (5) and (6) it is quite easy to determine the global distribution of the geographically anticorrelated orbit error implied by the coefficient adjustments, as shown in Plate 3a. Note that the implied

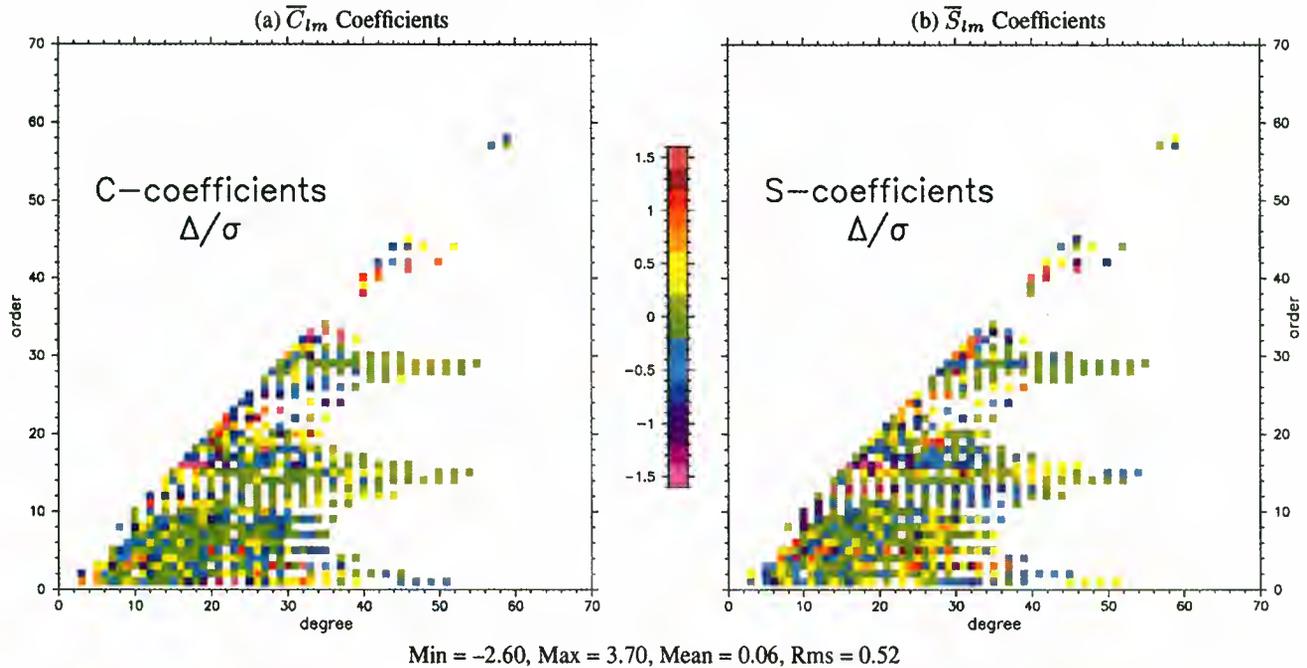


Plate 2. Estimated gravity coefficient adjustments to JGM 3 to obtain DGM-E04. All adjustments (Δ) are normalized through division by the formal error (σ) of the original JGM 3 coefficients.

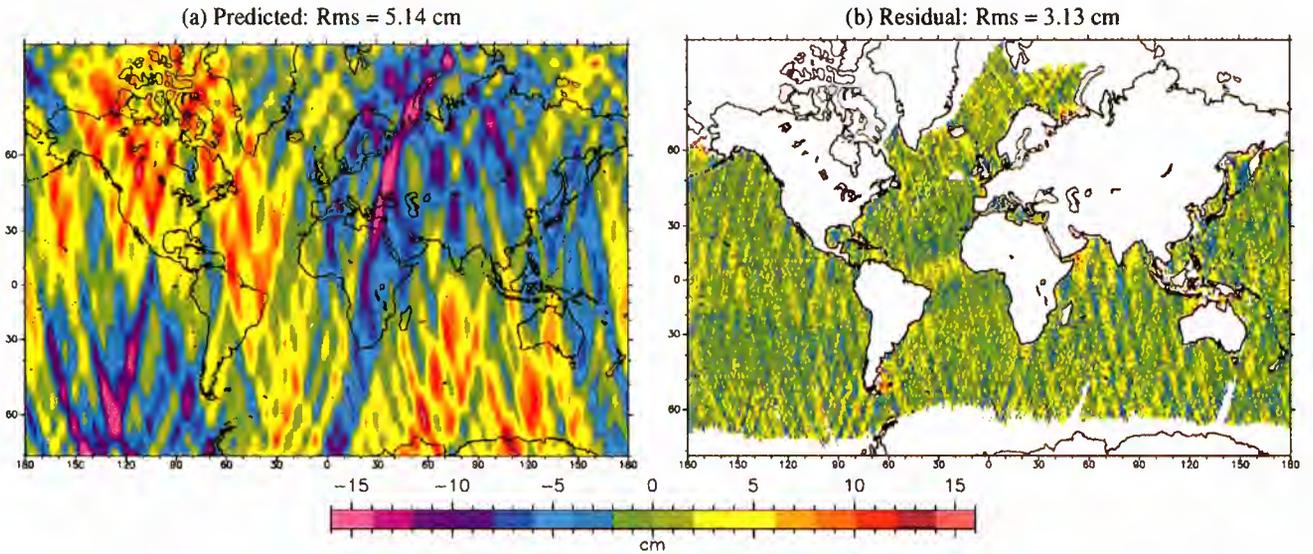


Plate 3. (a) Geographically anticorrelated orbit error implied by the differences between the JGM 3 and DGM-E04 gravity models. Values are multiplied by 2 to resemble the average XDs in Plate 1a. (b) Residual average XDs when subtracting Plate 3a from the average XDs in Plate 1a.

errors over the continents are not significantly different from ocean areas, which again indicates that the constraint was successfully applied. From this it is an easy step to project how the remaining average XDs would look when replacing JGM 3 for DGM-E04 in the POD (Plate 3b). It reduces the global rms of the average XDs from 5.98 to 3.13 cm, which implies that the geographically anticorrelated radial orbit error reduces from about 3.0 to 1.6 cm.

3.4. Independent Assessment of DGM-E04

Since the French Earth observation satellite SPOT-2 is in a similar orbit as ERS-1 and ERS-2, orbit determination ought to also improve for this satellite when adopting DGM-E04 instead of JGM 3. The only means of verification is, in this case, the data fits of the Doppler orbitography and radiopositioning integrated by satellite (DORIS) tracking data. Because these range-rate measurements are mostly sensitive to along-track orbit errors, it gives us some feel of the orbit error in this direction as well, in contrast to the vertical sense of ERS altimeter crossovers. The observed reduction of rms of the DORIS data fits from 0.70 to 0.66 mm s⁻¹ is encouraging.

Frank Lemoine (GSFC, personal communication, 1997) tested the suitability of DGM-E04 for TOPEX orbit determination. Having a completely different orbit from ERS, its sensibility to the gravity field is likewise dissimilar. This means that adjustments to some coefficients that had little effect on the ERS orbit might produce significant errors in the TOPEX orbit. Table 3 shows the results of computing a random set of four TOPEX orbits based on SLR and DORIS tracking data and each using three different gravity models: JGM 2, JGM 3, and DGM-E04. The data fits appear least sensitive to the choice of gravity model. Comparisons of the SLR/DORIS orbits with Global Positioning (GPS) reduced-dynamic orbits produced by the Jet Propulsion Laboratory (JPL), however, show a clear preference for the DGM-E04 model. Because the GPS reduced-dynamic orbits are virtually unaffected by gravity model errors, this test would indicate that the one that comes closest has the least gravity-induced orbit error. In 11 out of the 12 cases it is the DGM-E04 orbit that fits the GPS orbit best in radial, cross-track, and along-track direction.

Richard Rapp (OSU, personal communication, 1997) verified the lower degree and order coefficients (up to 14) of DGM-E04 by applying it as a long-wavelength reference geoid for the extraction

of a global ocean dynamic topography from TOPEX/POSEIDON sea surface heights. The resulting topography differs from the numerical ocean model POCM-4B by 12.55 cm rms. The same analysis with JGM 3 gives a similar rms residual of 12.62 cm. This suggests that for this application, DGM-E04 indeed acts within the error margins of JGM 3.

4. Gravity-Induced Orbit Errors

The third-generation orbits (section 2.3) cover the entire ERS Tandem Mission, starting on April 29, 1995, with the switch on of

Table 3. Statistics of TOPEX Orbit Computation

Gravity Model	Data Fits		Rms Difference With GPS Arc, cm		
	SLR, cm	DORIS, mm s ⁻¹	Radial	Cross	Along
<i>Cycle 10</i>					
JGM 2	4.58	0.580	3.02	6.85	11.89
JGM 3	4.60*	0.577*	2.18	6.45	8.97
DGM-E04	4.63	0.577*	2.14*	5.99*	8.28*
<i>Cycle 19</i>					
JGM 2	4.42	0.551	3.71	5.17	10.51
JGM 3	4.24*	0.546*	3.00	3.52	7.66
DGM-E04	4.29	0.546*	2.98*	3.43*	7.45*
<i>Cycle 21</i>					
JGM 2	3.05	0.539	3.20	5.40	7.93
JGM 3	2.86*	0.537*	2.46*	5.03	5.86
DGM-E04	2.89	0.537*	2.48	4.73*	5.83*
<i>Cycle 46</i>					
JGM 2	3.00	0.565	2.48	6.76	7.80
JGM 3	2.73*	0.563*	1.74*	4.53	5.11
DGM-E04	2.74	0.563*	1.74*	3.63*	4.76*

TOPEX orbit computations with the JGM 2, JGM 3, and DGM-E04 gravity models was performed at GSFC (Frank Lemoine, personal communication, 1997). Parameterization of the orbits is the same as in the second-generation TOPEX precise orbits, as documented by Marshall et al. [1995]. Listed are the SLR and DORIS tracking data residuals and orbit differences with the JPL reduced-dynamic orbits based on GPS tracking data.

*Lowest values for each cycle.

Table 4. Results of the Simultaneous ERS Orbit Determination for 104 Orbital Arcs of the Tandem Mission

Data Type	Number	Orbit Solution		
		JGM 3	EGM96	DGM-E04
<i>ERS-1</i>				
SLR residuals (rms), cm	112,144	5.63	5.59	4.94
XD residuals (rms), cm	196,325	9.58	9.16	8.06
Altimeter time tag bias (mean±variance), ms	104	-1.24±0.21	-1.45±0.20	-1.52±0.19
<i>ERS-2</i>				
SLR residuals (rms), cm	110,428	5.49	5.50	4.78
XD residuals (rms), cm	187,212	9.62	9.20	8.07
Altimeter time tag bias (mean±variance), ms	104	-1.05±0.22	-1.25±0.23	-1.30±0.22
<i>ERS-1 Minus ERS-2</i>				
Altimeter range bias (mean±variance), cm	104	-2.14±0.65	-2.14±0.65	-2.14±0.65

The XD residuals pertain to each single component of a XO, so each XO is counted twice and no differentiation is made between single- and dual-satellite XOs. The altimeter time tag biases and relative range bias are estimated independently for each orbital arc. Only arcs without significant orbital maneuvers are considered.

the ERS-2 altimeter and ending on June 2, 1996, with the switch off of the ERS-1. We have chosen this period to compare the results obtained with the DGM-E04, JGM 3, and EGM96 gravity models and to analyze the respective orbit accuracies. Because only 30% of the Tandem Mission overlaps with the data period of the DGM-E04 tailoring, the verification can be considered to be nearly independent. Naturally, we expect the gravity model tuning to have the same effect on the orbits for this particular period as on any other.

In this section we will first show the impact of the gravity model selection on the tracking data residuals, followed by an assessment of the geographically correlated orbit errors based on ERS single- and ERS/TOPEX dual-satellite XOs, and finally look at the orbit differences. In section 5, collinear altimeter profiles are analyzed to determine the orbit errors due to nonconservative forces.

4.1. Precise Orbit Determination Results

The tracking data residuals (measured minus computed SLR ranges and crossover height differences) are a measure for the orbit accuracy. These residuals should be interpreted with care, since the data that have been used in the POD are likely to underestimate the

actual orbit error. Moreover, the SLR range residuals are a measure of the orbit accuracy in all three directions, whereas the XDs depend on the radial orbit error only.

Table 4 presents the results for 104 arcs of 5.5 days, in which ERS-1 and ERS-2 orbits are determined simultaneously, alternatively computed with the JGM 3, EGM96, and DGM-E04 gravity models. Listed are the statistics of the SLR range residuals, the statistics of the satellite-specific components of the XD residuals (so each single and each dual are counted twice), and the estimates of the apparent altimeter time tag bias and relative altimeter range bias.

Note that the XD residuals are clearly the smallest in the DGM-E04 solutions. This is to be expected since XDs have been used to tailor the gravity model, starting from JGM 3. More relevant is the fact that DGM-E04 even performs better than the new EGM96 model, not only in terms of the rms XD but also judging from the SLR residuals, and this while SLR data have not been used to tune the model.

The POD also provides estimates for deficiencies in the datation of the altimeter data. Although the estimated parameters could also

Table 5. Statistics of Apparent SLR Range Biases of Near-Vertical Overflights

Site	Location	Number		JGM 3			EGM96			DGM-E04		
		of Passes		Mean			Mean			Mean		
		Asc	Des	Asc	Des	Var	Asc	Des	Var	Asc	Des	Var
7090	Yarragadee, West-Australia	69	46	-1.0	0.4	1.7	-1.0	0.8	1.4	-0.2	0.7	1.3
7105	Greenbelt, Maryland	23	13	-0.7	-1.4	1.5	-2.4	-1.0	2.0	-0.6	0.0	1.4
7109	Quincy, California	34	15	-0.2	3.0	2.1	-0.4	3.4	1.9	0.2	0.2	1.6
7110	Monument Peak, California	74	38	0.7	1.9	2.7	0.1	2.7	3.1	0.9	-0.5	2.7
7403	Arequipa, Peru	59	30	-0.9	-0.7	1.5	-2.5	0.3	2.2	0.5	2.6	1.9
7835	Grasse, France	38	7	-4.0	-9.2	2.1	-2.7	-5.8	2.4	-3.1	-5.2	1.8
7836	Potsdam, Germany	89	5	-3.3	-4.8	2.5	-2.9	-2.1	2.1	-2.4	-1.7	2.3
7839	Graz, Austria	61	18	-3.9	-5.5	1.8	-3.0	-4.8	2.1	-2.9	-2.7	1.8
7840	Herstmonceux, England	60	42	-1.7	-2.7	1.8	-1.6	-1.7	1.9	-2.0	-1.6	1.6
7843	Orroral Valley, Victoria	10	9	3.0	5.1	2.7	1.0	3.5	1.8	-0.4	4.6	1.6
Total (rms)		517	223	2.4	4.3	2.0	2.0	3.1	2.1	1.7	2.6	1.8
(Des+Asc)/2 (rms)				3.3				2.4			2.0	
(Des-Asc)/2 (rms)				1.2				1.1			1.0	

The apparent range bias is estimated simultaneously with a timing bias for each pass with at least eight measurements before and after the culmination point and a highest elevation of at least 65°. For all three orbit solutions the average range biases along ascending passes (Asc) and along descending passes (Des) and the variance (Var) of the range biases around the respective means are shown. Values are in centimeters.

absorb part of the 2-cpr errors in the restituted orbital altitude as well as constant along-track errors, the larger part will indeed be related to the altimeter datation. Table 4 clearly shows that the OPR altimeter time tags are systematically early by 1.3 to 1.5 ms for ERS-2 and ERS-1, respectively.

4.2. High-Elevation SLR Passes

Despite the aforementioned limitations, the SLR residuals of high-elevation overflights of laser ranging stations can provide at least some indication of the orbit's vertical accuracy. Along near-vertical overflights the cross-track orbit error does not affect the residuals, and system noise and along-track orbit error can be separated from the radial by fitting a two-parameter curve through the SLR residuals. The two parameters, range bias and timing bias, more or less directly relate to the radial and along-track orbit error. Table 5 lists the statistics of these apparent range biases for near-vertical overflights of some high-performance SLR stations.

When we distinguish between ascending and descending passes, the average of the apparent range biases per station becomes a measure for the local geographically correlated radial orbit error along each pass, i.e., $-\delta r^{\text{asc}}$ and $-\delta r^{\text{des}}$, where the minus signs come from the fact that the SLR range residuals are "observed minus computed" and the orbit errors are "computed minus true." The station-by-station averaging removes the time variant part of the orbit error associated with nonconservative forces (listed in Table 5

under "Var") and isolates the time invariant geographically correlated part (under "Mean").

The mean of a station's ascending and descending averages would resemble the geographically fully correlated orbit error ($-\delta r^c$), if it were not for constant system biases and station coordinate errors persisting in this value. Computing half the difference between the descending and ascending averages eliminates the constant errors and provides a better estimate for the geographically anticorrelated orbit error ($-\delta r^a$). The rms of these linear combinations are listed at the bottom of Table 5 and are indicative of the improvement of DGM-E04 over JGM 3 and even over EGM96 but should not be regarded as an accurate representation of the actual geographically correlated orbit error.

4.3. Geographically Anticorrelated Orbit Error

To demonstrate that the reduction of the geographically anticorrelated orbit error is also apparent in the actual computed orbits for the Tandem Mission, we have generated local averages and variances of XDs like in section 3.2 and alternatively took the GFZ PGM055 orbit from the OPR data or substituted our third-generation orbits. In each case the $\sim 870,000$ XOs covering the Tandem Mission are reduced to some 48,000 locations.

The statistics and graphs in Plate 4 indeed show that the average XDs are by far the smallest with DGM-E04. Both PGM055 and,

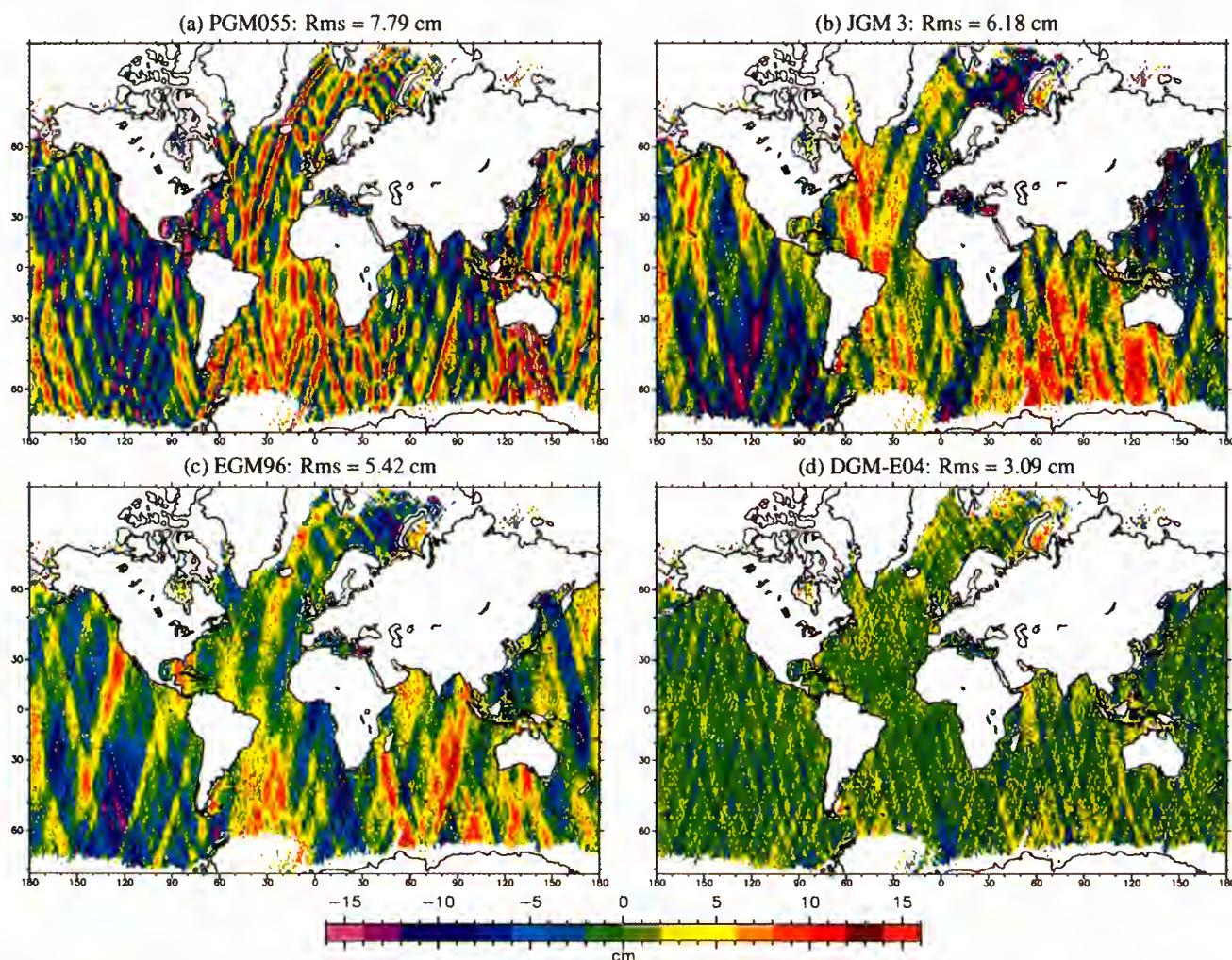


Plate 4. Averages of stacks of ERS-1 and ERS-2 single-mission XDs (Tandem Mission). The various graphs pertain to the following different ERS orbit solutions: (a) GFZ PGM055 and (b) DUT JGM 3 and (c) EGM96 and (d) DGM-E04. In each case the best fitting apparent altimeter time tag bias is applied.

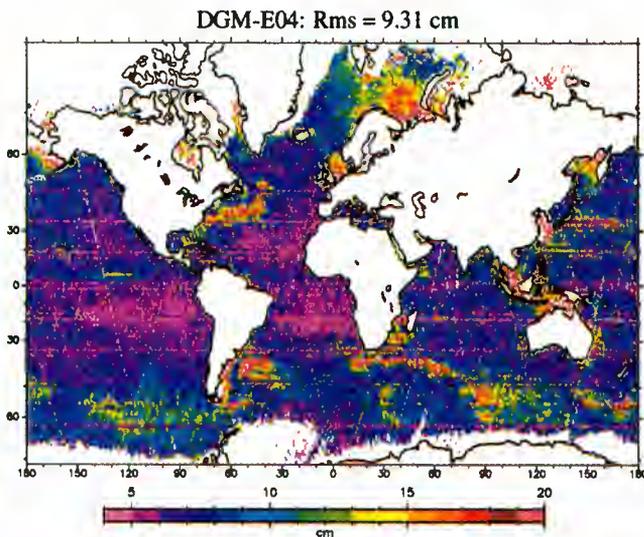


Plate 5. Local variance of stacks of ERS-1 and ERS-2 single-mission XDs based on DGM-E04 orbits (Tandem Mission).

to a lesser extent, JGM 3 display large geographically correlated orbit errors in distinctive narrow patterns, 10° to 20° wide. These patterns can be associated with errors in the near-resonant terms of the gravity field. These terms appear slightly better conditioned in the EGM96 model, in which broader features, however, still persist. In the DGM-E04 solution all of these features are absent. What remains is a chaotic pattern of small spatial scales of which the origin is not clear.

The rms of the geographically anticorrelated radial orbit error equals half the rms of the average XDs displayed in Plate 4. Note

that the result for DGM-E04 is virtually the same as was predicted in section 3.3 and Plate 3b.

The variance of the XDs displayed in Plate 5 is for DGM-E04 but is quite indifferent between the various orbit solutions. The larger part of the variance can be attributed to sea level variability.

4.4. Geographically Fully Correlated Orbit Error

Until now we have mainly considered the geographically anticorrelated orbit error, which is observed in single-satellite XDs and is used to tune the gravity model. But, as a rule of thumb, the fully correlated orbit error (δr^c) which is invisible in XDs has about the same variance [Rosborough, 1986]. To verify whether the tuning also reduced the geographically fully correlated orbit error, or at least did not increase it, we have computed ERS/TOPEX dual-altimeter crossovers for the same period during the Tandem Mission as before. The TOPEX data are extracted from the recent release of merged geophysical data records (cycles 98–136) [Archiving, Validation and Interpretation of Satellite Oceanographic data (AVISO), 1996], applying the appropriate corrections.

A justifiable assumption is that TOPEX orbit errors are small compared to those of ERS [Le Traon et al., 1995; Marshall et al., 1995], such that ERS/TOPEX dual satellite XDs display ERS' orbit errors rather than TOPEX's. Again, local averaging of XDs can be used to root out time variant errors. Because ascending and descending passes of ERS do not meet in the same crossover with TOPEX passes, it is actually more practical to average on an equi-rectangular grid. The acquired average ERS/TOPEX XD is representative of the geographically fully correlated radial orbit error for ERS.

Plate 6 shows averages of ERS-1/TOPEX and ERS-2/TOPEX XDs for four different ERS orbit solutions: PGM055, JGM 3,

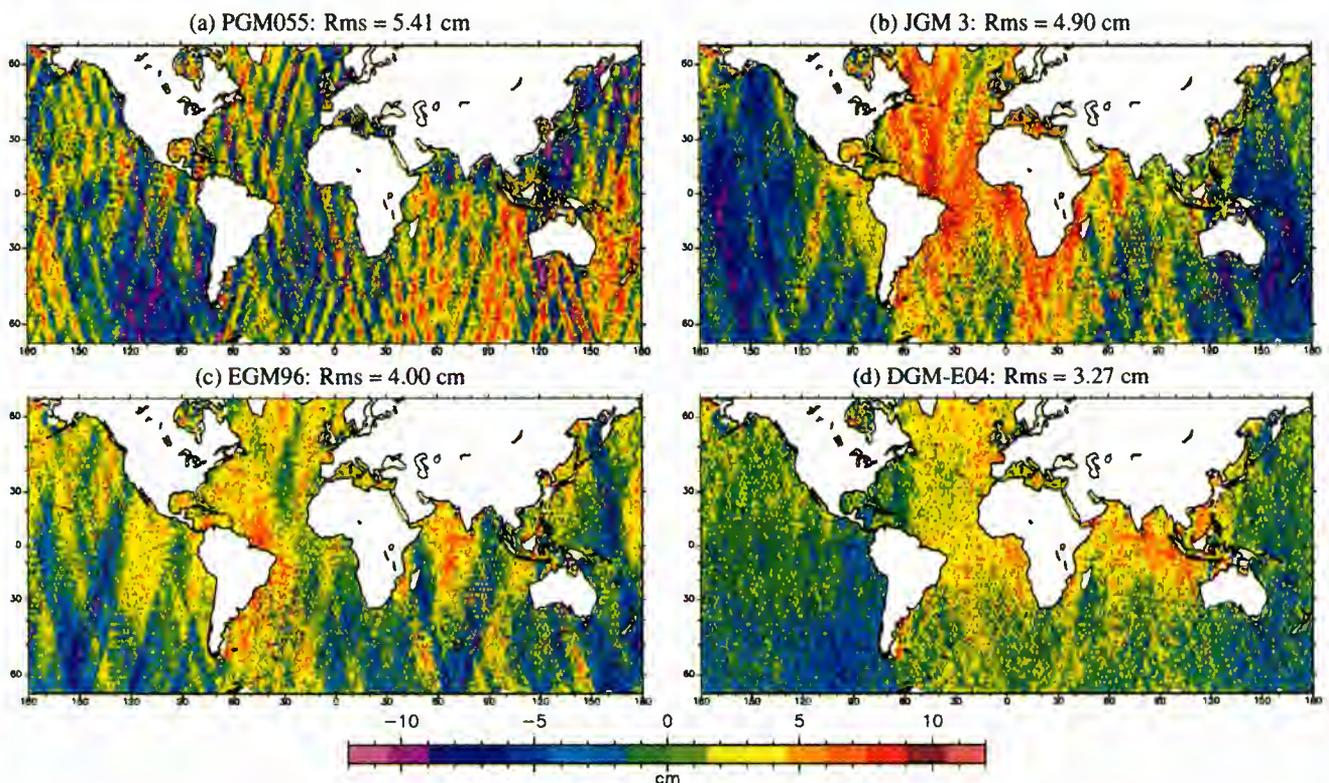


Plate 6. Locally averaged ERS–TOPEX dual-satellite crossover height differences for the ERS Tandem Mission. The graphs pertain to the following different orbit solutions for ERS: (a) GFZ PGM055 and (b) DUT JGM 3 and (c) EGM96 and (d) DGM-E04. In each case the best fitting apparent altimeter time tag and range biases for ERS-1 and ERS-2 is applied.

EGM96, and DGM-E04. Appropriate time tag biases and range biases for ERS-1 and ERS-2 are applied to give a good match with TOPEX. PGM055 is again the one to show the most evident track-like pattern associated with near-resonant terms. Striking is the broad structure that seems to persist throughout JGM 3, EGM96, and DGM-E04 orbit solutions. It even appears to be the only significant structure remaining in the DGM-E04 solution. The reduction of the variance from 5.41 (JGM 3) to 3.27 cm (DGM-E04) remains a remarkable achievement, since the fully correlated orbit error was not involved in the gravity field tailoring process. It demonstrates again that the tailored model indeed constitutes an improvement, not merely a reduction of one observable.

It is not certain that the remaining pattern for DGM-E04 is a true reflection of the actual geographically correlated orbit error of ERS. The ERS/TOPEX XDs will be partly corrupted by TOPEX orbit errors and any time invariant or geographically correlated difference between ERS and TOPEX altimetry. Even though we have attempted to harmonize the altimeter biases, sea state bias, and the ocean tide corrections, persistent differences in modeling propagation corrections may cause part of the effect. More likely though, what we observe is due to reference frame offsets or is caused by errors in a few (zonal) gravity coefficients that have not been adjusted in the development of DGM-E04.

Finally, we would like to emphasize that for many applications of altimetry, such as the monitoring of ocean currents and computation of marine gravity, the slope error is more important than the absolute error. This makes the DGM-E04 orbits even more favorable because slopes in the orbit error are almost an order of magnitude smaller than in the PGM055 orbits (order 0.2 versus 1 μrad , respectively).

4.5. Orbit Differences

Figure 3 depicts the predicted and observed spectra of the radial orbit differences between the three DUT orbit solutions. The predicted spectra are according to LPT and the differences in the

gravity field coefficients ($\Delta\bar{C}_{lm}, \Delta\bar{S}_{lm}$). The observed spectra are periodograms of the actual differences between the orbit solutions.

The JGM 3/DGM-E04 difference has a major peak at 0.93014 cpr (= 1 cpr - 1 cpd), and corresponds to $l - 2p = 1$ and $m = 1$ in (2), i.e., gravity coefficients of odd degree and order 1. This 1-cpr signal with a daily modulation is precisely the near-resonant signal we held responsible for the track-like patterns in the average XDs with JGM 3 (section 4.3). This peak is indeed markedly smaller in the EGM96/DGM-E04 difference, explaining the reduced track pattern in the average XDs with EGM96.

Because we have not excluded the near-resonant terms in the predicted spectra, they have large peaks close to 0 and 1 cpr caused by the harmonics of order 43. Having periods or a modulation longer than the length of the orbital arc, they are effectively absorbed by the state vector or the daily empirical along-track accelerations, as a result of which they do not show up in the observed spectra. The JGM 3/DGM-E04 spectra show a one-to-one match between predicted and observed. Near-resonant terms are absent here because they are simply copied from JGM 3 into DGM-E04. Orbit differences caused by nonconservative forces, with a more or less continuous distribution of power around 1 cpr, are minute.

Table 6 shows that the orbit solutions based on the three gravity models are quite close. Irrespective of which combination of DUT orbit solutions is compared, the rms orbit difference is around 4 cm in radial and around 15 cm in along- and cross-track direction. Because they are fully independent, the GFZ PGM055 and DUT DGM-E04 orbit solutions differ quite a bit more. Yet the radial orbit difference of only 7 cm is a very encouraging figure, likely to indicate that either orbit solution is at least as accurate as that. Encouraging is also that, coming from JGM 3, the EGM96 and DGM-E04 solutions appear to converge.

5. Nonconservative Forces

The dominant nonconservative forces acting on the satellite are atmospheric drag and solar radiation. Both forces are the sum of

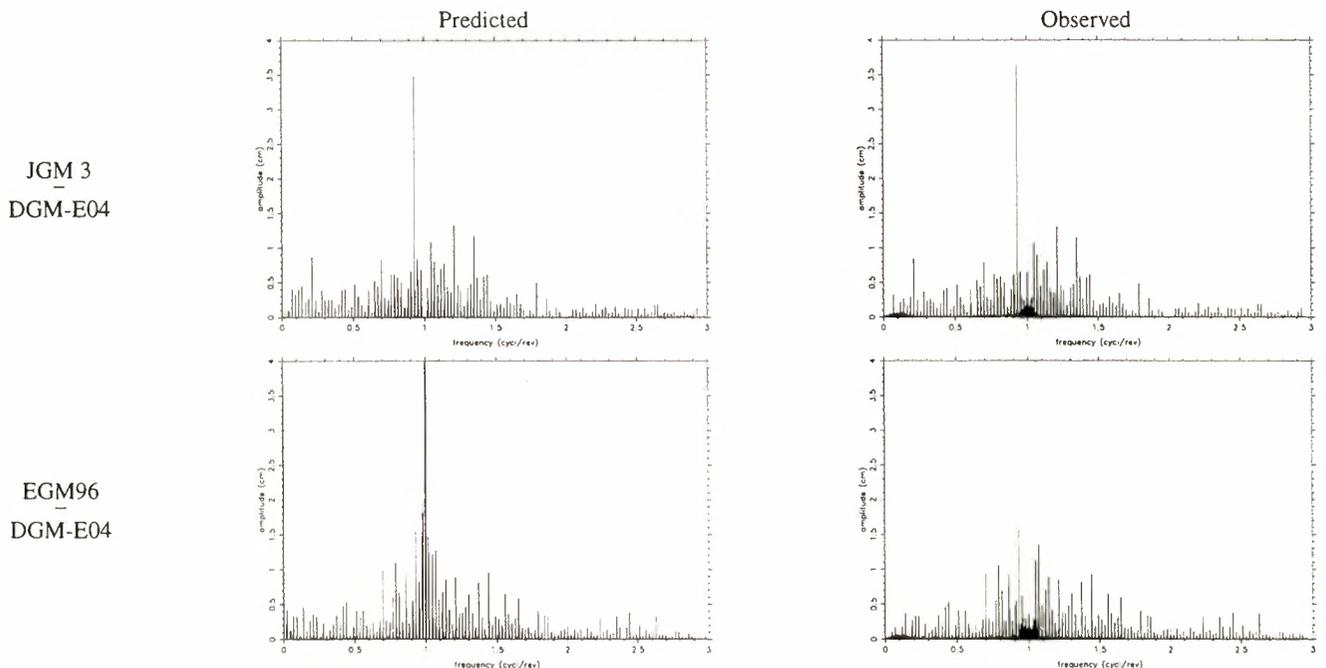


Figure 3. Predicted and observed spectra of the radial differences between ERS orbits computed with the JGM 3, EGM96, and DGM-E04 gravity models.

Table 6. Differences Between GFZ (PGM055) and DUT Orbit Solutions

Orbit Solutions	Orbit Differences rms			
	Radial	Cross	Along	Total
<i>ERS-1 Orbits</i>				
JGM 3 – EGM96	3.93	15.18	14.48	21.34
JGM 3 – DGM-E04	4.29	11.72	17.53	21.51
EGM96 – DGM-E04	3.81	16.04	14.55	21.99
PGM055 – JGM 3	8.00	19.79	28.50	21.34
PGM055 – DGM-E04	6.69	19.52	23.69	31.42
<i>ERS-2 Orbits</i>				
JGM 3 – EGM96	4.04	13.73	15.01	20.75
JGM 3 – DGM-E04	4.40	10.90	17.29	20.91
EGM96 – DGM-E04	4.03	15.33	14.47	21.46
PGM055 – JGM 3	8.30	17.53	29.19	21.34
PGM055 – DGM-E04	6.99	17.06	23.70	30.03

All values are in centimeters. A 3.5- σ editing is imposed on the total orbit difference to remove bad orbits.

fect of particles or photons exchanging momentum on the (quite complex) satellite surface, and this requires accurate modeling of the satellite geometry and surface properties.

5.1. Modeling of the Nonconservative Forces

The satellite geometry of ERS-1 and ERS-2 is modeled by two macromodels consisting of 10 flat panels. Six panels form the satellite bus and payload module, two panels form the front and aft side of the SAR, and a further two rotating panels model the front and

aft side of the solar array. At first each of the panels had been given a size, orientation, and reflective properties, which were representative of their true geometry and the properties of various subelements defined in a more detailed micromodel, consisting of 48 panels. A major advance was made by adjusting the properties of the panels of the macromodel, such that each of them might not conform to the true geometry or collective reflective properties of that panel but that their overall consistency with the micromodel in describing the actual forces acting on the whole satellite was optimal, judging from Monte Carlo ray-tracing experiments on both the macromodel and micromodel. Because of the attachment of the Global Ozone Monitoring Experiment (GOME), the ERS-2 macromodel differs slightly from the one of ERS-1.

Further deficiencies in the modeling of the nonconservative forces are absorbed by estimating 6-hourly drag coefficients, which scale the atmospheric drag forces, and a set of 22-hourly empirical accelerations with a 1-cpr modulation.

5.2. Collinear Tracks

As indicated before, gravity-induced orbit errors are the same along two collinear passes and cancel when differencing the two altimetric sea surface profiles. The remaining nonconservative forces cause orbit errors around 1 and 2 cpr and are easily separated from the short-scale sea height differences associated with measurement and correction errors and sea level variability.

Figure 4 gives an example of a pair of collinear tracks of ERS-1 and ERS-2 with a time interval of only 1 day. Figure 4a shows the locations of the measurements. Figure 4b gives the relative sea surface height profile with respect to the OSU MSS95 mean sea

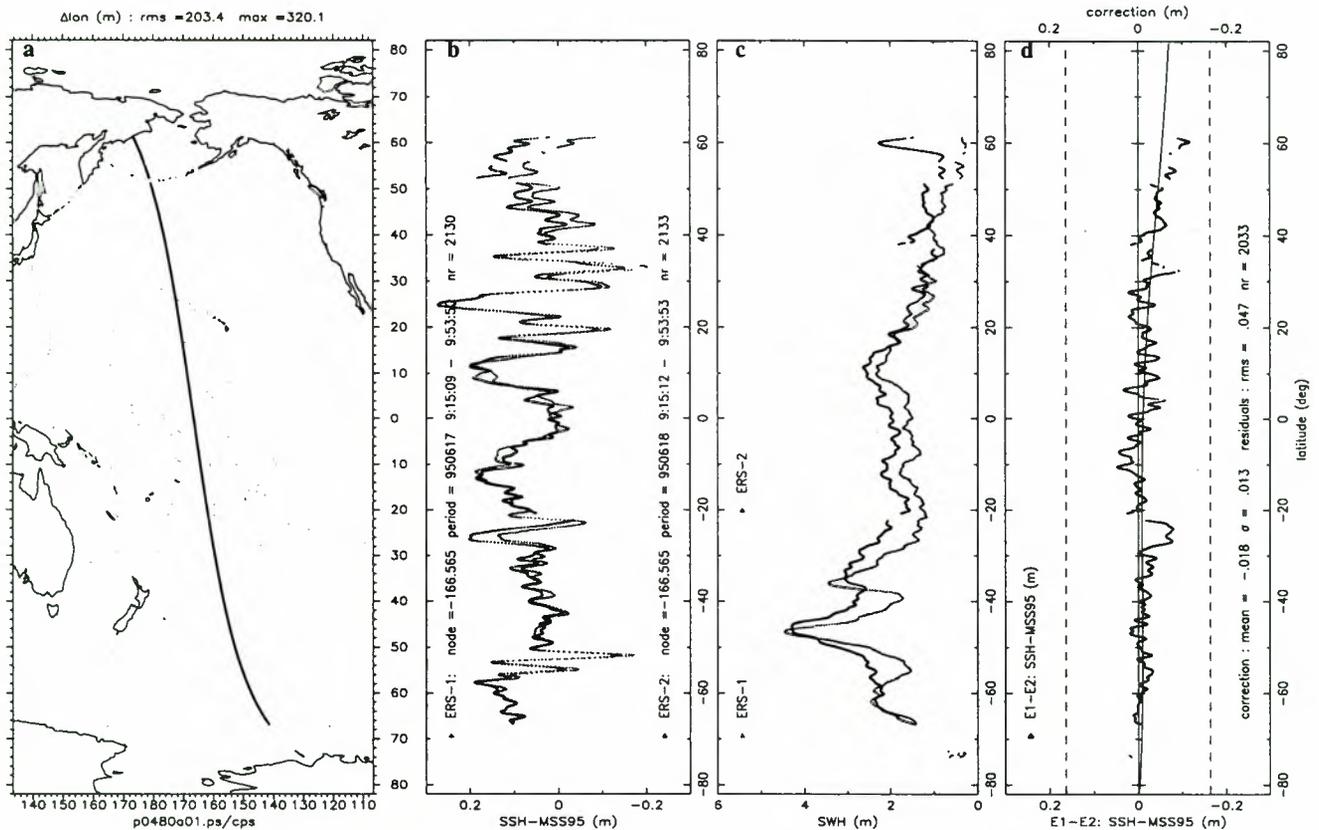


Figure 4. A collinear pair of sea surface height profiles of ERS-1 (grey) and ERS-2 (black) with a time interval of 1 day. Ascending pass 480, May 13-14, 1995. (a) Location, (b) sea surface height with respect to MSS95, (c) SWH, and (d) height difference, including five-parameter fit.

surface model. A running average filter is applied to remove the altimeter noise. Figure 4c gives the Significant Wave Height (SWH) for both collinear tracks derived by the two altimeters. Here, a running average filter is also applied. Figure 4d shows the residual difference between the smoothed sea surface profiles. A five-parameter orbit error model (displayed as the thin solid line) is fitted through the residuals and absorbs constant, 1-cpr and 2-cpr signals. The deviations from the fit are a result of sea level variability and errors in the geophysical corrections to the altimeter data and amount (in this case) to 4.7 cm rms; the orbit error model itself has an rms of 1.3 cm.

The top part of Table 7 gives the statistics of over 10,000 collinear pairs of ERS-1 and ERS-2 with a 1-day time interval. Obviously, because of the clear separation between orbit error and short-wavelength errors the rms value of the residual sea height differences (5.62 cm) is independent of the gravity model used in the orbit computation. The rms value of the five-parameter fits is significantly higher for the GFZ/PGM055 orbits than for the DUT (JGM 3, EGM96, and DGM-E04) orbits, which suggests that the modeling of the nonconservative forces at GFZ is less optimal. Yet, unexpectedly, also the choice of the gravity model appears to affect the long-wavelength fits slightly (2.87 cm for DGM-E04 and a few millimeters more for others). This indicates that a small part of the gravity-induced orbit error is aliased into the nonconservative force parameters (drag coefficients and empirical forces). When the aliasing would be purely geographically correlated or otherwise invariant from repeat cycle to repeat cycle, we would still not detect it in the collinear track differences. Elimination of a larger part of the gravity induced orbit error thus appears to reduce the aliasing and makes the absorption of nonconservative forces more efficient. Again, Table 7 demonstrates the superiority of DGM-E04 in this respect over the other gravity models.

When going from the short time interval of 1 day of dual-satellite collinears to a sizeable interval of 35 days of single-satellite collinears, the short-wavelength height differences are significantly larger (around 10.5 cm rms) because of the decorrela-

tion of correction errors (mainly wet tropospheric correction) and increasing sea level variability.

But, also, the orbit errors tend to be highly correlated over a 1-day time lag because ERS-2 then still senses about the same atmospheric conditions as ERS-1 1 day before. This could be used to an advantage by coupling the nonconservative force parameters between the two satellites during the orbit determination. For 35-day time intervals the rms value of the five-parameter fit thus increases to about 5.0 cm (for DGM-E04, Table 7). Assuming full decorrelation, this indicates that the radial orbit error due to nonconservative forces is $\sqrt{2}$ times smaller, 3.5 cm.

6. Error Budget

Table 8 summarizes the results obtained in the previous sections and lists the three components of the radial orbit error (geographically anticorrelated, fully correlated, and nonconservative) for each of the orbit solutions discussed in this paper. Table 8 also gives the contributions of the orbit error to XDs with a short time interval (typically up to 1 day) and a long time interval (typically 2 weeks or more). Note that the fully correlated part does not contribute to the XDs, that the anticorrelated part adds to the XDs at double efficiency, and that the nonconservative part has a certain decorrelation time.

After adding contributions for errors in the various altimeter geophysical corrections and sea level variability we obtain an overall budget for the rms XDs, ranging from the short to long time intervals. Plate 7 confirms these results and shows a dramatic reduction of the XDs from the PGM055 orbits to DGM-E04 orbits. The latter even coincides with the level of TOPEX, which is shown for reference and indicates that at this point the altimeter correction errors and sea surface variability have a significantly larger contribution to the rms XDs than the orbits. In conclusion, for the ERS Tandem Mission the radial rms accuracy of the GFZ PGM055 orbits is 7.7 cm; the DUT third-generation precise orbits have a radial rms accuracy of 6.8 cm (JGM 3), 6.0 cm (EGM96), and 5.0 cm (DGM-E04).

Table 7. Orbit Error Differences and Sea Height Residuals Between Collinear Pairs of ERS-1 and ERS-2

Orbit Solution	Number of Collinear Pairs	Number of Data Points	Orbit Error rms, cm	Height Residuals rms, cm
<i>ERS-1 Minus ERS-2 (1-Day Interval)</i>				
PGM055	10,103	14,699,314	3.69	5.62
JGM 3	10,193	14,826,690	3.09	5.62
EGM96	10,214	14,844,032	3.03	5.62
DGM-E04	10,220	14,861,823	2.87	5.62
<i>ERS-1 Minus ERS-1 (35-Day Interval)</i>				
PGM055	10,509	15,186,450	5.39	10.44
JGM 3	10,516	15,186,699	5.12	10.44
EGM96	10,537	15,211,211	5.11	10.44
DGM-E04	10,564	15,237,874	5.00	10.44
<i>ERS-2 Minus ERS-2 (35-Day Interval)</i>				
PGM055	8,586	12,156,457	5.92	10.55
JGM 3	8,825	12,485,460	5.31	10.56
EGM96	8,832	12,487,222	5.26	10.57
DGM-E04	8,862	12,522,016	5.11	10.57

Radial orbit error differences are based on a five-parameter fit of the sea surface height differences; height residuals are indicative of measurement and correction errors and sea surface variability. Pairs are edited out when the orbit difference exceeds 3.5 times the nominal.

7. Conclusions and Outreach

We have demonstrated the limitations of currently available gravity models for the computation of ERS orbits. Focusing on the local average (time invariant part) of XDs reveals structures that can be linked to errors in various gravity field coefficients, especially those that can be associated with orbit errors of frequencies close to 1 cpr. The PGM055 gravity model appears to be an unfortunate choice for the generation of the orbits on the OPR. This model causes geographically anticorrelated orbit errors much larger even than the older JGM 3 model.

The time invariant part of the XDs forms the basis for the development of an ERS-tailored gravity model DGM-E04 within the error margins of JGM 3. This technique has demonstrated its strength: XDs reduce exactly as predicted by the adjustments in the gravity coefficients. Also, the acquired gravity model performs significantly better in terms of SLR tracking residuals, improves orbit determination for SPOT-2, and is competitive with JGM 3 for TOPEX orbit determination. Crossing ERS and TOPEX altimetry demonstrates a reduction of the geographically fully correlated radial orbit error on top of the forced reduction of the anticorrelated part. Remarkably, the nonconservative force modeling errors also appear to have diminished in the DGM-E04 orbits compared to their JGM 3 counterpart. The stepwise improvement of the orbit

Table 8. Contribution of Orbit Errors, Altimeter Corrections, and Sea Surface Variability to the Error Budget of Altimetric Sea Surface Heights and Crossovers

Source	Height Budget (Crossover Budget)				
	GFZ/PGM055	JGM 3	EGM96	DGM-E04	TOPEX/JGM 3
<i>Radial Orbit Errors</i>					
Geographically anticorrelated*	3.9 (7.8–7.8)	3.1 (6.2–6.2)	2.7 (5.4–5.4)	1.5 (3.1–3.1)	1.0 (2.0–2.0)
Geographically fully correlated†	5.4 (0.0–0.0)	4.9 (0.0–0.0)	4.0 (0.0–0.0)	3.3 (0.0–0.0)	1.0 (0.0–0.0)
Nongravitational‡	3.8 (3.7–5.4)	3.6 (3.1–5.1)	3.6 (3.0–5.1)	3.5 (2.9–5.0)	2.0 (1.5–2.8)
Total	7.7 (8.6–9.5)	6.8 (6.9–8.0)	6.0 (6.2–7.4)	5.0 (4.2–5.3)	2.4 (2.5–3.4)
<i>Mesoscale Features (Common to All)</i>					
Dry tropospheric correction			1.0 (1.0–1.4)		
Wet tropospheric correction			1.4 (1.0–2.0)		
Ionospheric correction			1.0 (1.0–1.4)		
Ocean tides			3.0 (4.2–4.2)		
Solid Earth tides			0.5 (0.7–0.7)		
Sea state bias			1.0 (1.0–1.4)		
Instrument errors			1.0 (1.0–1.4)		
Sea surface variability			4.0 (2.0–5.6)		
Total			5.6 (5.2–7.8)		
<i>Total Error Budget</i>					
Total	9.5 (10.1–12.3)	8.8 (8.6–11.2)	8.2 (8.1–10.8)	7.5 (6.7–9.4)	6.1 (5.8–8.5)

The ranges within brackets relate to crossovers with a short time interval (1 day) to a long time interval (weeks). Values (in centimeters) are given for various ERS orbit solutions, and TOPEX results are given for reference. Values for TOPEX are from Marshall et al. [1995].

*On the basis of Plate 4.

†On the basis of Plate 6.

‡On the basis of Table 7.

determination for ERS seems now to have culminated in a radial rms accuracy of 5.0 cm for the DGM-E04 orbits.

7.1. Room for Improvement

When studying ERS/TOPEX dual-satellite XO differences, the geographically fully correlated orbit error appears to be larger than the anticorrelated. It is not yet certain that the observed differences are fully accountable to gravity model errors. When this is clarified, this data type can easily be included in the gravity model tailoring

process. Likewise, we can add SLR and finally PRARE data as well and use an averaging process similar to what is shown in section 4.2 to isolate the gravity-induced part of the residuals. This paves the way to include also the zonal gravity coefficients in the tuning process. With the long-term operation of ERS-1 and ERS-2, sufficient data will be available for a further gravity model tuning to these satellites, which is then equally suitable for their successor: EnviSat.

By far the largest contributor to the orbit error, however, is the time variant part. This suggests that there is margin for improvement of the surface forces (drag and solar radiation). The time variant part, however, also comprises the gravitational effect of solid Earth and ocean tides, which so far have not been considered. Yet this does not affect our results, since all time variant orbit errors are simply mingled into the one that we labeled “nonconservative” or “nongravitational.”

7.2. Remaining Considerations

A widely distributed legend argues that when orbit computations use the same gravity field for all altimeter satellites, there is no mismatch between the gravity-induced orbit errors and hence they cancel when differencing the different data sets. This is, however, not true. Because of their distinct inclination, repeat cycle, altitude, and choice for orbital arc length, gravity model errors impact differently on the computed orbital altitude. The best choice of gravity model is the one that introduces the least errors for each particular mission. Tailored models which are tuned to a particular satellite mission are therefore the best candidates for adoption in the POD, as long as these act within the error margins of a general-purpose model.

This is irrevocably demonstrated when differencing ERS and TOPEX altimetry in Plate 6 of section 4.4. Nearly the worst of all performance is given by using the same gravity model (JGM 3) for both satellites. The best match is obtained using DGM-E04

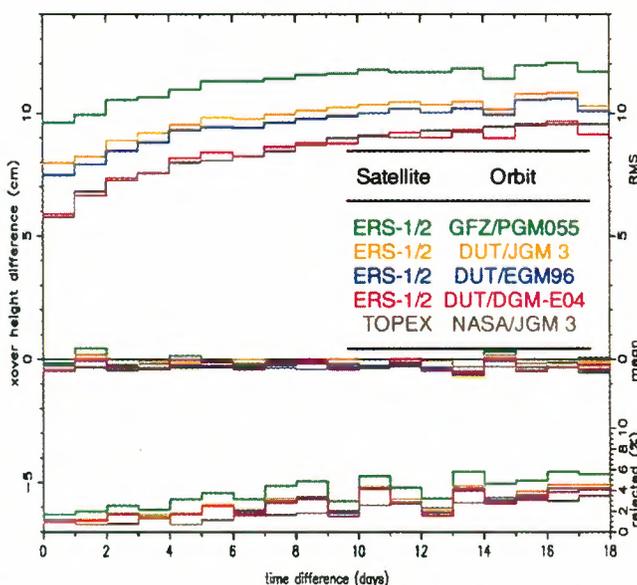


Plate 7. The rms and mean XD (ascending–descending) as a function of time interval. Data beyond 66° latitude and in high-variability areas are excluded. A 3.5-σ editing is applied.

for ERS and JGM 3 for TOPEX. In fact, JGM 3 is very much optimized for TOPEX, which is also reflected in the error variance-covariance matrix of JGM 3. The tailoring to ERS thus mainly concerns the weakly determined (lumped) coefficients that are best observed in ERS XDs. Consequently, for TOPEX orbit determination, DGM-E04 and JGM 3 are very similar, as was also demonstrated in section 3.4.

Acknowledgments. The ERS orbit determination at DUT would not have been such a success without the support from many of our colleagues. In particular, we like to thank Bob Cheney of NOAA for supporting the operational orbit determination financially and to colleagues at GSFC for providing numerous gravity models and our workhorse, the GEODYN orbit determination software. We also acknowledge the support of the precise orbit determination through the EC project ESAMCA, led by Duncan Wingham of UCL/MSSL. Furthermore, we are grateful to those whose inputs and discussions were very valuable throughout the years: C. K. Shum, John Ries, Pierre-Yves Le Traon, Carl Wagner, Michael Anzenhofer, and Arthur Smith. Special thanks go to John Lillibridge for hosting the DUT orbits on the NOAA ftp site and to Frank Lemoine (GSFC) and Richard Rapp (OSU) for their independent assessments of the DGM-E04 model.

References

- Aksnes, K., P. H. Andersen, S. Hauge, and H. Erlandsen, ERS-1 orbit calculation with the Geosat software, in *Proceedings of the Second ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-361*, pp. 741–745, 1994.
- Archiving, Validation, and Interpretation of Satellite Oceanographic data, AVISO user handbook: Merged TOPEX/POSEIDON products, *Rep. AVI-NT-02-101-CN*, edition 3.0, Cent. Natl. d'Etudes Spatiales, Toulouse, France, 1996.
- Barlier, F., C. Berger, J. Falin, G. Kockarts, and G. Thuillier, Atmospheric model based on satellite drag data, *Ann. Geofis.*, 34, 9–24, 1977.
- Centre ERS d'Archivage et de Traitement, Altimeter products, User manual, *C1-EX-MUT-A21-01-CN*, issue 2.6, Plouzané, France, 1994.
- Centre ERS d'Archivage et de Traitement, Altimeter and microwave radiometer ERS products, User manual, *C2-MUT-A-01-IF*, version 2.2, Plouzané, France, 1996.
- Engelis, T., Radial orbit error reduction and sea surface topography determination using satellite altimetry, *Rep. 377*, Dep. of Geod. Sci. and Surv., Ohio State Univ., Columbus, 1987.
- Engelis, T., On the simultaneous improvement of a satellite orbit and determination of sea surface topography using altimeter data, *Manuscr. Geod.*, 13, 180–190, 1988.
- Gaspar, P., and F. Ogor, Estimation and analysis of the sea state bias of the ERS-1 altimeter, *report of task B1-B2, IFREMER contract 94/2.426 016/C*, Collect., Localisation, Satell., Ramonville St. Agne, France, 1994.
- Gaspar, P., and F. Ogor, Estimation and analysis of the sea state bias of the new ERS-1 and ERS-2 altimeter data (OPR version 6), *report of task 2, IFREMER contract 96/2.246 002/C*, Collect., Localisation, Satell., Ramonville St. Agne, France, 1996.
- Gruber, T., A. Bode, C. Reigber, and P. Schwintzer, D-PAF global earth gravity models based on ERS, in *Proceedings of the Third ERS Symposium, Eur. Space Agency Spec. Publ., ESA SP-414*, in press, 1997.
- Kaula, W. M., *Theory of Satellite Geodesy*, Blaisdell, Waltham, Mass., 1966.
- Lemoine, F. G., et al., The development of the NASA GSFC and NIMA Joint Geopotential Model, in *International Association of Geodesy Symposia, Gravity, Geoid and Marine Geodesy*, vol. 117, pp. 461–469, Springer-Verlag, Berlin, 1997.
- Le Traon, P.-Y., and F. Ogor, ERS-1/2 orbit improvement using TOPEX/POSEIDON: The 2 cm challenge, *J. Geophys. Res.*, this issue.
- Le Traon, P.-Y., P. Gaspar, F. Bouysse, and H. Makhmara, Using TOPEX/POSEIDON data to enhance ERS-1 data, *J. Atmos. Oceanic Technol.*, 12, 161–170, 1995.
- Marsh, J. G., et al., The GEM-T2 gravitational model, *NASA Tech. Memo. 100746*, 91 pp., 1989.
- Marshall, J. A., N. P. Zelensky, S. M. Klosko, D. S. Chinn, S. B. Luthcke, and K. E. Rachlin, The temporal and spatial characteristics of TOPEX/POSEIDON radial orbit error, *J. Geophys. Res.*, 100(C12), 25,331–25,352, 1995.
- Massmann, F.-H., C. Reigber, R. König, and J. C. Raimondo, ERS-1 orbit information provided by D-PAF, in *Proceedings of the Second ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-361*, pp. 765–770, 1994.
- Nerem, R. S., et al., Gravity model development for TOPEX/POSEIDON: Joint Gravity Models 1 and 2, *J. Geophys. Res.*, 99(C12), 24,421–24,448, 1994.
- Rosborough, G. W., Satellite orbit perturbations due to the geopotential, *Rep. CSR-86-1*, 154 pp., Cent. for Space Res., Univ. of Tex. at Austin, 1986.
- Scharroo, R., K. F. Wakker, B. A. C. Ambrosius, R. Noomen, W. J. van Gaalen, and G. J. Mets, ERS-1 precise orbit determination, in *Proceedings of the First ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-359*, pp. 477–482, 1993a.
- Scharroo, R., K. F. Wakker, B. A. C. Ambrosius, R. Noomen, W. J. van Gaalen, and G. J. Mets, ERS-1 precise orbit determination, *Adv. Astronaut. Sci.*, 84, 293–307, 1993b.
- Scharroo, R., K. F. Wakker, and G. J. Mets, The orbit determination accuracy of the ERS-1 mission, in *Proceedings of the Second ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-361*, pp. 735–740, 1994.
- Scharroo, R., G. J. Mets, and P. N. A. M. Visser, TOPEX-class orbits for the ERS satellites (abstract), *Ann. Geofis.*, 14, suppl. 1, C 262, 1996a.
- Scharroo, R., P. N. A. M. Visser, G. J. Mets, and B. A. C. Ambrosius, TOPEX-class orbits for the ERS satellites (abstract), *Eos Trans. AGU*, 77(17), Spring Meet. Suppl., S76, 1996b.
- Schrama, E. J. O., The role of orbit errors in processing of satellite altimeter data, Ph.D. dissertation, Dep. of Geod., Delft Univ. of Technol., Delft, Netherlands, 1989.
- Schutz, B. E., B. D. Tapley, and C. Shum, Evaluation of the Seasat altimeter time tag bias, *J. Geophys. Res.*, 87(C5), 3239–3245, 1982.
- Shum, C. K., B. D. Tapley, B. J. Kozel, P. Visser, J. C. Ries, and J. Seago, Precise orbit analysis and global verification results from ERS-1 altimetry, in *Proceedings of the Second ERS-1 Symposium, Eur. Space Agency Spec. Publ., ESA SP-361*, pp. 747–752, 1994.
- Smith, A. J. E., and P. N. A. M. Visser, Dynamic and non-dynamic ERS-1 radial orbit improvement from ERS-1/TOPEX dual-satellite altimetry, *Adv. Space Res.*, 16(12), 123–130, 1995.
- Tapley, B. D., TOPEX/POSEIDON precision orbit determination standards, memorandum, Bur. of Eng. Res., Univ. of Tex. at Austin, 1992.
- Tapley, B. D., and G. W. Rosborough, Geographically correlated orbit error and its effect on satellite altimetry missions, *J. Geophys. Res.*, 90(C6), 11,817–11,831, 1985.
- Tapley, B. D., et al., The Joint Gravity Model 3, *J. Geophys. Res.*, 101(B12), 28029–28049, 1996.
- Visser, P. N. A. M., The use of satellites in gravity field determination and model adjustment, Ph.D. dissertation, Delft Univ. of Technol., Delft, Netherlands, 1992.
- Visser, P. N. A. M., ERS-1 precise orbit determination using TOPEX/ERS-1 dual satellite crossover differences, *Tech. Memo CSR-TM-93-07*, Cent. for Space Res., Univ. of Tex. at Austin, 1993.
- Visser, P. N. A. M., Gravity field model adjustment from ERS-1 and TOPEX altimetry and SLR tracking using an analytical orbit perturbation theory, *Adv. Space Res.*, 16(12), 143–147, 1995.

R. Scharroo and P. N. A. M. Visser, Delft Institute for Earth-Oriented Space Research, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, Netherlands (e-mail: remko.scharroo@lr.tudelft.nl; pieter.visser@lr.tudelft.nl)

(Received January 23, 1997; revised October 8, 1997; accepted November 5, 1997.)

Global marine gravity field from the ERS-1 and Geosat geodetic mission altimetry

Ole Baltazar Andersen and Per Knudsen

Kort- og Matrikelstyrelsen, Copenhagen, Denmark

Abstract. Satellite altimetry from the Geosat and the ERS-1 Geodetic Missions provide altimeter data with very dense spatial coverage. Therefore the gravity field may be recovered in great detail. As neighboring ground tracks are very closely distributed, cross-track variations in the sea surface heights are extremely sensitive to sea surface variability. To avoid errors in the gravity field caused by such effects, sea surface variability needs to be carefully eliminated from the observations. Initially, a careful removal of gross errors and outliers was performed, and the tracks were fitted individually to a geoid model and crossover adjusted using bias and tilt. Subsequently, sea surface heights were gridded using local collocation in which residual ocean variability was considered. The conversion of the heights into gravity anomalies was carried out using the fast Fourier transform (FFT). In this process, filtering was done in the spectral domain to avoid the so-called “orange skin” characteristics. Comparison with marine gravity was finally carried out in three different regions of the Earth to evaluate the accuracy of the global marine gravity field from ERS-1 and Geosat.

1. Introduction

Satellite altimetry from the geodetic missions of the Geosat and the ERS-1 provides the opportunity for geodesists to make very detailed mappings of the marine gravity field. The newly released Geosat geodetic mission altimetry contains data from the first 18 months of the Geosat satellite mission (1985-1986) covering marine regions at latitudes between -72° and 72° . The satellite was operated in a nonrepeating orbit yielding a very dense, though not completely homogeneous coverage of observations. ERS-1 completed its 336-day geodetic mission in March 1995. Altimetry from this mission covers all oceans between -82° and 82° latitude and provides a very dense and homogeneous coverage. At the equator the spacing between the ERS-1 satellite's ground tracks is constantly 8.3 km.

Mapping of the Earth gravity fields from different data sources has previously been presented by, e.g., *Haxby* [1987], *Balmino et al.* [1987], *Sandwell* [1992], *Sandwell and Smith* [1997], *Knudsen* [1991], *Knudsen et al.* [1992], *Tscherning et al.* [1993], and *Andersen et al.* [1995]. In this study the method used by *Andersen and Knudsen* [1995, 1996] is used with special emphasis on the latest improvement in data availability and in the determination of the global marine gravity field.

2. Data Editing and Initial Data Processing

The altimeter data from both geodetic missions are obtained as the usual 1 s mean values with an along-track spacing of about 7 km. To enhance the quality of the altimetry, both data sources applied the newest set of orbits similar for the two geodetic missions. Here Geosat applied the new recomputed orbits based on the JGM-3 gravity model, [*Tapley et al.*, 1996] and ERS-1 applied the JGM-3 orbits provided by the Delft Institute of Technology. The observations were corrected for

atmospheric path delay using values provided in the ERS-1 Ocean Product Records (OPR) and in the Geophysical Data Records (GDR) for Geosat for wet and dry tropospheric path delay and ionospheric path delay. Similarly, the provided correction for Earth tides was applied. Finally, the ocean and load tide contributions were updated using the global *Andersen-Grenoble* (AG95.1) ocean tide model [*Andersen et al.*, 1995]. The AG95.1 model was recently investigated by *Shum et al.* [1997]. It was chosen as it has been carefully tapered onto the Finite Element Solution (FES94.1) ocean tide model [*Le Provost et al.*, 1994] around the 66° parallels to avoid discrepancies at this latitude. Such discrepancies can be found in many recent global ocean tide models based on TOPEX/POSEIDON altimetry (e.g., the Center of Space Research (CSR3.0) model [*Shum et al.*, 1997]).

Initially, data were removed if any of the applied range corrections were absent except the ocean tide correction. Subsequently, sea surface height observations differing by more than 10 m from the joint National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC) and National Imagery and Mapping Agency (NIMA) Earth Gravity Model (EGM96) geoid model complete to degree and order 360 [*Lemoine et al.*, 1997] were removed. Finally, data were removed if the standard deviation of the height observations exceeded 0.3 m. This resulted in about 30 million altimeter data from the Geosat and about 20 million altimeter data from the ERS-1 geodetic missions.

From this point on all subsequent processing and derivation of the gravity field was carried out in small cells of the size of 2° latitude by 10° longitude. Globally, this mosaic has 80 times 36 cells (latitude by longitude) equivalent to 2880 cells. However, observations were extracted in cells extended by a border zone of 0.5° latitude by 1° longitude to a size of 3° latitude by 12° longitude. The selection of such small subareas was essential to the modeling of orbit errors and sea surface variability. A choice of larger cells caused problems in the removal of these signals, while smaller cells may corrupt parts of the regional geoid signal.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02198.
0148-0227/98/97JC-02198\$09.00

Initially, the number of tracks was limited at high latitude where the track distance becomes extremely narrow. Here, the number of tracks can be limited without much loss of information. Therefore only the altimeter data along the 800 longest tracks were selected.

To reduce the effects of residual orbit errors and sea surface variability, the tracks were then fitted individually to the joint NASA GSFC and DMA geoid model EGM96 by estimating bias and tilt terms to each track, thus removing all signals with a wavelength longer than about 3° - 4° . Subsequently, a crossover adjustment of the tracks was carried out, also using bias and tilt terms [e.g., *Knudsen and Brovelli, 1993*]. To avoid problems with rank deficiency, a minimum variance criterion was applied.

Hereafter, each height observation was compared with an estimated height to detect outliers and gross errors. As in work by *Tscherning [1990]*, the estimate was obtained using least squares collocation. That is,

$$\hat{h}_p = \mathbf{c}_p^T (\mathbf{C} + \mathbf{D})^{-1} \mathbf{h}, \quad (1)$$

where \mathbf{C} and \mathbf{D} are signal and error covariance matrices, \mathbf{c}_p is a vector of the covariances between the estimate and the observations, and \mathbf{h} contains the 20 nearest observations (5 in each quadrant around the estimation location). The observation itself entered the estimation with its standard deviation increased to 1 m. As a covariance function, a second-order Markov covariance function is used like

$$c(r) = C_0 \left(1 + \frac{r}{\alpha}\right) e^{(-r/\alpha)}, \quad (2)$$

where r is the lag and C_0 is the signal variance. The parameter α was fixed so that the correlation length (where a 50% correlation is obtained) was 30 km to ensure a relatively smooth interpolation.

Then a screening for gross errors, e.g., observations affected by sea ice, was carried out. This was done by removing observations with a discrepancy between the observed and the estimated height larger than 1 m.

Finally, the fitting to the geoid model EGM96 and the crossover minimization were repeated, this time using the original data edited for outliers and gross errors that might have affected the initial adjustment. Signals of wavelengths longer than 3° - 4° are removed through this process. Consequently, most of the permanent parts of the sea surface topography have been reduced.

3. Mapping of the Gravity Field

The estimation procedure for the marine gravity field was chosen to derive a global map of the gravity field within a reasonable computational effort. Because of the very large amount of observations, rigorous methods such as least squares collocation is not feasible. Hence a less rigorous but very efficient approximation method based on the fast Fourier transform (FFT) was chosen, though it requires that the data are distributed in a regular grid. The result may be sensitive to cross-track gradients caused by sea surface variability arising as the distance between parallel tracks becomes very small in the geodetic mission altimetry. Such effects may be reduced by using altimetric slopes or second-order derivatives [e.g., *Hwang and Parsons, 1995; Rummel and Haagmans, 1990*]. However, it was decided to stay with the heights as the observations and elaborate existing software in the GRAVSOFT package [*Tscherning et al., 1992*] for gridding and FFT manipulations.

Then the edited and adjusted altimeter data within each extended cell used for processing (3° latitude by 12° longitude) were interpolated onto a regular grid using the collocation technique described. To filter out remaining sea surface variability that may cause erroneous cross-track gradients between parallel tracks, an additional covariance function for this error was introduced. This error covariance function was applied to observations on the same track only, assuming the error to be temporally uncorrelated. Hence, for observations on the same track, a covariance function like

$$c(r) = C_0 \left(1 + \frac{r}{\alpha}\right) e^{(-r/\alpha)} + D_0 \left(1 + \frac{r}{\beta}\right) e^{(-r/\beta)}, \quad (3)$$

was used. The parameters D_0 and β were empirically determined to a variance of $(0.1 \text{ m})^2$ and a correlation length of 100 km, respectively. For observations on different tracks, D_0 was set to zero yielding an expression similar to (2). In both cases, C_0 was fixed at $(0.2 \text{ m})^2$, and the parameter α was fixed so that the correlation length was 15 km.

To enhance the filtering of the sea surface variability the estimation was carried out relative to a local weighted average of the nearest 100 observations. A weight function $1/(r^2 + r_0^2)$ with $r_0 = 0.1^\circ$ was applied so the local weighted average is a smooth representation of the surface heights that is less affected by the sea surface variability. In the collocation estimation the 48 nearest observations are used to secure redundant geoid information at crossing tracks. The result of the gridding is a $1/16^\circ$ by $1/16^\circ$ grid, having a size of 4° latitude by 16° longitude. Compared with the area used to select data, this grid was extended by a border zone of 1° latitude by 3° longitude to avoid spectral leakage in the following steps. This grid of sea surface heights will be treated as geoid undulations through the following steps.

The gravity anomalies, Δg , was derived from the geoid undulations, N , using FFT techniques [*Schwarz et al., 1990*]. In the frequency domain (u, v) , that is,

$$\Delta \tilde{g}(u, v) \approx \omega \gamma \tilde{N}(u, v) F(\omega), \quad (4)$$

where $\omega^2 = u^2 + v^2$ and γ is the normal gravity. Such a transformation from geoid undulations into free air gravity anomalies is a differentiation that enhances the high frequencies. Consequently, it is sensitive to noise. Therefore a Wiener filtering function, $F(\omega)$, was introduced in (4). This

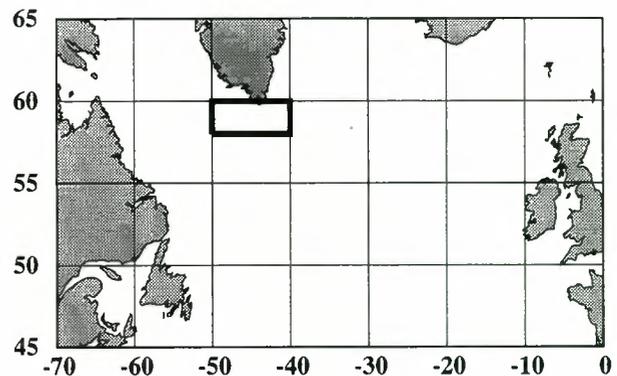


Figure 1. Location of test area south of Greenland in the North Atlantic Ocean.

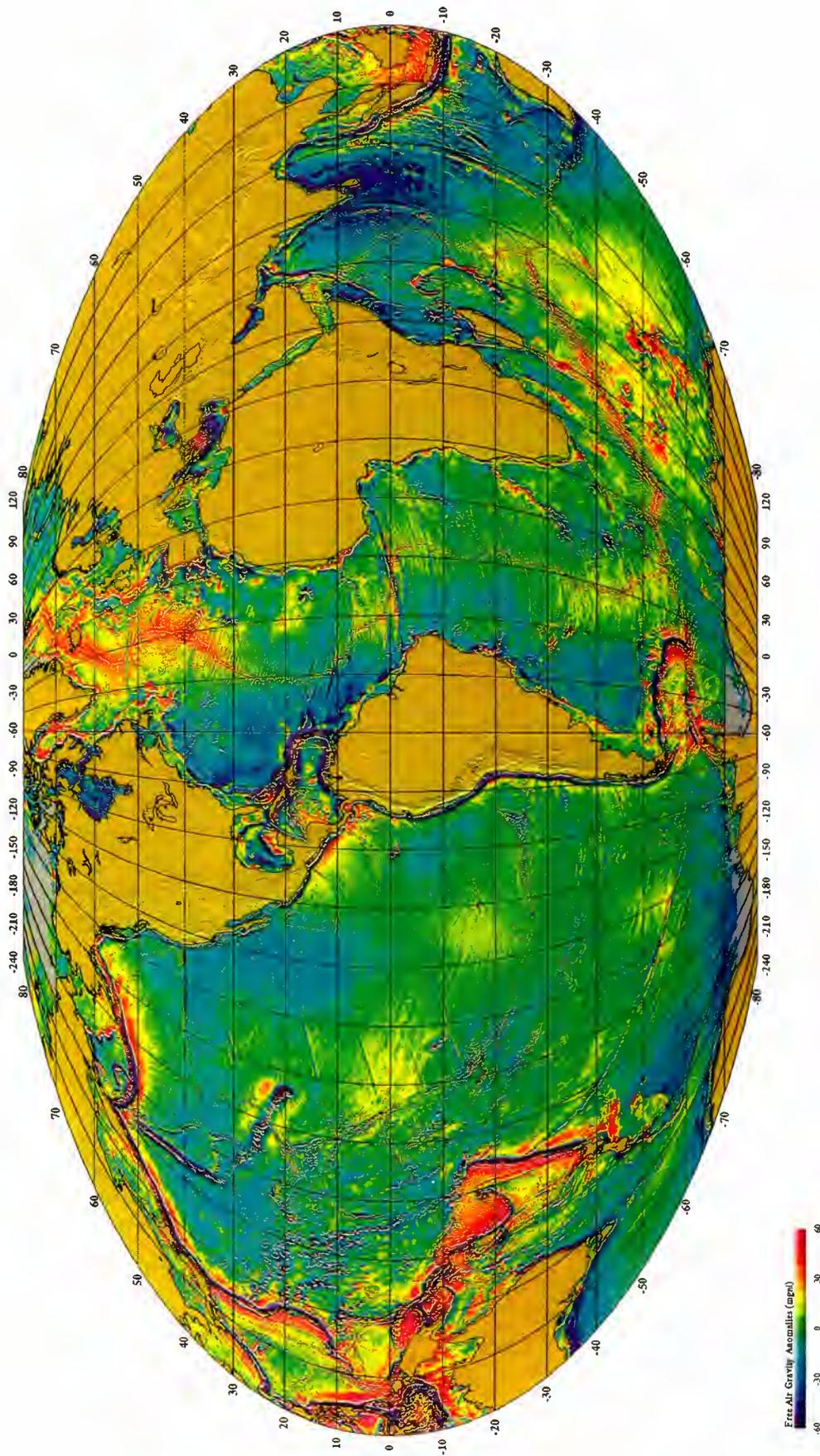


Plate 1. Global marine gravity field from the Geosat and the ERS-1 geodetic missions.

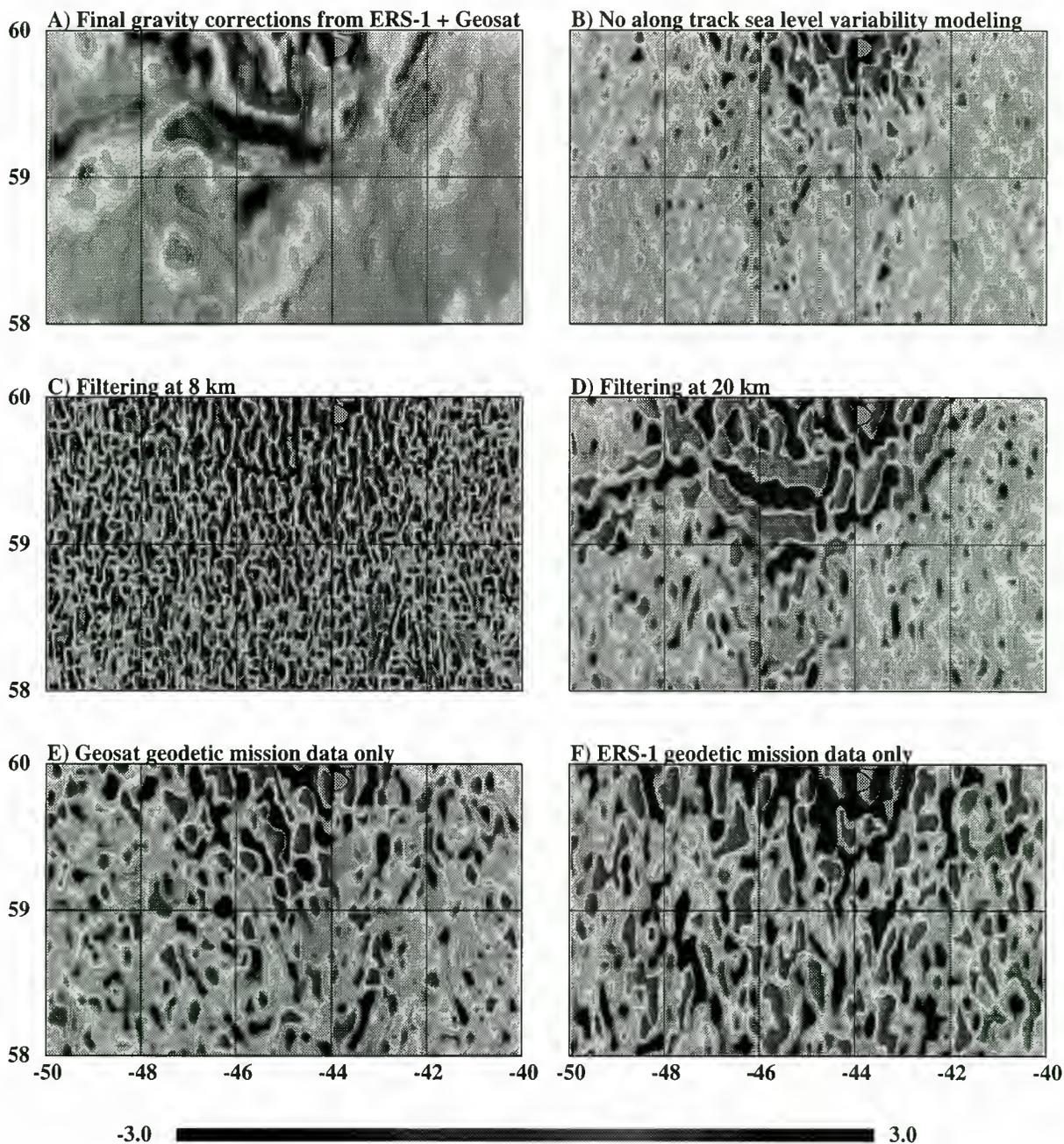


Figure 2. Comparison of different gravity field solutions in the test area shown in Figure 1. All units are mGal. 2(a) The “final” derived gravity residuals from ERS-1 and Geosat with respect to the Earth Gravity Model (EGM96) [Lemoine *et al.*, 1997]. The greyscale should be multiplied by a factor of 10 to range between -30 and 30 mGal for this picture. 2(b) Gravity difference between the final gravity field in Figure 2a and a gravity field estimated without along-track sea level variability modeling. 2(c) Gravity difference between the final gravity field in Figure 2a and a gravity field estimated with a filtering at 8 km. 2(d) Gravity difference between the final gravity field in Figure 2a and a gravity field estimated with a filtering at 20 km. 2(e) Gravity difference between the final gravity field in Figure 2a and a gravity field estimated using data from the Geosat Geodetic Mission (GM) data only. 2(f) Gravity difference between the final gravity field in Figure 2a and a gravity field estimated using data from the ERS-1 GM only.

filter function is equivalent to a collocation filter that assumes Kaulas rule to be valid and that assumes uncorrelated noise [details by Forsberg and Solheim, 1988]. That is,

$$F(\omega) = \frac{\omega_c^4}{\omega^4 + \omega_c^4}. \quad (5)$$

In this case the “cutoff” frequency, ω_c , where the filter is 0.5, was empirically determined to a wavelength of 12 km, which roughly corresponds to about 10 cycles per degree or harmonic degree 3600.

Before the Fourier transform of the geoid grid was computed, a cosine taper was applied in the outer 0.5° parts of the grid.

Table 1. Statistics of Different Gravity Field Solutions in a Test Area South of Greenland

	Mean	Standard Deviation	Minimum	Maximum
Final residual gravity to EGM96 from ERS-1 and Geosat; the final residual gravity field is shown in Figure 2a	-0.80	8.91	-41.44	36.34
Gravity difference between solutions computed with and without along-track sea level variability modeling; the gravity differences are shown in Figure 2b	0.01	1.05	-9.30	9.83
Gravity at wavelength between 8 and 12 km as shown in Figure 2c	0.00	3.63	-19.40	18.79
Gravity at wavelength between 12 and 20 km as shown in Figure 2d	0.01	2.04	-10.80	12.23
Gravity difference between final gravity field and gravity field computed using Geosat geodetic mission (GM) data only; the gravity differences are shown in Figure 2e	-0.00	1.69	-10.51	14.21
Gravity difference between final gravity field and gravity field computed using ERS-1 GM data only; the gravity differences are shown in Figure 2f	-0.09	2.03	-9.86	11.52

The location of the test area is shown in Figure 1. All units are mGal.

This was done to avoid spectral leakage caused by wavelengths that are not periodic within the area. Subsequently, the contribution of the EGM96 gravity field was restored to obtain the free air gravity anomalies, and the gravity field was isolated within the small 2° by 10° cell. Finally, all cells were collected to give the global gravity field.

4. Results

The procedure described above is well known, but in this analysis the various parameters have been fine-tuned to obtain an optimal product from a combination of Geosat and ERS-1 geodetic mission altimeter data in most of the world's oceans. The tuning of the parameters was based on empirical criteria with the scope of recovering as much detail and as much reliable gravity field information as possible. To demonstrate important phases in the derivation of the gravity field process, a 2° by 10° cell south of Greenland was selected. The location of this test site can be seen in Figure 1.

Figure 2 shows the result of the different steps in the fine-tuning of parameters used for the derivation of the gravity field.

Figure 2a shows the gravity field from ERS-1 and Geosat. The gravity anomalies are shown with respect to the EGM96 as it comes out when the fine tuned procedure is used. This field will in the following be called the final gravity field.

Figure 2b illustrates the importance of considering the residual sea surface variability in the gridding process. It shows the difference between a solution obtained without the error covariance function in (3) and the final gravity field in Figure 2a. The statistics of the difference are shown in Table 1. Even though, the standard deviation of the difference is only around 1 mGal. Figure 2b clearly shows anomalies with a clear track pattern. This so-called "trackiness" affects the gravity field recovery if the residual sea surface variability was ignored in the gridding process.

The role of the filter, $F(\omega)$ in (5), is sketched in Figures 2c and 2d. Figure 2c shows which parts of the solution filtered out when the cutoff is changed from a wavelength of 8 to 12 km. The differences look noisy with no coherency with the structures shown in Figure 2a. If the cutoff is changed from a wavelength of 12 to 20 km, then the pattern changes. Figure 2d

Table 2. Comparison With 23,663 Selected BGI Gravity Observations in the North Atlantic Ocean

Solution	Altimetry	Mean Difference	Standard Deviation
<i>Andersen and Knudsen</i> [this issue], computed with respect to the EGM96 geoid model	ERS-1	-4.4	6.2
	Geosat	-4.4	6.1
	ERS-1 and Geosat	-4.4	5.8
<i>Sandwell and Smith</i> [1997]	ERS-1 and Geosat	-4.4	7.6
<i>Andersen and Knudsen</i> [this issue], computed with respect to the OSU91A geoid model	ERS-1 and Geosat	-4.6	5.6

Data are from *BGI* [1996]. Locations can be seen in Figure 3. All units are mGal.

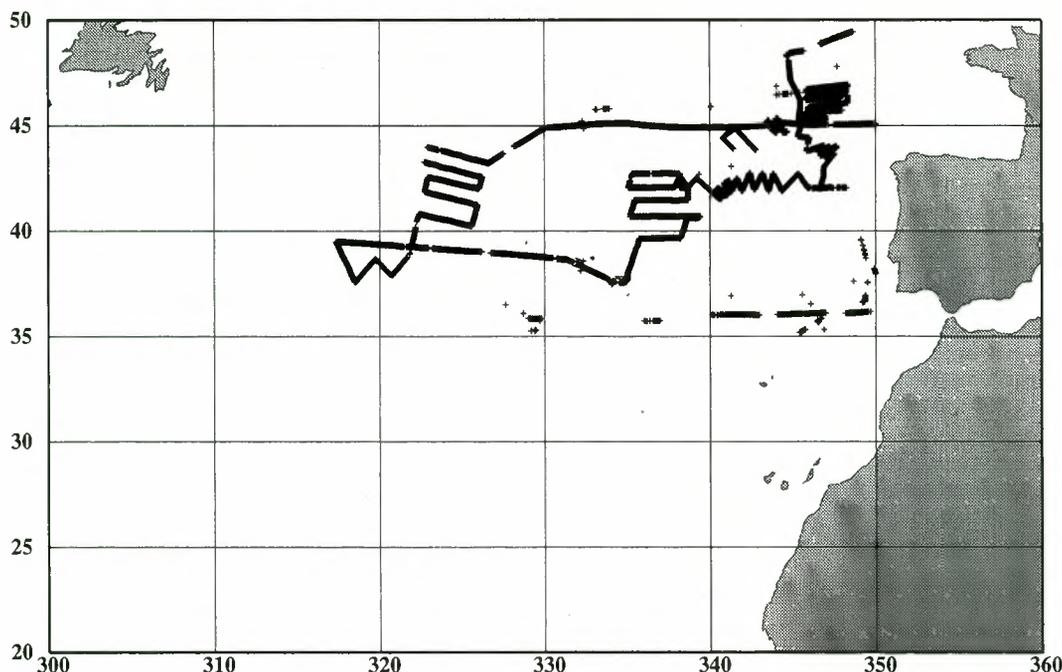


Figure 3. Location of 23,663 selected gravity observations in the North Atlantic Ocean from BGI [1996].

shows gravity differences that are clearly coherent with the gravity structures shown in Figure 2a. The standard deviations of the two sets of gravity differences are 3.6 and 2.0 mGal, respectively. Therefore valuable gravity field information has been removed. Wavelengths shorter than 12 km appear to be dominated by noise and cause the so-called "orange skin effect" in altimeter-derived gravity fields.

Figures 2e and 2f deal with the importance of using a combination of ERS-1 and Geosat geodetic mission altimetry rather than just ERS-1 or Geosat altimetry. In Figure 2e the difference between the solution in Figure 2a from ERS-1 and Geosat altimetry and a solution from Geosat altimetry alone is

shown. Figure 2f shows the similar difference with a solution from ERS-1 only. In both figures a track pattern related to Geosat tracks (Figure 2e) and ERS-1 tracks (Figure 2f) can be observed. Table 1 shows that the standard deviation of the differences are 1.7 and 2.0 mGal for Geosat and ERS-1 data, respectively. The standard deviation is highest for the ERS-1 alone differences as more Geosat data entered the gravity field determination. However, in both comparisons (Figures 2e and 2f), the combined use of both ERS-1 and Geosat improves the crossover analysis and thus reduces the trackiness.

The global recovery of the marine gravity is shown in Plate 1. This marine gravity field covers most ocean regions of the world. However, a few regions at extreme latitudes exist where no data were available. These are shown with grey in Plate 1. In most regions of the world, this gravity field method works very well. However, some marine regions exist where the method does not work as well. Problems as "trackiness" may be expected in regions with a large ocean variability and in shallow water regions where the ocean tide models may not be sufficiently accurate.

5. Comparison With Marine Gravity

Comparisons with marine free air observations were made in three different regions of the world. These regions were chosen as they have very different oceanographic and gravity signals. One region is in the middle of the Atlantic Ocean toward the Azores, which has very large oceanographic variability. The second is close to the coast of Antarctica in a partly ice-covered region. The third and final region is in the eastern Mediterranean Sea, which has a very large gravity signal.

The first comparison was made with a selection of gravity measurements from the Bureau Gravimetric International (BGI) database. In the region bounded by 10° and 50°N and 60° and 10°W, 23,663 marine gravity observations were selected from BGI [1996]. Only validated data after 1989 were selected. The selected marine gravity anomalies ranges from -100 mGal to 123 mGal. The result of this comparison is presented in Table 2 with the locations of the observation shown

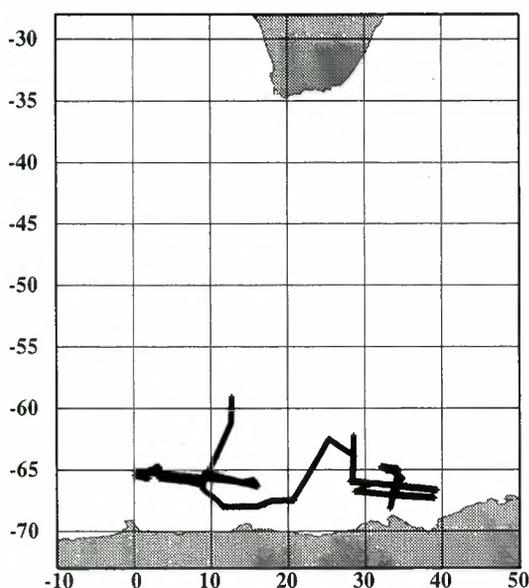


Figure 4. Location of 3642 gravity observations close to Antarctica. The data have been measured by Bundesanstalt für Geowissenschaften und Rohstoffe (BGR).

Table 3. Comparison With 3642 Gravity Observations in the Antarctic Region

Solution	Altimetry	Mean Difference	Standard Deviation
<i>Andersen and Knudsen</i> [this issue], computed with respect to the EGM96 geoid model	ERS-1	0.8	4.8
	Geosat	0.8	4.1
	ERS-1 and Geosat	0.8	3.8
<i>Sandwell and Smith</i> [1997]	ERS-1 and Geosat	2.3	3.8
<i>Andersen and Knudsen</i> [this issue], computed with respect to the OSU91A geoid model	ERS-1 and Geosat	0.2	4.0

The locations are shown in Figure 4. All units are mGal.

in Figure 3. An additional comparison with a previous global gravity field by *Andersen and Knudsen* [1995] derived with respect to the OSU91A geopotential model [*Rapp et al.*, 1991] has been added.

The similar comparison was done with the global gravity field by *Sandwell and Smith*, [1997]. This gravity field is one of their most recent gravity fields (version 7.2), and it will hereinafter be referred to as the Sandwell and Smith gravity field. This gravity field is derived from Geosat and ERS-1 geodetic mission altimetry using slopes. The distributed gravity field has a $1/30^\circ$ longitude spacing and a variable latitude spacing. This was subsequently interpolated onto a $1/16^\circ$ by $1/16^\circ$ grid using collocation.

The standard deviations for these five global gravity field solutions range between 5.6 and 7.6 mGal as shown in Table 2. The gravity field from ERS-1 and Geosat derived with respect to the OSU91A geopotential model has the lowest standard deviation of 5.6 mGal, and the Sandwell and Smith gravity field have the highest standard deviation of 7.6 mGal. All satellite-derived models agree on a systematic mean difference (ship gravity anomalies minus altimetry derived anomalies) of 4.5 mGal. However, deciding the cause of these systematic errors

is difficult, as no information was available on the processing of the BGI marine gravity data.

The second region of comparison was selected close to Antarctica in a partly ice-covered region between 60° - 70° S and 0° - 40° E. In this comparison, 3642 observations made by the ship *Polar Stern* in 1990, were provided to us by the Bundesanstalt für Geowissenschaften und Rohstoffe (BGR) in Germany. Comparison with the similar data set has previously been reported by *Kim* [1996], *Trimmer and Manning* [1996], and *Rapp and Yi* [1997]. The observed marine gravity observations range from -70 to 85 mGal in this region. The location of the observations is shown in Figure 4 and the comparison is reported in Table 3. The global gravity field from ERS-1 and Geosat has the same standard deviation with the marine data set as the Sandwell and Smith global gravity field, namely 3.8 mGal. However, the *Sandwell and Smith* [1997] gravity field has a mean difference of 2.2 mGal. Similar comparisons with different subsets of the same data set were carried out by *Rapp and Yi* [1997].

The final comparison with marine gravity data was done with a set of 4151 gravity observations in the eastern Mediterranean Sea by *Morelli et al.* [1975]. The location of the stations is

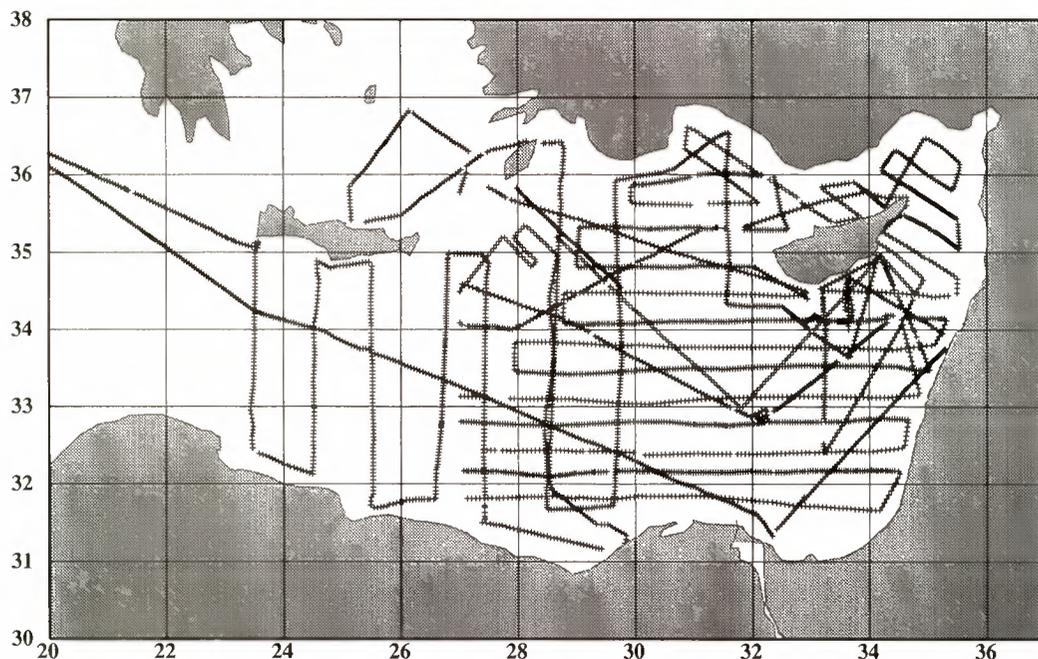


Figure 5. Location of 4151 marine gravity observations in the eastern Mediterranean Sea from *Morelli et al.*[1975].

Table 4. Comparison With 4151 Gravity Observations in the Eastern Mediterranean Sea

Solution	Altimetry	Mean Difference	Standard Deviation
<i>Andersen and Knudsen</i> [this issue], computed with respect to the EGM96 geoid model	ERS-1	-1.5	10.2
	Geosat	-1.3	10.3
	ERS-1 and Geosat	-1.1	9.7
<i>Sandwell and Smith</i> [1997]	ERS-1 and Geosat	-6.6	18.9
<i>Andersen and Knudsen</i> [this issue], computed with respect to the OSU91A geoid model	ERS-1 and Geosat	-3.1	9.9

Locations can be found in Figure 5. All units are mGal.

shown in Figure 5 and the comparison is tabulated in Table 4. This region has a very large gravity signal and the marine observations range from -218 to 116 mGal. Therefore it is expected that the comparison will have a larger standard deviation in this region. The standard deviations with this data set are generally around 10 mGal with the *Sandwell and Smith* [1997] data set having a very large standard deviation of 18.9 mGal. Similarly, the Sandwell and Smith gravity field has a systematic mean difference of 6.6 mGal. In the previous comparison, close to Antarctica, a systematic difference with marine gravity was also found. The cause of this has not been investigated, but the estimation of the medium to long wavelength parts of the gravity field may suffer when using slopes as done by Sandwell and Smith. Comparisons of the Andersen and Knudsen, and Sandwell and Smith gravity fields with marine gravity in the whole of the Mediterranean Sea by *Behrend et al.* [1996] shows similar result.

6. Discussion

The method used by *Andersen and Knudsen* [1995, 1996] has been fine-tuned to obtain an optimal gravity field from a combination of Geosat and ERS-1 geodetic mission altimeter data. This technique has been used in the global recovery of the marine gravity shown in Plate 1. The comparison with three different sets of marine gravity showed very promising results. However, some marine regions may exist where the method does not work as well. Problems with trackiness may be expected in regions with a large ocean variability and in shallow water regions where the ocean tide models may not be sufficiently accurate.

Much effort was focused on the removal of trackiness which primarily is caused by sea level variability. Usually, this oceanographic variability shows up as smooth signatures along a ground track and causes no problems before it is combined with the variability along other tracks. A method was derived to model this error by introducing an additional covariance function for this error in the interpolation of the data. Such signals may be filtered out more easily if sea surface slopes are used as observations instead of sea surface heights. However, the estimation of the medium to long wavelength parts of the gravity field may suffer when using slopes.

Acknowledgments. This research was partly supported by the Danish Space Board. The Geosat data were kindly provided by NOAA. The ERS-1 data were obtained from ESA to which DUT/DEOS kindly provided the JGM-3 orbits. The authors wishes to thank R. H. Rapp, Ohio State University, for valuable suggestions.

References

- Andersen, O. B., and P. Knudsen, Global gravity field from the ERS-1 geodetic mission, *Earth Obs. Q.*, 47, 1-5, 1995.
- Andersen, O. B., and P. Knudsen, Altimetric gravity field from the full ERS-1 geodetic mission, *Phys. Chem. of the Earth*, 21(4), 337-341, 1996.
- Andersen, O. B., P. Knudsen, and C. C. Tscherning, Investigation of methods for global gravity field recovery from dense ERS-1 geodetic mission altimetry, in *Global Gravity Field and Its Temporal Variations*, IAG Symp., vol. 116, edited by R. H. Rapp, A. Cazenave, and S. Nerem, pp. 218-226, Springer-Verlag, New York, 1996.
- Andersen, O. B., P. L. Woodworth, and R. A. Flather, Intercomparison of recent global ocean tide models, *J. Geophys. Res.*, 100(C12), 25,261-25,282, 1995.
- Balmino, G., B. Moynot, M. Sarrailh, and N. Valès, Free air gravity anomalies over the oceans from Seasat and Geos 3 altimeter data, *Eos Trans. AGU*, 68(2), 17-18, 1987.
- Behrend, D., H. Denker, and K. Schmidt, Digital gravity data sets for the Mediterranean Sea derived from available maps, *Bull. d'Inf.*, 78, 32-41, 1996.
- Bureau Gravimétrique International (BGI), Global gravity data, CD-ROM sea data, version 1.0, CNES, Toulouse-Cedex, France, 1996.
- Forsberg, R., and D. Solheim, Performance of FFT methods in local gravity field modelling, *Chapman Conference on Progress in the Determination of the Earth's Gravity Field*, pp. 100-103, AGU, Ft. Lauderdale, Fla., 1988.
- Haxby, W. F., Gravity field of the worlds oceans (Seasat altimetry), map, Nat. Geophys. Data Cent., Boulder, Colo., 1987.
- Hwang, C., and B. Parsons, Gravity anomalies derived from Seasat, Geosat, ERS-1, and TOPEX/POSEIDON altimetry and ship gravity: A case study over the Reykjanes Ridge, *Geophys. J. Int.*, 122, 551-568, 1995.
- Kim, J.-H., Improved recovery of gravity anomalies from dense altimeter data, *Rep. 444*, 130pp. Dept. of Geod. Sci. and Surv., Ohio State Univ., Columbus, 1996.
- Knudsen, P., Simultaneous estimation of the gravity field and sea surface topography from satellite altimeter data by least squares collocation, *Geophys. J. Int.*, 104, 307-317, 1991.
- Knudsen, P., and M. Brovelli, Collinear and cross-over adjustment of Geosat ERM and Seasat altimeter data in the Mediterranean Sea, *Surv. Geophys.*, 14(4-5), 449-459, 1993.
- Knudsen, P., O. B. Andersen, and C. C. Tscherning, Altimetric gravity anomalies in the Norwegian-Greenland Sea: Preliminary results from the ERS-1 35 days repeat mission, *Geophys. Res. Lett.*, 19(17), 1795-1798, 1992.
- Lemoine, F. G., et al., The development of the NASA GSFC and DMA joint geopotential model, in *Proceedings of the International Symposium on Gravity Geoid and Marine Geodesy*, Springer-Verlag, New York, in press, 1997.
- Le Provost, C., M. L. Genco, F. Lyard, P. Vincent, and P. Canceil, Spectroscopy of the world ocean tides from a finite element

- hydrodynamic model, *J. Geophys. Res.*, 99(C12), 24,777-24,797, 1994.
- Morelli, C., M. Pisani, and C. Gantar, Geophysical studies in the Aegean Sea and in the eastern Mediterranean, *Boll. Geofis. Teor. Appl.*, 18, 127-168, 1975.
- Rapp, R. H., and Y. Yi. Role of ocean variability and sea surface topography in the recovery of the mean sea surface and gravity anomalies from satellite altimeter data, *J. Geod.*, in press, 1997.
- Rapp, R. H., Y. M. Wang, and N. K. Pavlis, The Ohio State 1991 geopotential and sea surface topography harmonic coefficient models, *Rep. 410*, Dep. of Geod. and Sci. Surv., Ohio State Univ., Columbus, 1991.
- Rummel, R., and R. H. N. Haagmans, Gravity gradients from satellite altimetry, *Mar. Geod.*, 14, 1-12, 1990.
- Sandwell, D. T., Antarctic marine gravity field from high-density satellite altimetry, *Geophys. J. Int.*, 109, 437-448, 1992.
- Sandwell, D. T., and W. H. F. Smith, Marine gravity anomaly from Geosat and ERS-1 satellite altimetry, *J. Geophys. Res.*, 102(B5), 10,039-10,054, 1997.
- Schwarz, K.-P., M. G. Sideris, and R. Forsberg, Use of FFT methods in physical geodesy, *Geophys. J. Int.*, 100, 485-514, 1990.
- Shum, C. K., et al., Accuracy assessment of recent ocean tide models, *J. Geophys. Res.*, in press, 1997.
- Tapley, B. D., et al., The Joint Gravity Model 3, *J. Geophys. Res.*, 101(B12), 28,029-28,049, 1996.
- Trimmer, R., and D. Manning, The altimetry derived gravity anomalies to be used in computing the joint DMA/NASA earth gravity model, in *Global Gravity Field and Its Temporal Variations, IAG Symp.* vol. 116, edited by R. H. Rapp, A. Cazenave, and S. Nerem, pp. 198-204, Springer-Verlag, New York, 1996.
- Tscherning, C. C., A strategy for gross-error detection in satellite altimeter data applied in the Baltic Area for enhanced geoid and gravity determination, in *Determination of the Geoid - Present and Future, IAG Symp.*, vol. 106, pp. 3-12, edited by R.H. Rapp and F. Sansó, Springer-Verlag, New York, 1990.
- Tscherning, C. C., P. Knudsen, S. Ekholm, and O. B. Andersen, An analysis of the gravity field in the Norwegian sea using ERS-1 altimeter measurements, in *Proceedings of the First ERS-1 symposium, European Space Agency Spec. Publ. ESA SP-359*, 413-418, 1993.
- Tscherning, C. C., R. Forsberg, and P. Knudsen, The GRAVSOFT package for geoid determination, in *Proceedings of the First Workshop on the European Geoid*, Prague, pp. 327-334, Springer-Verlag, New York, 1992.

Ole Baltazar Andersen and Per Knudsen, Kort- og Matrikelstyrelsen, DK-2400 Copenhagen NV, Denmark. (e-mail: oa@kms.min.dk; pk@kms.min.dk)

(Received December 31, 1996; revised June 30, 1997; accepted July 30, 1997.)

Improved techniques for the derivation of sea surface temperatures from ATSR data

Ian J. Barton

Division of Atmospheric Research, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Aspendale, Victoria, Australia

Abstract. A new technique is developed for the analysis of satellite data to provide accurate measurements of seasurface temperature (SST). Satellite data sets are partitioned into subsets depending on the value of a selected parameter (for example, latitude, total water vapor, and water vapor content of an atmospheric layer) to provide a suite of algorithm coefficients that reduces the errors associated with the derivation of SST. For data sets obtained with the along-track scanning radiometer (ATSR) the data themselves can be used to modify the coefficients used in the SST algorithm, but for other instruments it may be necessary to use additional data sets to select a correct set of algorithm coefficients. A simulated set of ATSR data is used to develop algorithm coefficients for different atmospheric conditions, and the improvement in SST derivation is demonstrated. For ATSR, when all the data for three infrared channels in both the nadir and forward views are available, the improvement is marginal, but for situations when there are limited data the improvement is considerable. The analysis suggests that in the future the best SST analyses will be obtained by developing an interactive system, where the satellite data are ingested into a numerical weather forecast model so that algorithms can be selected and applied with a forecast or analysis of the atmospheric state. The new techniques are tested by application to a large global data set of ATSR brightness temperatures. This analysis highlights the future need for large SST validation data sets that include coincident satellite and surface-based measurements.

1. Introduction

Over the last few years there has been a change of emphasis in the science of deriving sea surface temperature (SST) using infrared satellite data. In the past the derivation of algorithm coefficients has concentrated on the removal of atmospheric effects caused by the absorption of the surface-emitted infrared radiation by water vapor and other constituents. This has provided a large number of SST algorithms, some of which are claimed to be applicable globally, while others are for specific geographic locations and seasons. In a recent review, Barton [1995] compared many of these algorithms and found that they were all basically similar and that all of them performed reasonably well in average conditions. Most of these algorithms are applicable to data from the advanced very high resolution radiometers (AVHRRs) on the operational meteorological satellites, mainly because the data are freely available and techniques for data calibration and instrument performance are easily applied. The latest operational AVHRR algorithms now include a dependence on either total water vapor content or a surrogate of that parameter. The accuracy of the SST derived using these algorithms is now reach-

ing the theoretical limit (given the noise temperature on each channel used, the accuracy of the onboard calibration, the 10-bit data digitization, and normal atmospheric variability) of 0.6–0.8 K [May, 1993]. However, there are still situations where the errors in SST from AVHRR data are unacceptably large (> 1.0 K).

With the advent of data from the along-track scanning radiometer (ATSR), which have reduced noise temperatures on the infrared data, 12-bit data digitization, and improved onboard absolute calibration, the theoretical limit to SST derivation has been reduced to < 0.2 K [Závody *et al.*, 1995]. With this new accuracy the derivation and validation of SST values is now focused on other areas rather than on simple atmospheric correction. For instance, it is now important to consider the variation in the sea surface emissivity with view angle and wind speed, the difference between the bulk SST and the thin radiative skin layer at the surface, and the detailed vertical structure of the atmospheric temperature and water vapor. This paper deals with the last of these three phenomena and develops a new technique for deriving SST algorithms that takes some account of the water vapor and temperature distribution in the atmosphere.

2. Hypothesis

The derivation of SST from infrared satellite data involves the application of the satellite measurements to

remove the effect of atmospheric absorption using an inversion procedure. In the simplest case the difference between two brightness temperatures at two different wavelengths can be used as a measure of the effect of atmospheric absorption, and a linear algorithm provides the derived SST value. The coefficients of the SST algorithm can be derived in several ways. A theoretical derivation involves the use of a radiative transmission model to derive theoretical satellite-measured radiances that are converted to brightness temperatures, and then a regression procedure is used to derive the algorithm coefficients [e.g., *Závodny et al.*, 1995]. Algorithm coefficients can also be derived from a data set of coincident satellite and surface-based measurements [e.g., *McClain*, 1984]. A third method involves the use of theoretical models to derive the form of the algorithm with an initial set of coefficients, and then a set of satellite and surface measurements is used to fine-tune the coefficients to give the best agreement between the derived SST and the surface measurements [e.g., *Walton*, 1988].

The application of these inversion procedures to derive SST from satellite data implies the use of a first guess to the state of the atmosphere between the satellite and the Earth's surface. The theory behind this concept is discussed further by *Rodgers* [1976] and *Eyre* [1987]. For each algorithm the first guess is related to the average vertical structure of the water vapor and other constituents in the atmospheres from which the coefficients are derived, whether it be the profiles used in a radiative transfer model or those associated with an ensemble of coincident satellite and ground-based data. In any situation, if the atmospheric structure matches the first guess, then a perfect retrieval will be made using the derived algorithm. It is also possible to derive a correct SST value, even if the first guess is not matched by the atmospheric structure. However, in many cases, if the structure is significantly different to the first guess, then an error in the derived SST will occur. The hypothesis investigated in this paper is that if some information on the vertical structure of the atmosphere (water vapor and temperature) can be incorporated into the analysis, then more accurate global fields of SST can be derived from satellite data.

Over the past decade, much effort has been invested into the development of SST algorithms for deriving SST from AVHRR and ATSR data. *Barton* [1995] has reviewed the status of these algorithms and finds that there has been little improvement in the accuracy of the derived SST. This is because the most recent algorithms do not address the remaining source of error in the SST, i.e., the situations where the atmospheric structure departs from that implied by the algorithms' first guesses. In this paper, techniques are developed whereby either the satellite data themselves or some external data source can be used to first select the appropriate algorithm coefficients (first guess) before the SST derivation is attempted. This can also be a two-step procedure; in the first instance, algorithms can be derived that include a dependence on (say) total water vapor content, and then in a second stage a dependence on the vertical water vapor profile can be introduced.

For ATSR data, where there are six brightness temperatures available for each pixel, the improvements are minimal, but for AVHRR derivation there can be significant improvements in the derived SST if anomalous atmospheric structure can be detected. In these cases, even with AVHRR data alone, it is possible to select a more appropriate algorithm and thus derive a more accurate SST.

In this analysis, satellite data sets (both real and simulated) are partitioned into discrete groups depending on the value of predetermined parameters. A separate SST algorithm is developed for each group, and the results are tested to see if there is any improvement in the derived SST accuracy. This procedure thus identifies those parameters that both determine the accuracy of the derived SST value and also that may be introduced into the algorithms to improve SST accuracy. The final SST algorithm may then either have coefficients that are dependent on the sensitive parameters or several discrete sets of coefficients may be developed and interpolation used to derive the final values. In some cases it may be sufficient to just use a small matrix of algorithm coefficients.

3. ATSR Data Sets

3.1. ATSR Instrument

The first ATSR, which was launched on the ERS-1 satellite in July 1991, was specifically designed to provide accurate measurements of SST by obtaining two views of the Earth's surface to improve the correction to be applied for absorption of infrared radiation by the atmosphere. As well as accounting for absorption by water vapor, it has also been found that the dual view allows an improved estimate of SST in the presence of large aerosol concentrations by providing dual-path measurements at the same wavelength [*Smith et al.*, 1994]. The dual view, coupled with actively cooled detectors, 12-bit data digitization, and accurate onboard absolute calibration, gives the ATSR an improved performance over other satellite radiometers. The ATSR also includes a two-channel passive microwave radiometer that can be used to measure the total water vapor column in the atmosphere beneath the ERS-1, and the ultimate SST algorithms used to derive SST will incorporate data from the microwave instrument. Further details on the ATSR are given in work by *Delderfield et al.* [1986].

Both the AVHRR and the ATSR use the same three thermal infrared wave bands that are located in spectral "windows", where there is a minimum of atmospheric absorption, but where there is a residual absorption mainly due to atmospheric water vapor. The radiance measured by the radiometer originates partly from the surface and partly from the intervening atmosphere. The relative contribution of these radiances is dependent on the wavelengths used, the distribution of the water vapor in the atmosphere, and the temperature of that water vapor. The ATSR provides six infrared measurements of the Earth's surface, three wavelengths at two view angles; *Barton* [1996] has shown that these

measurements can provide details of the water vapor distribution in the atmosphere as well as an estimate of the total water vapor column.

3.2. Simulated ATSR Data Set

The main ATSR data used in this analysis are a simulated set derived using an atmospheric transmission model with a collection of data from 885 maritime radiosonde ascents obtained in the Australian region [Baker and Dowd, 1978]. The radiosonde data set covers a range of latitudes from 10°N to 50°S. The transmission model has been supplied by the Rutherford Appleton Laboratory, and details are given by Závody *et al.* [1995]. The model uses a line-by-line procedure across the filters of the ATSR to produce simulated brightness temperatures at the satellite for given plane-parallel atmospheric vertical profiles of water vapor and temperature. The brightness temperatures were calculated in the nadir for the nadir view and a zenith angle at the surface of 55.15° for the forward view. The radiosonde data were also analyzed to provide a total water vapor content as well as the water vapor content of each 1 km layer between the surface and 6 km above the surface. For each radiosonde profile a linear lapse rate of temperature and an exponential lapse rate of water vapor were assumed between each radiosonde data point. For all the model calculations used in this work a noise temperature of 0.04 K was assumed for each of the infrared channels. This is an average value for the three infrared channels with the detectors maintained at a temperature of 91 K [Mason, 1991].

In this analysis, realistic variations in air-sea temperature difference were introduced by altering the surface temperature of the radiosonde profile without changing the remaining parameters. The surface temperature was increased and decreased by 2.0 K, which triples the size of the data set to 2655 profiles. The operational ATSR algorithms described by Závody *et al.* [1995] were used to provide an ATSR-derived SST for six, four, three, and two input channels of data. The data set itself was also used to generate similar algorithms with no account taken of the latitude of the radiosonde data. The infrared brightness temperature data used in each of the four algorithms are ZSST6 and BSST6: T12N, T11N, T37N, T12F, T11F, and T37F; ZSST4 and BSST4: T12N, T11N, T12F, and T11F; ZSST3 and BSST3: T12N, T11N, and T37N; and ZSST2 and BSST2: T12N and T11N, where 12, 11, and 37 refer to the 12, 11, and 3.7 μm channels and N and F refer to the nadir and forward views, respectively. Z (Závody) refers to operational ATSR latitude-dependent algorithms and B (Barton) refers to algorithms generated from the simulated data set used in this analysis. These latter algorithms were obtained by multiple linear regression using the technique described by Barton *et al.* [1989].

3.3. ATSR Data Set

In the past it has been found that techniques developed using simulated satellite data sometimes do not work when applied to real data supplied by satellite

instruments [Barton, 1995]. This can be due to several causes including the effect of data digitization, the signal fluctuations introduced by atmospheric inhomogeneities, and the inherent noise present in radiometric measurements. Digitization and atmospheric effects cannot necessarily be reduced by spatial averaging while radiometric noise can be improved, but assumptions regarding the homogeneous nature of the atmosphere and Earth's surface are required. Thus it is important that new techniques be applied to satellite data to ensure that these effects are not prevalent. The techniques developed in this paper can only be tested by applying them to a large high-quality set of combined ATSR and surface data. Unfortunately, such a data set does not exist. However, by using a large ATSR data set, it may be possible to test the applicability of the new SST derivation techniques.

A suitable ATSR data set has been compiled by selecting 13 ascending passes of the ERS-1 satellite that cover mainly oceanic tracks from the south pole to the north pole. Ten of these tracks were in the Pacific Ocean with two in the Indian Ocean and one in the Atlantic Ocean. The data were all received in the latter half of April 1992, a period that was chosen to ensure that data from the 3.7 μm channel were present (this channel failed during May 1992). Only ascending passes were used, as these occur during the night and the 3.7 μm data were not contaminated with reflected solar radiation.

Each pass was examined manually for cloudless areas that covered at least 21 by 21 pixels and lines. For each 11 by 11 pixels in the center of these areas the mean, maximum, and minimum values and the standard deviation of the brightness temperatures in each channel (3.7, 11, and 12 μm in both the nadir and forward views) were calculated. The total atmospheric water vapor was also calculated from the brightness temperatures in the two passive microwave channels using an algorithm provided by L. Eymard (personal communication, 1993). In all, a total of 2313 small areas covering a wide range of latitudes and pixel numbers (distance from the sub-satellite track) were analyzed and incorporated into the data set.

4. Criteria for the Partitioning of Data Sets

In this paper several different parameters are explored as a basis for partitioning the simulated satellite data. For each group of data a different set of algorithm coefficients (i.e., first guess) is used to derive the SST. To aid with this analysis, the errors associated with the derivation of SST using the simulated brightness temperature data were explored to discover any dependencies on different parameters. If the errors (the differences between the radiosonde surface temperature and the ATSR-derived SST) are dependent on a particular parameter, then that may give a clue as to how the analysis techniques may be improved. In the analysis in this section, only the data generated from the basic 885 profiles are used.

4.1. Latitude

The current operational SST algorithms for ATSR data [Závodny *et al.*, 1995] provide coefficients for tropical, midlatitude, and polar regions and linearly interpolate between the derived SST values depending on the latitude. When applying this algorithm (ZSST6) to the simulated ATSR data set, it was found that there was no relation between the residual SST errors and latitude. This is not surprising, as any latitude dependence should be removed by the derivation technique. However, when the SST was estimated with a simple six-channel algorithm derived from the simulated data set itself (the BSST6 algorithm), there is also no dependence of SST error on latitude. For example, the data with latitudes south of 25°S have a mean error of 0.002 K while the remaining data have a mean error of -0.002 K. Although this mean error difference of 0.004 K is extremely small (and it may not be statistically significant), it can nevertheless be used to give an indication of the importance of different parameters in determining the SST errors. The results of this analysis are included in Table 1. When the same latitude analysis was undertaken for the four- and three-channel algorithms, the mean error differences were not significant. Oddly enough, it was one of the latitude-dependent algorithms, ZSST2, which showed the largest difference. This suggests that the current technique of partitioning the algorithm coefficients based on latitude alone may not be the optimum technique.

A first impression may be that the magnitudes of these differences are not important when we are aiming to estimate SST values with an accuracy of 0.2-0.3 K. However, this analysis is based on simulated data, and the theoretical errors are quite small. For example, using the simulated data set, the standard error of the SST derived with the BSST6 is 0.060 K and that derived with ZSST6 is 0.043 K. The difference here is due to the different selection of the noise temperature on each channel. In this analysis the noise temperature in all channels was assumed to be 0.04 K. If this noise is reduced to 0.027 K, then the standard error of the BSST6 algorithm is reduced to that of the ZSST6 algorithm. For the two-channel algorithms the standard errors are 0.340 and 0.273 K for ZSST2 and BSST2, respectively. It is interesting to note that in this case the algorithm derived from the simulated data set performs better (on the data set) than the operational algorithm. With these small standard errors the mean error differences given in Table 1, in some cases, become significant. This is also evident when the results displayed in Figures 1 and 2 are observed.

4.2. Total Water Vapor Content

Atmospheric water vapor is the main absorber of thermal infrared radiation and is thus a likely candidate for the partitioning of data to improve SST derivation. Of course, total water vapor content is in some ways dependent on latitude as tropical atmospheres are

Table 1. Mean Standard Error Differences Obtained When the Simulated SST Errors are Partitioned by Various Parameters

Parameter	Threshold	Algorithm	Number < and > Threshold		Mean Error < Threshold, K	Mean Error > Threshold, K	Mean Error Difference, K
Absolute latitude	25°	ZSST6	449	436	-0.030	-0.037	0.007
		BSST6			-0.002	0.002	0.004
		ZSST4			-0.048	-0.082	0.034
		BSST4			0.008	-0.009	0.017
		ZSST3			-0.087	-0.051	0.036
		BSST3			-0.001	0.001	0.002
		ZSST2			-0.127	0.024	0.151
		BSST2			0.014	-0.015	0.029
Total water vapor	3.0 cm	ZSST6	611	274	-0.044	-0.008	0.036
		BSST6			-0.002	0.005	0.007
		ZSST2			0.037	-0.251	0.288
		BSST2			-0.025	0.056	0.081
	2.0 cm	ZSST6	415	470	-0.050	-0.018	0.032
		BSST6			-0.011	0.010	0.021
		ZSST2			0.099	-0.186	0.285
		BSST2			0.010	-0.009	0.019
Water vapor above 3 km	0.25 cm	ZSST6	431	454	-0.051	-0.016	0.035
		BSST6			-0.005	0.004	0.007
		ZSST2			0.167	-0.261	0.428
		BSST2			0.077	-0.074	0.151
Surface inversion	SST > 288 K and temperature > SST+0.5 K at any height	ZSST6	804	81	-0.029	-0.076	0.047
		BSST6			0.007	-0.066	0.073
		ZSST2			-0.049	-0.086	0.036
		BSST2			-0.015	-0.148	0.163

SST is sea surface temperature.

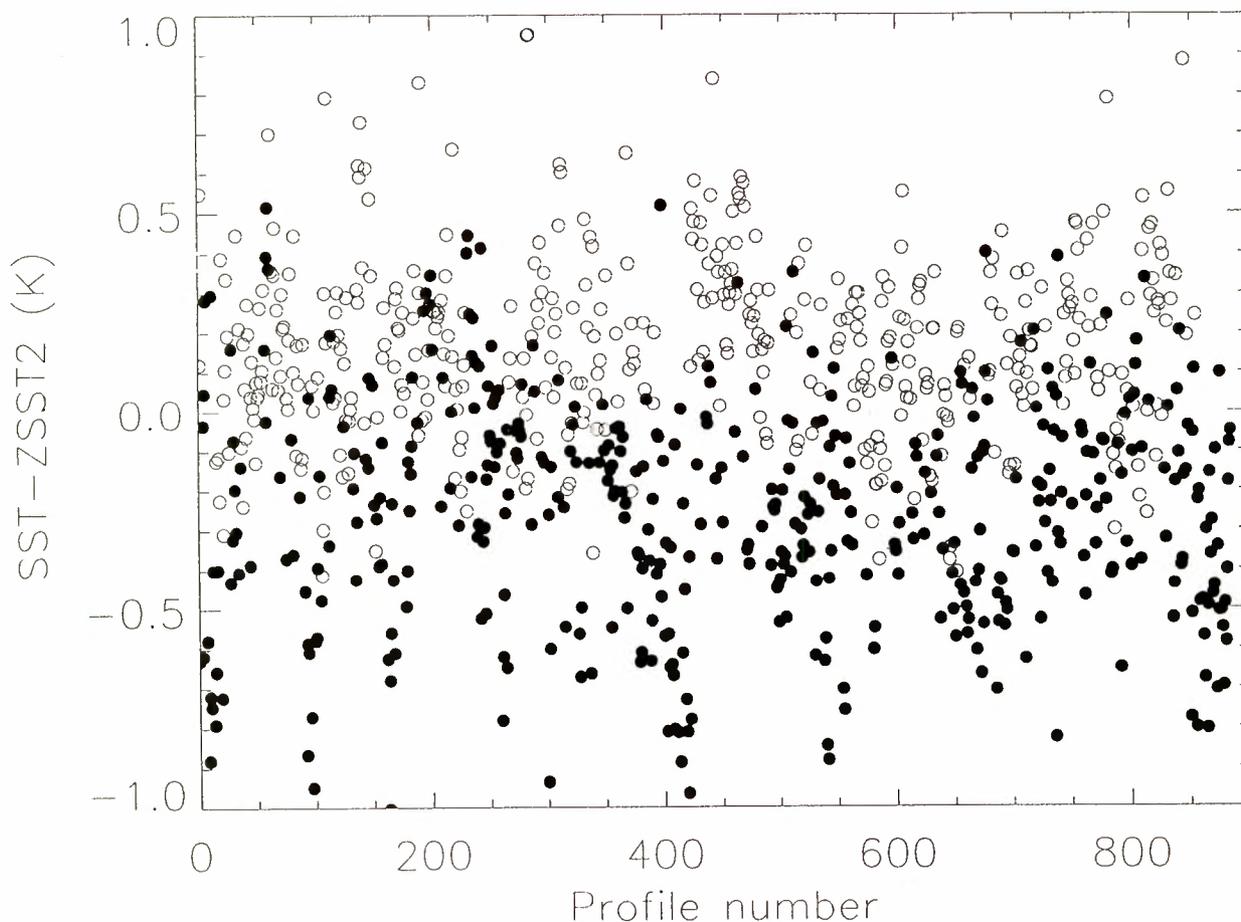


Figure 1. The partition of errors in the ZSST2 algorithm ($SST-ZSST2$) against the water vapor amount above 3 km. The open circles show those profiles where the water vapor amount is < 0.25 cm, and the solid circles show where the water vapor is > 0.25 cm.

generally more humid, but as water vapor is the main absorber, it is perhaps more logical to use this parameter in any partitioning. The major difficulty in this approach is that atmospheric water vapor content is a highly variable parameter and is difficult to measure accurately. When the data are partitioned into those cases that have a total water vapor content greater than and less than a certain threshold, the mean SST error for each group is different. These differences for the six- and two-channel algorithms are tabulated in Table 1. The results suggest that if the operational SST algorithms were divided into two sets, one for high water vapor and one for low, then there would be an improvement in the accuracy of the derived SST. For the BSST6 and BSST2 algorithms it appears that the benefits may not be so great.

4.3. Upper Level Water Vapor Content

Water vapor that exists in the upper atmosphere (above 3 km) usually has a temperature that is significantly different from the surface temperature and can thus have a greater effect on the satellite-measured radiance than the larger amounts of water vapor that are nearer the surface but have a temperature close to the surface temperature. A threshold of 0.25 cm

of water vapor above 3 km divides the data set into approximately equal numbers, and the mean error differences were calculated using this criterion. Again, the results show that the errors in the operational ATSR algorithms are strongly correlated with this parameter.

The distribution of the errors in the ZSST2 algorithm are shown in Figure 1. For the ZSST2 algorithm the mean error difference is considerably larger than the standard error of 0.340 K as specified above.

In later sections of the paper the upper level water vapor is also expressed as a ratio of the water vapor above that level to the total water vapor. For a height of 3 km a ratio of 0.1 is used to divide the data sets.

4.4. Surface Temperature Inversion

The radiosonde data set contains some tropical profiles that exhibit a surface temperature inversion. Given this anomalous atmospheric situation (the temperature inversion is often accompanied by a quite dry layer immediately above it), one may expect that the standard SST algorithms may not provide good estimates of SST in these conditions. In this analysis a surface inversion was deemed to exist if the surface temperature was > 288 K and the temperature at any height below 3 km

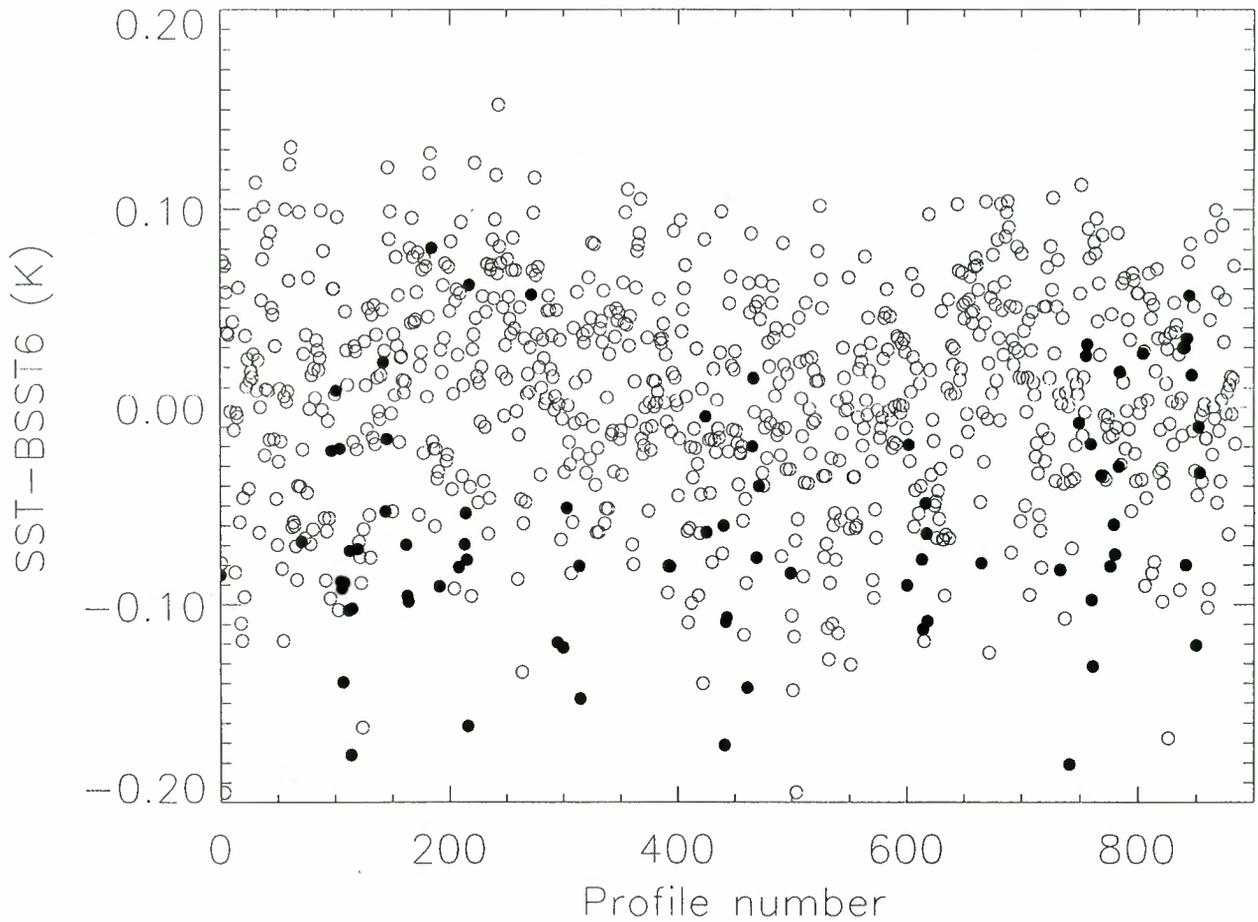


Figure 2. The partition of errors in the BSST6 algorithm (SST-BSST6) for surface temperature inversions. The solid circles are for those profiles where there is an inversion present, while the open circles are for the remaining profiles.

was greater than the surface temperature by more than 0.5 K. The results of this partitioning are given in Table 1. Here the results show a larger effect in the BSST2 and BSST6 algorithms than in the ZSST cases. However, when compared with the standard errors of the data sets given above, the effect appears to be larger in the six-channel algorithms. For the BSST6 case the mean error difference is greater than the standard error of 0.060 K. For this case the partitioning is shown in Figure 2.

5. First Stage of Data Analysis

In the first stage of the data analysis a linear SST algorithm was developed using the number of selected data inputs (channels) from the simulated data set, and the standard error associated with the SST derivation was calculated. The full data set was then partitioned depending on the value of a parameter that is either contained in the data set or derived from the data set. A linear SST algorithm was then derived for each set of partitioned data, and an associated standard error was also obtained. Finally, a standard error was derived for the whole data set (2655 cases) using the separate SST algorithms developed for each group. In most cases

there were two options for the calculation of the partition parameter and the estimate of the final standard error. First, the partition parameter used (say total water vapor) could be that contained in the simulated data set and is the value obtained from the radiosonde data. Alternatively, the parameter could be calculated from the input data that were available (for example, the total water vapor would be derived from a linear regression of water vapor with the data from the available channels). For each selection of input data (channels) the results are presented in a table that is accompanied by a short discussion.

5.1. All Simulated ATSR Data

In this section, all ATSR data are assumed to be available, and the results are given in Table 2. This includes a simulated measure of total water vapor content from the passive microwave radiometer as well as the six simulated ATSR infrared brightness temperatures. In this case the standard error (with the simulated data) is extremely small at 0.0585 K. When the data are partitioned by latitude or by nadir split-window temperature difference (T11N-T12N), there is no improvement in the overall retrieval accuracy. The slight increase in the combined error for T11N-T12N is due to the ef-

Table 2. Standard Errors in the Derivation of SST Using the Seven ATSR Channels for the Complete and Partitioned Data Sets

Partition Parameter	Range	Number	Standard Error, K	Combined Error, K
None	all data	2655	0.0585	0.0585
Absolute latitude	< 25°	1437	0.060	
	> 25°	1218	0.056	0.0582
T11N-T12N	< 1.0 K	1250	0.056	
	> 1.0 K	1405	0.061	0.0589
WV	< 2.00 cm	1245	0.052	
	> 2.00 cm	1410	0.061	0.0575
WV-W2 (total water vapor above 2 km)	< 0.30 cm	801	0.045	
	> 0.30 cm	1854	0.060	0.0560
WV-W2	< 0.50 cm	1200	0.047	
	> 0.50 cm	1455	0.061	0.0552
WV-W2	< 1.00 cm	1899	0.055	
	> 1.00 cm	756	0.062	0.0572
WV-W2*	< 0.50 cm	1285	0.052	
	> 0.50 cm	1370	0.062	0.0574
(WV-W3)/WV	< 0.10	1128	0.051	
	> 0.10	1527	0.060	0.0560
WV-W3	< 0.25 cm	1293	0.051	
	> 0.25 cm	1362	0.061	0.0561

ATSR is the along-track scanning radiometer. The seven ATSR channels include six infrared channels and water vapor. W2 and W3 refer to the total water vapor (WV) below heights of 2 and 3 km. The parameters are obtained from the full simulated data set; that is, WV-W2 is obtained from the radiosonde data.

* The partition parameter is derived from the available remotely sensed data. These are the six infrared channels and the total water vapor content from the radiosonde profile but assumed to be measured by the microwave radiometer.

fect of including a noise temperature of 0.04 K with all brightness temperatures. There is some improvement when the data are partitioned using total water vapor content, but the best improvement is for partitioning by the water vapor distribution using the total water vapor content above 2 km. The results in Table 2 also show that there are optimum values to be adopted for the partition parameter; a value of 0.5 cm for the total

water vapor above 2 km (WV-W2) gives greater benefits than a value of 1.0 cm. Also, as expected, the improvements are better when the partition parameter is independently available rather than being derived from the data themselves. Although the improvement here is small, it does show that improved estimates of SST are possible using careful partitioning of the data. The analysis has also been undertaken with only the

Table 3. Standard Errors in the Derivation of SST Using the Three ATSR Nadir Channels for the Complete and Partitioned Data Sets

Partition Parameter	Range	Number	Standard Error, K	Combined Error, K
None	All data	2655	0.076	0.076
Absolute latitude	< 25°	1437	0.078	
	> 25°	1218	0.068	0.074
T11N-T12N	< 1.0 K	1250	0.067	
	> 1.0 K	1405	0.079	0.073
WV	< 3.0 cm	1833	0.064	
	> 3.0 cm	822	0.084	0.071
WV*	< 3.0 cm	1935	0.069	
	> 3.0 cm	720	0.079	0.072
WV-W2	< 0.50 cm	1200	0.058	
	> 0.50 cm	1455	0.085	0.074
WV	< 1.5 cm	720	0.057	
	1.5 < WV < 2.5	876	0.067	
	2.5 < WV < 3.5	465	0.077	
	> 3.5 cm	594	0.078	0.069
WV*	< 1.5 cm	489	0.067	
	1.5 < WV < 2.5	1152	0.066	
	2.5 < WV < 3.5	518	0.077	
	> 3.5 cm	496	0.077	0.070

See Table 2 for denotation.

* The partition parameter is derived from the available remotely sensed data.

Table 4. Standard Errors in the Derivation of SST Using the Four ATSR Split-Window Channels for the Complete and Partitioned Data Sets

Partition Parameter	Range	Number	Standard Error, K	Combined Error, K
None	All data	2655	0.217	0.217
Absolute latitude	< 25°	1437	0.222	
	> 25°	1218	0.158	0.195
T11N-T12N	< 1.0 K	1250	0.146	
	> 1.0 K	1405	0.225	0.192
WV	< 3.0 cm	1833	0.172	
	> 3.0 cm	822	0.215	0.186
WV-W3	< 0.25 cm	1293	0.175	
	> 0.25 cm	1362	0.224	0.201
WV	< 1.5 cm	720	0.108	
	1.5 < WV < 2.5	876	0.174	
	2.5 < WV < 3.5	465	0.184	
	> 3.5 cm	594	0.198	0.167
WV*	< 1.5 cm	594	0.105	
	1.5 < WV < 2.5	925	0.171	
	2.5 < WV < 3.5	579	0.227	
	> 3.5 cm	557	0.213	0.182

See Table 2 for denotation.

* The partition parameter is derived from the available remotely sensed data.

Table 5. Standard Errors in the Derivation of SST Using the Two Nadir ATSR Split-Window Channels for the Complete and Partitioned Data Sets

Partition Parameter	Range	Number	Standard Error, K	Combined Error, K
None	All data	2655	0.297	0.297
Absolute latitude	< 25°	1437	0.317	
	> 25°	1218	0.158	0.256
T11N-T12N	< 1.0 K	1250	0.144	
	> 1.0 K	1405	0.307	0.244
T11N	< 292 K	1201	0.211	
	> 292 K	1454	0.355	0.279
WV	< 3.0 cm	1833	0.176	
	> 3.0 cm	822	0.348	0.242
WV*	< 3.0 cm	1908	0.188	
	> 3.0 cm	747	0.351	0.245
(WV-W3)/WV	< 0.10	1128	0.174	
	> 0.10	1527	0.305	0.257
WV-W3	< 0.25 cm	1293	0.169	
	> 0.25 cm	1362	0.310	0.251
WV	< 1.0 cm	141	0.066	
	1.0 < WV < 2.0	1104	0.138	
	2.0 < WV < 3.0	588	0.208	
	> 3.0 cm	822	0.348	0.235
WV	< 1.5 cm	720	0.094	
	1.5 < WV < 2.5	876	0.190	
	2.5 < WV < 3.5	465	0.267	
	> 3.5 cm	594	0.362	0.232
WV-W3	< 0.10 cm	789	0.131	
	0.1 < WV-W3 < 0.25	504	0.196	
	0.25 < WV-W3 < 0.6	669	0.237	
	> 0.60 cm	693	0.331	0.235
WV-W3*	< 0.10 cm	437	0.149	
	0.1 < WV-W3 < 0.25	658	0.164	
	0.25 < WV-W3 < 0.6	891	0.228	
	> 0.60 cm	669	0.352	0.242

See Table 2 for denotation.

* The partition parameter is derived from the available remotely sensed data.

six infrared channels, and again, the improvement by partitioning is minimal.

5.2. Nadir ATSR Infrared Data Only (Three Channels)

The standard error for deriving the SST from the three nadir ATSR channels is < 0.1 K when the data set is not partitioned. This low value is mainly due to the relative insensitivity of the $3.7 \mu\text{m}$ channel to water vapor variations. It is also possible to partition the data into more than two groups, and in this case the benefits of dividing the data into four groups have been explored. The results in Table 3 indicate some minor increase in accuracy for this four-way partitioning. There is, of course, a limit to the number of thresholds that can be used in this process, and this will only be determined after careful analysis. It also may be possible to derive coefficients that are interpolated from one group to the next (as is done with latitude in the current ATSR operational algorithms).

5.3. Nadir and Forward ATSR Split Window (Four Channels)

When the $3.7 \mu\text{m}$ data are not available (for example, during the day), the standard error for the full data set increases to a value larger than 0.2 K. When the data are partitioned using the total water vapor content, there is a considerable increase in the SST accuracy. The results in Table 4 show that in each of the partitioned groups there is an increase in accuracy over that for the case when there is no partitioning. However, the benefits are not so great when the data themselves are used to derive the partition parameter.

5.4. Nadir ATSR Split Window (Two Channels)

This case also applies to the application of AVHRR data to derive SST using the two split-window channels at wavelengths of 11 and $12 \mu\text{m}$. Here the SST accuracy for the ATSR case is near 0.3 K, and for the AVHRR split-window channels with a noise temperature of 0.1 K

Table 6. Results From the Application of the Partitioning Technique to the Simulated ATSR Data Set

Algorithm	Partition Parameter	Range	Number	Standard Error, K	Combined Error, K
NLSST	none	all data	2655	0.248	0.248
	Absolute latitude	$< 25^\circ$	1437	0.281	0.235
		$> 25^\circ$	1218	0.163	
	WV	< 3.0 cm	1833	0.173	0.223
		> 3.0 cm	822	0.305	
	WV-W3	< 0.25 cm	1293	0.139	0.228
		> 0.25 cm	1362	0.288	
	WV-W4	< 0.12 cm	1347	0.159	0.226
		> 0.12 cm	1308	0.279	
	WV-W5	< 0.08 cm	1560	0.173	0.227
		> 0.08 cm	1095	0.288	
	(WV-W3)/WV	< 0.10	1128	0.140	0.226
		> 0.10	1527	0.272	
	(WV-W4)/WV	< 0.05	1227	0.167	0.228
		> 0.05	1428	0.269	
	(WV-W5)/WV	< 0.03	1233	0.182	0.233
		> 0.03	1422	0.280	
T11N-T12N	< 1.0 K	1250	0.104	0.227	
	> 1.0 K	1405	0.281		
T11N	< 292 K	1201	0.142	0.217	
	> 292 K	1454	0.264		
WVSST	none	all data	2655	0.258	0.258
	WV-W3	< 0.25 cm	1293	0.117	0.223
		> 0.25 cm	1362	0.291	
	(WV-W3)/WV	< 0.10	1128	0.116	0.221
		> 0.10	1527	0.274	
	T11N	< 292 K	1201	0.156	0.250
> 292 K		1454	0.306		
QUSST	none	all data	2655	0.242	0.242
	WV-W3	< 0.25 cm	1527	0.265	0.222
		> 0.25 cm	1128	0.134	
	(WV-W3)/WV	< 0.10	1227	0.158	0.217
		> 0.10	1428	0.258	
	T11N-T12N	< 1.0 K	1250	0.148	0.242
> 1.0 K		1405	0.302		

The algorithms applied to the data are the NLSST, WVSST, and the QUSST described by May [1993] and Emery *et al.* [1994]. See Table 2 for denotation.

the standard error increases to 0.32 K. In this case the standard error is reduced when the data are partitioned in many different ways, but the best results were found for the cases when the partition parameter was the total water vapor and the water vapor above 3 km. The results are given in Table 5.

6. Second Stage of Data Analysis

In the first stage of the analysis it was found that all algorithms were improved if some dependency on total water vapor content was included in the algorithm specification. The improvement was most noticeable when using the split-window nadir channels (e.g., for AVHRR daytime data). This is in agreement with several recent reports where latitude, brightness temperature difference, brightness temperature magnitude, and SST itself has been used as a surrogate for total water vapor. For example, the current National Oceanic and Atmospheric Administration (NOAA) operational algorithm, the nonlinear SST (NLSST) described by *May* [1993], uses a forecast value of SST to modify the algorithm coefficients, while *Emery et al.* [1994] include a term based on the product of the total water vapor and the brightness temperature difference in their water vapor SST (WVSST). *Barton* [1995] has suggested the inclusion of a term based on the square of the brightness temperature difference in the standard linear SST algorithm. *Emery et al.* [1994] have called this latter algorithm the quadratic SST (QUSST). The form of these three algorithms (for nadir viewing only) are

$$\text{NLSST} = a_0 + a_1 T_{11N} + a_2 (T_{11N} - T_{12N}) \text{SSTf} \quad (1)$$

$$\text{WVSST} = b_0 + b_1 T_{11N} + b_2 T_{12N} + b_3 (T_{11N} - T_{12N}) W \quad (2)$$

$$\text{QUSST} = c_0 + c_1 T_{11N} + c_2 T_{12N} + c_3 (T_{11N} - T_{12N})^2 \quad (3)$$

where a , b and c are coefficients, W is total water vapor and SSTf is a forecast SST.

In this second stage of the analysis the technique developed above has also been applied to these three algorithms, and the results are given in Table 6. For the NLSST the forecast SST used in the algorithm was that derived using the six-channel linear algorithm derived from the full simulated ATSR data set, namely,

$$\begin{aligned} \text{SST} = & -0.2926781 - 0.5024930(T_{12N}) + 0.5285400(T_{11N}) \\ & + 1.422118(T_{37N}) - 0.08964696(T_{12F}) \\ & + 0.03976154(T_{11F}) - 0.3941395(T_{37F}). \end{aligned} \quad (4)$$

For the NLSST algorithm the best improvement was found when the data were partitioned using the 11 μm brightness temperature, but there was also improvement when total water vapor and upper level water vapor were used to divide the data. For both the WVSST

and the QUSST the only improvement found was when the data were partitioned using the upper level water vapor content. The ratio of water vapor content below 3 km to the total water vapor gave a slightly better improvement than the total water vapor above 3 km.

7. Test with Real ATSR Data

7.1. First Stage

The ATSR data set described previously has been used to ensure that the partitioning procedure is valid when applied to real ATSR data as well as those simulated using the atmospheric transmission model. The major concern with the ATSR data set is that an independent measure of SST is not available. To fully test the performance of the new technique, it will be necessary to use a full ATSR validation data set consisting of a large number of surface-based SST measurements with the coincident ATSR data. However, if it is assumed that the temperature derived using all six ATSR infrared channels is correct, then it is possible to test the procedure by then applying the technique developed for the two nadir split-window channels. This may not be the optimum procedure for testing the technique, but in the absence of a large validation data set it is perhaps the only way to proceed. For AVHRR data the technique may be tested using the NOAA data set of combined satellite data and SST values from their network of drifting buoys.

The global six-channel SST algorithm given by (4) was applied to the 2313 cases in the real ATSR data set. These SST values were then assumed to be correct. The partitioning technique using the two nadir split-window channels was then applied to this data set. First, the real ATSR data were partitioned using the values contained in the data set itself. This was possible when latitude, water vapor, brightness temperature, and brightness temperature differences were taken as the partitioning parameters, but for those cases when the parameter was not included in the basic data set (e.g., total water vapor above 3 km) or assumed to be absent (e.g., no microwave measurement of total water vapor) the parameter was derived using a linear algorithm developed using the simulated data set. The partitioned sets of data were then used to provide a value of SST, and then the data sets were combined to give a final standard error that was compared to that derived with no data partitioning. In this analysis of the real ATSR data set it was found that a bias of 0.4 K existed between the SST values derived using the six- and two-channel algorithms derived from the simulated data set. This difference is due to the difficulties associated with accurate modeling of the radiative transfer through the atmosphere; water vapor absorption coefficients in the thermal infrared bands are still not well defined. Similar biases have been reported in past ATSR validation data analyses [*Mutlow et al.*, 1994; *Barton et al.*, 1995]. This bias (0.4 K) was removed from the data before the standard errors were derived. The results are given in Table 7.

Table 7. Results From the Application of the Partitioning Technique to the Real ATSR Data Set

Partition Parameter	Range	Number	Standard Error, K	Combined Error, K
None	all data	2313	0.283	0.283
Absolute	< 25°	907	0.276	
latitude	> 25°	1406	0.179	0.237
WV	< 3.0 cm	1816	0.194	
	> 3.0 cm	497	0.324	0.243
WV*	< 3.0 cm	2054	0.209	
	> 3.0 cm	259	0.346	0.245
WV-W3*	< 0.25 cm	1177	0.205	
	> 0.25 cm	1136	0.292	0.269
(WV-W3)/WV*	< 0.10	576	0.222	
	> 0.10	1737	0.267	0.274
T11N-T12N	< 1.0 K	1399	0.193	
	> 1.0 K	914	0.300	0.255
T11N	< 292 K	1395	0.185	
	> 292 K	918	0.312	0.251

The algorithms applied to the data are those developed from the simulated data set using the two nadir split-window channels. The true SST is assumed to be that derived using the six infrared channel algorithm developed from the simulated data set. See Table 2 for denotation.

* The partition parameter is derived from the available remotely sensed data.

The real ATSR data set includes data from a wide range of latitudes as well as distances from the sub-satellite track. It is thus encouraging that when the partitioning technique is applied, the results are quite similar to those obtained using the simulated data set. Although not proving that the technique works, it suggests that it will supply improved SST values when implemented with large ATSR data sets.

To fully implement the technique, it will be necessary to select carefully the partitioning parameter and to determine the number and value of the thresholds applied. Care must be taken not to retreat to the derivation and application of a single global algorithm.

7.2. Second Stage

The NLSST, WVSST, and QUSST algorithms derived in the first stage of the analysis have also been applied to the real ATSR data set. Again, the SST values for this data set are assumed to be those produced by the six-channel algorithm given by (4). The results are presented in Table 8 and show that for the NLSST algorithm the standard error of the undivided data set was 0.329 K, which is larger than that for the simulated data set. There is some improvement in the SST accuracy when the data are partitioned using the measured water vapor or the 11 μm brightness temperature, while the use of the upper level water vapor increases the standard error. This increase is largely due to errors in partitioning. Basically, the NLSST algorithm uses only two terms (channels), namely, (T11N-T12N)(SSTf) and T11N, where SSTf is the forecast surface temperature (in degrees Celsius). These two terms are not sufficient to derive a reasonable estimate of the upper level water vapor, and thus the partitioning technique will not be applicable. Thus, for this algorithm, the upper level water vapor was determined using seven-channel algorithms (six infrared plus water vapor) derived from the

simulated data set and applied to the real ATSR data. When this was done, the partitioning of the data into two NLSST algorithms using upper level water vapor as the partitioning parameter showed some decrease in the standard error (see Table 8). When the WVSST and QUSST algorithms were applied to the real ATSR data set, the SST accuracy was only improved when the data were partitioned using the upper level water vapor amount.

8. General Discussion

The data analysis undertaken in this paper shows that there is an improvement in the accuracy of the derived SST if the data set can be partitioned either using a parameter contained in the data set, an external parameter, or one that is derived using the data set itself. In this latter case the improvement is not as great as when the parameter is explicitly defined or when it is available from an external source. The improvement is not significant for ATSR data sets in which all the infrared data are available to derive the SST, but if the 3.7 μm channel data are not available, then the improvement is considerable.

In all cases the most accurate SST values were obtained for those data sets with low total water vapor (high latitude, low brightness temperatures, and low upper level water vapor). The analysis also show that the best results will be obtained when the (partitioning) parameter can be specified from external data sources and not derived using the satellite data itself. This is most important for the AVHRR case where the water vapor is required for the WVSST algorithm. If this parameter is not available from other sources and is developed from the AVHRR data themselves, then the algorithm will probably be equivalent to the QUSST as

Table 8. Results of Testing the Partitioning Technique With the ATSR Data Set

Algorithm	Partition Parameter	Range	Number	Standard Error, K	Combined Error, K
NLSST	none	all data	2313	0.329	0.329
	WV	< 3.0 cm	1816	0.263	
		> 3.0 cm	497	0.328	0.278
	WV-W3*	< 0.25 cm	917	0.201	
		> 0.25 cm	1396	0.543	0.441
	(WV-W3)/WV*	< 0.10	530	0.203	
		> 0.10	1783	0.463	0.418
	T11N	< 292 K	1395	0.209	
		> 292 K	918	0.265	0.232
	WV-W3†	< 0.25 cm	1405	0.217	
		> 0.25 cm	908	0.347	0.275
	(WV-W3)/WV†	< 0.10	1587	0.222	
> 0.10		726	0.438	0.307	
WVSST	none	all data	2313	0.213	0.213
	WV-W3*	< 0.25 cm	1124	0.189	
		> 0.25 cm	1189	0.244	0.219
	(WV-W3)/WV*	< 0.10	582	0.209	
		> 0.10	1731	0.210	0.210
	T11N	< 292 K	1395	0.168	
		> 292 K	918	0.258	0.208
	WV-W3†	< 0.25 cm	1405	0.291	
		> 0.25 cm	908	0.117	0.207
	(WV-W3)/WV†	< 0.10	1587	0.173	
		> 0.10	726	0.216	0.188
	QUSST	none	all data	2655	0.242
WV-W3*		< 0.25 cm	1219	0.191	
		> 0.25 cm	1094	0.273	0.233
(WV-W3)/WV*		< 0.10	568	0.209	
		> 0.10	1745	0.255	0.245
T11N		< 292 K	1395	0.170	
		> 292 K	918	0.308	0.235
WV-W3†		< 0.25 cm	1405	0.181	
		> 0.25 cm	908	0.293	0.231
(WV-W3)/WV†		< 0.10	1587	0.196	
		> 0.10	726	0.300	0.234

The algorithms applied to the data are the NLSST, WVSST, and the QUSST described by *May* [1993] and *Emery et al.* [1994]. See Table 2 for denotation.

* The partition parameter is derived from the available remotely sensed data.

† The partition parameter is derived using a linear seven-channel algorithm derived from the simulated data set but applied to the real ATSR data.

the total water vapor content is likely to have a linear dependence on brightness temperature difference in the two channels.

In the first stage of the analysis the results indicate that for the split-window nadir case (AVHRR daytime) the primary source of error in simple linear algorithms is correlated with total water vapor content or one of its surrogates. This has now been well established by *May* [1993], *Emery et al.* [1994], *Barton* [1995], and others and has led to the development and application of the NLSST, WVSST, and QUSST algorithms for AVHRR data. In these cases it is not necessary to develop discrete algorithms for partitioned groups of data, but the sensitive parameter has been incorporated into the algorithm by including another term (as in the WVSST and QUSST), or the coefficients are dependent on water vapor or a surrogate (e.g., SSTf in the NLSST).

In the second stage of analysis, where the NLSST, WVSST, and QUSST algorithms are examined, there are some interesting results. The NLSST still appears to be sensitive to latitude, water vapor, and the 11 μm brightness temperature, while the WVSST and QUSST are only sensitive to the vertical distribution of water vapor. These results support the initial premise of this paper, that, after the dependence on total water vapor content is accounted for, the remaining errors in satellite-derived SST are due to anomalies in vertical profiles of water vapor and temperature; that is, situations where the first guess of the atmospheric state is markedly wrong. The strong role that upper level water vapor plays in determining SST derivation accuracy has also been discussed by *Minnett* [1986].

The preliminary analysis in this paper also suggests that separate algorithms should be developed for those

cases where a surface temperature inversion existed in tropical areas. A full analysis of this situation has not been attempted in this work, but the techniques applied to variable upper level water vapor content could also easily be applied to this phenomenon.

The technique can also be used to incorporate changes in aerosol content into the satellite data analysis. With AVHRR and ATSR 2 an independent (from the thermal channels) estimate of aerosol content is possible using the visible-near infrared channels, and SST algorithm coefficients can then be developed to account for variations in aerosol content. These algorithms should then provide reliable estimates of SST, even after violent volcanic eruptions that deposit large amounts of aerosol into the stratosphere.

9. Conclusions

A technique has been developed that enables the detection of those parameters which limit the accuracy of satellite-derived SST values. For simple linear algorithms, as proposed by *Závody et al.* [1995], the SST errors are mainly associated with total water vapor content or one of its surrogates (e.g., latitude, SST, brightness temperature difference, and water vapor content of a defined atmospheric layer). For the case when only the split-window channels in the nadir view are available (the AVHRR daytime case), algorithms have been developed that include a dependence on one of these terms. When a measure of water vapor is available, the WVSST may be the best algorithm to use, but in the absence of a water vapor measurement the QUSST appears to give better results than the NLSST. For all these algorithms the SST derivation can also be improved if information on the water vapor content of the atmosphere above 3 km is available.

For ATSR, when all its data are available, the benefits of introducing a new set of algorithms may be minimal as the SST is (theoretically) accurately derived. However, there may still be some improvements in refining the existing algorithms. The analysis suggests that it may be better to define algorithms for ranges of water vapor rather than for latitude as is done with the existing operational algorithms. Also it may be worth exploring the possible development of sets of algorithm coefficients depending on the vertical water vapor profile derived from the data itself (using the techniques described by *Barton* [1996]) or available from external sources.

For situations where $3.7 \mu\text{m}$ data are not available (ATSR and AVHRR during the day) the analysis procedure is clearer. First, an algorithm including a term based on total water vapor should be developed. The QUSST or WVSST algorithms are suggested depending on whether a reliable measure of water vapor is available. This is also true for the ATSR four-channel algorithm, where the inclusion of an extra term improves the algorithm performance. Once these algorithms have been developed, further improvements are possible if information on the water vapor content of the atmosphere above (say) 3 km is available.

In most situations the details of water vapor distribution will not be available from satellite data. If they are, such as in the case of ATSR with all channels available, then the extra data can be used directly to provide SST. However, if the satellite data can be analyzed together with a numerical weather model, it may be possible to include information on the state of the atmosphere in the SST determination. Ultimately, the analysis of all meteorological satellite data should be performed in such a manner so that information on the state of the atmosphere (clouds, aerosols, water vapor temperature, etc.) and the surface (roughness, wind speed, surface type, etc.) can be incorporated into the analysis of the satellite data.

Meanwhile, opportunities exist to further test the performance of the analysis procedure suggested above for AVHRR. Mid-level water vapor content can be derived using the TIROS-N operational vertical sounder (TOVS) suite of instruments on the NOAA satellites, and data from these instruments could be incorporated into SST analyses. The latest geostationary satellites include split-window channels at 11 and $12 \mu\text{m}$ as well as a channel located on the wing of the $6.3 \mu\text{m}$ water vapor absorption line. This channel is sensitive to the water vapor in the layers near 3 km, and so the data from these satellites should be well suited to proving the technique. If the data from this water vapor channel are useful in improving SST estimates, then such a channel could be incorporated into new designs of the AVHRR.

The analysis described in this paper has been developed using simulated satellite data that have been produced from a radiative transfer model and a collection of radiosonde data. The technique has been applied to ATSR data but with an assumption that the SST derived using the ATSR data itself is correct. Techniques developed using simulated data must always be tested on real satellite data to ensure that the assumptions that are necessary in the development of any simulated data set are valid. This highlights the need for the development of a large validation data set for all satellite instruments. Such a set exists for SST derived using AVHRR, but for ATSR the lack of a large number of high-quality validation cases restricts the final testing of several new techniques. This is the case for the use of ATSR data in deriving land surface temperature and water vapor distribution as well as SST.

Finally, the selection of optimum partition parameters and thresholds for different sets of data will require much careful analysis. However, when implemented, the new SST algorithms should account for a wider range of atmospheric conditions than those algorithms currently in operational use.

Acknowledgments. The ATSR data are used by courtesy of the European Space Agency. The Rutherford Appleton Laboratory (RAL) processed and provided the ATSR data products. The author would also like to thank Albin Závody (at RAL, United Kingdom) for supplying a version of his radiative transmission model, Laurence Eymard (at CETP, France) for the microwave data and the analysis algorithm, and Janice Bathols (at CSIRO, Australia) for her painstaking analysis of the ATSR data set.

References

- Baker, P.W., and A.M. Dowd, The occurrence of radio ducting, superrefractive and subrefractive conditions in Australian maritime regions, *Tech. Rep. ERL-0048-TR*, 26 pp. and 9 figs., Electron. Res. Lab., Def. Sci. and Technol. Organ., Dep. of Def., Adelaide, 1978.
- Barton, I.J., Satellite-derived sea surface temperatures: current status, *J. Geophys. Res.*, *100*, 8777-8790, 1995.
- Barton, I.J., Water vapour analyses using ATSR data, in Proceedings of the 1st Australian ERS-1 Symposium, pp. 1-10, CSIRO Office of Space Science and Applications, Canberra, 1996.
- Barton, I.J., A.M. Závody, D.M. O'Brien, D.R. Cutten, R.W. Saunders, and D.T. Llewellyn-Jones, Theoretical algorithms for satellite-derived sea surface temperatures, *J. Geophys. Res.*, *94*, 3365-3375, 1989.
- Barton, I.J., A.J. Prata and R.P. Cechet, Validation of the ATSR in Australian waters, *J. Atmos. Oceanic Technol.*, *12*, 290-300, 1995.
- Delderfield, J., D.T. Llewellyn-Jones, R. Bernard, Y. de Javel, E.J. Williamson, I.M. Mason, D.R. Pick, and I.J. Barton, The along track scanning radiometer (ATSR) for ERS-1, in *Instrumentation for Optical Remote Sensing From Space*, edited by J.S. Seeley et al., *Proc. SPIE Int. Soc. Opt. Eng.*, *589*, 114-120, 1986.
- Emery, W.J., Y. Yu, G.A. Wick, P. Schluessel, and R.W. Reynolds, Correcting infrared satellite estimates of sea surface temperature for atmospheric water vapor attenuation, *J. Geophys. Res.*, *99*, 5219-5236, 1994.
- Eyre, J.R., On systematic errors in satellite sounding products and their climatological mean values, *Q. J. R. Meteorol. Soc.*, *113*, 279-292, 1987.
- Mason, G., Satellite measurements of sea surface temperature, D. Phil. thesis, Oxford Univ., Oxford, England, 1991.
- May, D.A., Global and regional comparative performance of linear and nonlinear satellite multichannel sea surface temperature algorithms, *Rep. NRL/MR/7240-93-7049*, 34 pp., Nav. Res. Lab., Washington, D.C., 1993.
- McClain, E.P., Multi-channel sea surface temperatures from the AVHRR on NOAA-7, in *Proceedings of the Satellite-Derived Sea Surface Temperature: Workshop-II*, *JPL Publ.*, *84-5*, 1-8, 1984.
- Minnett, P.J., A numerical study of the effects of anomalous North Atlantic atmospheric conditions on the infrared measurement of sea surface temperature from space, *J. Geophys. Res.*, *91*, 8509-8521, 1986.
- Mutlow, C.T., A.M. Závody, I.J. Barton, and D.T. Llewellyn-Jones, Sea surface temperature measurements by the along-track scanning radiometer (ATSR) on the ERS 1 satellite: Early results, *J. Geophys. Res.*, *99*, 22,575-22,588, 1994.
- Rodgers, C.D., Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation, *Rev. Geophys.*, *14*, 609-624, 1976.
- Smith, A.H., R.W. Saunders, and A.M. Závody, The validation of ATSR using aircraft radiometer data over the tropical Atlantic, *J. Atmos. Oceanic Technol.*, *11*, 789-800, 1994.
- Walton, C.C., Nonlinear multichannel algorithms for estimating sea surface temperature with AVHRR satellite data, *J. Appl. Meteorol.*, *27*, 115-124, 1988.
- Závody, A.M., C.T. Mutlow, and D.T. Llewellyn-Jones, A radiative transfer model for sea surface temperature retrieval for the along track scanning radiometer (ATSR), *J. Geophys. Res.*, *100*, 937-952, 1995.

I. J. Barton, CSIRO Division of Atmospheric Research, CSIRO Marine Laboratories, P. O. Box 1538, Hobart, Tasmania 7001, Australia. e-mail: ian.barton@ml.csiro.au.

(Received December 18, 1996; revised July 9, 1997; accepted September 8, 1997.)

Actual and potential information in dual-view radiometric observations of sea surface temperature from ATSR

M. J. Murray, M. R. Allen,¹ C. T. Mutlow, and A. M. Závody

Space Science Department, Rutherford Appleton Laboratory, Chilton, England, United Kingdom

M. S. Jones and T. N. Forrester

James Rennell Division, Southampton Oceanography Centre, Southampton, England, United Kingdom

Abstract. The along-track scanning radiometer (ATSR) on ERS-1 has delivered a continuous global record of radiometric (skin) sea surface temperature (SST) since August 1991. We present a comprehensive analysis of the large-scale and low-frequency characteristics of the data set using direct comparison with other global SST analyses to develop a quantitative understanding of the various factors contributing to the accuracy of and sources of bias in the first 4 years of the ATSR SST record. Such a global analysis is a necessary complement to direct validation against in situ observations, since large-scale sources of bias may be indistinguishable from instrument noise or sampling uncertainty in individual validation campaigns. No large-scale features attributable to atmospheric contamination through aerosols or water vapor are discernible in the difference between the three-channel, dual-view ATSR SSTs and the National Oceanic and Atmospheric Administration (NOAA) blended analysis of advanced very high resolution radiometer (AVHRR) and in situ SST observations. Features in the difference field can be traced to known deficiencies in the data used in the NOAA operational analysis or tentatively related to skin-bulk temperature differences in certain regions. Atmospheric contamination is, however, evident in the ATSR two-channel SST retrieval (used in daytime and after the failure of the 3.7 μm channel) compared to the NOAA operational analysis. The Pinatubo aerosol plume is the dominant feature of the difference field in the first year of ATSR operation. In all cases, however, the amplitude of the atmospheric signature is significantly lower in dual-view than in corresponding single-view ATSR SSTs, indicating that the potential remains for unbiased two-channel SST retrieval even in the presence of aerosol.

1. Background

A consistent, continuous, and global sea surface temperature (SST) field with an absolute accuracy of around 0.3 K has been called for to satisfy the requirements of ocean and climate research [e.g., *Harries et al.*, 1983; *Allen et al.*, 1994]. Conventional observations made from ships and buoys are concentrated in shipping lanes and are sparse in the southern hemisphere. Observations made by satellite borne infrared radiometers provide near-global coverage, but SST retrieval is inevitably compromised by the difficulty of adequately

compensating for atmospheric effects such as clouds, water vapor, and aerosols.

More than 15 years of SST measurements have been made by the advanced very high resolution radiometer (AVHRR) series of instruments on the National Oceanic and Atmospheric Administration (NOAA) TIROS-N series of operational satellites. These have provided global SSTs with a nominal accuracy of ~ 0.7 K [*McClain et al.*, 1985], and the AVHRR multichannel sea surface temperature (MCSST) product consisting of weekly composites for the globe at ~ 18 km resolution is widely used in operational and research activities. Considerable effort is being devoted to the production of a consistent AVHRR SST record through the National Aeronautics and Space Administration (NASA) Pathfinder project.

The AVHRR thermal channels are calibrated using an internal blackbody target and a space view; the SST retrieval algorithm for the MCSST product is derived

¹Also at Department of Physics, University of Oxford, Oxford, England, United Kingdom.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC02180.
0148-0227/98/97JC-02180\$09.00

using a regression analysis between the AVHRR brightness temperatures and a quality-controlled buoy SST data set. This approach leads to a more accurate SST retrieval at the cost of loss of independence of the various data sources, which complicates the interpretation of long-term trends.

AVHRR SST retrievals were found to be compromised by atmospheric aerosol contamination after the eruption of Mount Pinatubo in June 1991. By September 1991 the number of daytime SST retrievals dropped close to zero in the tropics, as the cloud decontamination algorithms proved to be unreliable in these conditions; negative biases >1 K have been reported in those SST retrievals which were achieved [Reynolds, 1993]. This problem was addressed in late 1991 by the implementation of a revised SST retrieval algorithm which dealt specifically with the increased aerosol contamination.

The first along-track scanning radiometer (ATSR 1) is a four-channel, dual-view, self-calibrating, infrared radiometer with spatially coregistered spectral channels centered at 1.6, 3.7, 10.8 and 12.0 μm . The 1.6 μm channel is useful only in daytime, and its primary purpose is cloud identification. The key experimental feature of ATSR is that it views the same point on the ocean surface twice, over an interval of ~ 100 s, at two different angles through the atmosphere as the satellite passes overhead. This allows a "dual-view" retrieval in which two different atmospheric path lengths are utilized to quantify and correct for the effects of atmospheric emissions and absorption.

Radiometric calibration is achieved using two highly precise, ultrastable, onboard calibration targets maintained at around 265 and 305 K. Both targets are viewed during each ATSR scan (many times every second), and the data are used to continually calibrate the instrument, enabling the determination of single-channel equivalent temperatures correct to within 0.05 K [see Delderfield *et al.*, 1986; Edwards *et al.*, 1990]. A further feature of the ATSR design is the use of a Stirling cycle mechanical cooler, which maintained the detectors at <100 K in the early years of the mission [see Werrett *et al.*, 1985; Bradshaw *et al.*, 1985].

ATSR was designed to measure SSTs to an accuracy of 0.3 K without recourse to surface observations. A standard set of climatological radiosonde profiles was used in the generation of the coefficients in the retrieval algorithm [Zavody *et al.*, 1995], but no postlaunch corrections have been applied to bring the ATSR data into agreement with other SST observations; hence it represents a completely independent data set.

The absolute accuracy of ATSR SSTs can only be evaluated through validation against accurately collocated in situ observations of ocean skin temperature [Minnett, 1991]. Considerable effort has been made in this direction [see, e.g. Mutlow *et al.*, 1994; Barton *et al.*, 1995; Forrester and Challenor, 1995; Donlon and Robinson, 1997]. Nevertheless, the practical difficulty of making accurate in situ measurements of radiometric SST means that the size of validation data sets remains limited. An alternative approach, previously adopted

by Harrison and Jones [1994], is to compare the ATSR data with an independent SST analysis. This cannot provide an absolute measure of instrument accuracy since collocation noise would typically dominate individual ATSR-analysis differences, but it enables the examination of large-scale patterns of bias to search for the spatiotemporal signature of atmospheric contamination. We focus on a comparison with the NOAA operational SST analysis, also popularly known as "Reynold's analysis," as it is widely regarded as the most accurate SST analysis available on this resolution.

2. NOAA Operational Analysis

The NOAA operational SST analysis [Reynolds, 1988; Reynolds and Marisco, 1993; Reynolds and Smith, 1994], hereafter NOAA OA, uses a blend of in situ (ship and buoy) and MCSST data to produce global weekly mean SSTs on a 1° grid. These data are compiled using a two-step procedure in which satellite data are initially corrected for any (spatiotemporally varying) bias by using them to provide second derivative information in a solution of Poisson's equation anchored to in situ data. An optimal interpolation (OI) scheme is then used to produce a continuous analysis. Thus the absolute values of buoy and ship data are given a relatively high weight, and only the derivatives of the satellite are, in effect, used in the OI. In situ SSTs outside the range $-2^\circ\text{C} < T < 35^\circ\text{C}$ are discarded, as are those data for which the SST anomaly lies outside ± 3.5 times the climatological standard deviation. The contributing SSTs represent a mixture of daytime and nighttime observations.

3. ATSR data

A variety of products are derived from ATSR data [Bailey, 1995], but this analysis is based on the spatially averaged sea surface temperature (ASST) product. These data comprise half-degree, spatially averaged SSTs with associated temporal, positional, and confidence information. The data considered here were acquired during the 4 year period from August 1991 to July 1995, inclusive. These data have been made available on a CD-ROM set [Murray, 1995] available from the ATSR Science Team at the Rutherford Appleton Laboratory.

The ATSR SST retrieval scheme follows that of McClain *et al.* [1985], where SST is given by a linear combination of infrared brightness temperatures

$$\text{SST} = a_0 + \sum_{i=1}^N a_i T_i$$

where the a_i are constant (latitude-dependent) coefficients and T_i is the cloud-free scene brightness temperature as observed by ATSR either in the nadir or forward view. A range of retrieval algorithms are possible with this geometry; in the most accurate, a dual-view (that is, using both forward- and nadir-view data) three-spectral-channel (3.7, 10.8, and 12.0 μm) retrieval

which can be used only at night, i ranges from 1 to 6. Dual-view, two-channel (10.8 and 12.0 μm) retrievals are used for daytime observations when sunglint contaminates the 3.7 μm brightness temperatures and also for all retrievals after the premature loss of the 3.7 μm channel on May 27, 1992. Single-view (nadir only) SSTs are computed for comparison with the dual-view product and also to provide an estimate of SST where cloud cover or operational constraints make it impossible to obtain a spatially coregistered dual-view retrieval. Separate coefficients are derived from representative radiosonde data sets and an atmospheric radiative transfer model; three characteristic atmospheric profiles are used (polar, midlatitude, or tropical), and latitude-dependent coefficients are obtained by linear interpolation between them. The same SST retrieval coefficients have been used for the entire ATSR data set, with no attempt made to correct for variation in atmospheric aerosol. A complete description of the ATSR SST retrieval scheme is given by Závody *et al.* [1994, 1995].

ATSR data were averaged into a $1^\circ \times 1^\circ$, weekly spatiotemporal grid matching the spatial and temporal division of the NOAA OA. Only data within 6 K of the *Global Ocean Surface Temperature Atlas* (GOSTA) climatology were used [Bottomley *et al.*, 1990]; this requirement excluded $<0.3\%$ of daytime and $\sim 4\%$ of nighttime ATSR data. Rejection was overwhelmingly because the ATSR SSTs were too cold, probably because of deficiencies in the cloud decontamination algorithm. A more conservative climatological check would improve the nominal ATSR/NOAA OA agreement and might also help in the identification of more subtle sources of bias but would also introduce the risk of excluding valid data in regions of high interannual SST variability. Likewise, a consistency check between daytime and nighttime ATSR data, as used by Jones *et al.* [1996a, b], would also reduce bias and noise.

In the case of nighttime data most cold SSTs were observed in the Southern Ocean at latitudes south of 30°S , although a significant contribution from northern latitudes occurs in boreal summer. Too-cold daytime observations usually arise at midlatitudes in the northern hemisphere, again during boreal summer. Unsurprisingly, the cold bias associated with inadequate cloud decontamination is much less than with nighttime SSTs, since the availability of the 1.6 μm reflectances enables much better cloud identification in daylight. ATSR daytime and nighttime observations have been considered separately throughout this paper, as have dual-view (forward- and nadir-view) and single-view (that is, nadir-only) observations.

4. Comparison of ATSR and NOAA OA SSTs

Any infrared radiometric SST measurement represents the temperature of a water layer <0.1 mm thick; this is typically several tenths of a degree cooler than the temperature a few millimeters below it, which in

turn may be either cooler or warmer than the “bulk” temperature at 1 m depth, depending on the state of the near-surface thermocline. (See Schuessel *et al.* [1990] for a discussion of the factors influencing the sense and magnitude of this “skin-bulk” temperature difference.) In general, skin temperature is subject to much stronger diurnal variation than the bulk temperature as defined above, with a peak skin SST occurring in the late afternoon as a result of solar heating, at which time the ocean skin may be warmer than the bulk temperature. However, most ATSR daytime observations are made between 1030 and 1100 LT and nighttime observations are made about 12 hours later. At these times it is likely that, on average, the ocean skin will be cooler than the bulk temperature. As NOAA OA uses absolute temperatures from in situ SSTs and only the derivatives of the satellite SSTs, it is representative of ocean bulk temperature.

This comparison uses NOAA OA SSTs warmer than -1.8°C , the value used as an ice mask, and excludes data flagged as land or ice in the GOSTA climatology. Improved ice identification is an area of priority activity identified by this survey.

4.1. Zonal Mean Differences

Initially, we investigate time-latitude plots of zonally averaged, weekly mean SST fields (Hofmoller diagrams) (that is, we consider weekly-mean SSTs for a 1° wide latitude band covering all longitudes). Zonal mean fields are appropriate because the major sources of bias between ATSR and NOAA OA (i.e., aerosols, cloud contamination, and the skin-bulk effect) should, to a first approximation, be predominantly zonal in structure. Only ocean data have been included (no lakes or inland seas). All temperatures are given in Kelvin and shown in the range 268 to 304 K for absolute SSTs and -2.0 to 0.5 K for ATSR–NOAA OA differences; therefore a negative difference indicates that an ATSR SST is colder than that from NOAA OA. As explained above, ATSR SSTs might be expected to be slightly cooler than NOAA OA because of the skin-bulk effect and may be further affected by inadequate compensation for atmospheric contamination. Except where specified otherwise, the plates and figures cover the 4 year period from August 1991 to July 1995, inclusive.

Plate 1 represents the global SST field from NOAA OA, and Plate 2 shows the equivalent field for nighttime, dual-view ATSR data. As described above, where possible, nighttime SST retrievals utilize the 11, 12 and 3.7 μm data. However, no 3.7 μm data were transmitted to ground during the period from August 6 to September 13, 1991, and the 3.7 μm channel was permanently lost on May 27, 1992. Subsequent to the loss of the 3.7 μm data, nighttime SST retrievals were initially about 0.6 K cooler than the previous three-channel retrievals. Complications arising from the failure resulted in a short period of missing data around this time (evident as white horizontal lines).

Plate 3 shows the zonal mean of the difference between the ATSR dual-view nighttime SSTs and NOAA OA. Aerosol contamination caused by the Mount Pina-

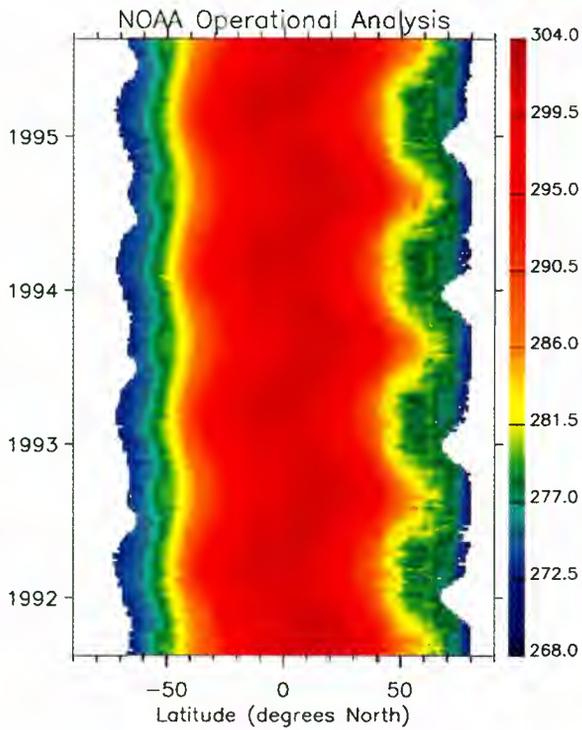


Plate 1. Time-latitude plot of zonal mean sea surface temperature (SST) from the National Oceanic and Atmospheric Administration operational analysis (NOAA OA).

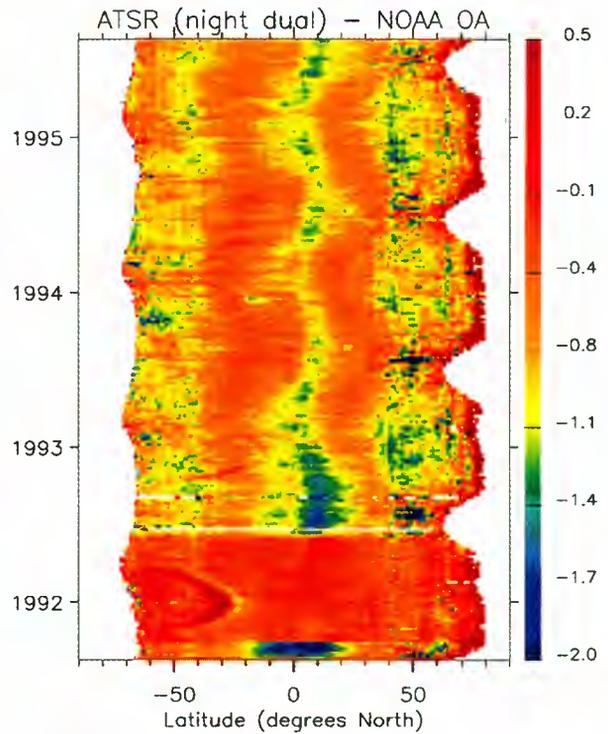


Plate 3. Zonal mean differences between ATSR dual-view nighttime SSTs and NOAA OA. Note revised color scale.

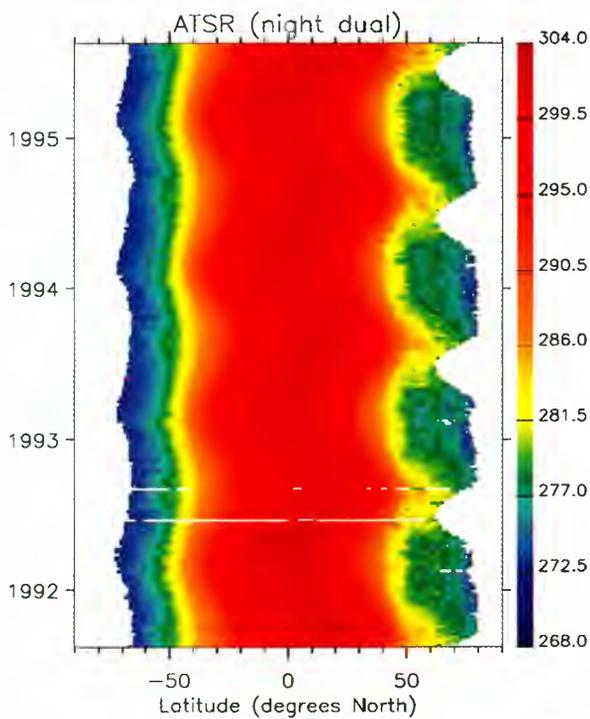


Plate 2. As Plate 1, with along-track scanning radiometer (ATSR) dual-view nighttime SSTs.

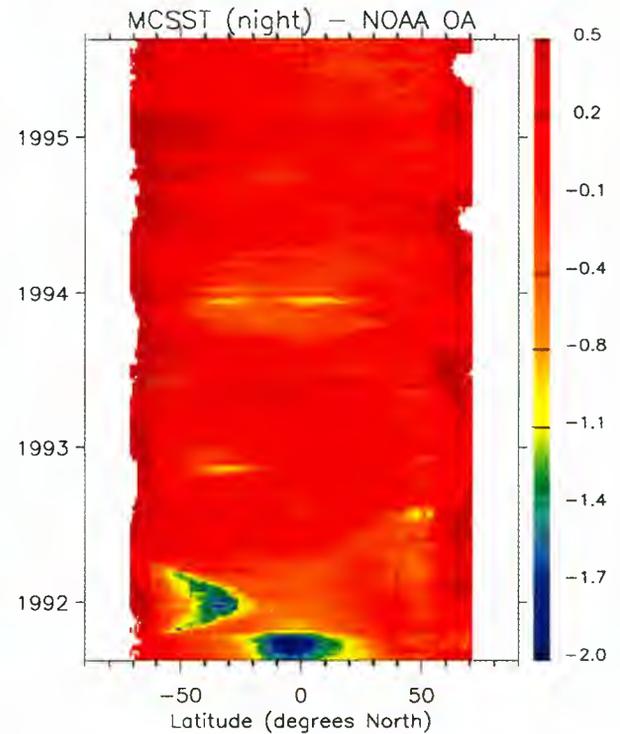


Plate 4. Zonal mean differences between multichannel sea surface temperature (MCSST) nighttime SSTs and NOAA OA.

tubo plume is apparent both at the start of the mission and subsequent to the failure of the 3.7 μm channel in May 1992. However, almost complete elimination of the Pinatubo aerosol is achieved when ATSR 3.7 μm data are available.

This comparison highlights a parabola-shaped anomaly in the southern hemisphere during October 1991 to March 1992. This feature is apparent in the NOAA OA, with affected SSTs being subject to a negative bias of up to 2 K in this region [Reynolds, 1993]. This is due to an AVHRR calibration anomaly affecting the 3.7 μm data and thus the nighttime SSTs. The problem occurs at certain solar-satellite geometries when the satellite moves out of Earth's shadow. The effect was aggravated by the aerosol-corrected algorithm in use at the time, the shortage of daytime SSTs, the local aerosol load from Mount Hudson (46.1°S, 72.6°W), and the usual lack of in situ data in the region.

Plate 4 shows the analogous plot of the difference between the AVHRR MCSST nighttime product and NOAA OA. Cold biases approaching 2 K characterize the tropical MCSST values at the beginning of the period. Subsequent to the implementation of the aerosol-corrected algorithm, the large-scale characteristics of the MCSST product match NOAA OA rather well, except in the region affected by the anomaly discussed above. This is, to some extent, to be expected since MCSST data are used as input to the NOAA OA.

Comparison of ATSR data with and without 3.7 μm data suggests that the standard deviation of the ATSR–NOAA OA differences rose from 0.7 to 0.9 K after the 3.7 μm failure. These figures are based on comparing 40 weeks of data immediately before May 27, 1992, with 40 weeks of data shortly afterward. Only SSTs north of 10°S were considered in both data sets, to exclude the effects of the anomaly in the NOAA OA. These figures give some indication of the impact of the loss of the 3.7 μm channel but are not an indication of the absolute accuracy of either data set, since the ATSR and NOAA OA data have not been accurately collocated; that is, we are comparing weekly 1°×1° areas with no guarantee that observations in the two data sets have been taken at the same time in that spatiotemporal region. Moreover, the smoothing of the NOAA OA would result in a nonzero standard deviation in the difference between the two data sets, even if the unsmoothed data were in perfect agreement.

The success of the ATSR SST three-channel, dual-view retrieval in conditions of severe aerosol contamination is clearly apparent, and the value of the dual-view observation technique can be established by comparing this result with the equivalent ATSR nadir-only SSTs. Plate 5 shows the ATSR–NOAA OA difference field using only ATSR nighttime nadir-view information in the SST retrieval. It is clear that the availability of the 3.7 μm data enables a much improved SST retrieval in the presence of aerosols. However, unlike the dual-view retrieval, aerosol-induced biases of up to 1.5 K are present in the nadir-only SSTs, as indicated by the equator-to-pole gradients in the difference field.

The difference fields shown in Plates 3 and 5 both exhibit a seasonal cycle in equatorial regions and at midlatitudes in the northern hemisphere. This cycle appears more marked in the case of the nadir-only SSTs and is least evident in the case of the dual-view SSTs to which the 3.7 μm data contributed. Since we would expect some seasonality in atmospheric contamination from both aerosol and water vapor, this result is not surprising. Further examination of this seasonal cycle in these biases might prove a useful approach to understanding their origins.

Plates 6 and 7 show the (ATSR–NOAA OA) difference fields using ATSR daytime dual-view SSTs and daytime nadir-view SSTs respectively. ATSR daytime SST retrievals use only the 11 and 12 μm brightness temperatures, with the 1.6 μm data used for cloud decontamination. This two-channel retrieval does not deal with aerosol contamination as effectively as the three-channel nighttime retrieval. Nevertheless, the improvement of the dual-view with respect to the nadir-only SST retrieval is again demonstrated.

Both the latter two plates exhibit noise at high latitudes in the northern hemisphere. The small area of sea contributing to the zonal mean in this region may be partially responsible for this, as may an inadequate ice mask. Another contributory factor is that in this initial processing phase, ATSR dual-view SSTs could not be processed over a 200 km wide strip of the northern Atlantic. This is simply because orbits were initially processed individually, making it impossible to compute a dual-view SST if the necessary data were acquired across the (arbitrary) junction between one orbit and the next. This will be rectified in the planned reprocessing. A strong seasonal cycle at midlatitudes in the northern hemisphere is again apparent and appears particularly clearly in the dual-view SST comparison.

4.2. Time Series From Selected Regions

To allow more quantitative comparison than is possible in Hofmoller diagrams, we show time series of weekly mean temperature differences over selected latitude ranges. High-latitude data have been omitted from the comparisons to avoid the effects of ice contamination, and only gridboxes where both data sets have valid SSTs are used in compiling regional mean differences, to minimize discrepancies caused by sampling.

Figure 1 shows the difference between ATSR dual-view SSTs and the NOAA OA in the latitude range from 70°S to 70°N, with ATSR daytime and nighttime SSTs shown separately. During the period when the 3.7 μm data were available, ATSR nighttime SSTs exhibit a cold bias of only 0.3 K with respect to the NOAA OA. Subsequent to the May 1992 failure, a sudden drop of around 0.7 K occurs, after which nighttime SSTs follow a similar trend to daytime SSTs but are consistently cooler than the latter by about 0.2 K.

Considering the daytime difference, the ATSR SSTs are ~0.8 K cooler than the NOAA OA SSTs in 1991. This is rather large to be associated solely with the skin-bulk difference, and the temporal evolution of the

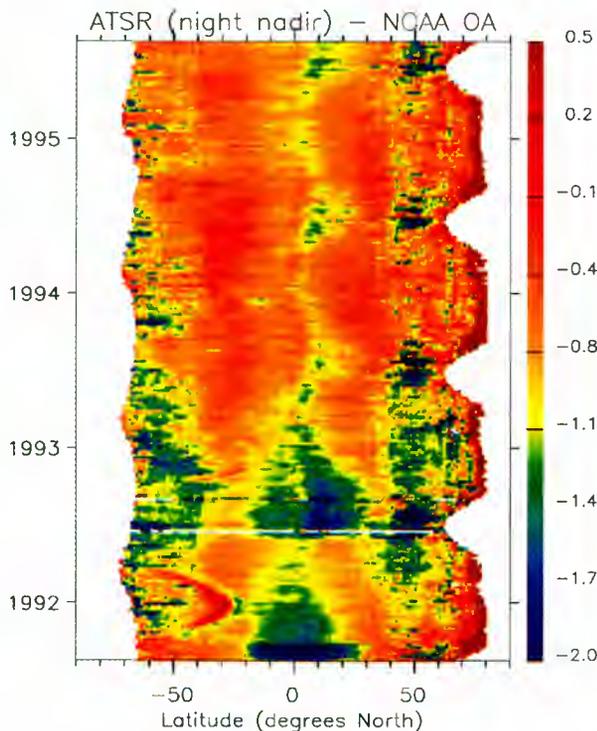


Plate 5. As Plate 3, but with ATSR single-view (nadir-only) nighttime SSTs.

bias suggests that it can be attributed primarily to inadequate compensation for atmospheric aerosols from Mount Pinatubo. In 1992 and the first half of 1993 the difference slowly reduces as the Pinatubo aerosols dissipate until it stabilizes at a mean (ATSR–NOAA OA) value of ~ -0.5 K until October 1994, when a further

drop of around 0.2 K is observed in the ATSR SSTs, as discussed below.

The same comparison for tropical regions (30°S to 30°N) is shown in Figure 2. Once again, the ATSR daytime SSTs become warmer with respect to the NOAA OA as the Pinatubo aerosols are reduced. As this is the region most affected by aerosols, the trend is stronger than in the wider latitude band considered in the previous figure. For the ATSR daytime SSTs the cold bias with respect to NOAA OA reduces from around -1.2 K in September 1991 to -0.4 K in October 1993. However, a relative downward trend in ATSR SSTs is apparent from the start of 1994. The exact origin of this trend remains to be firmly identified, but a major contribution may well be associated with the temperature of the ATSR detectors.

From early 1994 it became increasingly difficult for ATSR's onboard cooler to maintain the detector temperatures below 95 K. To reduce mechanical wear and maximize the life of the cooler, temperatures were allowed to rise gradually, reaching 110 K in early 1996. This warming affects the response of all the detectors, but the main difficulty is in the case of the $12\ \mu\text{m}$ channel, where the shift in response is sufficient to modify the ATSR long-wavelength filter cutoff. Generally, the result of this effective change in spectral response is to depress retrieved SSTs, particularly in humid tropical conditions. During the period from early 1994 to early 1996 the $12\ \mu\text{m}$ detector temperature rose from 97 to 110 K. This results in an expected reduction of about 0.2 K in the value of retrieved tropical SSTs, a value consistent with the cooling observed in Figure 2. The effect on midlatitude SSTs is <0.1 K, and the effect

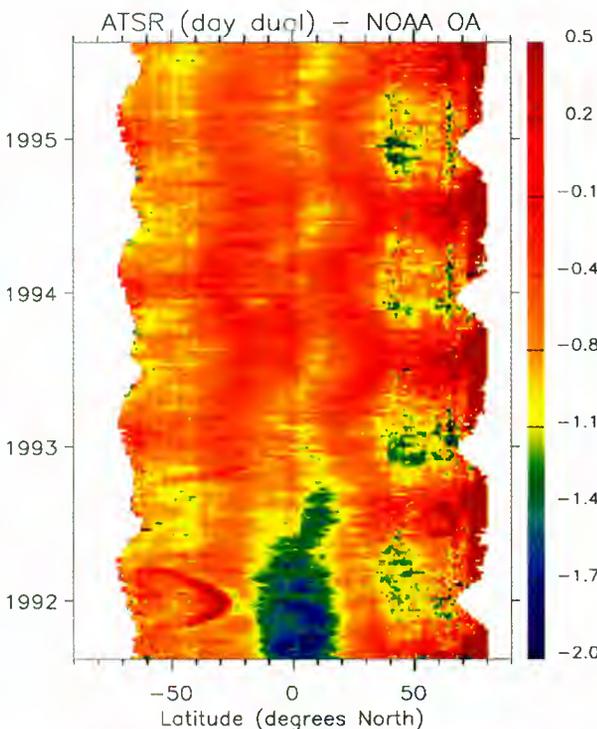


Plate 6. Zonal mean differences between ATSR dual-view daytime SSTs and NOAA OA.

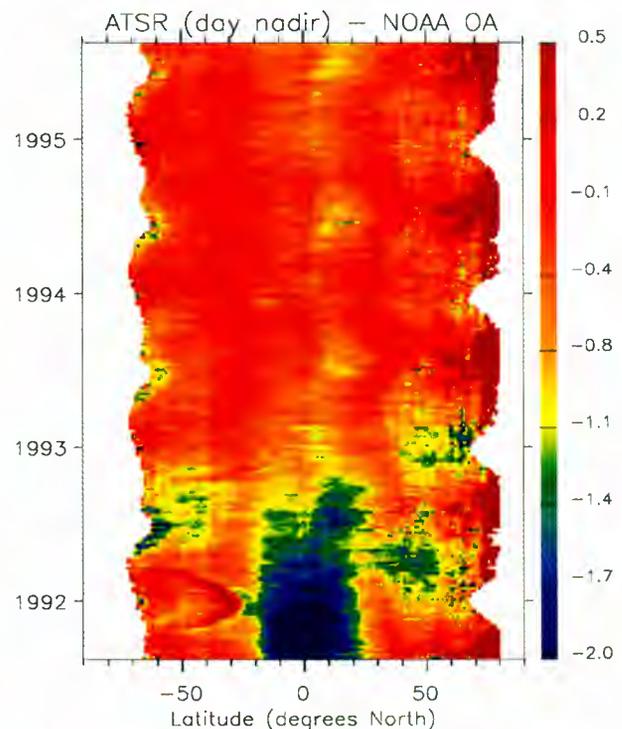


Plate 7. As Plate 3, but with ATSR single-view (nadir-only) daytime SSTs.

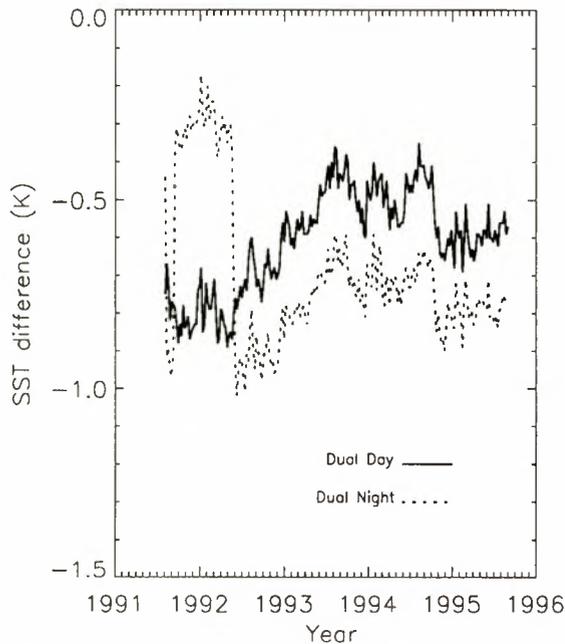


Figure 1. Area mean difference between ATSR dual-view SSTs and NOAA OA in the latitude range from 70°S to 70°N.

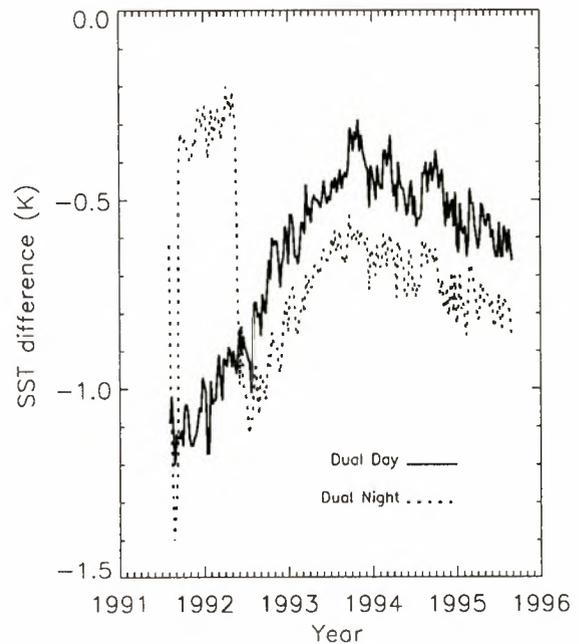


Figure 2. As Figure 1, using dual-view data in the latitude range from 30°S to 30°N.

for high-latitude SSTs is negligible (<0.01 K) as discussed in the appendix. We stress that as a result of the unique onboard calibration of ATSR and our detailed knowledge of the instrument's spectral response, we have the necessary information to model and correct for any long-term trend caused by changes in detector temperature, so this should not ultimately compromise the temporal integrity of the data set. The same trend is apparent in the difference time series for the ATSR nighttime SSTs, although the dominant signal here is again the sudden apparent cooling following $3.7 \mu\text{m}$ failure in May 1992.

The value of the dual-view retrieval is clear when we consider the corresponding time series using nadir-only ATSR SST retrievals. As shown in Figure 3, during the time of heaviest aerosol contamination in the tropics, nadir-only SSTs are cooler than the corresponding dual-view SSTs by up to 0.7 K.

The loss of the $3.7 \mu\text{m}$ data affected nighttime retrievals in two ways: a two-channel algorithm using only 11 and $12 \mu\text{m}$ data was necessarily adopted, and the cloud identification tests which used $3.7 \mu\text{m}$ data were lost. To estimate the relative importance of these two factors, we attempt to quantify the reduced effectiveness of the cloud decontamination by comparing all ATSR nighttime SSTs with those from relatively cloud-free regions.

Jones *et al.* [1996a, b] show that in specific geographical regions, residual cloud contamination exists in the nighttime ATSR SST data. This cloud contamination results in an increased nighttime variability and a cooler nighttime mean SST (relative to the daytime data). Both the seasonality and location of the contamination coincide with that of marine stratiform cloud [Klein and Hartmann, 1993]. A scheme to reduce the effect of such

contamination was proposed and tested by Jones *et al.* [1996a, b]. Briefly, an annual and semiannual cycle is fitted to the daytime ATSR SST data at each location; data are rejected if the residuals from this model exceed 3 times the standard deviation of the daytime residuals at that location. This results in the rejection of 4.7% of the SST data set, primarily nighttime data. An advantage of this scheme is that the rejection criteria change with region (depending on the natural ocean SST variability), and hence this should minimize the chance of filtering out real SST features. A further advantage is that the filtering scheme is not based on any external SST data sets. These may contain errors caused by either sparsity of sampling (in the case of in situ climatologies) or biases caused by different methods of SST retrieval (in the case of AVHRR-based climatologies).

Figure 4 shows a comparison of weekly mean dual-view nighttime SSTs in the latitude range from 70°S to 70°N; one average is computed using all SSTs, as in Figure 1, and the other uses only those SSTs which arose from "cloud-free" observations. (The latter are only available from January 1992 onward.) During the period when $3.7 \mu\text{m}$ data were available the weekly mean temperature based on all SSTs remains very close to that derived only from "clear" half-degree fields. After the $3.7 \mu\text{m}$ data were lost the weekly mean based on cloud-free cells is warmer than the mean of all the SSTs but by a factor of <0.1 K.

This suggests that although cloud contamination can be a serious problem in particular regions, it has only a small effect on the global mean scale. It seems likely that most of the drop in ATSR nighttime SSTs can be attributed to a bias between the three-channel and two-channel SST retrieval algorithms. This is encouraging, since eliminating this bias in the two-channel retrieval

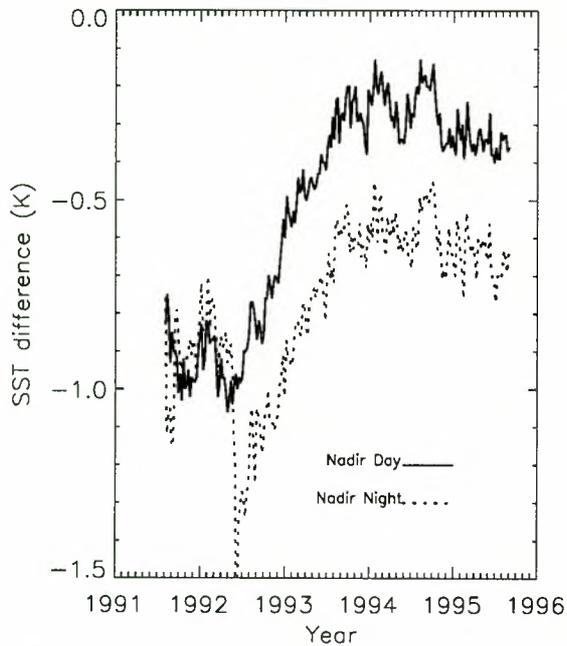


Figure 3. As Figure 1, using ATSR single-view (nadir-only) SSTs.

may simply require revised coefficients; this is clearly an urgent priority which will be addressed in the ATSR 1 reprocessing program.

Since these comparisons are dominated by the effects of aerosols in equatorial regions, we now consider the midlatitudes in both hemispheres. Figures 5 and 6 show the weekly mean ATSR–NOAA OA differences for the latitude bands 70° – 30° S and 30° – 70° N, respectively. As shown in Figure 5, ATSR SSTs average about 0.7 K cooler than NOAA OA values in midlatitudes in

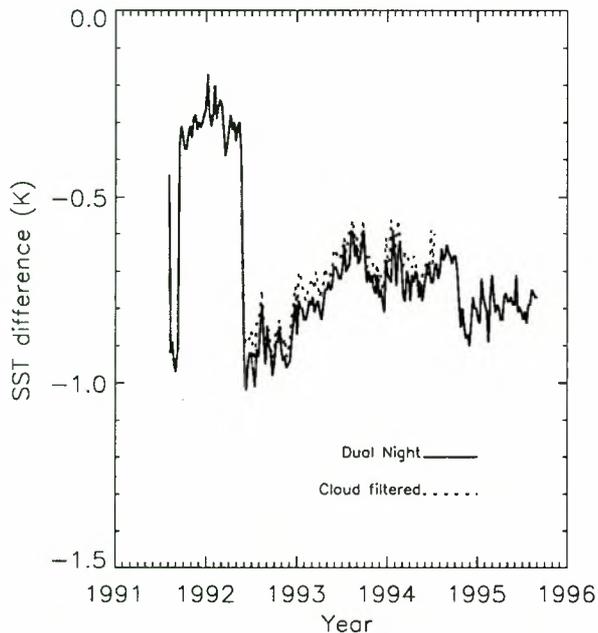


Figure 4. As Figure 1, comparing area mean difference averaged over all dual-view nighttime SSTs to the subset considered cloud-free by the Jones postprocessing scheme.

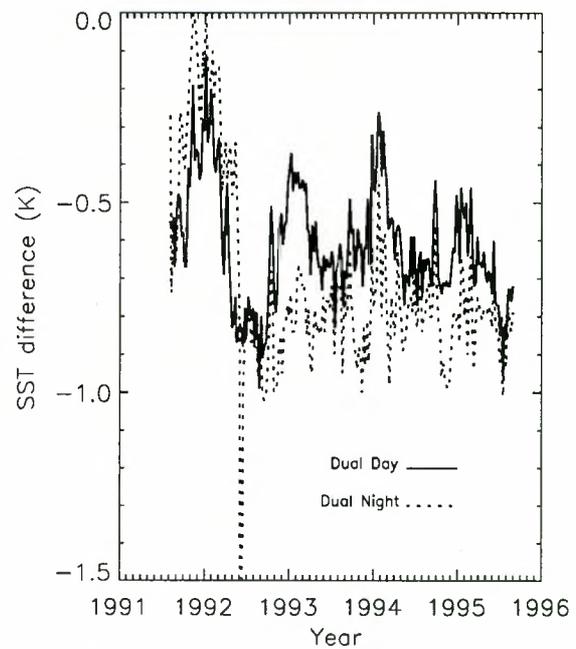


Figure 5. As Figure 1, using dual-view data in the latitude range from 70° to 30° S.

the southern hemisphere, with no clear aerosol-related trend. However, there is a weak seasonal cycle in the bias, particularly in the case of the ATSR daytime SSTs in the first 2 years of the mission, which appear warmest with respect to NOAA OA in the southern summer.

Figure 6 shows the corresponding plot for northern midlatitudes. A strong seasonal cycle is evident, with ATSR daytime SSTs peaking in the middle of northern summers at temperatures in excess of NOAA OA but showing a cold bias of about 1.0 K with respect to the latter in the winter. Nighttime SSTs evaluated when $3.7 \mu\text{m}$ data were available are ~ 0.4 K cooler than

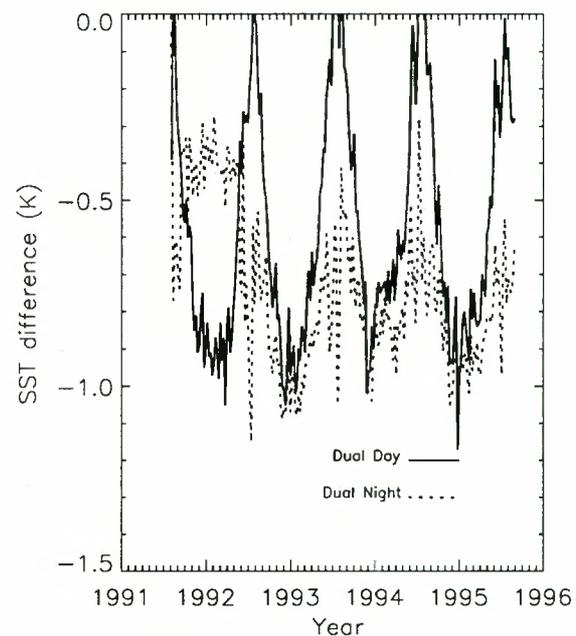


Figure 6. As Figure 1, using dual-view data in the latitude range from 30° to 70° N.

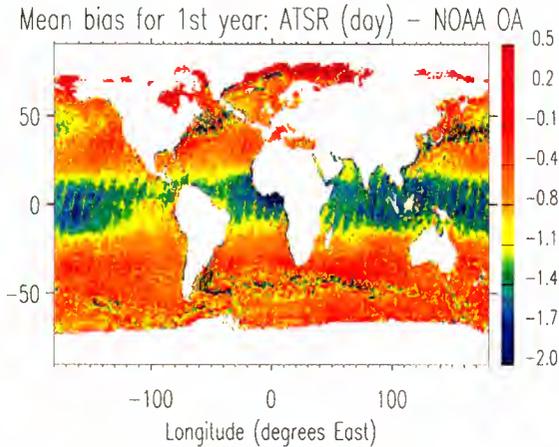


Plate 8. One year mean (August 1991 to July 1992) ATSR minus NOAA OA; ATSR daytime dual-view SSTs only.

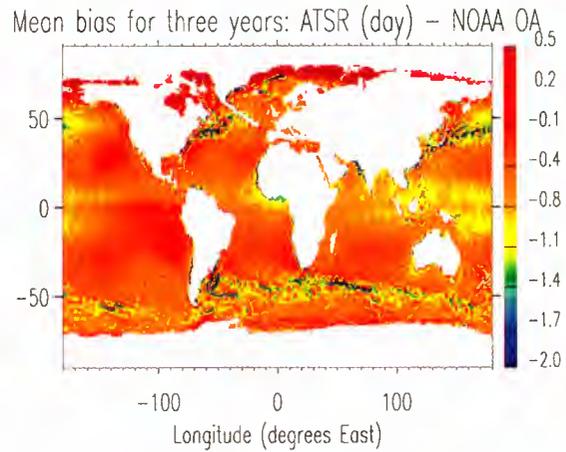


Plate 9. Three year mean (August 1992 to July 1995) ATSR minus NOAA OA; ATSR daytime dual-view SSTs only.

NOAA OA, with no sign of seasonal variation. After May 1992 these SSTs show a weak seasonal cycle, with a cold bias ranging from about 0.9 K in winter to 0.5 K in summer compared to NOAA OA. This points toward the need for seasonally dependent two-channel retrieval algorithms, discussed further in the conclusions.

4.3. Geographic Variation of ATSR–NOAA OA Bias

As a preliminary investigation of the geographic location of the differences in the ATSR and NOAA OA, we now consider global maps of time-averaged SST difference fields. Only ATSR dual-view data have been used. All temperatures differences are given in Kelvin and are shown in the range from -2.0 to 0.5 K.

Plate 8 shows the annual average of ATSR–NOAA OA using ATSR daytime dual-view SSTs for the first year of ATSR operation. The dominant feature of this map is the large negative bias exhibited by ATSR SSTs in the tropics; these values, which are depressed by more than 1 K with respect to NOAA OA, trace the extent of the Pinatubo plume during this period. The missing band of data from Greenland through the North Atlantic and into the Mediterranean arises because the ATSR data downlink occurs in this area, as mentioned above.

Plate 9 shows the same plot for the 3 year average beginning in August 1992, by which time atmospheric aerosol was close to the climatological level. In general, the ATSR SSTs show a cold bias of about 0.4 K with respect to NOAA OA. However, considerable geographic variation is evident. ATSR SSTs are around 1.0 K cooler than those from NOAA OA in the tropical west Pacific, possibly as a result of inadequate compensation for atmospheric water vapor. Note that retrieval coefficients are based on a prelaunch model calculation [Závodny *et al.*, 1995] in which the radiosonde data sets may be biased with respect to the real atmosphere and in which there is uncertainty in the magnitude of the water vapor continuum absorption. Saharan dust is

likely to be responsible for the cold bias in ATSR SSTs evident around Africa's Ivory Coast.

Extensive regions in the northwest Atlantic and northwest Pacific exhibit ATSR SSTs up to 2 K cooler than NOAA OA SSTs; this may be associated with residual cloud contamination of ATSR data or a systematic dependence of the ATSR retrieval on local meteorological conditions (given that these features correspond to the northern hemisphere storm tracks). It might also be evidence of tropospheric aerosol contamination from the industrialized regions of North America and east Asia. Alternatively, given the close correspondence of the pattern to climatological nonsolar surface heat flux, we may be seeing evidence of the skin-bulk effect. This clearly represents a promising avenue for further investigation.

ATSR SSTs are much warmer than the corresponding NOAA OA SSTs in the Arctic and Hudson Bay regions. This is likely to be due to inconsistencies in ice flagging between the two data sets.

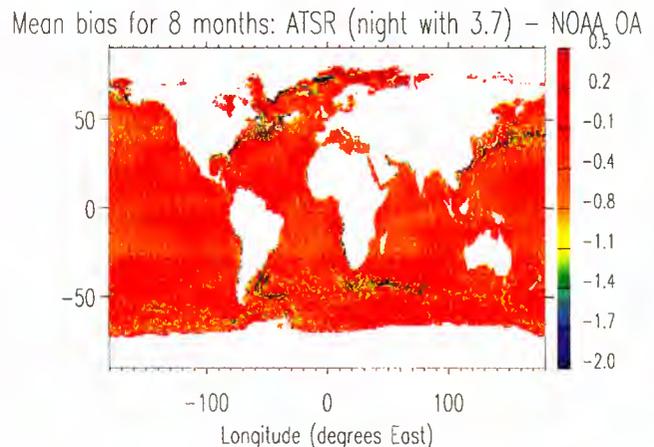


Plate 10. Nine-month mean (August 1991 to May 1992) ATSR minus NOAA OA; ATSR nighttime dual-view SSTs only.

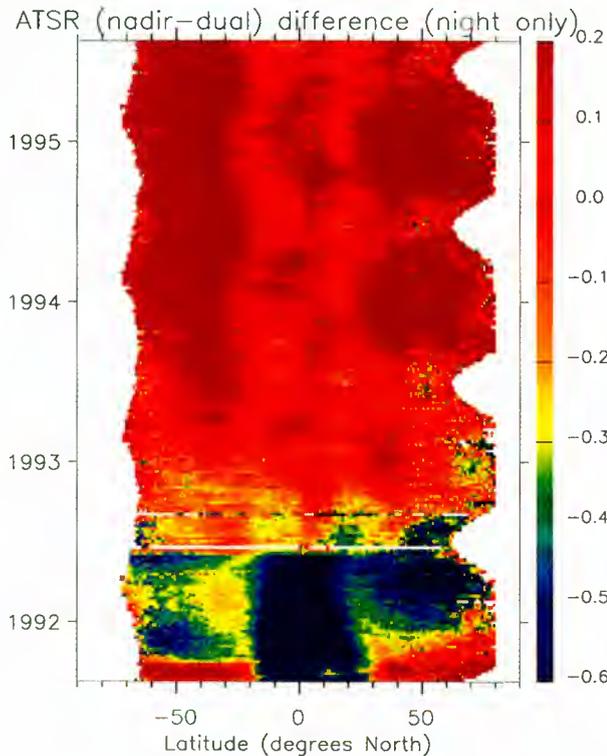


Plate 11. Time-latitude plot of zonal mean difference between ATSR dual-view and ATSR single-view (nadir-only) SSTs for nighttime data.

An intriguing aspect of the map is the undulations in ATSR-NOAA OA bias apparent off the U.S. eastern seaboard. Since this is a region of strong SST gradients, we may be observing an artefact of the optimal interpolation used in the NOAA OA, but the exact origins of this feature remain to be investigated. Similar structure is apparent in other regions where the SST is affected by western boundary currents, e.g., the Argentine and Japanese coasts.

Plate 10 shows the corresponding plot for those 9 months of ATSR nighttime SSTs for which the $3.7 \mu\text{m}$ data were available. This is considerably noisier than the daytime map (because of a much shorter integration time), and the ERS-1 3 day orbit repeat cycle (which dominated this period) is clearly visible. However, the pattern of bias owing to the Pinatubo plume which was so marked in Plate 8 is completely absent. Additionally, the zonal variation characteristic of water vapor contamination apparent in the daytime data (Plate 9) is missing from this plot. Some latitude dependence of the difference field is visible in the Tropics; the likely source of this weak pattern is the latitudinal variation of the retrieval coefficients. The success of the ATSR three-channel dual-view SST retrieval is confirmed by these comparisons.

4.4. Nadir-Dual Differences Within ATSR Data

In general, differences between the dual-view and nadir-only SST retrievals are attributable to absorp-

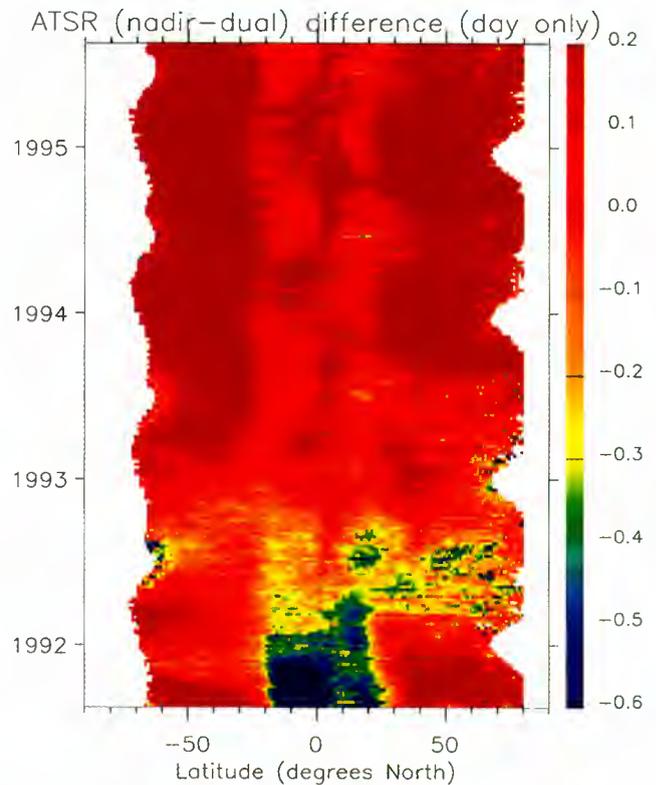


Plate 12. Time-latitude plot of zonal mean difference between ATSR dual-view and ATSR single-view (nadir-only) SSTs for daytime data.

tion by atmospheric constituents which were not included in the calculation of the retrieval coefficients. The nadir-dual difference can thus be used as a tracer for atmospheric aerosols (recalling that the ATSR retrieval coefficients are based on an assumption of climatological background aerosol loading) and to a lesser extent may be affected by changes in total column water vapor. Plate 11 shows the zonally averaged ATSR nadir-dual difference for nighttime data. Prior to the failure of the $3.7 \mu\text{m}$ channel and ignoring the first 3 weeks of operation when $3.7 \mu\text{m}$ data were not available, the effects of the Mount Pinatubo aerosol plume are clear, with dual-view SSTs more than 1 K warmer than the corresponding nadir-only values in the tropics. The heaviest aerosol load appears to be within 20° of the equator until November 1991. Until this time

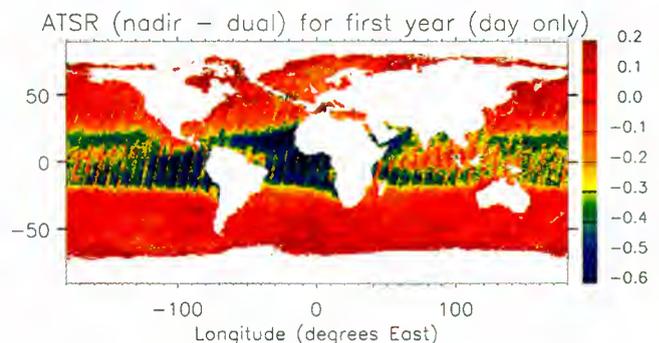


Plate 13. One year mean (August 1991 to July 1992) nadir minus dual difference for ATSR daytime SSTs.

the southern hemisphere appears more heavily contaminated than the northern hemisphere. However, in the first months of 1992 there appears to be stronger transport into the northern hemisphere, and by May 1992 the aerosol load appears to be distributed over all latitudes.

Plate 12 shows the corresponding plot for daytime data. The aerosol plume appears well-confined within 30° from the equator until March 1992, after which time there appears to be a rapid movement into the northern hemisphere. This is in broad agreement with the observations of *Baran and Foot* [1994].

Plate 13 shows the global pattern of the nadir–dual difference for daytime data during the first year of ATSR operation. The distribution of aerosol contamination apparent in this plot resembles that shown in the comparison between ATSR daytime SSTs and NOAA OA (Plate 8). This suggests that further use of nadir–dual differences may enable a better aerosol correction for the ATSR data set.

The close correspondence between the patterns of nadir–dual differences and those of ATSR–NOAA OA differences in the daytime and post $3.7\ \mu\text{m}$ failure nighttime two-channel retrievals (Plates 3 and 6) is encouraging since it implies that the information necessary to correct these two-channel retrievals is indeed present in the ATSR data.

5. Conclusions

The most important result to emerge from this comparison is the success of the ATSR nighttime dual-view SST retrievals using 3.7 , 11 , and $12\ \mu\text{m}$ data during a period of heavy atmospheric aerosol contamination. This result is particularly significant in that the coefficients in the retrieval algorithm were determined before launch and assume only climatological background levels of aerosol. The only concession to Pinatubo is a relaxation of the optimal thresholds in the various cloud decontamination schemes [*Závody et al.*, 1994, 1995]. The implication is that a three-channel, dual-view SST retrieval is genuinely insensitive to significant changes in atmospheric composition, as anticipated in the design of this mission. This contrasts with the need to develop a post hoc solution to the problem of aerosol contamination in the AVHRR SST retrieval and in ATSR two-channel retrievals.

During the time when all these data were available, ATSR nighttime SSTs were around $0.3\ \text{K}$ cooler than NOAA OA, an amount consistent with the skin-bulk temperature difference and the slightly cooler temperatures expected at nighttime relative to the NOAA OA diurnal average. The value of the dual-view retrieval is confirmed by the comparison with the nadir-only retrieval in which a cold bias of up to $2\ \text{K}$ can be attributed to the effect of aerosols and in which the location of the Pinatubo plume can be clearly seen.

As discussed above, the loss of the $3.7\ \mu\text{m}$ channel in May 1992 affected nighttime retrievals in two ways: a two-channel SST retrieval algorithm using only 11 and $12\ \mu\text{m}$ data was necessarily adopted, and the cloud iden-

tification tests which used $3.7\ \mu\text{m}$ data were lost. The two-channel algorithm appears to be subject to a systematic cold bias and is compromised by the presence of atmospheric aerosols, with cloud contamination being a problem in specific regions. An improved algorithm, which incorporates a correction for the Pinatubo aerosols and improved cloud tests, is under development.

ATSR daytime data are not considered to be seriously affected by cloud contamination. However, ATSR's daytime SSTs exhibit a larger cold bias with respect to NOAA OA SSTs than can be attributed to the skin-bulk effect, and the spatiotemporal signature of water vapor and aerosol contamination is clearly evident. This suggests that the coefficients used in the two-channel retrievals may not be optimal. As mentioned above, the initial 9 months of SSTs from the three-channel, nighttime observations show no evidence of atmospheric bias. These data may be used directly to provide a "ground truth" in the derivation of unbiased coefficients for a revised two-channel nighttime retrieval for use after the $3.7\ \mu\text{m}$ channel loss. In conjunction with a model of the diurnal cycle in skin SST we also intend to use the period when $3.7\ \mu\text{m}$ data were available to examine the origins (and scope for removal) of the observed bias in the two-channel daytime retrieval algorithm.

In the longer term several results presented here, particularly the seasonal cycle in the bias in the two-channel retrieval algorithms, point to the need for a more physically based approach to SST retrieval to exploit the full potential of the ATSR data. Information about total column and vertical profiles of atmospheric moisture and aerosol, while apparently unnecessary for accurate three-channel, dual-view SST retrieval, may significantly improve two-channel SSTs.

ATSR 2, the second instrument in the ATSR series, was launched on board the European Space Agency's (ESA) ERS-2 in April 1995. Global SSTs from this instrument will be available in early 1997, and the improved processing scheme developed for ATSR 2 data will be used to re-process ATSR-1 observations, with the resulting SSTs given at $10\ \text{arc min.}$ resolution. Both ATSR 1 and ATSR 2 observations from the period May–December 1995 will be available shortly, and the comparison of these data sets will help to establish the accuracy of both instruments.

Appendix: The Effect on ATSR SST Retrieval of an Increase in $12\ \mu\text{m}$ Detector Temperature

The ATSR detector temperatures were subject to regular assessment after launch, with the cooler's operation adjusted to keep operating temperatures below $\sim 95\ \text{K}$ in order to maintain the $12\ \mu\text{m}$ spectral response [*J. Delderfield*, personal communication, 1997]. As shown in Figure A1, this was no longer possible after early 1994, and the ATSR $12\ \mu\text{m}$ detector temperature increased gradually, reaching $110\ \text{K}$ in 1996. There are

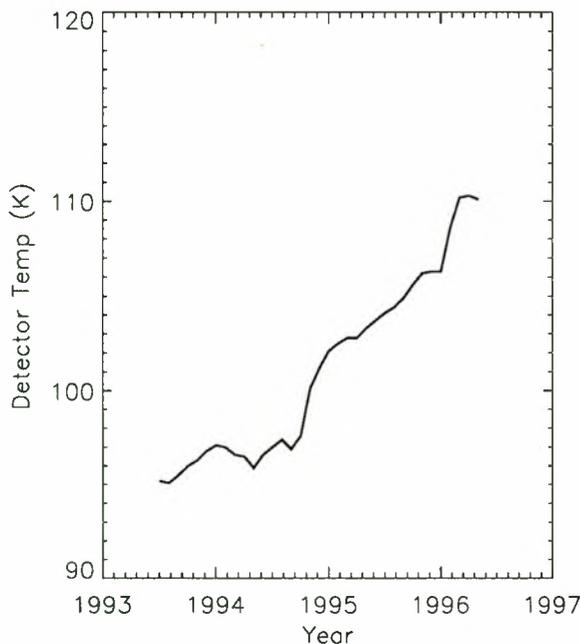


Figure A1. ATSR 12 μm detector temperature for 1993–1996.

significant orbital and annual variations superimposed on the general temperature drift. A similar increase occurred for the other detectors, but the effect is only important for the 12 μm channel as the shift in its detector spectral response characteristics results in attenuation of response within the long wavelength end of its filter's spectral band.

The spectral characteristics of the 12 μm detector at a range of operating temperatures have been tabulated by Mason [1991] and J. Delderfield (personal communication, 1997). These data have been used to investigate the effect of detector temperature on retrieved SST for a range of atmospheric conditions. The atmospheric profiles used in the simulations were the U.S. standard tropical, midlatitude (summer), and high-latitude

(summer) profiles (i.e., far from extreme conditions). In each case, values were computed for five air–sea temperature differences ($-4, -2, 0, 2,$ and 4 K).

The differences given in Table A1 are with respect to the simulated results for a 12 μm detector temperature ($T_{12\ \mu\text{m}}$) of 82 K (positive values indicate that the retrieved SST is cooler than that obtained for the case of the 82 K detector temperature). Humid tropical atmospheres are the most sensitive to the change in spectral response, with errors in excess of 0.2 K possible at the top of the range of observed detector temperature. Single-view and dual-view retrievals have approximately the same sensitivity to this effect.

In the planned reprocessing of ATSR 1 data, this will be addressed by interpolating between sets of SST retrieval coefficients computed for a range of detector temperatures. Thus the value of ATSR SSTs in investigating climate trends remains uncompromised by this effect.

Acknowledgments. The along-track scanning radiometer was provided to the European Space Agency (ESA) by a consortium consisting of Rutherford Appleton Laboratory (RAL), Mullard Space Science Laboratory, the University of Oxford, and the United Kingdom Meteorological Office, with contributions from Australia and France. ATSR data were provided courtesy of ESA and were processed at RAL. We are grateful to NOAA for the provision of MCSST data and the NOAA SST analysis. Funding from the Natural Environment Research Council (NERC) via the ATSR post-launch support and Ocean Dynamics funding lines is gratefully acknowledged. We would also like to thank Ian Barton, John Delderfield, Tim Nightingale, Richard Reynolds, Phil Watts, John Wright, and anonymous reviewers for helpful comments.

References

- Allen, M.R., C.T. Mutlow, G.M.C Blumberg, J.R. Christy, R.T. McNider, and D.T. Llewellyn-Jones, Global change detection, *Nature*, 370, 24–25, 1994.

Table A1. SST Retrieval Errors Caused by Warming of 12 μm Detector

$T_{12\ \mu\text{m}}, \text{K}$	Latitude	Single-View Retrieval					Dual-View Retrieval				
		-4	-2	0	2	4	-4	-2	0	2	4
90	tropical	0.01	0.01	0.02	0.02	0.03	-0.01	0.00	0.01	0.01	0.02
	midlatitude	0.00	0.00	0.01	0.01	0.01	-0.02	-0.01	-0.01	0.00	0.00
	high latitude	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
100	tropical	0.07	0.09	0.10	0.11	0.13	0.04	0.06	0.08	0.10	0.12
	midlatitude	0.04	0.05	0.06	0.07	0.09	0.00	0.01	0.03	0.04	0.05
	high latitude	-0.01	0.00	0.00	0.01	0.01	-0.01	0.00	0.00	0.00	0.00
110	tropical	0.17	0.19	0.22	0.24	0.26	0.11	0.15	0.19	0.22	0.26
	midlatitude	0.09	0.12	0.14	0.16	0.19	0.03	0.05	0.07	0.09	0.11
	high latitude	0.00	0.01	0.02	0.03	0.04	0.00	0.00	0.01	0.01	0.01

Differences are given with respect to the simulated results for $T_{12\ \mu\text{m}}=82$ K (positive values indicate that the retrieved SST is cooler than that obtained for the case of the 82 K detector temperature). Atmospheric profiles used in the simulations were the U.S. standard tropical, midlatitude (summer), and high-latitude (summer) profiles. In each case, values were computed for five air–sea temperature differences as shown below ($\Delta T_{\text{air-sea}} = -4, -2, 0, 2, 4$ K).

- Bailey, P., SADIST-2 v100 Products, *Tech. Rep. ER-TN-RAL-AT-2164*, 123pp., Space Sci. Dept., Rutherford Appleton Lab., 1995.
- Baran, A.J., and J.S. Foot, New application of the operational sounder HIRS in determining a climatology of sulphuric acid aerosol from the Pinatubo eruption, *J. Geophys. Res.*, **99**, 25,673–25,679, 1994.
- Barton, I.J., A.J. Prata, and R.P. Cecket, Validation of the ATSR in Australian waters, *J. Atmos. Oceanic Technol.*, **12**, 290–300, 1995.
- Bottomley, M., C.K. Folland, J. Hsiung, R.E. Newell, and D.E. Parker, *Global Ocean Surface Temperature Atlas: Joint Project, United Kingdom Meteorological Office and Massachusetts Institute of Technology*, 20pp., Her Majesty's Stn. Off., Norwich, England, 1990.
- Bradshaw, T.W., J. Delderfield, S.T. Werrett, and G. Davey, Performance of the Oxford miniature Stirling cycle cooler, *Adv. Cryog. Eng.*, **31**, 801–809, 1985.
- Delderfield, J., D.T. Llewellyn-Jones, R. Bernard, Y. de Javel, E.J. Williamson, I. Mason, D.R. Pick, and I.J. Barton, The along-track scanning radiometer (ATSR) for ERS-1, in *Instrumentation for Optical Remote Sensing From Space, Proc. SPIE Int. Soc. Opt. Eng.*, **589**, 114–120, 1986.
- Donlon, C.J., and I.S. Robinson, Observations of the oceanic thermal skin in the Atlantic Ocean, *J. Geophys. Res.*, **102**, 18,585–18,606, 1997.
- Edwards, T., et al., The along-track scanning radiometer – Measurement of sea surface temperature from ERS-1, *J. Br. Interplanet. Soc.*, **43**, 160–180, 1990.
- Forrester, T.N., and P.G. Challenor, Validation of ATSR sea surface temperatures in the Faeroes region, *Int. J. Remote Sens.*, **16**, 2741–2753, 1995.
- Harries, J.E., D.T. Llewellyn-Jones, P.J. Minnett, R.W. Saunders, and A.M. Závody, Observations of sea surface temperature for climate research, *Philos. Trans. R. Soc. London A*, **309**, 381–395, 1983.
- Harrison, D.L., and C.P. Jones, A user appraisal of ATSR near-real-time products, in *Proceedings of the First ERS-1 Symposium, ESA Spec. Pub.*, **SP-359**, 791–795, 1994.
- Jones, M.S., M.A. Saunders, and T.H. Guymer, Reducing cloud contamination in ATSR averaged sea surface temperature data, *J. Atmos. Oceanic Technol.*, **13**, 492–506, 1996a.
- Jones, M.S., M.A. Saunders, and T.H. Guymer, Global remnant cloud contamination in the along-track scanning radiometer data: Source and removal, *J. Geophys. Res.*, **101**, 12,141–12,147, 1996b.
- Klein, S.A., and D.L. Hartmann, The seasonal cycle of low stratiform clouds, *J. Clim.*, **6**, 1587–1606, 1993.
- Mason, G., Test and calibration of the along-track scanning radiometer, a satellite borne infrared radiometer designed to measure sea surface temperature, D. Phil Thesis, 248pp., University of Oxford, U.K., September 1991.
- McClain, E.P., W.G. Pichel, and C.C. Walton, Comparative performance of AVHRR based multichannel SSTs, *J. Geophys. Res.*, **90**, 11,587–11,601, 1985.
- Minnett, P.J., Consequences of sea surface temperature variability on the validation and applications of satellite measurements, *J. Geophys. Res.*, **96**, 18,475–18,489, 1991.
- Murray, M.J., *Sea surface Temperatures from ATSR (August 1991 to July 1995)*, [CD-ROM], Rutherford Appleton Lab., Chilton, England, 1995.
- Mutlow, C.T., A.M. Závody, I.J. Barton, and D.T. Llewellyn-Jones, Sea surface temperature measurements by the along-track scanning radiometer on the ERS-1 satellite: Early results, *J. Geophys. Res.*, **99**, 575–588, 1994.
- Reynolds, R.W., A real-time global sea surface temperature analysis, *J. Clim.*, **1**, 75–86, 1988.
- Reynolds, R.W., Impact of Mount Pinatubo aerosols on satellite-derived sea surface temperatures, *J. Clim.*, **6**, 768–774, 1993.
- Reynolds, R.W., and D.C. Marisco, An improved real-time global sea surface temperature analysis, *J. Clim.*, **6**, 114–119, 1993.
- Reynolds, R.W., and T.M. Smith, Improved global sea surface temperature analyses using optimum interpolation, *J. Clim.*, **7**, 929–948, 1994.
- Schuessel, P., W.J. Mery, H. Grassi, and T. Mammen, On the bulk-skin temperature difference and its impact on satellite remote sensing of sea surface temperature, *J. Geophys. Res.*, **95**, 13,341–13,356, 1990.
- Werrett, S.T., G.D. Peskett, G. Davey, T.W. Bradshaw, and J. Delderfield, Development of a small Stirling cycle cooler for spaceflight applications, *Adv. Cryog. Eng.*, **31**, 791–799, 1985.
- Závody, A.M., M.R. Gorman, D.J. Lee, D. Eccles, C.T. Mutlow, and D.T. Llewellyn-Jones, The ATSR data processing scheme developed for the EODC, *Int. J. Remote Sens.*, **15**, 827–843, 1994.
- Závody, A.M., C.T. Mutlow, and D.T. Llewellyn-Jones, A radiative transfer model for SST retrieval for the ATSR, *J. Geophys. Res.*, **100**, 937–952, 1995.

M.R. Allen, M.J. Murray and C.T. Mutlow, and A.M. Závody, Space Science Dept., Rutherford Appleton Laboratory, Chilton, Didcot, OX11 0QX, U.K. (e-mail: M.R.Allen@rl.ac.uk; Jo.Murray@rl.ac.uk; C.T.Mutlow@rl.ac.uk; A.M.Zavody@rl.ac.uk)

T.N. Forrester and M.S. Jones, James Rennell Division, Southampton Oceanography Centre, European Way, Southampton, SO14 3ZH, U.K. (e-mail: T.Forrester@soc.soton.ac.uk; Matthew.Jones@soc.soton.ac.uk)

(Received December 30, 1996; revised July 11, 1997; accepted July 30, 1997.)

Regime shifts in mesoscale deformations of sea ice during the winter

Reinert Korsnes

Norsk Polarinstitutt, Tromsø, Norway

Abstract. Large temporary rigid areas in the Arctic sea ice cover form and disintegrate in a few days during winter. When these rigid areas are shore-fast, they eliminate moment flux from air to water via wind forcing on the ice. Temporary shore-fast ice causes mixing of upper stratified waters below the ice in areas of strong tidal currents. Satellite data and observations by explorers in the Arctic show that large rigid areas tend to break up during the summer. We can conclude that formation and breakup of rigid areas are sensitive to temperature. The present study illustrates techniques for satellite observation of the dynamics of the Arctic sea ice cover during winter. Better understanding of the physical mechanisms involved will enhance codes for such satellite observations. The study indicates anomalous/“critical” dynamics and scale invariance in the formation and breakup of rigid areas. These indications come from direct satellite observations and are supported by generic conceptual models for ice dynamics during winter. Critical phenomena often form fractal structures with no natural lengths.

1. Significance of Deformations in the Ice Field

This paper reports indications of temporary regime shifts in the polygonal-block structure of polar sea ice during winter. One of these regimes is when the strength of the ice (single ice floes) is significant for ice drift; the other regime is when the strength of ice floes is not significant for ice drift. These situations occur when there is free ice drift or when interactions between ice floes are most significant for ice drift. This other regime is typical for the summer season.

The present indications of regime shifts come from satellite-based observations of deformations in the ice field east of Spitsbergen, Svalbard, between Nordaustlandet and Kong Karls Land during the period January 2 to March 31, 1994. The swath of an ERS-1 synthetic aperture radar (SAR) scene ($100 \times 100 \text{ km}^2$) limits the study area.

The present study also gives theoretical arguments for possible shifts in the ice drift regime. We can imagine one class of situations (regime) when the ice field consists of ice floes (temporary rigid areas) of maximum size far below the distance to the coast. We may then suppose that there is a homogenization scale over which we describe the ice field as a continuum. This is a key assumption behind many models of sea ice dynamics such as models by *Hibler* [1979] and others. The second regime is where the maximum sizes of the rigid areas are of the same order as the distance to the coast. These large rigid areas may last for just a few days, and they can be directly fixed to the coast or partly fixed via surrounding ice.

The part of the ice extent conforming to this second ice drift regime has large spatial and temporal variations. A lower limit for this part of the ice extent is clearly the areas of fast ice during the whole winter (such as in several fjords of Svalbard). However, there may be no clear upper spatial limit. *Thorndike and Colony* [1982] mention episodes from the Arctic Ice Dynamics Joint Experiment (AIDJEX) camps and examples of

data from the International Arctic Buoy Program (IABP) where the ice is motionless far from coastlines in the Arctic Ocean during the winter [see also *Thorndike and Cheung*, 1977; *Thorndike*, 1986; *Korsnes*, 1994]. However, these extreme events are rare and perhaps not significant for large-scale drift patterns over long periods of time.

Observation of formation and breakup of rigid areas in the ice field gives indications of temperature fluxes from and external forcing on the ice field. Hence more insight into such processes allows the possibility for better codes for radar satellite observations in the ice field.

Shifts in ice drift regimes are significant for currents below the ice, for ice drift, and for ice production in an area. When ice appears on the open ocean, it normally transfers a flux of momentum from the air to the water. It amplifies the wind forcing on the upper ocean since drift ice has larger wind friction compared to open water. However, when the ice field becomes rigid on large scales and constitutes temporary shore-fast ice, it eliminates wind forcing on the upper ocean. This situation also has a direct effect on the steering of currents in shallow waters and mixing in stratified upper water layers due to long and deep ice ridges.

This study indicates that formation and breakup of rigid ice areas can be critical to whether or not external forcing expires some critical values. Rigid areas form and disintegrate over much different timescales. This study focuses on a timescale of 3 days. Russian observations [*Gorbunov et al.*, 1983], for example, show large interannual variations of the extent of shore-fast ice in the Kara Sea.

2. Analyses of ERS-1 SAR Images

This work shows an analysis of a time series of ERS-1.SAR.PRI images from the European Space Agency (ESA) data archive covering the same area at the Kong Karls Land, Svalbard, every 3 days at 1051 UTC January 2 to March 31, 1994 (the data from January 23 were missing). Figure 1 illustrates the actual spatial coverage.

Accurate relative positioning is crucial for the estimation of

Copyright 1998 by the American Geophysical Union.

Paper number 97JC03332.
0148-0227/98/97JC-03332\$09.00

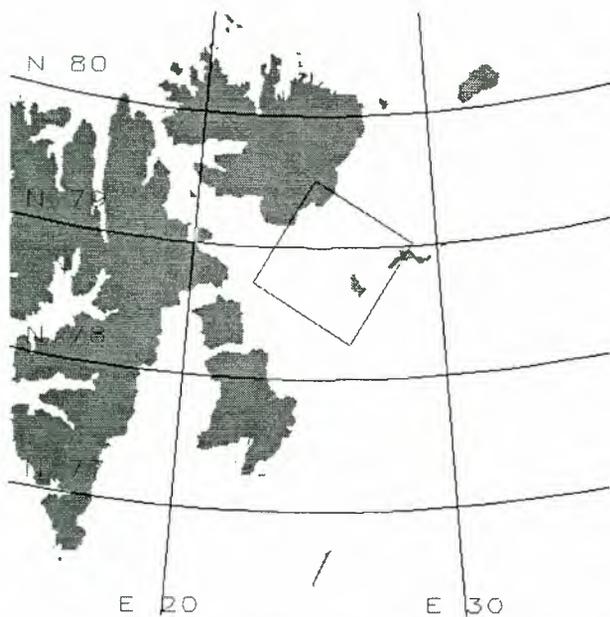


Figure 1. Data coverage at Kong Karls Land, east of Spitsbergen, Svalbard, for time series of ERS-1 synthetic aperture radar (SAR) images every 3 days at 1051 UTC January–March 1994 (given by the $100 \times 100 \text{ km}^2$ frame).

small deformations in the ice field based on ice tracking from satellite imagery. This analysis therefore comprises an estimation of the radar return from circular neighborhoods of diameter 112.5 m around points 12.5 m apart in a regular mesh corresponding to the pixels of the ERS-1.SAR.PRI product. Figure 2 illustrates the picture elements approximating these circular neighborhoods.

The averaging procedure included in (1) (compare $\langle I \rangle = (1/N) \sum_i DN_i^2$ below) reduces “noise” in the image data (i.e., it makes different coverages of the same ice more consistent and the application of ice tracking based on cross correlation better). The circular shape of the averaging window (as given by Figure 2) ensures independence of rotation of the ice (or rotation of image coverage) between different image times.

The procedure given by *Laur* [1992] provided the actual calibration and estimate of the backscattering coefficient σ^0 for a local incidence angle α

$$\sigma^0 = \frac{\langle I \rangle \sin \alpha}{K \sin \alpha_{\text{ref}}} \quad (1)$$

where $\langle I \rangle = (1/N) \sum_i DN_i^2$ is the square average of the instrument counts DN_i (pixel values) given for each picture element (pixel) in the ERS-1.SAR.PRI data product covering the actual distributed target (from Figure 2, $N = 69$). The reference incidence angle (midrange angle) for which the value of K is given in the ESA ERS-1.SAR.PRI data product is α_{ref} . Each pair of images 3 days apart, which are then based on these data products, were subject for ice tracking starting with a regular grid of points 500 m apart (giving a maximum of 40,000 ice displacement vectors). Identification of rigid areas (during 3 days) within the set of displacement vectors were made by sorting out subsets of vectors fitting the formula

$$z \rightarrow (z - z_0)e^{i\phi} + z_0 \quad (2)$$

where the complex number z gives the start position of a point and z_0 is a center for the rotation ϕ . Techniques and software

from the work by *Korsnes* [1994] provided this computing task. Let z_n , ω_n , $n = 1, 2, \dots, N$, respectively, be the start and end points of displacements from the ice-tracking algorithm. The actual criterion for the selection of rigid movement (shift/rotation) is

$$|\omega_n - [(z_n - z_0)e^{i\phi} + z_0]| \leq \epsilon \quad (3)$$

where $\epsilon = 25 \text{ m}$ in this case (close to the image data resolution). Plates 1, 2, and 3 illustrate examples of the result. The solid black line on these images represents the coastline. Plates 1, 2, and 3 show a distributed sample of 3-day displacement vectors for the sea ice. The smallest yellow dots give the start positions of the displacement vectors from the identification algorithm that are not classified as part of any of the 10 largest rigid areas. The similar but larger and overlapping dots (making continuous areas) with different colors give the start positions for displacement vectors for ice that are part of the 10 largest rigid areas (one color for each rigid area). Land and shore-fast ice are in red. Plate 1 shows data from a period with much deformation in the ice field. Plate 2 illustrates the formation of large rigid areas and shore-fast ice.

It is tempting to explain the formation and breakup of fast ice and large rigid areas by variations in currents below the ice and wind forcing. Figure 3 gives estimates of tidal currents (based on *Kowalik and Proshutinsky* [1993, 1994]) and hindcast (geostrophic) wind data from the Norwegian Meteorological Institute [*Reistad and Iden*, 1995] at the position $78^\circ 59' \text{N}$, $26^\circ 23' \text{E}$. Figure 3 also shows data from temperature measurements on the Hopen Island, and it gives estimates of the size of the largest rigid area in the given study area during 3 days (between the image times).

The solid black arrow on the Plates 1, 2, and 3 gives the hypothetical free ice drift (displacement) between the image times based on these estimates of the geostrophic wind speed vector \mathbf{G} . This hypothetical free ice drift (displacement) vector has length $0.0077 \cdot |\mathbf{G}|$ (where $|\mathbf{G}|$ is the length of the vector \mathbf{G}) and a direction 5° clockwise from \mathbf{G} . This is, according to

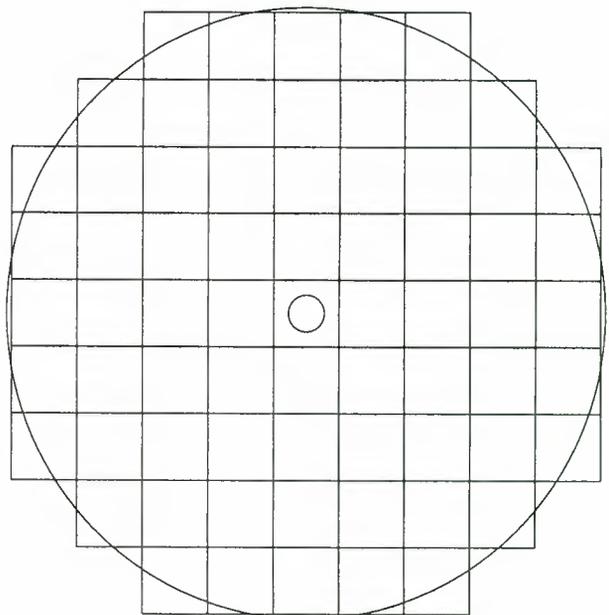


Figure 2. Approximating a circle of diameter 112.5 m by (square) pixels in an ERS-1.SAR.PRI image.

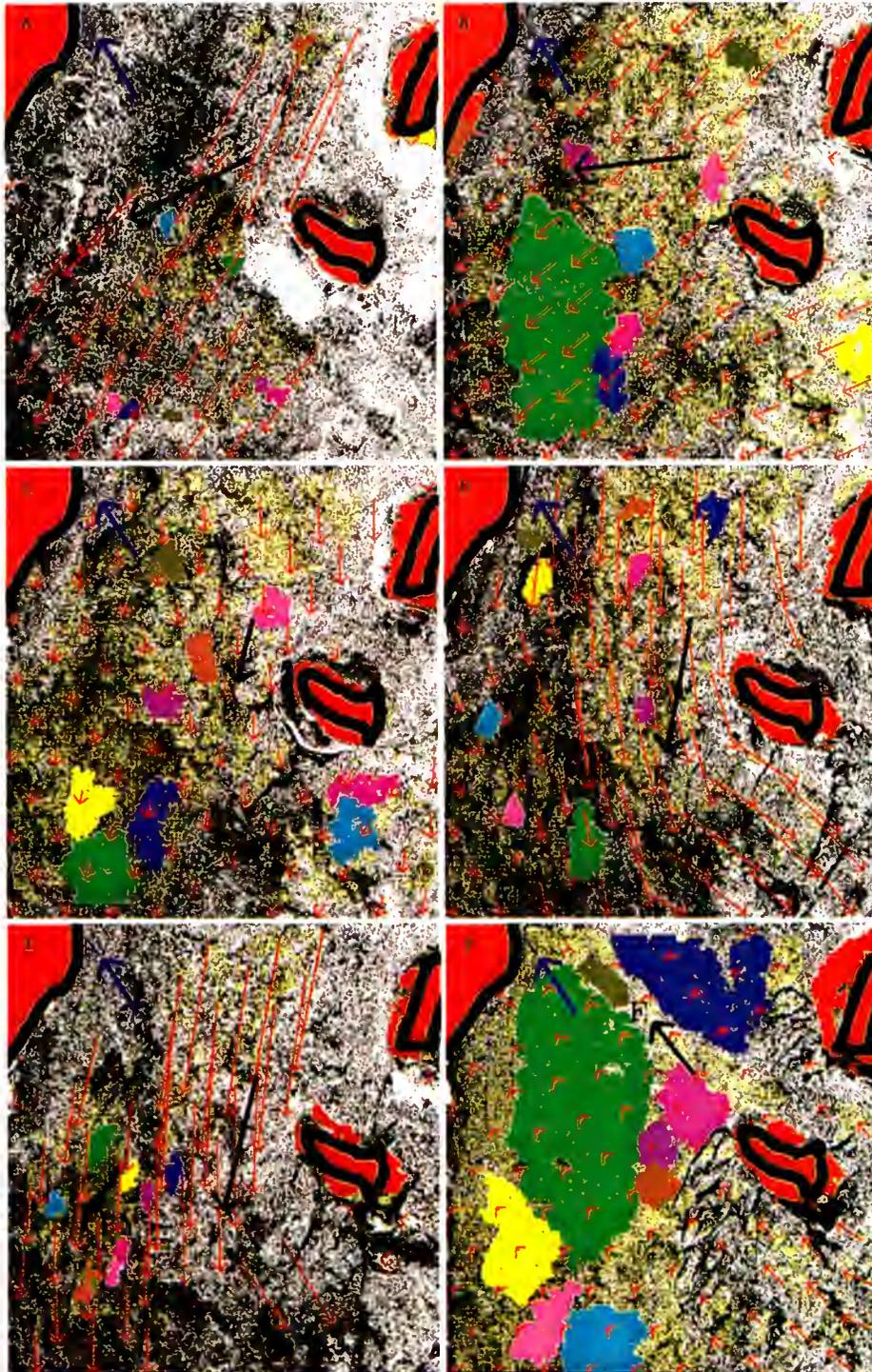


Plate 1. Identification of displacement (red vectors) and rigid areas in the ice field from between sequential ERS-1 SAR images covering the same area ($100 \times 100 \text{ km}^2$) 3 days apart starting January 2, 1994. The A–F labels in the upper left corners give the sequence. Blue arrow labeled “N” indicates north. Black arrow labeled “F” gives hypothetical 3-day free ice drift (displacement). Original Data ©ESA (1995), Distributed by Eurimage.

Thorndike and Colony [1982], a typical relation for winter situations in the central Arctic away from coastlines.

Note that the observed 3-day ice displacement often has direction up to 30° to the left (anticlockwise) of the hypothetical free ice drift. A possible explanation is tidal currents (such as associated with M2 shifting phase relative to the solar time over 3 days). Another perception is anisotropic forces through

the ice because of land. A third possibility, which is easy to illustrate, is resistance/friction due to motion and the fact that the ice moves in steps between the image times (due to formation of temporary large rigid areas). This will give an anticlockwise shift in drift direction compared to free drift. The conceptual argument below clarifies this. Assume the following relation in a normal (right angle) coordinate system describes

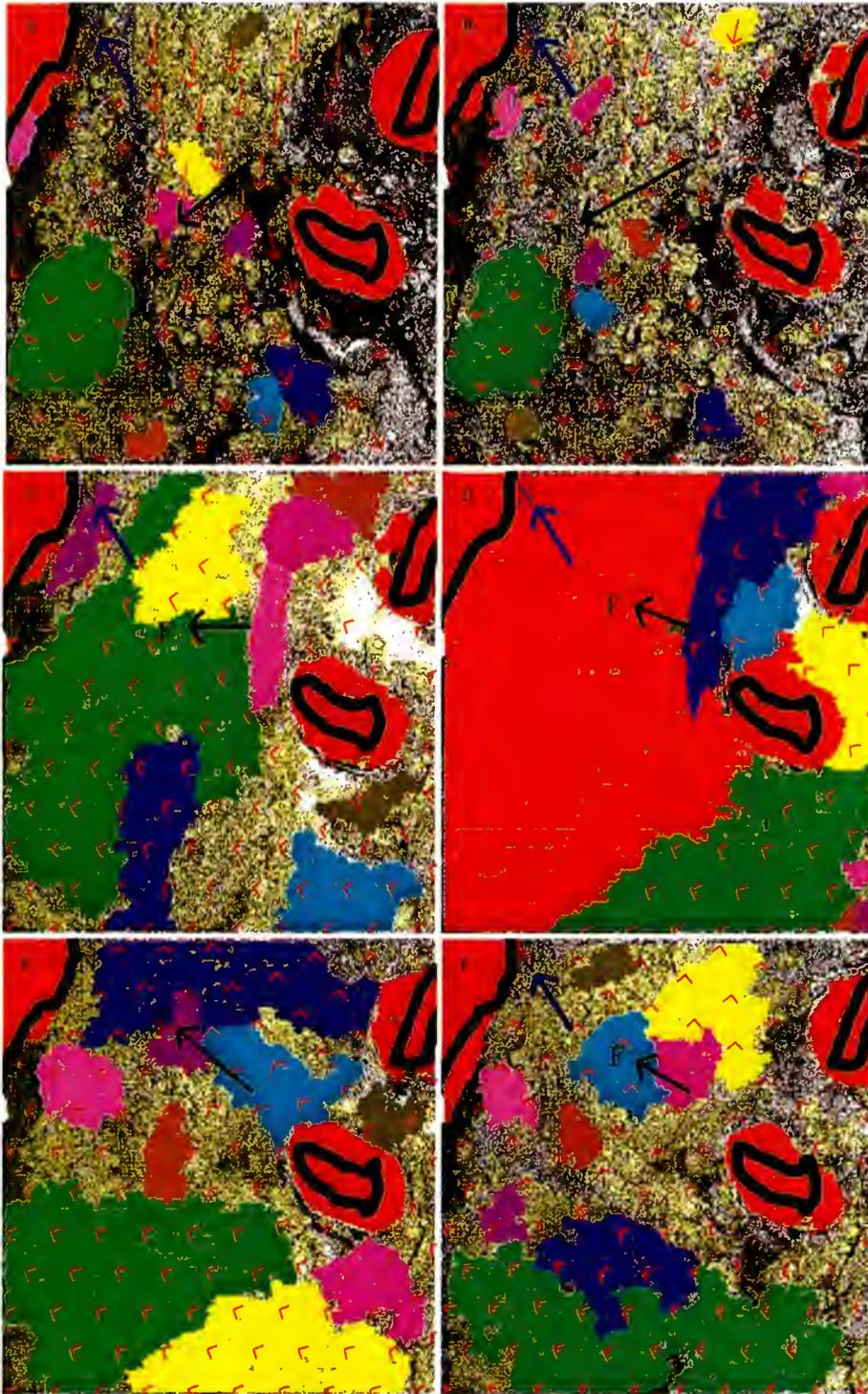


Plate 2. Identification of displacement (red vectors) and rigid areas in the ice field from between sequential ERS-1 SAR images covering the same area ($100 \times 100 \text{ km}^2$) 3 days apart starting March 12, 1994. The A-F labels in the upper left corners give the sequence. Blue arrow labeled “N” indicates north. Black arrow labeled “F” gives hypothetical 3-day free ice drift (displacement). Original Data ©ESA (1995), Distributed by Eurimage.

the velocity $\mathbf{v}(t)$ at the time t of an (ice) particle at rest in the origin at time $t = 0$

$$\mathbf{F} - \mathbf{f} \times \mathbf{v} - a\mathbf{v} = \frac{d\mathbf{v}}{dt} \quad (4)$$

\mathbf{F} is the (wind) force per unit mass on the particle (assumed constant parallel to the y axis), and \mathbf{f} is the local Coriolis

parameter; a is a “linear friction coefficient.” Figure 4 gives, for various values of the friction a , the particle track for $t = 0 \cdots T$, where $T = 3$ days, which is the same as the time step between the subsequential satellite images. $F \stackrel{\text{def}}{=} |\mathbf{F}| = 10^{-5} \text{ m s}^{-2}$ in this case. Note that the solution \mathbf{v} of (4) scales (linear) with the driving force \mathbf{F} , and hence the length of \mathbf{F} does not affect the form (directions) of the trajectory after a given time t .

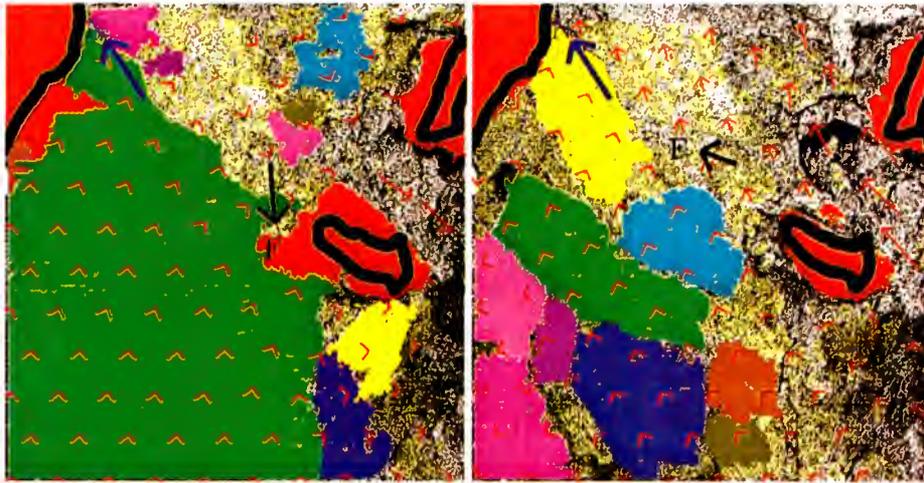


Plate 3. Identification of displacement (red vectors) and rigid areas in the ice field from between sequential ERS-1 SAR images covering the same area ($100 \times 100 \text{ km}^2$) 3 days apart starting on (left) January 26 and (right) January 29, 1994. Blue arrow labeled “N” indicates north. Black arrow labeled “F” gives hypothetical 3-day free ice drift (displacement). Original Data ©ESA (1995), Distributed by Eurimage.

The direction of the average velocity after a given time is given by the ratio between the friction coefficient a and the Coriolis parameter f (hence the value of a is given in units of the length f of f in Figure 4). Note from Figure 4 that the particle always starts parallel to the force F (i.e., parallel to the y axis). When the friction $a = 0$, the particle proceeds in a repeated pattern after stopping at the y axis (the period is the “inertial period” (24 hours)/(2 sin θ), which is about 12 hours at 79°N). The concept of theoretical free ice drift is based on a stationary situation where forces through the ice are not so significant as close to coasts during the winter. Hence we will expect its direction to be to the right (clockwise) of which moves in steps and/or is subject to significant “friction” compared to hypothetical free ice drift. Comparing Plates 1 and 2 with Figure 4, a typical value for the friction coefficient would be between f and $5f$ (assuming the model conditions and non-stagnant ice).

3. Theoretical Arguments: Conceptual Models

3.1. Universality: Critical Shifts in the Ice Field

We here search for universality in the description of motions in the ice field. We want to sort out relevant aspects leading to possible rapid formation and breakup of fast ice and large rigid areas. Aggregation and fracturing in the ice field form fractal structures. Figure 5 conforms to this idea, showing examples of number $N(x)$ of floes with area larger than the limit x [see Korsnes, 1993a]. The actual floes last, by definition, more than 3 days in this case. These size distributions indicate scaling relationships $N(x) \sim x^{-d_f}$ with different values of the exponent d_f . If large floes are more likely to deform than small floes, then the graphs in Figure 5 will tend to tilt (increasing d_f). A similar argument applies for aggregation/clustering of ice floes (decreasing d_f). Hence we cannot expect a universal “fractal dimension” for deformations in the ice field. Figure 5 demonstrates this. Floes often “aggregate” and break throughout the winter, and their definition depends on time as pointed out by Korsnes [1993b].

Note that the actual identification of rigid areas is sensitive to small deformations in the ice during 3 days and hence on the

parameter ε in (3). Such small deformations are outside the scope for this study.

We may expect that large groups of densely packed ice floes have bulk properties as “granular materials” (compare, e.g., Erlingsson [1988, 1991]). This means that deformations tend to concentrate in active zones covering a small fraction of the ice field area. Mandelbrot [1983] suggested turbulent dissipation takes place on a fractal set. We may similarly imagine that deformations take place on a “fractal” set (not filling up an area). Hence voids will emerge as rigid areas between many linear active zones.

Areas with the smallest deformation rates will tend to get rigid (freeze together) first when freezing takes place in the ice field. Those rigid areas also become stronger with time. Hence we get some irreversibility in the deformation process.

Freezing builds up rigidity in the ice field. However, the following mechanisms help make a situation of large rigid areas in the ice field unstable: (1) large rigid areas in the ice field weaken the heat flux from water to air; hence transport of heat from deeper waters below the ice causes dilution and breakup of large rigid areas; (2) formation of shore-fast ice, including deep and long ridges, produces sites of heavy vertical mixing in areas of strong tidal currents; this may bring heat up to the ice and dilute it; and (3) formation of large rigid areas in the ice field will directly affect the redistribution of the ice and produce singularities in the field of stress and strain. When the ice field consists of small floes, deformation of one floe will not directly produce deformations far off in comparison to the situation when the ice field consists of large rigid areas. Deformations will, in the latter situation, tend to spread over large distances. One explanation of the breakup of shore-fast ice includes melting (heat flux), and an alternative explanation is a mechanical phenomenon. We, in both cases, imagine that the rigidity in the ice field evolves toward an unstable state. We therefore look for features in the dynamics of sea ice during winter resembling the general phenomenon of self-organized criticality. Bak et al. [1988] introduced this concept by using a “sandpile” or “earthquake” model. It gives general insight into statistical and temporal aspects of natural systems tending to stay in a state of marginal stability.

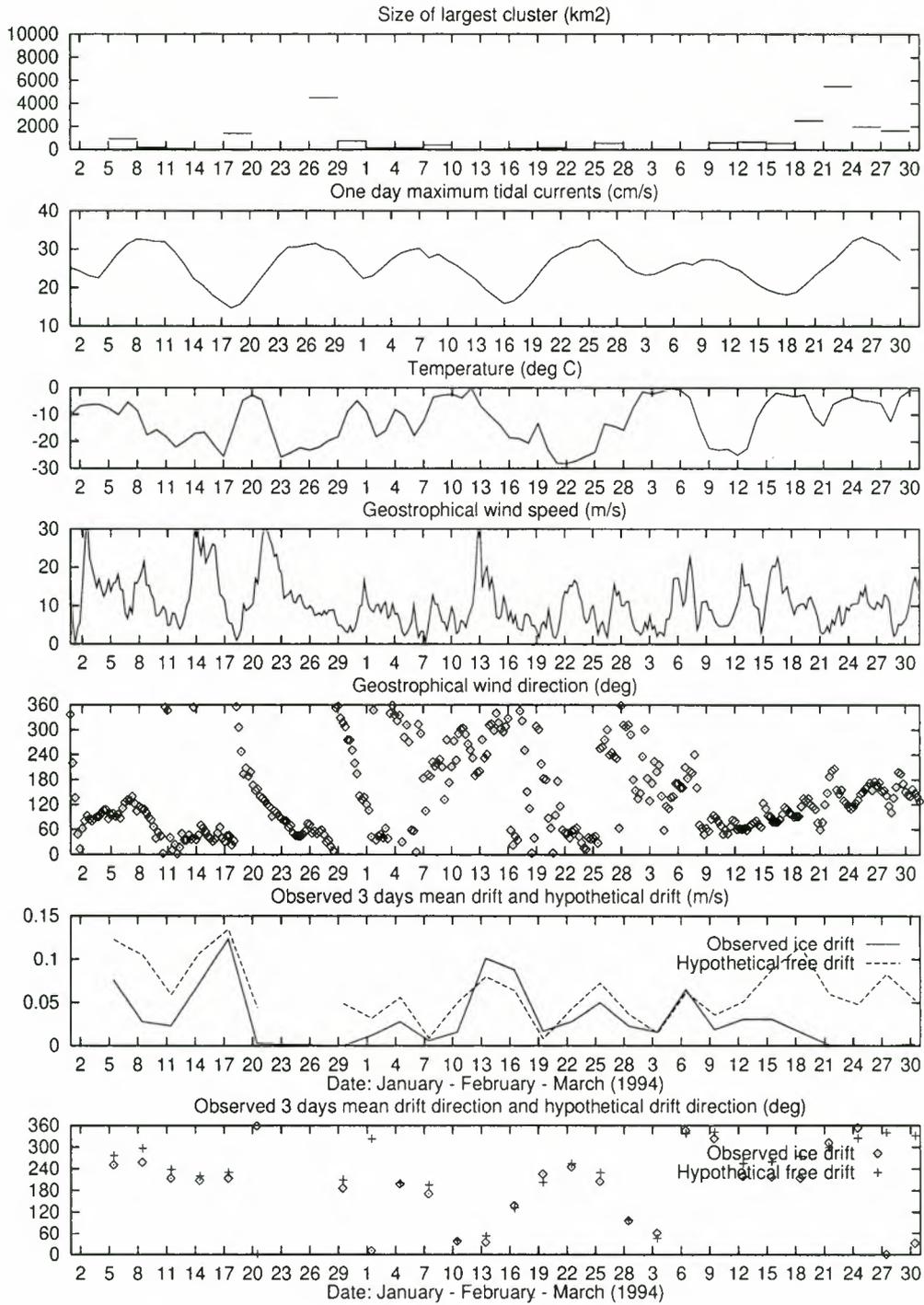


Figure 3. Size of the largest rigid area between subsequential image times, 3 days apart, estimates of maximum 1-day total tidal currents at selected positions 78°47'N, 25°48'E, and meteorological data for January, February, and March 1994. The tidal estimates are from the model by *Kowalik and Proshutinsky* [1993, 1994] consisting of the constituents M2, S2, K2, N2, K1, O1, P1 and Q1. The lowest two graphs show observed and hypothetical 3-day average free ice drift.

3.2. A Simple Granular Model

Plates 1 and 2 indicate that large rigid areas tend to form when the ice moves toward the northwest (pressed against Nordaustlandet) and that these areas break down when the ice moves from this direction. However, Plate 3 gives an example of the opposite. In this latter case a large rigid area has developed in a situation with stable wind from the northeast. A

general breakdown takes place during the 3 days after when the wind has turned (compare Figure 3).

A simple granular model for the bulk properties of the (homogenized) ice field gives perception to these observations. The well-known (yield) criteria for deformation is in this case

$$m \cdot T_n = (n \cdot T_n) \alpha \tag{5}$$

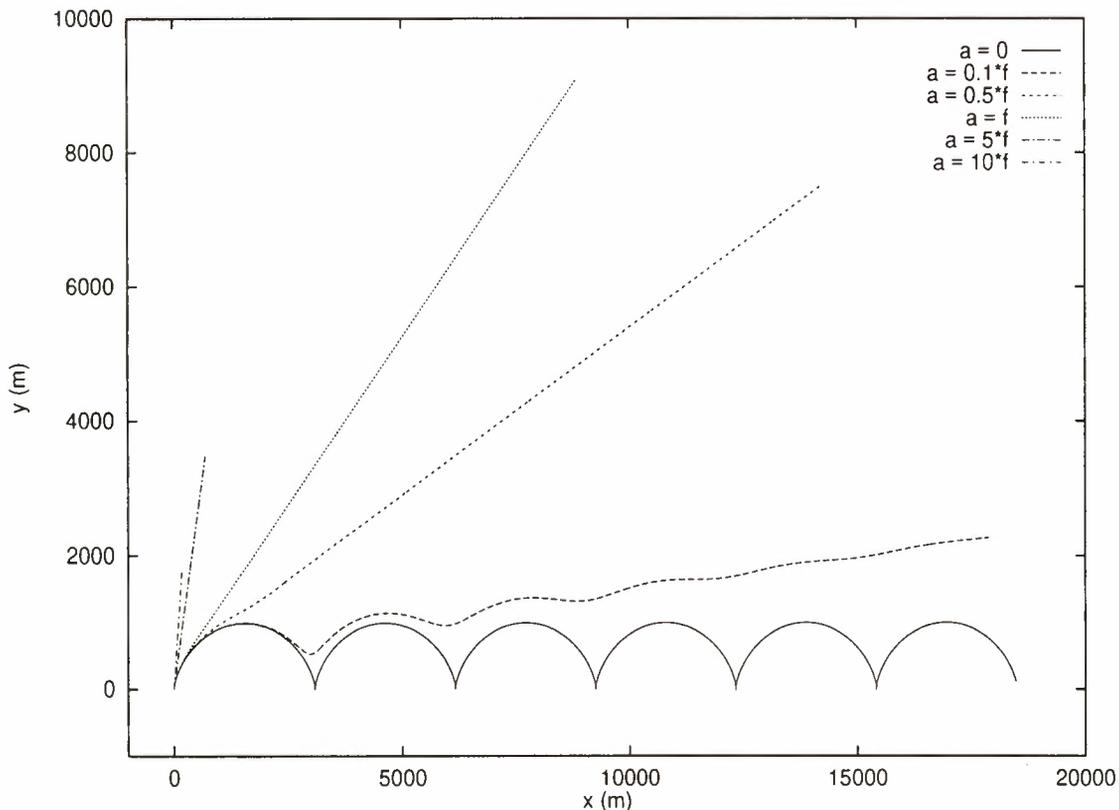


Figure 4. Track of particle moving according to (4) for various values of friction, the friction parameter a , and constant forcing $F = 10^{-5} \text{ m s}^{-2}$ parallel to the y axis. The particle starts at rest at the origin.

where \mathbf{T} is a tensor representing the horizontal (bulk) stress within the ice ($\mathbf{T}\mathbf{n}$ gives the total force per unit length on the line which is normal to the unit vector \mathbf{n} through the actual point). Here \mathbf{m} is a unit vector making \mathbf{m} and \mathbf{n} define a right-hand coordinate system. Also, $\sigma = \mathbf{n} \cdot \mathbf{T}\mathbf{n}$ and $\tau = \mathbf{m} \cdot \mathbf{T}\mathbf{n}$ are the normal and shear force on the line, respectively; α is the internal (bulk) friction coefficient. Figure 6 illustrates this situation as a brick on a slide (with unit mass). Assume that we increase the slope angle θ of the slide until the brick starts to slide because of the gravity \mathbf{g} at $\theta = \theta_f$. The friction coefficient α between the brick and the slope determines θ_f ($\tau = \alpha\sigma$)

$$\theta_f = \arctan \alpha \tag{6}$$

Convergence in the ice field indicates increasing pressure $\sigma = \mathbf{n} \cdot \mathbf{T}\mathbf{n}$ compared to shear forces. This shows that the ice field will tend to stand against a stronger shear force $\tau = \mathbf{m} \cdot \mathbf{T}\mathbf{n}$. In such a situation the ice field consolidates by freezing to make large rigid areas.

The packing of ice floes in an ice field as a result of a large-scale convergence depends, in general, on the small-scale movements during the convergence. Fast large-scale convergence compared to the small-scale random movements gives looser packing and weaker bulk strength properties than otherwise. This is in accordance with treating the same type of granular materials in different ways. It is then well known that the bulk properties vary depending on the development of packing and load [see, e.g., *Nedderman, 1992, chapter 6; Claudio and Imposimato, 1996*]. The mechanical properties of materials will in general depend on the organization and redistribution of the components. The polar sea ice has in this way a “memory.” It has a large set of possible stable states (like

sand). External forces must exceed critical values in order to deform rigid areas. We therefore expect deformations to happen in “bursts” like earthquakes. Plate 3 shows an event where a large rigid area breaks up in a few days (January 29 to February 1, 1994). A similar event takes place in late March (Plate 2) under different wind directions. The deformation patterns depend on the spatial distribution of forces and how these redistribute during deformations.

The expressions above are mainly a mechanical explanation. However, a general divergence in the ice field can be connected to upwelling in the ocean giving melting. This will also tend to produce breakdown in the ice field.

3.3. Percolation Model

The examples of ice drift given by Plates 1 and 2 show a dramatic change in the deformation (or rigidity) of the ice field during 3 days. Large rigid areas “spontaneously” appear. There is a large difference during January 8–11, for example, between observed ice drift and hypothetical free ice drift (the latter is more than 2 times larger than the first). The largest rigid area during this period has a diameter of about 15 km. We would imagine that if this situation were constant during the actual period (January 8–11, where the wind direction is in the range 9° – 120°), forces through these rigid areas would not significantly affect the ice drift. Hence we can assume that in this period there were events where larger temporary rigid areas transferred force through the ice. This section addresses generic concepts to explain how such rigid areas emerge and disappear with minor variations in external forcing on the ice field.

A simple self-similarity argument using the standard theory

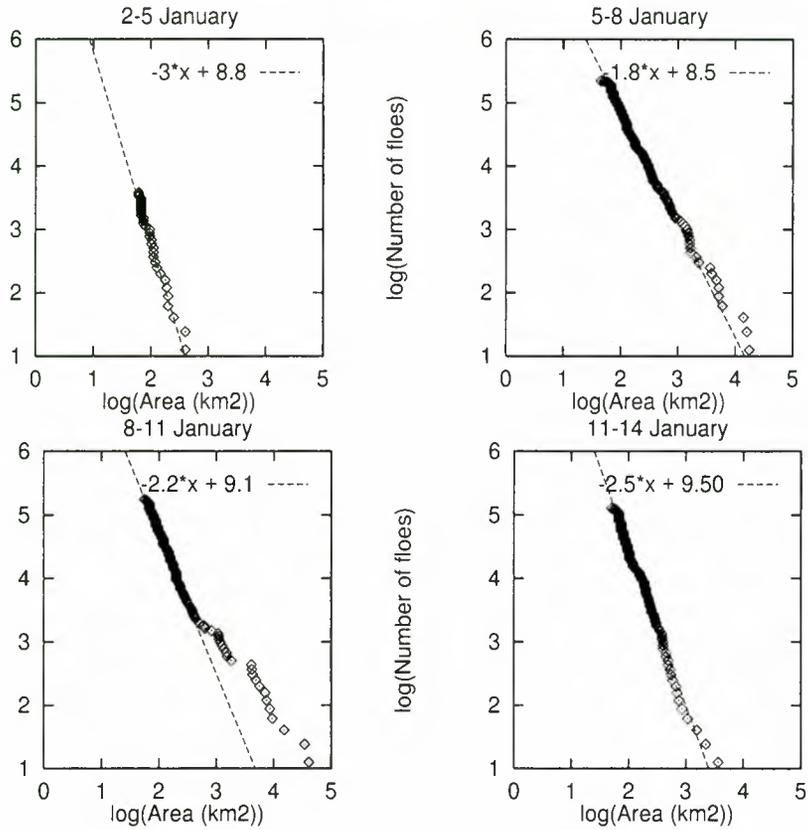


Figure 5. Size distributions for floes lasting (3 days) in the periods January 2–5, 5–8, 8–11, and 11–14, 1994. The graphs show the number of floes larger than given size together with straight lines with given slope.

of percolation [see, e.g., *Stinchcombe and Watson, 1976; Stauffer, 1985; Feder, 1988*] provides the conceptual framework for a regime shift in transfer of force through the ice field. The author emphasizes these concepts because of their simplicity and generality. Percolation theory applies in various situations.

The following arguments illustrate a critical condition for transfer of force through the ice field (following *Turcotte*

[1992]). Four similar square grids, each with (independent) percolation probability $P_1 = p$, are given. These squares are added to form a squared area 4 times larger. Figure 7 lists the possible configurations of permeable (filled) subsquares within this larger square. Let P_2 be the probability that the composed square is permeable. We get, after some rearrangement, that

$$P_2 = 2P_1^2 - P_1^4 \tag{7}$$

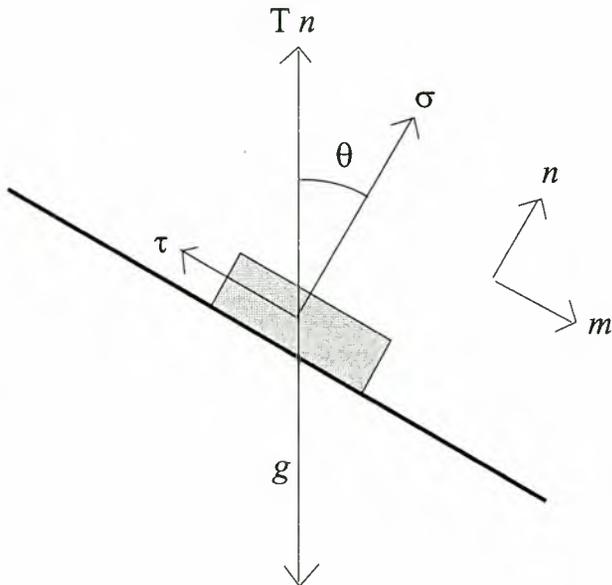


Figure 6. A brick on a slide.

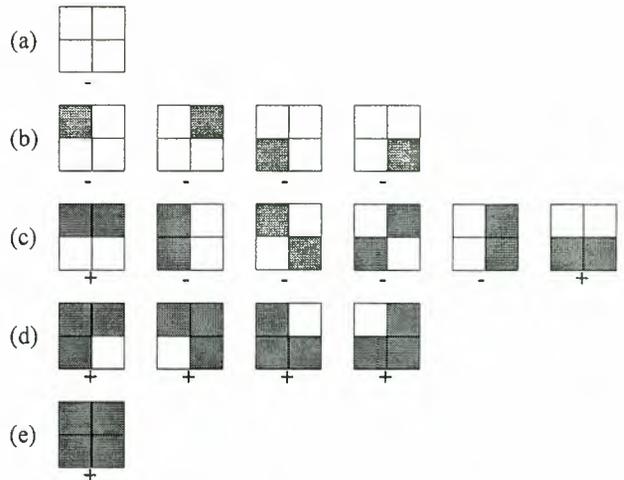


Figure 7. Possible configurations of permeable subgrids [from *Turcotte, 1992*] for calculation of probability of total permeability (crosses).

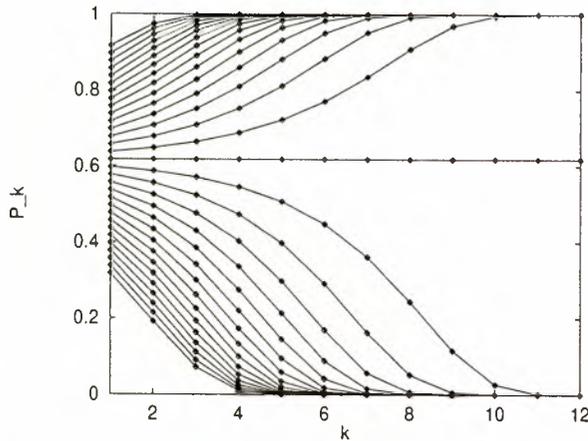


Figure 8. Probability P_k (as a function of k) for permeability of a combined square consisting of k^2 elementary subsquares with various probability $P_1 = p$ for percolation. It is a “regime shift,” or unstable point, at $p = 0.618$, where a small perturbation from this value gives different large-scale (bulk) properties of the combined body.

Recursively repeating the argument, we get

$$P_k = 2P_{k-1}^2 - P_{k-1}^4 \quad (8)$$

This gives three possible limit points for P_k within the range $[0, 1]$. These values are 0, 0.618, and 1. Here $p^* = 0.618$ is a (unstable) critical point. Figure 8 shows the probability P_k of percolation of a combined square consisting of k^2 elementary subsquares ($k = 0, \dots, 12$) for various probability $P_1 = p$ for percolation of these subsquares.

If $0 \leq P_1 < p^*$, then $P_k \rightarrow 0$, and $p^* < P_1 \leq 1$ gives $P_k \rightarrow 1$. Hence, as illustrated by Figure 8, there will be a dramatic shift in the long-range percolation properties of a network when the local percolation probability crosses the critical value p^* . If $p = p^*$, then $P_k = p^*$ for $k = 1, 2, \dots$. Hence scaling up a part of the system at this threshold/unstable situation gives a system similar to the whole system. This gives the idea of scale invariance and the appearance of fractals (self-similarity) when a system is in an unstable state. Note that breakdown in the ice field can result from an unstable state giving fractal (self-similar) ice floe size distributions as in Figure 5.

4. Conclusive Remarks

The present work focuses on ideas to explain observations of formation and breakdown of temporary large rigid areas in Arctic sea ice fields during winter. These observations fit into generic concepts of “critical behavior” where a system dramatically changes characteristics when some parameters enter critical sets of possible values (like temperature dependent phase shift of materials). The author feels that the physical mechanisms behind this type of behavior are open for debate and that simple generic concepts and analogies still have much value for understanding such complex systems. Generic aspects and simplifications are in general important for physical modeling and understanding.

The ice field on the actual scales of study seems to exhibit memory, and mechanical properties are not likely to be successfully expressed by a few parameters within a homogenization scale. There can be different processes for formation and

breakdown. One may expect small deformations and strong singularities of forces in the ice field to initiate breakdown. One possible critical shift in the ice dynamics is when freezing becomes effective enough to “repair” small faults in the ice field before they initiate large-scale breakdown. One may imagine an ongoing aggregation process in the ice field, parallel to breakdown. The result will be sensitive to the net heat flux from the ice field.

The result from identification of rigid areas in the ice field as given by Plates 1, 2, and 3 visually indicates small deformations in almost rigid areas during 3 days. The identification procedure for rigid areas (given by equation 2) gives significantly larger prediction errors for sea ice compared to land. This indicates small deformations in the (almost) rigid areas during 3 days. The observed large rigid areas seem only to last a few days. Hence there is a significant probability that almost rigid areas during 3 days have started a deformation process during this period. Plate 2 shows, in contrast, a rigid area lasting several 3-day periods. We expect this object to be more consolidated than more short-lived rigid areas. The prediction error for the model equation (2) in this case is more comparable to the prediction error for land.

Acknowledgments. This work results from the Norwegian Polar Institute (NP) project number 171301. I partly did the work during April and May 1996 on an inspiring research visit to the National Institute of Polar Research (NIPR) in Tokyo. The Japan Society for the Promotion of Science (JSPS) awarded me a fellowship under the JSPS Postdoctoral Fellowship Program for Foreign Researchers (short-term) in Japan for this research visit. The European Space Agency (ESA) provided the ERS-1.SAR.PRI data under the research usage price. Andrey Proshutinsky at the University of Alaska, Fairbanks, produced the tidal current estimates for Figure 3.

References

- Bak, P., C. Tang, and K. Wiesenfeld, Self-organized criticality, *Phys. Rev. A*, 38, 364–374, 1988.
- Claudio, P., and S. Imposimato, Time dependent mechanical behaviour of loose sands, *Mech. Cohesive Frictional Mater.*, 1, 45–73, 1996.
- Erlingsson, B., Two-dimensional deformation patterns in sea ice, *J. Glaciol.*, 34, 301–308, 1988.
- Erlingsson, B., Propagation of characteristics in sea ice deformation fields, *J. Glaciol.*, 15, 73–80, 1991.
- Feder, J., *Fractals*. Plenum, New York, 1988.
- Gorbunov, Y. A., I. D. Carelin, I. M. Kuznetsov, S. M. Losev, and A. L. Socolov, *Osnovy fiziko-statisticheskikh metodov ledovyykh prognozov i raschetov dlya arcticheskikh morey zablagovremennostyu do 30 sutok*, UF (The Basis of the Physical-Statistical Methods for Ice Forecasts and Calculations in the Arctic Seas Beforehand up to 30 Days). 288 pp., Gidrometeoizdat, Saint Petersburg, 1983.
- Hibler, W. I., A dynamic thermodynamic sea ice model, *J. Phys. Oceanogr.*, 9, 815–849, 1979.
- Korsnes, R., Quantitative analysis of sea ice remote sensing imagery, *Int. J. Remote Sens.*, 14, 295–311, 1993a.
- Korsnes, R., A mathematical approach to the segmentation of SAR remote sensing imagery, *Int. J. Remote Sens.*, 14, 313–332, 1993b.
- Korsnes, R., Some concepts for precise estimation of deformations/rigid areas in polar pack ice based on time series of ERS-1 SAR images, *Int. J. Remote Sens.*, 15, 3663–3674, 1994.
- Kowalik, Z., and A. Y. Proshutinsky, Diurnal tides in the Arctic Ocean, *J. Geophys. Res.*, 98, 16,449–16,468, 1993.
- Kowalik, Z., and A. Y. Proshutinsky, The Arctic Ocean Tides, in *The Polar Oceans and Their Role in Shaping the Global Environment*, *Geophys. Monogr. Ser.*, vol. 85, edited by O. M. Johannessen, R. D. Muench, and J. E. Overland, pp. 137–158, AGU, Washington, D. C., 1994.
- Laur, H., ERS-1 SAR calibration, Derivation of backscattering coefficient σ_0 in ERS-1.SAR.PRI products, technical note, issue 1, rev. 0, 17 pp., Eur. Space Agency, Frascati, Italy, 1992.

- Mandelbrot, B., *The Fractal Geometry of Nature*, W. H. Freeman, New York, 1983.
- Nedderman, R., *Statics and Kinematics of Granular Materials*, Cambridge Univ. Press, New York, 1992.
- Reistad, M., and K. Iden, Updating, correction and evaluation of the hindcast data base of air pressure, winds and waves for the North Sea, Norwegian Sea and the Barents Sea, *Res. Rep. 9*, Norw. Meteorol. Inst., Oslo, 1995.
- Stauffer, D., *Introduction to Percolation Theory*, Taylor and Francis, Bristol, Pa., 1985.
- Stinchcombe, R., and B. Watson, Renormalization group approach for percolation conductivity, *J. Phys. C Solid State Phys.*, **9**, 3221–3247, 1976.
- Thorndike, A., Kinematics of sea ice, in *The Geophysics of Sea Ice*, edited by N. Untersteiner, *NATO ASI Ser., Ser. B*, **146**, pp. 489–549, Plenum, New York, 1986.
- Thorndike, A., and J. Cheung, AIDJEX measurements of sea ice motion 11 April 1975 to 14 May 1976, *AIDJEX Bull.*, **35**, 149 pp., Div. of Mar. Resour., Univ. of Wash., Seattle, 1977.
- Thorndike, A., and R. Colony, Sea ice motion in response to geostrophic winds, *J. Geophys. Res.*, **87**, 5845–5852, 1982.
- Turcotte, D., *Fractals and Chaos in Geology and Geophysics*, Cambridge Univ. Press, New York, 1992.

R. Korsnes, Norsk Polarinstitut, 9005 Tromsø, Norway, (e-mail: Reinert.Korsnes@tromso.npolar.no)

(Received December 9, 1996; revised September 24, 1997; accepted November 10, 1997.)

Evolution of the passive and active microwave signatures of a large sea ice feature during its 2½-year drift through the Arctic Ocean

Francis Gohin, Alain Cavanié, and Robert Ezraty

Institut Français de Recherche pour l'Exploitation de la Mer, Brest, France

Abstract. Although estimation of the total ice concentration from special sensor microwave imagers (SSM/I) has proven to be successful, none of the various algorithms developed to discriminate new and older ice provide satisfying results. While the strong contrast between the emissivity of sea ice and that of open water can be utilized to provide reliable estimators of the total ice concentration, passive microwave characteristics of second-year and multiyear ice may locally evolve in different ways, even during the cold season. Scatterometers, as the active microwave instrument in wind mode (AMI-wind) on board the European Remote Sensing Satellites (ERS), provide backscatter data which have a higher sensitivity to the surface topography of ice and a better stability in time, at a resolution compatible with the SSM/I measurements. Here we present the evolutions of the microwave properties of an ice feature appearing along the shores of Novosibirskiye Ostrova (New Siberian Islands) at the end of July 1992 as the ice ages during its 3-year drift toward the Fram Strait. The track of this well-defined ice surface is easily followed on the maps of the backscatter coefficient provided by the AMI-wind during the cold season. In summer, because of melting, the ice undergoes critical changes which alter its microwave signatures and hamper automatic tracking. Moreover, on approaching the Fram Strait the resolution of the scatterometer is not sufficient to capture the complex and rapid transformations of the ice cover. To compensate for this, buoy data obtained from the International Arctic Buoy Program are used, alone during summers or together with satellite data, to build basin-wide ice displacement fields. These displacement fields, successively applied to each pixel of the ice feature selected, provide a series of Lagrangian observations. During the drift, which ends in May 1995, the active and passive signatures evolve coherently, except for the cold season 1992–1993 when unrealistic multiyear ice concentrations are deduced from the brightness temperatures, which, at that time, are much less stable than the backscatter coefficient over the ice surface tracked, identified as second-year ice.

1. Introduction

The concentration and the type distribution of sea ice can be considered as results of the global climate and statistically analyzed as indicators of its fluctuations [Gloersen, 1995] or presented as major elements of the world hydroclimatic system, since the export of fresh water and ice through the Fram Strait influences convection and deep water formation. The origin of the so-called Great Salinity Anomaly [Dickson *et al.*, 1988] has been attributed to an increase in sea ice extent at the end of the 1960s [Aagaard and Carmack, 1989]. On the basis of results of a simulation of the Arctic ice-ocean system for 1955–1975, Häkkinen [1993] attributes the widening of the transpolar drift stream in 1968 to a well-developed low-pressure area extending from the Iceland Sea to the Barents and Kara Seas opposed to the Beaufort High, which recovers after several years of weakening. This effect being amplified by advection of thicker ice north of Greenland, the author evaluates to 1600 km³ the excess of fresh water exported through the Fram Strait that year, leading to a total export about twice as large as the average.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC03333.
0148-0227/98/97JC-03333\$09.00

Considering that the sensible heat flux from the ocean to the atmosphere is 1 or 2 orders of magnitude greater over thin ice than over thick multiyear ice, the determination of the ice type is crucial for Arctic climate models. The special multichannel microwave radiometer on board Nimbus 7, launched in 1978, and its successor, the special sensor microwave imager (SSM/I) in 1987, have provided a long temporal record of data. These instruments offer a daily coverage over polar oceans with a fairly coarse resolution (30–50 km). Multichannel algorithms have been developed to estimate total sea ice concentration and the concentration of multiyear ice [Steffen *et al.*, 1992]. Estimating ice thickness from space is much more difficult, though local thresholding methods can be used to separate sea ice types representing thickness categories on images provided by synthetic aperture radars (SAR) [Haverkamp *et al.*, 1995]. However, sufficient sets of data over the whole Arctic Ocean are crucially missing, even if we agree that our understanding of the Arctic climate will depend more and more on modeling, guided by data assimilation. Walsh and Zwally [1990] consider that so many points have to be assessed in models and products derived from the observations that these two possible approaches to quantify the ice pack on the basis of numerical modeling or observations have still to be compared with each other and improved.

Although they are designed to measure winds at the ocean surface, scatterometers provide valuable information over sea ice with a resolution similar to that of the SSM/I. The active microwave instruments in wind mode (AMI-wind) flown aboard the European Remote Sensing Satellites (ERS) are C band (5.3 GHz), vertically (VV)-polarized scatterometers. AMI-wind data have been provided by ERS-1 from August 1991 to June 1996 and since then by ERS-2. Because the AMI-wind backscatter over sea ice is independent of the azimuth angle but varies as a function of incidence angle, maps of the backscatter coefficient at a fixed incidence angle of 40° [Gohin and Cavanié, 1994] have been produced and distributed by the Centre ERS d'Archivage et de Traitement (CERSAT) after projection on the same polar grid used by the National Snow and Ice Data Center (NSIDC) to present SSM/I data. Backscatter data can thus be conveniently compared to the brightness temperatures as well as to the derived products of the SSM/I (ice concentration and ice type).

In this study the variability of ice parameters retrieved from low-resolution microwave instruments is investigated by following the evolution, with time, of the microwave signatures of a large ice feature drifting toward the Fram Strait. Over the area studied, which was almost fully ice covered, the multiyear ice concentration is of particular importance. As noted by several authors [Comiso, 1986, 1990; Thomas, 1993; Kwok *et al.*, 1996], this multiyear ice concentration as obtained from SSM/I fluctuates during the cold season and is low compared to the ice concentration estimated at the end of the previous summer.

The present Lagrangian approach enlarges a concept already presented by Kwok *et al.* [1995], who proposed to determine the age distribution of sea ice from Lagrangian observations of ice displacements estimated from SAR data; they selected a timescale in days and a 5-km pixel size. Our study is based on a timescale of months and a pixel size of 25 km, which are adapted to the size of the Arctic Ocean, the mean residence time of sea ice from its formation to its evacuation, and the size of the selected ice feature. This feature, which covers about 400,000 km², is followed from its initial formation (as ice surviving its first summer) along the northern shores of Novosibirskiye Ostrova (New Siberian Islands) at the end of July 1992 until its evacuation through the Fram Strait in May 1995.

Eulerian fields of ice displacement, for 8-week periods, are calculated over the Arctic Ocean by a two-step procedure using both backscatter maps and buoy data from the International Arctic Buoy Program (IABP). By applying the successive displacement fields to each pixel of the selected ice feature a Lagrangian series of microwave observations is produced. This $2\frac{1}{2}$ -year time series sets in evidence the seasonal and interannual fluctuations of the active (AMI-wind) and passive (SSM/I) microwave signatures of ice as it progresses through the Arctic Ocean and ages, evolving from the state of essentially second-year ice to that of third- and fourth-year ice. Thus variations in active and passive microwave signatures are observed without any significant perturbation caused by ice advection.

2. Data

2.1. Buoy Data

The buoy data have been obtained from the Polar Science Center at the University of Washington. The data set contains daily positions in latitude and longitude as well as the identi-

fication code of the buoys. These positions have been transformed in horizontal and vertical coordinates (respectively columns and rows) on the 25- × 25-km resolution, 304- × 448-pixel grid of the stereopolar projection used at NSIDC [Gloersen *et al.*, 1992].

At the start day of this study (December 3, 1992), 45 buoys were deployed over the area selected for ice tracking (Plate 1). In March 1993, 11 buoys were added to the network, with an important concentration north of Novosibirskiye Ostrova, around the ice feature to be tracked. As about one tenth of the buoy positions are missing on a daily basis, the number of buoy displacements, calculated using pairs of individual positions 8 weeks apart, is reduced to <40.

2.2. Satellite Data

2.2.1. AMI-wind backscatter data. The AMI-wind has three antennas pointing 45° , 90° , and 135° to the right of the satellite ground track. Thus, because the satellite is yaw steered, the fore and aft beams point to the surface with the same incidence angles. The antenna swaths are composed of 19 cells, 25 by 25 km in size, starting 225 km from the ground track. Given a nominal flight altitude of 800 km, the incidence angles of the electromagnetic beams vary approximately from 25° to 59° for the fore and aft beam and from 18° to 45° for the central beam. The noise level of the individual backscatter measurements is about 6% and is almost constant over the whole range of incidence angles and backscatter values, σ^0 .

Over consolidated sea ice, within the range of incidence angle (θ) of the AMI-wind, σ^0 (in decibels) is well described by a linear function of θ , the slope depending on sea ice type [Gohin and Cavanié, 1994]. A similar behavior was observed by Drinkwater and Carsey [1991] over the Antarctic Ocean for Seasat A scatterometer system (SASS) at Ku band. This property is used to map backscatter values at the constant incidence angle of 40° (σ_{40}^0) over the whole polar ocean. Because of the grid size, of the ERS orbiting cycle (3-, 35-, or 135-day repeat cycle), of the 500-km swath of the AMI-wind, and of interruptions due to SAR imaging mode, it proved necessary to use 1 week of data in order to cover the whole polar ocean and obtain a reasonable stability of the final map. The following procedure is used: for each cell of the polar grid an estimation of $d\sigma/d\theta$ is performed using, on the one hand, the σ^0 at the incidence angle of the central antenna and, on the other hand, the average of the σ^0 of the fore and aft beams. For each pixel and over the whole week a mean slope $d\sigma/d\theta$ is estimated, and by applying the linear equation (1), σ_{40}^0 are calculated and averaged.

$$\sigma_{40}^0 = \sigma_{\theta}^0 + \frac{d\sigma}{d\theta} (40 - \theta) \quad (1)$$

During the cold seasons, using two consecutive maps and assuming that, over selected areas and pixel by pixel, the differences of σ_{40}^0 are only due to measurement and processing noise, it is possible to estimate the quality of the final image as the square root of half the variance of the relative differences. Such a computation indicates that the relative noise decreases from 8 to 2% when σ_{40}^0 increases from 0.0025 to 0.15. In the following, only σ_{40}^0 values will be used, thus the subscript "40" will be dropped in the notation.

Plate 1 presents a time series of σ^0 maps, from December 1992 to May 1995, on which are plotted the IABP buoy positions. During the summer periods, because of surface melt and

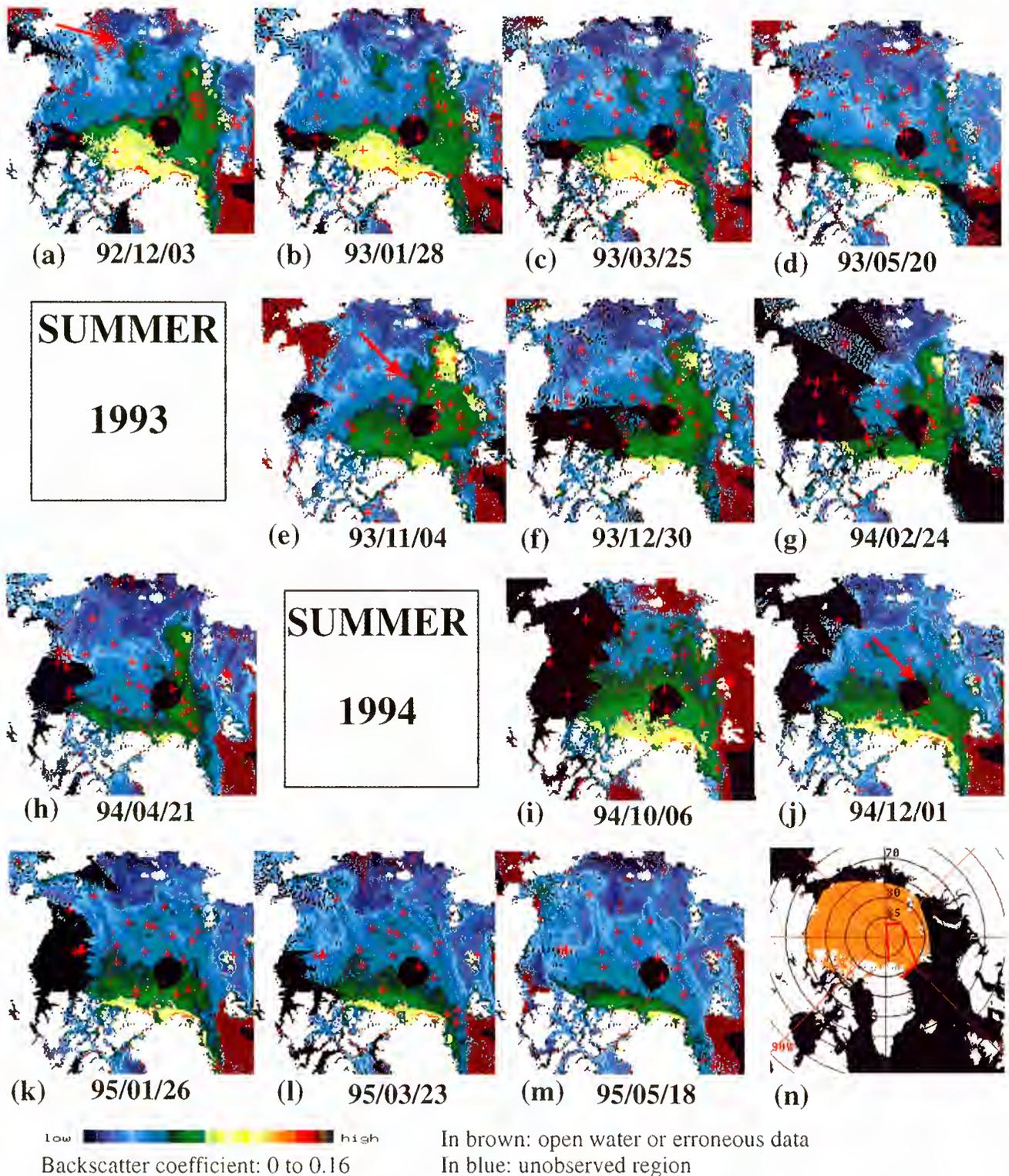


Plate 1. Weekly active microwave instrument in wind mode (AMI-wind) backscatter maps spanning the drift of the selected ice feature. Three arrows show its positions after the summers of 1992, 1993, and 1994. Images are separated by 8-week periods centered on the indicated dates. Red crosses indicate buoy positions. Plate 1n shows the area where displacements are calculated and the limit of the Fram Strait area, in red.

ponds. backscatter signatures deviate so much from those of winter sea ice that they are useless for tracking; therefore corresponding maps were skipped. Although year to year backscatter variations are evident, for example, north of Greenland, the annual signatures remain stable throughout the winter

periods with some indication of surface changes in late April or May. The stability of σ^0 allows us to identify the sea ice structure, marked by an arrow, on the December 1992 map and to follow it, visually, during its 1992–1993 winter drift until summer melt blurs the map.

2.2.2. SSM/I data. The SSM/I measures microwave emission at 19.35, 37.0, and 85.5 GHz in both vertical and horizontal polarizations and at 22.23 GHz in vertical polarization only. The footprint sizes are 55, 32, 13, and 49 km, respectively. The incidence angle is constant and equal to 53°.

The data used here are the daily Defense Meteorological Satellite Program (DMSP) F11 brightness temperature grids at 19 GHz (vertical and horizontal polarization) and 37 GHz (vertical polarization) provided by the NSIDC. These temperatures (T_b) are combined to compute the gradient and polarization ratios (GR and PR) defined as $GR = (T_{b_{37v}} - T_{b_{19v}})/(T_{b_{37v}} + T_{b_{19v}})$ and $PR = (T_{b_{19v}} - T_{b_{19h}})/(T_{b_{19v}} + T_{b_{19h}})$. These ratios damp surface temperature changes and are used as variables to characterize ice types and to estimate first-year, multiyear, and total ice concentration (C_{FY} , C_{MY} , C_T) according to the NASA team algorithm [Cavalieri, 1992]

$$C_{FY} = (a_0 + a_1PR + a_2GR + a_3PRGR)/D \quad (2)$$

$$C_{MY} = (b_0 + b_1PR + b_2GR + b_3PRGR)/D \quad (3)$$

$$C_T = C_{FY} + C_{MY} \quad (4)$$

where

$$D = c_0 + c_1PR + c_2GR + c_3PRGR \quad (5)$$

The coefficients a_i , b_i , and c_i are derived from constant brightness temperature of tie points corresponding to open water, first-year, and multiyear ice for the Arctic or open water, type A and type B ice for the Antarctic.

3. Estimation of Basin-Wide Ice Displacement Fields

3.1. General Presentation of the Two-Step Method

The method developed to build ice displacement fields is based on a two-step procedure somewhat similar to that employed by Fily and Rothrock [1987], adapted to our low-resolution satellite data. Fily and Rothrock track sea ice by a nested correlation method applied to Seasat SAR images. From highly smoothed images they calculate a first-guess displacement field, used to guide the search for maximum correlation between small areas of ice common to two successive high-resolution images. Here first-guess displacements are calculated on the NSIDC grid for one pixel out of four along both horizontal and vertical axes (100- × 100-km coarse grid). The horizontal and vertical displacements are obtained from IABP buoy displacements by an objective interpolation method, kriging [Olea, 1974]. For each displacement on the coarse grid, kriging yields an estimation of the local error. A detailed description of kriging is presented in section 3.2.

The second step of the procedure takes advantage of the stability of σ^0 in time. Ice located at point X (Plate 2e–2f) is supposed to have drifted to a new position located within a circle whose radius is proportional to the error estimation. The center of the circle, X_0 , corresponds to the initial position shifted by the first-guess displacement. On the fine grid, all 7- × 7-pixel arrays of the backscatter image 8 weeks later, whose center pixels fall within the circle, are searched to find the one that has the maximum correlation with the 7 × 7 array centered on X on the initial image. When a highly correlated array is found, the center of this array defines a new tie point X' associated to X. Limiting the search to circular areas con-

strained by the error estimations provided by the kriging procedure gives the tracking algorithm the required restriction for search (unlimited areas would lead to many fliers); for points in the vicinity of buoys the first-guess displacement will not be modified since the radius of the circle will be small due to the reduced error estimate. Details of this search procedure, applied automatically on every point of the coarse grid, are given in section 3.4. By kriging, a new estimation of the displacement field over the coarse grid can now be made using both buoy and backscatter image displacements. Finally, over the Arctic area selected (colored in orange on Plate 1n), a bilinear interpolation is applied to extend the displacement fields over the fine NSIDC grid.

The improvement due to the second step will not be uniform on the grid, as some areas have been deliberately excluded (the Fram Strait area within the red box, on Plate 1n) or are deprived of AMI-wind observations because of SAR operation (part of the Beaufort Sea on Plate 1g). The method will be most efficient in places where ice is strongly featured; such conditions are found around the ice feature originating from north of Novosibirskiye Ostrova, particularly during the 1992–1993 cold season.

3.2. Kriging Hypotheses

Considering the random variable $D(x)$, where D is the displacement along the X or Y axis, we have to set a minimum number of hypotheses that are necessary to obtain the estimations of D . Elementary hypotheses in geostatistics [Matheron, 1970] do not deal with the mean value of $D(x)$ at any location x , as the mean is unknown and different from one location to another, but deal with the increments $[D(x+h) - D(x)]$ of the random variable D between two locations x and $x+h$. They are formulated as

$$\mathcal{E}[D(x+h) - D(x)] = 0 \quad (6)$$

$$V(h) = \frac{1}{2} \mathcal{E}\{[D(x+h) - D(x)]^2\} \quad (7)$$

where \mathcal{E} is the expected value and V is the semivariogram. Equation (6) states that there is no trend in the increments of the displacement, and (7) states that their variance depends only on the distance, h , between their locations and not on their specific locations.

The semivariograms of the displacements along the X and Y axes have been modeled using the so-called spherical model [Curran, 1988].

$$V(h) = S[1.5(h/a) - 0.5(h/a)^3] \quad h \leq a \quad (8)$$

$$V(h) = S \quad h > a \quad (9)$$

where S and a are defined as the sill and the range, respectively.

Figure 1 shows the adjustment of parameters S and a to the experimental semivariograms. For both components of the displacement a unique range has been estimated equal to 100 pixels (2500 km). However, the X and Y displacement sills are very different, 20 and 45 square pixels, respectively. The higher sill observed for the Y component can be explained by its strong gradient north of the Fram Strait. As ice approaches the Strait, it moves slowly horizontally while it accelerates along the Y axis.

3.3. Estimating the Displacement by Kriging

For each point of the coarse grid over the selected Arctic ocean area the ice displacement is estimated by local kriging

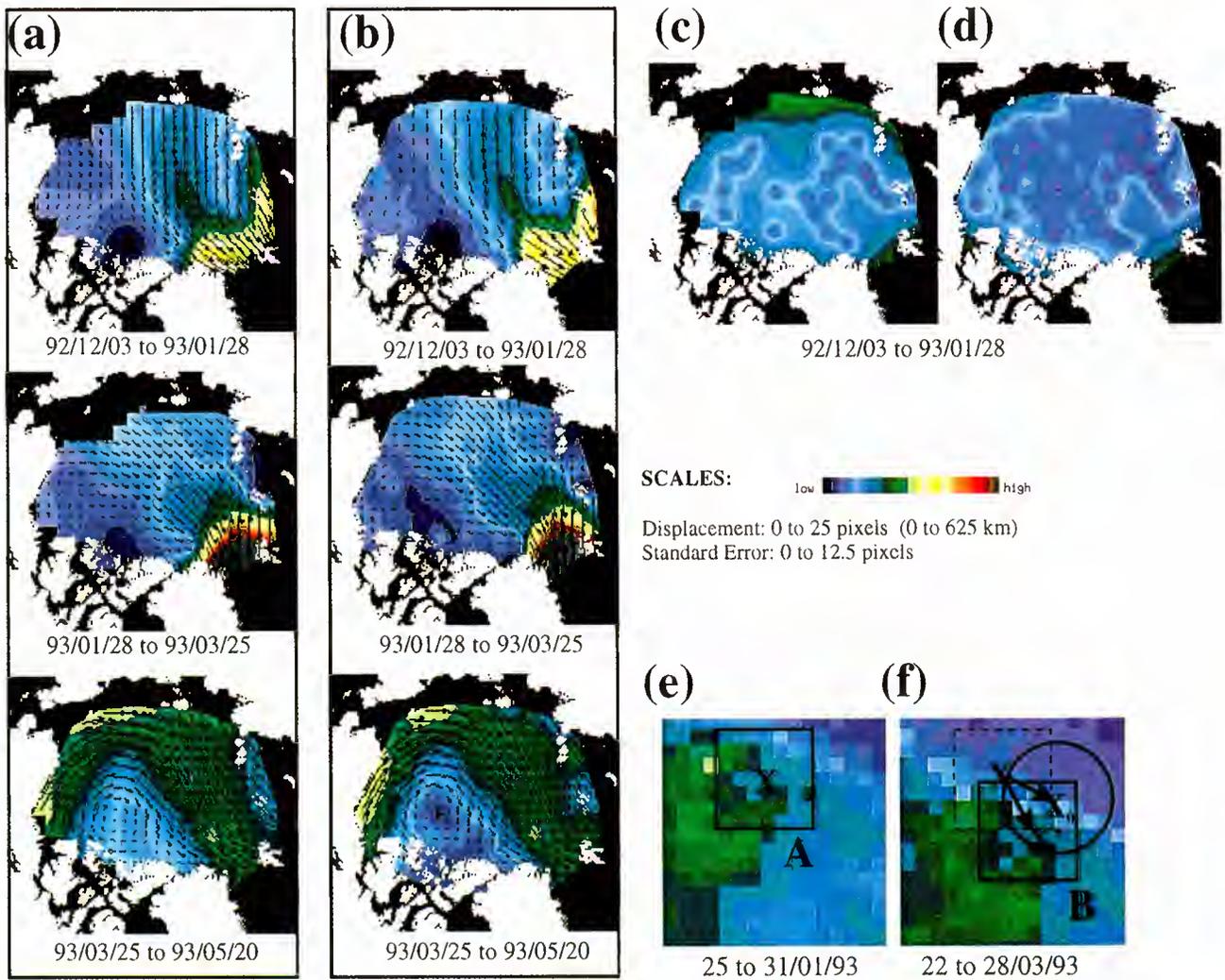


Plate 2. Ice motion estimated for the three initial periods of the study (a) from buoys alone and (b) from buoys and images. Standard error of the ice drift for the first period is presented (c) for buoys only and (d) for buoys and images. (e, f) The method to modify the initial displacement field calculated from the buoy displacements alone is illustrated. X is a location on the coarse grid to which is associated, by kriging of buoy data, position X_0 . This first-guess position is improved by searching for the 7×7 -pixel array B most correlated to the array A centered on X.

using the six closest displacements. The kriging estimator, $D^*(x)$, is a linear combination of $D(x_i)$

$$D^*(x) = \sum_{i=1,6} \lambda_i D(x_i) \tag{10}$$

The λ_i result from the minimization of the variance of errors, $\mathcal{E}\{[D^*(x) - D(x)]^2\}$, under the condition of no bias, $\mathcal{E}[D^*(x) - D(x)] = 0$. The latter condition is satisfied by choosing the weights, λ_i , such that their sum is equal to one. Then from (6)

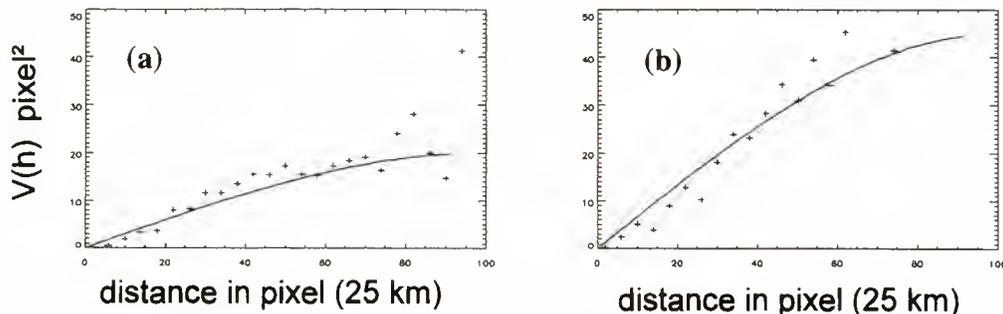


Figure 1. Semivariograms of ice displacements for 8-week periods along the (a) X and (b) Y axis of the National Snow and Ice Data Center (NSIDC) polar grid.

$$\begin{aligned} \mathcal{E}[D^*(x) - D(x)] &= \mathcal{E}\left[\sum_{i=1,6} \lambda_i D(x_i) - D(x)\right] \\ &= \sum_{i=1,6} \lambda_i \mathcal{E}[D(x_i) - D(x)] = 0 \end{aligned} \quad (11)$$

From (7) and the constraint on the sum of the λ_i the following equation results [Matheron, 1970]

$$\begin{aligned} \mathcal{E}\{[D^*(x) - D(x)]^2\} &= \sum_{i=1,6} \lambda_i \mathcal{V}[h(x_i, x)] \\ &\quad - \sum_{i=1,6} \sum_{j=1,6} \lambda_i \lambda_j \mathcal{V}[h(x_i, x_j)] \end{aligned} \quad (12)$$

The minimization of $\mathcal{E}\{[D^*(x) - D(x)]^2\}$ leads to a set of equations linear in λ_i , which, after inversion, yields unique λ_i solutions.

3.4. Extracting Displacement From Successive Backscatter Images

In order to improve the displacement field estimated from buoy data alone, displacements are also deduced from successive backscatter images and incorporated in the set of second-step displacements. Each pixel on the coarse grid is now considered as the center of a 7×7 -pixel array on the fine grid.

For a pair of 7×7 pixel arrays, A on the initial image and B on the following image, the correlation $C_{a,b}$ is defined as

$$C_{a,b} = \frac{\sum_{i=1,n} \sigma_{a,i}^0 \sigma_{b,i}^0 - n \overline{\sigma_a^0} \overline{\sigma_b^0}}{\sqrt{\sum_{i=1,n} (\sigma_{a,i}^0 - \overline{\sigma_a^0})^2 \sum_{i=1,n} (\sigma_{b,i}^0 - \overline{\sigma_b^0})^2}} \quad (13)$$

where $\overline{\sigma_a^0}$ and $\overline{\sigma_b^0}$ are the means of the arrays A and B and n is the number of validated pairs $(\sigma_{a,i}^0, \sigma_{b,i}^0)$ extracted from the arrays. Each pair $(\sigma_{a,i}^0, \sigma_{b,i}^0)$ is validated if both backscatter data have not been recognized as land, open water, or spurious observations (such cases occur during the AMI-wind/SAR switches). If n , the maximum value of which is equal to 49, is lower than 41, the displacement is not validated.

If the correlation between array A and array B is >0.75 , the initial displacement is kept unchanged. If not, the algorithm seeks the maximum correlation at successive distances increasing either from one to six pixels or twice the standard error of kriging, the smaller of the two being chosen; the standard error of kriging is defined as the square root of the sum of the error variances along the horizontal and vertical axes. When a correlation >0.75 is found, the search for maximum correlation does not extend beyond that distance; the displacement is validated and added to the set of buoy displacements. The displacement is not validated if the correlation coefficient does not reach a value of 0.75. These additional displacements, with an error variance set to 8 square pixels, are then used for kriging. This value of the error variance gives them little weight when buoy data are present in their vicinity.

The performance of the algorithm is linked to the evolution of the correlation, $C(d)$, as the lag, d , increases from 0 to 5 pixels. For A centered on a buoy position and B centered on the position of the same buoy, 8 weeks later, the correlation coefficient, noted $C(0)$, would be equal to 1 if σ^0 were perfectly stable and if ice were rigid, without convergence or divergence. Shifting B by a distance d (in pixels) in any direction around the true position of the buoy, the corresponding

correlation coefficient between A and B will decrease from $C(0)$ to $C(d)$. The results of such a numerical experiment applied to the pairs of buoy positions observed for the period December 3, 1992, to May 20, 1993, are presented in Figure 2 in the form of histograms of the correlation coefficient for artificial shifts, d , from 1 to 5 pixels. It can be seen that the modes of the distributions for $d = 0$ to $d = 2$ are quite similar (close to 0.75) and differ significantly from those for $d \geq 3$. The probability of a correlation >0.75 decreases rapidly as the distance increases and has the values 0.47, 0.36, and 0.24 respectively for $d = 0$ to $d = 2$, respectively. Although these results are encouraging, the quality of the displacements is, in practice, judged by their ability to retrieve the motion and the deformation of a selected ice feature, as described in section 4 of this study.

As an illustration of the results of the method, Plate 2 presents the estimated fields for three consecutive 8-week periods, starting on December 3, 1992, and ending on May 20, 1993, obtained by using the buoys alone (initial displacement field shown on Plate 2a) and after adding scatterometer data (Plate 2b). Displacement fields can be quite different, but the general features of Plate 2 are in good agreement with the mean circulation [Colony and Thorndike, 1984; Serreze et al., 1989], dominated by the transpolar flow and the anticyclonic gyre of the Beaufort Sea. Except in the Fram Strait area, in which the satellite data have been intentionally excluded, and the north pole area, where no satellite data are available because of scatterometer coverage, the average precision of the ice displacement estimation is improved but does not reach the low levels of error variance in the vicinity of the buoys (Plates 2c–2d). In particular, the coastal features of the displacement are more accurately determined when scatterometer data are used. Plate 2c also reveals the poor accuracy of the displacement field derived from the buoys to the north of Novosibirskiye Ostrova during the first two periods studied (from December 3, 1992, to March 25, 1993). The closest buoys, located in the Beaufort Sea and north of Severnaya Zemlya, have displacements relatively small compared to those of the area where our selected ice structure appears (Plate 2a). This explains why during 8-week periods, the estimated ice displacements are increased from one to two pixels (25 to 50 km) when using AMI-Wind data. Although individually small, these improvements summed over a cold season make a significant impact on the automatic tracking of ice features.

4. Formation and Displacement of the Selected Ice Mass

4.1. Formation in July 1992

Cyclones are frequent in summertime and induce divergence, which may decrease the ice concentration [Serreze et al., 1989], mainly in the Canadian Basin, where a maximum of cyclonic activity is generally situated. At the end of July 1992, however, a low-pressure system passed north of Novosibirskiye Ostrova [Colony and Rigor, 1993], pushing ice along the northern shores of the archipelago. From July 24 to August 6, 1992, the evolution of the brightness temperature at 19 GHz (H), Plate 3, shows how quickly the ice pack (characterized by its higher temperatures) was transformed, being broken, melted, and moved away as the cyclone progressed through the area. Deep lows moved east of 180°E , enhancing ice divergence north of Novosibirskiye Ostrova. In summer, σ^0 images cannot

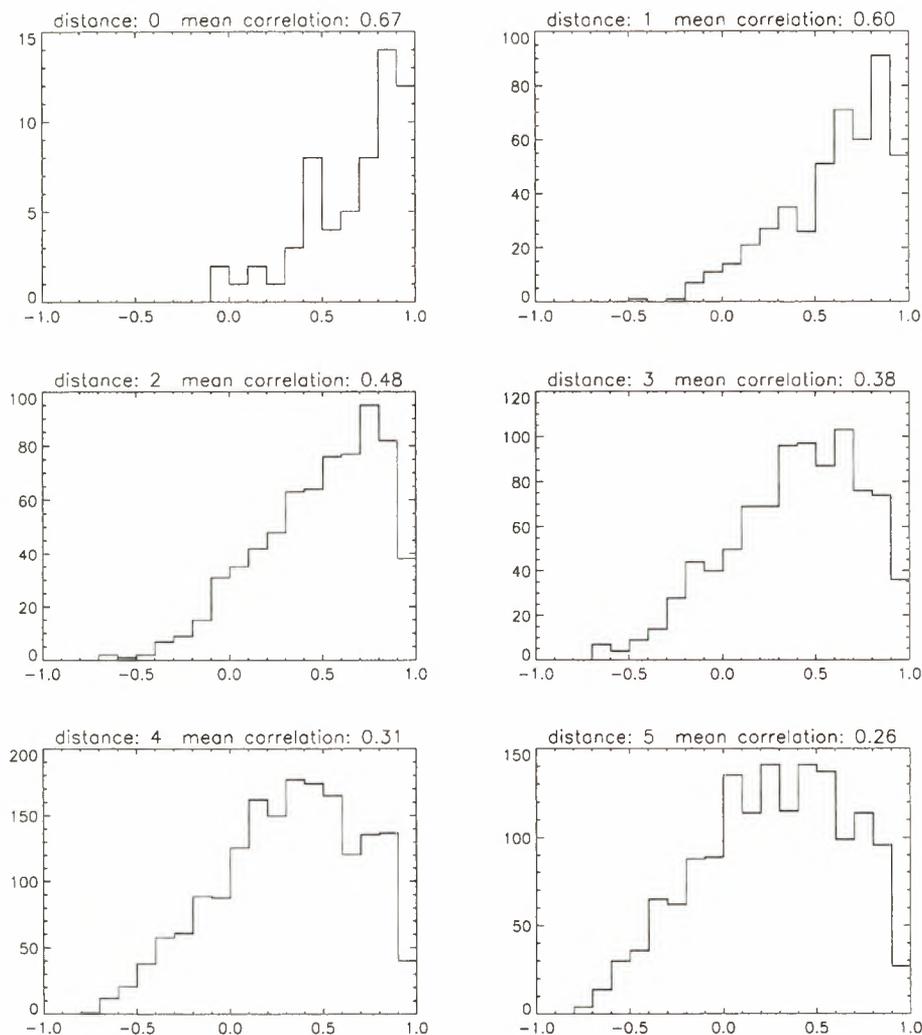


Figure 2. Histograms of correlation coefficients for 7×7 -pixel cell arrays A and B, centered on the positions of buoys 8 weeks apart (distance = 0), and for positions of array B artificially shifted from 1 to 5 pixels (distances from 1 to 5) in any direction.

be used directly because the contrast between open water and ice, covered by melt ponds, is very low. Less sensitive in summer than the backscatter coefficient and better related to ice concentration [Gohin, 1995], images of the derivative of the backscatter, $d\sigma/d\theta$, at 28° incidence angle (Plate 3d) show the same characteristic feature as seen by the SSM/I. Though poorly featured, the ice structure, contoured with a red ellipse on Plates 3c–3d, is already visible at the beginning of August 1992. During August and September, satellite imagery, not shown, confirms this initial situation, which evolves slowly until October (Plates 3e–3h), when the microwave properties are stabilized.

4.2. Locations With Time

From the weekly σ^0 map centered on December 3, 1992, a subimage is manually contoured. This ice area, visible on Plate 4a, is the ice feature selected for this study. It is essentially composed of second-year ice which had just begun drifting away from the Novosibirskiye Ostrova. Using the ice displacement fields previously calculated, the position of each subimage pixel is located on the following backscatter im-

age, and the corresponding subimage is extracted. Plate 4 presents the time series of the ice feature positions as obtained by this method. The quality of the restitution can be estimated by looking at the three first subimages selected from January to May 1993. Although σ^0 evolves, slowly in winter and more quickly in spring, the contours of the selected ice are well retrieved, as well as they could be obtained manually. In November 1993, σ^0 has increased, and the northern part of the feature is hardly discernible from the neighboring ice, while its southern part is quite distinguishable, characterized by high backscatter coefficients (Plate 1e). Ice displacement in the southern part of the feature is less than it is farther in the north. The difference increases with time as the feature is stretched. Three points, indicated on Plate 4a, have been selected on the images, and their displacements, every 8 weeks, are shown on Figure 3. The northern pixel, P1, constantly in the transpolar drift stream, reaches 80°N by August 1994. Its mean displacement is 215 km over 8 weeks, compared to 208 and 154 km for the central and southern pixels, P2 and P3, respectively.

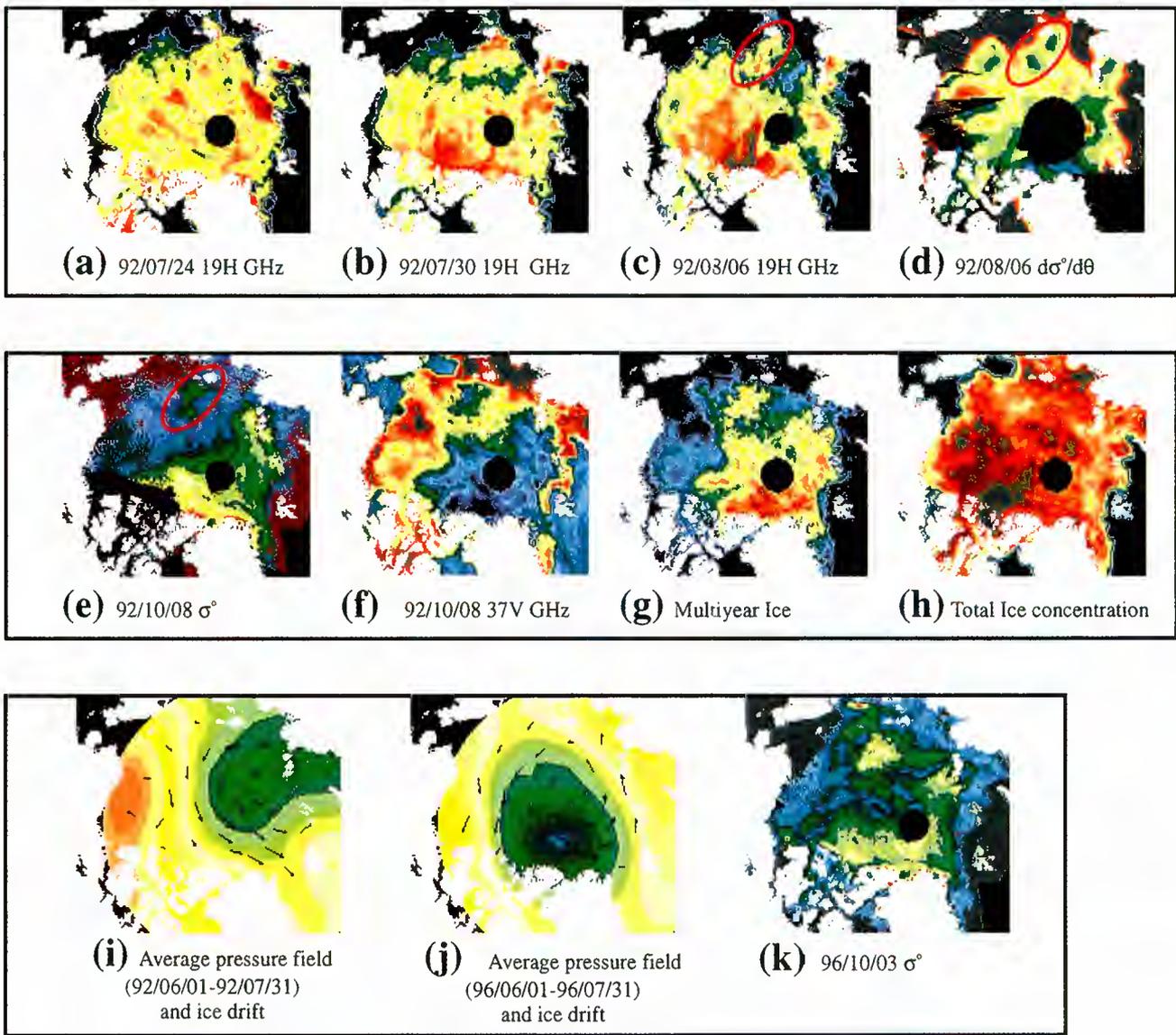


Plate 3. Summer data over the Arctic. (a–c) Images of brightness temperature at 19 GHz (H). (d) The derivative of σ^0 as a function of the incidence angle confirms the formation of the structure. At the beginning of October, the structure is visible (e) on the image of σ^0 , as well as (f) on the image of the brightness temperature. The pressure fields averaged over June and July (i) 1992 and (j) 1996, and the displacement fields produced by the International Arctic Buoy Program are shown. The meteorological situation in summer 1996 led to a backscatter distribution in (k) October, similar to that observed in 1992.

5. Evolution of Microwave Signatures and Multiyear Ice Concentration

5.1. Microwave Signatures

An opposite evolution in backscatter coefficient and in brightness temperature is expected [Ezraty *et al.*, 1994], as the parameters are respectively linked to the reflectivity, Γ , and the emissivity, ϵ , which, at a given wavelength and polarization, is

equal to $1 - \Gamma$. However, differences in wavelengths and incidence angles between active and passive microwave measurements generate significant variations.

Generally, the opposite trends expected in active and passive signatures are observed (Plates 3e–3f). However, a maximum in σ^0 occurs on that part of the ice feature closest to Novosibirskiye Ostrova, while the minimum in brightness temperature is observed farther north. This behavior is also confirmed by the

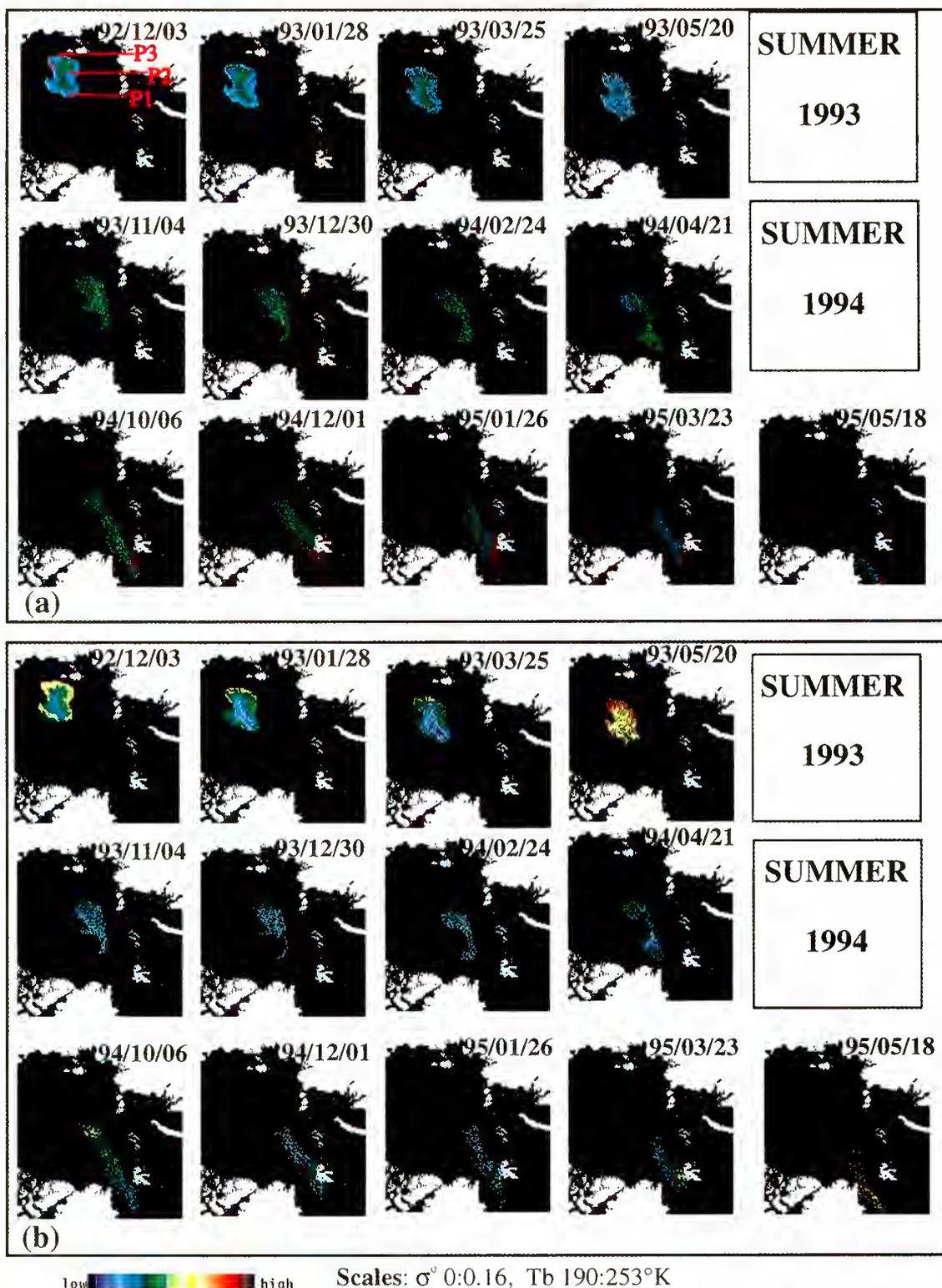


Plate 4. Detail of images along the track of the selected structure; (a) AMI-wind backscatter coefficient and (b) SSM/I brightness temperature at 37 GHz (V).

scatterplot of 37 GHz (V) brightness temperature versus σ^0 on December 3, 1992 (Figure 4d). On Figure 4d the cloud of points splits into two branches; the bottom one, which includes most of the cells of the feature, reaches a minimum in brightness temperature associated to the highest multiyear concentration while the upper one contains the highest values of σ^0 ,

associated with the cells in the vicinity of the coast. The topography of the ice surface there, likely to be rough as the pack has been pushed and compressed to the coast, plays an important role in the σ^0 increase. As expected from previous studies, [Wismann *et al.*, 1996], the time series of σ^0 averaged over the selected feature for a given cold season is much more stable in

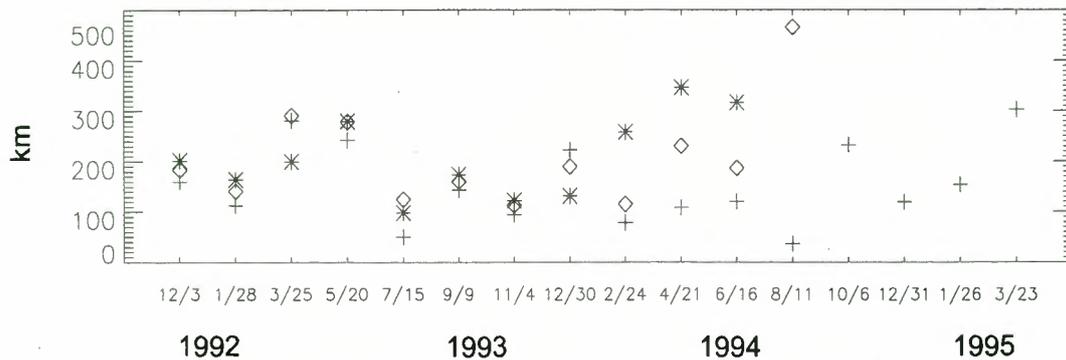


Figure 3. Time series of the estimated displacements over 8-week periods for points P1 (asterisk), P2 (diamond), and P3 (cross). Initial locations are 172°45'E, 82°20'N; 162°25'E, 79°50'N; and 154°E, 75°30'N, respectively.

time than the following passive microwave parameters: brightness temperature, GR, or multiyear ice concentration (Figure 5). This stability in σ^0 is attested not only by the regular level of the averaged backscatter coefficient during the first and the second season but also by the stability of the standard deviation observed on the selected feature (Figure 5b). A small decrease in σ^0 , appearing in spring, coincides with a relatively high increase in brightness temperature and a larger dispersion of

the data. In 1994 the decrease occurs in late April, as shown by the time series of the mean σ^0 ; this is also visible on the corresponding images of Plates 4a–4b.

A significant negative trend in σ^0 occurs only during the third cold season, but by this time the ice feature is exiting through the Fram Strait, and significant divergence, linked to the creation of open water areas, is a reasonable explanation of this evolution. This is confirmed by the clouds of points in the

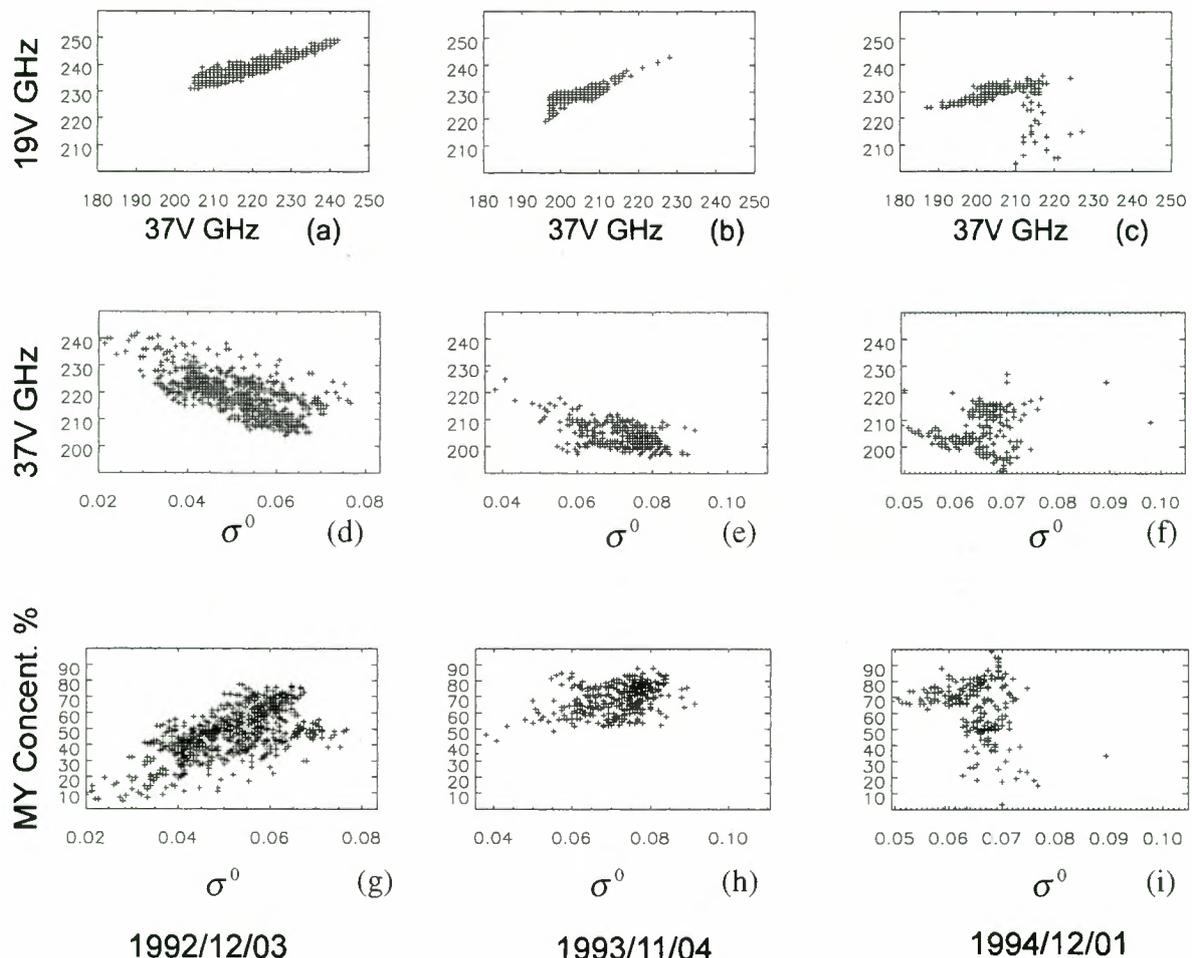


Figure 4. Scatterplots of 19 GHz (V) versus 37 GHz (V), 37 GHz (V) versus σ^0 , and multiyear ice concentration versus σ^0 for the sea ice structure at the beginning of the three cold seasons.

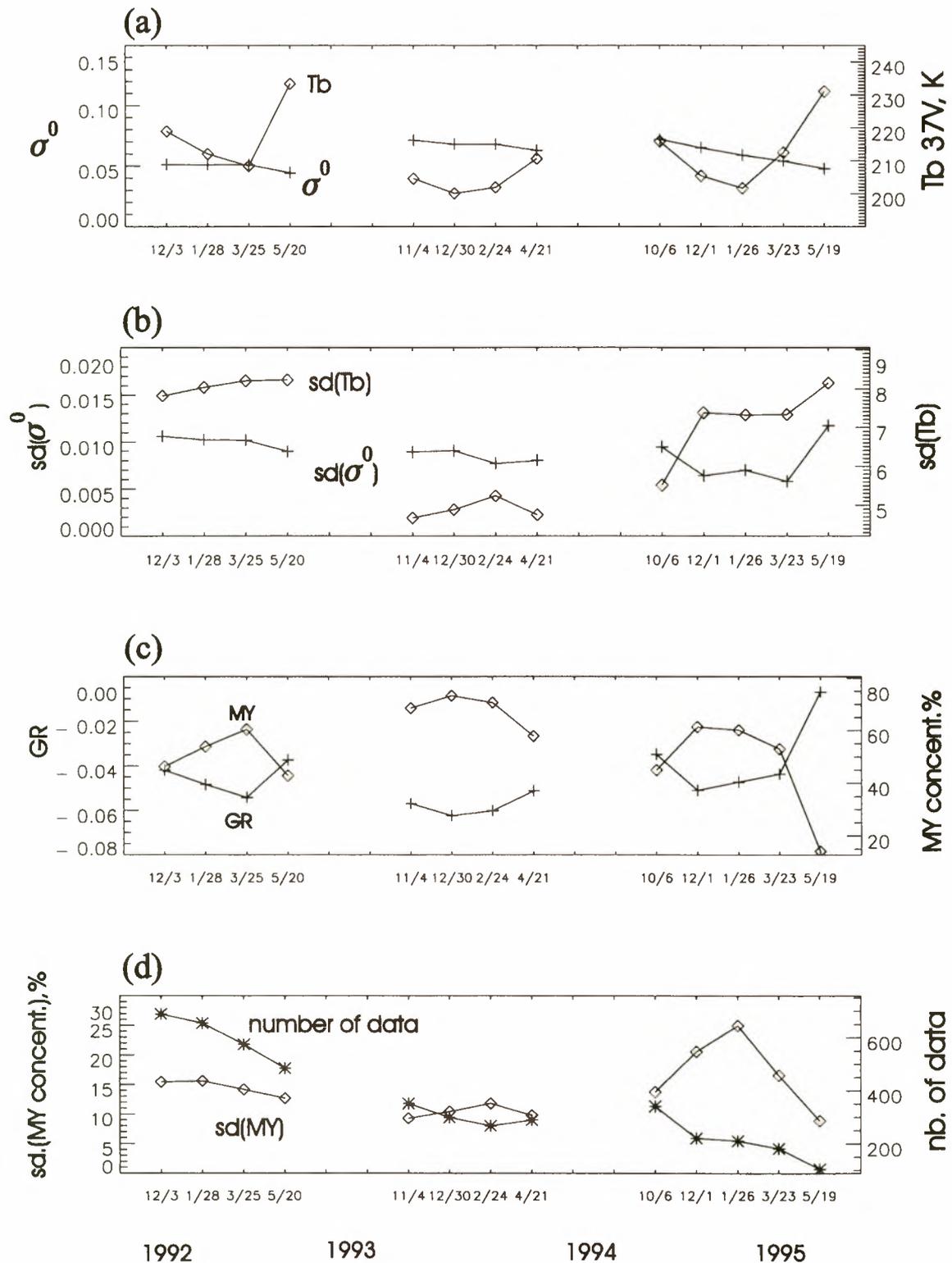


Figure 5. Time evolution of active microwave instrument in wind mode (AMI-wind) backscatter, special sensor microwave imager (SSM/I) brightness temperature and computed parameters over the selected ice structure. (a) Mean backscatter coefficient (cross) and 37-GHz vertically polarized (V) brightness temperature (diamond), (b) standard deviations of backscatter coefficient (cross) and 37-GHz (V) brightness temperature (diamond), (c) gradient ratio (cross) and multiyear ice concentration (diamond), and (d) standard deviation of multiyear ice concentration (diamond) and number of observed ice pixels during the drift (asterisk).

19-GHz (V) versus 37-GHz (V) brightness temperature plots, which depart from the well-known straight line behavior [Comiso, 1986] only during the 1994–1995 winter (Figure 4c).

5.2. Multiyear Ice Concentration

The gradient ratio (Figure 5c) decreased from December 1992 to March 1993 while, inversely, the multiyear ice concentration evolved from 45 to 60%. Because this increase in multiyear ice concentration cannot be justified by advection, it is clearly the result of transformations in the electromagnetic properties of the ice. As noted by Comiso [1990], multiyear ice floes could have first-year ice signatures, especially in zones subject to atmospheric and oceanic forcings; this was the situation encountered in summer 1992 in the area where the ice feature studied was formed. Drainage with time of the saline surface of this second-year ice and the decrease in air temperature due to the displacement toward the north pole and to the seasonal evolution have probably more affected the surface, including the snow cover [Cavaliere *et al.*, 1991] than the interior of the ice. The stability of σ^0 during that period shows that ice is not significantly modified, neither in its surface topography nor in its deeper layers sensed by the scatterometer at C band. As noted previously by several authors in different contexts [Comiso, 1986, 1990; Thomas, 1993; Kwok *et al.*, 1996] but here in a Lagrangian frame, multiyear ice concentrations, as given by the NASA team algorithm, fluctuate considerably in the course of the cold season (Figure 5c). The gradient ratio, far from being constant, is seen to be inversely correlated to the 37-GHz (V) brightness temperature. This important correlation had already been noted but in the context of fixed geographical areas, so that influence of advection could not be eliminated [Wismann *et al.*, 1996].

During the second year of observation, the feature, composed essentially of third-year ice drifting in the vicinity of the north pole, shows a more regular pattern in backscatter and brightness temperature. The multiyear ice concentration is stable and estimated to 75%, which is probably a little low but not unrealistic for third-year ice. The mean level of σ^0 has significantly increased after the summer and stays at a level equivalent to the highest observed during the previous year near the structure. A slow decrease of σ^0 occurs during the third year of observation. This decrease of σ^0 can be attributed to an increase of new ice concentration as ice diverges on approaching the Fram Strait. The multiyear ice concentration reaches 60% in January 1995 before decreasing from March to May, by which time most of the initial structure has passed Fram Strait (Plate 4) and melted or mixed with younger ice.

6. Summary and Conclusion

6.1. Ice Displacement Fields

Over the Arctic Ocean, scattered multiyear ice or contrasted boundaries in σ^0 within second-year ice areas, due to local roughness effects, create contours on backscatter maps that can be utilized for tracking; from year to year, backscatter maps show very different situations, some years being much more favorable for this purpose. Varying atmospheric conditions in summer are crucial as they produce features which will persist through the cold season. In the autumns of 1992 (Plate 3e) and 1996 (Plate 3k), many features of second-year ice appear on the backscatter images. As in 1992, the situation in 1996 can be explained by the atmospheric cyclonic circulation

in summer (Plate 3j), which has modified the usual ice distribution.

To follow the selected ice structure, we have constructed displacement fields of Arctic sea ice based on buoy positions combined with displacements estimated by automatic tracking using successive backscatter maps. Our aim was to follow an identified ice feature and not to provide a general algorithm for estimating the ice displacement over the whole Arctic Ocean in all seasons. For that purpose it would have been absolutely necessary to include wind data [Thorndike and Colony, 1982]. Our procedure is efficient enough to retrieve the ice displacements accurately in the vicinity of buoys and, during the cold season from October to May, at locations where the backscatter maps show high contrasts. This was the case in the vicinity of the ice feature that we selected. The Eulerian displacement fields, applied to the cells of the structure, defined in early December 1992, have allowed a realistic tracking of its displacements, verifiable during the first year.

In the vicinity of the Fram Strait, the method described is ill adapted to study sea ice dynamics because of the rapid sea ice evolution both in space and time. Divergence in this area of accelerated displacements and strong tidal processes on the Yermak Plateau lead to a complex mixing of old, new, and broken first-year ice. In such a region, daily SSM/I 85-GHz brightness temperatures, advanced very high resolution radiometer (AVHRR) observations, when atmospheric conditions are clear, or SAR data [Kergomard *et al.*, 1994; Kornes, 1994] much better describe the situation of the polar pack ice.

6.2. Evolution of Microwave Signatures

The active and passive microwave signatures evolve coherently during the drift of the structure, showing, on average, opposite trends. However, in its earlier stages the backscatter peak is shifted toward the coast, where ice is likely to have been crushed and ridged. During the first cold season the average σ^0 of the structure, which is classified as second-year ice, is constant while the brightness temperatures and the GR show a more erratic behavior; this leads to unrealistic fluctuations in the multiyear ice concentration as evaluated from the NASA team algorithm. We note an increase in multiyear ice concentration from 40 to 65% from December 1992 to March 1993 while convergence seems to be quite limited because of the σ^0 stability. Successive backscatter images show a stretching of the selected ice feature, its northern part drifting slightly faster, in an environment of lower multiyear ice concentration. It is interesting to note that during another cold season and for perennial ice of the Beaufort Sea, Kwok *et al.* [1996] have observed a decrease in the multiyear ice concentration. If the decrease in spring, observed at the end of the cold seasons (May 1993 and April 1994), can be attributed to an increase in atmospheric water vapor, the winter trend is difficult to explain without taking into account a change in emissivity or an effect of the physical temperature as ice is drifting northward.

We agree with the conclusions already published by Thomas [1993] that there are several “types” of multiyear ice, each with its own radiometric signature, itself dependent on the physical temperature which varies spatially. Although there exists a clear relationship between the GR and multiyear ice concentration, it seems to be an impossible task to accurately and coherently estimate the multiyear ice concentration from this parameter alone without taking into account the spatial and temporal continuity of the ice cover from the end of summer to spring.

Acknowledgments. The authors thank the National Snow and Ice Data Center, University of Colorado, for providing the SSM/I data; the Polar Science Center, University of Washington, for the buoy data; and the Centre ERS d'Archivage et de Traitement for processing and providing the AMI-wind backscatter maps. We are particularly grateful to I. G. Rigor (Polar Science Center), who provided the most recent buoy positions available from the IABP. Many comments to the initial draft made by the anonymous reviewers must be acknowledged and have helped to improve the document. This study was partly supported by a 1994 grant of the Programme National de Télédétection Spatiale.

References

- Aagaard, K., and E. C. Carmack, The role of sea ice and other fresh water in Arctic circulation, *J. Geophys. Res.*, **94**(C10), 14,485–14,498, 1989.
- Cavaliere, D. J., Sea ice algorithm, in *NASA Sea Ice Validation Program for the Defense Meteorological Satellite Program Special Microwave Imager: Final Report, NASA Tech. Memo.*, 104559, 25–31, 1992.
- Cavaliere, D. J., J. P. Crawford, M. R. Drinkwater, D. T. Eppler, L. D. Farmer, R. R. Jentz, and C. C. Wackerman, Aircraft active and passive microwave validation of sea ice concentration from the Defense Meteorological Satellite Program special sensor microwave imager, *J. Geophys. Res.*, **96**(C12), 21,989–22,008, 1991.
- Colony, R. L., and I. G. Rigor, International Arctic Buoy Program data report for 1 January 1992–31 December 1992, *Tech. Memo. APL-UW TM29-93*, 215 pp., Appl. Phys. Lab., Univ. of Wash., Seattle, 1993.
- Colony, R., and A. S. Thorndike, An estimate of the mean field of the Arctic sea ice motion, *J. Geophys. Res.*, **89**(C6), 10,623–10,629, 1984.
- Comiso, J. C., Characteristics of Arctic winter sea ice from satellite multispectral microwave observations, *J. Geophys. Res.*, **91**(C1), 975–994, 1986.
- Comiso, J. C., Arctic multiyear ice classification and summer ice cover using passive microwave satellite data, *J. Geophys. Res.*, **95**(C8), 13,411–13,422, 1990.
- Curran, P. J., The semi-variogram in remote sensing, An introduction, *Remote Sens. Environ.*, **24**, 493–507, 1988.
- Dickson, R. R., J. Meincke, S.-A. Malmberg, and A. J. Lee, The 'Great Salinity Anomaly' in the northern North Atlantic 1968–1982, *Prog. Oceanogr.*, **20**, 103–151, 1988.
- Drinkwater, M. R., and F. D. Carsey, Observations of the late-summer to fall sea ice transition with the 14.6 GHz Seasat scatterometer, in *Proceedings of IGARSS'91, Helsinki, Finland*, vol. 3, pp. 1597–1600, Inst. of Electr. and Electron. Eng. Geosci. and Remote Sens. Soc., Spring, Tex., 1991.
- Ezraty, R., F. Gohin, and A. Cavanié, Arctic sea ice during the winter to summer transition as seen by the AMI-Wind and ATSR/M of ERS-1, *EARSel Adv. Remote Sens.*, **3**, 100–111, 1994.
- Fily, M., and D. A. Rothrock, Sea ice tracking by nested correlations, *IEEE Trans. Geosci. Remote Sens.*, **GE-25**(5), 570–580, 1987.
- Gloersen, P., Modulation of hemispheric sea-ice cover by ENSO events, *Nature*, **373**, 503–506, 1995.
- Gloersen, P., W. J. Campbell, D. J. Cavalieri, J. Comiso, C. L. Parkinson, and H. J. Zwally, Arctic and Antarctic sea-ice, 1978–1987: Satellite passive-microwave observations and analysis, *NASA Spec. Publ.*, *NASA SP-511*, 290 pp., 1992.
- Gohin, F., Some active and passive microwave signatures of Antarctica sea ice from mid-winter to spring 1991, *Int. J. Remote Sens.*, **16**, 2031–2054, 1995.
- Gohin, F., and A. Cavanié, A first try at identification of sea ice using the three beam scatterometer of ERS-1, *Int. J. Remote Sens.*, **15**, 1221–1228, 1994.
- Häkkinen, S., An Arctic source for the Great Salinity Anomaly: A simulation of the Arctic ice-ocean system for 1955–1975, *J. Geophys. Res.*, **98**(C9), 16,397–16,410, 1993.
- Haverkamp, D., L. K. Soh, and C. Tsatsoulis, A comprehensive, automated approach to determining sea ice thickness from SAR data, *IEEE Trans. Geosci. Remote Sens.*, **33**(1), 46–57, 1995.
- Kergomard, C., J. C. Gascard, M. Fily, C. Delacourt, C. Lefort, and D. Darr, The transpolar ice drift north of Fram Strait studied from ERS-1 SAR imagery, in *Proceedings of the Second ERS-1 Symposium—Space at the Service of Our Environment*, Eur. Space Agency Spec. Publ., *SP-361*, 263–267, 1994.
- Kornes, R., An ice drift series from the Fram Strait January–March 1992 based on ERS-1 SAR data, *Polar Res.*, **13**(1), 55–58, 1994.
- Kwok, D., A. Rothrock, H. L. Stern, and G. F. Cunningham, Determination of the age distributions of ice motion, *IEEE Trans. Geosci. Remote Sens.*, **33**(2), 392–400, 1995.
- Kwok, D., J. C. Comiso, and G. F. Cunningham, Seasonal characteristics of the perennial ice cover of the Beaufort Sea, *J. Geophys. Res.*, **101**(C12), 28,417–28,439, 1996.
- Matheron, G., La théorie des variables généralisées et ses applications; Ecole Nationale Supérieure des Mines de Paris, *Cah. Cent. Morphol. Math. Fontainebleau*, **5**, 1–212, 1970.
- Olea, R. A., Optimal contour mapping using universal kriging, *J. Geophys. Res.*, **79**(5), 695–701, 1974.
- Serreze, M. C., R. G. Barby, and A. S. McLaren, Seasonal variations in sea ice motion and effects on sea ice concentration in the Canadian Basin, *J. Geophys. Res.*, **94**(C8), 10,955–10,970, 1989.
- Steffen, K., J. Key, J. Cavalieri, J. Comiso, P. Gloersen, K. St. Germain, and I. Rubinstein, The estimation of geophysical parameters using passive microwave algorithms, in *Microwave Remote Sensing of Sea Ice*, *Geophys. Monogr. Ser.*, vol. 68, edited by F. D. Carsey, pp. 201–231, AGU, Washington, D. C., 1992.
- Thomas, D. R., Arctic Sea ice signatures for passive microwave radiometry, *J. Geophys. Res.*, **98**(C6), 10,037–10,052, 1993.
- Thorndike, A. S., and R. Colony, Sea ice motion in response to geostrophic winds, *J. Geophys. Res.*, **87**(C8), 5845–5852, 1982.
- Walsh, J. E., and H. J. Zwally, Multiyear sea ice in the Arctic: Model- and satellite-derived, *J. Geophys. Res.*, **95**(C7), 11,613–11,628, 1990.
- Wismann, V., K. Boehnke, A. Cavanié, R. Ezraty, F. Gohin, D. Hoekman, and I. Woodhouse, Land surface observations using the ERS-1 windsat scatterometer, II, report, contract 11103/94/NL/CN, Eur. Space Agency, Paris, 1996.

A. Cavanié, R. Ezraty, and F. Gohin, IFREMER, Centre de Brest, BP 70 29280 Plouzané, Breizh, France. (e-mail: gohin@ifremer.fr)

(Received January 29, 1997; revised November 2, 1997; accepted November 10, 1997.)

Sea ice motion from satellite passive microwave imagery assessed with ERS SAR and buoy motions

R. Kwok,¹ A. Schweiger,² D. A. Rothrock,² S. Pang,¹ and C. Kottmeier³

Abstract. Observing the motion of sea ice from space is analogous to observing wind stress over the wet oceans; both provide surface forcing for modeling ocean dynamics. Ice motion also directly provides the advective component of the equations governing the mass balance of the sea ice cover. Thus its routine observation from space would be of great value to understanding ice and ocean behavior. To demonstrate the feasibility of creating a global multidecadal ice motion record from satellite passive microwave imagery and to quantitatively assess the errors in the estimated ice motions, we have tracked ice every 3 days in the Arctic Ocean and daily in the Fram Strait and Baffin Bay during the 8 winter months from October 1992 to May 1993 and daily in the Weddell Sea during the 8 winter months from March to October 1992. The method, which has been well used previously, involves finding the spatial offset that maximizes the cross correlation of the brightness temperature fields over 100-km patches in two images separated in time by from 1 to 3 days. The resulting ice motions are compared with contemporaneous buoy- and SAR-derived ice motions. The uncertainties in the displacement vectors, between 5 and 12 km, are better than the spatial resolution of the data. Both 85-GHz data with 12-km spatial resolution and 37-GHz data with 25-km resolution are tracked. These trials with the 37-GHz data are new and show quite surprisingly that the error is only about 1 km larger with these data than with the 12-km 85-GHz data. Errors are typically larger than average in areas of lower ice concentration; in the most dynamic regions, particularly near the ice edge in the Barents and Greenland Seas; and in zones of high shear. These passive microwave ice motions show a large increase in spatial detail over motion fields optimally interpolated from buoy and wind observations, especially where buoy data are virtually absent such as near coasts and in some passages between the Arctic Ocean and its peripheral seas. The feasibility of obtaining ice motion from the 37-GHz data in addition to the 85-GHz data should allow an important record of ice motion to be established for the duration of the scanning multichannel microwave radiometer (SMMR), special sensor microwave/imager (SSM/I), and future microwave sensors, that is, from 1978 into the next millenium.

1. Introduction

The circulation of sea ice determines the advective part of the ice balance and provides a velocity boundary condition on the ocean surface. The spatial gradient of this circulation, ice deformation, controls the abundance of thin ice and therefore the many surface processes dependent on thin ice, such as turbulent heat flux to the atmosphere and salt flux into the ocean. Convergent motion increases the local ice cover mass by rafting and ridging. Sea ice motion is readily observed in satellite imagery, the scale of the observed motion being dependent on the spatial resolution of this imagery. In recent years a number of procedures [Fily and Rothrock, 1987; Kwok *et al.*, 1990; Emery *et al.*, 1991] have been developed to recognize common features in various types of sequential images to ob-

tain displacement measurements. The quality of these measurements depends more on geometric fidelity and image resolution than on a thorough physical understanding of the ice signatures themselves. Imagery acquired by synthetic aperture radars (SAR) and the advanced very high resolution radiometer (AVHRR) are of particular importance for motion observations. The European Remote Sensing Satellites (ERS-1 and ERS-2), the Japanese Earth Resources Satellite (JERS-1), and the Canadian RADARSAT satellite all carry SARs; successive high-resolution SAR imagery ($\approx 10\text{--}100$ m) can provide almost all-weather, diurnal observations of ice motion. Prior to RADARSAT, large-scale motion measurements were limited by the narrow SAR swath widths, about 100 km. The wide-swath mode (500 km) of RADARSAT provides routine mapping of the Arctic in <3 days. To date, the SAR data records of the polar oceans, since the launch of ERS-1 in 1991, are still fairly sparse because of limitations in swath width, orbit geometry, and lack of onboard tape recorders on the ERS satellites. AVHRR imagery provides wide-swath observations (1000 km) at moderate resolution (≈ 1 km), but there are spatial gaps in the parts of the ice cover obscured by clouds [Emery *et al.*, 1991]. The AVHRR record provides more extensive coverage of the polar regions and dates back to 1981.

It has recently been demonstrated that despite its antenna

¹Jet Propulsion Laboratory, California Institute of Technology, Pasadena.

²Polar Science Center, Applied Physics Laboratory, College of Ocean and Fisheries Sciences, University of Washington, Seattle.

³Universität Karlsruhe Institut für Meteorologie und Klimaforschung, Karlsruhe, Germany.

Copyright 1998 by the American Geophysical Union.

Paper number 97JC03334.
0148-0227/98/97JC-03334\$09.00

Table 1. EFOV of the Relevant SSM/I Radiometer Channels

Frequency, GHz	EFOV, km
37.0	38 × 30
85.5	16 × 14

EFOV, effective field of view; SSM/I, special sensor microwave/imager.

footprints of 10 or more kilometers, sequential imagery from 85-GHz special sensor microwave/imager (SSM/I) can provide ice motion observations [Agnew *et al.*, 1997; Liu *et al.*, 1998]. We demonstrate here that one can obtain ice motion from the 37-GHz data as well. This offers an extraordinary opportunity to create an ice motion data record starting in 1978 with the scanning multichannel microwave radiometer (SMMR), which had a 37-GHz channel but no 85-GHz channel and continuing through the present and indefinitely into the future with planned SSM/I and other passive microwave radiometers. These measurements would be complementary to the higher-resolution satellite observations described above.

The objective of this paper is to provide a quantitative assessment of the ice motion observations obtained with 85-GHz SSM/I and with 37-GHz SMMR data. The utility of these measurements depends on an understanding of the errors in displacement estimates. Our approach is to characterize these uncertainties by comparing the derived motion with available buoy and ERS SAR-derived ice motions. In the following section the ice motion tracking procedure used to obtain the displacement estimates is described, including the temporal sampling issues of ice motion and the filtering of outliers. In section 3 we compare 8 months of ice motion from the Arctic and 8 months of ice motion from the Antarctic with buoy and SAR ice motions. In section 4 we examine spatial features visible in bimonthly mean drift patterns computed from the results of section 3. One of the best long-term records of Arctic ice motion comes from optimally interpolated buoy motions augmented by surface geostrophic wind fields [Rigor and Heiberg, 1997]. Section 4 also shows a comparison between these buoy-wind-derived motion fields and those from SSM/I. The latter show interesting spatial details missing in the smoothed buoy/wind fields. Summary remarks are given in section 5.

2. Ice Motion Tracking

2.1. Description of Tracker

The inputs to the ice motion tracker are sequential SSM/I brightness temperature fields (obtained from the National Snow and Ice Data Center) mapped onto rectangular grids in a polar stereographic projection. The SSM/I instrument has a swath width of 1394 km, and observations are made at an angle of incidence of $\sim 53^\circ$ at the Earth's surface. The gridded daily brightness temperature fields are synthesized from data collected from ascending and descending orbits. The 37-GHz daily averages are binned, using a nearest neighbor scheme, to square pixels with dimensions of 25×25 km and the higher-resolution 85-GHz observations are binned to 12.5×12.5 -km-size pixels. Brightness temperatures observed over a 24-hour period (midnight to midnight UT) are summed then divided by the total number of observations to obtain an average brightness temperature at that pixel. Regions of open

ocean are masked out using analyzed fields of ice concentration derived from the same passive microwave observations. Missing data values are replaced by the median value in the 3×3 -pixel neighborhood of that pixel. Sea ice displacements are obtained by a motion tracker that identifies common features in the successive brightness temperature fields. The procedure used here is based on the algorithm described by Kwok *et al.* [1990]. This routine was originally developed for production of ice displacement fields from SAR imagery. Only an overview of this tracking approach is given here. Instead, we provide a more detailed account of issues associated with tracking ice motion in passive microwave data.

In the motion tracker the magnitude of the normalized cross-correlation coefficient is used as the measure of similarity between features in the passive microwave data sets. The method, which has been well used previously, involves finding the spatial offset that maximizes the cross-correlation coefficient of the brightness temperature fields of features in two images separated in time. Features are $n \times n$ pixels of brightness temperature arrays that we extract from an image. Starting with an approximate location on the second image, we compute the correlation coefficient between an array from the first image and an array of the same size on the second image. This correlation value is recorded. The computation is repeated at each position as the array from the first image is shifted on a two-dimensional grid to obtain an array of correlation coefficients. The peak of this sampled surface is regarded as the location of maximum similarity between the arrays from the two images. The displacement vector is then given by the difference in geographic location of the centers of the two arrays. This procedure is repeated for each feature extracted from the first image. The process involves two different steps which utilize this image matching procedure: the first step is performed at the nominal resolution to find the approximate location of these features and the second step focuses on refinement of the location of these features in the oversampled data set (discussed below). Fast Fourier transforms are used to improve the efficiencies in the computation of these correlation surfaces. After each step, filters are used to discard outliers and inconsistent displacement vectors based on the statistics of the displacement field. Details of these filter can be found in work by Kwok and Cunningham [1993]. After the outliers are removed, there are holes in the output field. We do not interpolate to fill them. Below, we describe three modifications to this process for measuring displacements in the passive microwave data.

2.1.1. Location of correlation peak. Oversampling of the correlation surface could be used to provide subpixel location of the correlation peak. If the input data are optimally sampled, we should be able to estimate the peak location to better than the nominal spatial resolution of the sensor (see Table 1). An alternative to oversampling the correlation surface is oversampling the input data, and this is the more efficient approach when the data volumes are small as is the case here. Either approach should give the same results. In our implementation we use an oversampling factor of 6 for the input data, resulting in a sample pixel spacing of ~ 2.1 km for the 85-GHz data and 4.2 km for the 37-GHz data. With this oversampling factor, the uncertainty in the location of the peak of the correlation surface is determined by the pixel spacing of the oversampled data. Even with an oversampling factor of 6, the quantization noise of several kilometers is still significant for each vector observation. To reduce the uncertainty due to sampling, the

Arctic 930408_930411 85V(black) buoy(white) motion vectors

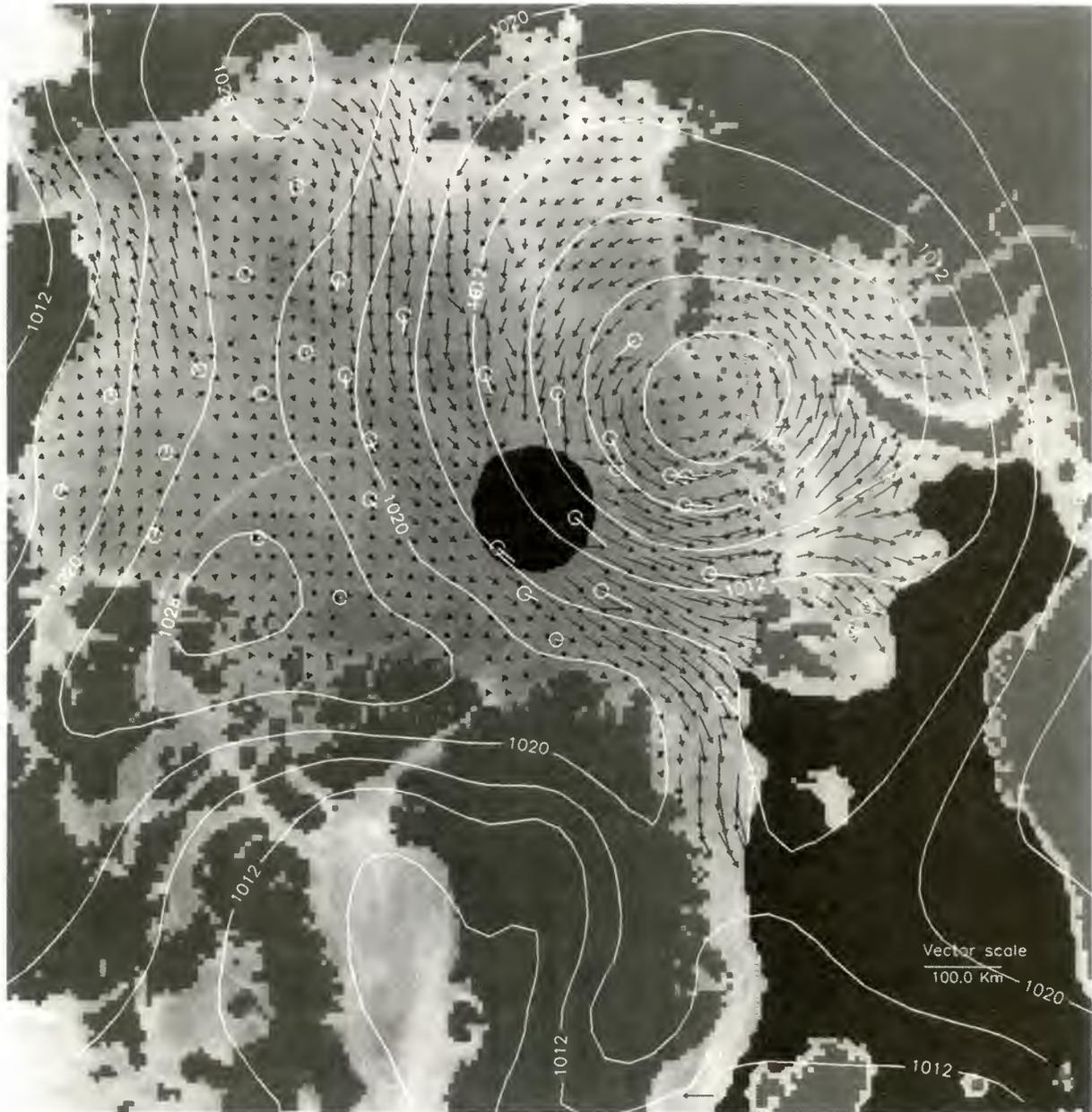


Figure 1. Three-day ice motion in the Arctic Basin from 85-GHz passive microwave data. Contours are isobars overlaid on the motion field. Three-day buoy motions are shown in white. The pressure contour interval is 4 mbar. The grid spacing is 80 km.

subpixel location of the peak is estimated using the eight neighborhood correlation values around the peak of the sampled correlation surface. A biquadratic surface of the form $f(x, y) = \sum_{i=0}^8 a_i x^i y^j$ is fit to the nine correlation values; the maxima of the function are then computed using the method of steepest descent. This effectively reduces errors caused by spatial sampling of the correlation surface. This combination of oversampling factor and interpolation is a compromise between data volume and computational throughput.

2.1.2. Weather/surface change filter. The normalized correlation coefficient as a measure of similarity is insensitive to the absolute brightness levels of the arrays or features being compared. Both the 85- and 37-GHz channel observations are affected by clouds, atmospheric water vapor content, and

changes in surface conditions (such as snow properties), and fairly significant changes in brightness temperature levels are associated with some of these phenomena. To avoid the detection of a false match between features using solely the normalized correlation coefficient, a match between features has to satisfy an additional condition that the difference between the mean brightness temperature levels of the two arrays is $<10\%$ of the observed brightness temperature. This threshold, which is a rather arbitrary choice, seems to be effective in removing areas that are dissimilar in brightness temperature or contaminated by atmospheric emissions. We find that this filter typically removes one to two observations per vector field. At times when there is severe weather, especially in the Weddell Sea, it could remove erroneous motion estimates from a large region.

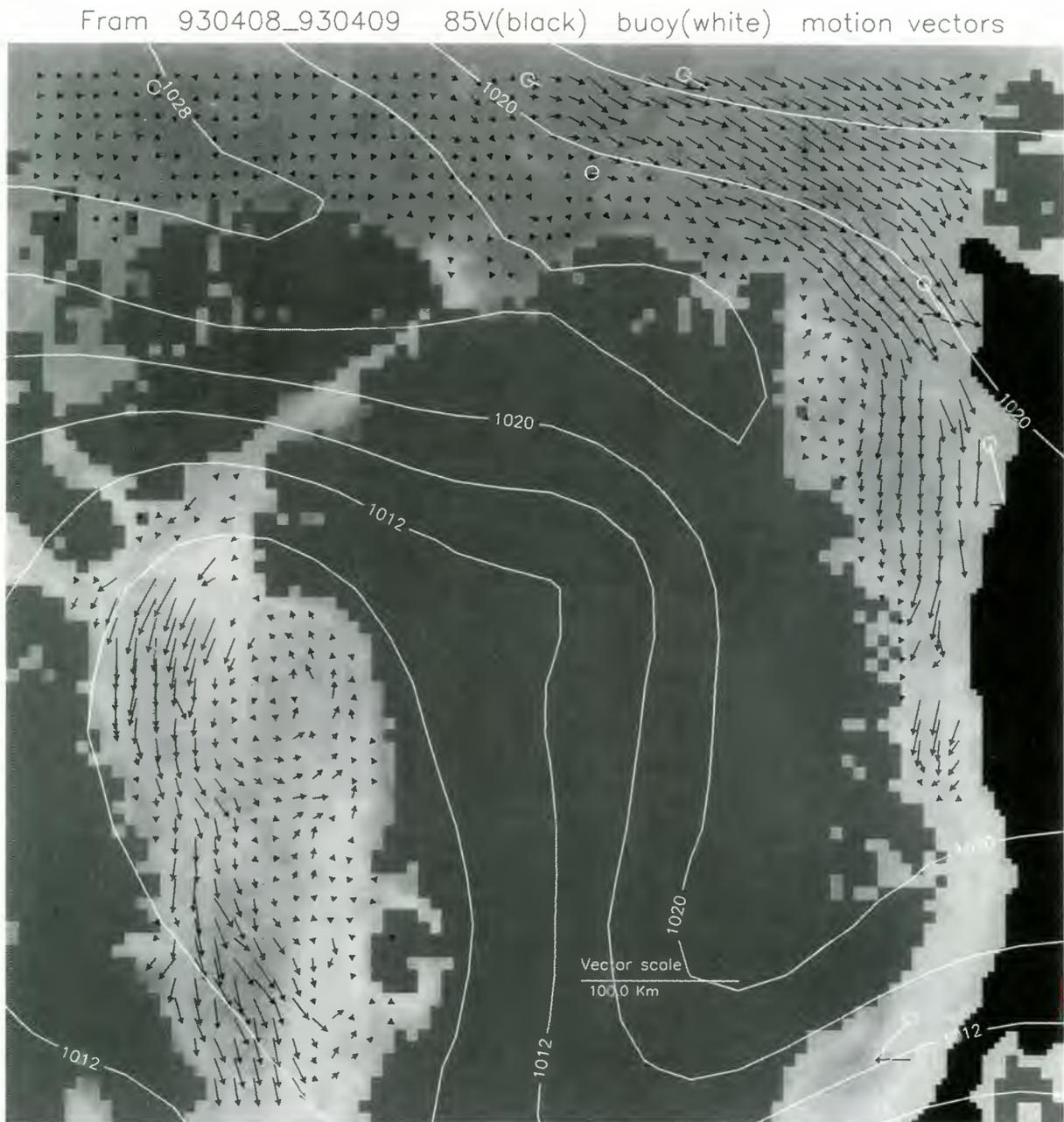


Figure 2. One-day ice motion in the Fram Strait and Baffin Bay and surroundings from 85-GHz passive microwave data. Contours are isobars overlaid on the motion field. One-day buoy motions are shown in white. The pressure contour interval is 4 mbar. The grid spacing is 40 km.

2.1.3. Filtering of outliers. The output displacement vector fields are further filtered using the wind field. We use the wind direction to constrain the direction of the output vectors. Displacement vectors that are not within 90° of the wind direction are discarded. The tolerance of this filter is set rather high because of uncertainty in the quality of the wind fields and because ice motion is not entirely wind driven. We use the reanalyzed wind fields from the National Center for Environment Prediction for this purpose.

2.2. Sampling Issues

Although the SSM/I products provide daily maps of the polar regions, the coarse resolution of the sensor limits the observability of displacements that are comparable to the foot-

print (Table 1). In other words, the resultant field would be quite noisy if we attempt to derive daily ice motion fields in the winter Beaufort Sea where the climatologically mean sea ice motion is between 2 and 5 km/d, small compared to the 12–15-km sensor resolution. Lengthening the time separation would enlarge the signal (displacement relative to the noise, resolution). Another consideration, however, is the change in the passive microwave signature of the features being tracked. The temporal correlation of the data tend to decrease as the areal fraction of sea ice types change with time due to advection, opening/closing, and ice growth. The determination of the temporal sampling interval of the motion field should be based on a balance of these factors.

Weddell 920623_920624 85V(black) buoy(white) motion vectors

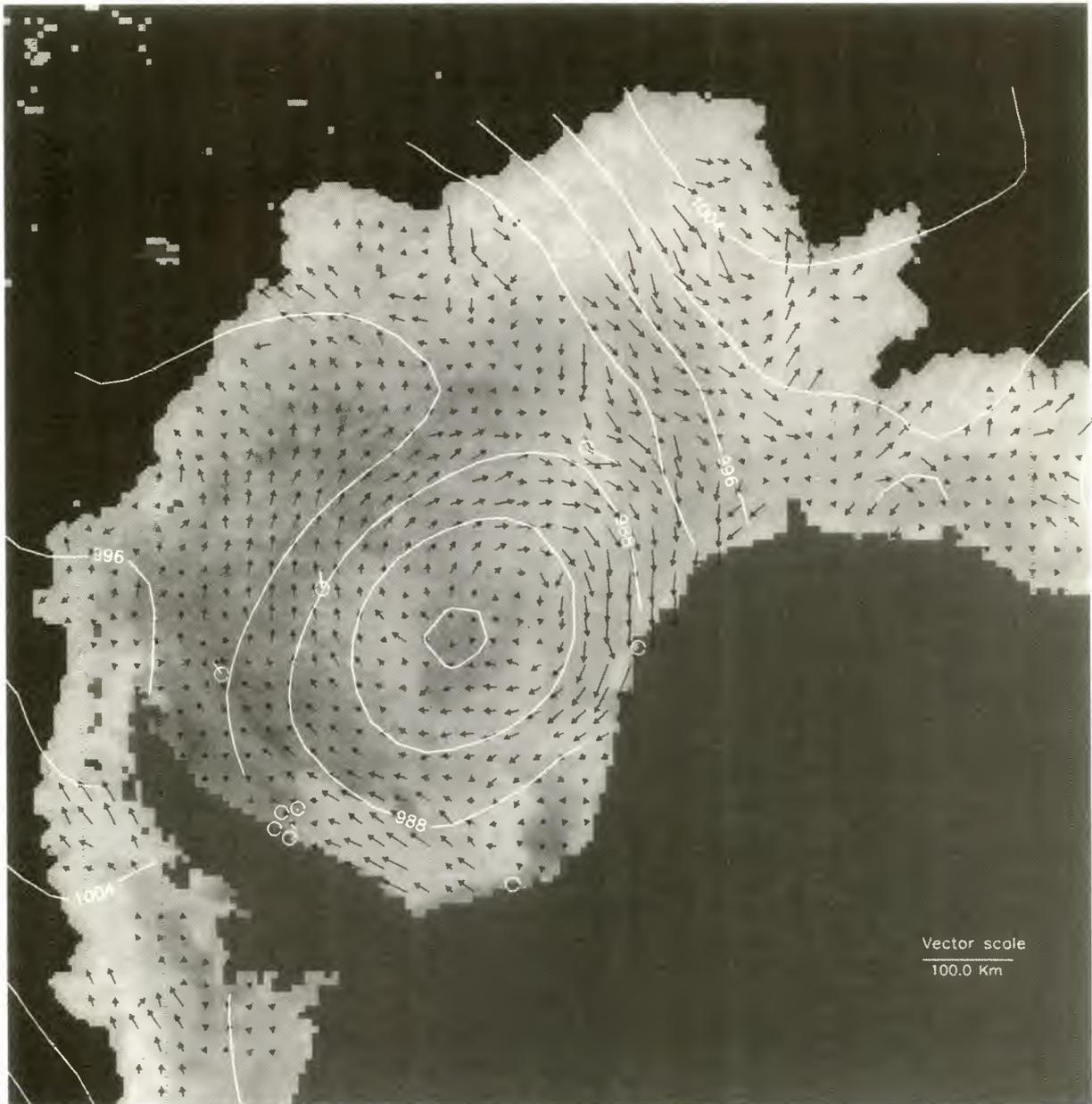


Figure 3. One-day ice motion in the Weddell Sea from 85-GHz passive microwave data. Contours are isobars overlaid on the motion field. One-day buoy motions are shown in white. The pressure contour interval is 4 mbar. The grid spacing is 80 km.

Our approach is to subset the daily SSM/I products into different regions such that the sampling interval of these displacement fields are based appropriately on the expected motion of sea ice in these regions. In the comparisons and examples that follow we use 1- and 3-day sampling intervals for ice motion in two regions of the Arctic and one in the Weddell Sea.

2.3. Summer Ice Motion

Summer passive microwave data are typically not suitable for ice motion observations because of the opacity of the atmosphere during these months (June–September in the Arctic and November–February in the Antarctic). The observed 85- and 37-GHz brightness temperatures tend to have large con-

tributions from the atmospheric emissions due to the increase in water vapor and cloud liquid water content in the summer. Melt/freeze cycles also increase the variability of surface emissivity. The derived ice motion, on the basis of our approach, tends to be unreliable between the onset of melt in the spring and freeze in the fall. This introduces a gap in the annual ice motion record which needs to be filled by some other means. In the following section we restrict our attention to passive microwave ice motion estimates from the winter months.

2.4. Regional Motion Fields

We select three regional subsets of data from the gridded brightness temperature fields for our illustration here and for our comparative analysis in the next section. The first region

Table 2. Comparison of 85-GHz Ice Motion With Buoy Motion for October 1992 to May 1993 (Arctic) and March 1992 to October 1992 (Antarctic)

Region	Channel, GHz	Interval, days	Δ , km	e_{x_s} , km	e_{y_s} , km	θ , km	Scale	ρ	Number of Points
<i>All Points</i>									
Arctic	85 V	3	-1.0 (11.1)	1.1 (8.9)	0.8 (10.8)	0.5 (46.5)	6.6 (47.1)	0.76	2119
	85 H	3	-0.9 (11.7)	1.2 (9.9)	1.2 (11.2)	3.0 (50.4)	6.7 (49.2)	0.72	2097
Fram Strait/Baffin Bay	85 V	1	0.5 (5.3)	0.3 (4.3)	0.2 (5.2)	-2.7 (65.4)	4.2 (9.5)	0.76	1054
	85 H	1	0.3 (6.0)	0.4 (4.7)	0.8 (6.0)	1.7 (66.7)	4.6 (13.3)	0.67	1028
Weddell Sea	85 V	1	2.1 (7.8)	0.6 (7.4)	0.1 (8.2)	1.7 (59.6)	2.5 (7.8)	0.67	1553
<i>Excluding Points With Small Displacements</i>									
Arctic	85 V	3	-1.2 (6.7)	0.9 (7.3)	0.7 (7.8)	1.7 (30.6)	1.0 (0.4)	0.86	1440
	85 H	3	-1.1 (7.5)	0.6 (8.5)	0.9 (7.9)	3.2 (32.8)	1.0 (0.4)	0.83	1419
Fram Strait/Baffin Bay	85 V	1	0.4 (3.9)	0.3 (4.7)	-0.4 (5.0)	2.8 (29.8)	1.1 (0.4)	0.84	222
	85 H	1	0.1 (4.0)	0.0 (4.8)	0.0 (5.1)	-4.2 (31.5)	1.0 (0.4)	0.83	205
Weddell Sea	85 V	1	-0.6 (6.7)	-0.1 (8.3)	0.4 (8.8)	2.6 (44.0)	1.1 (0.5)	0.78	719

In each set of numbers the first is the mean, and the number in parentheses is the standard deviation. V is vertically polarized, and H is horizontally polarized.

encompasses the entire Arctic Basin, which we sample at 3-day intervals. The second region covers Fram Strait and Baffin Bay on the east and west coasts of Greenland. The last region is the Weddell Sea. The ice motion in the latter two regions are characterized by large ice displacements and deformation with active ice growth/ablation in the winter. We use a 1-day sampling period in these regions. Examples of the 1- and 3-day 85-GHz displacement fields from these regions are shown in Figures 1, 2, and 3. Note that the density of observations obtained is still quite high after the filtering process. For comparison, buoy displacement vectors are shown along with the observed displacement vectors. The agreement between the displacement observations is quite remarkable; quantitative evaluation of these fields is provided in the next section. Overlay of the surface pressure field over the derived motion field also shows the circulation pattern to be generally tangent to the isobars. We note here that all the vectors in the motion field were generated by the tracking process, and no attempt was made to fill in these holes by interpolation; there are grid points where the tracker did not provide a motion observation.

3. Error Analysis

In order to quantify the uncertainties in the ice motion from passive microwave data and to understand the utility of these measurements, we compare the motion results with those derived from drifting buoys and the analysis of sequential SAR images. Motion information from buoys, available at 3-hour intervals are used for comparison purposes because the higher sampling rate allows us to better match the observation times of the passive microwave data set without resorting to temporal interpolation of the position measurements. SAR-derived ice motion is used in areas where no or few buoys are available. This comparison study is done with SSM/I motion results from the Arctic and Antarctic winters of 1992 and 1993. We also compare the ice motion results obtained using the 37-GHz and the 85-GHz channel data sets. The results show that we can reliably obtain ice motion, albeit at a slightly higher level of uncertainty, from the 37-GHz brightness temperature fields. The three regions used in our analysis, described in the previous section, have been selected because they have different motion characteristics and buoy and SAR motions are available for comparisons.

3.1. Source of SAR and Buoy Motion

The SAR ice motion data set used here is obtained by tracking common features in sequential ERS-1 radar imagery using a tracking approach similar to the one described here. We compiled SAR motion measurements from two regions of the Arctic: Beaufort Sea and Fram Strait. The gridded SAR ice motion data from the Beaufort Sea are produced by the Geophysical Processor System [Kwok *et al.*, 1990] at the Alaska SAR Facility. The ERS SAR motions are typically produced from images separated by 3 days. In total, 200 3-day displacement fields on a 5-km grid covering an area of $\sim 100 \times 100$ km were available during this period. The Alaska SAR Facility reception mask, however, does not cover the Fram Strait region, and no equivalent facility providing ice motion processing exists for data outside its coverage area.

SAR ice motions for the Fram Strait region are computed from images obtained from the European Space Agency. Low-resolution quicklook images and high-resolution image products were reformatted to allow the computation of displacements by a copy of the Geophysical Processor System located at the Jet Propulsion Laboratory. Displacement errors are estimated to be < 300 m. Approximately 1000 image pairs at time intervals varying from 1 to 7 days are obtained for this region. However, the whole data set is not used here because some motion pairs did not have suitable start times or correct time separations between observations.

Buoy data are obtained from two different sources. Arctic data were provided by the International Arctic Buoy Program (IABP). The buoy positions at 3-hour intervals are interpolated from positions sampled at 1–2-hour intervals by fitting a polynomial to the time series data. The position accuracy is estimated to be ≈ 350 m [Thomdike and Colony, 1980]. Buoy data for the Weddell Sea were collected and preprocessed by the Alfred Wegener Institute (AWI). Some of the buoys are equipped with Global Positioning System (GPS) receivers. Errors in positions are ~ 350 m for the buoys using the ARGOS positioning system and better than 50 m for those using the GPS system [Kottmeier *et al.*, 1992].

3.2. Comparison With Buoy/SAR Ice Motion

Though spatially sparse, ice motions from the IABP buoys are the best measurements for assessing the quality of passive microwave motion because of its continuous temporal cover-

Table 3. Comparison of 37-GHz Ice Motion With Buoy Motion for October 1992 to May 1993

Region	Channel, GHz	Interval, days	Δ , km	e_x , km	e_y , km	θ , km	Scale	ρ	Number of Points
<i>All Points</i>									
Arctic	37 V	3	-0.4 (10.6)	3.7 (8.7)	0.2 (9.2)	-2.7 (49.1)	4.4 (19.8)	0.79	2242
	37 H	3	0.6 (11.3)	3.7 (9.8)	0.5 (10.1)	-1.9 (51.6)	4.4 (22.1)	0.77	2181
Fram Strait/Baffin Bay	37 V	1	-1.7 (7.6)	1.5 (5.1)	2.1 (6.9)	-2.1 (61.4)	2.5 (6.2)	0.61	590
	37 H	1	-1.9 (8.1)	1.4 (5.3)	2.4 (7.1)	3.4 (62.5)	2.7 (6.0)	0.57	573
<i>Excluding Points With Small Displacements</i>									
Arctic	37 V	3	-0.3 (7.0)	3.7 (7.9)	-0.4 (7.6)	-0.8 (25.2)	1.0 (0.3)	0.89	842
	37 H	3	0.2 (7.7)	3.5 (8.7)	0.0 (8.5)	-0.3 (28.2)	1.1 (0.4)	0.87	853
Fram Strait/Baffin Bay	37 V	1	-0.2 (5.5)	1.2 (5.6)	0.2 (5.8)	-2.0 (34.9)	1.1 (0.5)	0.77	105
	37 H	1	0.0 (5.2)	0.7 (5.5)	0.3 (5.7)	2.0 (36.3)	1.1 (0.5)	0.78	106

In each set of numbers the first is the mean, and the number in parentheses is the standard deviation.

age. To date, SAR-derived ice motions are dependent on available reception facilities, and satellite coverage of the Arctic and Antarctic are limited to certain regions and seasons. Results from the comparison of 8 months of Arctic and Antarctic SSM/I ice motions with buoy and SAR ice motions are shown in Tables 2, 3, and 4. Scatterplots of the passive microwave and buoy motions are shown in Figures 4 and 5. In Tables 2-4 we measure the quality of the displacement estimates with the following quantities:

$$\begin{aligned}
 \mathbf{e} &= \mathbf{u}_{SSM/I} - \mathbf{u}_{ref} \\
 \Delta &= |\mathbf{u}_{SSM/I}| - |\mathbf{u}_{ref}| \\
 \theta &= \arccos \frac{\mathbf{u}_{SSM/I} \cdot \mathbf{u}_{ref}}{|\mathbf{u}_{SSM/I}| |\mathbf{u}_{ref}|} \\
 \rho &= \frac{\sum \mathbf{u}_{SSM/I} \cdot \mathbf{u}_{ref}}{\sum |\mathbf{u}_{SSM/I}| |\mathbf{u}_{ref}|} \\
 scale &= |\mathbf{u}_{SSM/I}| / |\mathbf{u}_{ref}|
 \end{aligned}$$

where \mathbf{e} is the vector difference between $\mathbf{u}_{SSM/I}$ and \mathbf{u}_{ref} ; where ref denotes either buoy or SAR motion data. Here e_x and e_y are the components of \mathbf{e} in the two orthogonal directions defined by the SSM/I polar stereographic map projection; the abscissa and ordinate are defined by the 45° and 135°E meridians, respectively. The angular or directional difference between the two vectors is θ ; ρ is the correlation between two displacement data sets, and scale is the ratio of the displacement magnitudes of the motion estimates. The standard deviations of the differences in each case are measures of uncer-

tainty in the derived displacement vectors. Each quantity measures a different aspect of the displacement estimates.

The dominant sources of error in the measurement of motion are errors in geolocation of the pixels, σ_g (standard deviation); location errors due to binning of the data, σ_b ; and errors in ice motion tracking, σ_f . Assuming that the errors are normally distributed, unbiased, and uncorrelated, the expected errors in the displacement estimates, σ_u , are given by [Holt et al., 1992]

$$\sigma_u^2 = 2(\sigma_g^2 + \sigma_b^2) + \sigma_f^2$$

The expected magnitude of these and other errors are shown in Table 5. Using these values, σ_u is then ~16 and 25 km for the 85- and 37-GHz motion estimates, respectively. If any of these errors are correlated, the displacement uncertainties would be reduced. We also note a shortcoming of the input data used in motion tracking. Errors in the motion estimates are introduced by the manner in which current gridded brightness temperature data sets are constructed; ascending and descending observations at different times are binned and averaged to create the gridded field. Brightness temperatures observed over a 24-hour period (midnight to midnight UT) are summed then divided by the total number of observations to obtain an average brightness temperature at a pixel. So features are spatially distorted, and these fields of pixels are not associated with a distinct instant in time. This creates a spatial and temporal smearing or blurring of the brightness temperature fields, which is dependent on the magnitude of the local ice motion. Accounting for the contribution these factors, we

Table 4. Comparison of SSM/I Ice Motion with SAR Motion for October 1992 to May 1993

Region	Channel, GHz	Interval, days	Δ , km	e_x , km	e_y , km	θ , km	Scale	ρ	Number of Points
<i>All Points</i>									
Arctic	85	3	-0.9 (7.3)	1.9 (11.1)	-0.8 (7.1)	1.2 (44.2)	1.5 (3.7)	0.74	108
	37	3	0.0 (7.3)	4.0 (12.3)	-2.0 (7.7)	3.6 (46.9)	1.3 (2.0)	0.75	122
Fram Strait/Baffin Bay	85	1	0.3 (4.8)	-0.2 (4.7)	0.2 (5.5)	-2.8 (43.0)	1.1 (0.6)	0.84	105
	37	1	0.1 (5.3)	-0.1 (3.9)	0.6 (6.4)	-0.7 (50.2)	1.2 (0.7)	0.82	92
<i>Excluding Points With Small Displacements</i>									
Arctic	85	3	-1.8 (6.1)	1.9 (5.9)	-1.2 (6.1)	2.5 (15.3)	1.0 (0.3)	0.93	81
	37	3	-0.3 (6.4)	4.0 (6.1)	-2.1 (6.5)	3.9 (18.1)	1.0 (0.3)	0.92	68
Fram Strait/Baffin Bay	85	1	0.8 (3.9)	0.1 (3.7)	-0.1 (5.0)	-6.0 (25.9)	1.1 (0.4)	0.86	45
	37	1	0.5 (3.4)	-0.0 (2.7)	-0.1 (5.4)	-7.3 (24.3)	1.1 (0.3)	0.87	39

In each set of numbers the first is the mean, and the number in parentheses is the standard deviation.

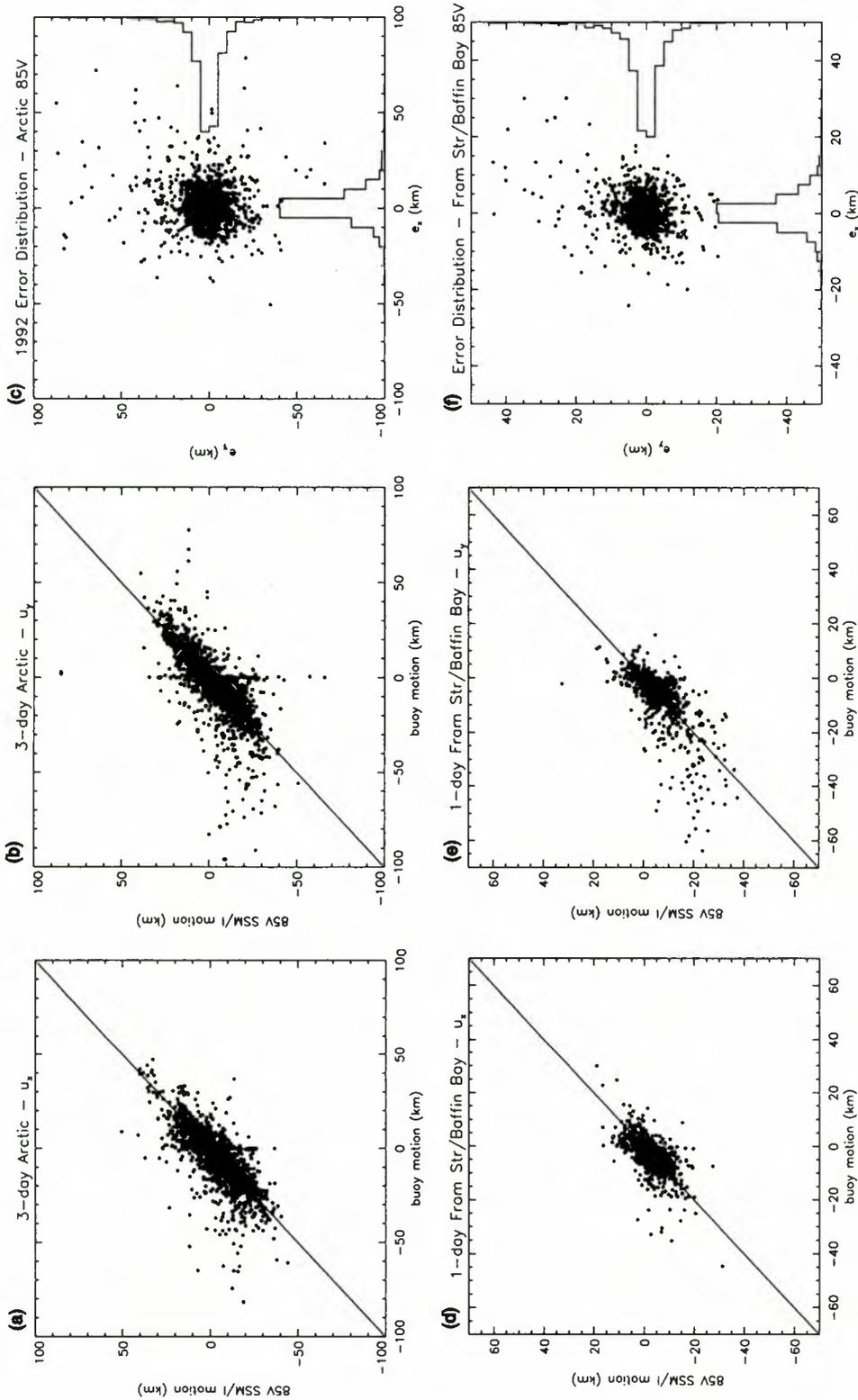


Figure 4. Comparison of the 85-GHz ice motion with buoy displacements. Scatterplot of u_x, u_y and error distributions e_x, e_y . Here e_x and e_y are the components of $\mathbf{e}(\mathbf{u}_{SSM/I} - \mathbf{u}_{buoy})$ in the two orthogonal directions defined by the special sensor microwave/imager (SSM/I) polar stereographic map projection. (a) Scatterplot of $u_x^{SSM/I}$ versus u_x^{buoy} (3-day Arctic), (b) scatterplot of $u_y^{SSM/I}$ versus u_y^{buoy} (3-day Arctic), (c) plot of e_x versus e_y (3-day Arctic), (d) scatterplot of $u_x^{SSM/I}$ versus u_x^{buoy} (1-day Fram Strait/Baffin Bay), (e) scatterplot of $u_y^{SSM/I}$ versus u_y^{buoy} (1-day Fram Strait/Baffin Bay), (f) plot of e_x versus e_y (1-day Fram Strait/Baffin Bay), (g) scatterplot of $u_x^{SSM/I}$ versus u_x^{buoy} (1-day Weddell Sea), (h) scatterplot of $u_y^{SSM/I}$ versus u_y^{buoy} (1-day Weddell Sea), and (i) plot of e_x versus e_y (1-day Weddell Sea).

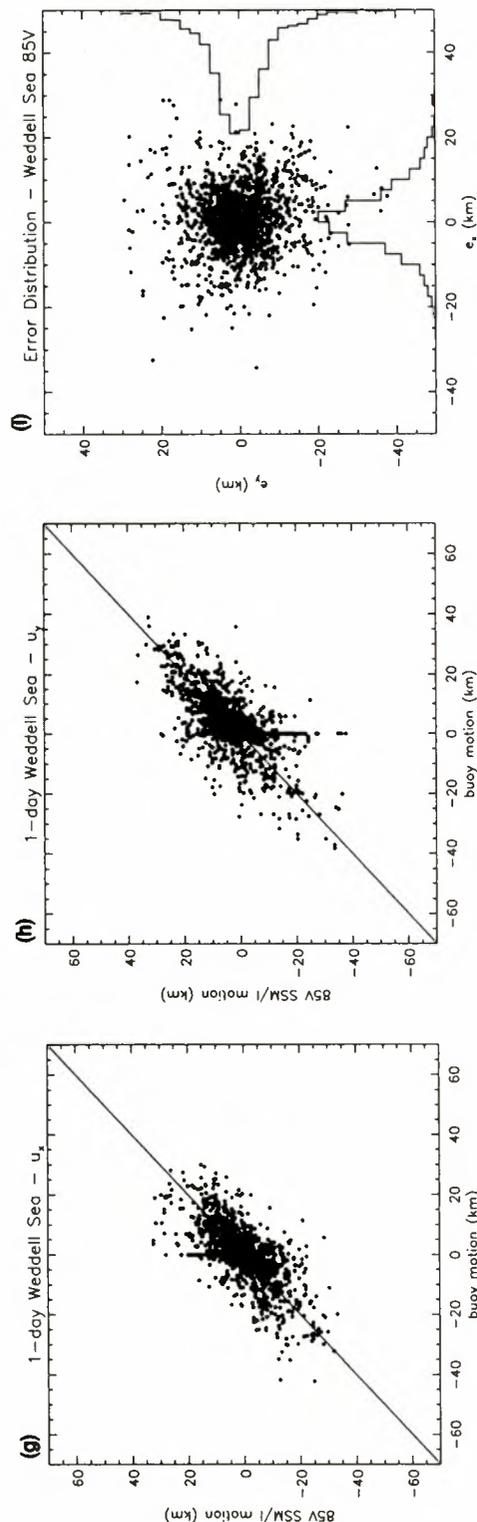


Figure 4. (continued)

expect displacement errors observed here to be larger than our estimate of σ_u above.

Here our measurement process also contributes additional uncertainties to the observation of σ_u . Since we use buoy/SAR displacements as truth, errors in these quantities add to the uncertainty of σ_u . However, the errors in the SAR and buoy ice motions by themselves are relatively insignificant compared to the sensor resolution, location, and tracking errors. There are measurement errors, however, due to the spatial and temporal misregistration of the passive microwave displacement vectors and the buoy/SAR motions. In our comparison we select grid points that are spatially (within 40 km) and temporally closest to available buoy/SAR observations for comparisons. We do not spatially or temporally interpolate the data sets to collocate the observations. Hence the measurements could be up to 12 hours (see discussion above) out of phase in time and/or separated by a distance of 40 km. Any temporal misregistration introduces variability into the comparison especially since sea ice motion on the average decorrelates fairly quickly with time, 1 day [Thorndike, 1984]. Variability could also be caused by the difference between the meaning of displacement between SSM/I motion measurements and buoy/SAR measurements. Motion derived from the passive microwave data are areal averages of displacements over the array of points ($\sim 100 \times 100$ km and 200×200 km in the 85- and 37-GHz data, respectively) used in the tracking procedure while buoy displacements are computed using sequential point position measurements on the ice cover. We choose not to interpolate the passive microwave motion to the buoy locations because the observations already represent the mean motion over a large area on the ice cover. Decorrelation of the buoy measurements and passive microwave measurements is expected if there are large local velocity gradients. Altogether, the measurement of σ_u is contaminated by the errors discussed here.

The results in Tables 2, 3, and 4 indicate that the magnitude of the uncertainties in displacement are slightly better than that computed above. As given in Tables 2 and 3, the standard errors are 13 km for Arctic Ocean, 7 km for Fram Strait, and 9 km for the Weddell Sea. The distribution of errors (Figure 4i) from the Weddell Sea results seem to have a larger scatter because some of the buoys near the coast were motionless at certain times during the season (see Figure 3) and atmospheric contamination may be more of a problem in the Weddell Sea. We also note that there are latitudinal differences between the gridded fields used in the motion tracking; in the Arctic most of the ice cover is above 70°N , and in the Antarctic most of the ice cover is north of 70°S . The number of pixels used to compute an average brightness temperature is smaller in the lower latitudes because of the divergence of the orbits; there are more missing passive microwave pixels in the Weddell Sea fields compared to the Arctic Ocean.

The small biases in e_x and e_y are discussed later in this section. The directional errors between the passive microwave ice motion and the buoy/SAR displacement vectors are small and unbiased. However, the average ratios of the displacement magnitudes (scale) of the measurements, which in the absence of noise we expect to be unity, seem unreasonably large. In the derived motion field we expect subresolution scale displacements to be noisy and would give large "scale" errors, i.e., the division of two numbers with low signal content. The mean scale is greater than unity because of the larger uncertainty of the passive microwave measurements in the numerator. Here

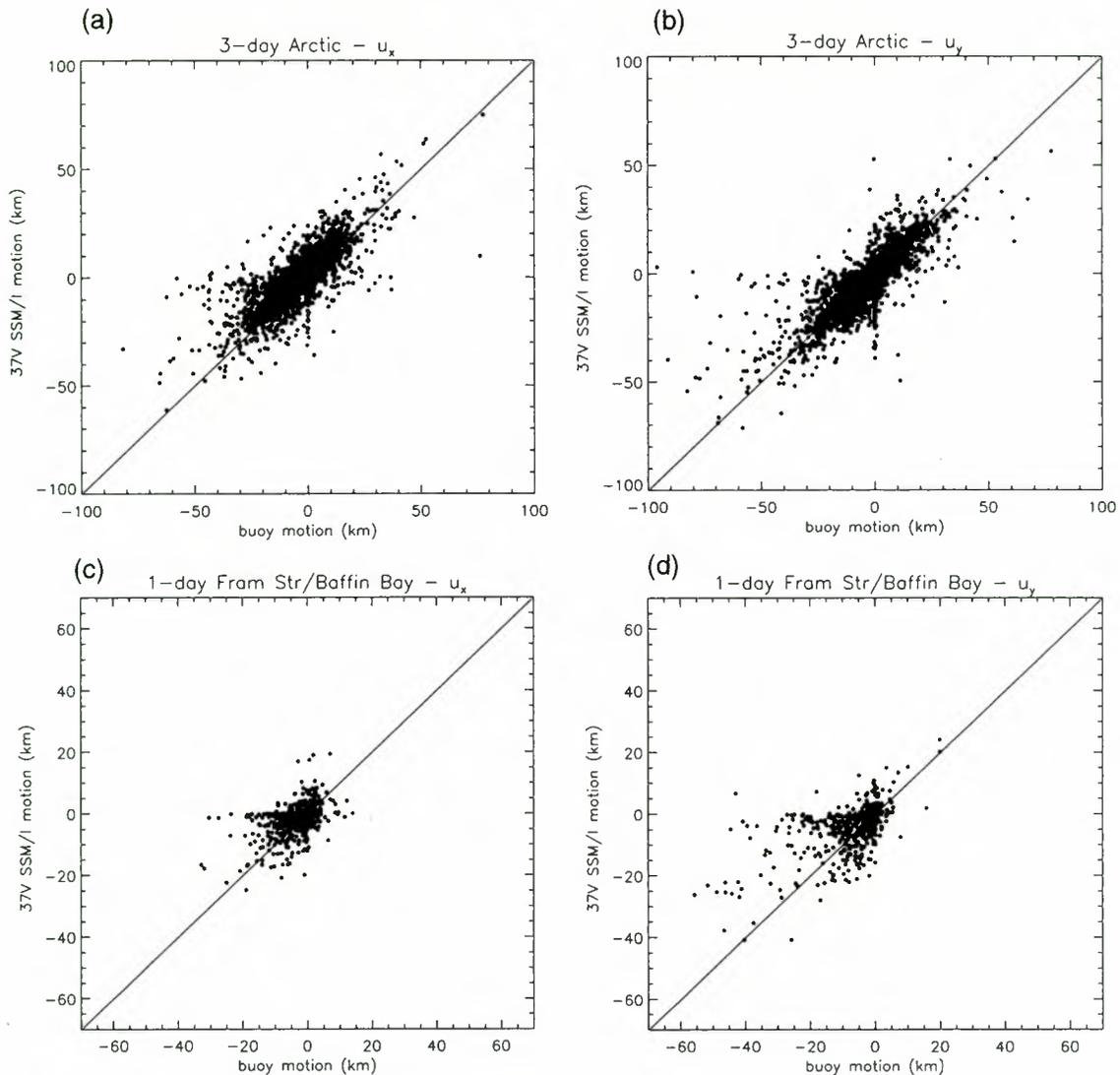


Figure 5. Comparison of the 37-GHz ice motion with buoy displacements. (a) Scatterplot of $u_x^{SSM/I}$ versus u_x^{buoy} (3-day Arctic), (b) scatterplot of $u_y^{SSM/I}$ versus u_y^{buoy} (3-day Arctic), (c) scatterplot of $u_x^{SSM/I}$ versus u_x^{buoy} (1-day Fram Strait/Baffin Bay), and (d) scatterplot of $u_y^{SSM/I}$ versus u_y^{buoy} (1-day Fram Strait/Baffin Bay).

the uncertainty of the quantity in the numerator is on the order of 5–10 km while the uncertainty of the denominator is ~ 350 m, a difference of more than an order of magnitude. To examine the contribution of noisy vectors to the comparisons, we excluded all data points where the displacements were less than half the pixel spacing (6.25 km for 85-GHz data and 12.5

km for 37-GHz data) and recomputed the statistics. More points are excluded from the 37-GHz results. This choice of threshold is not based on the data; rather, the expected error in the displacement estimates is from the tracker. The results in Tables 2, 3, and 4 show mean scale values close to unity as well as a decrease in the variability in the directional differences between the measurements. The decrease in the number of points used, after we excluded the points with small displacements, indicates the number of observations with motion which are less than the above thresholds. These results illustrate quite well the limitations of the observations; when the motions are small compared to the spatial resolution of the data, the displacements are noisy because they are comparable to the noise amplitude. We do not suggest the deletion of small displacement vectors from motion fields because the noise contribution to individual estimates would be reduced (averaging process) in the creation of mean fields of motion.

It is also remarkable that the uncertainty in the 37-GHz motion (in Table 3) is only slightly worse than the 85-GHz estimates, even though the spatial resolution is approximately

Table 5. Sources of Displacement Errors

Source	Expected σ
Geolocation error (σ_g) [Goodberlet, 1990], km	8
SSM/I binning error (σ_b)	0.5 pixel spacing
Tracking error (σ_f)	≈ 0.5 resolution element
Buoy motion (ARGOS/GPS), m	350/50
SAR motion (ERS), m	300
Misregistration between SSM/I and buoy/SAR ice motion	see text
Temporal smearing of brightness temperature field	see text

The expected σ is standard deviation.

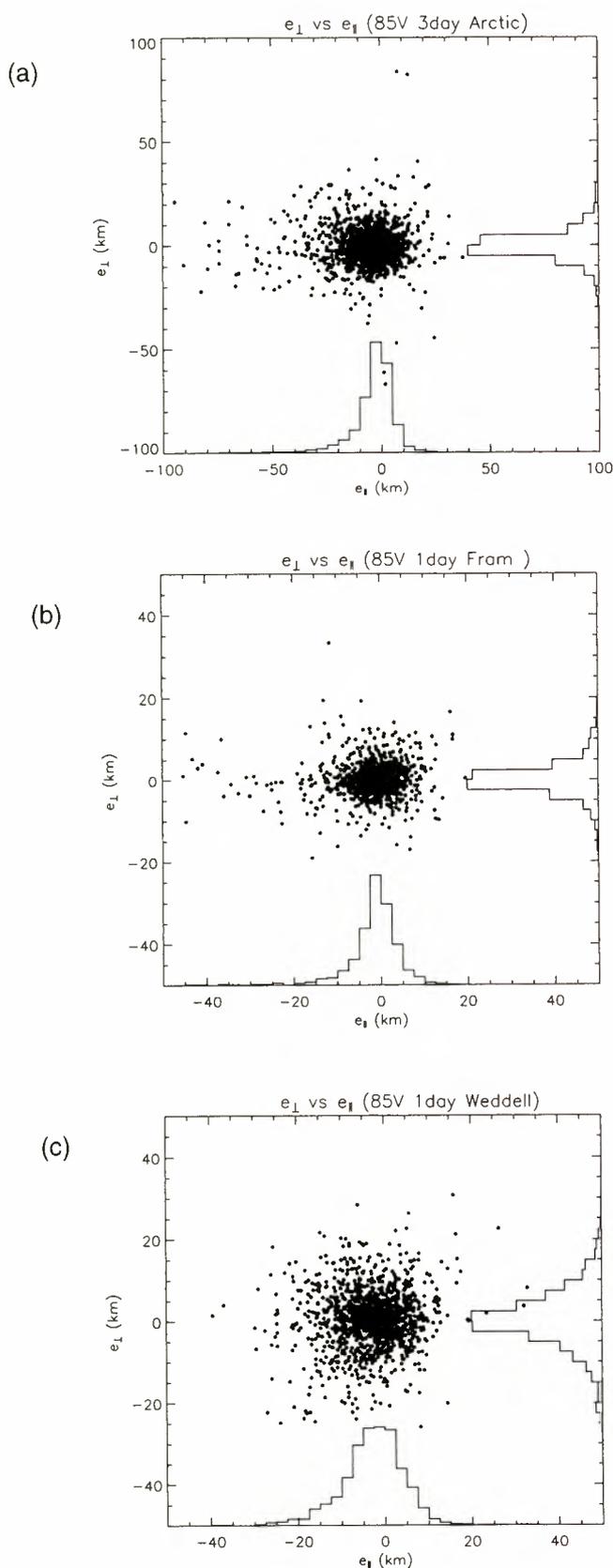


Figure 6. Plot of the displacement errors in the directions perpendicular, e_{\perp} ($=u_{\perp}$), and parallel, e_{\parallel} ($=u_{\parallel} - |\mathbf{u}_{\text{buoy}}|$), to the direction of buoy motion; (a) 3-day Arctic, (b) 1-day Fram Strait/Baffin Bay, and (c) 1-day Weddell Sea.

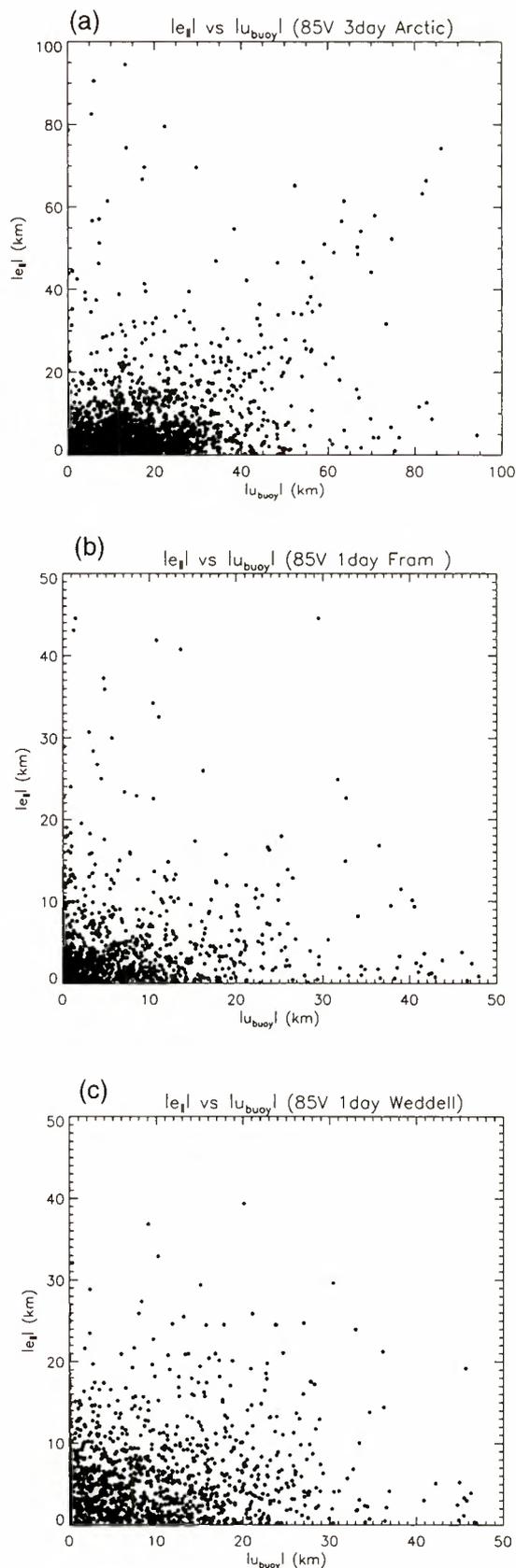


Figure 7. Plot of the displacement errors in the directions perpendicular, e_{\perp} ($=u_{\perp}$), and parallel, e_{\parallel} ($=u_{\parallel} - |\mathbf{u}_{\text{buoy}}|$), to the direction of buoy motion; (a) 3-day Arctic, (b) 1-day Fram Strait/Baffin Bay, and (c) 1-day Weddell Sea.

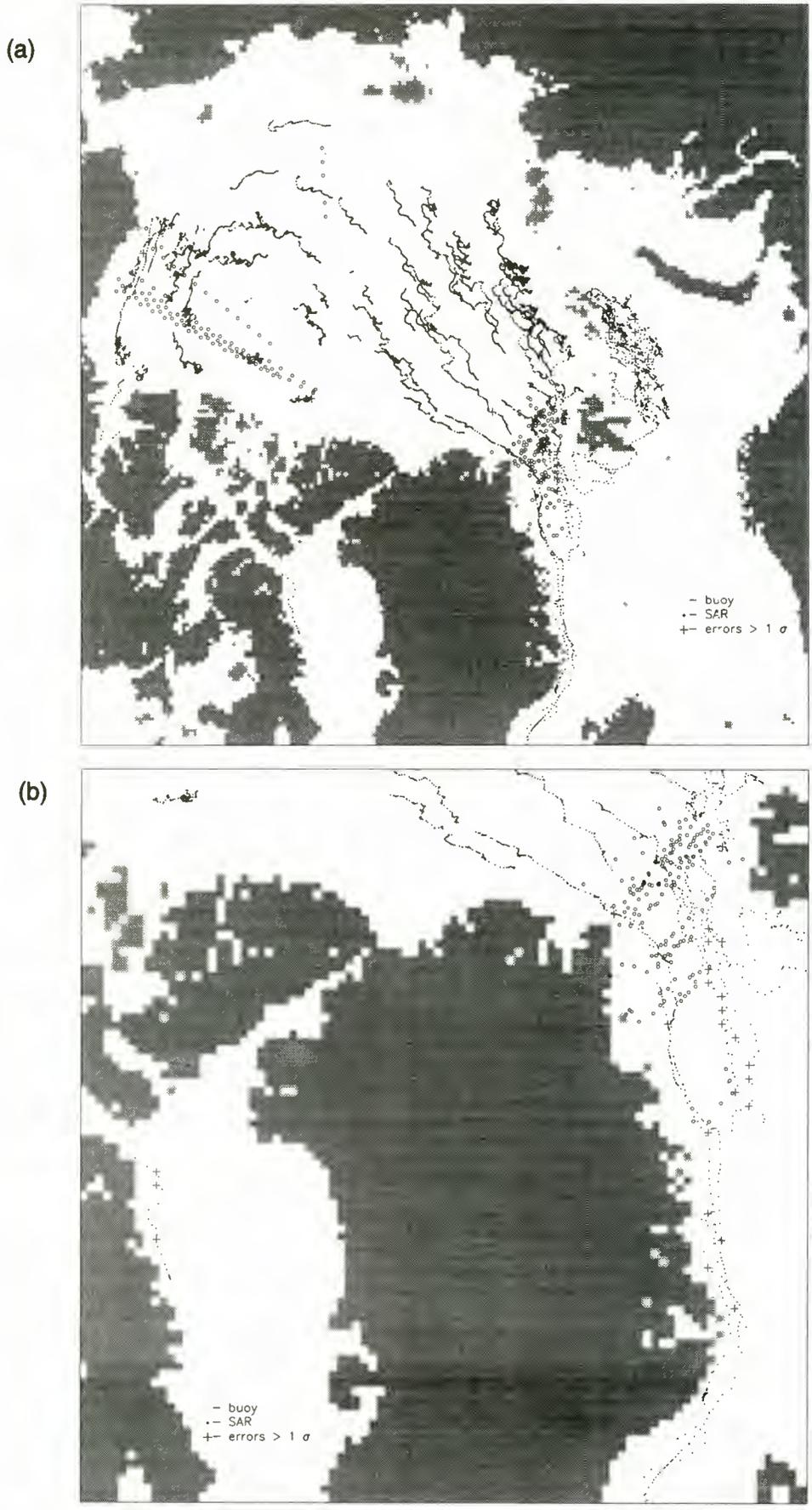


Figure 8. Location of the passive microwave ice motion estimates with errors $>1\sigma$ (crosses) from the mean; (a) Arctic, (b) Fram Strait/Baffin Bay, and (c) Weddell Sea. Buoy positions (dots) and locations where SAR ice motion are available (open circles) are also shown.

(c)

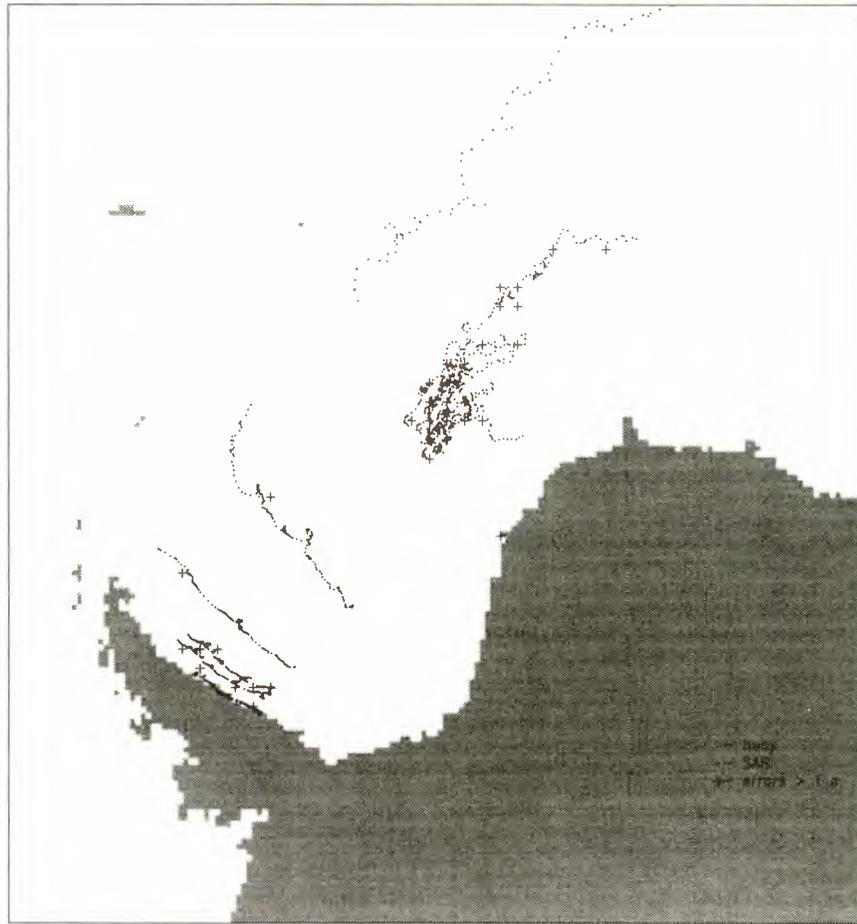


Figure 8. (continued)

twice as coarse. At 37 GHz the tracker uses a brightness temperature field 4 times the area as that of the 85-GHz data (200×200 km instead of 100×100 km). This additional area could be significant in determination of the average location or offset using the cross-correlation technique. We also note that the 37-GHz signal is relatively unaffected by atmospheric noise, such that the resolution advantage of 85-GHz could be partly offset by its sensitivity to atmospheric emissions.

Table 4 shows results of the comparison between 37- and 85-GHz ice motion and SAR ice motion. Each SAR ice motion observation represents the average ice motion over a ERS-1 100×100 -km image frame. As discussed above, this may be a more compatible comparison as the passive microwave motion also measures spatially averaged motion rather than the motion of discrete particles on the ice cover. Indeed, the results indicate, albeit with a smaller number of points available, that the passive microwave motion compares more favorably with SAR ice motion than the buoy ice motion. The average correlation between the motion estimates are higher and the directional differences are lower than results from the comparisons between passive microwave motion and buoy motion. The uncertainty and biases in the displacements are, however, similar to the buoy comparisons.

Of greater concern are the biases in the displacement estimates shown in Tables 2, 3, and 4. These statistically significant biases (highest in the Arctic, 1 km), although quite small in comparison to the resolution of the sensors, are similar when compared to buoy and SAR ice motions. In Figures 4 and 5 we

see that the distribution of errors, e_x versus e_y , appear to be isotropic; that is, they have no particular orientation in space. Figure 6, however, shows that there is structure when the errors are projected in directions which are parallel ($e_{\parallel} = u_{\parallel} - |u_{\text{buoy}}|$) and perpendicular ($e_{\perp} = u_{\perp}$) to the direction of buoy motion. Here we have taken u_{buoy} to be our best estimate of the true displacement and u to be the SSM/I displacement. The e_{\parallel} versus e_{\perp} plots clearly indicate that the biases are introduced by errors which are in the direction parallel to the motion rather than perpendicular to it. Hence this noise process did not manifest itself as directional biases in the comparison statistics. The orientations of these biases are not resolved in the e_x versus e_y plots because the average ice motion over a large region like the Arctic typically has no clear directional preference. The mean error of the population, e_{\parallel} , seems to be biased by a small sample population of SSM/I motion observations, which underestimated the actual motion as evidenced by the tail of the histogram. Figure 7 shows the dependence of the errors on the magnitude of motion. There seems to be a small dependence of the 3-day Arctic SSM/I motion estimates on the magnitude of displacement (Figure 7a), although this dependence is not obvious in the 1-day Fram Strait/Baffin Bay and Weddell Sea motion estimates. Next, we examine the spatial location of the points with errors $> 1\sigma$ of the uncertainty shown in Tables 2 and 3. When we first mapped these erroneous vectors, we thought they must be associated with a shear zone near Greenland and near the ice edge, but Figure 8 shows that the locations of the errors are determined simply

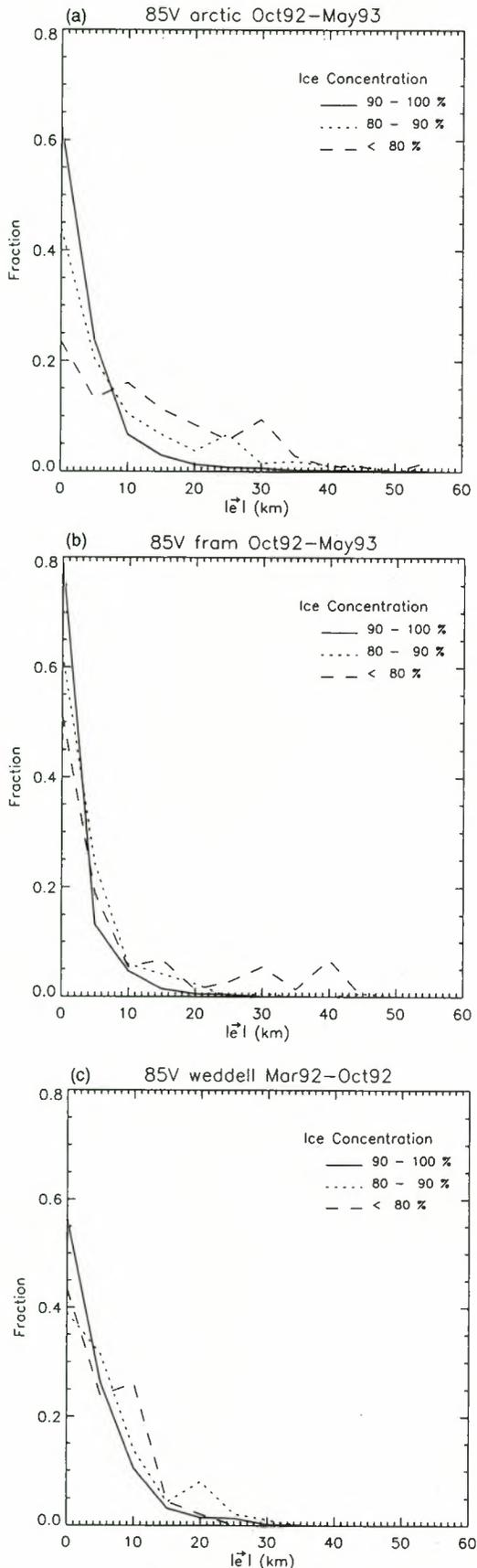


Figure 9. Histograms showing the error distribution as a function of three ranges of ice concentrations: (a) Arctic, (b) Fram Strait/Baffin Bay, and (c) Weddell Sea.

by where corroborative buoy tracks are located. Figure 8a, however, shows that the least reliable tracking occurs in the first-year ice of the Barents and eastern Arctic Ocean, where the motion is dynamic and where the brightness temperature fields are affected by storm systems that are active in this region. Tracking in the central Arctic seems more reliable. The regions with the least reliable results are also areas with typically lower ice concentration. In Figure 9 we plot the distribution of errors at different ranges of ice concentration. In the Arctic Ocean we find that the error distributions are broader as the ice concentration decreases; in other regions the effect does not seem to be as strong. In areas of lower ice concentrations it is possible that the ice motion from drifting buoys are less representative of the area-averaged ice motion obtained from SSM/I. This hypothesis (also suggested by *Kottmeier et al. [1992]*), however, cannot be easily tested with available data sets. If, indeed, the ice concentration is a good indicator of larger uncertainty in the tracking, it could be used to further cull the SSM/I motion fields of unreliable estimates.

In summary we made the following observations about the error characteristics: larger errors seem to occur in conjunction with larger motion in the Arctic motion fields; this small population of large errors is not centrally distributed but skewed in such a way as to introduce small biases in the population mean; and the tracking results obtained in the certain regions and areas of lower ice concentration are generally less reliable. Why is the tracker output biased toward smaller ice motion estimates for large displacements? In regions with large displacements, large velocity gradients, or low ice concentration the spatial distribution of the brightness temperature in the fields used for tracking tends to decorrelate with time, thus reducing the likelihood of the tracker finding the correct offset. In such cases, where the brightness temperature field is indistinguishable from adjacent fields, an incorrect peak in the correlation surface could be located by the tracker. If this estimated offset satisfies all the criteria in the filtering process, it would be accepted and would be output as a displacement estimate. Typically, it is more difficult to remove erroneous estimates of smaller displacements because unreasonably large displacement errors are usually less coherent with the local fields of motion and inconsistent motion estimates are more easily discarded by the filtering process. The more effective filtering of erroneous large displacement estimates introduces a bias in the error statistics. Thus this could be caused by an asymmetry in the filtering of errors in the motion tracker. In any case the errors in the motion estimates are intimately dependent on the tracking and filtering processes. Some filters, which may effectively remove outliers, could inadvertently introduce biases in the measurements.

3.3. Comparison of Buoy- and SSM/I-Derived Trajectories

One helpful way to visualize the quality of the SSM/I-derived ice motion is to derive trajectories and compare them with buoy trajectories. Figure 10 shows the location of 16 buoy trajectories in the Arctic Ocean spanning the 8 months of our SSM/I motion data (October 1992 to May 1993). Both buoy and SSM/I trajectories are shown in Figures 11a–11p. To create an SSM/I ice trajectory, we start at the same time and location as that of a buoy. Subsequent locations of the SSM/I track are derived from the gridded 3-day SSM/I motion fields. A motion estimate at an arbitrary location is determined by interpolation using available grid points within a 50 km radius of that location. For each comparison (Figure 11) we show the

Buoy Tracks - Oct 92 through May 93

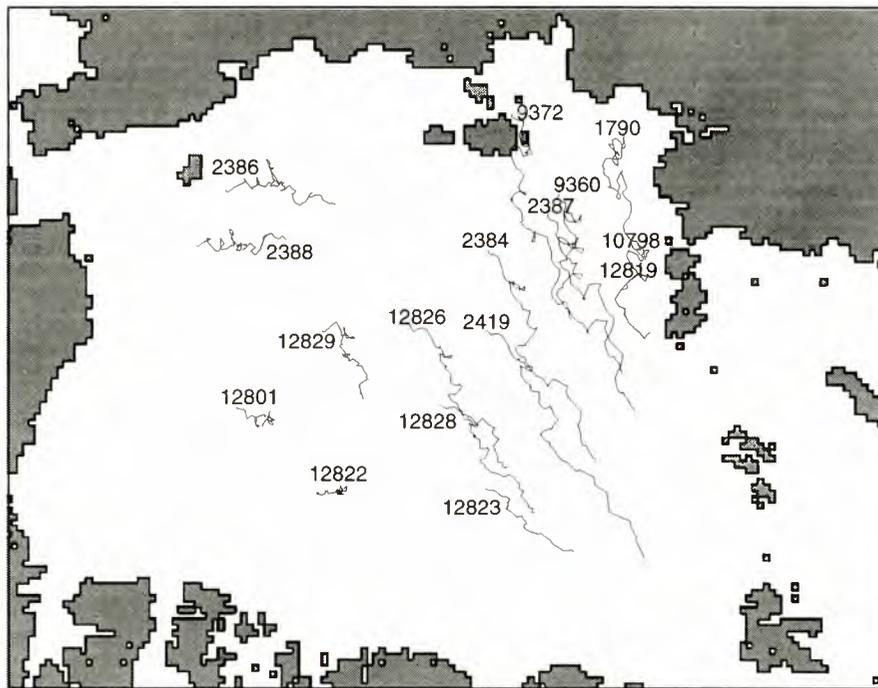


Figure 10. Location of buoy trajectories in the Arctic Ocean.

two trajectories and the corresponding differences in motion estimates during each 3-day period. Note that each plot has a different scale (shown in upper right-hand corner) that is dependent on the net displacement over the whole period. Locations where there were no SSM/I motion estimates (see Figures 11b and 11f) are indicated by zero differences between buoy and SSM/I displacements. In general, given the errors in the SSM/I motion estimates, the tracks agree reasonably well. If the errors were biased, we expect the derived tracks to diverge. The largest discrepancies are the tracks shown in Figures 11j, 11l, and 11p. These buoys (12801, 12822, 12829) are located in areas with relatively very little ice motion throughout the year (see scale on each plot). It is not surprising that these trajectories are noisy since these SSM/I motion estimates are dominated by noise. This suggests that maybe 6-day SSM/I ice motion fields should be considered or 3-day motion fields should be averaged to reduce the noise contribution in these slow-moving areas.

3.4. Polarization and Cross-Channel Comparisons

The vertically polarized (V) and horizontally polarized (H) channels at 85 and 37 GHz provide slightly different views of the ice cover as well as different sensitivities to atmospheric emissions (cloud liquid water, water vapor). In the winter Arc-

tic the ice cover could be modified by precipitation, storm systems, and opening/closing of leads. The presence of open water tends to decrease the brightness temperature observed by the H channel more than that of the V channel, although the contribution of open water to the total brightness temperature is probably rather small in the winter Arctic. Storm systems and precipitation could modify the passive microwave signature of the surface in a unpredictable ways. Ideally, for the purposes of ice tracking one would select the channel/polarization which is least sensitive to the atmosphere and provides the highest contrast between surface features. The additional condition is that these surface features remain relatively stable over the period of observation. However, it is not clear at the outset whether there are any differences in the quality of the motion data obtained from the V or H channels. The comparison of the motion results from the H and V channels with buoy motion showed the uncertainties in the observed motion to be similar. Comparison of the motion of the H and V channels (shown in Table 6) shows that they are consistent and their errors are smaller than that when compared to buoys. In summary, there is relatively little difference between motion derived from the H and V channels.

Table 7 shows the consistency in the ice motion between 85

Table 6. Differences Between unfiltered Vertically and Horizontally Polarized Ice Motion for October 1992 to May 1993

Region	Channel, GHz	Interval, days	Δ , km	e_x , km	e_y , km	θ , km	Scale	ρ	Number of Points
Arctic	85	3	-0.4 (1.0)	-0.3 (1.2)	-0.2 (0.9)	0.2 (2.7)	1.1 (0.1)	0.87	46852
	37	3	-1.0 (1.4)	-0.3 (1.8)	-0.2 (1.9)	0.4 (3.5)	1.1 (0.1)	0.85	28325
Fram Strait/Baffin Bay	85	1	0.0 (0.7)	-0.1 (0.8)	-0.4 (0.8)	0.0 (3.9)	1.0 (0.1)	0.89	48933
	37	1	0.1 (2.3)	0.0 (2.7)	-0.3 (3.2)	-1.0 (10.9)	1.0 (0.1)	0.86	10871

In each set of numbers the first is the mean, and the number in parentheses is the standard deviation.

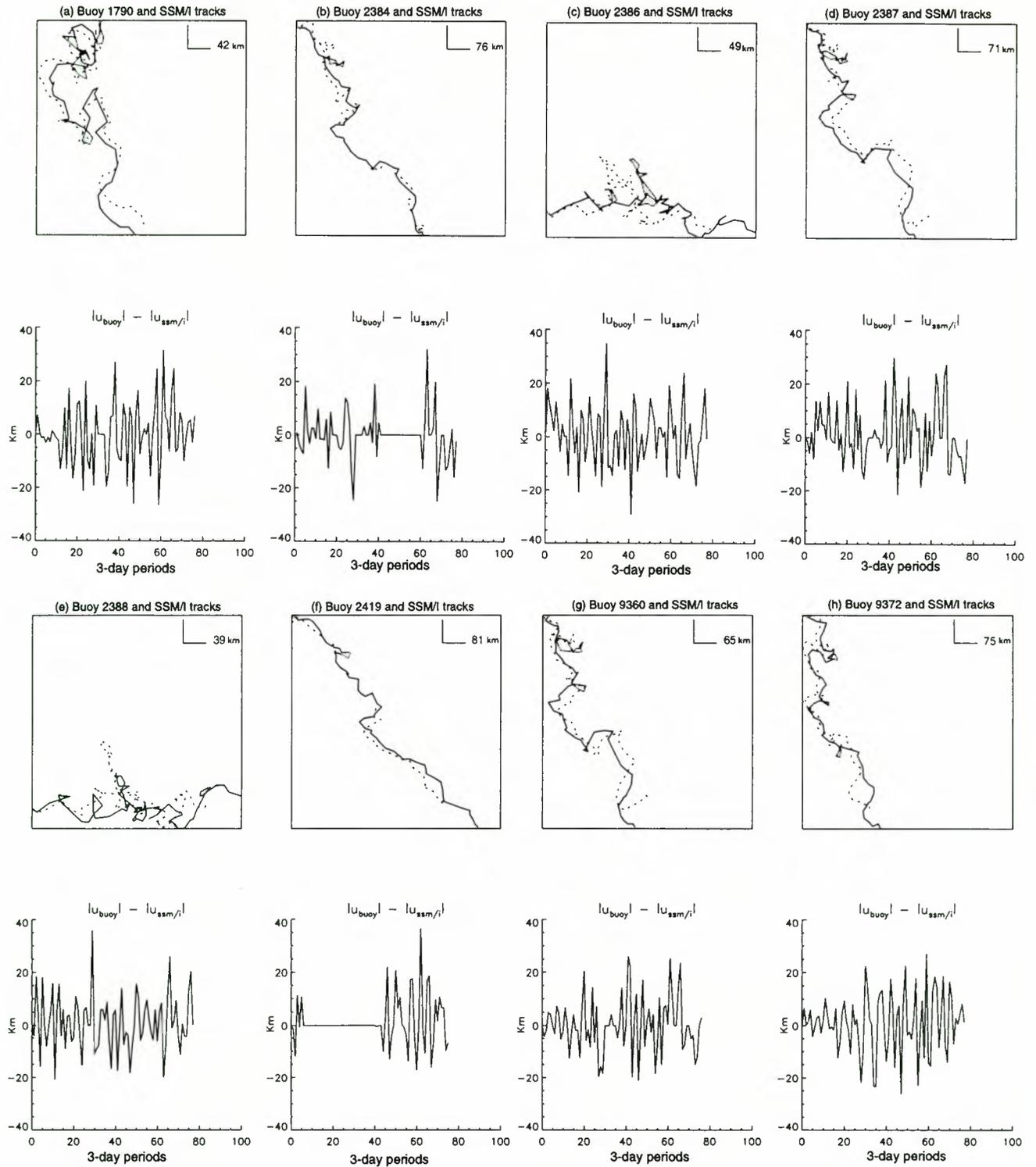


Figure 11. Comparison of 3-day buoy (solid line) and SSM/I derived (dashed line) tracks. (a) Buoy 1790, (b) buoy 2384, (c) buoy 2386, (d) buoy 2387, (e) buoy 2388, (f) buoy 2419, (g) buoy 9360, (f) buoy 9372, (i) buoy 10798, (j) buoy 12801, (k) buoy 12819, (l) buoy 12822, (m) buoy 12823 (n) buoy 12826, (o) buoy 12828, and (p) buoy 12829. Note scale on the top right-hand corner of each plot. c

and 37 GHz. Tables 2 and 3 already showed that there is only a slight degradation in the quality of the motion observations when one uses data from the lower-resolution channel. The biases in the observations between the two frequency channels

are small. It is remarkable that the variance in the displacement estimates between the two channels are also small. This shows that it is feasible to obtain ice motion from the 37-GHz as well as the 85-GHz brightness temperature fields.

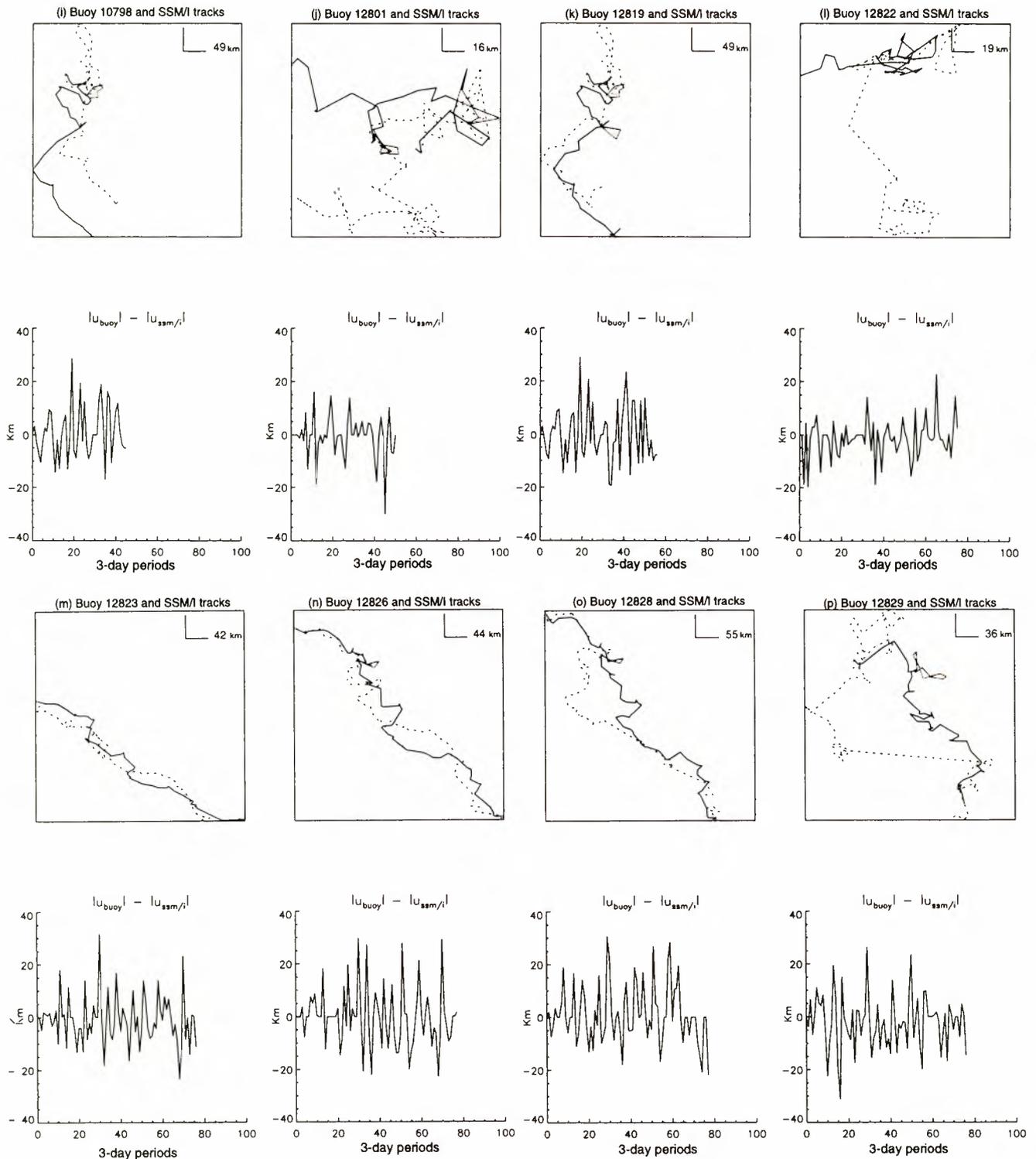


Figure 11. (continued)

4. Features in SSM/I Ice Motion Fields

We summarize the circulation patterns during the winter of 1992–1993 by grouping the Arctic passive microwave observations into fields of mean motion for October–November, December–January, February–March, and April–May and the Antarctic passive microwave observations into fields of mean motion for March–April, May–June, July–August, and September–October of 1993. The 2-month mean ice motion from

the three regions is shown in Figures 12, 13, and 14. For the Arctic case we compare these mean fields with motion estimates using buoy and wind data; these fields were provided by the National Snow and Ice Data Center (NSIDC). The following subsection describes how the mean fields are computed from daily or 3-day motion fields. Then we describe the new perspective gained with the higher spatial detail of the motion fields.

Table 7. Differences Between SSM/I 37- and 85-GHz Horizontally or Vertically Polarized Ice Motion for October 1992 to May 1993

Region	Channel, GHz	Interval, days	Δ , km	e_x , km	e_y , km	θ , km	Scale	ρ	Number of Points
Arctic	85V-37V	3	-1.3 (2.0)	-3.4 (1.6)	-0.1 (2.0)	3.6 (6.2)	1.1 (0.1)	0.75	39603
	85H-37H	3	-1.9 (2.0)	-3.3 (1.6)	0.1 (1.5)	3.3 (6.7)	1.2 (0.1)	0.76	39329
Fram Strait/Baffin Bay	85V-37V	1	0.18 (1.7)	-1.6 (1.5)	-1.0 (1.9)	4.6 (9.3)	1.1 (0.1)	0.81	26186
	85H-37H	1	0.1 (1.6)	-1.2 (1.9)	-1.0 (1.8)	3.4 (10.6)	1.1 (0.1)	0.79	25731

In each set of numbers the first is the mean, and the number in parentheses is the standard deviation.

4.1. Creation of the Mean Motion Fields

The following procedure is used to fill in holes in the data before the averaging process. At grid points where the tracker did not provide any observations we use an interpolation procedure to provide a motion estimate. A displacement estimate at a point is created by normalized weighting of at least three observations within a 100-km neighborhood with their relative weights determined by a spatial autocorrelation function. We model this function with a decaying exponential, namely

$$R(x) = \exp\left(-\frac{|x|}{l}\right)$$

where l is a length scale derived for each motion field and x is the distance from the point of interest. We find that this functional form fits the data quite well, although the results are relatively insensitive to the shape of the autocorrelation function because in most cases the correlation of the measurements remains fairly high over a distance of 100 km. Sample correlation functions from the three regions are shown in Figure 15. We compute the data points in the following manner: At each point on our gridded motion field we compute the distance-dependent correlation of that point with other grid points within a radius of 600 km; the points on the plot represent the mean correlation values from the application of this procedure to all the grid points on the motion field. The length scales used in the interpolation of the motion fields are computed from the data itself. The variability of this parameter for the different regions over the winter period of 1992–1993 are 685 ± 332 km (Arctic Basin), 569 ± 365 km (Fram Strait/Baffin Bay), and 875 ± 346 km (Weddell Sea).

The above procedure does not fill in all the holes in the data. There are remaining gaps in the time series of displacement vectors at a given grid point, because either there were no observations in the neighborhood for estimation of the motion or the sea ice concentration was $<15\%$ at that point. We compute an average displacement vector only if 80% of the observations are available over the 2-month period. For instance, the requirement is that there be at least 50 observations out of a maximum of 60 observations (1-day sampling) over a 2-month period before that grid point is included in the output field. Typically, this excludes points close to the ice edge as well as grid points in regions like Baffin Bay, where sea ice coverage remains low until late October (see Figure 13a).

4.2. Arctic Basin Circulation

Figures 12a, 12d, 12g, and 12j show the SSM/I-derived bimonthly mean fields of motion for the Arctic Basin between October 1992 and May 1993. The feature known as the Transpolar Drift is present, although its orientation and to a lesser extent its strength vary considerably. The other well-known

feature of the long-term mean field, the Beaufort Gyre, is never fully formed in the bimonthly fields, although it is somewhat present in April and May of 1993. There is never significant flow parallel to the Canadian archipelago, whereas there are substantial westward flows along the Alaskan coast. The other major feature, the outflow through Fram Strait, can be seen to come from quite different parts of the ocean at different times; this would suggest that the ice thickness in the Fram Strait comes from a variety of source regions and could be quite variable. The February–March field shows a well-developed cyclonic circulation centered between the New Siberian Islands and Severnaya Zemlya. There is quite a significant alternating exchange between the Arctic Ocean and the Kara and Barents Seas. The Kara Sea appears to be a net exporter of ice, sometimes into the eastern Arctic Ocean and other times into the Barents Sea. The entire 8-month period shows flow out of the Laptev Sea, which is a primary source of ice in the Transpolar Drift stream. The New Siberian Sea is alternately a source or a sink of ice. Many of these features have never been observable in such spatial and temporal detail.

To put these fields into perspective, we compare these fields with ice motion fields obtained from interpolation of buoy observations coupled with estimates from a wind-driven ice drift model [Pfirman *et al.*, 1997]. Early practice was to optimally interpolate a field of motion from about 10 to 20 buoy displacements, taking a “background” field of zero velocities. More recent practice has been to use geostrophic surface winds estimated from the gradient of surface pressure field as additional input to the motion interpolation. The correlation functions, used in the interpolation, have length scales of ~ 1500 km and represent motion on a large scale. At this length scale, smaller-scale structures in the motion field are smoothed. This scheme does not account well for the behavior of ice motion close to coasts or near islands in the Arctic, and the accuracy of these fields are dependent on the quality of the analyzed wind fields. Here we use fields interpolated from buoy and winds (provided by NSIDC) as the best source of motion estimates without satellite image tracking and compare these interpolated fields of motion with fields derived from SSM/I. The mean ice motion from the buoy fields and the differences between the passive microwave motion field and buoy-derived field are shown in Figures 12b, 12c, 12e, 12f, 12h, and 12i. The difference fields represent $\mathbf{u}_{\text{buoy/wind}} - \mathbf{u}_{\text{SSM/I}}$.

The October–November buoy/wind field (Figure 12a) shows an anomalous drift pattern in the southern Beaufort Sea compared to the SSM/I motions. This large eastward advection of ice near the coast is probably due to the existence of a high over the continent affecting the wind drift model. Also, a strong local anticyclonic motion pattern local to the Beaufort Sea can be observed. The SSM/I observed motion shows a very

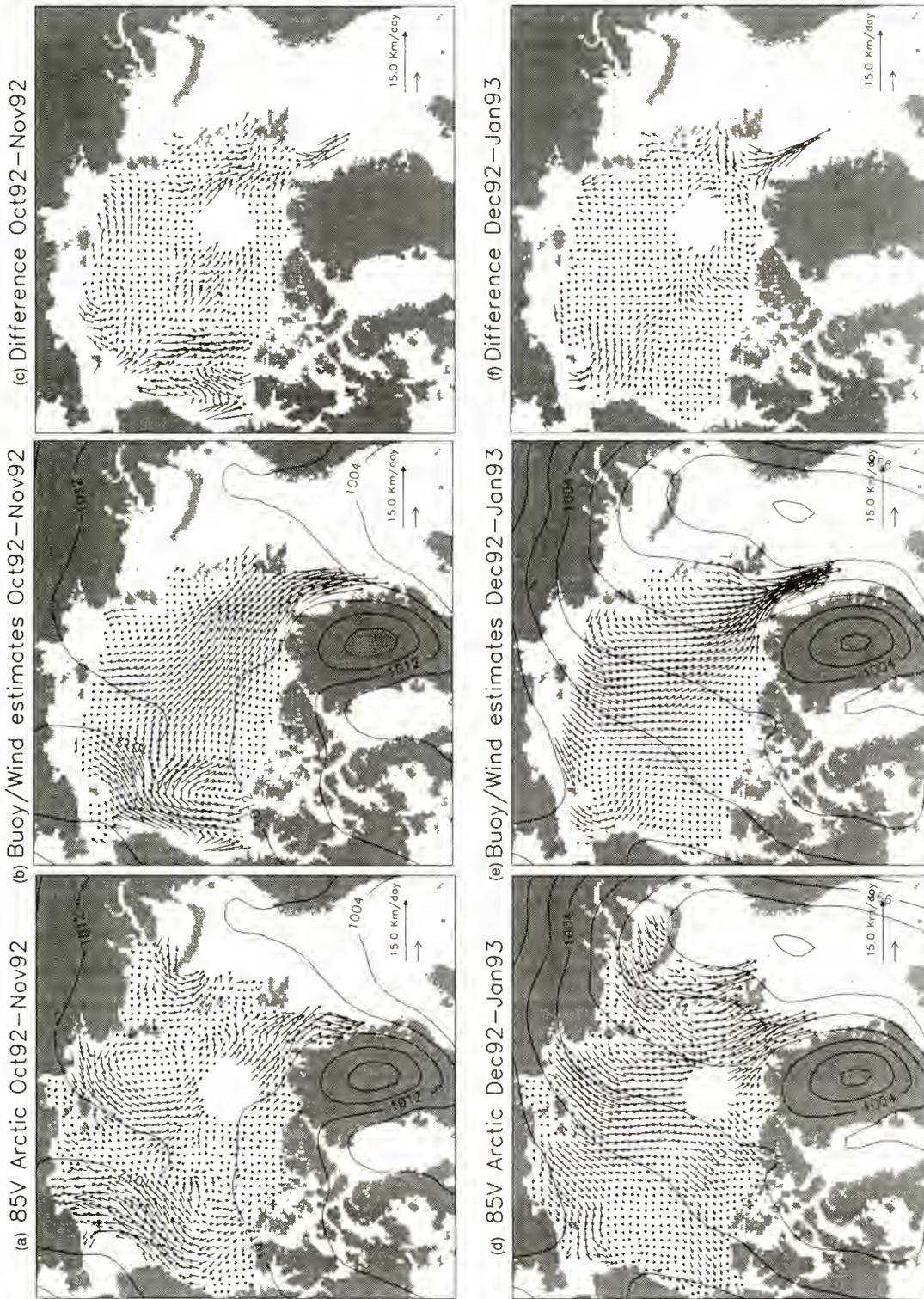


Figure 12. Comparison of the 85-GHz ice motion with interpolated buoy/wind motion field. The difference fields represent $u_{\text{buoy/wind}} - u_{\text{SSM/I}}$, that is, the buoy wind displacement minus the SSM/I displacement. The pressure contour interval is 4 mbar. Grid spacing is 80 km. (a) Two-month motion (October–November 1992) from SSM/I, (b) 2-month motion (October–November 1992) from buoy/wind estimates, (c) difference field (October–November 1992), (d) 2-month motion (December 1992 to January 1993) from SSM/I, (e) 2-month motion (December 1992 to January 1993) from buoy/wind estimates, (f) difference field (December 1992 to January 1993), (g) 2-month motion (February–March 1993) from SSM/I, (h) 2-month motion (February–March 1993) from buoy/wind estimates, (i) difference field (February–March 1993), (j) 2-month motion (April–May 1993) from SSM/I, (k) 2-month motion (April–May 1993) from buoy/wind estimates, and (l) difference field (April–May 1993).

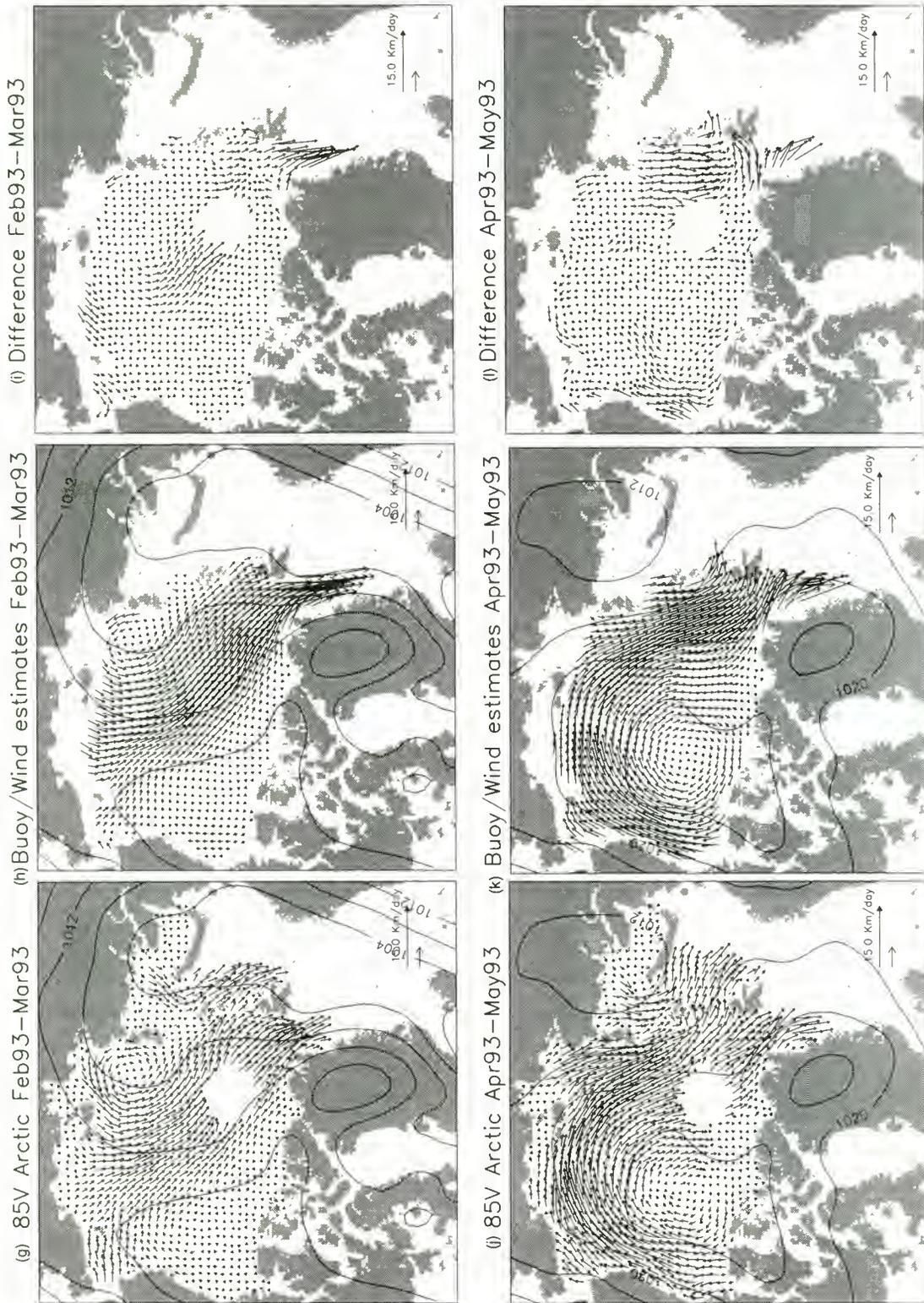


Figure 12. (continued)

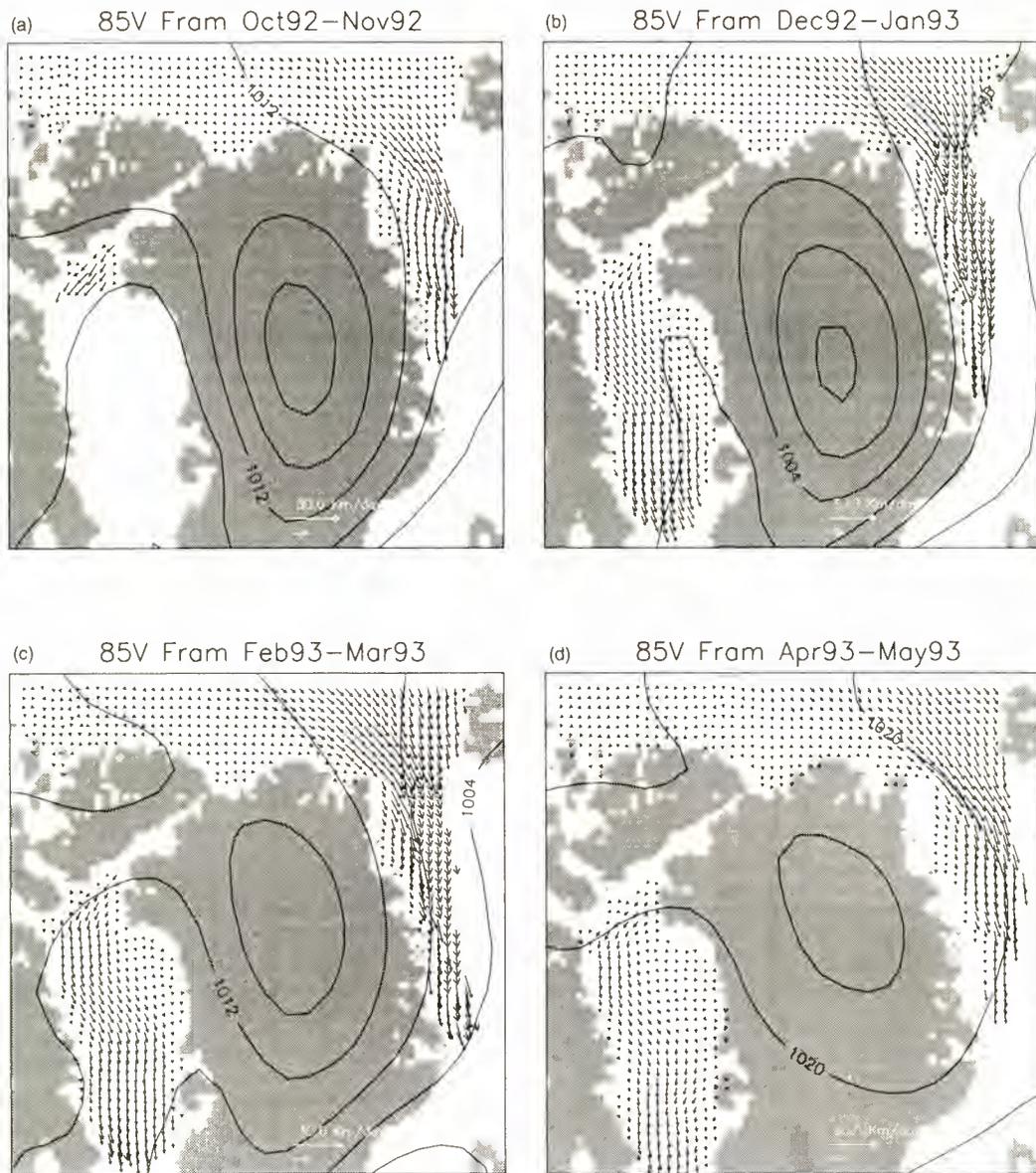


Figure 13. Averaged bimonthly ice motion (Fram Strait/Baffin Bay) between October 1992 and May 1993. The pressure contour interval is 4 mbar. The grid spacing is 40 km. (a) Two-month motion (October–November 1992) from SSM/I, (b) 2-month motion (December 1992 to January 1993) from SSM/I, (c) 2-month motion (February–March 1993) from SSM/I, and (d) 2-month motion (April–May 1993) from SSM/I.

different view of the drift pattern. Figure 12b shows a situation with a saddle point in the pressure field in the Chukchi Sea. Any inaccuracy in the placement of the saddle causes the structure of the wind inputs to the buoy/wind estimate to cause some error. This case shows clearly the buoy/wind overestimate of ice flux through Fram Strait. Notice too that the SSM/I motions show a distinct immobile region with distinct edges in the Lincoln Sea, whereas the buoy/wind estimate has smoothed this stagnated region. Figure 12d shows a buoy/wind estimate of a field of motion for April and May of 1993 in which the pressure field consists of a well-developed anticyclone, not unlike the long-term mean field, and the ice motion is fairly vigorous everywhere. The smoothness in the field is due to the rather long length scale assumed in the optimal interpolation. The comparable ice motion derived by tracking SSM/I shows more spatial structure, particularly near Spitzbergen and Franz

Josef Land. Note the existence of SSM/I-tracked motion data in coastal regions of the Siberian shelf and in the Barents and Kara Seas, where buoy data are virtually absent. The difference field in Figure 12 (which can only be computed where both data sets contain data) shows that in some regions, particularly in Fram Strait, the buoy/wind estimates are too strong. This is thought to be due to the use of wind drift to estimate motions where buoy data are absent and ice stresses impede ice motion and invalidate the “free drift” approximation by which wind data are incorporated into the buoy/wind optimal interpolation. The motion differences between buoy/wind estimates and the SSM/I estimates near the passages between Severnaya Zemlya, Franz Josef Land and Svalbard are most pronounced in Figure 12d. The buoy/wind estimates assume unconstrained flow around the islands and seem to overestimate the motion in this situation.

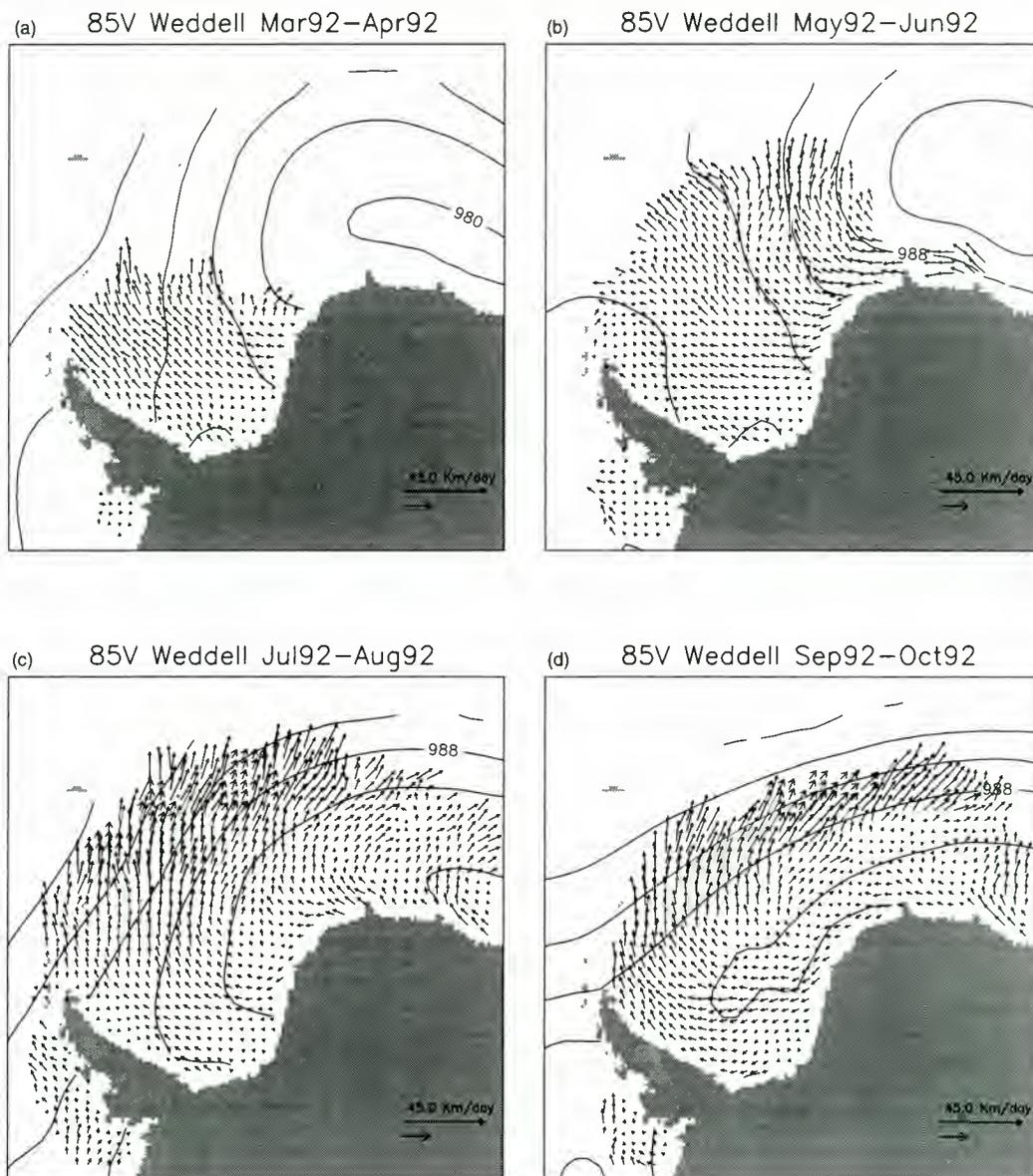


Figure 14. Averaged bimonthly ice motion (Weddell Sea) between March and October 1992. The pressure contour interval is 4 mbar. The grid spacing is 80 km. (a) Two-month motion (March–April 1992) from SSM/I, (b) 2-month motion (May–June 1992) from SSM/I, (c) 2-month motion (July–August 1992) from SSM/I, and (d) 2-month motion (September–October 1992) from SSM/I.

4.3. Fram Strait/Baffin Bay and Weddell Sea

The bimonthly mean sea ice motion in Fram Strait and Baffin Bay are depicted in Figure 13. An increase in the velocity of the sea ice is characteristic of the drift through the Fram Strait. Over the northeastern continental shelves of Greenland an anticyclonic feature of sea ice circulation can be observed. Mean motion north of Ellesmere Island is relatively small during all four periods. The ice motion west of Spitsbergen tends to be southwest and merges with the drift pattern through the Strait. The motion through the Strait is generally divergent. The ice motion in Baffin Bay is rather large, and the drift pattern is along the east coast of Baffin Island and southward toward Davis Strait into the Labrador Sea. On the west coast of Greenland the motion tends to be northwest but joins with the southward drift stream in Baffin Bay.

The net motion of sea ice in the Weddell Sea is shown in

Figure 14. The large-scale circulation pattern of sea ice in the Weddell Sea is not as well-characterized as the drift patterns in the Arctic Basin. The cyclonic circulation of sea ice, known as the Weddell Gyre, is a major feature in this region. East of the Antarctic Peninsula, the net ice drift is to the north along the peninsula, and the sea ice tends to turn east as it gets entrained in the Antarctic Circumpolar Current. Because of the persistent motion north and away from the coast the motion is highly divergent as new and old ice are carried northward.

5. Summary/Discussion

We have demonstrated that ice motion can be obtained not only from 85-GHz SSM/I data but also from the lower-resolution 37-GHz data. By comparing 8 months of SSM/I-derived motions with contemporaneous buoy and SAR ice

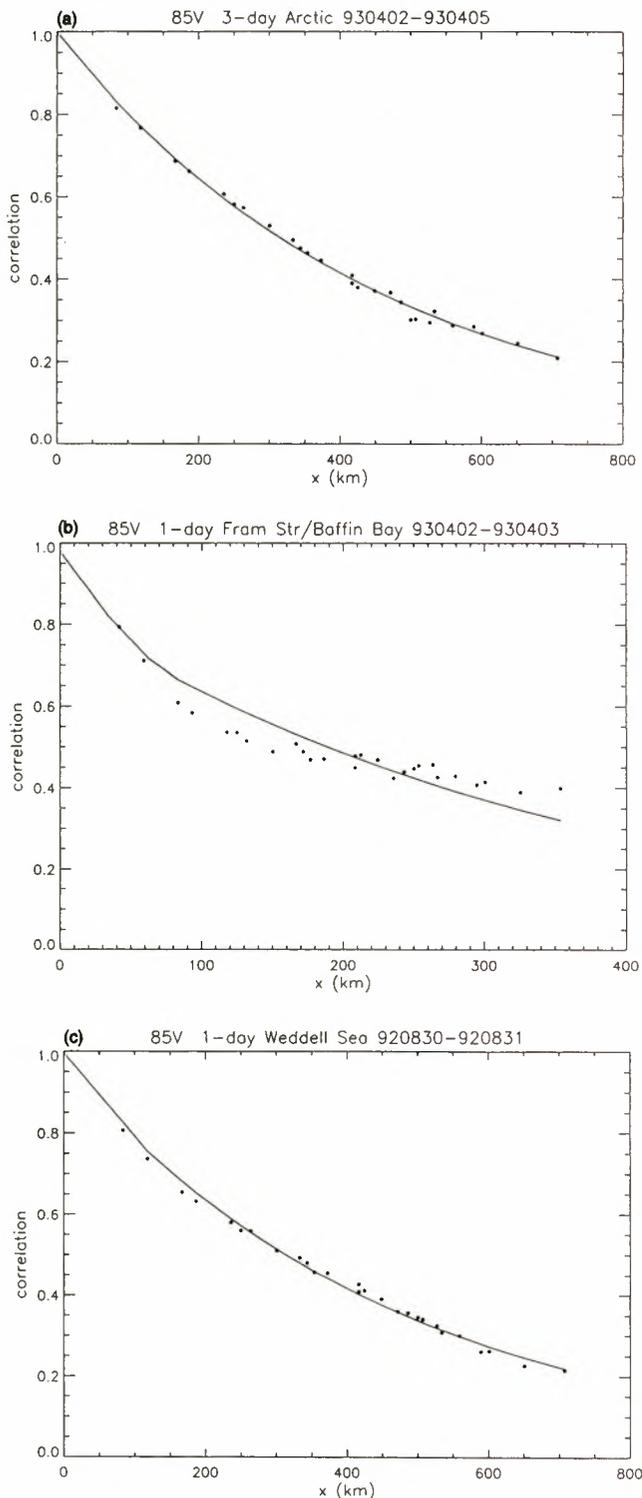


Figure 15. Sample autocovariance functions from the three regions: (a) Arctic, (b) Fram Strait/Baffin Bay, and (c) Weddell Sea.

motion observations, we find that the standard deviation of the sampling error ranges between 5 and 12 km. These errors are less than the nominal resolutions (12 and 25 km) of the two radiometer channels. Even though the uncertainties are substantial compared to daily ice displacements, the patterns and trajectories of ice drift are represented well in the 8 months of data that we have examined. If these uncertainties can be

characterized as unbiased additive Gaussian noise, as suggested by the data here, then averaging would reduce the noise contribution to the estimates of mean motion; the decrease being proportional to $1/\sqrt{N}$, where N is the number of observations averaged in the mean displacement. For example, a weekly average of daily motion fields would reduce the standard deviation by a factor of 2.6. Our bimonthly averages of 3-day displacements reduce the standard deviation by a factor of 4.5. An alternate approach is to increase the displacement signal by lengthening the time interval between observations. There are a number of alternatives for tracking ice motion with the passive microwave data sets; we believe that regionally subsetting the gridded daily fields of brightness temperature for motion analysis would yield the best results. This approach would allow one to select spatial and temporal sampling parameters to maximize the signal-to-noise ratio based on our understanding of the ice motion in each region. The currently available daily gridded SSM/I data sets are not ideal for ice motion analysis because the “drop in the bucket” sampling procedure averages observations over a day and somewhat smears the spatial resolution. Any ice motion tends to distort the brightness temperature field and introduces an error that depends on the magnitude of ice motion giving spatially non-uniform error statistics. Orbital data would better preserve the location and time of each observation.

Given the quality of the passive microwave motion estimates, these data sets are best used for the understanding of the synoptic and longer-term drift patterns rather than the detailed characteristics of daily motion. The daily fields, without further filtering, are not well-suited for computing daily deformation (e.g., divergence, shear, etc.) because the errors would be intolerable. Averaged fields are more suitable for computation of longer-term deformation.

In spite of these limitations the motion fields show spatial details not previously observable. We see flows near the coast that buoys have never provided; in fact, motion fields estimated from buoy motions and winds tend to overestimate many coastal flows. The exchange of ice between the Arctic Ocean and its peripheral seas is newly observable. There are approximately 20 years of SMMR and SSM/I data. The feasibility of extracting ice motion from 85- and 37-GHz data offers an extraordinary opportunity to create an ice motion data record starting in 1978 and continuing through the present and indefinitely into the future with planned SSM/I and other passive microwave instruments.

Acknowledgments. The SSM/I data and interpolated ice motion fields were provided by World Data Center A for Glaciology/National Snow and Ice Data Center, University of Colorado, Boulder. ERS-1 SAR PRI images were provided under grant AO2.USA164-1. We thank the staff of ESA ESRIN and UK-PAF for their support. Ignatius Rigor of the Polar Science Center, University of Washington provided the IABP buoy data set. The buoy measurements in the Weddell Sea were acquired under a project of the University of Bremen funded by the German Ministry of Research and Technology under grant 07KFT6171. R. Kwok and S. Pang performed this work at the Jet Propulsion Laboratory, California Institute of Technology, sponsored by the National Oceanic and Atmospheric Administration and the National Aeronautics and Space Administration. D. A. Rothrock and A. Schweiger performed this work at the Applied Physics Laboratory, University of Washington as part of Polar Exchange at the Sea Surface (POLES) under NASA contract NAGW-2407.

References

Agnew, T., H. Le, and T. Hirose, Estimation of large scale sea ice motion from SSM/I 85.5 GHz imagery, *Ann. Glaciol.*, in press, 1997.

- Emery, W. J., C. W. Fowler, J. Hawkins, and R. H. Preller, Fram Strait satellite image-derived ice motions, *J. Geophys. Res.*, 96(C5), 8917–8920, 1991.
- Fily, M., and D. A. Rothrock, Sea ice tracking by nested correlations, *IEEE Trans. Geosci. Remote Sens.*, GE-25(5), 570–580, 1987.
- Goodberlet, M. A., Special sensor microwave/imager calibration/validation, Ph.D. thesis, Univ. of Mass., Amherst, 1990.
- Holt, B., D. A. Rothrock and R. Kwok, Determination of sea ice motion from satellite images, in *Microwave Remote Sensing of Sea Ice*, *Geophys. Monogr. Ser.*, vol. 68, edited by F. D. Carsey, pp. 344–354, AGU, Washington, D. C., 1992.
- Kottmeier, C., J. Olf, W. Freiden, and R. Roth, Wind forcing and ice motion in the Weddell Sea region, *J. Geophys. Res.*, 97(D18), 20,373–20,383, 1992.
- Kwok, R., and G. F. Cunningham, Geophysical processor system data users guide, *Doc. D-9526*, Jet Propul. Lab., Pasadena, Calif., 1993.
- Kwok, R., J. C. Curlander, R. McConnell, and S. Pang, An ice motion tracking system at the Alaska SAR facility, *IEEE J. Oceanic Engineering*, 15(1), 44–54, 1990.
- Liu, A., D. J. Cavalieri, and C. Y. Peng, Sea ice drift from wavelet analysis of DMSP SSM/I data, *Int. J. Remote Sens.*, in press, 1998.
- Pfirman, S. L., R. Colony, D. Nürnberg, H. Eichen, and I. Rigor, Reconstructing the origin and trajectory of drifting Arctic sea ice, *J. Geophys. Res.*, 102(C6), 12,575–12,586, 1997.
- Rigor, I., and A. Heiberg, International arctic buoy program data report, 1 January 1996–31 December 1996, *Rep. APL-UW TM 05-97*, Appl. Phys. Lab., Univ. of Wash., Seattle, 1997.
- Thorndike, A. S., Kinematics of sea ice, in *The Geophysics of Sea Ice*, edited by N. Untersteiner, *NATO ASI Ser., Ser. B*, 146, 489–550, 1984.
- Thorndike, A. S., and R. Colony, Arctic ocean buoy program, 19 January 1979–31 December 1979, data report, Polar Sci. Cent., Univ. of Wash., Seattle, 1980.
- C. Kottmeier, Universität Karlsruhe, Institut für Meteorologie und Klimaforschung, Kaiserstrasse 12 (Physikhochhaus), 76128 Karlsruhe, Germany.
- R. Kwok and S. Pang, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109. (e-mail: ron@rgps1.jpl.nasa.gov)
- D. A. Rothrock and A. Schweiger, Polar Science Center, Applied Physics Laboratory, College of Ocean and Fisheries Sciences, University of Washington, Seattle, WA 98195.

(Received March 21, 1997; revised October 17, 1997; accepted October 28, 1997.)

