Application of hierarchical estimation methods to interpolation of SST data

M.J. Murray[1], M.R. Allen[1,2] & P.W. Fieguth[3]

1. RAL, Chilton, Didcot, 0X11 OQX, U.K.
2. Dept. of Physics, University of Oxford, U.K.
3. University of Waterloo, Ontario, N2L 3G1, Canada.

Introduction

Production of continuous fields from irregularly and sparsely sampled observational data is essential for many scientific applications of remote sensing. For example, atmospheric and oceanic models often require densely-gridded sea surface temperature (SST) and other geophysical fields. Traditional interpolation techniques, such as the Fast Fourier Transform, are not suited to irregularly sampled data, and optimal interpolation by brute-force matrix inversion is prohibitively expensive for the array sizes involved (around 2000×1000 for a global field at ten arcminute resolution). Instead, many approaches rely on heavy smoothing, essentially low-pass filtering the data set, which limits the scientific usefulness of the gridded fields.

For example, the NOAA blended analysis of AVHRR and in situ data represents the best available continuous SST dataset. These weekly SST fields at 1^o spatial resolution are produced using heavy smoothing (correlation length scales of $\sim$600–800km) which obscures small-scale structures and produces artefacts such as Gibbs ripples around coastlines Murray et al. (1998a). Each week's analysis is derived only from data collected in that week, and is thus independent of the previous week's analysis.

Since August 1991 the Along Track Scanning Radiometer series (ATSR-1 and ATSR-2 on board ERS-1 and ERS-2 respectively) has provided a near-continuous global sea-surface temperature (SST) record with a precision of approximately 0.3K (Murray et al. 1998a,b). This accuracy is associated with spatially-averaged SSTs which are generated at ten arcminute resolution. Daily coverage is around one third of the Earth's surface, with approximately one week's observations required to achieve a fairly full global SST field (allowing for 50cover). Typical coverage associated with three days of nighttime ATSR observations is shown in Fig 1a. Figure 1b shows the NOAA SST field corresponding to the period covered in Figure 1a. Also shown (Figure 1c) are the results from an ocean model run forced with observed winds and other meteorological data (from the European Centre for Medium Range Weather Forecasting). Of particular interest is the simulation of Tropical Instability Waves evident around the equator in the East Pacific. Comparison of these features with real observations requires the preservation of SST variability on a much finer scale than is associated with the NOAA analysis.

We describe below the dynamic estimation of the SST field using a recursive estimation algorithm which emulates the Kalman filter. This approach uses a recently-developed multiscale estimation algorithm for the update step, and makes simplifying assumptions about the surface dynamics leading to a computationally efficient prediction step. This approach facilitates the preservation of a much higher spatial and temporal resolution (initial results suggest that we can produce continuous SST fields at ten arcminute spatial resolution for three-day periods), but requires a fast optimal interpolation step. The multiscale estimation approach of Fieguth et al. (1998) provides a computationally efficient solution, enabling the interpolation, analysis and smoothing of extremely large data sets with realistic error estimates. Unlike other multiscale approaches such as wavelet decomposition, explicit treatment of inhomogeneities due to coastlines and boundary currents presents no major problems. This technique has already been successfully applied to altimetry data but the extension to SST, meteorological parameters, and other geophysical datasets presents substantial new challenges.

Multiscale Estimation

Observations are described using a hierarchical statistical structure or *tree* which divides into *subtrees* at *nodes* on a range of scales. Essentially the approach is to conditionally decorrelate the subtrees branching from each node so that these smaller subtrees can be processed independently, see Figure 2. (This requires knowledge of the statistics of the prior model underlying the observations; for example, a simple inverse correlation with distance might be assumed, although more complicated covariance models may be adopted following assessment of initial results.) In physical terms this corresponds to assuming that for each subtree, the influence of the external SST field can be completely represented by knowledge of SST along the subtree boundary. Completely sampling the boundary would result in an optimal solution, but a useful approximation can be achieved by sub-sampling this boundary, offering a computationally efficient method for interpolation of extremely large data sets.

Method

Key steps are outlined below:

Removal of underlying mean SST field. The 'mean' SST field for a particular day was derived by adding all SST observations for the 35-day period centred on the date and smoothing and interpolating the resultant field with a smoothing filter (see Figure 3).

Characterization of spatial and temporal statistics appropriate to SST in the Tropical Pacific. We have already conducted a study of observed lag correlations in ATSR data to determine the correlation length scales associated with SST in this region (Murray et al., 1999). This study has demonstrated that in addition to specifying different length scales zonally and meridionally, it is desirable to use a geographically varying correlation relation. This analysis also suggests a random noise measurement below 0.1K can be associated with ATSR $1/6°$ SSTs. (Repetition of this analysis with AVHRR data is necessary to confirm a meaningful input to the data fusion scheme outlined below.) (**Note:** Observations from the Caribbean were included in the SST field, and although the fields were estimated simultaneously, the two ocean basins were explicitly decoupled in the estimation.)

Establishment of suitable treatment of daytime and nighttime data. Ideally this would include the correct treatment of diurnal thermocline effects which are associated with high variability in daytime SST. However the present results were obtained by adding a bias of -0.2K to daytime SSTs.

Assimilation of data over time. A constant noise term was added to time-lagged observations to effect temporal decorrelation. The limited validation exercise carried out to date suggests that this approach is adequate for the tropical Pacific, but a more detailed examination will be required for more temporally variable regions. We may experiment with more realistic dynamics based on the advection-diffusion equation, initially using climatological ocean current information, or currents inferred from observed lag correlations in ATSR SST.

Results

Figure 4 shows the mean-removed ATSR SST observations for a single 3-day period centred on 16th July 1992, together with the estimate (assuming a $1/f$ spatial decorrelation with a meriodional correlation length of 50km and a zonal correlation length of 25km). Also shown is the 'mean' field removed and the error associated with the SST estimate. Subsequent to ascertaining this initial estimate, this was used as the 'mean' field for the next 3-day period, a procedure which was carried forward several months and then back to the starting time. Figure 5 shows several consecutive SST fields obtained using this Kalman filter approach. Figure 6 shows a cross section across the field at 14.5N; this shows the ability of the scheme to preserve fine detail where the observations permit. Note that the decoupling of the Pacific and Caribbean basins avoids the smearing apparent in the Reynolds interpolation scheme.

Conclusions

We have demonstrated the use of dynamic estimation to SST fields, and developed the following:

- Capability of assimilating data over time.
- Capability of dealing with spatially-varying prior model
- Ability to merge estimates from sources with different distributions of observations and error statistics.
- Ability to merge quantities related by a defined model (spatially varying if required).

ATSR/AVHRR/Buoy data fusion

Realisation of the power of this method will come with its application to the assimilation of multiple datasets. Clearly it is desirable to assimilate ATSR SST data with data from AVHRR and also include information from in situ measurements. A fundamental requirement for this data fusion exercise is the successful characterization of the relationship between the various measurements, and a precise description of the errors

and noise associated with each dataset. With this objective in mind, we are currently conducting a detailed comparison of ATSR and AVHRR observations in order to develop a methodology for assimilation of these data into a consistent, continuous, and accurate global skin SST analysis at a high spatial and temporal resolution. In particular, we aim to demonstrate the feasibility of producing this SST analysis operationally. We anticipate concentrating on a restricted data set (several months of data in 1993, covering the tropical Pacific and one other region).
Many factors complicate the ATSR/AVHRR/Buoy data fusion exercise, not least that whereas infrared radiometers sense the 'skin' SST characteristic of the uppermost millimetres of the ocean, in situ 'bulk' measurements are typically made at approximately 1m depth. As this bulk−skin difference varies from around $-0.5 - 4$K it is desirable to correct for this when assimilating observations. We have derived an empirical model relating bulk and skin SST in terms of local windspeed and heat flux (see Figure 7) and the estimation scheme enables this to be incorporated directly into the SST assimilation.

Future plans

- Production of time series of *global* SST fields with reliable error estimates for one or more ocean basins.
- Generation of a skin SST climatology. A requirement of this project is the generation of 'mean' temporally-averaged SST fields. A natural extension of this work is the generation of a monthly skin SST climatology. This will be representative of nighttime SST, which is not subject to the highly-variable diurnal warming (although adjusted daytime SSTs will be included).
- Evaluation feasibility of global operational assimilation of SST data.
- Development of better spatial and temporal statistical models.
- Better characterization day/night variability

References

Fieguth, P.W., M.R. Allen and M.J. Murray, "Hierarchical methods for global-scale estimation problems", *IEEE Transactions on geoscience and remote sensing*, 1998.
Merchant, C.J., A.R. Harris, M.J. Murray and A. Zavody, "Towards the elimination of bias in satellite retrievals of sea surface temperature: Part 1: Theory, modelling and inter-algorithm comparison", *J. Geophys. Res.*, (accepted), 1999.
Murray, M.J. et al., "Actual and potential information in dual-view radiometric observations of SST from ATSR", *J. Geophys. Res.,* 103, C4, 8153–8166, 1998a.
Murray, M.J., M.R. Allen, C.J. Merchant, and A.R. Harris, "Potential for improved ATSR dual-view SST retrieval," *Geophysical Research Letters,* 25, 17, 3363–3366, 1998b.
Murray M.J., M.R. Allen and P.W. Fieguth, "Optimal interpolation and statistical analysis of ATSR data', Nice EGS, 1998c.

Murray, M.J. et al., "Towards an improved global SST analysis" *J. Geophys. Res.,* (to be submitted), 1999.

Reynolds, R.W., and T.M. Smith, "Improved global sea surface temperature analyses", *J. Clim.,* 7, 929–948, 1994.