Improved Cloud Detection and Cross-Calibration of ATSR, MODIS and MERIS Data

James J. Simpson

Digital Image Analysis Laboratory, Scripps Institution of Oceanography
University of California, San Diego
La Jolla, CA 92093-0237, U.S.A.

ABSTRACT: Unsupervised classification methods, when combined with robust information vectors and feed forward neural networks for labelling, provide a basis for very accurate cloud detection in satellite data. Preliminary results, obtained by applying these methods to satellite data, show the usefulness of these methods for cloud detection. Anticipated directions for our ERS-II, ENVISAT and EOS-AM investigations, which will utilize a combination of data from the family of ATSR instruments, MERIS and MODIS, will emphasize a suite of complementary atmospheric, terrestrial and oceanic applications.

I. INTRODUCTION

A. Importance of Accurate Cloud Detection in Satellite Scenes over the Ocean and Cloud/Snow Separation Over Land

1. General Considerations

The ability to accurately and automatically segment clouds in satellite (ATSR, MODIS and MERIS) scenes over the ocean and land is important for a wide range of disciplines in the earth sciences (JSC Working Group, 1997). Clouds, for example, significantly affect the net heating and cooling of the atmosphere and underlying ocean by modifying the short-wave (Hobbs and Deepak, 1981) and long-wave (Hunt, 1982) radiation. This net radiative heating governs the thermodynamics and dynamics of the atmosphere, which in turn influences the formation and dissipation of clouds (Matveev, 1984). The potential feedback effects associated with this cloud-radiation interaction are one of the greatest sources of uncertainty in determining the relation between changes in climate and changes in external conditions such as solar radiation and atmospheric carbon dioxide concentration (Henderson-Sellers, 1984; Ramanathan, 1987).

2. Importance for Climate and Global Change Modeling

Many of the above cited issues are important for improved understanding of global change processes and are being studied using climate models. Such models generally parameterize the atmospheric radiative transfer process, and hence also parameterize the effects of clouds (Henderson-Sellers, 1984; Ramanathan, 1987). Moreover, clouds are the most important transient phenomenon incorporated into climate models (Henderson-Sellers, 1984). Unfortunately, both oceanic and land cloud climatologies suffer from deficiencies which affect their usefulness in climate models. The widely used cloud climatology of London (1957), for example, was compiled from surface-based observations and the oceanic data are particularly sparse and uncertain. Moreover, terrestrial cloud climatologies, while more common than their oceanic counterparts, suffer from the more difficult problems associated with accurate cloud detection in satellite scenes over highly heterogeneous land surfaces.

3. *Importance for Sea Surface Temperature*

One of the fundamental diagnostic parameters for climate change is sea surface temperature (SST). Many factors, however, can affect the accuracy of SST retrievals derived from satellite-observed radiances. Atmospheric aerosols, for example, scatter radiation and directly affect the observed brightness temperatures (BT) in the visible region of the spectrum (Zege *et al*., 1991). These effects can compromise cloud detection especially if visible data (near the blue region of the spectrum) are used. Atmospheric water vapor is known to attenuate infrared signals and hence degradate SST retrievals (McMillin, 1975; Deschamps and Phulpin, 1980). Perhaps the most significant source of residual error in SST retrievals, however, is the presence of undetected clouds in the scene (Henderson-Sellers, 1984: Robinson, 1985; Simpson and Humphrey, 1990). In fact, cloud contamination has been cited as one of the major factors, limiting more accurate MCSST retrievals even from improved sensors such as the Along Track Scanning Radiometer (ATSR) on the ERS-1/ERS-2 satellites (Harris *et al.*, 1995; Jones *et al.*, 1996a, b; Simpson *et al.*, 1998a). Moreover, cloud-free satellite-derived MCSST fields are also important for analyzing a wide variety of oceanic processes (large-scale surface circulation (Yan and Breaker, 1993), mesoscale eddy dynamics (Koblinsky *et. al*, 1984), and mixed layer dynamics (Yan *et. al*, 1990)).

4. *Importance of Cloud/Snow Cover Separation to Climate and Global Change Processes*

Snowpack that accumulates in the western United States each winter is a critical resource for water supplies and ecosystems. Year-to-year variations in the likelihood of plentiful snowpack, of sudden and catastrophic snowmelt and early or late snowmelt, are associated with variations of the global and, especially, the Pacific climate system. The best understood of these climate variations are the El Niño Southern Oscillation (ENSO) processes (Allan *et al.*, 1996) that form on an irregular basis in and above the tropical Pacific Ocean, influencing weather (Ropelewski and Halpert, 1987; 1989), snowpack (Cayan, 1996) and streamflow (Cayan and Webb, 1992; Dettinger *et al.*, in press) throughout the western United States. Traditionally, accurate estimates of areal extent of snow cover from satellite data have been difficult to obtain (Simpson *et al.*, 1998b). The new wavelength bands on MODIS and the ATSR family of sensors provide a potentially fruitful way to obtain improved estimates of areal extent of snow cover. New snow-related products to be derived from these data are of critical importance to climate and global change models (effects planetary albedo) and to several practical applications such as water resource management in the western United States, Europe, and parts of Africa (Morocco).

B. Overall Approach and Economies of Scale

Clearly, the stability of the entire satellite retrieval process (calibration, cloud detection, geophysical retrieval, validation) is important for: 1) improving the quality of satellite-derived skin SST; 2) conversion of the skin SST to bulk oceanic mixed layer temperature; 3) improving our understanding of ocean circulation; 4) improving the statistical reliability of satellite-based cloud climatologies over both land and ocean; and 5) improving our estimates of areal extent of snow cover. Of special importance is the need to validate each product with the best available *in situ* data. Validation provides both a way to quantify residual error in products and a feedback mechanism to improve the stability of the entire satellite retrieval process. These issues relate directly to the ongoing objectives of many international programs such as the International Geosphere-Biosphere Program (IGBP).

Economies of scale provide a basis for producing the above set of computationally related (although geophysically distinct) variables from ATSR, MODIS and MERIS data. Clouds must be identified in order to produce valid SST. Residual information in the cloud field can be simultaneously processed with minimum additional effort to produce cloud climatology. Optimal cloud detection over the ocean, in turn, requires that the land and clouds over land be segmented from the ocean and clouds over ocean data. Given the availability of 1.6 μm and other visible and mid-infrared data, it seems imprudent not to simultaneously process the land components of the scene as well. Moreover, concurrent processing of these different geophysical fields from the same original satellite data greatly reduces I/O inefficiencies and overall computational time compared to the computational resources that would be required if each group of products (ocean, atmosphere, land) were processed separately from the original data. Hence, our overall plan enhances the cost/benefit ratio to the climate change community.

II. DIFFICULTIES WITH ACCURATE CLOUD DETECTION IN SATELLITE SCENES

The simplest approach for cloud detection in a scene is to apply a set of static thresholds (albedo, temperature) to every pixel in the scene. This method can fail for several reasons: 1) subpixel clouds and cloud-pixel misalignment (i.e., the field of view of the radiometer falls on the edge of the cloud) can lead to errors because the distribution of radiances is non-uniform within the pixel (Shenk and Salomonson, 1972; Simpson and Humphrey, 1990); 2) variations in BT result from pixel specific variations in viewing and illumination geometry (Dalu, 1985; Foody, 1988); 3) sensor aging (Duggin, 1985); and 4) the spectral response of clouds varies with cloud type and height (Liljas, 1987).

Spatial coherence methods (Coakley and Bretherton, 1982) have an advantage over static threshold methods because they utilize the local spatial structure of the infrared radiance field to determine cloud-free and cloud-covered pixels and infer partially filled fields of view. Spatial coherence methods can fail, however, when: 1) the cloud system in the image is multilayered (which often is the case); 2) the clouds everywhere in the scene are smaller than the instrument's field of view; 3) the clouds have variable emissivity (cirrus clouds); and 4) cloud-free strong ocean thermal gradients exists in the scene (see examples given by Gallaudet and Simpson (1991)).

Other approaches (Saunders and Kriebel, 1988) combine static thresholds, spatial coherence tests and geometric viewing criteria. This approach can be limiting because the thresholds used are both static and regionally specific. Moreover, Gallaudet and Simpson (1991) have shown that such methods (like spatial coherence methods) often erroneously interpret regions of cloud-free strong ocean thermal gradient as cloud.

III. CLOUD AND CLOUD SHADOW: SIGNAL VERSUS NOISE

A. Cloud and Cloud Shadow as Signal

For atmospheric applications, both cloud and cloud shadow in satellite data are major signals. Accurate retrievals of short-wave and long-wave fluxes require information on clouds. Cloud shadow can be used to estimate both cloud base (Gurney, 1982; Berendes et. al., 1992) and cloud top height (Simpson et. al., 1999a). Cloud shadow is also important for proper numerical simulation of mesoscale atmospheric circulations that lead to major convective storm systems (Bailey et. al., 1981; Segal et. al., 1986; Lipton, 1993; McNider et. al., 1995).

B. Cloud and Cloud Shadow as Noise

A significant source of residual error in sea surface temperature (SST) retrievals is the presence of undetected clouds in the scene (Henderson-Sellers, 1984; Robinson, 1985; Simpson and Humphrey, 1990; Simpson et. al., 1998a). Even for data taken with advanced instruments, such as the Along Track Scanner Radiometer (ATSR) family of sensors, cloud cover has been reported as one of the major factors limiting SST retrieval accuracy (Jones et. al., 1996a, b), especially the detection and retrieval of low-lying marine stratiform cloud. This cloud type, in particular, predominates in certain geographical regions—namely continental up-wellings, which provide conditions of warm continental air overlying cold water. Likewise, the effects of subpixel cloud contamination on SST retrieval are important as documented by Harris et. al., (1995).

Both cloud and cloud shadow can affect the accuracy of satellite-derived vegetation estimates. Such estimates are important in global change models because terrestrial vegetation affects the climate system (through hydrometeorological feedback loops) on a wide range of spatial and temporal scales (Verstraete and Pinty, 1991). While early studies have emphasized the effects of aerosols, Rayleigh scattering, dust, and clouds on vegetation products (Normalized Difference Vegetation Index (NDVI)), more recent work has also highlighted the importance of undetected cloud shadow as a contaminant in satellite-based vegetation estimates (Simpson and Stitt, 1998)

IV. APPROACH

A. Cloud Detection

A hybrid cloud detection procedure is under development. It consists of an unsupervised segmentation of the scene into its natural groups using a new clustering procedure developed by Simpson et. al. (1999b) and labelling of the segments into geophysical classes (e.g. cloud, clear land, clear ocean) using a feed-forward neural network (FFNN).

The clustering algorithm uses a new splitting procedure which allows for arbitrary orientation of the splitting decision surface relative to the decision space. This feature overcomes problems in the historical ISODATA procedure of Ball and Hall (1964) and significantly improves average execution time (see Simpson et. al., 1999b for details).

The FFNN labeller has been developed as part of the Geostationary Meteorological Satellite (GMS) Pathfinder Project. A procedure for efficient training set development and subsequent learning by the FFNN has also been implemented (Simpson et. al., 1999c).

B. Data Fusion and Inter-Sensor Calibration

Data fusion involves any mathematical/statistical process that maps data taken on different space-time grids onto a uniform space-time grid with known error estimates. Data fusion makes possible an interpretation of the scene physics not obtainable from a single sensor and/or reduces the uncertainty associated with data from individual sensors. The diversity of satellite and *in situ* data types to be used requires the use of sound data fusion methods for meaningful results.

We will determine and implement the optimal data fusion method(s) for this study. Wavelet transformations, multi-resolution analysis (relational filter banks), and pixel level image data fusion methods will be considered. Data fusion also can be done successfully under a Bayesian formalism. This approach requires *a priori* information about the likelihood of change between the acquisition times of the different types of images.

As part of the data fusion process we will cross-calibrate MODIS with ATSR-2/AATSR using a nadir viewing, colocator technique recently developed as part of GMS Pathfinder (LeMarshall, Simpson and Jin, 1999). Moreover, use of high temporal resolution (every 30 minutes) GOES data greatly reduces the reliance on *a priori* information in the proposed application. Therefore, we will also combine GOES data (10 bit thermal resolution) with ATSR (12-bit thermal resolution) to examine high frequency variation (diurnal) in cloud cover and SST in climatically important regions as well as the temporal evolution of rapid snowmelt events.

V. EXAMPLES

Selected examples (atmospheric, terrestrial and oceanic) will be shown at the ATSR workshop in Frascati, Italy June 23-25, 1999. Examples will also be available for web downloading after 6 July 1999. Authorization to download these examples can be obtained by email request to admin@landlub.ucsd.edu after 6 July 1999.

VI. CONCLUSIONS

Accurate cloud, cloud shadow detection and calibration are important steps for building a stable retrieval process for geophysical products from MODIS, MERIS and ATSR data. Throughout this project we will implement the procedures cited herein to improve the overall accuracy and stability of geophysical products derived from MODIS, MERIS and ATSR data. Economies of scale allow for simultaneous processing and analysis of a set of complimentary atmospheric, terrestrial and oceanic products.

Acknowledgements

This work was funded by the Marine Life Research Group of the Scripps Institute of Oceanography and a NASA EOS Direct Broadcast (DB) grant. ESA has made data available under a data grant for the ENVISAT program Grant # A0-ID448, C. J. Koblinsky and J. J. Simpson CO-PI's

References

Allan, R., Lindesay, J. and Parker, D., 1996: El Niño Southern Oscillation and Climate Variability. Collingwood, Australia: CSIRO Publishing, 405pp.

Bailey, M. J., Carpenter, K. M., Lowther, L. R., and Passant, C. W., 1981: A mesoscale forecast for 14 Aug. 1975-The Hampstead Storm: *Meteorol. Mag*, **110**, 147-161.

Ball, G. H., and Hall, D. J., 1964: ISODATA: A novel method of data analysis and pattern classification. Menlo Park, California. Technical Report of the Stanford Research Institute, pp 61 and appendices.

Berendes, T., Sengupta, S. K., Welch, R. M, Wielicki, B. A., and Navar, M., 1992: Cumulous cloud base height estimation from high spatial resolution Landsat data: A Hough transform approach, *IEEE Trans. Geosci. Remote Sensing*, **30**, 430 - 443.

Cayan, D.R., 1996, Climate variability and snow pack in the western United States: *J. Climate*, 9, 928-948.

Cayan, D. R., and Webb, R. H., 1992: El Niño/Southern Oscillation and streamflow in the western United States. In H.F. Diaz and V. Markgraf (eds.), El Niño: Historical and Paleoclimatic Aspects of the Southern Oscillation. Cambridge: Cambridge University Press, 29-68.

Coakley, J.A., Jr. and F.P. Bretherton, 1982: Cloud cover from high resolution scanner data: Detecting and allowing for partially filled fields of view. *J. Geophys. Res.*, **87**, 4917 - 4932.

Dalu, G., 1985: Emittance effect on the remotely sensed sea surface temperature. *Int. J. Remote Sens.*, **6**, 733 - 740.

Deschamps, P.Y., and T. Phulpin, 1980: Atmospheric correction of infrared measurements of sea surface temperature using channels at 3.7, 11 and 12 µm. *Boundary-Layer Meteor.*, **18**, 131 - 143.

Dettinger, M.D., Cayan, D.R., McCabe, G.M., and Marengo, J.A., in press, Multiscale streamflow variability associated with El Niño/Southern Oscillation: Chapter 5 in Diaz, H.F., and Markgraf, V. (eds.), El Niño: Multiscale Variability and Societal Impacts, Cambridge University Press.

Duggin, M.J., 1985: Factors limiting the discrimination and quantification of terrestrial features using remotely sensed radiance. *Int. J. Remote Sens.*, **6**, 3 - 27.

Foody, G.M., 1988: The effects of viewing geometry on image classification. *Int. J. Remote Sens.*, **9**, 1909 - 1915.

Gallaudet, T.C., and J.J. Simpson, 1991: Automated cloud screening of AVHRR imagery using split-and-merge clustering. *Rem. Sens. Environ.*, **38**, 77 - 121.

Gurney, C. M., 1982: The use of contextual information to detect cumulous clouds and cloud shadows in Landsat data, *Int. J. Remote Sensing*, **3**, 51 - 62.

Harris, A.R., M.A. Saunders, J.S. Foot, K.F. Smith, and C.T. Mutlow, 1995: Improved sea surface temperature measurements from space. *Geophys. Res. Lett.*, **22**, 159-216.

Henderson-Sellers, A., 1984: *Satellite Sensing of a Cloud Atmosphere: Observing the Third Planet.* Taylor & Francis, Ltd.: Philadelphia, USA, 340 pp.

Hobbs, P.V. and A. Deepak, 1981: *Clouds: Their Formation, Optical Properties and Effects*, see Figure 2, p. 285 in contribution by G.F. Hunt, Academic Press: New York, 497 pp.

Hunt, G.E., 1982: On the sensitivity of a general circulation model climatology to changes in cloud structure and radiative properties. *Tellus*, **34**, 29-38

Jones, M.S., M.A. Saunders, and T.H. Guymer, 1996a: Reducing cloud contamination in ATSR averaged sea surface temperature data. *J. Atmos. Oceanic Tech.*, **13**, 492-506.

Jones, M.S., M.A. Saunders, and T.H. Guymer, 1996b: Global remnant cloud contamination in the Along Track Scanning Radiometer data: Source and removal. *J. Geophys. Res.*, **101**, 12141-12147.

JSC Working Group on Radiative Fluxes, 1997: *Radiation and Climate: Report of the First Session*, WCRP-10, WMO/TD-No. 235, 22 pp., and appendices.

Koblinsky, C.J., J.J. Simpson and T.D. Dickey, 1984: An off-shore eddy in the California Current System, part II: surface manifestation. *Prog. Ocean*, **13**, 51-69.

LeMarshall, J., J.J. Simpson, and Z. Jin, 1999: Satellite Calibration Using a Collocated Nadir Observation Technique: Theoretical Basis and Application to the GMS-5 Pathfinder Benchmark period. *IEEE Trans. Geosci. Rem. Sens.*, **37**, 499-507.

Liljas, E., 1987: '*Multispectral Classification of Cloud, Fog, and Haze in Remote Sensing Applications in Meteorology and Climatology*, R.V. Vaughn, Ed., NATO ASI Series, Vol. 201, 480 pp.

Lipton, A., 1993: Cloud shading retrieval and assimilation in a satellite-model coupled mesoscale analysis system, *Mon. Weath. Rev.*, **121**, 3062-3081.

London, J., 1957: '*A Study of Atmospheric Heat Balance*'. Report Contract AF19, 122-165, College of Engineering, New York University, NTIS No. 119227.

Matveev, L.T., 1984: '*Computer Processing of Remotely Sensed Images: An Introduction*'. Wiley, New York.

McMillin, L.M., 1975: Estimation of sea surface temperatures from two infrared window measurements with differential absorption, *J. Geophys. Res.*, 5113-5117.

McNider, R. T., Song, J. A., and Kidder, S. Q., 1995: Assimilation of GOES-derived solar insolation into a mesoscale model for studies of cloud shading effects, *Int. J. Remote Sensing*, **16**, 2207-2231.

Ramanathan, V., 1987: The role of earth radiation budget studies in climate and global circulation research. *J. Geophys. Res.*, **92**, 4075 - 4095.

Robinson, I.S., 1985: *Satellite oceanography: An introduction for oceanographers and remote-sensing scientists*. Ellis Horwood Limited: Chichester, England, 455pp.

Ropelewski, C. F., and Halpert, M. S., 1987: Global and regional scale precipitation patterns associated with the El Niño/Southern Oscillation (ENSO). *Monthly Weather Review*, 115, 2352-2362.

Ropelewski, C. F., and Halpert, M. S., 1989: Precipitation patterns associated with the high-index of the Southern Oscillation. *J. Climate*, 2, 268-284.

Saunders, R.W., and K.T. Kriebel, 1988: An improved method for detecting clear sky and cloudy radiances from AVHRR data. *Int. J. Remote Sens.*, **9**, 123 - 150.

Segal, M. J., Purdam, F. W., Song, J. L., Pielke, R. A., and Maher, Y., 1986: Evaluation of cloud shading effects on the generation and modification of mesoscale circulations, *Mon. Weath. Rev.*, **114**, 1202 - 1212.

Shenk, W.E., and V.V. Salomonson, 1972: A simulation study exploring the effects of sensor spatial resolution on estimates of cloud cover from satellites, *J. Appl. Meteorol.*, **12**, 214-220.

Simpson, J.J., and C. Humphrey, 1990: An automated cloud screening algorithm for daytime AVHRR imagery. *J. Geophys. Res.*, **95**, 13457 - 1348.

Simpson, J.J., and Stitt, J. R., 1998: A procedure for the detection and removal of cloud shadow from AVHRR data over land, *IEEE Trans. Geosci. Rem. Sens.*, **65**, 1 – 24.

Simpson, J.J., A.B. Schmidt, and A. Harris, 1998a: Improved Cloud Detection in Along Track Scanning Radiometer (ATSR) Data over the Ocean. *Rem. Sens. Environm.*, 65: 1-24.

Simpson, J.J., J.R. Stitt, and M. Sienko, 1998b: Improved estimates of the areal extent of snow cover from AVHRR data. *J. Hydro.*, 204: 1-23.

Simpson, J.J., T. McIntire, Z. Jin, and J. R. Stitt, 1999a: Improved cloud height estimation under arbitrary viewing and illumination conditions using AVHRR data. *Rem. Sens. Environ.*, submitted.

Simpson, J.J., T. McIntire, and M. Sienko, 1999b: An improved hybrid clustering algorithm for natural scenes. *IEEE Trans. Geosci. Rem. Sens.*, submitted.

Simpson, J.J., T. McIntire, and D. Tarkowski, 1999c: A hybrid clustering/neural network classifier for natural scenes. Manuscript in preparation.

Verstraete, M. M., and Pinty, B., 1991: The potential contribution of satellite remote sensing to understanding of arid land processes, *Vegetation*, **91**, 59 - 72.

Yan, X.-H., J.R. Schubel and D.W. Pritchard, 1990: Oceanic mixed layer depth determination from satellite data. *Rem. Sens. Environm.*, **32**, 55-74.

Yan, X.-H., L.C. Breaker, 1993: Surface circulation estimation using image processing and computer vision methods applied to sequential satellite imagery. *Photogramm. Engr. Rem. Sens.*, **59**, 407-413.

Zege, E.P., A.P. Ivanov, and I.L. Katsev, 1991: *Radiative Transfer Through a Scattering Medium*. Springer-Verlag: Berlin, Heidelberg.