Technical Report
RAL-TR-2002-017

CLRC

Assimilation Techniques for Blending GOES and POES SST Retrievals

M J Murray

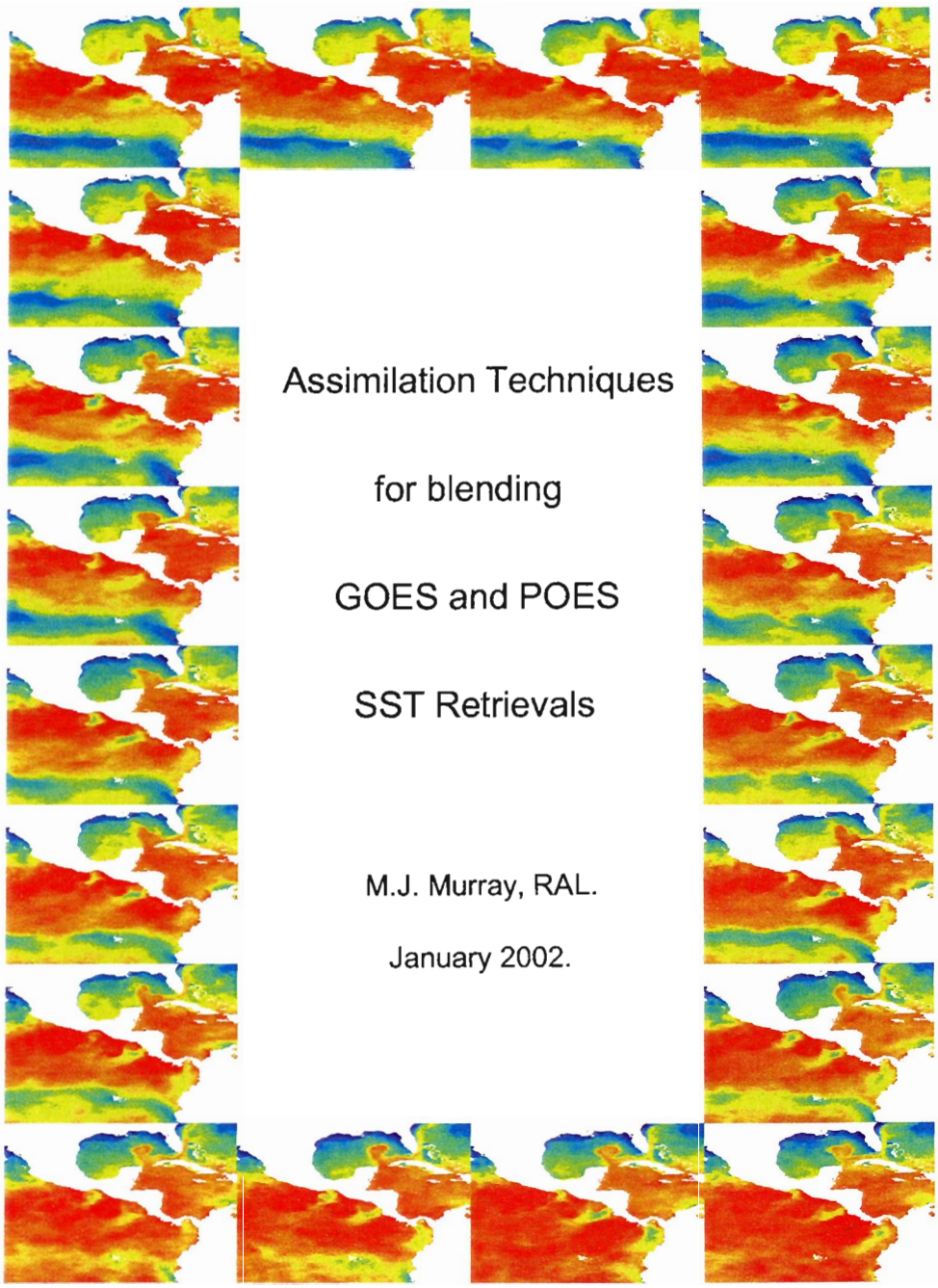
10th June 2002

COUNCIL FOR THE CENTRAL LABORATORY OF THE RESEARCH COUNCILS

Enquiries about copyright, reproduction and requests for additional copies of this report should be addressed to:

The Central Laboratory of the Research Councils
Library and Information Services
Rutherford Appleton Laboratory
Chilton
Didcot
Oxfordshire
OX11 0QX
Tel: 01235 445384 Fax: 01235 446403
E-mail library@rl.ac.uk

ISSN 1358-6254



Assimilation Techniques

for blending

GOES and POES

SST Retrievals

M.J. Murray, RAL.

January 2002.

Contents

1 Introduction **4**

2 Data Description **4**

2.1 POES Data 4

2.2 GOES Data 7

2.3 Ancillary datasets 9

3 Methodology **9**

3.1 Development of sequential assimilation scheme. 11

3.2 Determination of appropriate SST spatial correlation model. 11

3.3 Characterization of observation errors 15

3.4 Treatment of biases in SST Datasets 17

4 Results **20**

5 Conclusions and recommendations **34**

6 Appendix: Assimilation of data over time. **35**

7 References **36**

List of Figures

1 *Pathfinder SST coverage.* 6

2 *GOES SST Characteristics.* 8

3 *Multiscale Estimation on a quad tree.* 10

4 *Spatial Variability in SST gradient* 12

5 *Spatial variability of SST gradient in Caribbean* 13

6 Comparison of SST estimates generated using long and short correlation lengths. 14

7 *Observed mean and variance for December 1998 nighttime data.* 16

8 Geographical distribution and extent of diurnal warming. 18

9 *Diurnal Warming.* 19

10 *Initial 'mean' SST field* 22

11 *Static Estimation for single day.* 23

12 *Time series of SST estimates.* 24

13 *SST anomaly field evolution* 25

14 *Evolution of SST anomaly fields featuring TIWs.* 26

15 *Time-Longitude sections featuring TIW propagation* 27

16 *Buoy SST comparison.* 28
17 *Evolution of SST anomaly fields in Caribbean.* 29
18 *Comparison with Reynolds SST* 30
19 *Time evolution of errors.* 31
20 *Across-swath error propagation.* 32
21 *SST cross section across ocean basins.* 33

1 Introduction

The primary objective of this study is to develop two assimilation techniques for the blending of SST retrievals from NOAA GOES and POES platforms, and evaluate the feasibility of using these methods to generate global SST analyses. The two approaches are summarized below.

- A flexible platform to produce high-quality, high-resolution, continuous SST analyses with realistic error estimates from multiple SST datasets has recently been developed in a collaborative effort between Rutherford Appleton Laboratory and the University of Waterloo. This scheme uses a multiscale recursive estimation prodedure which is particularly suitable for the estimation of large-scale dynamic systems (Fieguth et al., 1999). Unlike conventional matrix inversion methods which scale as n^3, the multiscale method scales as n log$n$ - $n^{\frac{3}{2}}$, where n is the number of variables estimated. This is a vital consideration when considering global estimation at high spatial resolutions (*e.g* at 9km spatial resolution $\sim 4000 \times 2000$ pixels). The software is sufficiently fast to permit investigation of the statistical models used to describe the SST spatio-temporal variability and the observational errors. Unlike other multiscale methods such as wavelet decomposition, explicit treatment of inhomogeneities due to coastlines and boundary currents presents no major problems.

- An adaptive data-assimilation scheme based on an intermediate-complexity ocean model has been developed at Oxford University.

In this initial phase of the project the emphasis has been on developing the methods for use with POES data only. We have initially concentrated on a restricted data set comprising several months of data in 1998 -1999, covering the tropical Pacific and Caribbean.

2 Data Description

2.1 POES Data

The NOAA/NASA AVHRR Oceans Pathfinder SST dataset is derived from the AVHRR instruments on board the NOAA -7, -9, -11 and -14 polar orbiting satellites. The Pathfinder SST dataset currently covers the period 1985–1999 and the required data were downloaded from the website of the Physical Oceanography Distributed Active

Archive Center (PODAAC) at JPL (`www.podaac.jpl.nasa.gov`). These data were chosen as they represent an easily accessible source of POES data, and the same SST retrieval algorithm using only the brightness temperatures from AVHRR channels 4 and 5, was used for both daytime and nighttime SST retrieval (see Walton, 1988). Daily SST fields for both daytime (ascending pass) and nighttime (descending pass) are available on equal-angle grids of 4096×2048 pixels (nominally referred to as the 9km resolution). An example of the coverage associated with both the daytime and nighttime AVHRR overpasses is shown in Figure 1. These data were resampled using bilinear interpolation on to $\frac{1}{6}^{\circ}$ spatial resolution, as this is more tractable both in terms of size (2160×1080 pixels) and for ease of comparison with other datasets. These latter data were used for most of the experiments described below, although investigations of the statistics associated with the data used the 9km resolution datase. Additionally, the results of 3-month assimilation using the higher-resolution data for a region centred on Central America are shown on the front page. These cover the period 15th December 1998–14th March 1999 at 4-day intervals.

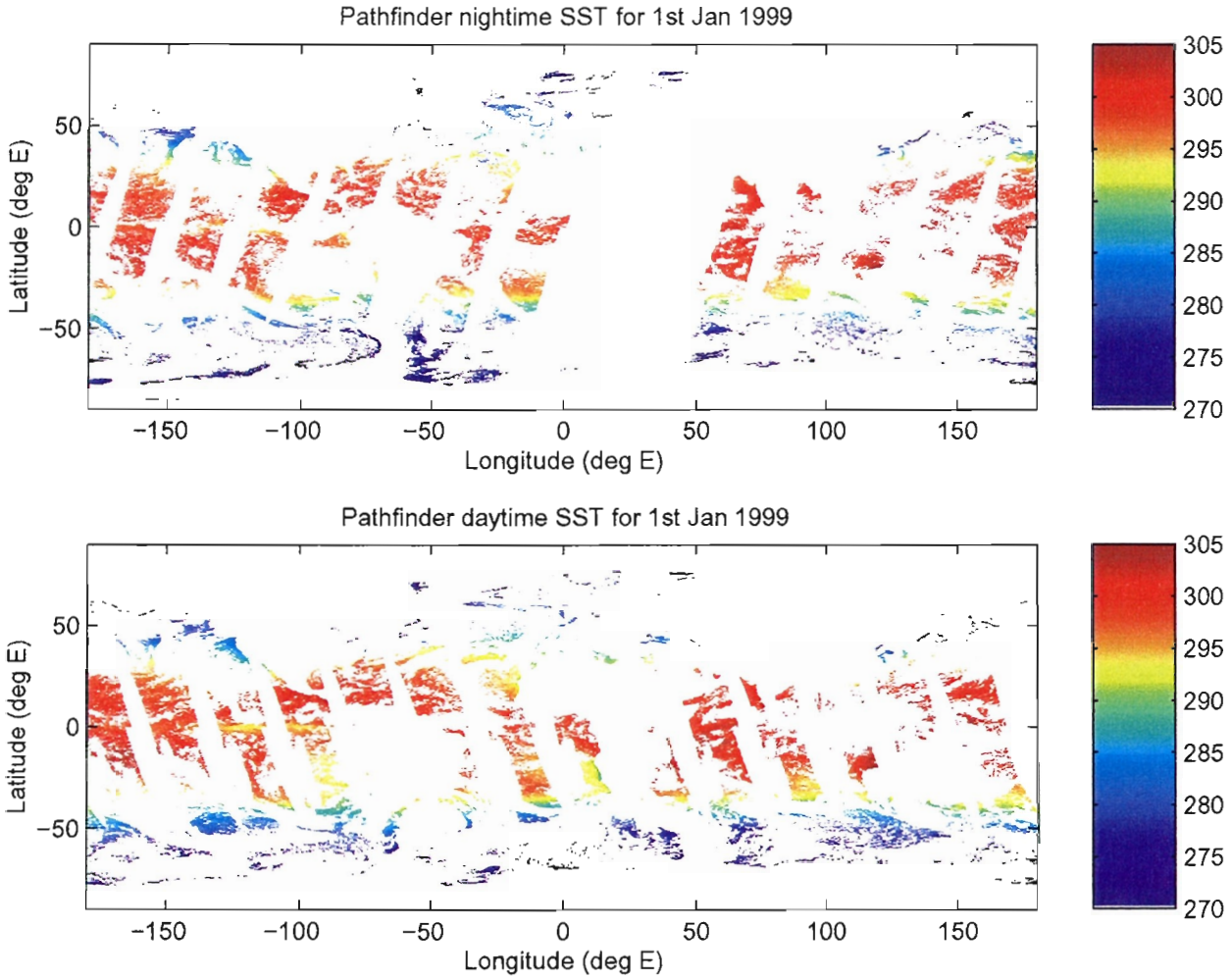


Figure 1: *Typical daily coverage for the daytime Pathfinder overpass (top panel) and the nighttime overpass (lower panel) (day shown is 1st January 1999).*

2.2 GOES Data

The National Environmental Satellite, Data and Information Service (NESDIS) has been producing hourly GOES SST estimates since December 1998. Whilst the SST retrieval quality does not match that of AVHRR retrievals, the frequent sampling associated with GOES data offers many benefits. These include improved monitoring of SST features and the diurnal warming cycle, and the potential for enhanced cloud detection. A sample GOES datset based on merged GOES-East and GOES-West has been acquired for the period August–December 2001. These are hourly data, at 1/5 degree spatial resolution and extend from 60°N – 45°S and 180°W – 30°W. Several characteristics of the GOES hourly data are illustrated in Figure 2.

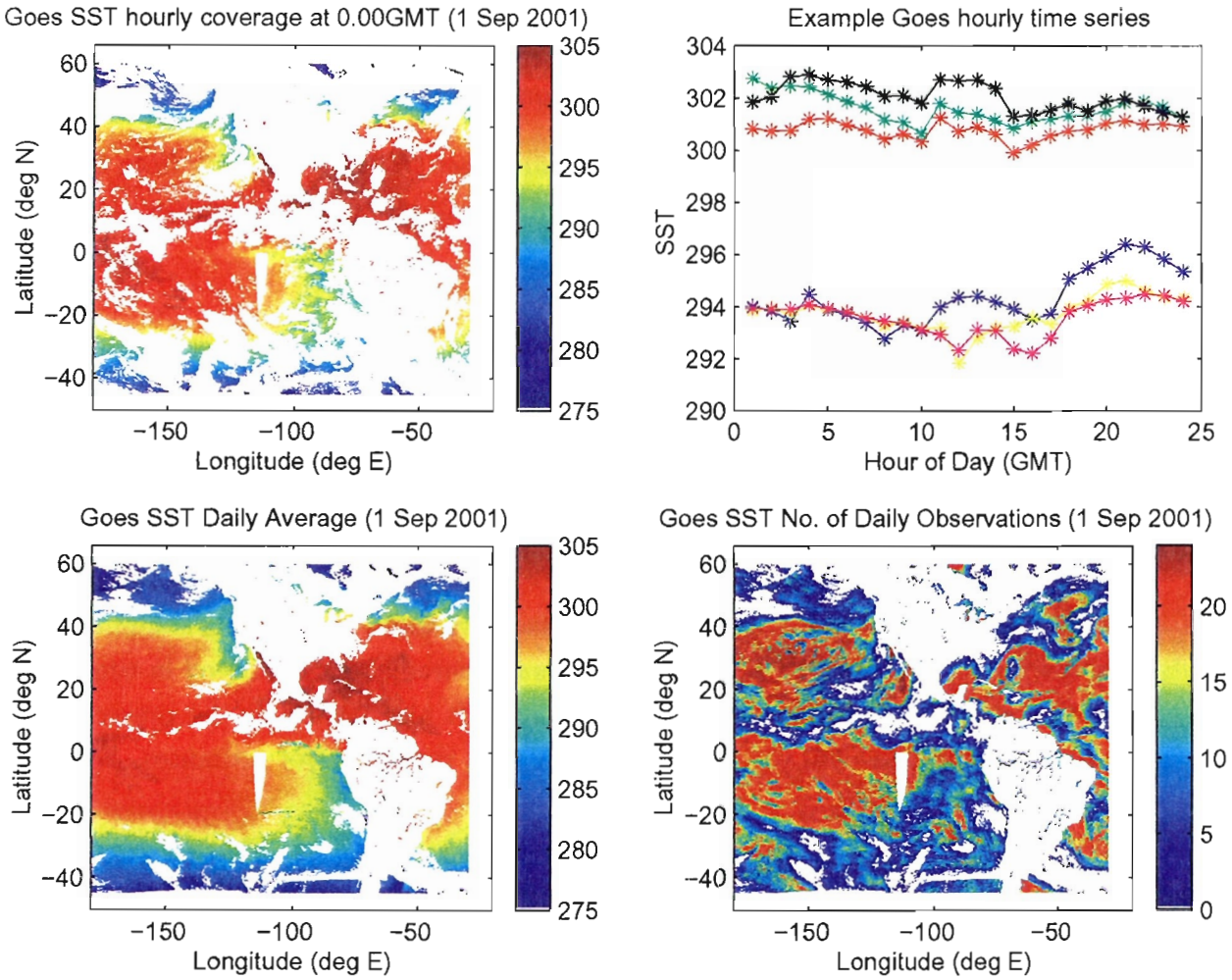


Figure 2: *Typical hourly coverage associated with GOES data is shown in the top left panel. Example hourly time series for a number of positions exhibiting a range of diurnal variation is shown in the top right panel. The coveraged associated with the daily average is shown (bottom left). Characteristic cloudy areas remain unobserved throughout the day. The number of hourly observations in the day is also shown (bottom right). All data are from 1st Septemder 2001. (Note hour of day given in x-axis is GMT, not local solar time.)*

2.3 Ancillary datasets

TRMM Microwave Instrument (TMI) SST and windspeed data at 0.5° spatial resolution were acquired from the TMI website (`www.ssmi.com`). The SST data are used for the validation and exploration of error characteristics in infrared SSTs. Windspeed data are used to correct for SST biases arising from ocean surface phenomena such as diurnal warming and the skin-bulk effect (see for example, Murray et al., 2000).

SST data from the Along Track Scanning Radiometer (ATSR), available from the ATSR website, (`atsr.rl.ac.uk`) have also been used for validation (see Murray et al., 1998a,b, Merchant et al., 1999).

3 Methodology

We describe below the dynamic estimation of the SST field using a recursive estimation algorithm which emulates the Kalman filter. This method uses a recently-developed multiscale estimation algorithm (Fieguth, 1998) for the update step, (see Figure 3) and makes simplifying assumptions about the surface dynamics leading to a computationally efficient prediction step. This approach allows measurements distributed in time to contribute to the SST estimate for a particular day. It also facilitates the preservation of high spatial and temporal resolution, together with realistic error estimates.

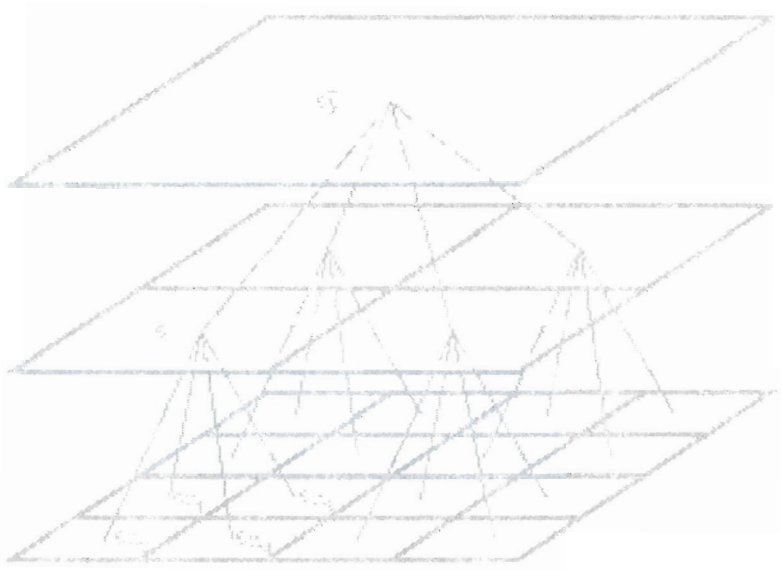


Figure 3: *The standard quad-tree hierarchical structure on which a statistical model is built. Observations are described using a hierarchical statistical structure or tree which divides into subtrees at nodes on a range of scales. For two-dimensional fields a quad-tree is used as the basis for multiscale modeling. An explicit statistical relationship is specified between those tree nodes connected by thin lines. Essentially the approach is to conditionally decorrelate the subtrees branching from each node so that these smaller subtrees can be processed independently. This requires knowledge 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.*

3.1 Development of sequential assimilation scheme.

We have developed a sequential assimilation scheme in which daily sets of SST observations are used to update the SST estimate from the previous (or subsequent) day. Following an estimate based on the initial 'mean' field as derived above, the final state from this step used as the initial guess for the next. (That is, the new SST estimate is the initial mean field with the estimated SST anomaly field added.) In this way, information is propagated in time throughout the chosen window. The difficulty is determining how to penalise the update step (i.e. choosing the background error covariance). As we have a relatively fast estimation scheme, it is possible to experiment with a variety of approaches, including making several passes through the dataset, refining the treatment of background error at each pass. In the sequential estimation attempted here, the estimated error field associated with the previous estimate is inflated by adding a positive constant and used as the input error estimate for the subsequent step (see Appendix 1). Subsequent to ascertaining a 'forward' estimate, a backwards pass can be performed by stepping back in time to the initial day. In this way observations from the entire window under consideration contribute to each day's estimate, and the final estimates should not be critically dependent on the initial mean field.

3.2 Determination of appropriate SST spatial correlation model.

The estimation method requires the field variability to be described with a single zonal and a single meridional length scale. Initially, these correlation length were estimated using a lag correlation analysis. However, the correlation lengths calculated in this way were averaged over $1^\circ \times 1^\circ$ degree boxes, leading to an underestimate of the true range of correlation length scale. An alternative method is the correlation length implied by the instantaneous gradient of the SST field. Figure 4 shows the SST estimate for 1st February 1999, and the spatial variability of the associated gradients. Two regions are characterized by much stronger than average gradients, these are the equatorial Pacific and the Gulf Stream regions, and the SST gradients associated with the latter are shown in Figure 5. A major part of the variability is associated with the zonal correlation length scale.

As shown in Figure 6, the sensitivity to correlation length in the Gulf Stream coastal region can be very large, (4K in this case). The problem is exacerbated by the persistent cloudiness, as days can pass without a clear AVHRR overpass. The value of TMI observations in addressing this problem is limited by the unavailability of microwave SSTs within $\sim$ 100km of coastlines, and the limited spatial resolution, hence the value of GOES data in allowing real observations to contribute.

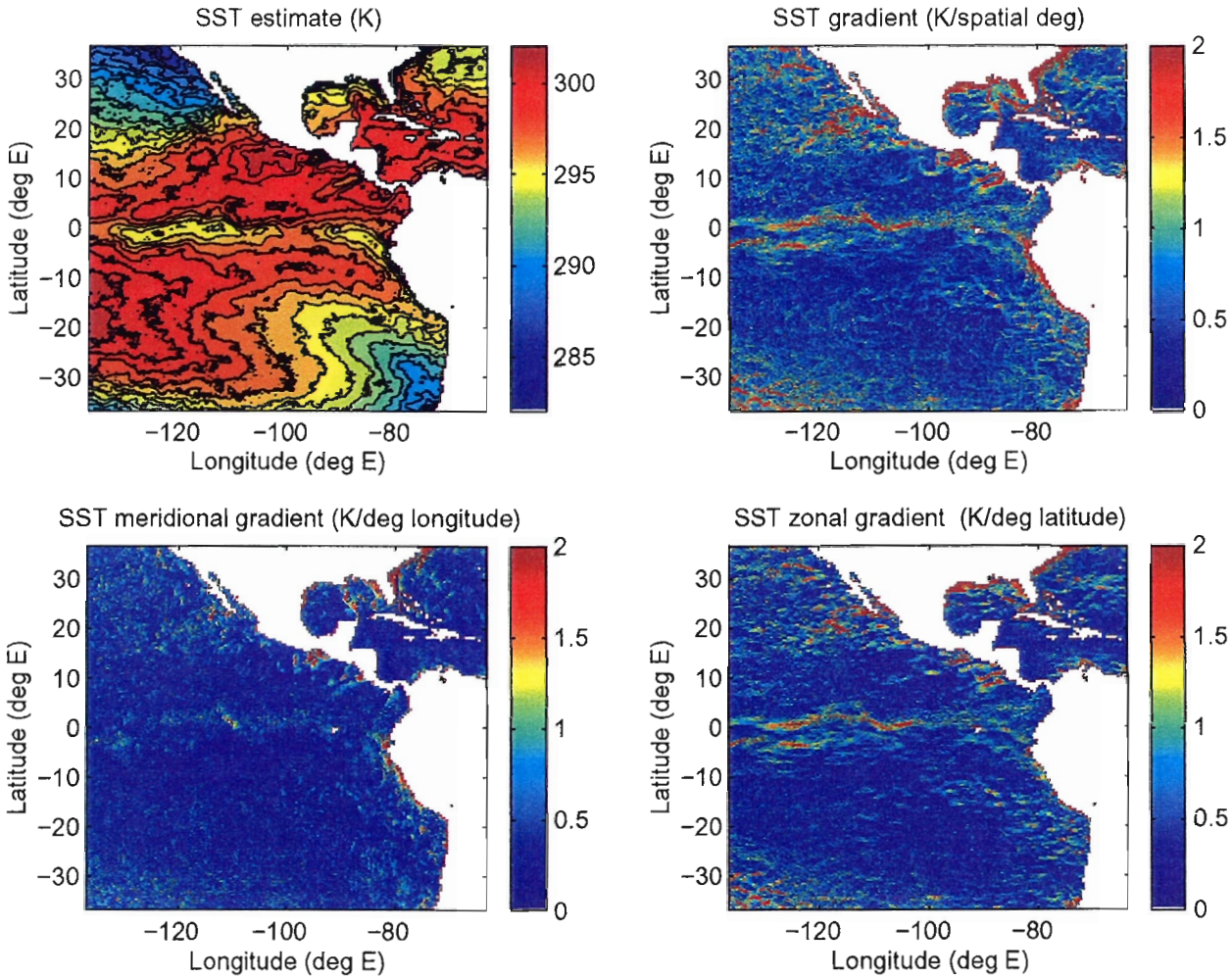


Figure 4: *Variability in SST gradient. Panel (a) shows the SST estimate for 1st February 1999. The SST gradient for the region is shown in panel (b), and the zonal and meridional components of the gradient are shown in panels (c) and (d) respectively. A major part of the variability is associated with the zonal correlation length scale. The equatorial Pacific and the Gulf Stream region are characterized by much stronger than average gradients.*

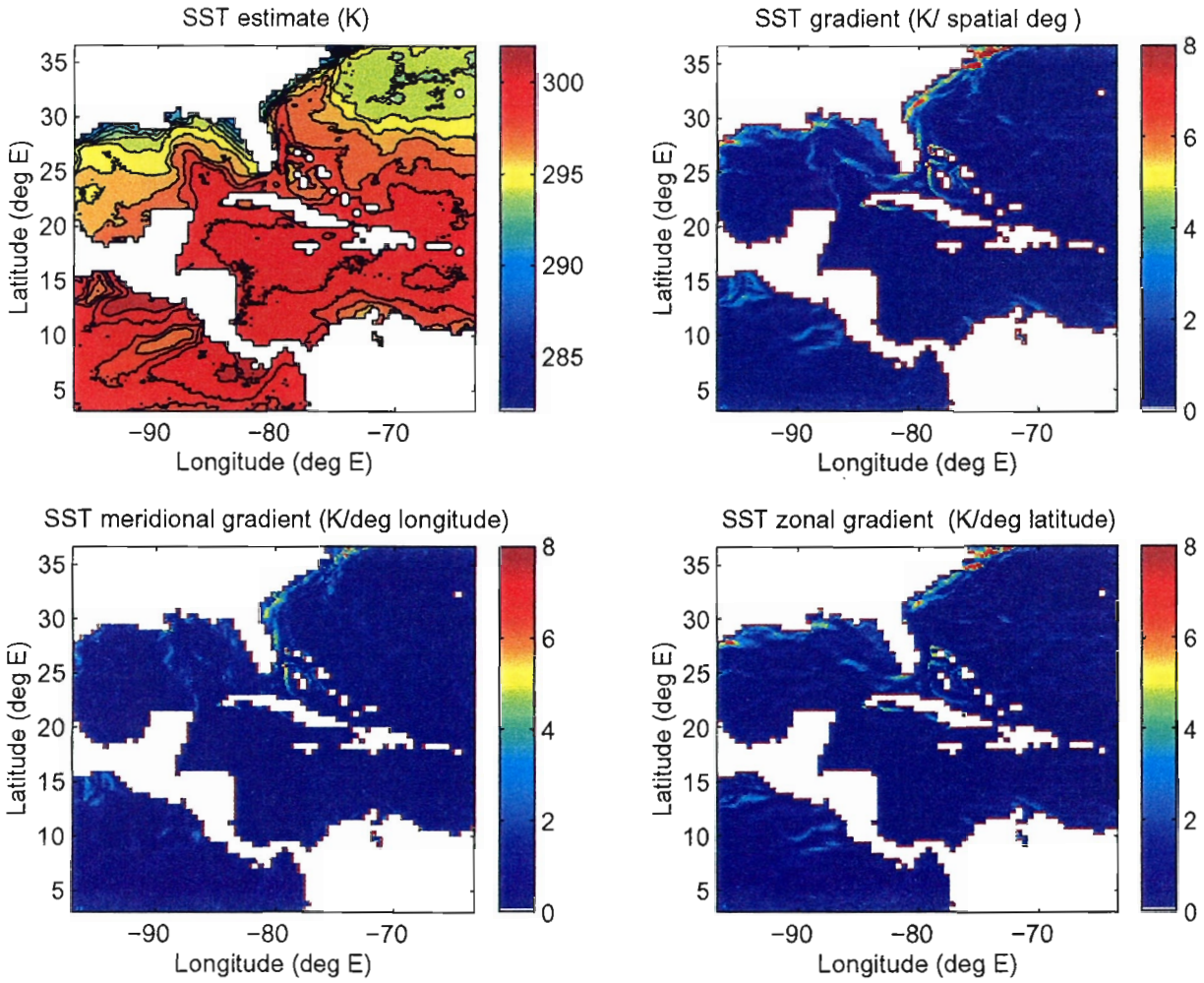


Figure 5: *Analagous to the previous figure, the variability in SST gradient associated with the Gulf Stream region is shown.*

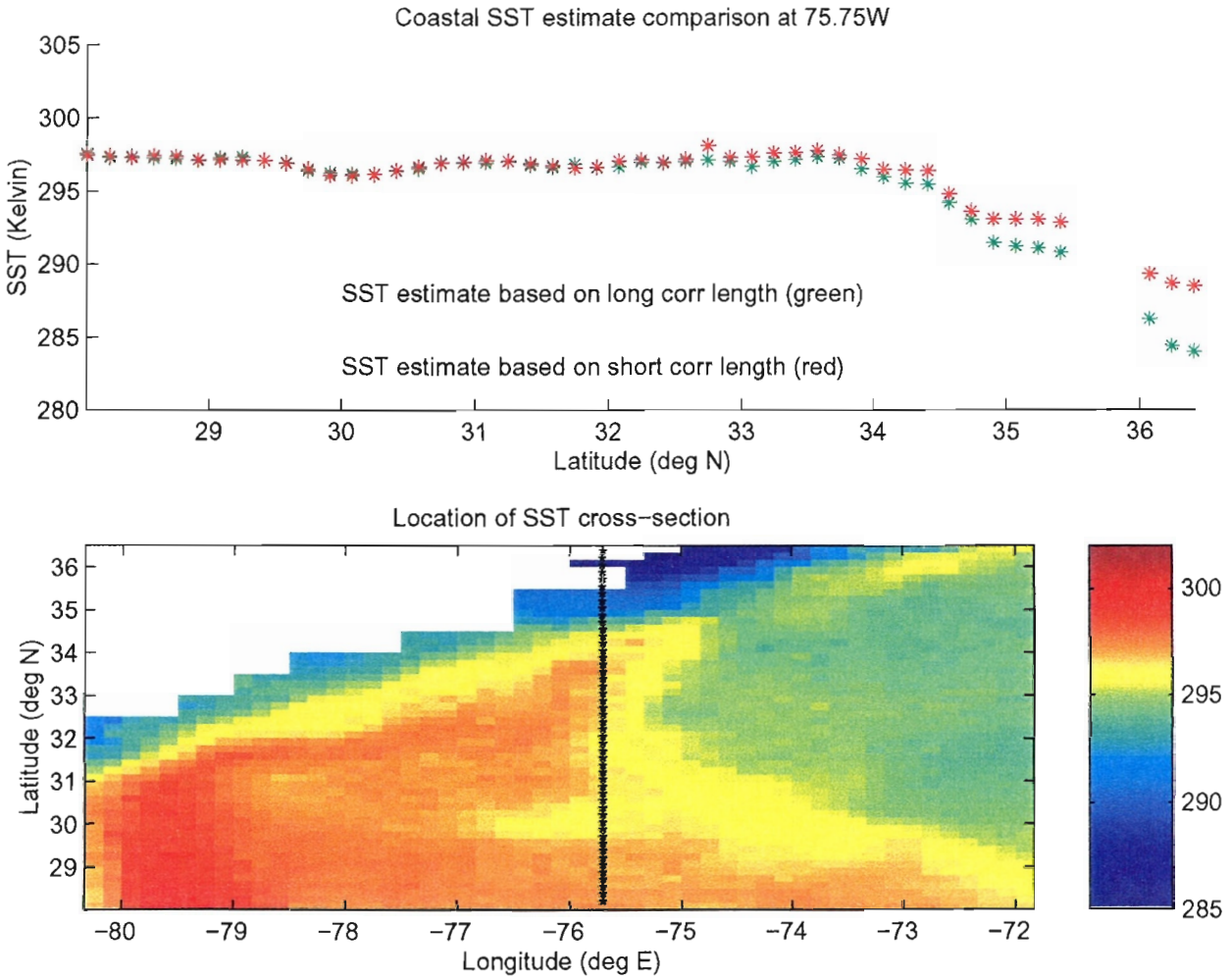


Figure 6: *Comparison of SST estimates generated using long and short correlation lengths. The region shown is a cross-section at 75.75W, which extends south from the N. Carolina coast. In the coastal region, which is characterised by strong SST gradients, the choice of correlation length is critical, and in the absence of SST observations for several days, the deviation between SST estimates can exceed 3K. The precise location of the chosen SST cross-section is indicated by the line of asterixes in the lower panel.*

3.3 Characterization of observation errors

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 observational errors associated with each dataset. The temporal and geographic variation of observational errors has been estimated by empirical assessment of lag correlations in the datasets, and by looking at the minimum variation in SST characteristic of quiescent patches of sea. Comparison with other SST datasets suggests the Pathfinder SSTs are subject to significant zonal bias. Clearly, it is appropriate to use a spatially-varying initial observational noise estimate. However in the results given below a single observation error (0.3K) has been assigned as initial SST noise estimate in a series. This choice simplified the interpretation of the propagation of error information in the sequential data assimilation described below.

The observed mean and variance in night-only SST measurements based on comparison with a smoothly varying mean SST are shown in Figure 7. Active regions display higher variance, and are thus associated with shorter temporal correlation length scales. However, as in the case with spatial length scales, to facilitate interpretation, a single value is used to characterize the entire region of interest.

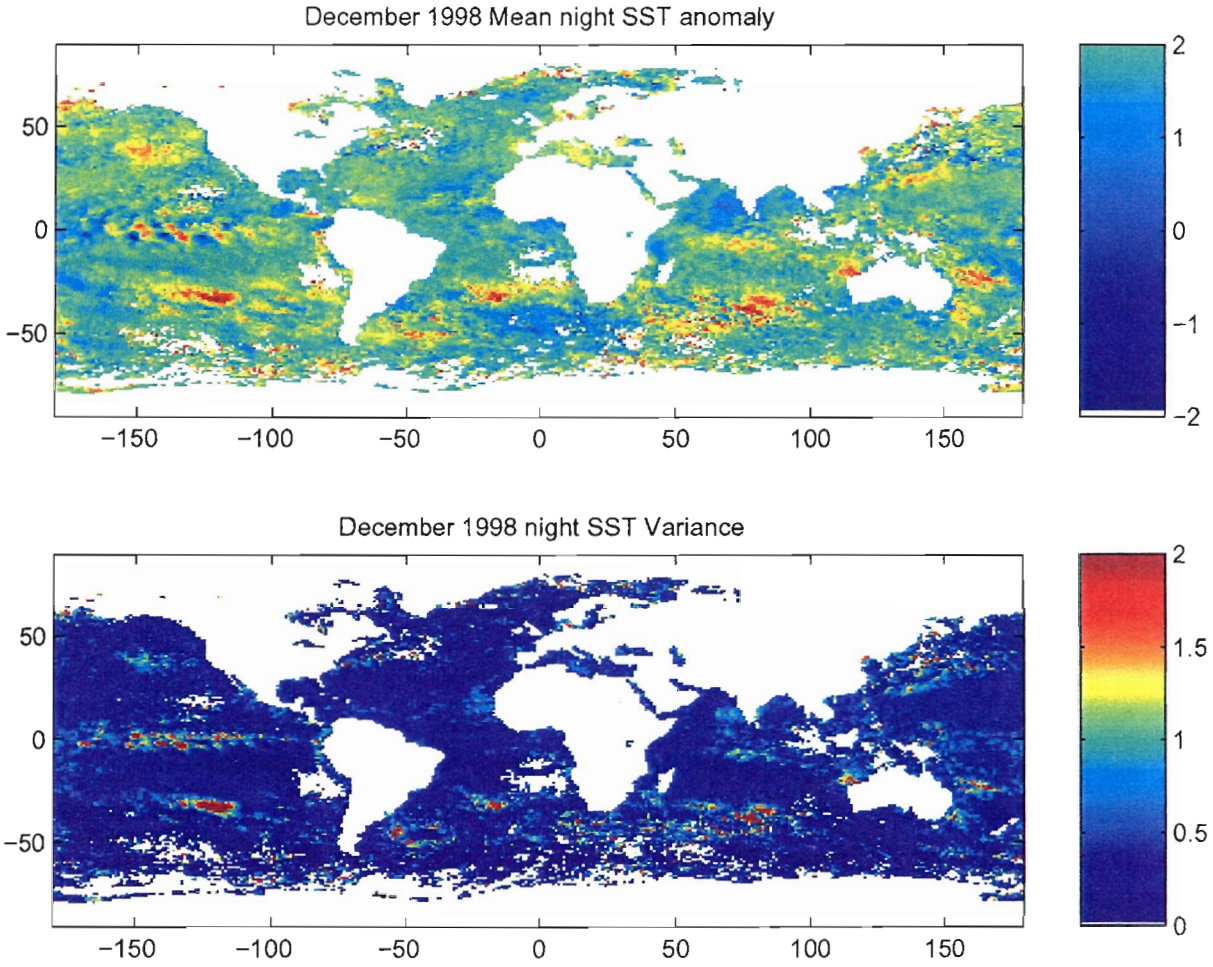


Figure.7: *Observed mean and variance for nighttime SST in December 1998.*

3.4 Treatment of biases in SST Datasets

SST observations from different sources are likely to be subject to spatially-correlated differences. For example, in the case of Pathfinder SST data, daytime observations exhibit a spatially-varying diurnal warming with respect to nighttime data. Other sources of such correlated bias between satellite SST measurements include retrieval sensitivity to water vapour, atmospheric aerosol and cloud contamination. The thermal structure of the near-surface ocean complicates the comparison of remotely sensed 'skin' SSTs to the buoy SSTs which are typically measured at a depth of $\sim$ 1m. In order to facilitate exploration and treatment of these effects, the estimation software has been modified such that the estimator treats one set of measurements as representative of the underlying SST field plus an (uncorrelated) observational noise term; the second set of measurements also represents the underlying SST field plus a suitable (uncorrelated) noise term but also a spatially-correlated term associated with the differences between the two sets of measurements.

For example, in the case of Pathfinder daytime and nighttime data, we chose to treat the nighttime SST as the field to be estimated, with daytime SST measurements considered as comprising an extra, spatially-correlated term associated with the diurnal warming signal. That is:
Night SST measurement = night SST + uncorr. noise
Day SST measurement = night SST + correlated noise + uncorr. noise

An estimate of the variance between the day and night measurement fields and an isotropic correlation length scale must be attributed to this correlated noise term. The latter should be representative of the scales associated with the processes contributing to the correlated difference term. A value of $\sim$ 900km was chosen arbitrarily for these experiments. This information is used to generate an estimate of the night SST, together with an error estimate, and the estimated field between the two sets of measurement. Figure 8 shows a composite of the observed difference between the Pathfinder night and day SST observations for June 1999, together with the associated SSMI windspeed data. Figure 9 shows the difference plotted as a function of windspeed for a range of latitude regions and two times of year. The latter two figures illustrates the possibility of developing day-night bias correction in order that the daytime observations can be allowed to have a greater effect on the SST estimates.

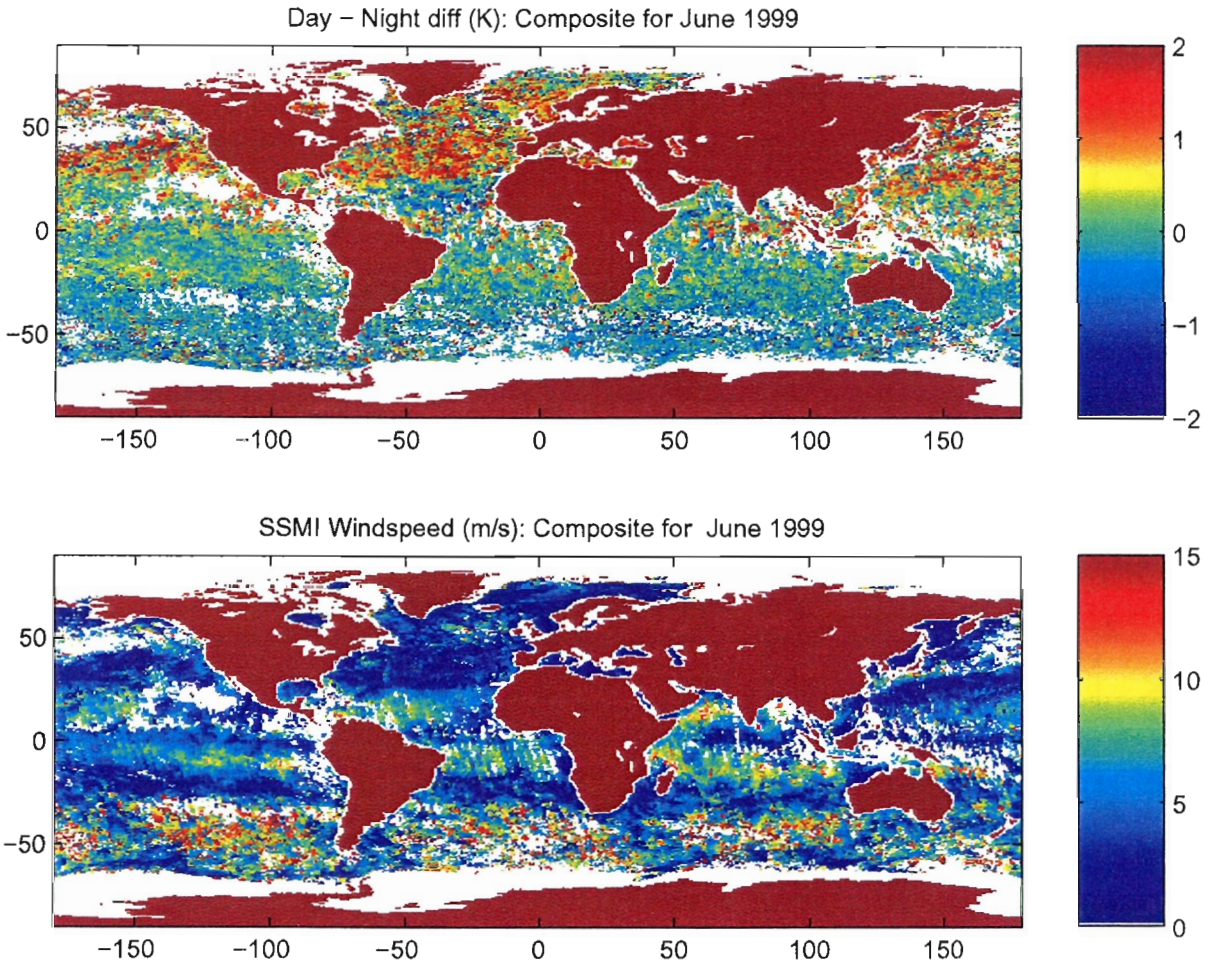


Figure 8: *The top panel shows a composite of the difference field between the night and day SST observations for June 1999. As the same Pathfinder SST retrieval algorithm is used for both day and night retrievals, this difference field is mainly attributable to diurnal warming, although differences in day/night cloud contamination may also feature. For each pixel the most positive difference value observed during the month is shown. The expected association between low windspeed and large diurnal excursion is observed.*

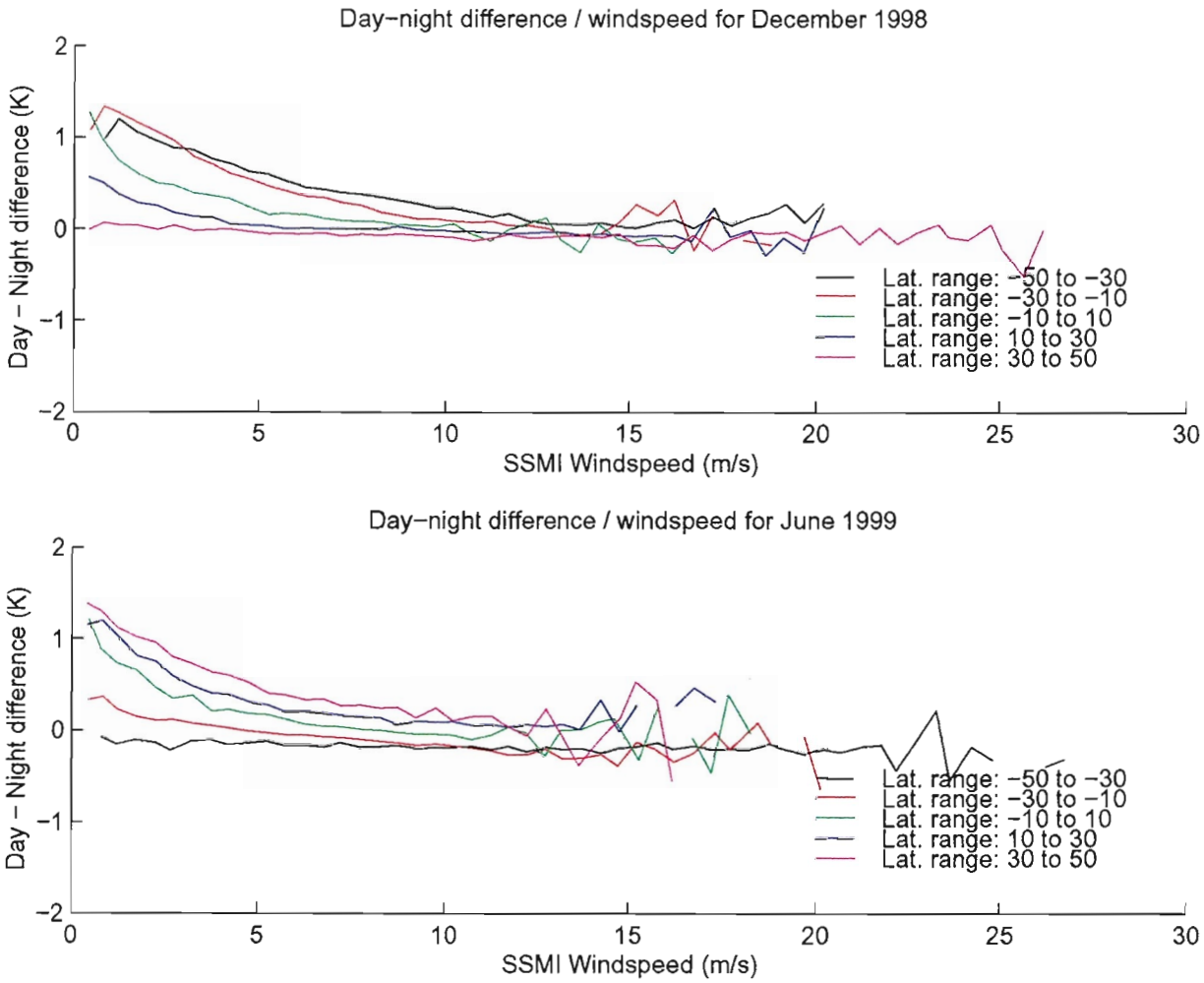


Figure 9: *Diurnal warming is shown for a range of latitude regions for both December 1998 (left) and June 1999 (right).*

4 Results

The sequential estimation scheme was used to generate SST estimates for selected geographical regions for the period 15th December 1998 – 1 May 1999. The major focus was on the area covering the Eastern Tropical Pacific basin and Caribbean Sea (35S-35N, 140W-70W). This region was chosen as it includes the GOES SST data coverage and has good-quality ancillary *in situ* datasets for use in validation.

In order to assess the sensitivity of the scheme to these issues, the SST estimates discussed below have been based on different combinations of Pathfinder night and day SSTs, different characterizations of both the spatial and temporal variability of the underlying SST field, and different attributions of observational error.

Figure 10 shows the 'mean' field centred on 15th December 1998 used as the initial SST estimate for this date. Figure 11 shows the SST estimate which results from using a single day of Pathfinder nighttime SST observations to modify this initial guess.

SST estimates for the first two weeks are shown in Figure 12 with the associated cumulative anomaly fields shown in Figure 13. Both the SST estimates and the evolution of the anomalies appear physically realistic.

Figure 14 focuses on the equatorial Pacific and the modifications occasioned by the activity of Tropical Instability Waves (TIWs) can be clearly seen. The evolution of these features in the SST estimate during a three-month period is shown in Figure 15 together with the contributing day and night SST observations. Figure 16 shows the good agreement with in situ buoy SSTs.

The evolution of small-scale structures in the Caribbean is shown in Figure 17. Figure 18 shows three SST estimates at weekly intervals for this region, together with the Reynolds weekly SST centred on the same dates (see Reynolds and Smith, 1994). This comparison highlights the preservation of fine-scale structure in the current scheme, relative to the smoothing associated with the Reynold's OI methodology.

Adequate propagation of error information in the estimation scheme relies not only on correct attribution of observational errors and the spatial variability model, but also on choosing the appropriate temporal decorrelation factor. As described above, this is incorporated by inflating the error estimates from the previous step by a constant additive factor. The time variability of SST varies geographically but in the experiments reported here a single constant was used to simplify interpretation of the error propagation. Figures 19 and 20 show the propagation of errors associated with the first few estimation steps. The potential for reducing errors in SST estimates using the sequential assimilation scheme is clear. However, these results are only representative of the nature of

the error propagation, since the temporal decorrelation factor used is arbitrary. Indeed, the results may underestimate the error reduction associated with quiescent regions, but smaller improvements in error would result in more active SST regions, where a larger temporal decorrelation factor is appropriate.

Figure 21 (red line) shows a cross section across the estimated SST field at 14.5N. Note that the decoupling of the Pacific and Caribbean basins avoids the smearing across basin boundaries which occurs in the Reynolds interpolation scheme (green line).

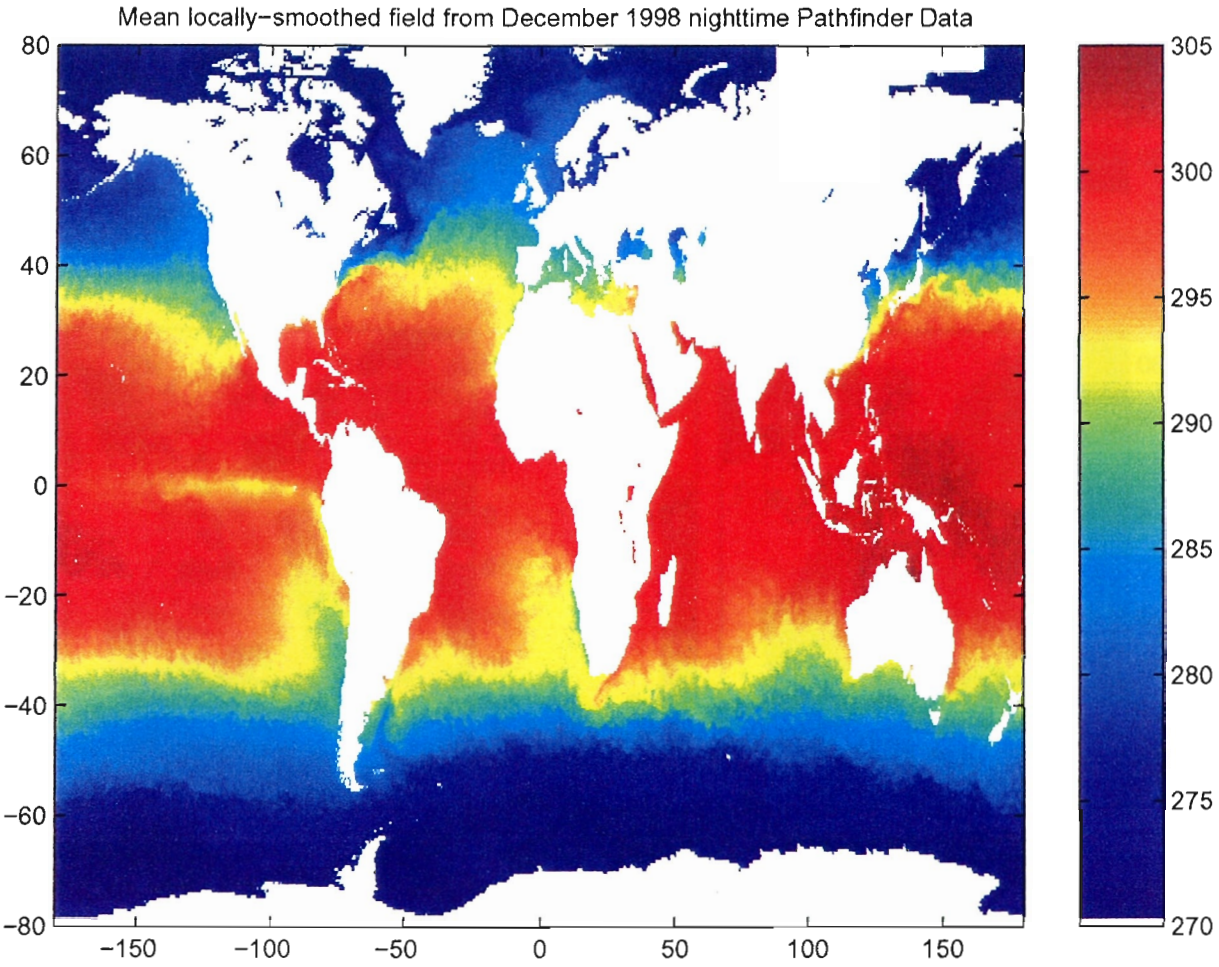


Figure 10: *Initial mean smoothed field for 15th December 1998. This was generated by performing a weighted mean of all December 1998 Pathfinder SST nighttime observations, and performing local smoothing to create a smooth and continuous initial estimate.*

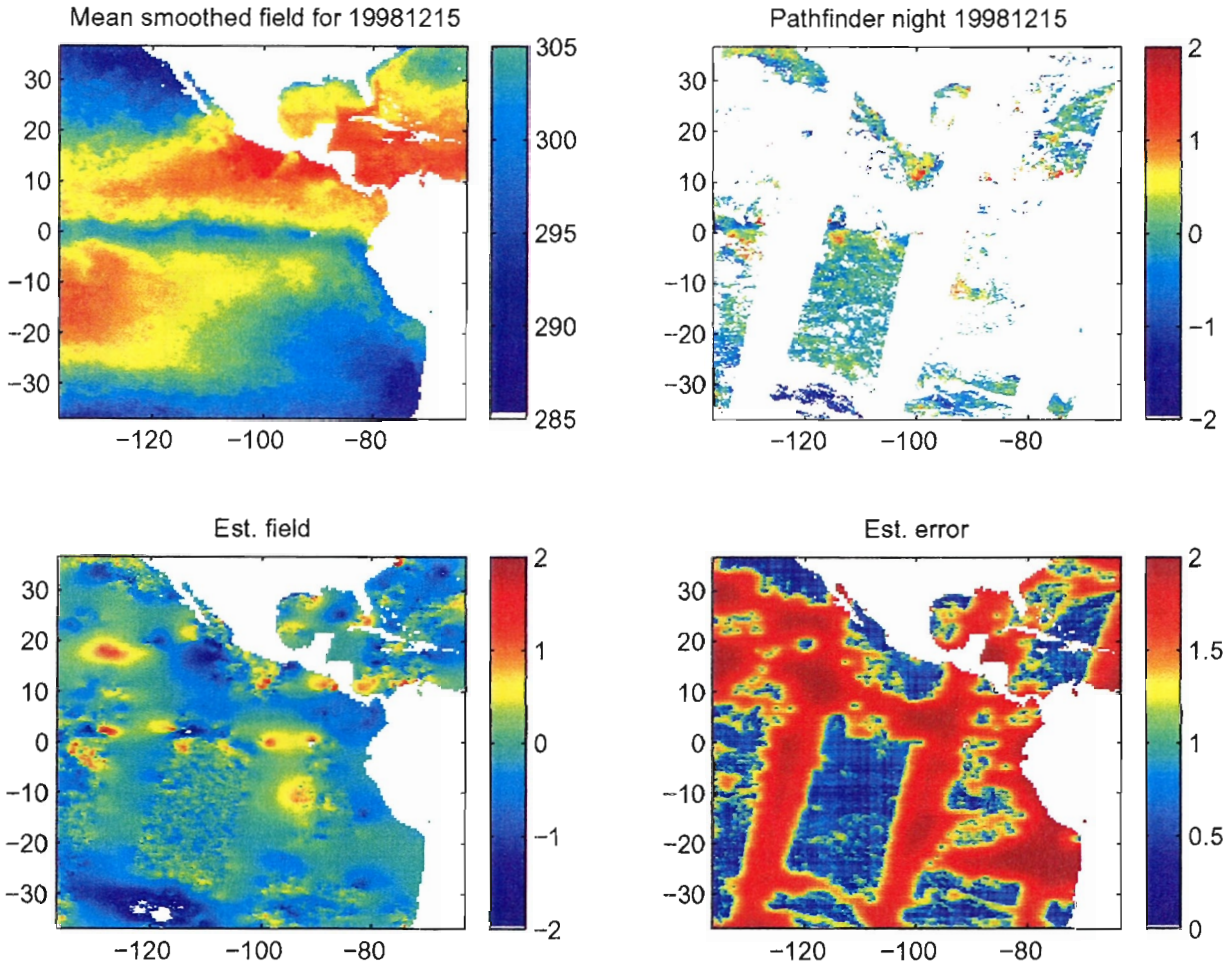


Figure 11: *Initial input data and estimated SST anomaly field and error for 15th December 1998. Panel (a) shows the initial mean field for the selected region, together with (b) a set of mean-removed Pathfinder nighttime observations for 15th December 1998. Panel (c) shows the estimated SST anomaly field (assuming a $1/f$ spatial decorrelation with a meridional correlation length of 200km and a zonal correlation length of 360km). Interpretation of this initial estimate is trivial as only local effects based on the warm and cold anomalies in the input data are apparent, and the estimate relaxes to zero in the data gaps. The estimated error associated with the SST estimate is shown in panel (d). As might be expected, the estimated error is approximately equal to the observational error attributed to the input data (0.3K) where close to measurements, but is much larger in data gaps. The rectangular artifacts apparent in the error estimate are due to approximations introduced by the multiscale models. Several possibilities for dealing with this issue are currently under consideration.*

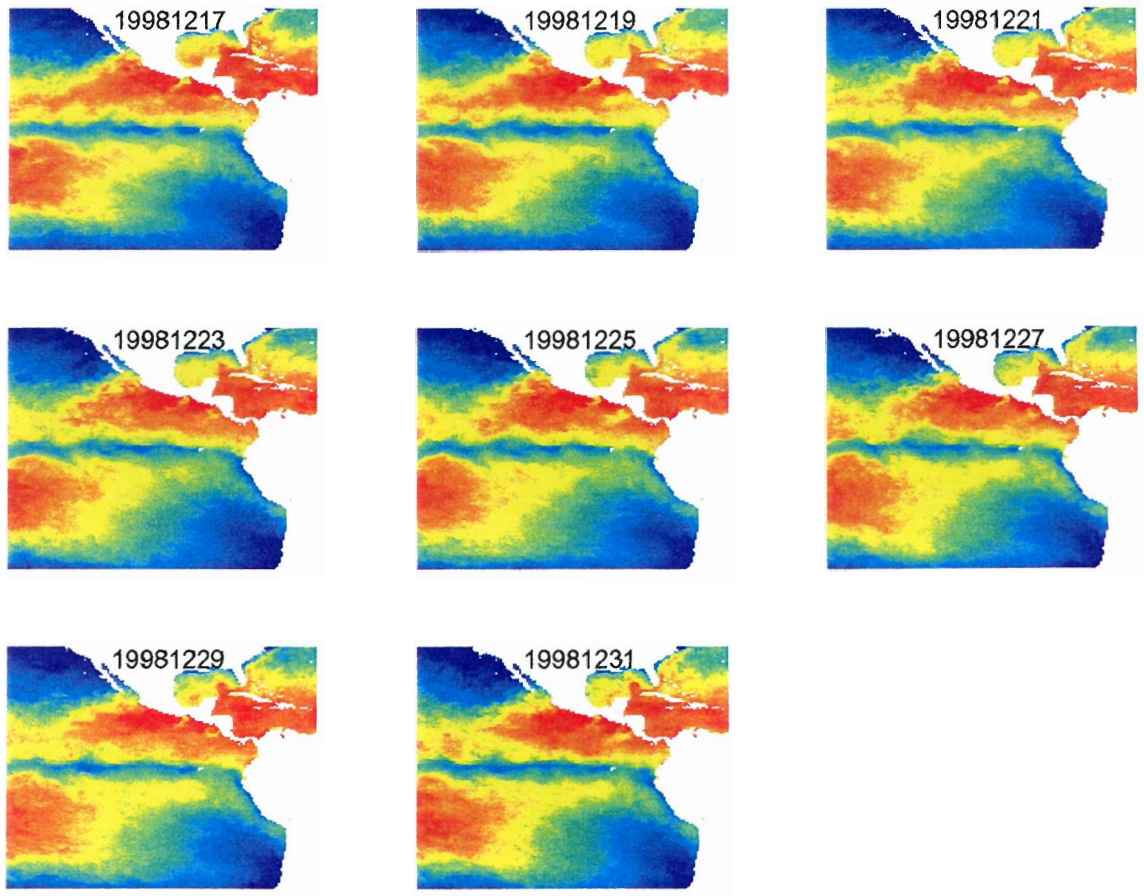


Figure 12: *SST estimates at 2-day intervals for the period 15th-31st December 1998. (Dates are given in the form yyyymmdd in the plot labels, e.g. 19981215 represents 15th December 1998.)*

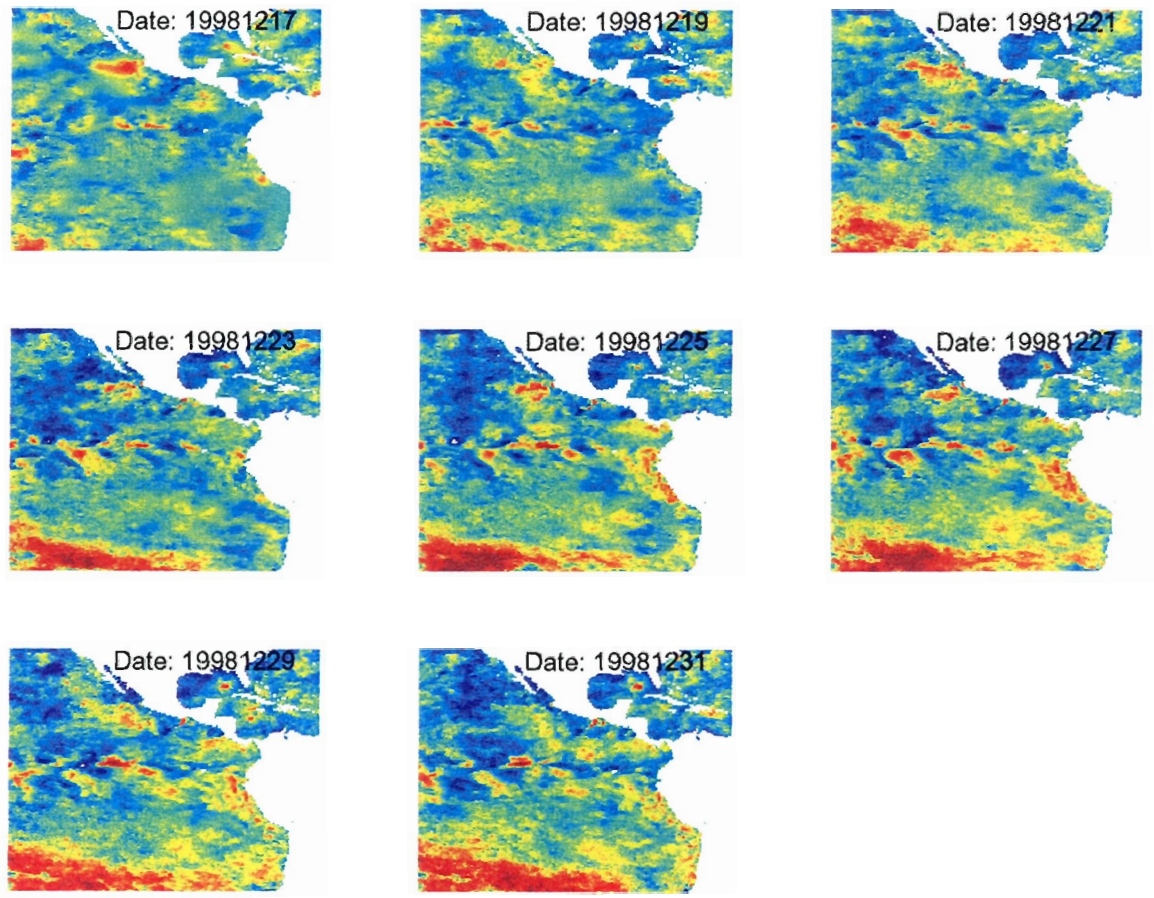


Figure 13: *The evolving anomaly fields (with respect to the initial mean field) are shown at 2-day intervals for the period 15th-31st December 1998. The evolution of the anomalies appears to be physically realistic. Apart from the large-scale features such as the warming evident in the south of the region, there are interesting smaller-scale features, particulary in the equatorial Pacific and Caribbean Sea.*

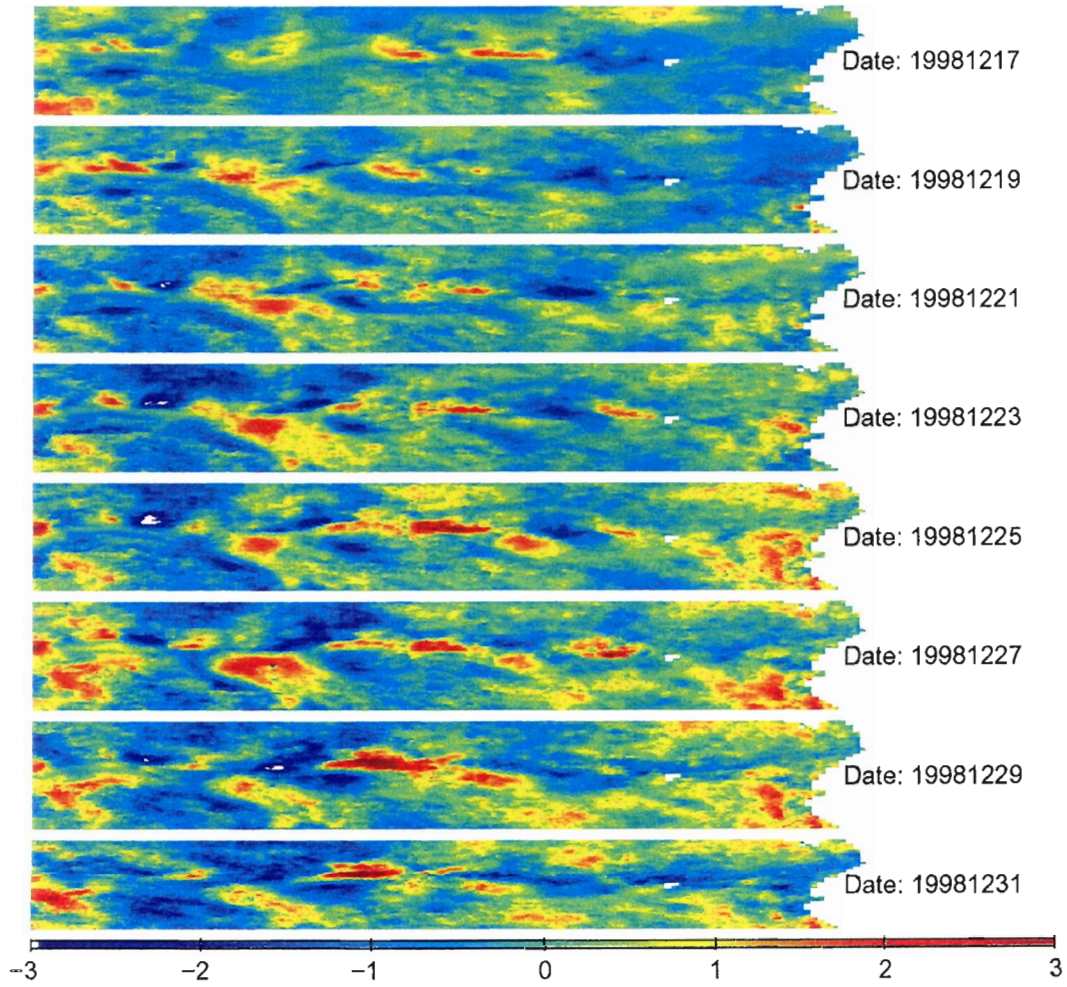


Figure 14: *Plots show the developing SST anomaly field in the equatorial Pacific. Plots are at two-day intervals in the date range 17th-31st December 1998.*

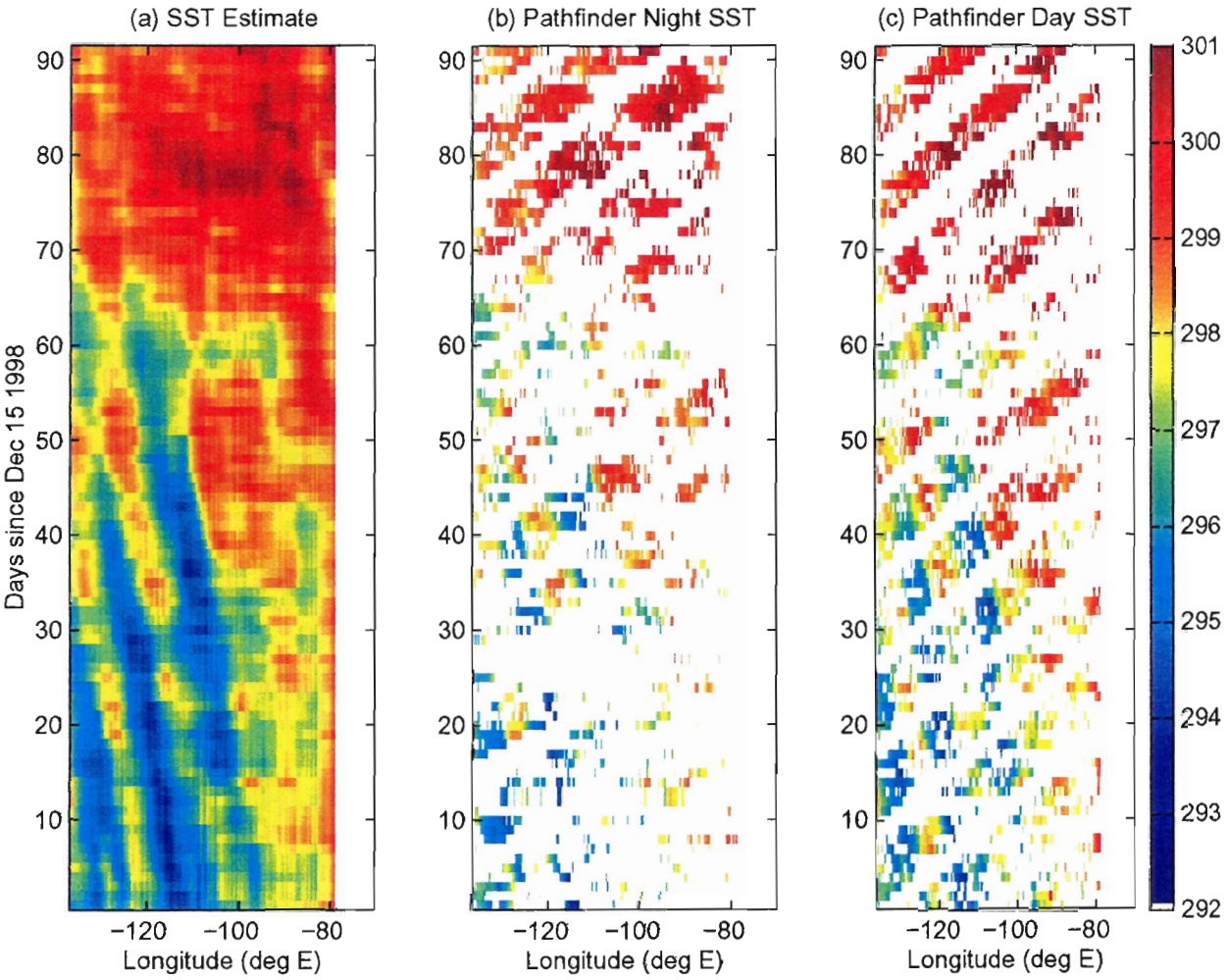


Figure 15: *Time-Longitude section across the Eastern Pacific basin at 2N for the period December 1998 - March 1999. Estimated SST is shown in the first panel with the Pathfinder nighttime and daytime observations shown in panels (b) and (c) respectively. Prominent westward-propagating waves are visible for the first two months of the datset. The continuous estimate permits accurate estimation of the velocity of these TIWs.*

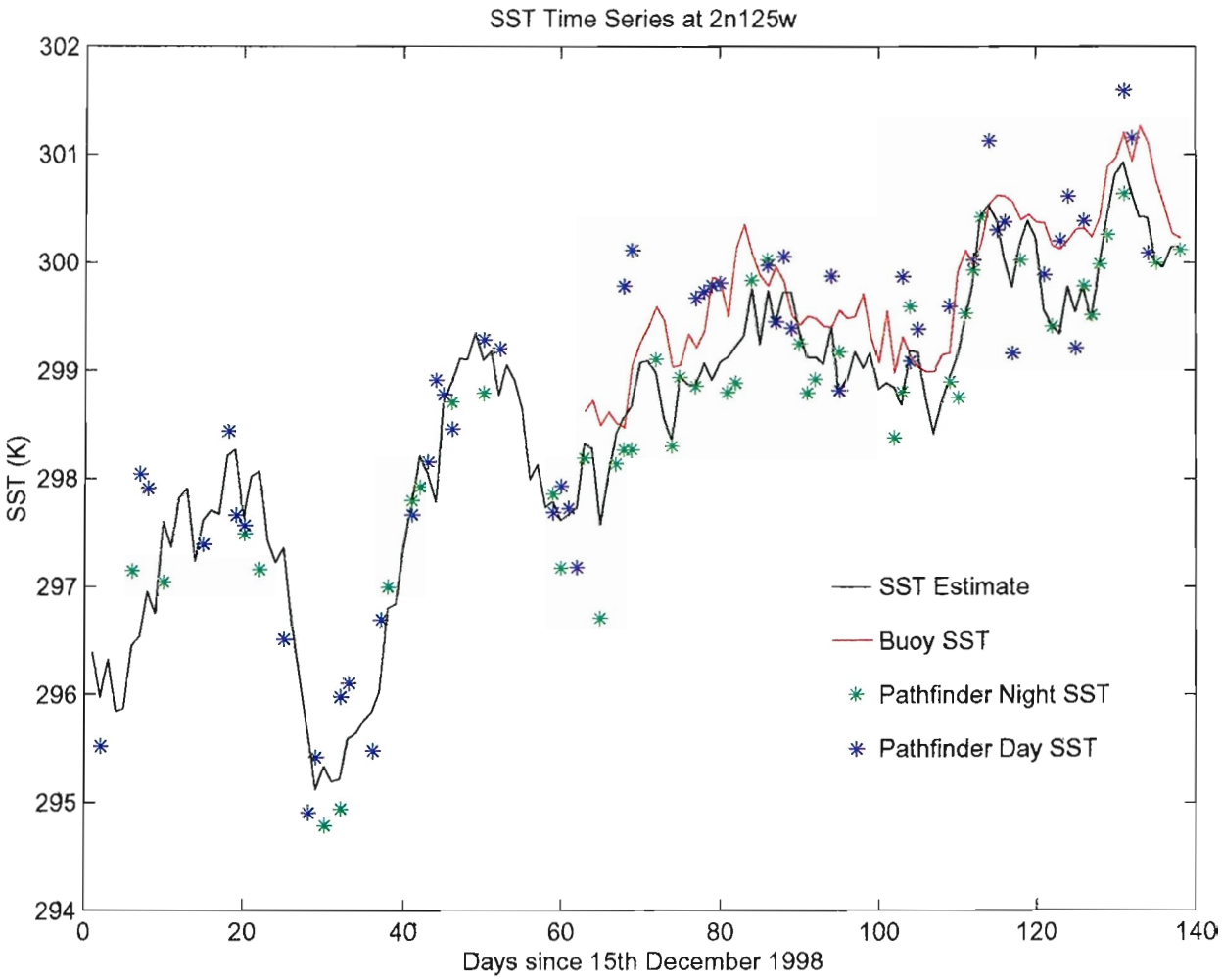


Figure 16: *Comparison of SST estimate with TAO Buoy SST at 2N 125W.*

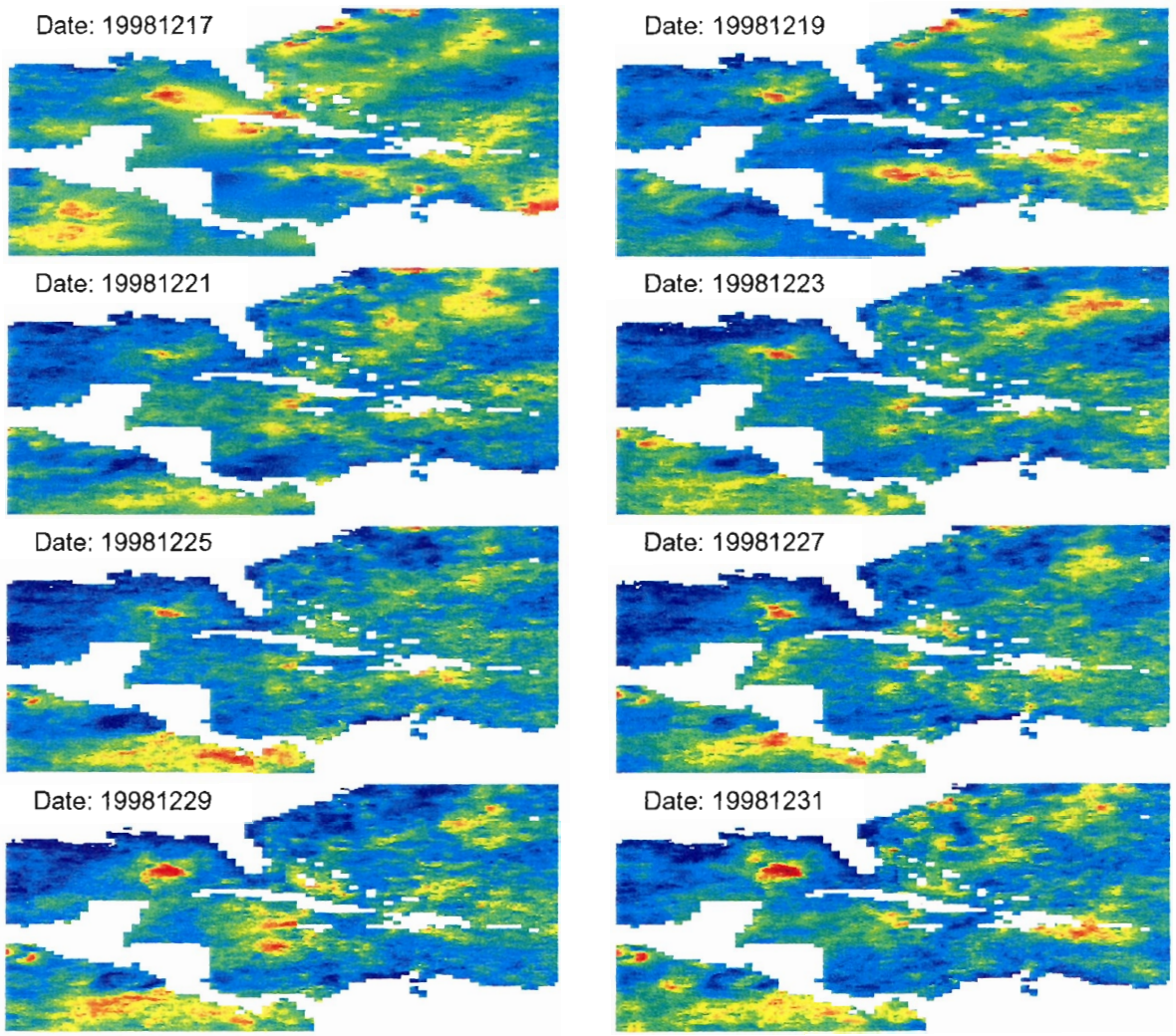


Figure 17: *Plots show the developing SST anomaly field in the Caribbean. Plots are at two-day intervals in the date range 17th-31st December 1998.*

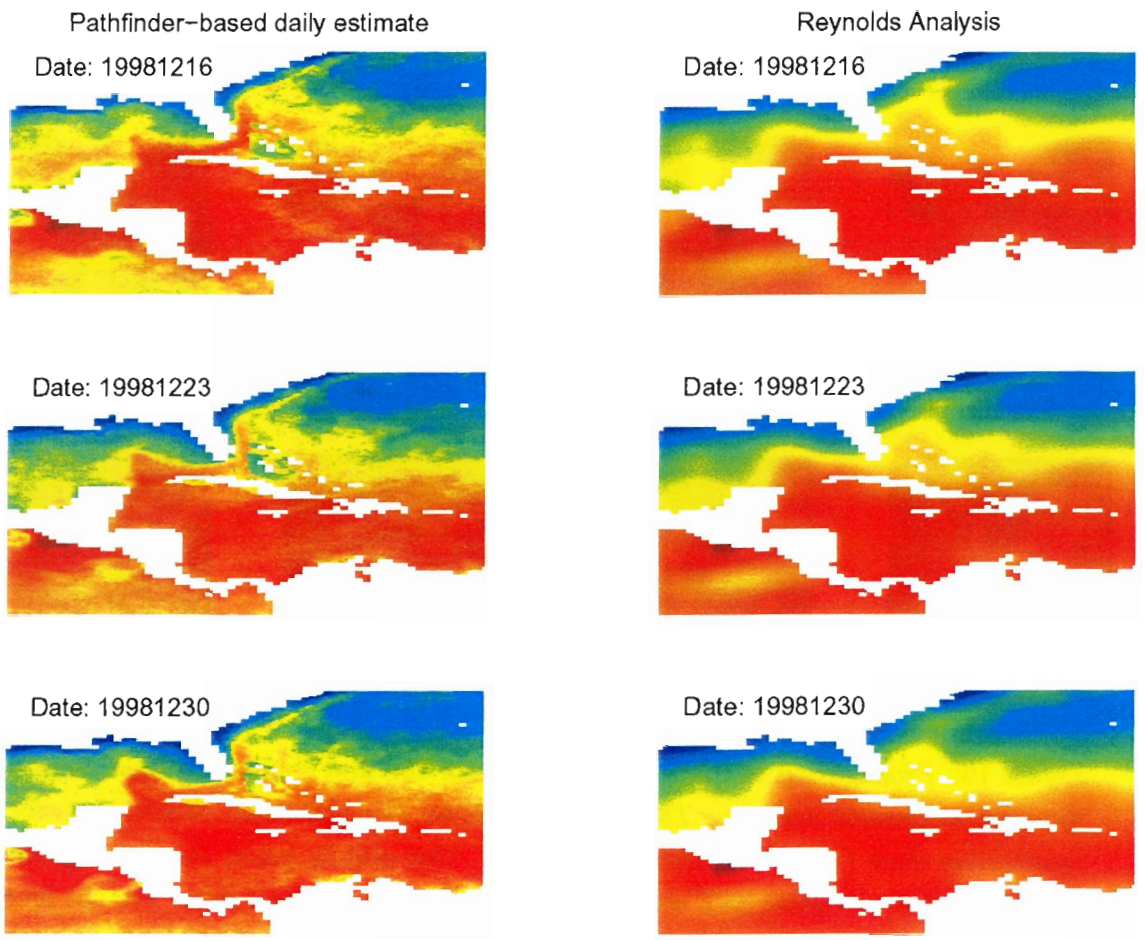


Figure 18: *Three daily SST estimates at weekly intervals (left hand panels) are compared with the weekly Reynolds SST analyses centred on the same dates.*

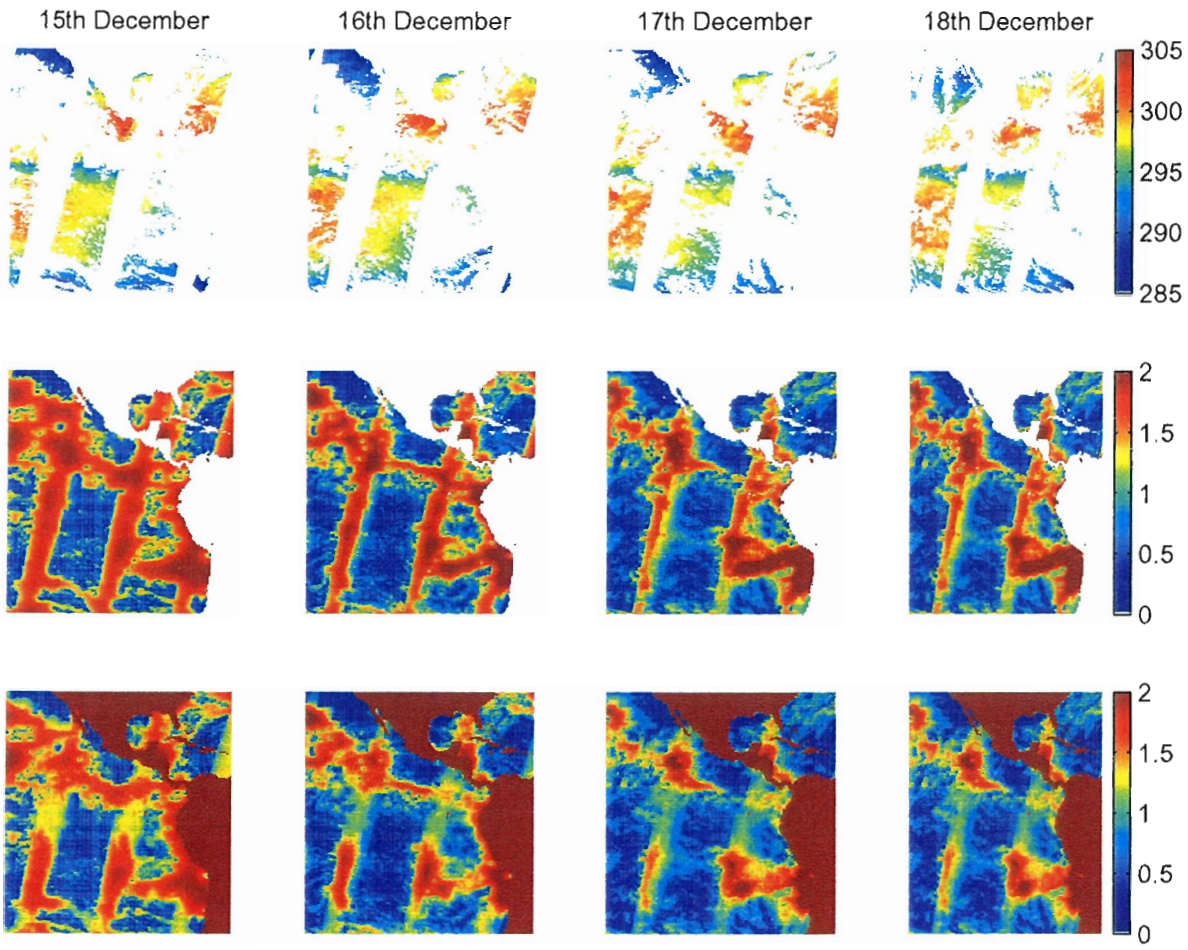


Figure 19: *Propagation of errors associated with the first four estimation steps of the dynamic estimation scheme. The top panel shows the nighttime observations for four consecutive days (15th-18th December 1998). The middle panel shows the estimated error for four consecutive SST estimates based only on nighttime data. For the first day, (middle left), the estimated error is small within the swath but rapidly increases away from observations as might be expected. However for the second and subsequent days, it is apparent that the error estimate beyond the western edges of the swaths is reduced as this area is associated with observations from the previous day. The bottom panel show the error estimates for the case where both daytime and nighttime data contribute to the SST estimate; these show areas of slightly reduced error coincident with daytime observations (which are not shown). The influence of daytime data in reducing error is much less than nighttime data due to the extra variance term associated with the former to allow for diurnal warming.*

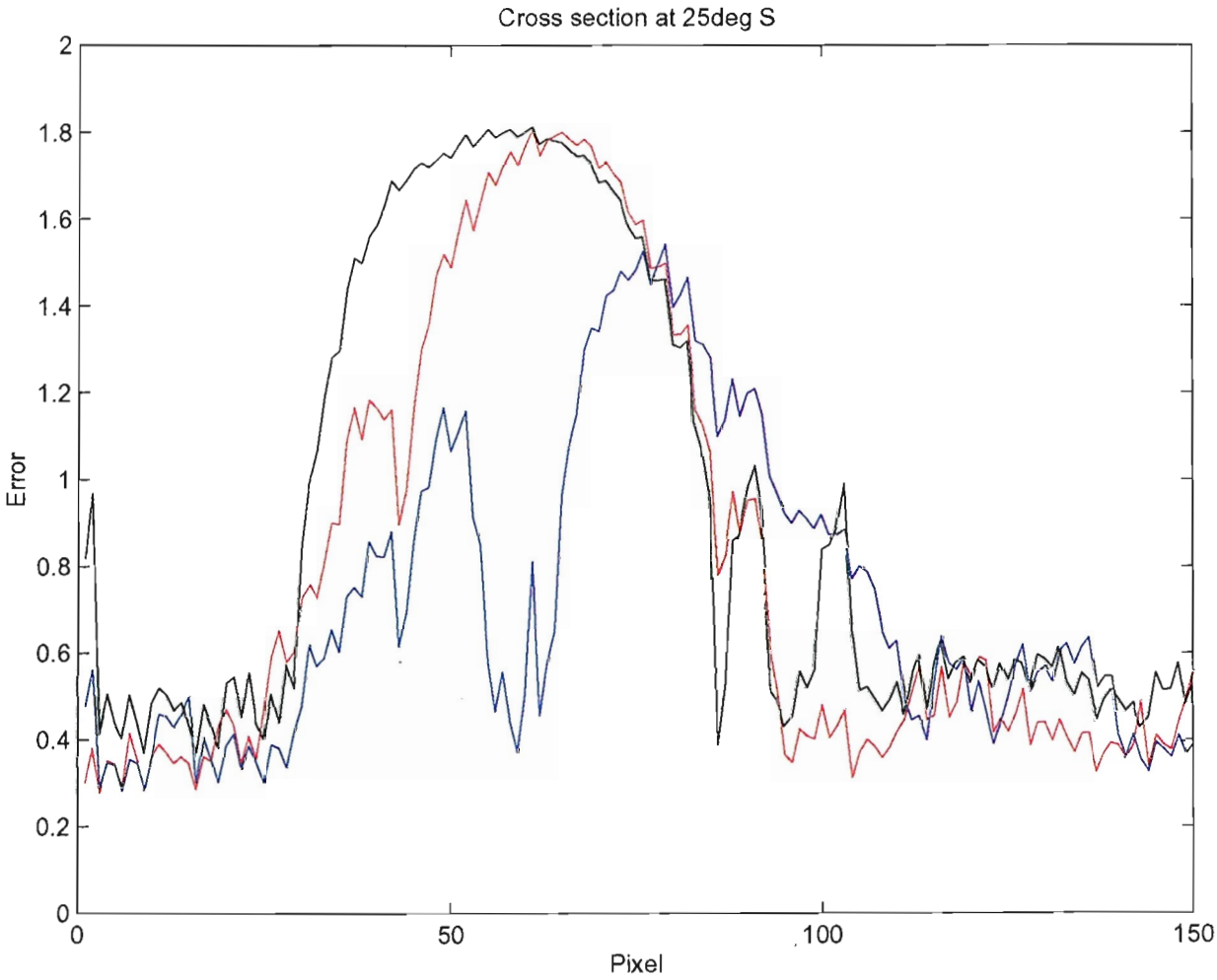


Figure 20: *Cross section of estimated error at 20°S. The error associated with the first day's estimate is shown in black; this is small ($\sim 0.3K$) within the observational swath, but large outside. For the second day, the error estimate (red line) shows the influence of the previous day's observations, with a further improvement obvious in the third day (blue line).*

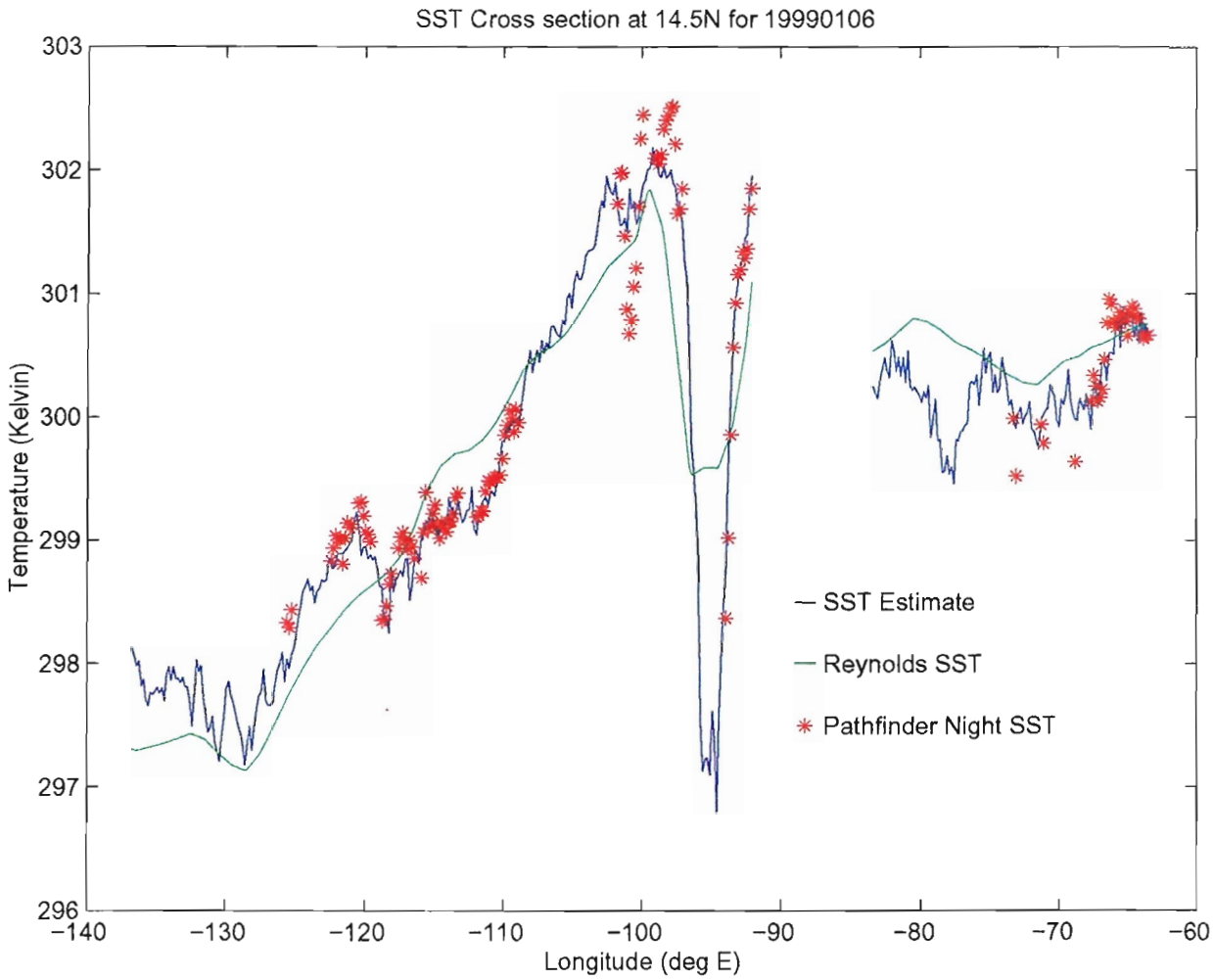


Figure 21: *Cross-section at 14.5N showing decoupling of Pacific and Caribbean basins. The red line shows the SST estimate, and the green line the corresponding Reynolds SST. The latter smooths over SST variation, and interpolates across the land leading to an underestimate of SST on the west coast and overestimate on the east coast of Panama.*

5 Conclusions and recommendations

The SST estimation scheme considered here potentially provides a fast and flexible basis upon which a global, high-resolution (~10km), near-real-time operational data assimilation system can be developed. Areas which require further consideration are as follow:

- Adaptive spatial-correlation model. The current scheme permits the use of only a single zonal and single meridional correlation length scale for the entire geographic region considered. Clearly a priority task is the implementation of a spatially-varying adaptive model, ideally with minimum impact on the computational efficiency of the assimilation scheme.

- Ingestion of multiple datasets. Currently only two datasets can be used simultaneously, but adaption of the scheme to permit multiple datasets is trivial. However enabling the ingestion of datasets at arbitrary spatial resolutions is less straightforward, and it is anticipated that we continue resampling datasets onto a single spatial resolution in the next phase of the project. The implications of this resampling for the error characteristics of the datasets will be considered, particularly in view of the non-Gaussian nature of SST observations associated with cloud contamination.

- Bias corrections. Biases between SST datasets arise from a wide range of sources, including cloud contamination, aerosol contamination, and the variability of the near-surface ocean temperature structure. We anticipate the development of appropriate methods to enable corrections for these effects. These corrections will involve the utilization of ancillary datasets such as windspeed and heat flux estimates. Quantification of the diurnal cycle will be much aided in areas covered by the GOES SST dataset. Partially reducing the sources of SST bias is sufficient to improve the utilization of affected datasets, provided the variance associated with the residual bias can be adequately quantified.

- Improved empirical determination of various error covariance terms. The observed mean and variance in night-only SST measurements based on December 1998 measurements are shown in Figure 7. Rather than using spatially varying observational errors, however, initial estimates have used constant observational errors to simply interpretation of error propagation features.

- Evaluation of the effect of assuming Gaussian error statistics in the SST estimation process.

- Treatment of boundary conditions. No high-latitudes SST data have been considered thus far, however in principle the treatment of ice-edge boundaries is straightforward provided these can be adequately identified.

- The simple persistence model used here could be improved by adding some knowledge of ocean dynamics. However it is not currently proposed to attempt this.

6 Appendix: Assimilation of data over time.

The dynamic evolution of SST is assumed to obey the linear discrete dynamic model given by:

$$x(t+1) = Ax(t) + Bw(t) \tag{1}$$

where $w(t) \sim \mathcal{N}(0, Q)$ is an uncorrelated Gaussian noise process, with zero mean, diagonal covariance Q.

The Pathfinder measurements are linearly related to SST:

$$y(t) = C(t)x(t) + v(t) \tag{2}$$

where $v(t) \sim \mathcal{N}(0, R)$ is an uncorrelated Gaussian noise process, zero mean, diagonal covariance R.

For physical systems governed by models (1), (2), the Kalman filter can be used to obtain filtered estimates for the state $x(t)$ at time t based on the data available up to time t.

Alternatively, it may be preferable to use a smoothing filter, in which the state $x(t)$ is estimated acausally, based on data both preceding and following time t. A smoothing filter can lead to substantial reductions in estimation error, particularly when the measurements are sparse or noisy, and when the physical process is strongly correlated over time. Smoothed estimates are computed based on all the available data $0 \leq t \leq T$ in two passes, a forward pass, identical to the Kalman filter, and a backward pass, initialized by the filtered estimates at end time T.

Our approach is to *emulate* the update and prediction steps of the Kalman filter without trying to solve them exactly (which is computationally prohibitive). The update step is static in nature, and we propose to use an existing, efficient technique based on a multiscale (hierarchical) framework. This technique can accommodate the irregular measurements and nonstationary prior models encountered in the SST problem; it attains computational and storage efficiency by constructing *models* for the needed statistics, rather than generating large covariance matrices. For two-dimensional fields a quadtree is used as the basis for multiscale modeling.

The prediction step of the Kalman filter incorporates the process dynamics, carrying the estimation statistics from one time to the next, however the computational complexity and storage requirements become problematic for large state dimensions; the size of the

SST problem makes it impossible to directly compute predicted estimates and associated error statistics. We assume that the ocean dynamics are very slow, so that a very simple dynamic model (each pixel independently evolves randomly) is adopted. This model implies the following simple estimate prediction

$$\hat{x}(t|t-1) = \hat{x}(t-1|t-1) \tag{3}$$

It remains to propagate the estimation error statistics. The prior model is embedded in the selected multiscale model, so explicit changes to the prior at every time step are inconvenient; instead, we exploit the *duality* between priors measurements, modifying the prior implicitly by introducing new "measurements." Specifically, the measurements at time t consist of one or two independent components: the satellite SST measurement (if any) and the predicted estimates from the previous time step, where $R(t)$ is the covariance of the SST measurement error, and $f(\cdot)$ represents a simplified prediction step. Since only the diagonal elements of $\tilde{P}$ are readily available, we define f as

$$f(\tilde{P}(t-1|t-1)) = \alpha \,\mathrm{diag}(\tilde{P}(t-1|t-1)) + Q \tag{4}$$

where Q is the process noise covariance. Because the predicted covariance is diagonal, the errors in the predicted estimates are treated as independent, which is clearly incorrect, since the errors are correlated. Instead, a positive constant α is introduced to inflate the predicted variances, such that the impact of the inflated variances of the independent measurements is comparable to the desired uninflated, correlated ones.

7 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., Harris A.R., Murray M.J., Zavody A.M. Toward the elimination of bias in satellite retrievals of sea surface temperature 1. Theory, modeling and interalgorithm comparison, *J. Geophys. Res.*, 104, C10, 23565-23578, 1999.

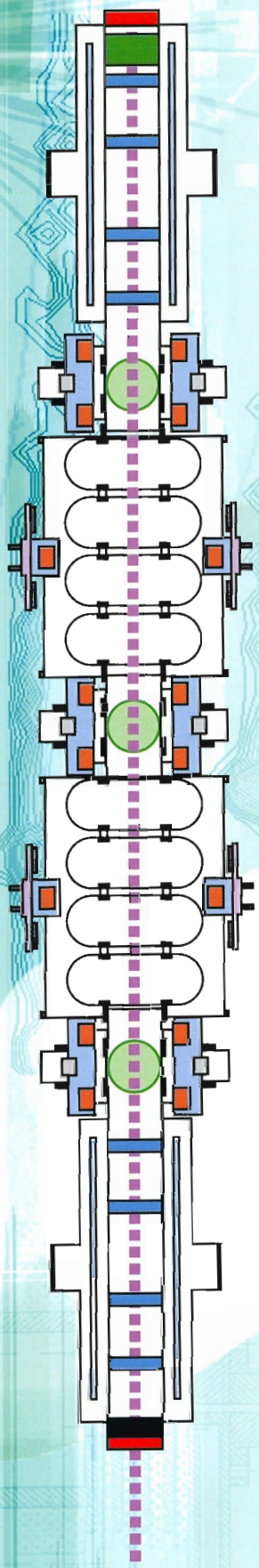
Murray, M.J. et al., "Actual and potential information in dual-view radiometric observations of SST from ATSR", *J. Geophys. Res.*, 103, C4, 8153–8166, 1998a.

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., M.R. Allen, C.J. Merchant, A.R. Harris, and C.J. Donlon, "Direct Observations of skin-bulk SST variability," *Geophysical Research Letters, 27, 1171-1174,*, 2000.

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



Report on the first three years of the initiative in

Accelerator R&D for Particle Physics

August 2002