A Statistical Multiscale Approach to Mapping Altimetric Data

Paul W. Fieguth, William C. Karl, Alan S. Willsky

Laboratory for Information and Decision Systems
Department for Electrical Engineering and Computer Science
Massachusetts Institute of Technology, Cambridge, MA 02139

ABSTRACT

A challenging problem in oceanography is the dense estimation of the surface of the ocean in a statistically meaningful manner, given sparse and irregularly sampled measurements of the surface. A previously developed highly efficient multiscale estimation framework is shown to be an appropriate tool for this task, and we demonstrate the manner of application and present experimental results. The algorithm is capable of computing 250,000 surface estimates of the ocean with error statistics in five seconds on a Sparc workstation.

Keywords: assimilation, smoothing, multiscale algorithms, multiscale estimation.

1 INTRODUCTION

This work deals with the application of a random field multiresolution methodology to the particular problem of the efficient estimation of ocean height. A map of ocean height is of considerable value to ocean modelers, as it allows the identification and correction of errors in computer simulations of global ocean circulation models. The primary reason for the recent increase in interest in the processing of such data is due to the launch of the joint American/French Topex/Poseidon (TP) altimeter: a satellite-based platform capable of characterizing the shape of the ocean surface to unprecedented accuracy (about 5cm). The challenge of our research lies in the necessity to perform statistically based, but highly efficient, computation of dense sets of ocean height estimates *and* error statistics based upon sparse measurements. Figure 1 indicates the region of interest in this initial work: the north-eastern Pacific basin from Hawaii to Alaska. What Figure 1 illustrates is the pattern of data samples provided by the satellite over each ten day period (the pattern repeats every ten days). Observe the sparsity, irregular sampling, and occasional data dropouts in the measurements. Such irregularity in the data pattern presents a major challenge,[4,21] as there is no regular structure that can be used to advantage. Furthermore, there are sources of nonstationarity not only in the expected variability of the sea level in different parts of the ocean, but also in the quality of measurements provided by the satellite. In particular, the altimeter provides direct measurements of the distance from the ocean surface to the satellite; what is actually desired is a measurements of ocean height relative to the geoid[13,14] (the equipotential surface of the earth's gravitational field). Thus the raw altimetric data are subjected to a series of corrections, most importantly for deviations of the geoid from an ellipsoid. Errors in the geoid translate directly into errors in the sea level estimates produced by the corrected altimetry measurements.

The data set is moderately large (20,000 measurements over the north-Pacific region every ten days), however what makes the task difficult is the number of ocean surface estimates desired. For the region shown in Figure 1, in excess of 100,000 surface estimates are computed. Extrapolating this number to a full ocean basin, or to the entire global surface, yields a formidable problem (particularly considering the requirement for estimation error statistics).

Finally, *scale* is a critical issue. Although most natural processes are traditionally described on a single scale

0-8194-1628-2/94/$6.00

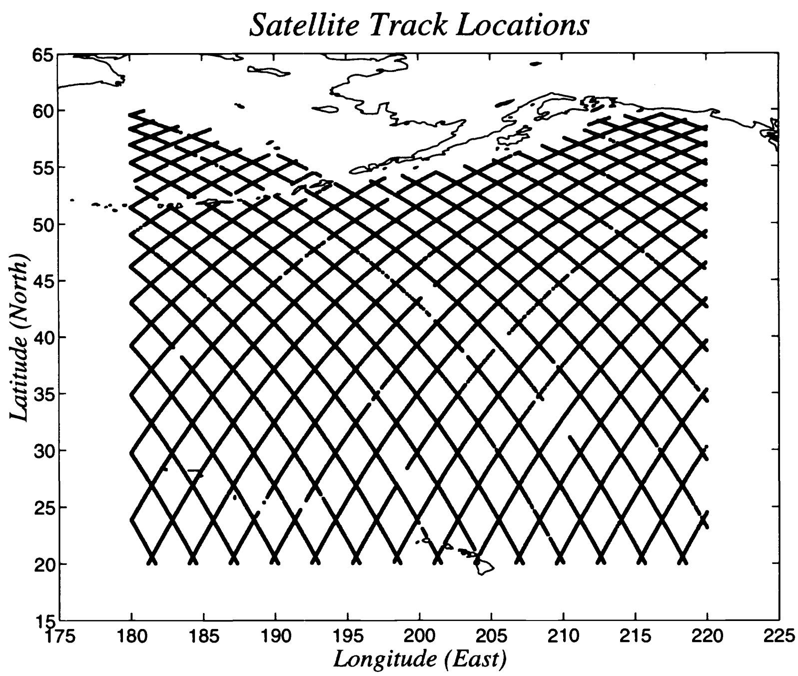


Figure 1: Set of Topex/Poseidon measurement tracks in north Pacific

(via a differential equation, for example), the recent activity in exploring the fractal and self-similar properties of many natural systems suggests that the richness of many such systems manifests itself by its interactions across scales. The ability to capture this richness should ensure a more capable processing of observations made upon this system.

A number of smoothing and data assimilation algorithms (e.g., Objective Analysis,[17] Optimal Interpolation,[5] Kriging[16]) have been developed, each of which have emphasized increasing degrees of statistical structure or computational efficiency. Each of these methods have addressed some of the challenges of the previous paragraphs, however they are unable to deal, in a computationally feasible manner, with the enormous size of problems which we wish to consider.

The key to our approach is the explicitly multiscale nature of our model: rather than operating exclusively on a single scale, we estimate aggregate variables on a hierarchy of scales. Such an approach has the combined benefit of computational efficiency, as well as an opportunity to explicitly describe the scale to scale nature of the system being observed. Furthermore, our modeling framework does not require a special regularity or homogeneity of the measurements or the prior model; efficient estimation is achieved in the presence of heterogeneous measurement locations, error variances, and multiscale model.

The purpose of this paper then is to demonstrate a novel application of a statistical multiscale modeling framework. Section 2 describes the application of scale dependent signal processing to TP data processing and briefly outlines our multiscale estimation framework. Section 3 presents various of the experimental oceanographic results obtained from our efforts. Conclusions and a description of ongoing efforts are summarized in Section 4.

2 MULTISCALE ESTIMATOR OVERVIEW

There are three basic properties of the ocean estimation problem that motivate the application of our estimation methodology:

1. The oceans are extremely large, and need to be imaged at a fine (about 0.2 degree) resolution in order to observe features of interest. This motivates the use of efficient algorithms.
2. Although subject to considerable uncertainty, we do have some understanding of the spatial statistics of the ocean surface. This motivates the use of methods which make explicit use of our knowledge of ocean statistics.
3. The calculation of estimation error covariances is required if the altimetric observations are to be incorporated into global circulation models. The interest in such models motivates the application of smoothers which are not only statistically based, but also capable of generating error statistics.

The general class of problems to which we address our efforts is the class of maximum a posteriori (MAP) problems, described as follows:

$$x \sim \mathcal{N}(0, P_x) \tag{1}$$

$\mathcal{N}(\mu, \sigma^2)$ represents a normal distribution of mean μ and variance σ^2. P_x represents the prior model (or, equivalently, the *a priori* covariance) of the state vector x.

$$y = Cx + v \qquad v \sim \mathcal{N}(0, R) \tag{2}$$

C describes the measurement model, where y is a set of measurements (linear functions of x) corrupted by Gaussian noise v. The general solution to the MAP problem is

$$(P_x^{-1} + C^T R^{-1} C)\hat{x} = C^T R^{-1} y \tag{3}$$

$$E\left[(x - \hat{x})(x - \hat{x})^T\right] = (P_x^{-1} + C^T R^{-1} C)^{-1} \tag{4}$$

This problem formulation and solution is also known as optimal interpolation.[5] In the context of the ocean surface estimation problem, x represents the height field of the ocean, stacked into a vector; y represents the set of measurements obtained from the satellite.

Simple smoothing algorithms (such as local least squares, low-pass filtering or interpolation etc.) chosen for their efficiency tend to be based upon model parameters chosen in a relatively subjective or ad-hoc manner. The model assumed by such smoothers is implicit, and need not bear any resemblance to known ocean statistics. Furthermore these methods do not produce error covariances. Among more general accelerated methods, most of these achieve efficiency through constraints which we feel to be inappropriate for problems such as large-scale data assimilation. For example, most accelerated FFT methods require a regularly sampled grid of measurements - nothing like the sparse irregularly sampled data set of TP.

A direct implementation of Equation 3 is possible only for vectors $\hat{x}$ of relatively low order. Practical implementations of optimal interpolation typically estimate only a small subset of $\hat{x}$ at a time, thus only a relatively small degree of the correlations contained in the prior model P_x are represented in the computed estimates. The novelty of our approach[2,3,6–8] lies in the fact that the prior model P_x represents the state x over multiple scales or resolutions. That is, the elements of x describe the behavior of the system represented by x at various resolutions; P_x is the model which describes the correlation of states over and across scales.

In our framework the prior model P_x is not specified explicitly, rather it is built up implicitly in a scale-recursive manner. Specifically, a prior covariance P_o is specified at the root (top) of the tree, and is built recursively

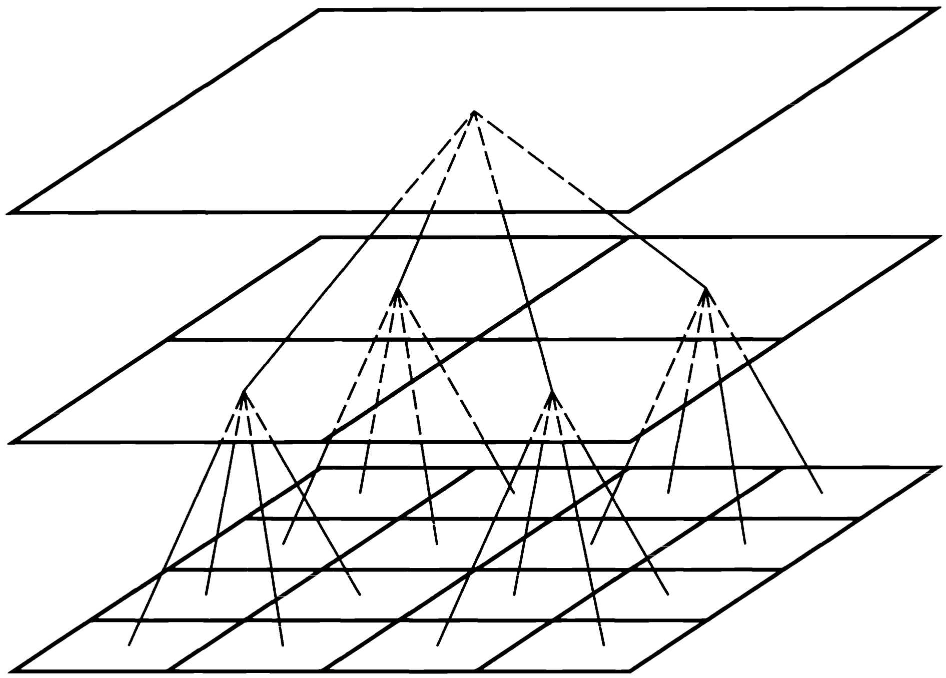


Figure 2: Simple multiscale tree example showing the connections between nodes on three different scales

towards finer scales. Such an implicit specification of P_x is not an unfamiliar concept. For example a difference equation driven by white noise

$$x = Qx + w \tag{5}$$

where Q describes the model for x (e.g., nearest neighbor) and w is white noise, implicitly generates the covariance structure $(I - Q)^{-1}$ for x.

The efficiency of our approach is exacted by insisting upon a special structure of P_x which spans a considerable class of models, including $1/f$ processes and Markov random fields. For prior models P_x possessing this special structure, we are able to solve the associated MAP problem exactly; that is, we are able to perform optimal smoothing, but in an efficient manner. Of even greater importance, the estimation error covariance is calculable. The full error covariance matrix is far too large to calculate explicitly: our approach calculates the diagonal elements of this matrix explicitly, and contains a formulation for the calculation of arbitrary off-diagonal terms.

The multiscale model is built on a tree, of which Figure 2 is a simple two-dimensional example. The efficiency of this scheme is gained via a sort of divide-and-conquer strategy: conditioned on any node on the tree, each of the subtrees connecting to this node are conditionally independent. Rather than jointly estimating a great number of values in one single state vector, leading to enormous covariance matrices, the estimates of the values at given space scale are divided among the tree nodes at that scale (with correspondingly smaller covariance matrices at each node). The joint statistics between estimates distributed across a tree scale are captured by estimates of aggregated values at nodes on coarser scales.

The statistical prior model that describes the system of interest is explicitly multiscale: the state vector at each tree node is written in terms of its parent node with an additive process noise term. The multiscale model may be written as

$$x_{m+1} = A_m x_m + B_m w_m \qquad w_m \sim \mathcal{N}(0, I) \tag{6}$$

where x_m represents the collection of all state vectors on scale m, and w_m is a white noise driving process for the m^{th} scale; $m = 0$ represents the coarsest, or root, scale; $m = M$ represents the finest scale. Similarly there is a multiscale measurement model:

$$y_m = C_m x_m + v_m \qquad v_m \sim \mathcal{N}(0, R_m) \tag{7}$$

where y_m is a vector of measurements corrupted by noise v_m. The measurement model permits the observation of states at any scale; that is, this framework permits the incorporation of measurements at various resolutions. Estimates of x are computed at all points on the tree using the multiscale analogy of a fixed interval smoother such as the Rauch-Tung-Striebel (RTS) algorithm, a non-iterative approach which performs a single upwards pass from the leaves to the root of the tree, followed by a downwards pass. The computational efficiency of this estimation framework is excellent: a grid of 250,000 estimates with error covariance information can be computed from 20,000 surface measurements of the ocean in about 5 seconds on a Sun Sparc-10. Such performance times make very real the possibility of gridding observations from a much larger area, such as the whole Pacific ocean.

A considerable amount of effort[2,7] has been devoted towards exploring the class of systems which may be represented by multiscale models of the form of Equation 6. One of the most important classes, and the one of direct interest here, is the class of $1/f$ models.[19] $1/f$ processes (such as Brownian motion) possess many interesting properties, relating to fractals and self-similarity over scales – the sort of property which we are interested in capturing in natural systems. Selection of the multiscale process values to vary as

$$B_m = \beta 4^{-\mu m/2} \cdot I \tag{8}$$

where I represents the identity matrix and m represents the scale of the node in question allows the approximation of a $1/f^{\mu}$ spectrum.[20]

One additional benefit of our framework is the ability to perform multiscale-model hypothesis testing.[7] That is, given a multiscale prior model and a set of observations, we can compute the relative likelihoods of a set of competing multiscale models. This algorithm permits, for example, a determination of the appropriate values for β and μ in Equation 8.

3 EXPERIMENTAL RESULTS

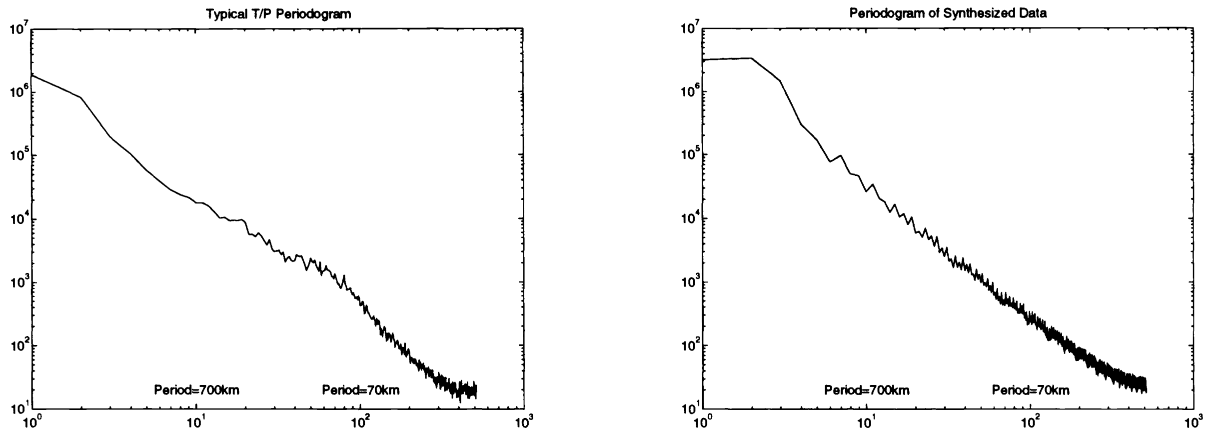
3.1 A. Multiscale Ocean Model Selection

The experimental results of this section are based upon TP altimetric data, repeats nine through thirty-nine. The usual corrections are applied to the data: atmospheric,[10] tidal,[15] orbital,[11] inverse barometer (or atmospheric pressure loading), and geoidal.[13,14] Given a collection of corrected raw satellite data and a multiscale framework from Section 2, the task which remains is the determination of the specific parameters within the multiscale model (i.e., the A_m, B_m, C_m, and R_m of Equations 6 and 7).

Figure 3 shows a sample power spectrum determined from TP data. The process noise of a multiscale model which reproduces this power spectrum is

$$x(\text{child}) = x(\text{parent}) + 52 \cdot 2^{-m(\text{child})} w(\text{child}) \tag{9}$$

A sample power spectrum of data synthesized from this model is also shown in Figure 3. Roughly speaking, this model asserts that the expected height offset between two points on the ocean surface is an increasing function of their separation. The prior covariance P_o, representing the prior knowledge of the mean value at the root node (corresponding roughly to the aggregate mean height of the ocean), is set to be very large ($\approx 10^5$). In other words we presume no prior knowledge of the overall mean ocean height.



Empirical Power Spectrum | Synthesized Power Spectrum

Figure 3: Comparison of power spectra

The measurement model is particularly simple, since our observations are direct measurements of a subset of x_M, where $m = M$ corresponds to the finest scale. That is,

$$C_m = 0 \qquad 0 \leq m < M \tag{10}$$

$$(C_M)_i = \begin{cases} 0 & \text{The } i^{th} \text{ component of } x_M \text{ does not correspond to a TP observation.} \\ 1 & \text{The } i^{th} \text{ component of } x_M \text{ corresponds to an observation point.} \end{cases} \tag{11}$$

We assume two sources of measurement noise:

- The error in estimating the distance from the satellite to the ocean surface, assumed to be 5cm white Gaussian noise.
- The error in the geoid model.

The highest quality geoid models currently available are quite effective at capturing large scale and moderate scale geoid fluctuations, but are less accurate in regions of sharp local changes. Such a result is not surprising: geoid models are constructed as a truncated spherical harmonic expansion, which exhibits larger errors near abrupt changes. Furthermore, navigation errors in the satellite in areas of steep geoid gradient lead to greater uncertainty in the geoid correction than in other regions in which the geoid is smoothly varying. As a result, altimetric measurements in the vicinity of steep geoid slopes represent a less accurate assessment of the corrected ocean surface. Consequently we argue the following measurement noise model:

$$R_M = (5\text{cm})^2 + g(\text{Geoid Slope}) \tag{12}$$

where $g()$ is an increasing function (and is detailed in the next section).

The spatial position of the multiscale tree on the ocean is somewhat arbitrary; that is, there is no particularly natural orientation for the multiscale tree, and the validity of the multiscale estimates should not change if the multiscale tree is shifted a few degrees east or west. Figure 4 shows a given field with four possible tree orientations. Given the multiscale model of Equation 9, the four children of some tree node s are correlated only through their common parent. The motivation behind the use of multiple trees is the reduction of artifacts due to decorrelation across coarse boundaries of the tree. In order to attenuate possible artifacts we compute ocean surface estimates for each of ten trees (each shifted with respect to the others) and average the results. It should be emphasized that this is not at all like spatial low-pass filtering or interpolation.

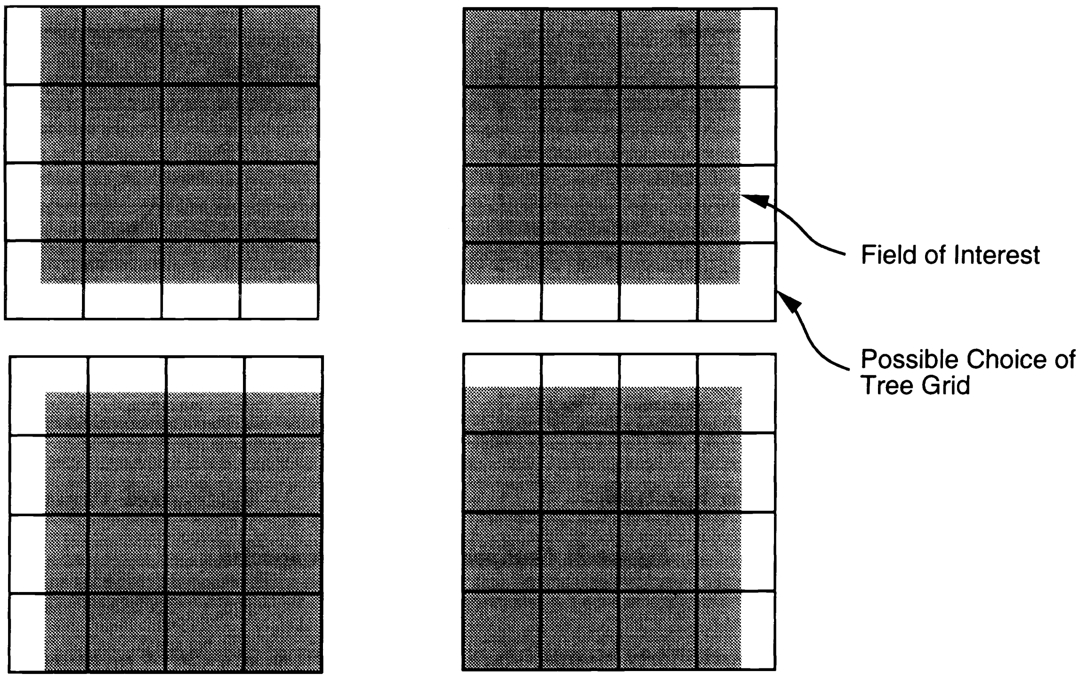


Figure 4: Four possible tree positions, shifted with respect to one another.

3.2 Gridding Results

Having determined a multiscale model suitable for processing ocean altimetric data we are in a position to perform preliminary data assimilation tests. A sample map of ocean surface estimates is shown in Figure 5. This map is based upon a single repeat cycle, or ten days, of data (about 20,000 data points). The 250,000 estimates and associated estimation covariance information were computed in about 1 minute on a Sun Sparc-10 (the map shows the combination of ten trees of estimates; each multiscale tree requires a computation time of about 5 seconds).

The ocean height variations shown in the figure are consistent with the known large-scale oceanographic behavior of the region: there is a predominant gradient in the north-south direction with surface height offset on the order of one meter. The estimates offer far higher resolution than has heretofore been available for ocean modelers for use as part of their ocean circulation studies. It is this very leap in resolution that makes the assessment of our results difficult - we have come across no other altimetric maps of sufficient resolution to support or refute our plots.

A set of estimate error variances corresponding to the above estimates is shown in Figure 6. The distribution of measurement dropouts in this data set can be inferred from Figure 1. As before, the results are computed as the average over ten multiscale trees. Because of the spatially varying uncertainty in our measurements due to geoid error, the occurrence of data dropouts, and the irregular pattern of data collection, we would expect that the uncertainty pattern in the optimal estimate of our ocean height map would be highly variable and would, to some extent, reflect these features. The map of the estimation error variances (Figure 6) supports this argument. In particular, observe that the regions of lowest uncertainty (the lightly shaded regions in the figure) correspond with the points at which we have satellite measurements; a careful inspection of the figure will also reveal occasional darker breaks along these regions, corresponding to data dropouts. In addition, because of our spatially-varying noise model, the measurements near the Aleutian and Hawaiian chains (which induce a significant geoid gradient) are modeled as being noisier, resulting in elevated covariance values. The large region

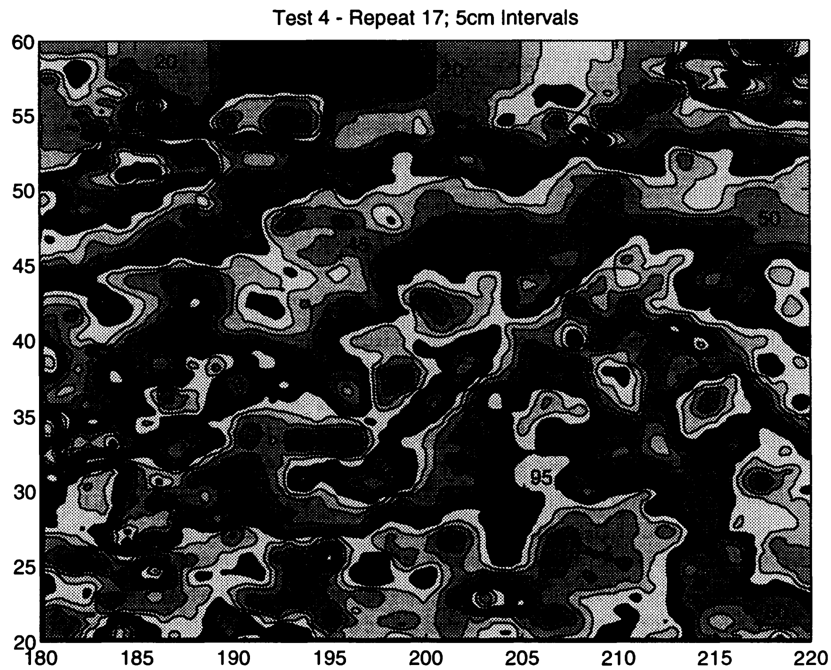


Figure 5: Estimation of ocean mean circulation based on a single ten day set of data

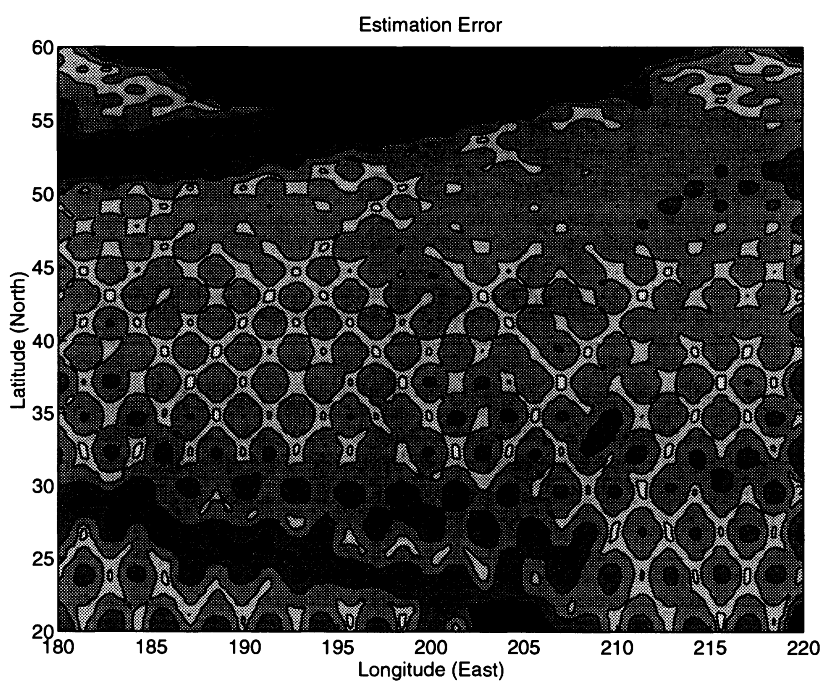


Figure 6: Estimation error variances based on one repeat cycle of data; darker regions represent greater uncertainty

of uncertainty at the top of the figure is due to the Alaskan land mass.

The examination of measurement residuals, the differences between the satellite measurement and the ocean surface estimates, can serve to verify the validity of our multiscale models. In particular, by normalizing these residuals with respect to their expected standard deviations (computed from the error covariance produced by our algorithm) we can isolate statistically significant outliers. For example, the nature of the geoid slope dependent term in the measurement error (Equation 12) can be examined. A set of estimates and residuals were computed using a measurement noise of

$$R_m = (5\text{cm})^2 \tag{13}$$

The resulting distribution of statistically large residuals is shown, superimposed upon a map of the geoid gradient in Figure 7. The correlation between large residual locations and steep geoid slope is unambiguous, and argues convincingly in favor of our geoid slope-corrected measurement noise model. As an additional comparison, the same locations of large residuals are shown superimposed on a plot of ocean bathymetry contours (the shape of the ocean bottom, which heavily influences the shape of the geoid) in Figure 8. Computing the RMS value of residuals as a function of geoid slope allows a quantitative assessment of the geoid-slope dependent term in the measurement noise model (function $g()$ of Equation 12). Since the efficiency of our framework is derived from a particular structure of the prior and measurement models and not from a regularity of measurement locations or error variances, the implementation of a spatially varying measurement noise (due to the heterogeneity of $g()$) is easily implemented and costs little computationally. Similarly spatial inhomogeneities in the prior model (that is, in the scale to scale process noise weights B_m), due perhaps to the prior knowledge of the location of a strong ocean current, are just as easily incorporated with essentially no computational burden.

4 CONCLUSIONS

We have demonstrated the application of a multiscale estimation framework to a significant estimation problem in oceanography. The results of this paper are of interest from a statistical estimation point of view, since every successful application broadens the class of problems which may be addressed by the methodology. This paper concentrates on the methodology and application of the multiscale framework, rather than on the details of the application itself. We are collaborating with colleagues in oceanography, an anticipate making more direct contributions to that field in the near future.

5 ACKNOWLEDGEMENTS

Research support was provided in part by the Office of Naval Research under Grant N00014-91-J-1004, by the Draper Laboratory under Grant DL-H-467133, and by the Air Force Office of Scientific Research under Grant F49620-92-J-0002. P.W.F. was also supported by an NSERC-67 fellowship of the Natural Sciences and Engineering Research Council of Canada. The support of NASA in providing the altimetric data is appreciated.

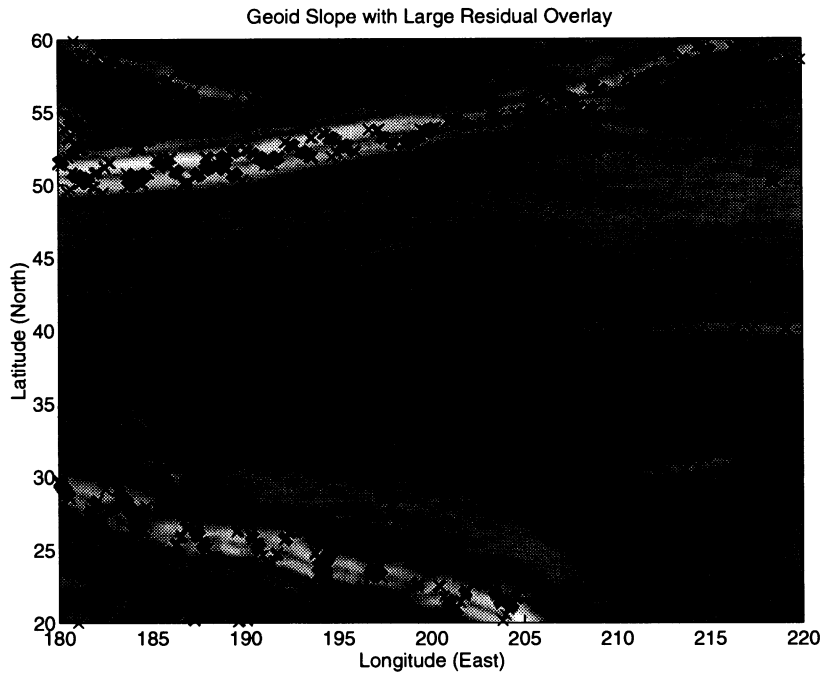


Figure 7: Overlay of geoid gradient map and sites of large residuals; regions of lighter shading represent steeper geoid gradient.

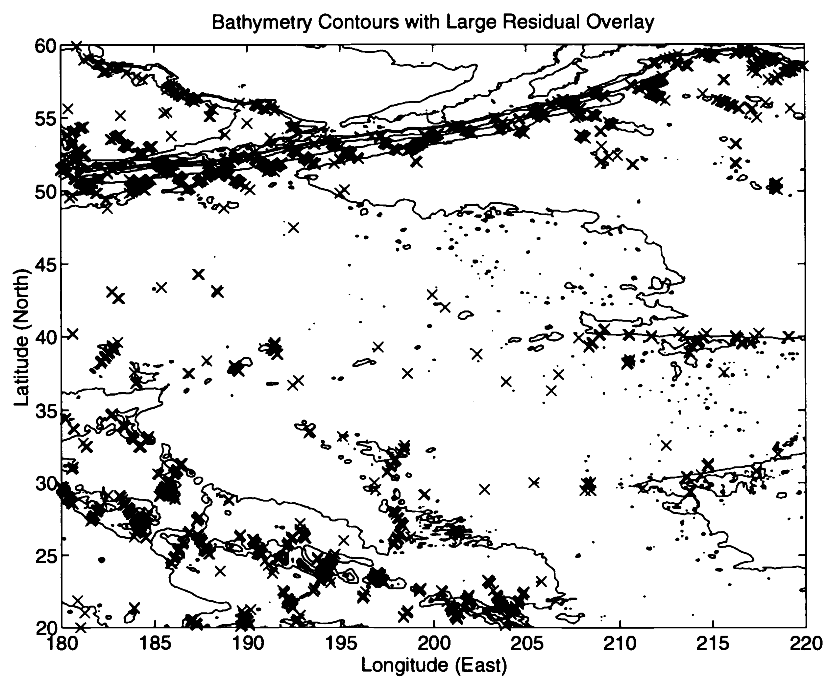


Figure 8: Overlay of ocean bathymetry contours and sites of large residuals

6 REFERENCES

[1] D. Chelton, M. Schlax, D. Witter, J. Richman, "Geosat Altimeter Observations of the Surface Circulation of the Southern Ocean", *J. Geophys. Research* (95), pp.17877–17903, 1990

[2] K. Chou, "A Stochastic Modeling Approach to Multiscale Signal Processing", PhD Thesis, Dept. of EECS, Massachusetts Institute of Technology, 1991

[3] K. Chou, A. Willsky, A. Benveniste, "Multiscale Recursive Estimation, Data Fusion, and Regularization", *IEEE Trans. on Automatic Control*, to appear.

[4] P. Gaspar, C. Wunsch, "Estimates from Altimeter Data of Baryotropic Rossby Waves in the Northwestern Atlantic Ocean", *J. Physical Oceanography* (19)#12, pp.1821–1844, 1989

[5] M. Ghil, P. Malanotti-Rizzoli, "Data Assimilation in Meteorology and Oceanography", *Advances in Geophysics* (33), pp.141–266, 1991

[6] M. Luettgen, W. Karl, A. Willsky, R. Tenney, "Multiscale Representations of Markov Random Fields", *IEEE Trans. Signal Processing* (41) #12, pp.3377–3396, 1993

[7] M. Luettgen, "Image Processing with Multiscale Stochastic Models", PhD Thesis, Dept. of EECS, Massachusetts Institute of Technology, 1993

[8] M. Luettgen, W. Karl, A. Willsky, "Efficient Multiscale Regularization with Applications to the Computation of Optical Flow", *IEEE Trans. on Image Processing* (3) #1, pp.41–64, 1994.

[9] P. Marth et al., "Prelaunch Performance of the NASA Altimeter for the TOPEX/POSEIDON Project", *IEEE Trans. on Geoscience and Remote Sensing* (31) #2, pp.315–332, 1993

[10] F. Monaldo, "Topex Ionospheric Height Correction Precision Estimated from Prelaunch Test Results", *IEEE Trans. Geoscience and Remote Sensing* (31) #2, pp.371–375, 1993

[11] R. Nerem et al., "Expected Orbit Determination Performance for the Topex/Poseidon Mission", *IEEE Trans. Geoscience and Remote Sensing* (31) #2, pp.333–354, 1993

[12] J. Nystuen, C. Andrade, "Tracking Mesoscale Ocean Features in the Caribbean Sea Using Geosat Altimetry", *J. Geophys. Research* (98)#C5, pp.8389–8394, 1993

[13] R. Rapp, Y. Wang, N. Pavlis, "The Ohio State 1991 Geopotential and Sea Surface Topography Harmonic Coefficient Models", Report #410, Dept. of Geodetic Science and Surveying, Ohio State University, 1991

[14] R. Rapp, "Geoid Undulation Accuracy", *IEEE Trans. Geoscience and Remote Sensing* (31) #2, pp.365–370, 1993

[15] R. Ray, "Global Ocean Tide Models on the Eve of Topex/Poseidon", *IEEE Trans. Geoscience and Remote Sensing* (31) #2, pp.355–364

[16] B. Ripley, *Spatial Statistics*, Wiley, 1991

[17] T. Schlatter, G. Branstator, I. Ihiel, "Testing a Global Multivariate Statistical Objective Analysis Scheme with Observed Data", *Mon. Weather Review* (104), pp.765–783, 1976

[18] R. Tokmakian, P. Challenor, "Observations in the Canary Basis in and Azores Frontal Region Using Geosat Data", *J. Geophys. Research* (98)#C3, pp.4761–4773, 1993

[19] G. Wornell, "Synthesis, Analysis, and Processing of Fractal Signals", PhD Thesis, Dept. of EECS, Massachusetts Institute of Technology, 1991

[20] G. Wornell, "Wavelet-Based Representation for the $1/f$ Family of Fractal Processes", *Proc. IEEE*, Sept. 1993

[21] C. Wunsch, "Sampling Characteristics of Satellite Orbits", *J. of Atmospheric and Oceanic Tech.* (6)#6, pp.891–907, 1989