Multiresolution Stochastic Processing of Topex / Poseidon Oceanographic Altimetry

Paul W. Fieguth, Student Member, IEEE
William C. Karl, Member, IEEE
Alan S. Willsky, Fellow, IEEE
Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
77 Massachusetts Ave.
Cambridge, Massachusetts 02139

Abstract - **Remarkable advances in remote sensing are producing measurements on increasingly finer grids, offering the possibility of a far better understanding and characterization of the ocean surface. The sheer quantity of measurements has rendered traditional data smoothing/assimilation algorithms obsolete however, even on high performance workstations; the basis of this paper is the application of an extremely fast multiresolution smoothing framework to ocean altimetry data.**

Using the previously developed multiresolution framework, a simple multiscale model is developed which captures the characteristic power-law spectrum of the ocean surface. Smoothing altimetric data and computing error variances on a grid of 512 by 512 points is accomplished in about 10 seconds on a SUN Sparc-10. Finally, further applications of this framework within the oceanographic context are discussed, including model identification and anomaly detection.

Abstract - **Le remarquables progrès en surveillance par satellites permettent des mesures de plus en plus précis, et offrent une possibilité de compréhension et caractérisation de la surface de l'océan, nettement meilleure que celle d'il y a quelques années. L'abondance de données a beaucoup limité l'utilisation des algorithmes traditionelles de lissage, même avec les ordinateurs de haute puissance. Donc, cet article concentre sur l'application d'un algorithme multiéchelle à grand vitesse aux données altimetrique de l'océan.**

En utilisant le formalisme multirésolution, précédemment developpé, un simple modèle multiéchelle possedant un spectre $1/f^{\mu}$ - le spectre characteristique de la surface de l'océan - est proposé. Utilisant un SUN Sparc-10 ordinateur, nous pouvons lisser des données sur un quadrillage de 512 x 512 points en moins de dix seconds. Nous décrivons quelques applications océanographiques de cet algorithme - l'identification des modèles multiéchelles, et la détection d'anomalies.

The principle author was supported by a NSERC-67 fellowship of the Natural Sciences and Engineering Research Council of Canada, 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.

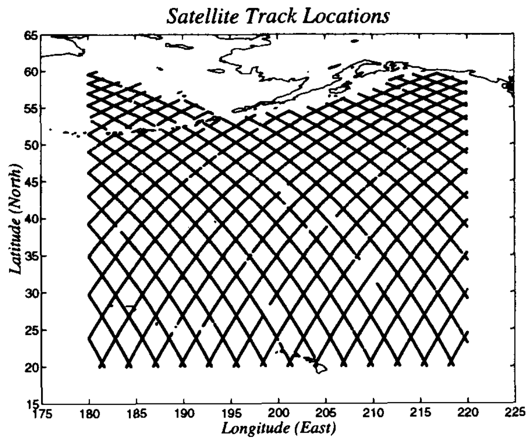

Figure 1: Set of Topex / Poseidon measurement tracks in north Pacific

I. INTRODUCTION

The need to perform sophisticated remote sensing of the ocean has increased substantially in recent decades. Not only is the significance of the role of the ocean on global warming and other meteorological trends becoming increasingly apparent, but the ocean circulation models meant to predict such events are running at increasing levels of precision and require correspondingly augmented quantities of reliable oceanographic data.

The cooperative international American/French Topex / Poseidon[5] mission has taken a significant step in providing this high quality data. Ocean surface height measurements, accurate to about 5cm, are available on densely sampled tracks spaced 300km apart; complete coverage of the earth is obtained every ten days.

The goal of our research is to assimilate Topex / Poseidon altimetric measurements into forms that allow meaningful quantitative assessment by oceanographers and that can drive dynamic ocean circulation models.

0-7803-2056-5 © 1994 IEEE

There are a number of aspects that make the development and application of a statistical model challenging in this specific problem:

- There is an enormous amount of data to be assimilated. In particular, every ten day cycle produces $\approx 20\,000$ measurements within the small region shown in Figure 1, which need to be gridded onto upwards of 100 000 points.
- The space/time sampling pattern of the data is unusual, and it is difficult to develop efficient estimators which can cope with it[8]. In particular, certain common acceleration methods (e.g., using the FFT) require a degree of regularity and structure not found here.
- The temporal and spatial scales of the ocean are coupled.
- The gravitational equipotential (or geoid, the shape that the ocean surface would assume in the absence of currents), is approximated by a model which is subject to considerable uncertainty[6]. The geoid model is subtracted from Topex / Poseidon observations to yield estimates of ocean altimetry, so the geoid model error contributes directly to altimetric errors.
- The prior statistics of the ocean surface are uncertain. Indeed, the determination of such statistics remains an important active oceanographic research objective.
- Not only ocean surface estimates, but also surface estimation *error* statistics are required in order to drive ocean circulation models. Error statistics are generally computationally prohibitive to calculate in large estimation problems such as the one considered in this paper.

We meet this challenge by using a novel multiscale statistical estimation framework. In this paper we will outline the nature of our estimator, and demonstrate its performance on the Topex / Poseidon altimetry assimilation problem.

The region of the ocean which forms the focus of our study in this paper is the north Pacific region shown in Figure 1. The choice of this region is partly motivated by the quiescence of the ocean in this area; i.e., the time evolution of the ocean is sufficiently slow to permit a moderate decoupling of time and space scales. A more general treatment of time remains an active research endeavor.

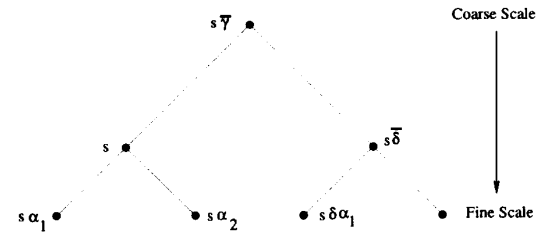

Figure 2: Multiscale tree example - in this case a dyadic tree is shown; the multiscale framework permits considerably more general tree structures however.

II. MULTISCALE ESTIMATION

A multiscale estimation methodology has recently been developed[1, 4] that offers the promise of efficient optimal data assimilation and smoothing algorithms. The methodology effectively generalizes the Kalman filter and Rauch-Tung-Striebel smoother to operate on a tree structure such as in Figure 2; that is, an explicitly multiscale system model is specified on the tree as

$$x(s) = A(s)x(s\bar{\gamma}) + B(s)w(s) \qquad (1)$$

where $x(s)$ represents the state at tree node s, and $B(s)w(s)$ represents the process noise term from parent state $s\bar{\gamma}$ to state s. The framework is very flexible and is applicable to multidimensional problems.

This multiscale methodology has a number of powerful features, several of which directly address the challenges listed earlier:

- A multiscale model is a natural means of representing fractal or self-similar processes, which include many natural systems. Specifically, the set of models which can be represented within our multiscale framework include $1/f$ processes (i.e., those processes having a $1/f^{\mu}$ spectrum).
- The measurements to be assimilated are permitted to be sparse, multiscale, and heterogeneous. Furthermore the multiscale model *itself* may be heterogeneous; this permits nonuniformities such as localized ocean currents or anticipated regions of geoid model error to be incorporated into the model.
- The resulting scale-recursive algorithm is extremely fast and offers a constant computational time per pixel as the number of pixels is scaled (i.e., the computation burden is directly proportional to the total number of pixels). A 512 by 512 array of estimates *and* error variances was computed in only ten seconds on a SUN Sparc-10. Furthermore the algorithm is easily parallelizable for additional computational gains.

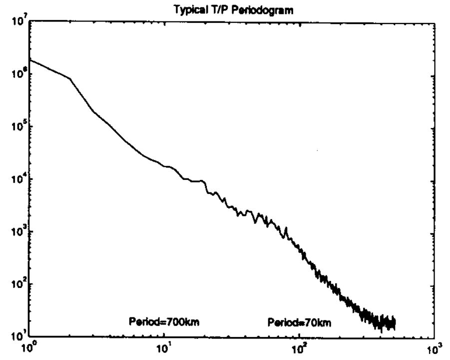

Figure 3: Empirical power spectrum based on Topex / Poseidon data

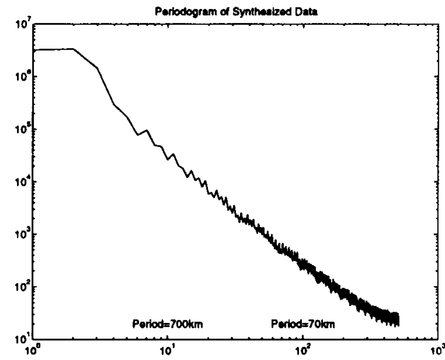

Figure 4: Power spectrum from simulations of a multiscale model (the model was chosen to approximate the spectrum of Figure 3)

- Both estimation error covariances and *cross*-covariances may be computed between arbitrary states (even across scales), an impossibility with other approaches such as those using Markov random fields. Such error covariances are crucial in allowing the assimilation of data into ocean circulation models.
- Multiscale error covariances allow us to directly assess the scale of reconstruction supported by the quality and quantity of data. That is, the estimation error variance may be minimized at some scale other than the finest scale if we have insufficient data to make meaningful estimates at the finest scale.
- Multiscale likelihood ratios can be computed, facilitating multiscale model identification. Using hypothesis testing on an appropriate class of models, these likelihood ratios allow us to determine various ocean surface statistics.

III. APPLICATION TO OCEAN DATA

Our research involves the simultaneous exploration of interdisciplinary scientific concerns in both multiscale stochastic algorithms and oceanography:

1. A simple model of the ocean height field is a fractal or self-similar one (e.g., a $1/f^2$ spectrum), which is supported by the form of the empirical spectrum as shown in Figure 3. Such a model takes the form

$$x(s) = x(s\bar{\gamma}) + B2^{-\mu m(s)}w(s) \tag{2}$$

where $w(s)$ is unit variance white noise, $m(s)$ measures the scale of corresponding node s, B and μ are parameters that control the offset and slope of the spectrum. The empirical spectrum is closely matched using $B = 52cm$, $\mu = 0.5$; the spectrum associated with this multiscale model is shown in Figure 4. Applying this model to one cycle (ten days) of Topex / Poseidon data yields altimetry estimates and estimation error variances such as those shown in Figures 5 and 7.

In (2), each multiscale tree node represents some aggregate ocean height over some region by a single scalar, and the prior model is $1/f$-like with a constant power law. Three variations of this model have been considered:

i. Permit a higher order representation (e.g., height and gradient) of the ocean at each tree node. Such a model is particularly useful if the smoothed altimetric measurements are to be fed into a global circulation model, for which surface gradients are required.

ii. Permit a greater degree of freedom in the process noise parameters, allowing μ in (2) to vary with scale; this could be used to model a system having a piecewise constant power law $1/f$ spectrum, for example. Since an algorithm to compute multiscale model likelihoods has been developed[4], applying this algorithm to a model parameterized in terms of several degrees of freedom in μ permits a certain characterization of the manner of spectral rolloff as a function of frequency.

iii. The ocean *anomaly* is the difference between the instantaneous ocean height and the yearly mean height. Since the size of the yearly mean is very much greater than the day to day fluctuations of the ocean surface, anomalies are commonly computed in order to observe these fluctuations. Anomaly maps may be computed directly using a multiscale model with μ selected to fit the anomaly

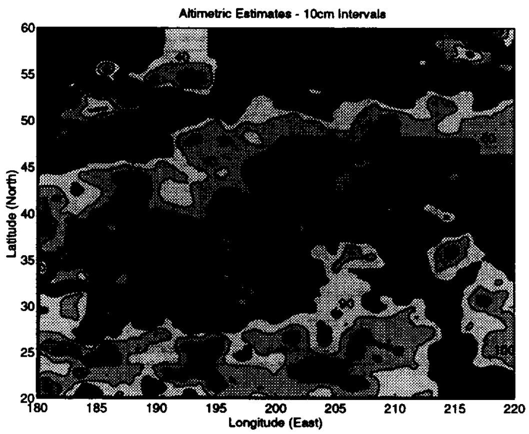

Figure 5: Sample Ocean Surface Estimates
Surface estimate based on assimilation of ten days of data; numbers indicate surface height in centimeters. The overall trend is a north-south gradient, as expected.

spectrum (Figure 9); the anomaly map corresponding to Figure 5 is shown in Figure 10.

2. Uncertainties in our knowledge of the earth's geoid and in the topography of the ocean floor directly translate into large estimation residuals. Figures 6 and 8 display the stunning correlation between large error residuals in our estimates, and areas of steep geoid gradient (which are correlated with greater errors in the geoid model) and the ocean bathymetry (the shape of the ocean floor). Our methodology can jointly estimate these quantities and detect statistically significant irregularities by

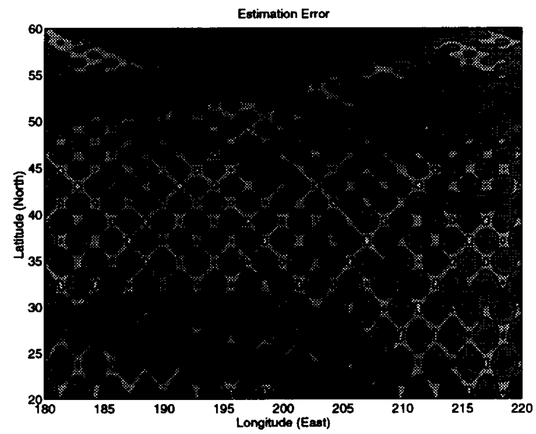

Figure 7: Estimation Error Variances
Darker regions represent *greater* uncertainty.
The light diagonal bands coincide with the measurement orbits of the satellite. The large uncertain area in the north is due to the loss of data over land (Alaska and the Aleutians). The band of uncertainty in the south-west is due to the Hawaiian island chain.

examining the measurement residuals. This is directly related to the theoretical problem of multiscale statistical whitening and anomaly detection.

3. Finally, a problem of great interest both in oceanography and in multiscale estimation theory is the relationship between spatial and temporal behaviors. Some effort has been made to determine a model which evolves in time, incorporating new measurements at each time step, but which also retains a multiscale structure

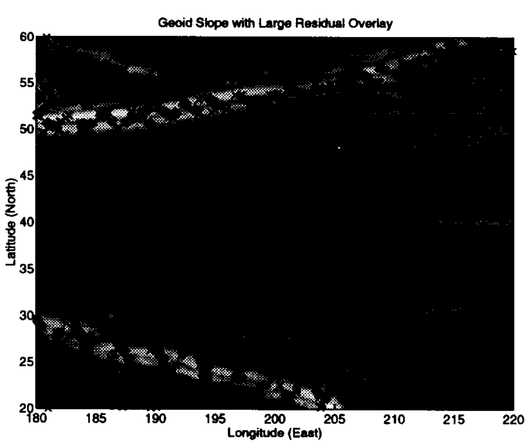

Figure 6: Large Residuals vs. Geoid Slope
Locations of large estimation residuals are marked with an X.
The background shading shows geoid slope: lighter areas represent steeper slopes.

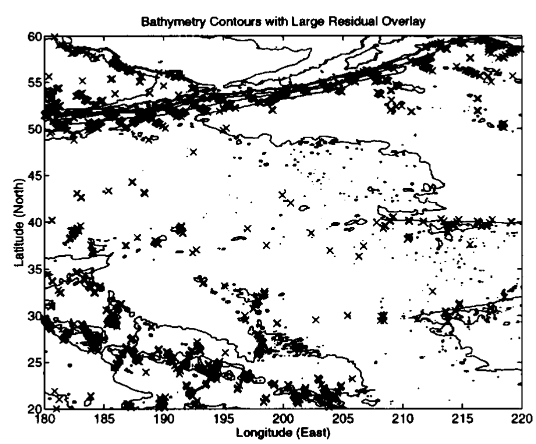

Figure 8: Large Residuals vs. Bathymetry
Locations of large estimation residuals are marked with an X.
The background contours show bathymetry (depth of ocean bottom). Deepest contour at 5000m.

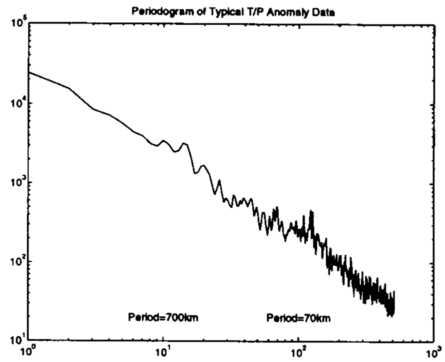

Figure 9: Power spectrum based on Topex / Poseidon anomaly data (the anomaly field is the difference between the instantaneous height and the yearly mean)

such as (2) at each time step. The class of time dynamics discovered so far which satisfies these requirements is too small to be of interest, however this remains a direction of interest.

Rather than a model, linear in time and multiscale in space (as just described), one may consider models multiscale in both space and time. That is

$$x(s) = A(s)x(s\bar{\gamma}) + B(s)w(s) \tag{3}$$

where s now indexes over both space and time. It is the determination of appropriate model parameters $(A(s), B(s))$ that presents a challenge here. We are hoping to gain insight into the modeling challenge by improving our intuition through consideration of the evolution of the ocean surface.

IV. CONCLUSIONS

This paper has outlined the application of a novel and efficient multiscale smoothing framework to ocean altimetric data. A number of simple models have been employed with encouraging results. Our research efforts involve the development and understanding of higher order multiscale models, and a continuing effort to efficiently incorporate time dynamics into a multiscale model.

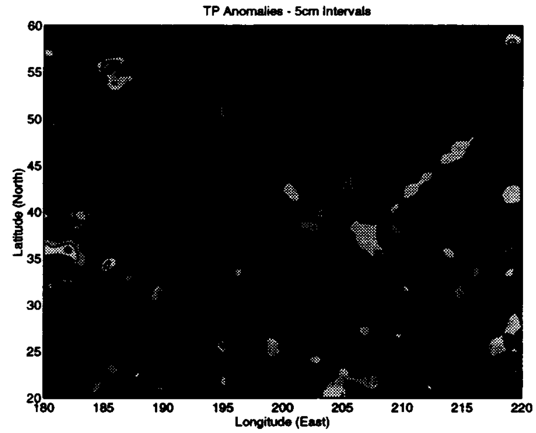

Figure 10: Anomaly field corresponding to Figure 5

References

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

[2] 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

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

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

[5] 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

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

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

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