

**“Assimilation of TOPEX Sea Level Measurements
with a Reduced-Gravity Shallow Water Model
of the Tropical Pacific Ocean”**

b y

Ichiro Fukunori

**Jet Propulsion Laboratory, California Institute of Technology,
Pasadena CA 91109**

January 25, 1995

(submitted for publication; *J. Geophys. Res.*)

Abstract

Sea surface height variability measured by TOPEX are analyzed in the tropical Pacific Ocean by way of an assimilation into a wind-driven, reduced-gravity shallow water model using an approximate Kalman filter and smoother. The analysis results in an optimal fit of the dynamic model to the observations, providing a dynamically consistent interpolation of sea level and estimation of the circulation. Nearly 80% of the expected signal variance is accounted for by the model within 20° of the equator, and estimation uncertainty is substantially reduced by the voluminous observation. Notable features resolved by the analysis include seasonal changes associated with the North Equatorial Countercurrent, and equatorial Kelvin and Rossby waves. Significant discrepancies are also found between the estimate and TOPEX measurements especially near the eastern boundary. Improvements in the estimate made by the assimilation are validated by comparisons with independent tide gauge and current meter observations. The employed filter and smoother are based on approximately computed estimation error covariance matrices, utilizing a spatial transformation and an asymptotic approximation. This analysis demonstrates the practical utility of a quasi-optimal filter and smoother,

1. Introduction

In a recent study, Fu et al. (1993) analyzed GEOSAT sea level observations in the tropical Pacific Ocean using a wind-driven equatorial wave model. The analysis was based on an assimilation of the measurements using a Kalman filter and smoother. Data assimilation allows quantitative analyses of observations in a dynamical framework, and in particular, Fu et al. (1993) found 68% of the signal variance of the sea level observation to be consistent with wind-driven equatorial waves. The limitation to Fu et al.'s (1993) analysis was in the simplified dynamics of their model. Variations not associated with equatorial Kelvin and Rossby waves were not resolved, and their simplified model geometry did not allow incorporations of the coastal geometry and was limited to analyzing data within 10° of the equator.

In this present study, we will utilize a reduced-gravity shallow water model to analyze TOPEX sea level observations (Fu et al., 1994) from September 24, 1992 till September 25, 1993. The objective of this study is to analyze the skill of a model that also incorporates other physics besides equatorial waves in accounting altimetric observations, and in turn to evaluate and to explore the accuracies and utility of TOPEX measurements. The third goal of this study is to demonstrate the efficacies of an approximate Kalman filter and smoother in making useful estimates of the ocean circulation.

Reduced-gravity shallow water models have frequently been used to study the sea level response in the tropical oceans to changes in wind forcing since the early study of Busalacchi and O'Brien (1981). The oceanic response to wind changes is an essential part of the physics of El Niño, and such a shallow water model is also a major component in the model of Zebiak and Cane (1987) in simulating and predicting El Niño. One of the dominant signals observed in the TOPEX measurements are the seasonal changes associated with the equatorial current system (Stammer and Wunsch, 1994). Such changes are in a quasi-steady-state balance with the wind that cannot be simulated with an equatorial wave model, but should be resolvable by a shallow water model.

Although TOPEX/POSEIDON observations have unprecedented accuracies, there are still remaining errors in the sea level data that are comparable to the signal of some of the oceanic phenomena of interest. One of the largest remaining error sources at

present is believed to be in the oceanic tidal corrections. The standard tidal corrections provided with the TOPEX GDR are based on the models of Cartwright and Ray (1990) and Schwiderski (1980). Fu and Pihos (1994, unpublished manuscript) have computed empirical corrections to some of the tidal constituents of these two models using TOPEX measurements. While assimilation corrects model deficiencies, the estimate in turn can be used to evaluate these different tidal corrections by comparing which corrections lead to the best agreement between data and model.

The larger dimensionality of the present shallow water model ($\approx 12,000$ variables) in comparison with the model used by Fu et al. (1993) (≈ 400) results in a significant increase in the computational requirements such that direct application of Kalman filtering becomes prohibitive. For example, the amount of computation for matrix multiplications involved in Kalman filtering is proportional to the cube of the state's dimension. Fukumori and Malanotte-Rizzoli (1995) have recently proposed an approximate filter employing a transformation that approximates the model state with one that has fewer degrees of freedom, thus effectively reducing the size of the estimation problem and the amount of computations involved. Fukumori and Malanotte-Rizzoli (1995) applied such a filter to a nonlinear primitive equation model with a state dimension exceeding 170,000 elements in a twin experiment. We will utilize their approximations for the present shallow water model, and thereby examine the approximation's utility in analyzing real observations.

2. Data

The analyzed observations are temporal sea level variabilities measured during the first year of the TOPEX/POSEIDON mission (cycles 1 to 37) corresponding to measurements taken from September 24, 1992 till September 25, 1993. Data are based on the TOPEX Geophysical Data Record (GDR) (POSEIDON data were not used in this study) with all standard environmental corrections applied including solid earth and ocean tides, water vapor, dry tropospheric and ionospheric delays, and an inverse barometer correction for atmospheric pressure loading of the sea surface. The Cartwright and Ray (1990) model is used for ocean tides, plus an additional correction of its M2 and K1 constituents computed by Fu and Pihos (1994, unpublished manuscript). The Fu and Pihos correction is an empirical estimate of the residual tidal error that is computed in a similar manner as Schrama and Ray (1994). To avoid

the effects of the model's artificial boundaries at 30°N and 30°S in the analysis, the analyzed data will be limited within 20° of the equator in the tropical Pacific Ocean. Temporal variabilities (sea level anomalies) are computed relative to the 1-year mean sea level to avoid uncertainties associated with the marine geoid.

3. Model

The model used in this study is based on a finite-difference shallow water model by R. Pacanowski of the Geophysical Fluid Dynamics Laboratory, NOAA. The model is a reduced-gravity shallow water model, linearized about the state of rest with parameters corresponding to typical values of the first baroclinic mode of the equatorial Pacific Ocean (Cane, 1984); The layer depth and gravity are 281 'm and 3.02 cm/s^2 , respectively, which give a wave speed of 2.91 m/s. The model domain (Figure 1) extends zonally across the Pacific basin, but limited meridionally between 30° S and 30°N. The model grid resolution is 2° zonally and 10 meridionally, amounting to a total state dimension of 11,940 variables. All horizontal boundaries are treated as impermeable with no-slip boundary conditions. Horizontal Laplacian friction is used with a viscosity of $10^7 \text{ cm}^2/\text{s}$, except within 5° of the boundaries, where it is linearly increased by a factor of 10. At the western boundary viscosity is increased by a factor of 100 within 10° of the boundary to dampen short waves unresolvable by the model grid, The model is forced by the NMC wind analyses. To reduce the storage requirements of these winds, a 3 day bin averaging was applied to the 12-hourly winds, which was then linearly interpolated in time during model integration.

4. Assimilation Method

A quasi-optimal Kalman filter and smoother will be employed to analyze the altimetric observations, Although statistically optimal, direct application of Kalman filtering (e.g., Gelb, 1974) to ocean data analyses are computationally prohibitive because of its requirement to evaluate the time evolving estimation error covariance matrix that make up the filter. Fukumori and Malanotte-Rizzoli (1995) explored a Kalman filter employing a transformation that approximates the model state with one that has fewer degrees of freedom, thus effectively reducing the size of the estimation problem. The transformation was chosen so as to resolve the large-scales of the model, with the aim of making the largest improvement in the model with the minimum amount of

computation by estimating scales containing the dominant part of the uncertainties. An additional approximation was made by using the asymptotic limit of the time-evolving error estimate in the filter (Fukumori et al., 1993) thereby further reducing the storage and the amount of computation in performing estimation. While these approximations lead to near optimal estimates under the assumptions, the main advantage of the method is that it provides a practical yet objective method to compute the filter and smoother.

In essence, the reduced-dimension static approximate filter improves the dominant scales of the model estimate by mapping the data-model misfit (innovation) to such scales for which statistical quantities are easier to evaluate because of their smaller dimension. Let us define a linear transformation \mathbf{B} relating the model state \mathbf{x} to one that has a smaller dimension \mathbf{x}' such that,

$$\mathbf{x}(t) \approx \mathbf{B}\mathbf{x}'(t) + \bar{\mathbf{x}} \quad (1)$$

Here $\bar{\mathbf{x}}$ is some prescribed time-invariant reference state and t is time. Then the filter can be approximated using the asymptotic error covariance matrix of the filtered \mathbf{x}' estimate (\mathbf{P}') in place of the error of \mathbf{x} ($\mathbf{P}(t)$) as

$$\mathbf{K}(t) = \mathbf{P}(t)\mathbf{H}^T(t)\mathbf{R}^{-1}(t) \approx \mathbf{B}\mathbf{P}'\mathbf{B}^T\mathbf{H}^T(t)\mathbf{R}^{-1}(t) \quad (2)$$

Matrix $\mathbf{H}(t)$ is an operator such that $\mathbf{H}(t)\mathbf{x}(t)$ is the model's equivalent of the observations and $\mathbf{R}(t)$ is the data error covariance matrix. Error \mathbf{P}' is evaluated from a model for \mathbf{x}' that is obtained by combining (1) with the dynamic model for \mathbf{x} . Namely, denoting the state transition matrix for \mathbf{x} as \mathbf{A} (assuming a linear model for simplicity), the transition matrix for \mathbf{x}' is $\mathbf{B}^*\mathbf{A}\mathbf{B}$ where \mathbf{B}^* is the pseudo inverse of \mathbf{B} .

While Fukumori and Malanotte-Rizzoli (1995) only examined an approximate filter, a similar approximation can be made to the smoother \mathbf{S} as well. The static fixed interval smoother of Fukumori et al. (1993) improves filtered estimates extracting information contained in formally future observations by sequentially mapping differences between smoothed estimates and dynamic updates backwards in time. Using (1) and its inverse, these differences can be mapped via the coarse space (\mathbf{x}') similar in fashion to the filter (2) by approximating the asymptotic smoother as

$$\mathbf{S} = \mathbf{P}\mathbf{A}^T\mathbf{P}^{-1} \approx \mathbf{B}\mathbf{P}'(\mathbf{B}^*\mathbf{A}\mathbf{B})^T[\mathbf{P}'_+]^{-1}\mathbf{B}^* \quad (3)$$

Matrix $\mathbf{P}'_-(\mathbf{P}_-)$ is the asymptotic limit of the dynamic updated error estimate of \mathbf{x}' (\mathbf{x}).

For the present shallow water model, the coarse state \mathbf{x}' was chosen based on scales of the model's variability y . As in Fukumori and Malanotte-Rizzoli (1995), empirical orthogonal functions for each independent state variable were computed based on results of a model simulation, and analyzed to determine the dominant scales of variability. Approximately 20 zonal and 14 meridional modes are necessary to resolve 95% of the variability y , and a coarse grid was chosen accordingly as shown by the dots in Figure 1. The coarse grid has 22 zonal and 15 meridional grid points with a 7.5° and 4° longitudinal and latitudinal resolution, respectively. The meridional resolution is increased to 3° near the equator. The alternate state consists of state variables on this coarse grid with a resulting total dimension of 831. Transforming the state on the coarse grid to the model's $2^\circ \times 1^\circ$ grid (B) is performed by objective mapping (Bretherton et al., 1976), using Gaussian correlation functions with zonal and meridional correlation distances of 7.8° and 4° , respectively.

5. Model and Data Error Estimates

Estimates of data error and model error (and in turn the process noise) used in deriving the filter and smoother can be made from a comparison between observations and model simulation (i.e., model run without assimilation) [Fu et al., 1993]. The model is a projection of the ocean physics with infinite degrees of freedom (d.o.f.) to one that has a finite d.o.f. Then model errors (viz., errors of the model variables) are the discrepancies between this projection and the model state as computed by the model algorithm, and comprises the effects of any error Meeting the computed evolution of the model state. For example, these include effects of forcing error, error in parametrization (e.g., layer thickness in a linear shallow water model), neglect of nonlinearity in linear models, and numerical truncation error in finite differencing, which are all in essence "(errors of model physics".

Data errors on the other hand are the discrepancies between the measurements and the model equivalent of the observations in the absence of model errors. This not only includes errors due to the measuring system but also results of physics not included in the model ("missing physics"). For example, for the present shallow water model, sea

level changes associated with instability waves (Perigaud, 1990) are likely present in the data but absent in the model and thus are part of data error.

A simulation was carried out from January 1, 1992 till October 31, 1993 following a spin-up over 1500 days forced by time-mean winds of the analysis period. Instantaneous model fields are saved every 3-days and are spatially interpolated and compared with TOPEX measurements in 3-day windows assuming the two are temporally coincident. The 3-day window was chosen because it is close to one of the subcycle periods of TOPEX. The model equivalent of sea surface height anomalies were computed relative to the simulation's time-mean sea surface corresponding to the period of the observations.

Figure 2 shows time series of the variance estimates of data error and the simulation's sea level error (Fu et al., 1993, equations (20) and (21)). Also shown are variances of the measurements and the simulation. On average, the model error variance is as large as half the total model variance and the measurements are dominated by data error. The time-averaged data and simulation error estimates are 35.3 cm^2 and 9.6 cm^2 , whereas the average of data and simulation variance are 45.6 cm^2 and 19.8 cm^2 , respectively. By assimilation, we aim to create a more accurate model estimate by dynamically and statistically averaging the model and data.

For simplicity, errors of data and simulation will be both modeled as being time-invariant, with mean variances equaling the corresponding estimates above. The measurement error part of data error was modeled as having a Gaussian covariance function with a 1300 km e-folding scale along the satellite ground track and a 23.65 cm^2 variance, comprised of orbit error (2.5 cm root-mean-square), tides (3.8), water vapor (1), pressure loading (1), and EM bias (1). Error covariances between data from different satellite passes were assumed to be independent. The remaining part of data error (11.65 cm^2) is ascribed to missing physics and was treated as being spatially uncorrelated.

Errors of the simulation depend on model dynamics and its process noise (the incremental error of the model). Process noise was modeled in the form of wind error with a Gaussian covariance among the pseudo stress components. Correlation distances are assumed to be 10° zonally and 2° meridionally and the variance (W) was modeled as a function of wind speed error σ as (Miller and Cane, 1989)

$$w = 4\sigma^2 u_0^2 + 3\sigma^4 \quad (4)$$

where U is the true wind speed. The correlation between errors of meridional and zonal stresses were assumed to be independent of each other, and the time-mean NMC winds were used in place of U . The asymptotic error was computed assuming a typical 3-day observation pattern (centered on May 16, 1993) with winds completely correlated over 3-days but independent from one three-day period to the next. A comparable mean model error to that shown in Figure 2 is obtained with $\sigma = 2.2$ (m/s).

6. Results of Assimilation

Figures 3, 4, and 5 show sea level anomalies in a longitude versus time plot along the equator, 12.5°N , and 12.5°S , respectively. The gaps and noise in the observations are readily apparent in the $2^\circ \times 10$ spatial and 3-day temporal resolution of the plots. The smoothed estimate filters out these noise and interpolates over the gaps resolving most features in the data, such as the eastward propagating signals associated with Kelvin waves and westward propagating Rossby waves. Seasonal differences are also evident at 12.5°N associated with changes of the North Equatorial Countercurrent, which are evidenced in the sea level anomaly maps shown in Figure 6.

Estimates made by the simulation show several qualitative similarities with TOPEX data, but significant quantitative differences are also evident. For example, "several signals are stronger in the smoothed estimate than in the simulation, such as the negative eastward propagating anomalies along the equator (days 51 OW600, Figure 3) and the seasonal changes in Figure 4. On the other hand, the simulation overestimates changes along 12.5°S (Figure 5) especially those east of 250°E , which are associated with Rossby waves generated by reflecting equatorial Kelvin waves at the eastern boundary. These reflected Rossby waves are not as strong in the smoothed estimate as they are in the simulation (Figure 6).

Correlation maps between the estimates and TOPEX data are shown in Figure 7. The simulation has significant correlations with the altimetric measurements in the central equatorial Pacific which are largely due to the model's ability to simulate wind-driven equatorial waves in these regions (Figures 3 and 4). However, there are vast areas over which the correlations are small and even negative. In contrast, the assimilation modifies the model in accordance with what is observed resulting in higher correlations over the entire region.

Although much improved, there are somewhat lower correlations in some regions even after assimilation. Larger discrepancies found in the western end of the model is likely due to effects of islands not resolved by the model and the excess damping required for the model's numerical stability. The southeastern region also has smaller correlation than the rest, but the signal variance in this area is small (Figure 8) and the discrepancy is likely due to the smaller signal to noise ratio. On the other hand, the area along the eastern boundary has small correlation even though there are significant amounts of observed sea level variability, which suggests possibly larger model deficiencies in this area than assumed by the modeled process noise.

Statistical consistencies must be examined to assess the validity of the estimates. Otherwise, for instance, models can be fit arbitrarily close to the observations by assuming an excessively large model error \mathbf{P} for the filter [Equation (2)]. The first requirement is the consistencies between the assumed data and process noise with the differences between model simulation and observations, which was performed in Section 5. The second check is an a posteriori test of the statistical consistencies. The differences between the dynamic prediction from the filtered sea level and data are the sum of the filter's error and data noise. The filter's residual has a variance of 34 cm^2 whereas the sum of its expected errors and data error are 37 cm^2 . Thus the filter is actually performing slightly better than what is expected and indicates a first order consistency with the assumed error statistics.

Figure 9 shows the amount of sea level variance accounted for by the different estimates as a function of time. Explained variance is estimated as data variance minus the variance of the residuals (i.e., differences between data and model) [e.g., Fu et al., 1993]. Also shown in Figure 9 is an estimate of signal variance, computed as data variance minus measurement error variance. The simulation occasionally has positive skill, but on average accounts for only 3% of the expected signal. In contrast, the smoothed estimate has a positive skill throughout the one-year period with an average skill of 78%. The increased skill is largely due to the dynamic model carrying information consistently from the observations forwards and backwards in time. For instance, the effect of the data at each instant (data update; Fu et al., 1993) is small compared with the total skill of the smoothed estimate. Thus, the skill of the smoother is not dominated by instantaneous measurements and the model is not excessively forced towards the observations.

Accuracies of the assimilated estimates can be further assessed through comparisons with independent ancillary observations. For example, Figure 10 compares sea level measurements from a tide gauge at Majuro ($171^{\circ}\text{E}, 7^{\circ}\text{N}$) with those from the simulation and smoother. While the simulation misses most events, variabilities estimated by the smoother coincide with many of those seen in the tide gauge record. The correlation coefficient of the tide gauge record at Majuro with the smoothed and simulated estimates are 0.78 and 0.22, respectively. Table 1 summarizes correlations of the simulated sea level and the smoothed estimate with daily tide gauge records obtained from the TOGA tide gauge center. The comparisons are limited to tide gauges within 20° of the equator within the model domain, which positions are shown in Figure 7. Practically all tide gauge measurements have higher correlation with the smoothed estimate than the simulation and are comparable with correlations between the model estimates and TOPEX observations (Figure 7). Those that do not improve much are tide gauges where the model simulation already performs well. The improvements demonstrate the affect of the assimilation in correcting model simulation. However, although improved, some correlations remain small even after smoothing and are underlined in Table 1. The corresponding locations are shown in Figure 7 by triangles. The three near New Guinea (Port Moresby, Lae, Booby Island) are likely due to the lack of the model's spatial resolution in resolving the coast line and islands in this region, and the model's comparison with TOPEX (Figure 7) are also small in this vicinity. The result at Yap is somewhat unclear but may also be due to the lack of spatial resolution in the model. Correlations between model and TOPEX are high in the area surrounding Yap, as well as that between model and tide gauge at nearby Malakal.

Comparisons of the assimilated estimates with tide gauge measurements are encouraging and indicate an increased accuracy by the assimilation. However, in spite of different error sources and the importance of dynamic effects in reconciling altimetric and tide gauge measurements (eg., Mitchum, 1994), tide gauges and altimeters are to some extent redundant in that both essentially measure the same quantity, namely sea level. An additional test of the model estimate's accuracy would be a comparison with a quantity physically independent of what has been assimilated. Figure 11 shows such a comparison between model velocity at 140°W on the equator with current meter data from the TOGA TAO array. Even though only sea level measurements are assimilated, the smoothed estimate's velocity is in better accord with the direct current

measurement than the simulation's result is. Correlation coefficients of this velocity measurement with the estimates made by the smoother and simulation are 0.44 and -0.18, respectively.

Correlation with other TAO current meter measurements are summarized in Table 2. As with the tide gauge comparisons, correlations with in situ current measurements are generally higher with the smoothed estimates than with simulated estimates, demonstrating the improvements made by the assimilation. Variations in the estimates' skill are to a large extent due to what the model dynamics can resolve. For example, vertical variations of the correlation coefficients are due to the model's ability to only resolve the first baroclinic mode, while the current meter observes cent ribut ions of higher modes as well. Correlations for the meridional components (v) are generally smaller than for the zonal components (u), and are also due to the model's physics, Figure 12 shows a comparison of meridional velocity at 140°W . Unlike the zonal component (Figure 11), meridional velocity is dominated by high frequency variations barely resolvable by the 3-day analysis interval employed in this study. The variations of the meridional velocity are also smaller than those of the zonal component. These differences between u and v reflect the zonal and meridional asymmetry in equatorial dynamics. In the low-frequency, long zonal-scale limit, v is necessarily much smaller than u (e.g., Cane and Patton, 1984). These differences are also reflected in the accuracy estimates given in Table 3, which are computed according to the approximations described in Sections 4 and 5. The expected improvements as a result of TOPEX sea level assimilation are much larger for zonal velocity than those for the meridional velocity. The smaller amplitude and higher frequency makes improvements of meridional velocity along the equator by altimetric assimilation more difficult than improving zonal circulation.

The nature of the estimate's velocity variations can be inferred further from a longitude vs time plot along the equator shown in Figure 13. The time and space scales of the zonal velocity (u) are similar to those of the sea level variations (h , Figure 3-b). Coincident eastward propagating signals have the same sign between u and h whereas westward propagating variations have opposite sign, both consistent with being associated with equatorial Kelvin and Rossby waves, respectively. In contrast, meridional velocity (v) is dominated by high frequency variations as was noted in Figure 12. The dominant variations in v are eastward propagating signals with a speed of about 1.0 (m/s), which is suggestive of being associated with a mixed Rossby-gravity

wave; A mixed Rossby-gravity wave with a ~ 3500 km wave length (westward phase propagation) and a 12.5-day period has an equivalent eastward group velocity.

While different data sets are useful in evaluating the estimates, the model in turn can be used to assess accuracies of different data sets themselves. One of the primary error sources in the TOPEX GDR is believed to be in oceanic tidal corrections. The results presented above were based on the Cartwright and Ray model plus an empirical residual tidal correction of its M2 and K1 components. Table 4 summarizes an analysis of the present shallow water model's consistency with TOPEX measurements using different tidal corrections. Tidal models of Cartwright and Ray (1990) and Schwiderski (1980) are examined along with two-component (M2 and K1) and four-component (M2, K1, S2, O1) empirical corrections derived separately for the two tidal models by Fu and Pihos (1994, unpublished manuscript). Despite large differences in the tidal corrections (data variance), the amount of variance explained by the assimilated estimates have relatively little differences, which demonstrates the robustness of the estimation. Yet, there are small but discernible differences among the different model estimates. Figure 14 shows a map of the differences in the amount of explained variance between using the two-component and four-component empirical correction of the Cartwright and Ray model. The former is more consistent with the model than the latter in the central Pacific suggesting that the four-component empirical correction may have removed some of the nontidal variability from the data. On average, the Cartwright and Ray tidal corrections plus the empirical 2-component residual tidal correction appears to have least tidal errors with most nontidal variations preserved.

7. Summary and conclusions

The first year of the TOPEX altimetric sea level observations over the tropical Pacific Ocean were analyzed by way of assimilating the measurements into a wind-driven reduced-gravity shallow water model. Although the model simulation has qualitative similarities with the altimetric measurements, significant quantitative differences are found. These were in general a too weak a sea level variation in the central Pacific and a too strong a response near the eastern boundary. The assimilation corrects the model estimate by extracting dynamic signals from the data that are consistent with model physics resulting in a space and time continuous dynamic interpolation and extrapolation of the measurements. The analysis is effective in resolving observed

sea level variations associated with equatorial Kelvin and Rossby waves and seasonal changes due to variations in the strength of the North Equatorial Countercurrent. A reasonable amount of wind error (2.2 m/s) can account for most of the model's deficiencies, and on average, 78% of the observed sea level variance can be explained by the wind-driven shallow water model.

The assimilation is performed with an approximate Kalman filter and smoother. While Fukumori and Malanotte-Rizzoli (1995) studied the approximate filter in a twin experiment, this study is the first demonstration of the utility of such approximation and the application of the approximate smoother in analyzing real observations. Although the model is relatively small and simple, the state dimension (11940) is extremely large for estimation. Without the approximations, each state error covariance matrix would have required 143 Mw of storage (119402 variables) and evaluation of its 1-year time evolution would have required 266 CPU hours on a Cray Y-MP. (The model requires 40 CPU seconds to run a one year simulation.) In comparison, the approximation requires 0.7 Mw to store \mathbf{P}' and 350 CPU seconds to compute its asymptotic limit.

The assimilated estimates were shown to be consistent with the observations within their expected uncertainties and comparisons with independent observations from tide gauges and current meters validate the improvements made by the assimilation. These results demonstrate the first order dynamic and statistical consistencies of the estimation method.

Although the model was shown to account for a major portion of the observed altimetric sea level variation, regions near the eastern boundary had larger model deficiencies than the rest even after assimilation (Figure 7) for which no simple explanation was identified. The eastern boundary region is also the area where model simulation had anomalously large Rossby waves reflected from the eastern boundary but not observed in the data (Figures 5 and 6). Although much of these deficiencies were accounted for by the assumed amount of wind error, the remaining differences may still be due to wind error in excess of what was modeled. For instance, wind along Peru may have larger errors than other regions due to the tall Andes mountains that are difficult to resolve in truncated atmospheric spectral models. Unaccounted errors in the altimetric data are always possible, but comparisons of TOPEX data with available tide gauge measurements are good (Mitchum, 1994). On the other hand, even though the height

measurements are accurate, there may be excess errors in the tidal corrections in these regions. For instance, a comparison of the 2 and 4-component empirical correction of the tidal residuals showed a larger model skill along the eastern boundary when the 4-component correction was used (Figure 14). Yet another possible reason for the model deficiency in the eastern region may be errors in the linearized reduced-gravity dynamics themselves. Baroclinic waves may be scattered by changes in bottom topography, which the reduced-gravity model cannot simulate. Spatial variations in the vertical density stratification may also affect the evolution of baroclinic waves, as well as nonlinear interactions with the background mean flow. These error sources are difficult to discern with the present model and data set, and resolution will have to wait for additional observations including better tidal models as well as using models with more physics for the analyses. More accurate tidal models are under development as part of the TOPEX/POSEIDON project and will soon become available. Global observations of winds will also become available with the launch of the NASA Scatterometer (NSCAT) on board the Japanese satellite ADEOS (Advanced Earth Observation Satellite) in 1996.

In spite of the assimilation's success in improving velocity estimates from altimetric data alone, the model is obviously too simple to fully resolve the circulation. A three-dimensional nonlinear model will be necessary to simulate effects due to the full suite of equatorial currents, especially their effects on the vertical dependence of the circulation. The first baroclinic mode resolved in the present study is just one component among the vertical degrees of freedom, but other components are also observable from altimetric observations (Fukumori et al., 1993). The approximate filter and smoother make application of estimation theory to models with larger dimension and more physics feasible, and are currently under way.

On the other hand, estimation of the circulation's transient meridional component may remain challenging even with the enhanced model dynamics because of its dominance of high frequencies in comparison with the sampling characteristics of a single satellite. The addition of data from other satellites flying simultaneously during the TOPEX/POSEIDON mission such as ERS-1 and the scheduled ERS-2 and GFO (Geosat Follow-On) would compensate the amount of sampling and might be effective. In situ data such as the TOGA TAO array provide a high rate of temporal sampling and may also be assimilated instead of merely being used as a means of validation.

Although the present analysis was performed at three-day intervals partly for computational convenience, a time-continuum analysis, namely assimilation at the exact times that the measurements are made, may prove effective as well. The approximate filter and smoother, Equations (2) and (3), can be applied at a much finer temporal interval without increase in computational requirements.

The present results are encouraging and prospects for estimating the three dimensional circulation are promising.

Acknowledgements

The author is grateful to Lee-Lueng Fu for providing processed TOPEX data and the empirical tidal corrections. Greg Pihos prepared the corrected TOPEX data over the model domain. The author is also indebted to Yi Chao for his help with the shallow water model. Tide gauge data were provided by the International TOGA Sea Level Center, Univ. of Hawaii. TAO data were made available by the TOGA-TAO Project Office, Dr. Michael J. McPhaden, Director. This research was carried out in part by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Computations were performed on the Cray Y-MP of the JPL Supercomputing Project, which is sponsored by the NASA Offices of Mission to Planet Earth, Aeronautics and Space Science.

References

- Bretherton, F. P., R. E. Davis, and C. B. Fandry, 1976. A technique for objective analysis and design of oceanographic experiments applied to MODE-73, *Deep-Sea Research*, **23**, 559-582.
- Busalacchi, A. J., and J. J. O'Brien, 1981. Interannual variability of the equatorial Pacific in the 1960's, *Journal of Geophysical Research*, **86**, 10,901-10,907.
- Cane, M. A., 1984. Modeling sea level during El Niño, *Journal of Physical Oceanography*, **14**, 1864-1874.
- Cane, M. A., and R. J. Patton, 1984. A numerical model for low-frequency equatorial dynamics, *Journal of Physical Oceanography*, **14**, 1853-1863.
- Cartwright, D. E., and R. D. Ray, 1990. Oceanic tides from Geosat altimetry, *Journal of Geophysical Research*, **95**, 3069-3090.
- Fu, L.-L., L. Fukumori and R. N. Miller, 1993. Fitting dynamic models to the Geosat sea level observations in the Tropical Pacific Ocean. Part II: A linear, wind-driven model, *Journal of Physical Oceanography*, **23**, 2162-2181.
- Fu, L.-L., E. J. Christensen, C. A. Yarnarone Jr., M. Lefebvre, Y. Ménard, M. Dorrer, and P. Escudier, 1994. TOPEX/POSEIDON mission overview, *Journal of Geophysical Research*, **99**, 24,369-24,381.
- Fukumori, I., J. Benveniste, C. Wunsch and D. B. Haidvogel, 1993. Assimilation of sea surface topography into an ocean circulation model using a steady-state smoother, *Journal of Physical Oceanography*, **23**, 1831-1855.
- Fukumori, I., and P. Malanotte-Rizzoli, 1995. An approximate Kalman filter for ocean data assimilation; An example with an idealized Gulf Stream model, *Journal of Geophysical Research*, (in press).
- Gelb, A., 1974. "Applied Optimal Estimation", M. LT. Press, Cambridge, MA 374 pp.
- Miller, R. N., and M. A. Cane, 1989. A Kalman filter analysis of sea level height in the tropical Pacific, *Journal of Physical Oceanography*, **19**, 773-790.
- Mitchum, G. T., 1994. Comparison of TOPEX sea surface heights and tide gauge sea levels, *Journal of Geophysical Research*, **99**, 24,541-24,553.

- Perigaud, C., 1990. Sea level oscillations observed with Geosat along the two shear fronts of the Pacific North Equatorial Countercurrent, *Journal of Geophysics/ Research*, **95**, 7239–7248.
- Schrama, E. J. O., and R. D. Ray, 1994. A preliminary tidal analysis of TOPEX/POSEIDON altimetry, *Journal of Geophysical Research*, **99**,24,799-24808.
- Schwiderski, E. W., 1980. Global ocean tides. 1: A detailed hydrodynamic interpolation model, *Marine Geodesy*, **3**, 161--217.
- Stammer, D., and C. Wunsch, 1994. Preliminary assessment of the accuracy and precision of TOPEX/POSEIDON altimeter data with respect to the large-scale ocean circulation, *Journal of Geophysical Research*, **99**, 24,584-24,604.
- Zebiak, S. E., and M. A. Cane, 1987. A model El Niño–Southern Oscillation, *Monthly Weather Review*, **115**, 2262-2278.

Table Captions

Table 1: Correlation coefficients of daily tide gauge measurements with model simulated and smoothed estimates. Check marks on the final column denote stations with improved correlation after assimilation. Geographic position of the tide gauges are denoted in Figure 7 as crosses. Underlined values in the last column are places where the correlation remains small and are denoted in Figure 7 as triangles.

Table 2: Correlation coefficients of TOGA-TAO daily current meter measurements with model simulated and smoothed estimates. Check marks on the smoothed estimates denote current meters with improved correlation after assimilation.

Table 3: Error estimates of model velocities (cm/s) at the TOGA-TAO current meter mooring locations. Within each parentheses, the first and second numbers correspond to the zonal and meridional components of the velocities, respectively.

Table 4: Data variance and the amount of which explained by the assimilation using different tidal corrections (cm²). The tidal models are Cartwright and Ray (1990; C&R) and Schwiderski (1980; Sch) plus additional empirical corrections for residual errors. See text for a description.

Figure Captions

Figure 1: Model domain and coarse grid. The solid line denotes the boarder of the model domain. The dots are locations of the coarse grid on which the reduced model state is defined. The total number of coarse grid points is 277.

Figure 2: Estimates of error variance (cm^2) computed over each 3-day period; a) data error, b) simulation error. The dashed curves are the data variance (a) and simulation variance (b), respectively.

Figure 3: Longitude vs time plot of sea level anomalies along the equator; a) TOPEX, b) smoothed estimate, c) simulation. Color scale is between ± 10 cm. Gray denotes missing values either because of land or missing data (in case of TOPEX). The resolution of the plots are 2° by 10 horizontally and 3-days in time. TOPEX data shown in a) were generated by averaging all available data within the $2^\circ \times 10$ and 3-day bins.

Figure 4: Same as Figure 3, except along 12.5°N .

Figure 5: Same as Figure 3, except along 12.5°S .

Figure 6: Sea level anomaly estimates for June 20, 1993; (a) simulation, (b) smoothed estimate.

Figure 7: Correlation maps between TOPEX observations and model sea level variations; a) simulation, b) smoothed estimate. The crosses and triangles denote tide gauge stations (Table 1).

Figure 8: Variance of TOPEX data (cm^2).

Figure 9: Amount of sea level variance accounted for by model estimates; simulation (thick dotted), smoother (thick solid). Also shown are signal variance (thin solid) and effect of data update (thin dotted). The data update is the effect of the data at each instant,

Figure 10: Comparison of sea level variability at Majuro ($171^\circ\text{E } 7^\circ\text{N}$) with model estimates; tide gauge (thin solid), smoother (thick), simulation (thick dotted). Daily tide gauge measurements are corrected for effects of atmospheric pressure loading assuming an inverse barometer, and a time-mean of the result is further removed from the data. Atmospheric pressure data are derived from the NMC estimate.

Figure 11: Comparison of zonal velocity at 140°W : TAO mooring at 25 m (black), smoother (blue), simulation (red). The TAO data are variations from its time-mean, while model estimates are relative to the simulation's time-mean.

Figure 12: Comparison of meridional velocity at 140°W : TAO mooring at 200 m (black), smoother (blue), simulation (red).

Figure 13: Longitude vs time plot of smoothed velocity anomalies (cm/s); a) zonal component, b) meridional component, Note that the color scales are different between the two,

Figure 14: Differences in the smoother's skill (cm^2) using different empirical corrections to the Cartwright and Ray tidal model. The two corrections are a 2-component (M2, K1) correction and a 4-component (M2, K1, S2, O1) correction. Positive values indicate that the model explains more variance with the former than the latter.

Station	Position	Sire.	Smth	
Pohnpei	(7.0°N, 158.2°E)	0.65	0.63	
Tarawa	(1.4°N, 172.9°E)	0.61	0.60	
Baltra	(0.4°s, 90.3°W)	0.50	0.70	✓
Nauru	(0.5°S, 166.9°E)	0.39	0.66	✓
Majuro	(7.1°N, 171.4°E)	0.22	0.78	✓
Malakal	(7.3°N, 134.5°E)	0.56	0.61	✓
Yap	(9.5°N, 138.1°E)	-0.10	<u>-0.06</u>	✓
Honiara	(9.4°s, 160.0°E)	-0.06	0.45	✓
Rabaul	(4.2°S, 152.2°E)	0.36	0.75	✓
Christmas	(2.0°N, 157.5°W)	0.54	0.74	✓
Kanton	(2.8°S, 171.7°W)	0.45	0.74	✓
Penrhyn	(9.0°S, 158.1°W)	0.34	0.59	✓
Funafuti	(8.5°S, 179.2°E)	0.46	0.64	✓
Saipan	(15.2°N, 145.7°E)	0.73	0.68	
Kapingarnarangi	(1.1°N, 154.8°E)	0.54	0.72	✓
Santa Cruz	(0.8°S, 90.3°W)	0.51	0.71	✓
Nuku Hiva	(8.9°S, 140.1°W)	0.18	0.24	✓
Wewak	(3.6°S, 143.6°E)	0.32	0.57	✓
Port Moresby	(9.5°S, 147.1°E)	-0.50	<u>-0.02</u>	✓
Manus	(2.0°S, 147.3°E)	0.54	0.65	✓
Madang	(5.2°S, 145.8°E)	-0.13	0.46	✓
Lae	(6.7°S, 147.0°E)	-0.32	<u>0.03</u>	✓
Kavieng	(2.6°S, 150.8°E)	0.39	0.59	✓
Alotau	(10.2°S, 150.4°E)	-0.09	0.45	✓
Anca	(18.5°S, 70.3°W)	0.57	0.55	
Lobos de Afuera	(6.9°S, 80.7°W)	0.51	0.51	
Quepos	(9.4°N, 84.2°W)	-0.03	0.37	✓
Callao	(12.1°s, 77.1°W)	0.54	0.65	✓
Balboa	(9.0°N, 79.6°W)	-0.20	0.33	✓
Booby Island	(10.6°S, 141.9°E)	-0.66	<u>-0.58</u>	✓

Table 1

Longitude	Depth	Simulation		Smoothed	
		u	v	u	v
165°E	10 m	0.63	0.38	0.79 ✓	0.424
	50 m	0.71	0.49	0.844	0.50 ✓
	100 m	-0.20	0.22	0.37 ✓	0.34 ✓
	150 m	0.31	0.34	0.51 ✓	0.33
	200 m	0.24	0.27	0.43 ✓	0.27
	250 m	0.03	-0.01	0.21 ✓	-0.03
140°W	25 m	-0.18	-0.19	0.44 ✓	-0.05 ✓
	80 m	-0.46	-0.03	0.19 ✓	0.05 ✓
	120 m	0.18	0.05	0.53 ✓	0.06 ✓
	200 m	0.50	0.11	0.46	0.19 ✓
110°W	80 m	0.22	0.12	0.34 ✓	0.21 ✓
	200 m	0.43	0.13	0.60 ✓	-0.01

Table 2

Position	Simulation	Smoothed
165°E	(14.7, 4.5)	(3.5, 2.8)
140°w	(15.3, 8.4)	(4.7, 3.5)
110°W	(13.2, 10.7)	(4.89 3.6)

Table 3

Applied Tidal Correction	Data Variance	Model Skill
Cartwright and Ray	56.2	16.7
C&R + M2-K1	45.6	17.1
C&R + M2-K1-S2-O1	40.8	16.4
Schwiderski	63.7	15.5
Sch + M2-K1	47.1	16.3
Sch + M2-K1-S2-O1	42.1	16.1

Table 4

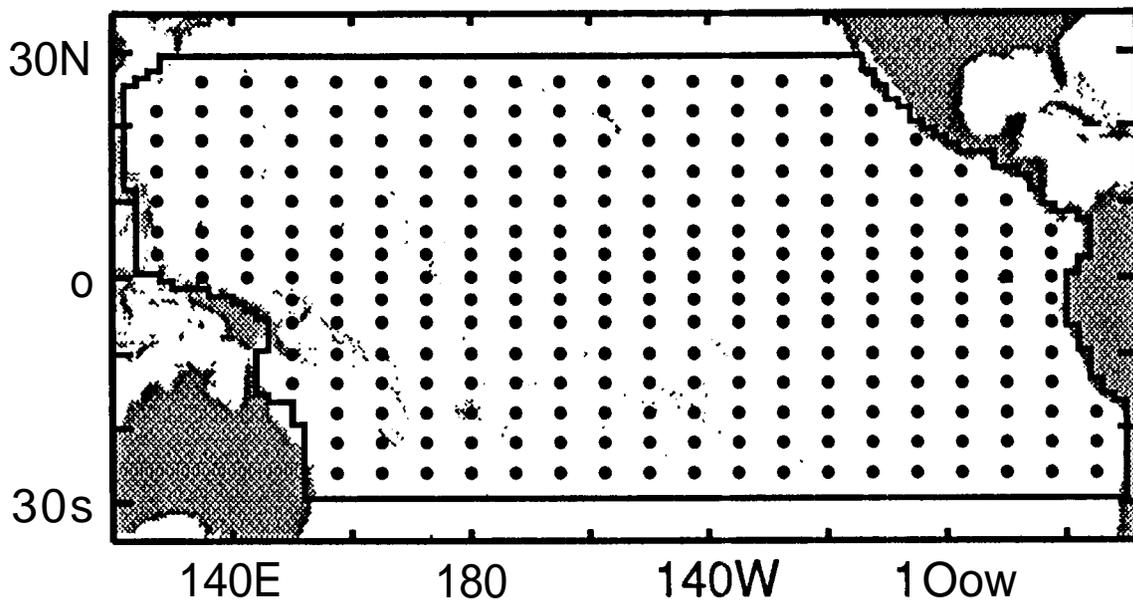


Fig. 1

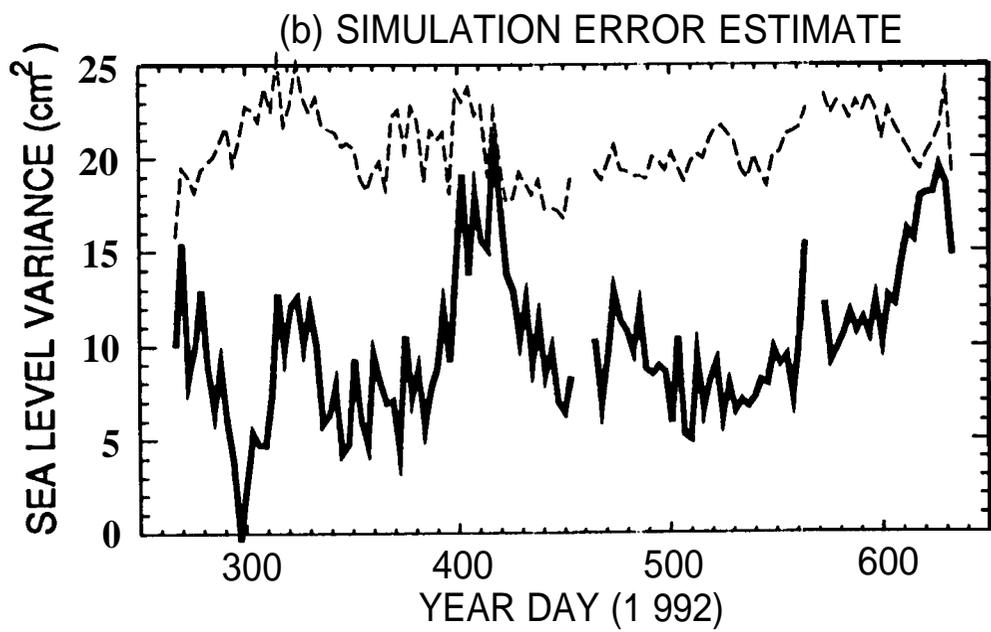
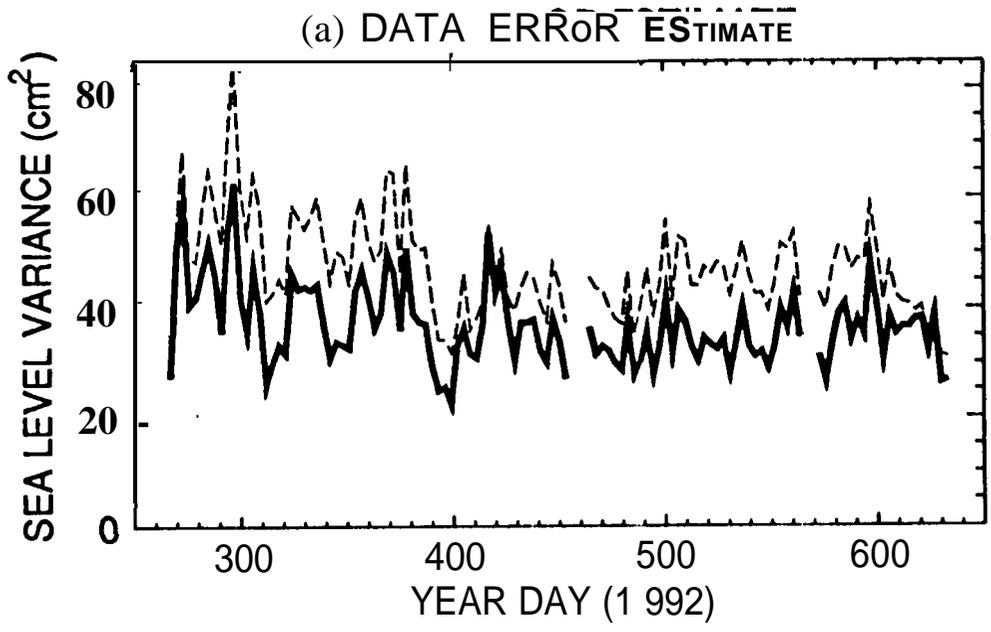


Fig. 2

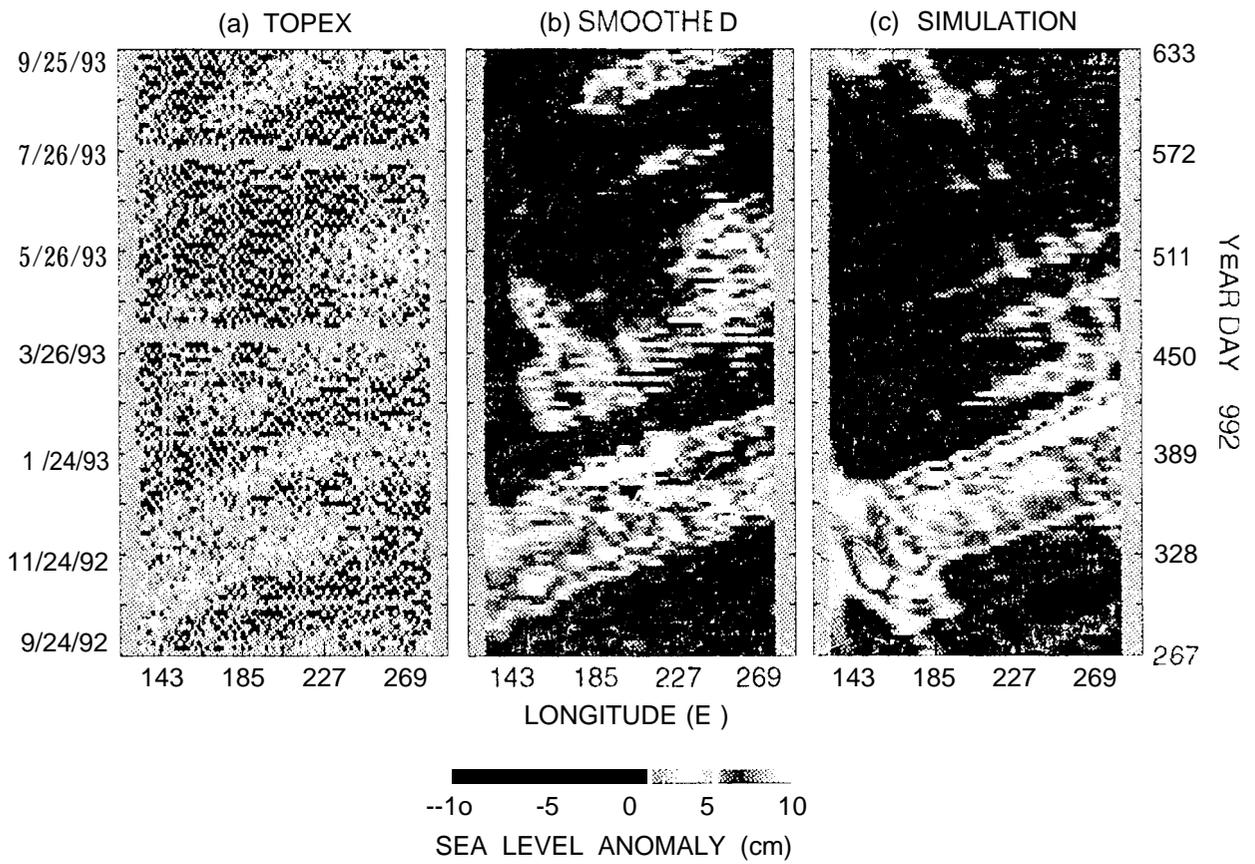


Fig. 3

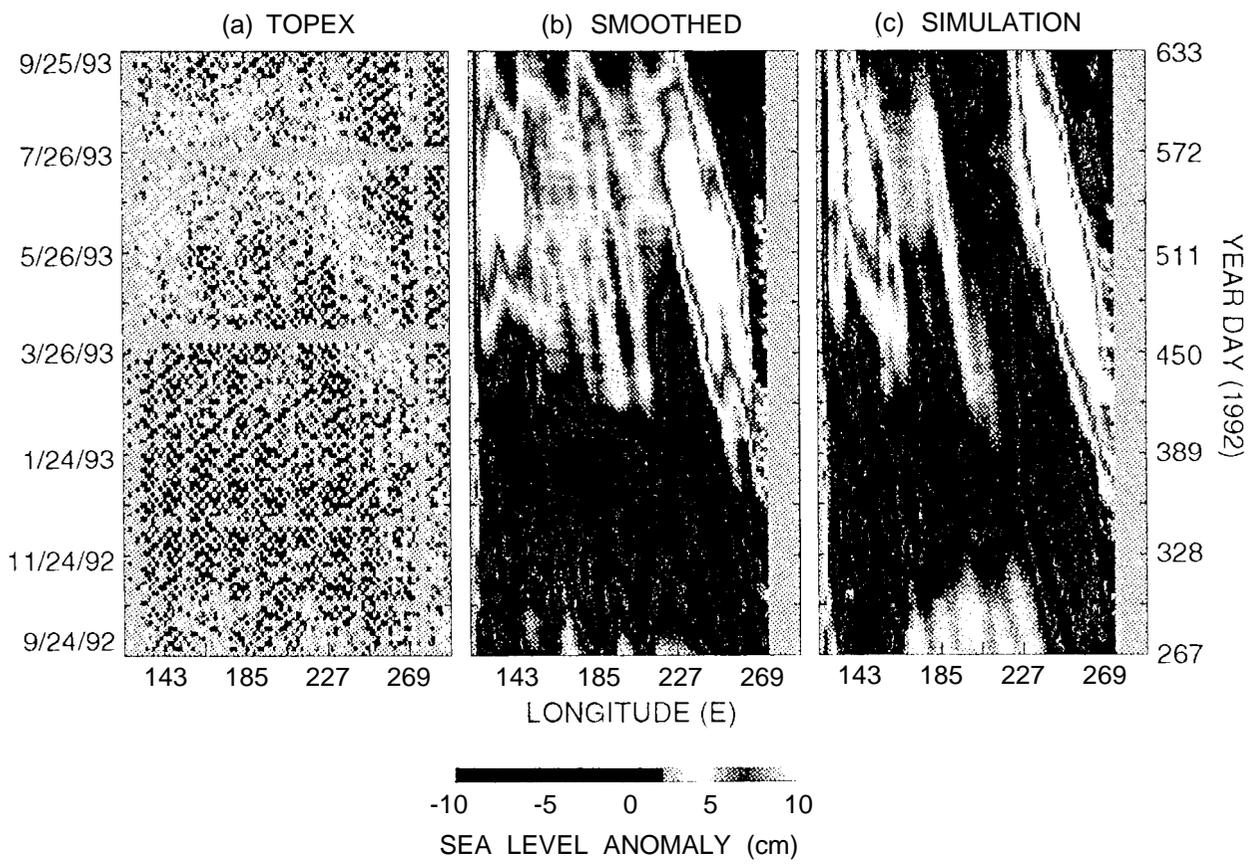
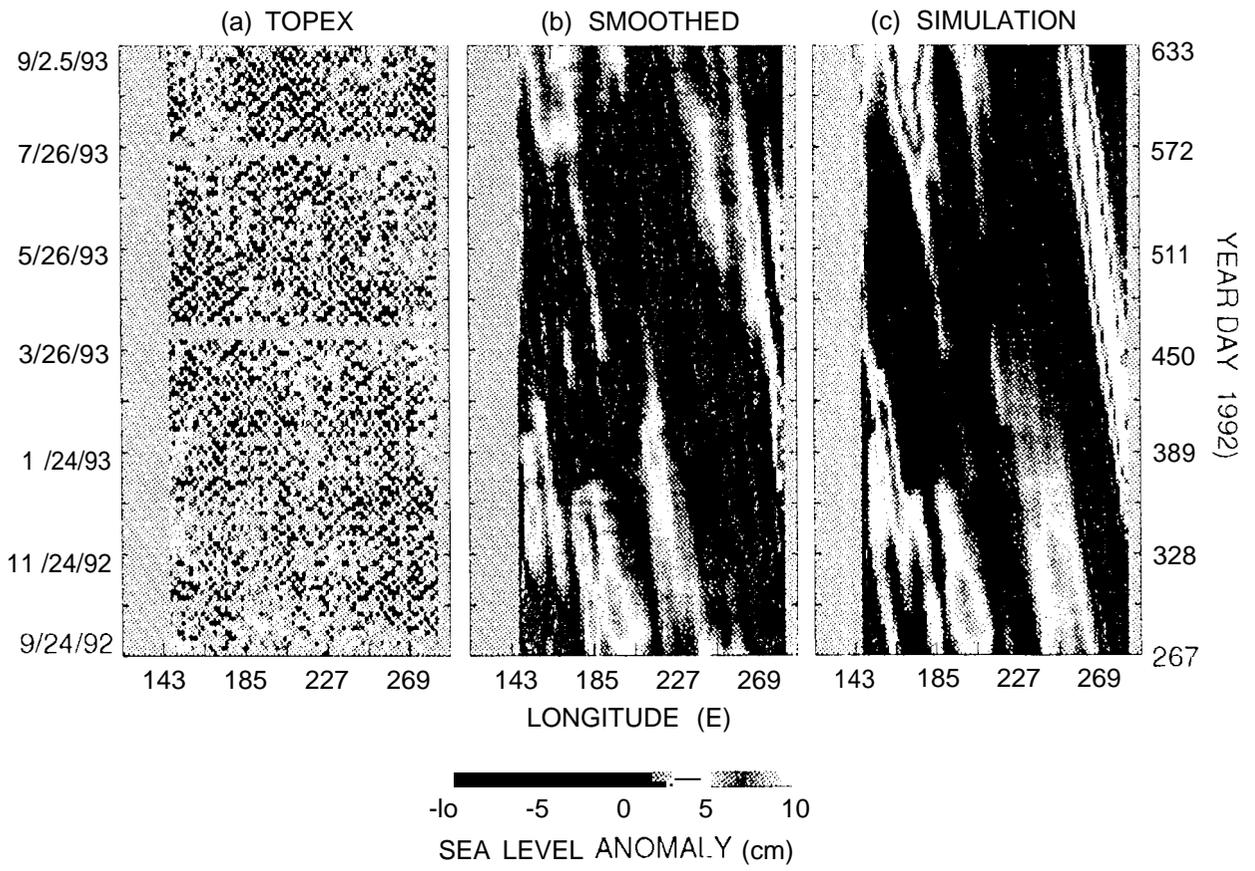
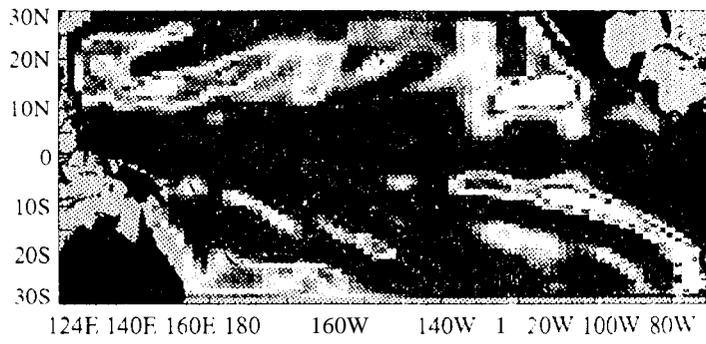


Fig. 4



(a) SIMULATION



(b) SMOOTHED

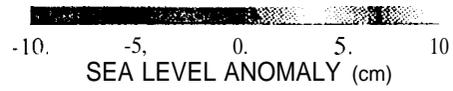
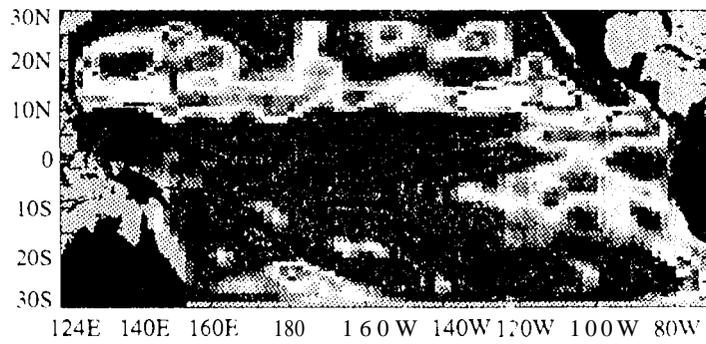


Fig. 6

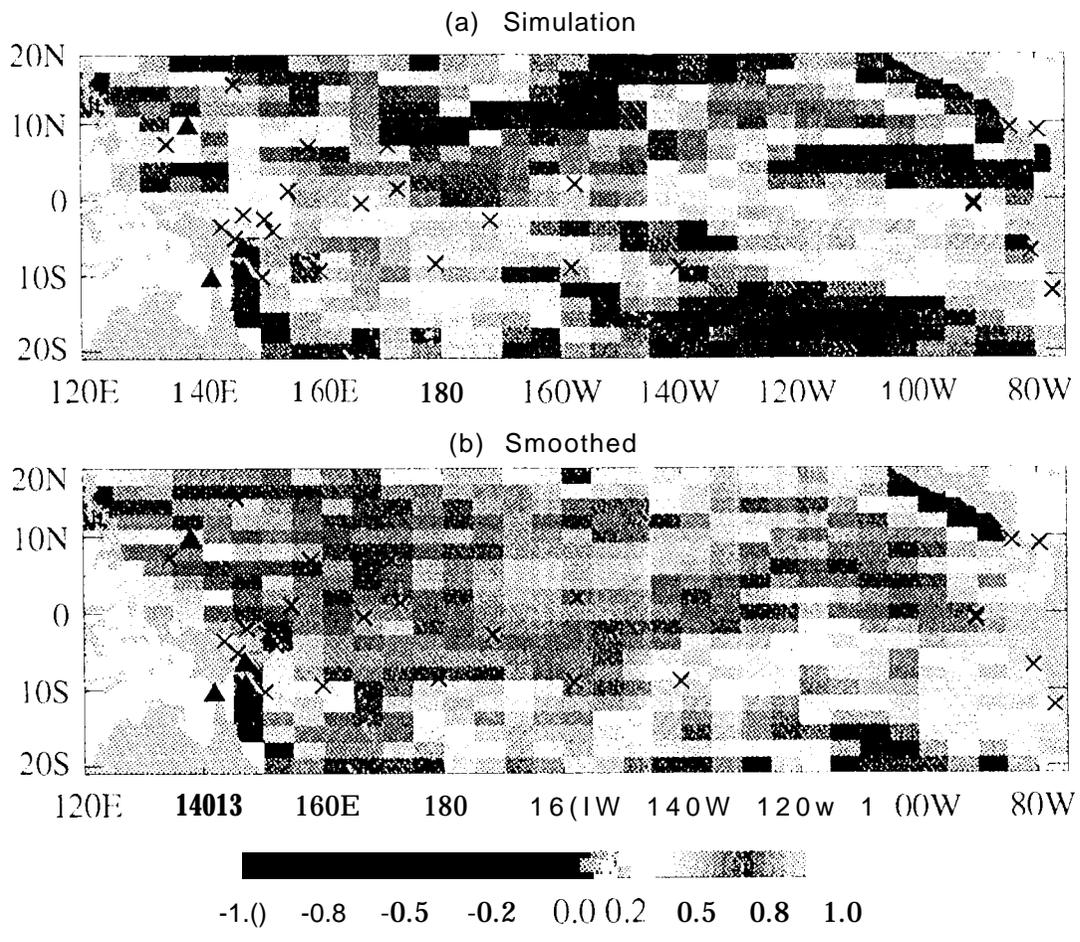


Fig. 7

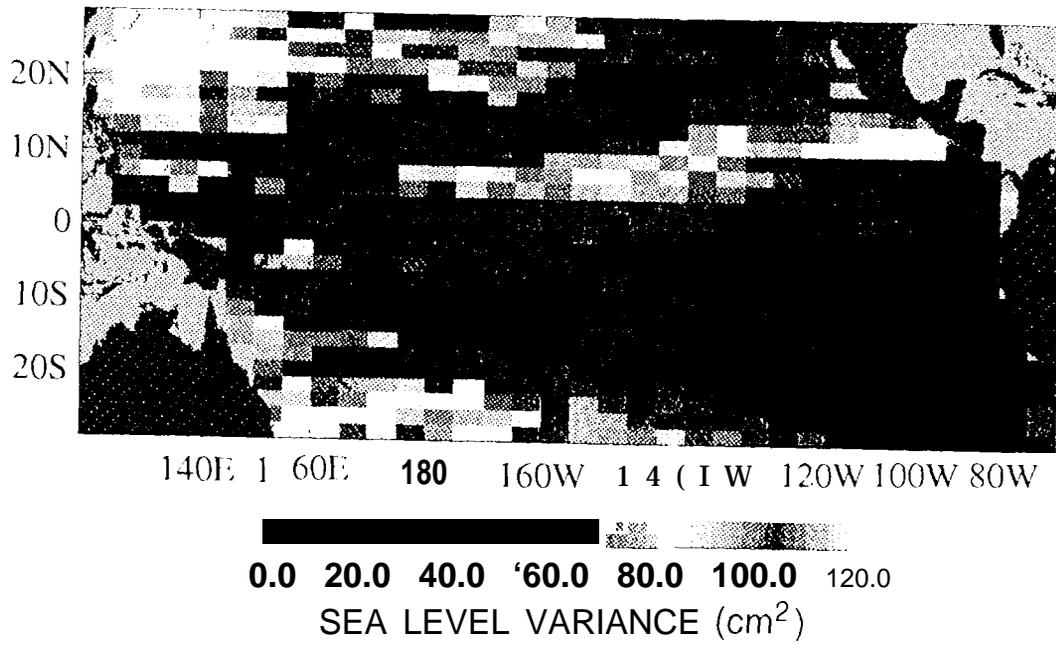


Fig. 8

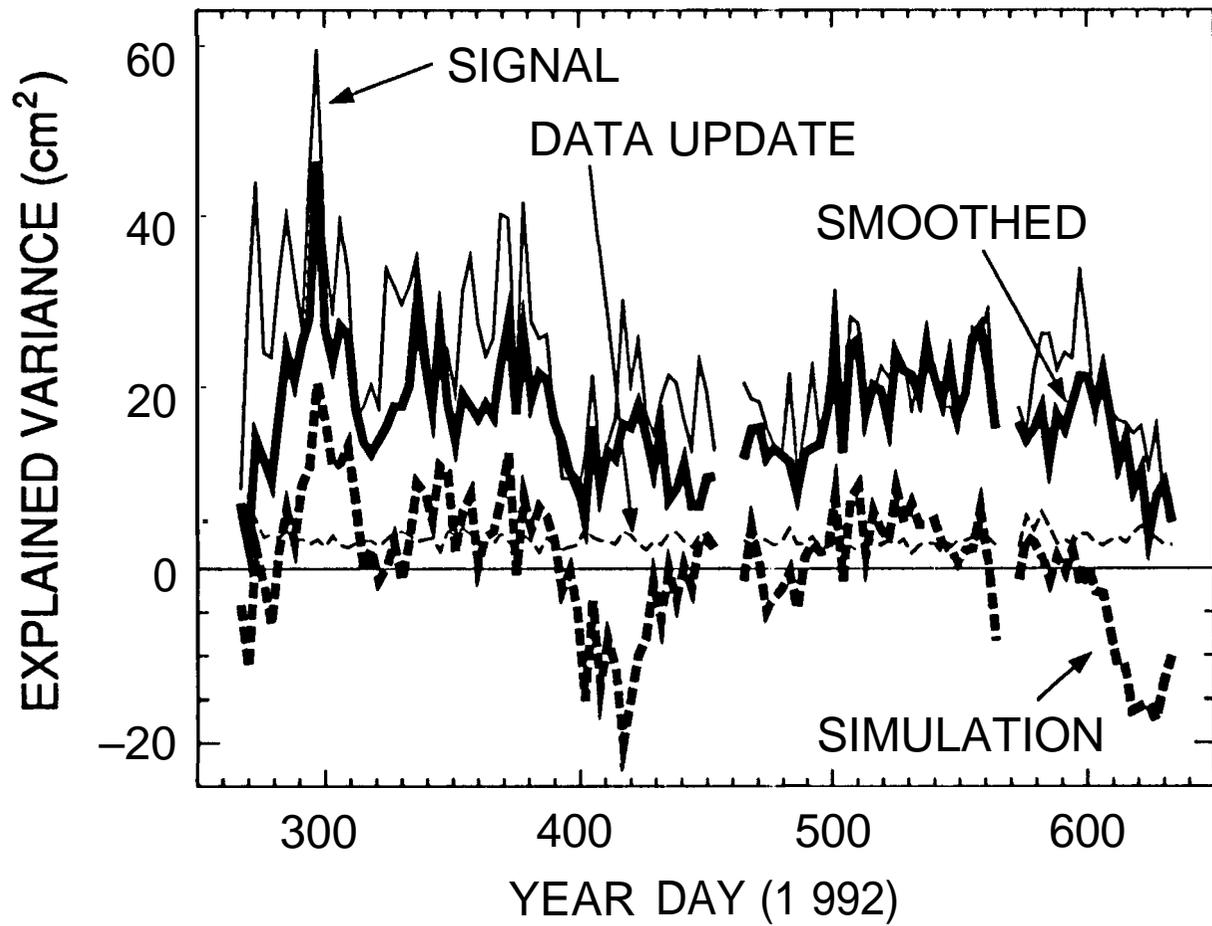
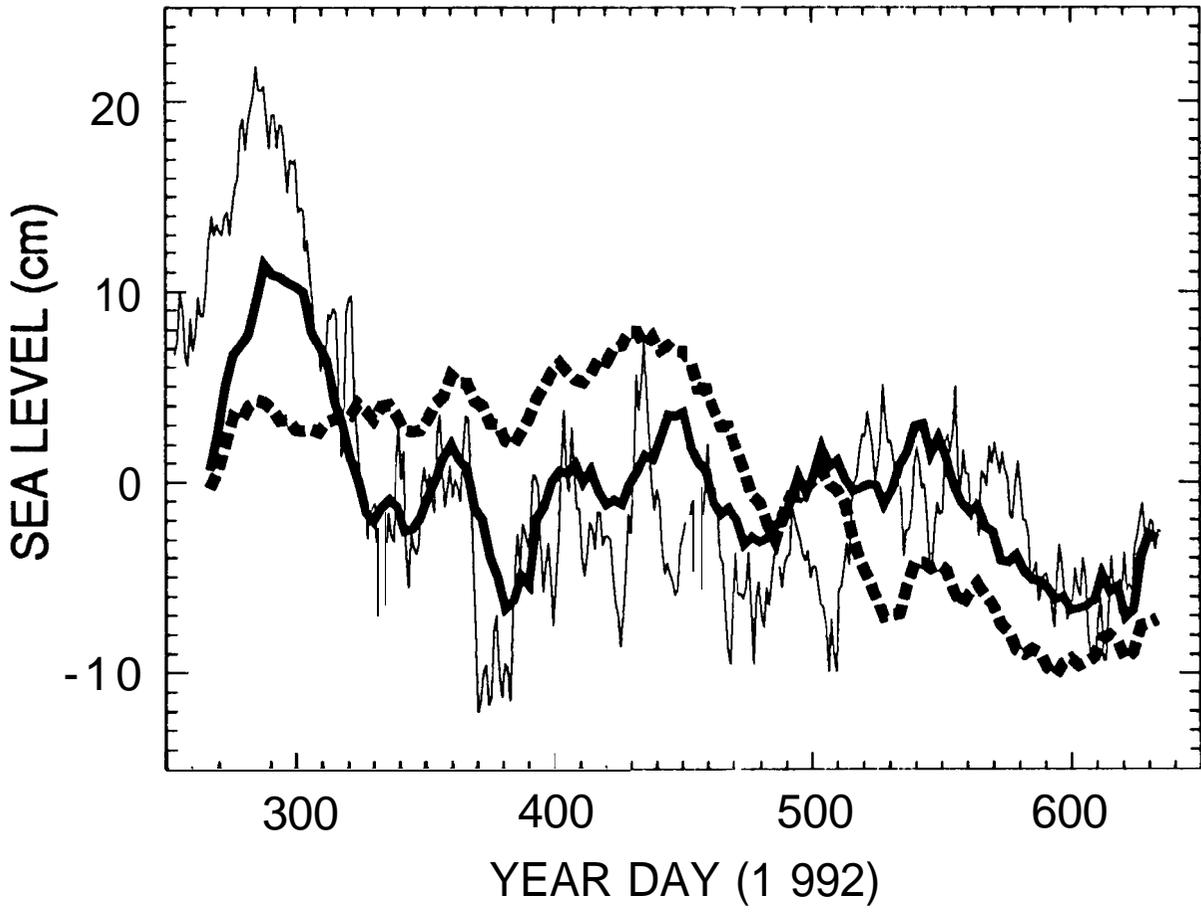


Fig. 9



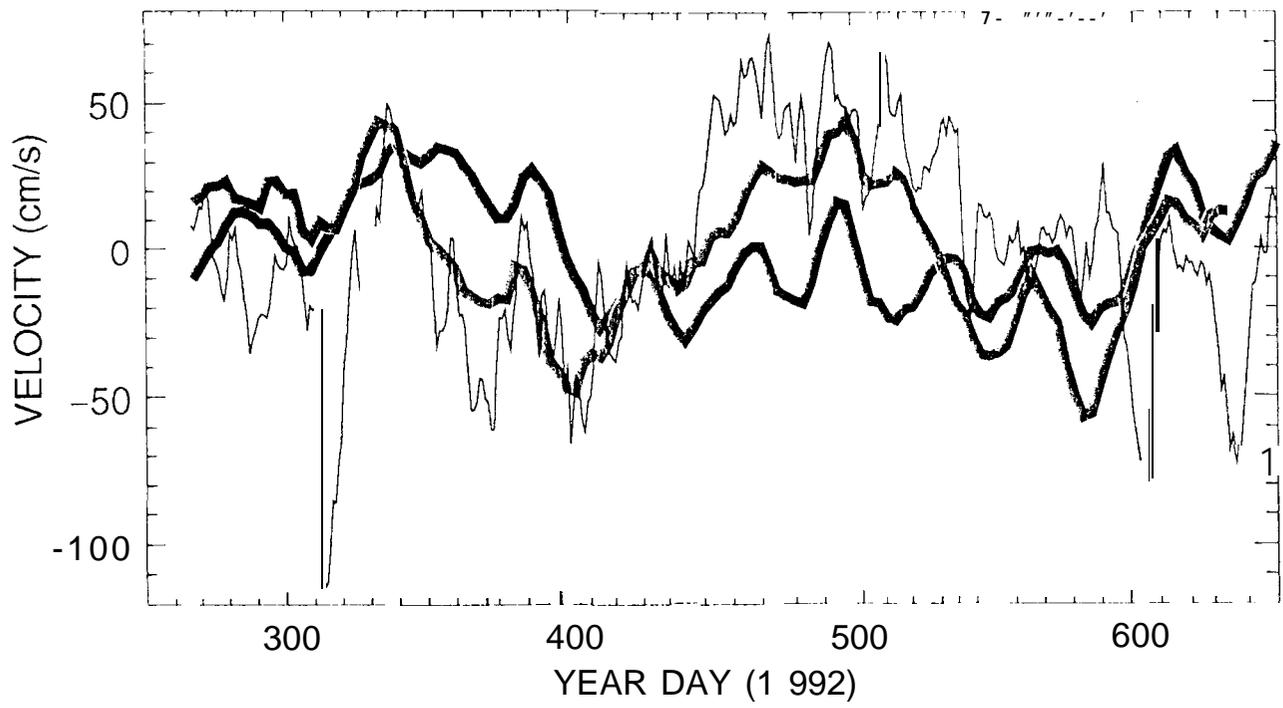


Fig. 11

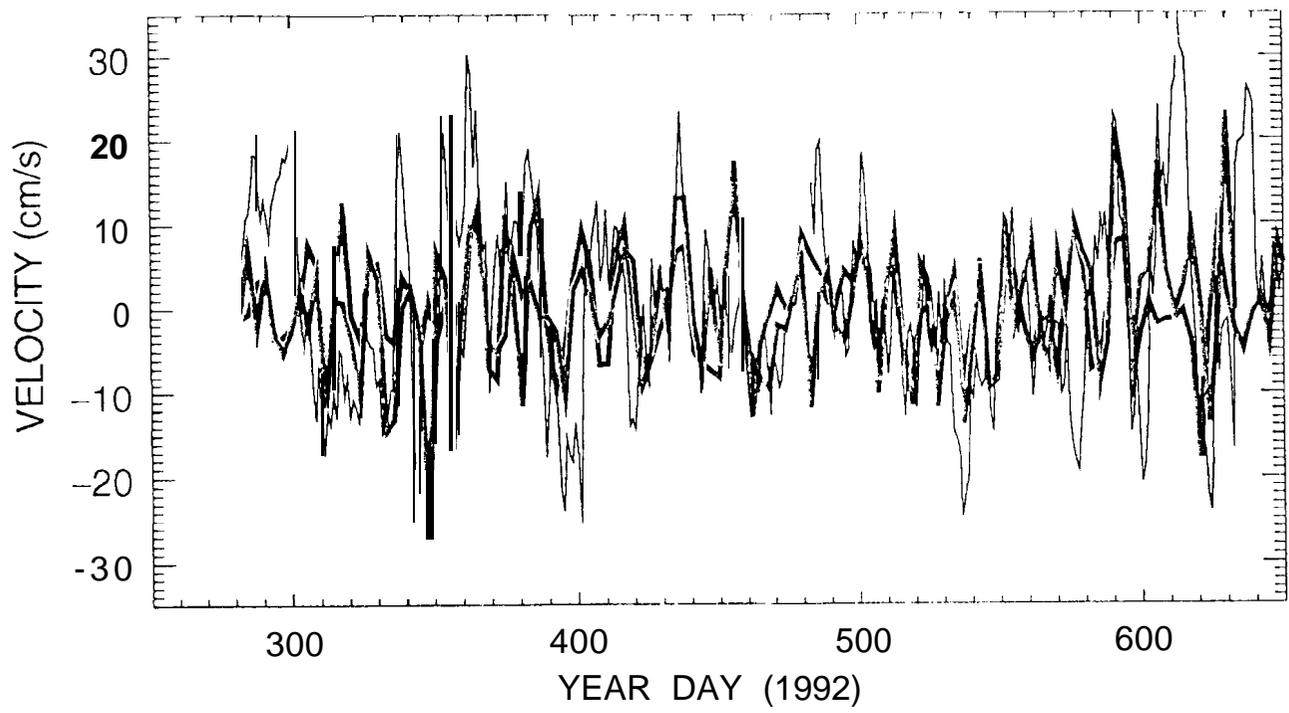


Fig.12

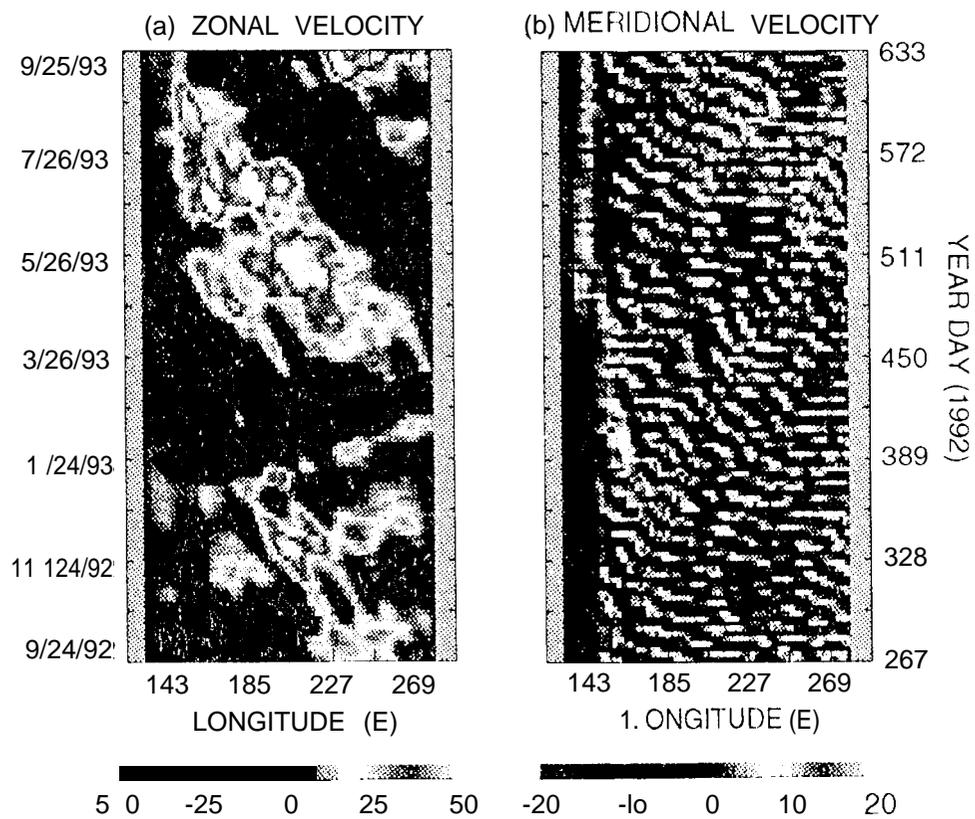


Fig. 13

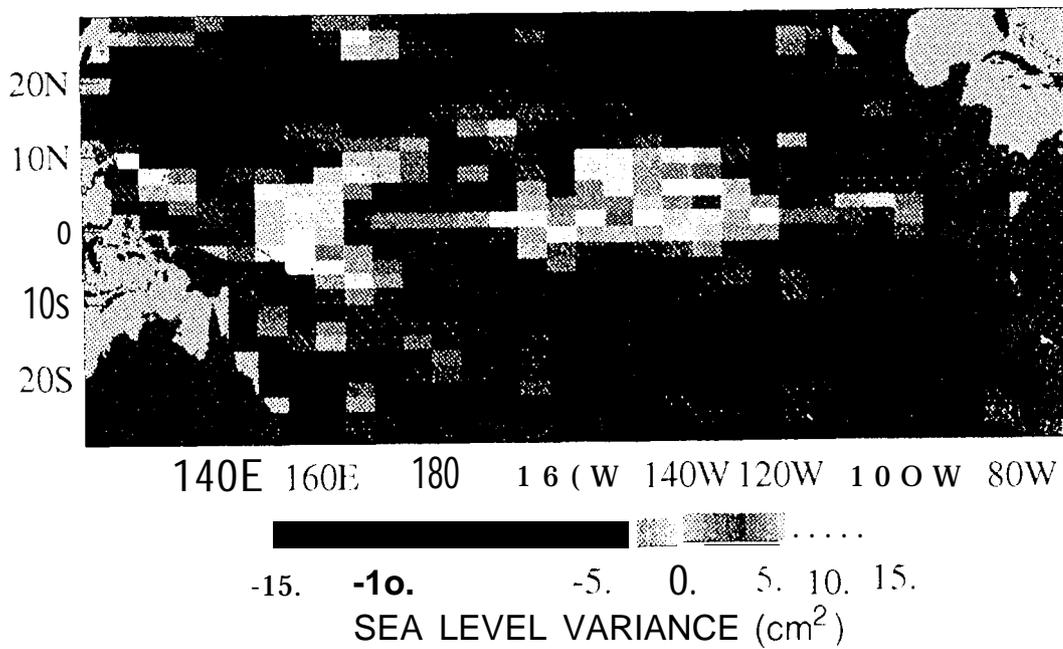


Fig.14