JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 101, NO. C6, PAGES 14,147–14,159, JUNE 15, 1996

Impact of altimeter, thermistor, and expendable bathythermograph data on retrospective analyses of the tropical Pacific Ocean

James A. Carton,[1] Benjamin S. Giese,[2] Xianhe Cao,[1] and Laury Miller[3]

Abstract. This study explores the relative impact of three major components of the tropical Pacific Ocean observing system on data assimilation analyses. The extensive observing components that are available during the 2-year period from October 1992 through September 1994 include sea level derived from TOPEX/POSEIDON altimetry, thermistor data from the Tropical Ocean/Global Atmosphere Tropical Atmosphere-Ocean (TOGA TAO) mooring array, and expendable bathythermographs (XBTs) from the Volunteer Observing Ship program. In the first part of this study, methods are introduced to assimilate these data into a primitive equation model using optimal interpolation. The resulting velocity and thermal analyses are then compared with a numerical simulation that has no data assimilation. The most innovative aspect of the assimilation procedure is the development of a statistical model that uses satellite altimetry to update the subsurface thermal field. To determine the impact of individual components of the observing system, three additional experiments are conducted. In each experiment, one component is withheld from the assimilation. The resulting analysis is then compared with the missing data set. Our results show (1) To resolve major features of the seasonal cycle, it is necessary to have either the altimeter or the XBT data. In the absence of XBTs the analysis appears to develop a significant temporal drift in its thermal structure. (2) Intraseasonal variability such as tropical instability waves is best resolved using altimetry because of that data set's superior spatial resolution. (3) Thermistor data from the TOGA TAO mooring array are helpful, but not as crucial, in resolving the seasonal cycle within ±8° latitude as the XBTs. The mooring array is too coarsely spaced to allow this analysis system to resolve tropical instability waves.

1. Introduction

This is a study of the usefulness of major components of the ocean observing system in constructing an analysis of temperature variability in the tropical Pacific Ocean. Although simulations of large-scale ocean circulation by models have improved steadily over the years [e.g., *Harrison et al.*, 1990], they still develop substantial errors in the distribution of heat storage. The same simulations contain even larger errors in the horizontal gradients of the thermal anomalies and thus in the velocity field. The need to reduce these errors has been a primary motivation for recent expansions of the observing system in the tropical Pacific. However, despite these expansions, in situ observations alone still have insufficient spatial and temporal coverage to allow estimation of the basin-scale ocean circulation. For subsurface temperature and currents, simply expanding the observing system is unlikely to solve this problem. The need for retrospective analyses of ocean circulation as well as the need to provide initial conditions to coupled atmosphere-ocean models can only be addressed by using available observations to constrain a numerical circulation model by way of data assimilation.

The most important elements of the current observing system consist of sea level provided by the TOPEX/POSEIDON (T/P) altimeter [*Cheney et al.*, 1994], subsurface temperature data provided mainly by the Volunteer Observing Ship (VOS) expendable bathythermograph (XBT), program and 66 moorings of the Tropical Ocean-Global Atmosphere program (TOGA) Tropical Atmosphere-Ocean (TAO) array [*McPhaden*, 1993]. One aspect of this paper is to introduce an algorithm for assimilation of sea level data along with the other data types. The approach we take is based on developing a statistical model that relates differences between observed and predicted sea level tendency to differences between observed and predicted thermocline depth tendency.

Much of the previous work on the important topic of sea level assimilation has been conducted with "synthetic" data produced by the same model that is later used for assimilation [*Hurlburt*, 1986; *Hurlburt et al.*, 1990; *De Mey and Robinson*, 1987; *Berry and Marshall*, 1989; *Holland and Malanotte-Rizzoli*, 1989; *Miller*, 1990; *Mellor and Ezer*, 1991]. Studies of this type are of limited applicability because of the different error structures of observed and synthetic data. Many studies of sea level assimilation assume midlatitude geostrophic dynamics [e.g., *White et al.*, 1990], and others use coarse vertical resolution. Both approximations simplify the problem tremendously because they tend to increase the correlation between surface pressure variations, surface velocity, and upper ocean density.

[1]Department of Meteorology, University of Maryland, College Park.
[2]Department of Oceanography, Texas A&M University, College Station.
[3]Geosciences Laboratory, NOAA, Silver Spring, Maryland.

Paper number 96JC00631.
0148-0227/96/96JC-00631$09.00

Unfortunately, these approximations are not appropriate in the tropics or coastal regions.

Instead, we adopt the scheme of *Carton and Hackert* [1990]. This scheme explicitly solves the optimum interpolation equations, minimizing the expected error of the solution under a variety of statistical assumptions *Derber and Rosati* [1989] solve the same equations implicitly), using a primitive equation ocean model based on the Geophysical Fluid Dynamics Laboratory (GFDL) modular ocean model (MOM) code. We extend this scheme by introducing a statistical model that relates vertical movements of sea level to changes of thermocline depth. This model turns out to be more stable in the tropics than statistical models relating sea level height to temperature fluctuations at fixed depth, as has been proposed by *Mellor and Ezer* [1991].

We conduct our study using data from the 2-year period from October 1992 to September 1994. Both were El Niño years, with unusually deep thermoclines and high sea levels in the east, although El Niño conditions were more pronounced in the spring of 1993 than 1994. The 2-year period was chosen because of the recent availability of high-quality wind products, as well as the vastly expanded ocean observing system.

The main focus of this paper is to determine the relative impact of the three components of the ocean observing system on the analysis of subsurface temperature, and to a lesser extent currents, during the 2-year period. Our experiment design is as follows. We carry out an analysis using all three data types. We then conduct a series of experiments. In the first we withhold altimeter sea level data. Comparison of this first experiment with the analysis reveals the contribution of altimeter data. In the second and third experiments we retain the altimeter sea level data but withhold the thermistor or XBT data. As before, we compare these experiments with the analysis to reveal the contribution of the thermistor or XBT data sets. Additionally, by comparing the three experiments with the data sets that are withheld, we are able to quantify the sea level and temperature error within the experimental analyses, thus providing a lower bound on the accuracy of the analysis.

2. Data Sets

Altimetry

Processing of 72 cycles of T/P altimeter data into sea level variations has been carried out by the National Oceanic and Atmospheric Administration (NOAA) Geosciences Laboratory, and the following is a description of their collinear difference analysis method. The data are derived from the T/P geophysical data records provided by Jet Propulsion Laboratory [*Callahan*, 1992]. The equatorial spacing of the orbit tracks, 2.83°, defines the minimum cross-track spatial scales that are resolved.

Standard geophysical corrections are applied (troposphere, ionosphere, inverse barometer, and sea state bias, described by *Callahan* [1992]). Ocean tidal signals were removed using the tide model of D. E. Cartwright and R. D. Ray [*Ray*, 1993] derived from earlier Geosat altimeter data, with an empirical correction to account for the residual tidal signal observed in the first year of TOPEX data [*Wagner et al.*, 1994]. We have found that unless the latter correction is made, sea level in the southeastern Pacific contains a distinct 4-cm, 60-day oscillation.

Sea level height tendencies are generated by averaging the sea level heights along 1° latitude segments along track and forming the collinear differences. This is done to avoid making any assumptions about the poorly known geoid. The sea level data are not gridded before assimilation. For each segment the time rate of change of sea level height between successive 9.92-day samples is directly assimilated into the model. Orbit precision for the T/P satellite is 5 cm, with most of the error at planetary wavelengths. We conducted several experiments that compare altimeter sea level with dynamic height estimates from the TOGA TAO array. These comparisons show that it is unnecessary to remove an along-track trend, as has been required with previous altimeter data sets. Thus no filtering has been done to correct for orbit error.

The accuracy of the data in the tropical Pacific has been examined recently in four studies. *Cheney et al.* [1994] find the root-mean-square difference between monthly averaged altimetric observations and monthly averaged tide gauge data to be 2–3 cm in the tropical Pacific. *Mitchum* [1994], *Busalacchi et al.* [1994], and *Giese et al.* [1994] find a somewhat higher error of 4–5 cm when 10-day averaged data are compared at a variety of tropical and subtropical stations. In the tropical Pacific, sea level variability during this 2-year period exceeds 10 cm in the northeast because of a combination of seasonal variations, tropical instability waves, and eddies and in the southwest because of the unusual conditions in spring 1994. In other parts of the tropical Pacific, sea level variability is typically 5–6 cm.

Temperature Observations

Subsurface temperature observations are provided by two major sources. The first is the extensive VOS network. These data also include other subsurface data transmitted over the Global Telecommunications System. We obtain the data from the Coupled Modelling Project of the NOAA National Centers for Environmental Prediction (NCEP), which has carried out preliminary error checking and correction. This error checking includes checks on static stability and subjective comparison with climatology [*Ji et al.*, 1995]. No attempt has been made to account for the drop rate problems of the T4 XBTs [*Levitus and Boyer*, 1994]. As a result of this error, many XBTs may have depth estimates that are too shallow by as much as 5%. The largest source of random error is the uncertainty in depth calibration corresponding to 1–2% of depth. This may cause as much as a 2°C error at thermocline depths.

The second major data set is provided by thermistor chains from 66 moorings of the TOGA TAO array. The data are obtained directly from the NOAA Pacific Marine Environmental Laboratory (PMEL) and are provided to us via their anonymous ftp distribution as daily averages. The thermistor data are checked and processed in the same way as the data obtained over the GTS. Figure 1 shows that the data coverage follows major shipping routes and mooring sites and is poorest in the eastern Pacific. The TOGA TAO moorings at 125°W are the main source of data for a region spanning 30° of longitude. Data were not included from two equatorial moorings at 110°W and 140°W because the data at these locations did not report in real time. After averaging of the data into superobservations based on 1° × 1° × 5 day bins, the 2-year data set contains 6909 thermistor and 14,346 XBT profiles.

In this study we use sea surface temperature (SST) as a proxy for mixed layer temperature. SST is also provided by the Coupled Modelling Project based on a combined analysis of satellite advanced very high resolution radiometer (AVHRR) radiances and buoy and ship reports [*Reynolds and Smith*, 1995]. Error estimates for these data range from 0.5° to 1°C.

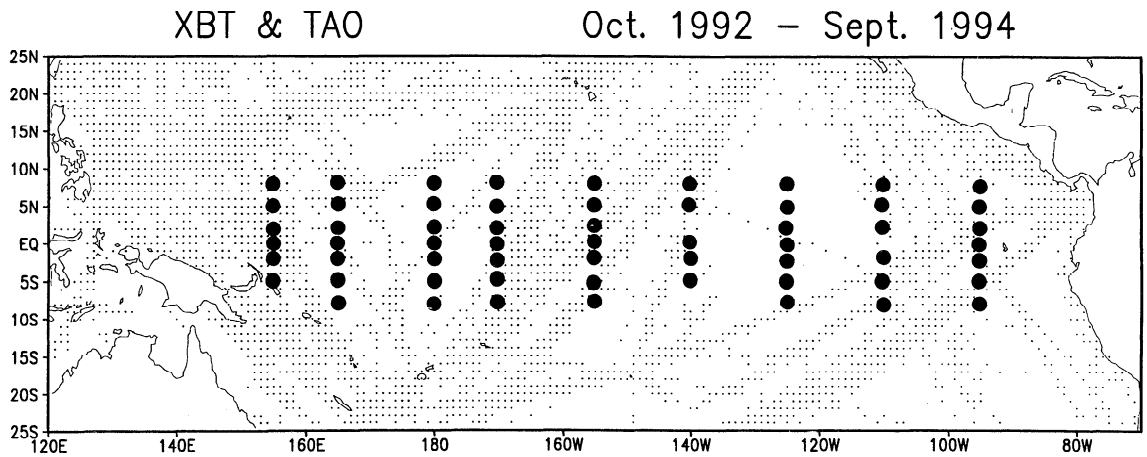

Figure 1. Available subsurface temperature observations between 25°S and 25°N. The data have been gridded on a 1° × 1° grid.

3. Assimilation Scheme

The statistical interpolation scheme described below has many features in common with assimilation schemes implemented in the 1980s by the European Center for Medium Range Weather Forecasts [*Shaw et al.*, 1987] and NCEP [*Dey and Morone*, 1985] and in the 1990s by the Data Assimilation Office of the Goddard Space Flight Center. The theoretical basis is discussed by *Daley* [1991]. A first guess, or background field, is produced at the analysis time by forcing a numerical model with observed surface forcing. Updating of the first guess is then carried out using available temperature and sea level height data.

The ocean model used in this study is based on the GFDL MOM software. The horizontal grid resolution is 0.5° × 2.5° in the tropics, expanding poleward of 5° to a uniform 1.7° × 2.5° grid at midlatitudes. The basin domain extends from 70°S to 60°N globally for a total of 146 × 91 horizontal grid points. The vertical resolution is 15 m in the upper 150 m, expanding toward the deep ocean for a total of 20 vertical levels. Vertical mixing is Richardson number dependent, while horizontal diffusion and viscosity are a constant 2×10^7 cm^2/s. The time step of the model is 2 hours. The winds for this study are provided by weekly 1000-mbar wind analyses of the NCEP converted to stress using a drag coefficient of 1.3×10^{-3}.

We consider a total of K observations of two types: the first $K - L$ observations directly measure temperature, T_O, while the last L observations infer the change in temperature between analysis times from observations of sea level, η_O. The observations are represented by a single observation vector $\mathbf{T}_O(\mathsf{r}_k)$. The corresponding background field for the first K-L observations is given by the model forecast temperature. The background field for the remaining L observations is given by the difference between the forecast temperature and the analysis temperature at the previous analysis time.

The expected error variance that is minimized has two sources: errors in the background estimate and errors in the observations. The background error covariance matrix B contains the effects of errors at the previous analysis time plus some growth term and a small error caused by interpolation from the analysis grid to the observation point. The observation error covariance matrix O contains a combination of three errors: instrument measurement error, inference error introduced if the measurement is not directly of temperature, and additional error because of aliasing of unresolved scales of motion. The observation, background, and analysis fields are written as vectors where the elements of the vectors are the variables at each grid point. Minimization of the expected analysis error leads to an algorithm for computing the analysis temperature T_A at horizontal location r_i and time t^n from observations at r_k (time will be t^n unless otherwise indicated).

$$\mathbf{T}_A(\mathsf{r}_i) = \mathbf{T}_B(\mathsf{r}_i) + \mathbf{B}(\mathsf{r}_i)^T(\mathsf{B} + \mathsf{O})^{-1}[\mathbf{T}_O(\mathsf{r}_k) - \mathbf{T}_B(\mathsf{r}_k)] \tag{1}$$

where the observation and background vectors are

$$\begin{aligned}\mathbf{T}_O^T(\mathsf{r}_k) = \{&T_O(\mathsf{r}_1), T_O(\mathsf{r}_2), \cdots, T_O(\mathsf{r}_{K\text{-}L}), H(\mathsf{r}_{K\text{-}L+1}) \\ &\cdot [\eta_O(\mathsf{r}_{K\text{-}L+1}) - \eta_O^{n-1}(\mathsf{r}_{K\text{-}L+1})], H(\mathsf{r}_{K\text{-}L+2}) \\ &\cdot [\eta_O(\mathsf{r}_{K\text{-}L+2}) - \eta_O^{n-1}(\mathsf{r}_{K\text{-}L+2})], \cdots, H(\mathsf{r}_k) \\ &\cdot [\eta_O(\mathsf{r}_K) - \eta_O^{n-1}(\mathsf{r}_K)]\}\end{aligned} \tag{2a}$$

$$\begin{aligned}\mathbf{T}_B^T(\mathsf{r}_k) = [&T_B(\mathsf{r}_1), T_B(\mathsf{r}_2), \cdots, T_B(\mathsf{r}_{K\text{-}L}), T_B(\mathsf{r}_{K\text{-}L+1}) \\ &- T_A^{n-1}(\mathsf{r}_{K\text{-}L+1}), T_B(\mathsf{r}_{K\text{-}L+2}) \\ &- T_A^{n-1}(\mathsf{r}_{K\text{-}L+2}), \cdots, T_B(\mathsf{r}_K) - T_A^{n-1}(\mathsf{r}_K)]\end{aligned} \tag{2b}$$

The background error vector $\mathbf{B}$ contains the covariances of analysis and observation errors

$$\mathbf{B}^T(\mathbf{r}_i) = [C_{1,i}, C_{2,i}, \cdots, C_{K\text{-}L,i}, C_{K\text{-}L+1,i}, \cdots, C_{K,i}] \tag{3}$$

where, for example, for $k \leq K\text{-}L$ $C_{k,i} = \langle [T_A(\mathsf{r}_k) - \tilde{T}(\mathsf{r}_k)] [T_O(\mathsf{r}_i) - \tilde{T}(\mathsf{r}_i)] \rangle$, and where $\tilde{T}$ is the true value of variable T.

The function $H(\mathsf{r})$ in (2a) represents the relationship between sea level tendency errors and subsurface temperature tendency errors. In the tropical upper ocean, changes in temperature generally result from vertical displacement of the thermocline. Figure 2 shows the relationship between sea level height and thermocline depth based on 2 years of simultaneous observations. The depth of the thermocline has a correlation with sea level of 0.6–0.8 (Figure 2a). The two variables are poorly correlated in the southeast and west because of the low amplitude of variability.

If the vertical displacement is small compared with the vertical scale of the temperature stratification, then to a good

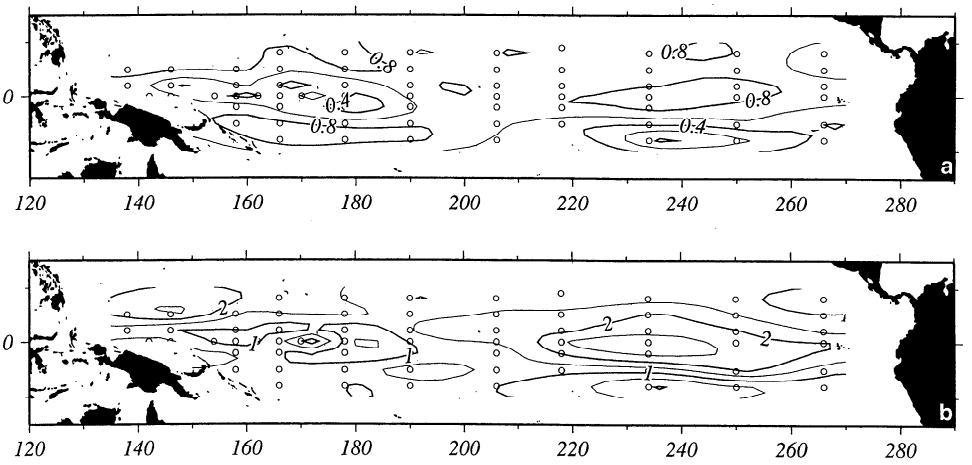

Figure 2. Relationship between altimeter sea level and thermocline depth as indicated by the depth of the 20°C isotherm depth at the TOGA TAO moorings, computed between October 1992 and September 1994. (a) Correlation of altimeter sea level height and thermocline depth. (b) Coefficient of proportionality, α, scaled by 100. A scaled value of $\alpha = 2$ means that a 1-cm rise in sea level corresponds to a 2-m deepening of the thermocline.

approximation, temperature at depth D is given by the vertical displacement δD multiplied by the vertical gradient plus terms of the order of the square of the displacement.

$$\delta T(z = D) = \overline{\frac{\partial T}{\partial z}}(z = D)\delta D + O\left[\frac{\partial^2 T}{\partial z^2}(\delta D)^2\right] \quad (4)$$

The overbar represents an average over the time interval when the displacement occurs. Introducing a linear relationship between small vertical displacements of the thermocline and sea level height gives

$$\delta T(x, y, D) \approx \alpha(x, y)\overline{\frac{\partial T}{\partial z}}(x, y, D)\delta\eta(x, y) \quad (5)$$

where α is an empirical coefficient of proportionality between sea level variations and thermocline depth variations. Multiplying both sides of (4) by $\delta\eta$ and taking the ensemble average gives a formula for α:

$$\alpha = \frac{\langle \delta\eta\ \delta D\rangle}{\langle \delta\eta\ \delta\eta\rangle} \quad (6)$$

Our estimates of α based on 2 years of simultaneous altimeter observations and the moored thermistor observations range from 100 to 250, with lower estimates in regions where the correlations are low (compare Figures 2a and 2b). These estimates are consistent with previous estimates by *Rebert et al.* [1985] at two tropical islands. For the experiments in this study, we choose a midpoint estimate of $\alpha = 200$ and apply it at all locations, meaning that everywhere the thermocline is assumed to deepen by 2 meters for each centimeter rise in sea level.

We assume that the observation error covariance matrix O is diagonal, meaning that the observational errors are uncorrelated. *Carton and Hackert* [1990] estimate the noise variance of XBT temperature errors in the tropical Atlantic to be 20% of the total error variance, so that the root-mean-square (rms) signal-to-noise ratio is $(0.8/0.2)^{1/2} = 2$. This is consistent with *White et al.* [1982], who previously estimated the signal-to-noise ratio to be 1.8. The most conservative estimate that we have found is that of *Meyers et al.* [1991], who derive a value of 0.75. For the XBT data we assume a signal-to-noise ratio of 1.2 in this study. Thermistor data are more accurate than XBT data because of their high temporal resolution and better calibration, and thus are given a higher signal-to-noise ratio of 2. Recent analyses by scientists at PMEL suggest that this may be an underestimate (W. Kessler et al., personal communication, 1995). Throughout the tropics the 10-day averaged altimeter sea level data appear to have an accuracy of 2–5 cm. If sea level variations were perfectly reflected in thermocline variations, then this would imply a 4- to 10-m error in thermocline depth. Because of the spatial variability and uncertainty of this relationship, the error roughly doubles to 8–20 m. We assign the altimeter-derived temperature tendency data a signal-to-noise ratio of 0.5 to reflect these substantial uncertainties.

The zonal and temporal scales of the temperature error have been estimated from a preliminary assimilation analysis using the TAO thermistor data as "truth" at two depths, 35 m and 100 m. We find a decrease in the horizontal scales with depth, similar to that found by *Carton and Hackert* [1990]. For simplicity we assume first-order exponential dependence for the covariance, with a zonal scale of 500 km and a latitudinal scale of 250 km. The temporal scale estimated from the same data is 14 days. In contrast, *Meyers et al.* [1991] derive broader zonal (15°) and meridional (2°–6°) scales of the background error in the tropics and a timescale of 1–3 months.

The numerical solution of (2) is carried out locally in a series of 5 × 5 grid point patches. In practice, it is desirable not to solve (1) every model time step. Updates are carried out once every 7 days, while the data time window extends to 1 month on either side of the update time (5 days for the extensive and repetitive altimeter data set). Because we are interested mainly in the tropical upper ocean, updating is confined to the upper 483 m in a range of latitudes between 25°S and 25°N.

4. Results

The analysis is carried out for the period from October 1992 to September 1994 using all available surface and subsurface thermal measurements, surface temperature, and altimeter sea level height information. Initial conditions are provided by an assimilation analysis using the XBT and thermistor data

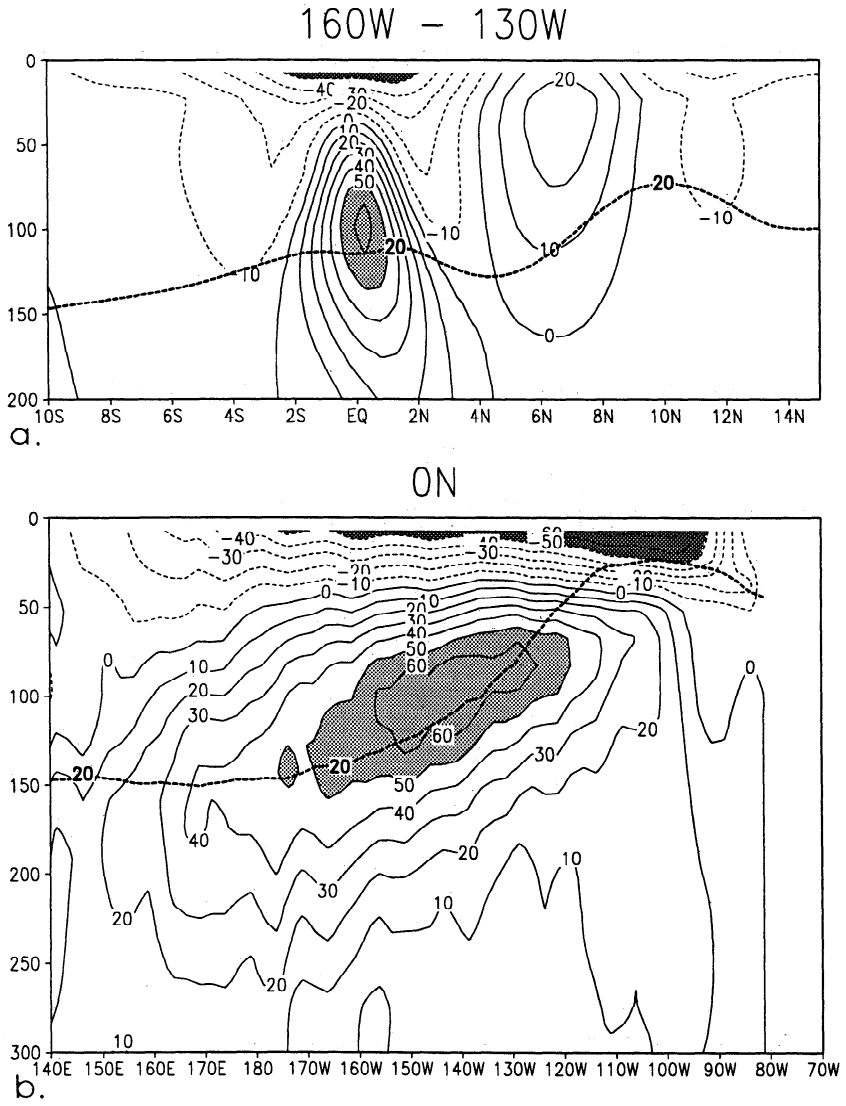

Figure 3. Time-averaged velocity from analysis using altimeter, TAO, and XBT data. (a) Zonal velocity at the longitudes of the high-velocity core of the Equatorial Undercurrent (160°–130°W). Annual mean depth of the thermocline is shown by the 20°C contour. (b) Zonal velocity with longitude along the equator. Annual mean depth of the thermocline is shown by the 20°C contour.

started from climatological conditions in January 1992. We begin with a comparison between this analysis and a simulation that includes no updating. Velocity measurements from the TOGA TAO current moorings represent an independent check of the analysis near the equator. We carry out a comparison of the simulation and the analysis with observations at the longitude of maximum Equatorial Undercurrent velocity (140°W). We then discuss three additional experiments. In experiment 1 the analysis is repeated excluding the altimeter sea level data, allowing us to examine the impact of this data set on the analysis. In experiments 2 and 3 the analysis is repeated including sea level data but excluding either the thermistor or the XBT data.

The major features of the zonal velocity around 140°W are the eastward North Equatorial Countercurrent and Equatorial Undercurrent (Figure 3). The former lies between 5°N and 10°N and is confined above the thermocline. The analysis annual average velocity during this 2-year period has a core maximum of 20 cm/s. The latter lies within 2° latitude of the equator and has much higher core velocities of 60–80 cm/s in our analysis, with maximum velocities within the thermocline at 120 m depth. The near-surface currents within a few degrees of the equator flow westward as part of the South Equatorial Current.

A current meter has been deployed at 140°W as part of the TOGA TAO array. We obtained the data from PMEL for a 1-year period of overlap beginning in April 1993 at six depths between the surface and 200 m. We have smoothed the data using a 7-day boxcar average and have overplotted the observations and analysis in Figure 4. Comparison of the two shows that the analysis Equatorial Undercurrent is weaker than that observed by 15 cm/s and that the South Equatorial Current is stronger than that observed by the same amount. The vertical shear above the undercurrent is in agreement with the observations, but the shear below the undercurrent is too weak. When no updating is carried out, the undercurrent speed drops by an additional 15 cm/s. In either case, the South Equatorial Current is too strong.

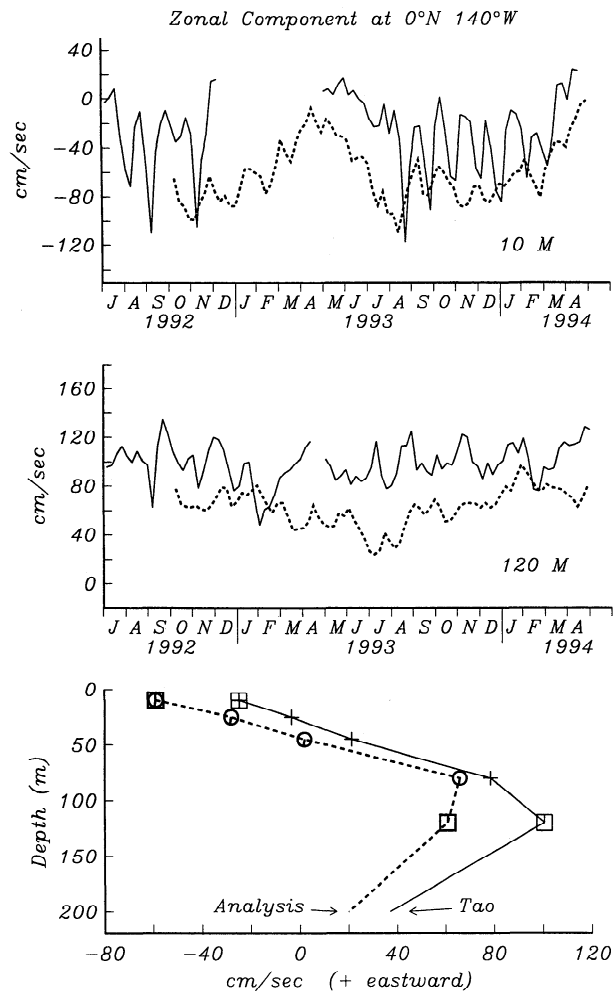

Figure 4. Zonal velocity at 140°W. (top) Observed (solid lines) and analysis (dashed lines) velocity with time at 10 m. (middle) Observed and analysis velocity at the depth of the mean Equatorial Undercurrent. (bottom) One-year averaged eastward current with depth. Boxes are used to mark the depths of moorings whose time series are given in the upper panels. Mean is computed over the period May 1993 to April 1994.

An examination by *Wacongne* [1990] of the balance of forces governing the speed of the Equatorial Undercurrent shows that at 140°W the primary term driving the undercurrent is the zonal pressure gradient, with viscous forces acting to decelerate the flow. Viscous forces should be similar in magnitude whether or not updating occurs. Thus it seems likely the intensification of the undercurrent resulting from the updating procedure is caused by a strengthening of the zonal pressure gradient. Updating does not affect the South Equatorial Current, whose high velocities are thus not the result of inaccurate subsurface thermal structure.

For the remainder of this section we focus our attention on the heat content variability of the three "impact" experiments, expressed in terms of integrated temperature between 0 and 483 m (483 m is the depth of the model level nearest the bottom of the XBT casts). The gross variability of the analysis heat content is shown in Figure 5a, and the relative effects of separately omitting the altimeter, TAO, and XBT data from the analysis are presented as rms differences, experiment minus analysis, in Figures 5b, 5c, and 5d. The maximum rms of the analysis occurs in two zonal bands, one at 5°N between 180°W and 120°W, and the other at 11°N between 160°W and the coast of Central America. Most of the variability at 5°N is associated with 20- to 30-day-period tropical instability waves. These waves were first observed by *Legeckis* [1977] as meridional perturbations of the northern edge of the equatorial cold tongue. They exist on the boundary between the South Equatorial Current and the North Equatorial Countercurrent in all seasons except spring. The variability in the 11°N band has only recently been identified by *Giese et al.* [1994] as a combination of wind-driven annual heat storage change and westward propagating eddies generated in the Central American gulfs of Papagayo and Tehuantepec. Annual heat content is high during August through January, when the Intertropical Convergence Zone is displaced northward, and low during other seasons.

Except for the extreme northern and southern margins of the study area, the experiment minus analysis heat content variability is typically 100°C m when the altimeter or XBT data are omitted (experiments 1 and 3) and 50°–100°C m when the TAO data are omitted (experiment 2). Thus the TAO data appear to contribute somewhat less to the analysis than do either the altimeter or XBT data. The omission of altimetry (Figure 5b) causes relatively small changes (<100°C m) along the equator and in the southeast, both regions of small variability in the analysis (Figure 5a), and also in the northwest, where the correlation between sea level and thermocline depth was shown in section 3 to be poor (Figure 2a). The omission of TAO data (Figure 5c) produces small changes (≈50°C m) throughout the equatorial region, except between 130°W and 100°W along the southern boundary, where the impact is somewhat greater (between 50 and 100°C m), because of the absence of XBTs (see Figure 1). Finally, the omission of XBT data causes moderate changes throughout the zonal band covered by the TAO array, except directly on the equator, where the effect is greater west of 130°W. The relative impacts of the different data sets are examined in greater detail below.

Altimeter Impact

The impact of altimetry on the annual cycle is best seen by looking at seasonal changes in heat content averaged between 160°W and 110°W, where the seasonal signal is primarily zonal and its amplitude is relatively large. Note that by averaging over 50° of longitude, we effectively suppress the energetic tropical instability waves. Figure 6a shows the deviations of the analysis with respect to the mean heat content as a function of time and latitude, while Figures 6b, 6c, and 6d present the different impact experiments, omitting altimeter, TAO, or XBT data and subtracting the result from the analysis. If multiplied by −1, these plots can be interpreted as the contribution of the data set that is omitted to the analysis.

The analysis shows a strong annual cycle north of 5°N. A maximum of 300°C m occurs as early as July at 10°N and as late as October at 5°N. South of 5°N, interannual signals dominate. Heat content is high during January–July 1993 compared with to January–July 1994 as a result of anomalous warming associated with the 1993 El Niño event. As Figure 6b shows, the impact of the altimeter data is greatest north of 5°N. There the amplitude of the annual cycle increases by as much as a factor of 2 when altimetry is included. In contrast, Figures 6c and 6d show that neither the TAO nor XBT data contribute much to the annual signal in this latitude range. South of 5°N, the altimetry still has a sizable impact. Along the equator the

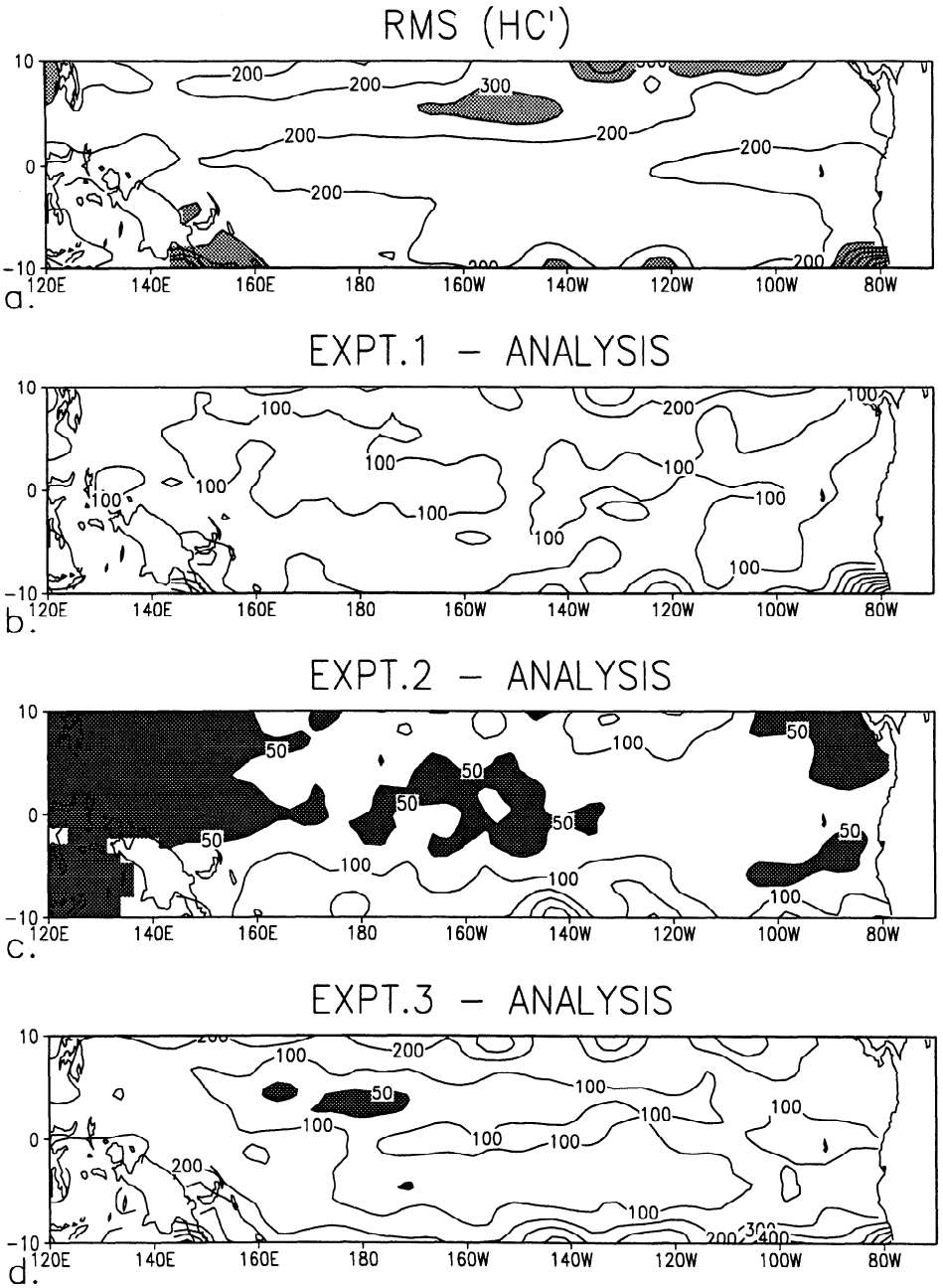

Figure 5. Root-mean-square heat content (0/483 m) variability. (a) Analysis rms heat content variability. Areas exceeding 300°C m are shaded. (b) Plot of rms difference between experiment 1 (omitting altimeter data) and analysis heat content. Areas less than 50°C m are shaded heavily. (c) Plot of rms difference between experiment 2 (omitting TAO thermistor data) and analysis heat content. Areas less than 50°C m are shaded heavily. (d) Plot of rms difference between experiment 3 (omitting XBT data) and analysis heat content. Areas less than 50°C m are shaded heavily.

altimeter acts to deepen the thermocline in fall and winter of 1993 and to raise it again in fall of 1994. In contrast, the TAO and XBT contributions act to force the analysis in the opposite direction. As Figure 2b shows, the coefficient (α) relating sea level change to thermocline change varies considerably along the equator. Using a constant coefficient causes altimeter data to overdrive the thermocline in some areas and underdrive the thermocline in other areas. One possible explanation for the fact that the altimeter data are pulling the analysis in the opposite direction of the TAO and XBT data is that our estimate of α is too large in this eastern region.

To highlight the impact of altimetry on short-wavelength/short-timescale phenomena inside the domain of the TAO array, Figure 7 presents a set of time-longitude plots of heat content variability between 110°W and 160°W along 5°N. The full analysis (Figure 7a) is dominated by tropical instability waves in fall and winter. When altimeter data are withheld (Figure 7b), the instability waves are weak. However, when either TAO or XBT data are withheld (Figures 7c and 7d) the instability waves are strong. From this we conclude that the altimeter data do a better job of resolving instability waves at 5°N than do either the TAO or XBT data. The amplitude of

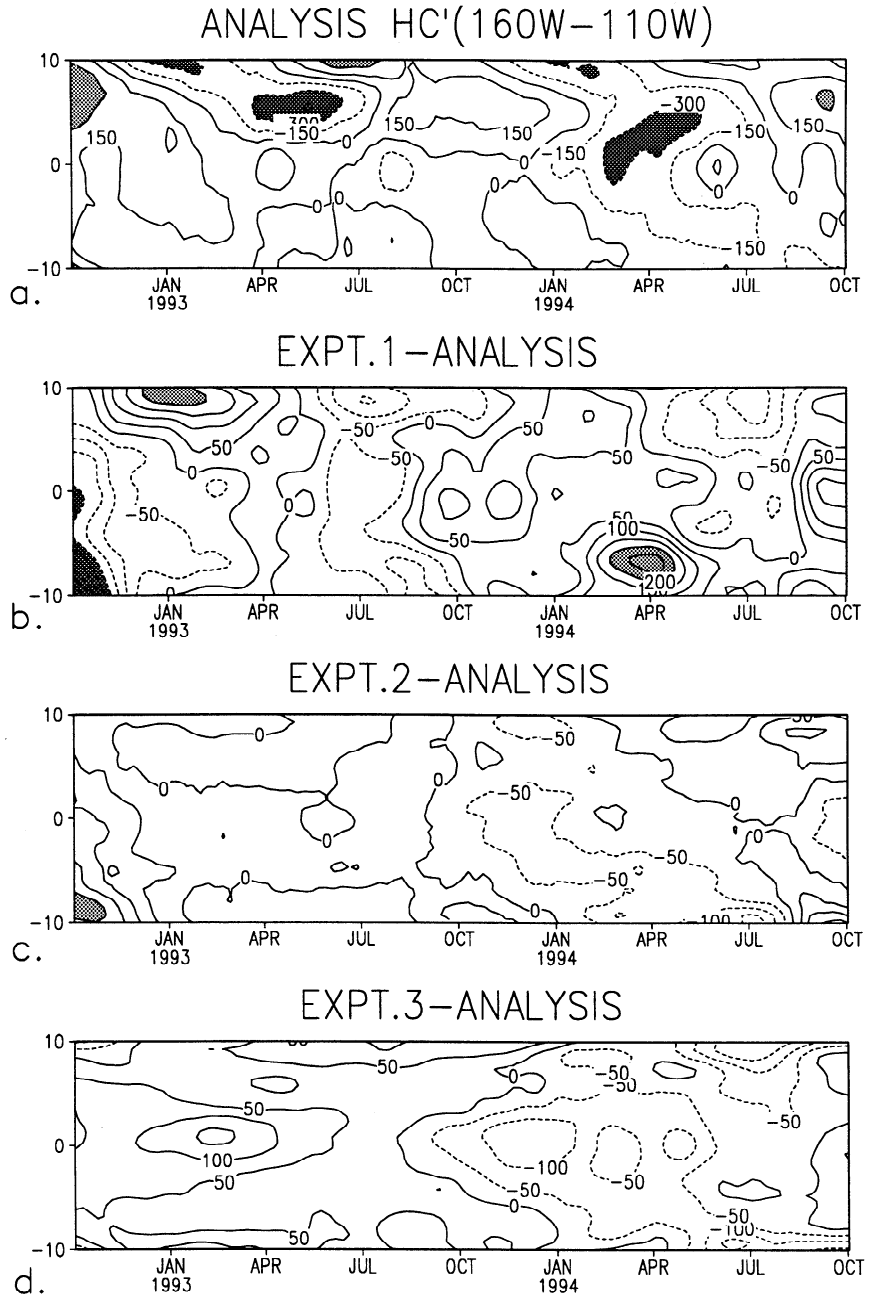

Figure 6. Analyses of heat content (0/483 m) in degree Celsius meters) averaged in longitude between 160°W and 110°W, with latitude and time. (a) Analysis heat content anomaly. Values exceeding ±300°C m are shaded. (b) Difference between experiment 1 and analysis heat content anomaly. Values exceeding ±200°C m are shaded. (c) Difference between experiment 2 and analysis heat content anomaly. (d) Difference between experiment 3 and analysis heat content anomaly.

the analyzed waves is also a function of the updating interval (Figure 8). When the interval is reduced from 7 to 3.5 days, the amplitude increases. However, the change is not enough to resolve the waves properly in the absence of altimeter data.

The impact of altimetry on short-wavelength/short-timescale phenomena outside the domain of the TAO array is evident in Figure 9. This shows heat content time series for the analysis and the three impact experiments at 10°N, 110°W, a location strongly influenced by the passage of eddies emanating from the coast of South America. When altimeter data are withheld, the annual cycle is reproduced reasonably well, but not the 1- to 2-month eddy fluctuations of the eddies. In experiment 3, when altimetry is retained but XBT data are withheld, the short-timescale variations are resolved, but the analysis gradually cools, indicating that altimeter data are unable to maintain the mean thermal structure.

TAO Impact

On the basis of the variability maps presented in Figure 5, it appears that the TAO data have a smaller impact on the analysis in the central and western equatorial Pacific than do either the altimeter or XBT data. Throughout these two areas,

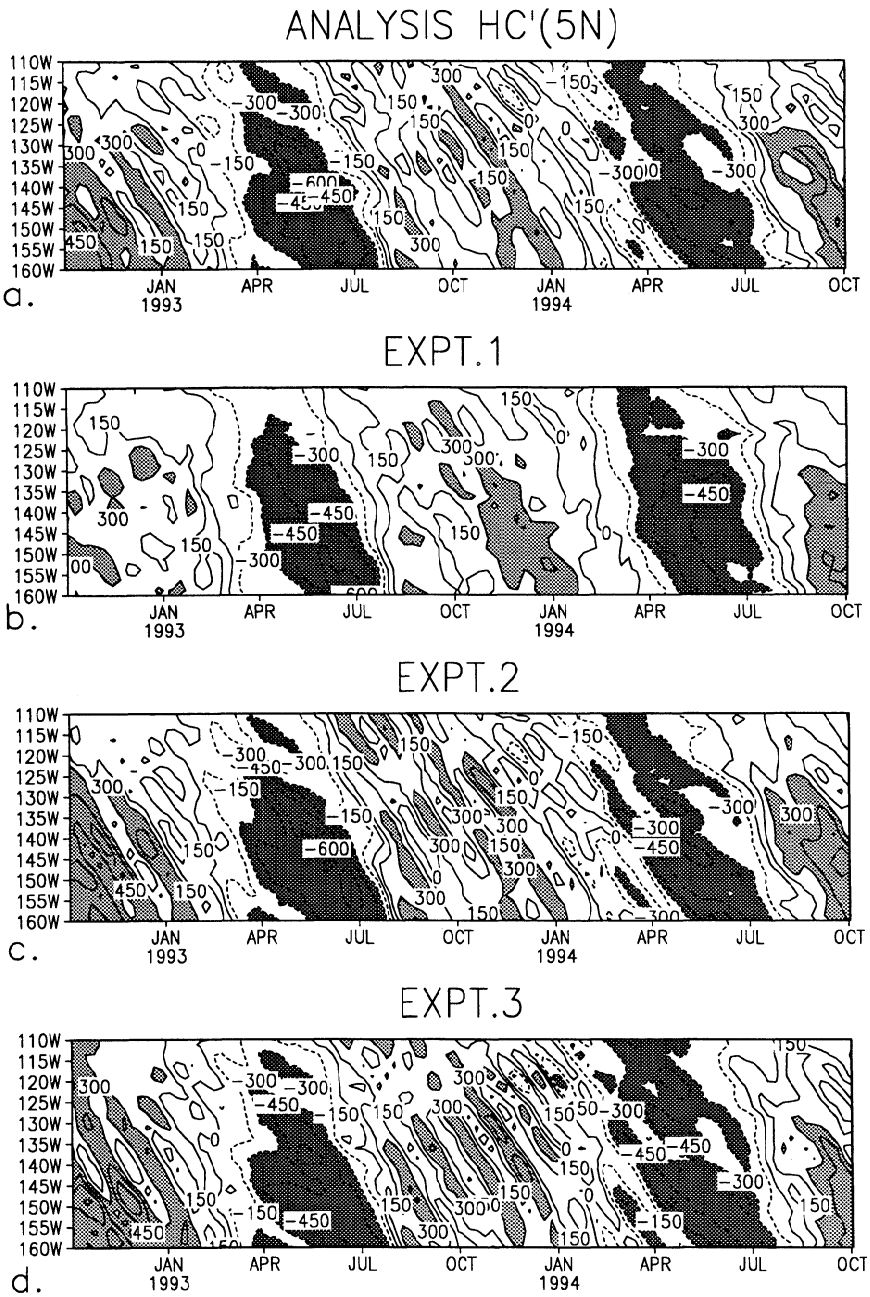

Figure 7. Analyses of heat content at 5°N (in degree Celsius meters) with longitude and time. Values exceeding ±300°C m are shaded. (a) Analysis heat content anomaly. (b) Experiment 1 heat content anomaly. (c) Experiment 2 heat content anomaly. (d) Experiment 3 heat content anomaly.

TAO contributes less than half of the rms variability. A comparison of Figure 5c with Figure 1 reveals that along the equator, areas of minimum TAO impact west of 160°E and between 180°E and 140°W coincide with areas of heavy VOS traffic, while areas of greater TAO impact between 170°E and 180°E, and between 130°W and 100°W coincide with areas of light traffic. Thus the spatial distribution of the XBT data influences the relative impact of the TAO data on the analysis.

Thermistor chain data have a great advantage over altimetry in that they provide direct information about the vertical structure of temperature. Figures 10a and 10b show the effect of the TAO data on the vertical structure of temperature as a function of longitude along the equator and latitude at 140°W. The largest errors occur in the eastern basin within a few degrees of the equator, with a second maximum at 10°S. The increase in variability at 10°S results from the presence of a particularly strong mesoscale feature that originates near the South American coast. A 0.5°–1.0°C temperature difference at thermocline depths corresponds to a difference of 4–8 m in the depth of the thermocline. This result provides an upper bound on temperature error in the analysis.

The impact of thermistor data on the time evolution of temperature is shown at one mooring located midbasin, 0°N, 155°W (Figure 10c). When thermistor data are withheld (experiment 2), the analysis reproduces the observed temperature to within 0.5°C. The major difference between the analysis

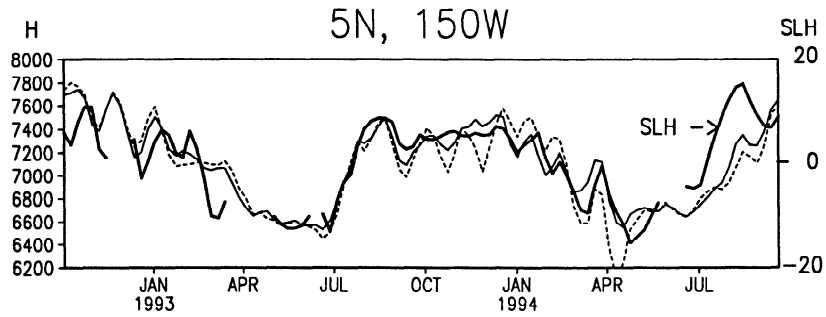

Figure 8. Time series at 5°N, 150°W of altimeter sea level (bold line, in centimeters) and heat content from analysis (in degree Celsius meters) with 7-day updating (solid line) and 3.5-day updating (dashed line).

without thermistor data and the observed temperature is that the analysis is persistently too cold in the upper 50 m. This allows the analysis surface temperature to drop below 25°C during the upwelling season of January–March. Farther east, the experiment 2 thermocline is noticeably too shallow by 5–10 m.

XBT Impact

On the basis of the variability maps presented in Figure 5, the impact of omitting XBT data (experiment 3) is similar in magnitude to that of withholding altimeter data. However, the impact is most pronounced at longer timescales and larger space scales than those associated with the altimeter data. For example, in the absence of XBTs, the eastern Pacific undergoes a warming in 1993 followed by a gradual cooling at a rate of 50°C m/yr in late 1993 and 1994 (Figures 6c and 9). This is equivalent to a 0.3°C/yr change over the upper 125 m. The cause of this drift is a combination of model thermal drift and accumulated updating errors associated with the use of sea level tendency, rather than sea level. As longer altimeter records become available, it should be possible to determine a meaningful reference level for sea level and thus eliminate one source of drift.

On intraseasonal timescales, the absence of XBTs produces enhanced tropical instability wave amplitudes of 450°C m. The phase of the waves in this experiment is similar to the phase of the waves in the analysis (Figure 7). This result, together with a similar result for experiment 2, indicates that sea level information is largely responsible for the proper resolution of tropical instability waves in the analysis. Adding XBT data reduces the amplitude of the waves, presumably because the update algorithm spreads the sparsely sampled XBT information over too broad a space/time domain to resolve the waves. At 10°N, experiment 3 is able to resolve the eddies with an amplitude similar to that in the analysis, suggesting that the XBTs do not contribute to the determination of the mesoscale variability in this location.

5. Discussion and Conclusions

The tropical Pacific observing system has three major components: subsurface temperature from the VOS XBT program, thermistor data from 66 moorings of the TOGA TAO Atlas mooring array, and sea level height estimates from TOPEX/POSEIDON altimetry. In this study we examine the impact of the three components on analysis of temperature and circulation of the tropical Pacific during a 2-year period. Although not a prognostic variable of the model, sea level influences the circulation through its relationship to thermocline fluctuations. We quantify this relationship by a comparison of simultaneous sea level and temperature time series at the Atlas mooring sites.

We focus our attention on the ability of the analysis to represent two major timescales of variability in the tropical Pacific: the seasonal cycle and intraseasonal fluctuations. When altimetry is withheld, major aspects of the seasonal cycle are still resolved, but the meridional gradients of heat storage are weakened. The root-mean-square difference between this experimental analysis dynamic height and observed sea level

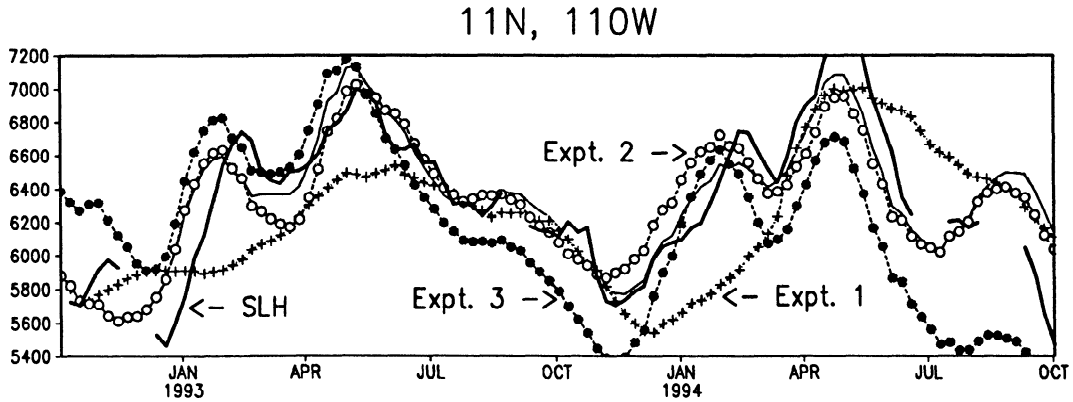

Figure 9. Time series of heat content (0/483 m in degree Celsius meters) at 10°N, 110°W: analysis (bold line), experiment 1 (crosses), experiment 2 (open circles), experiment 3 (solid circles), and altimeter sea level (solid line).

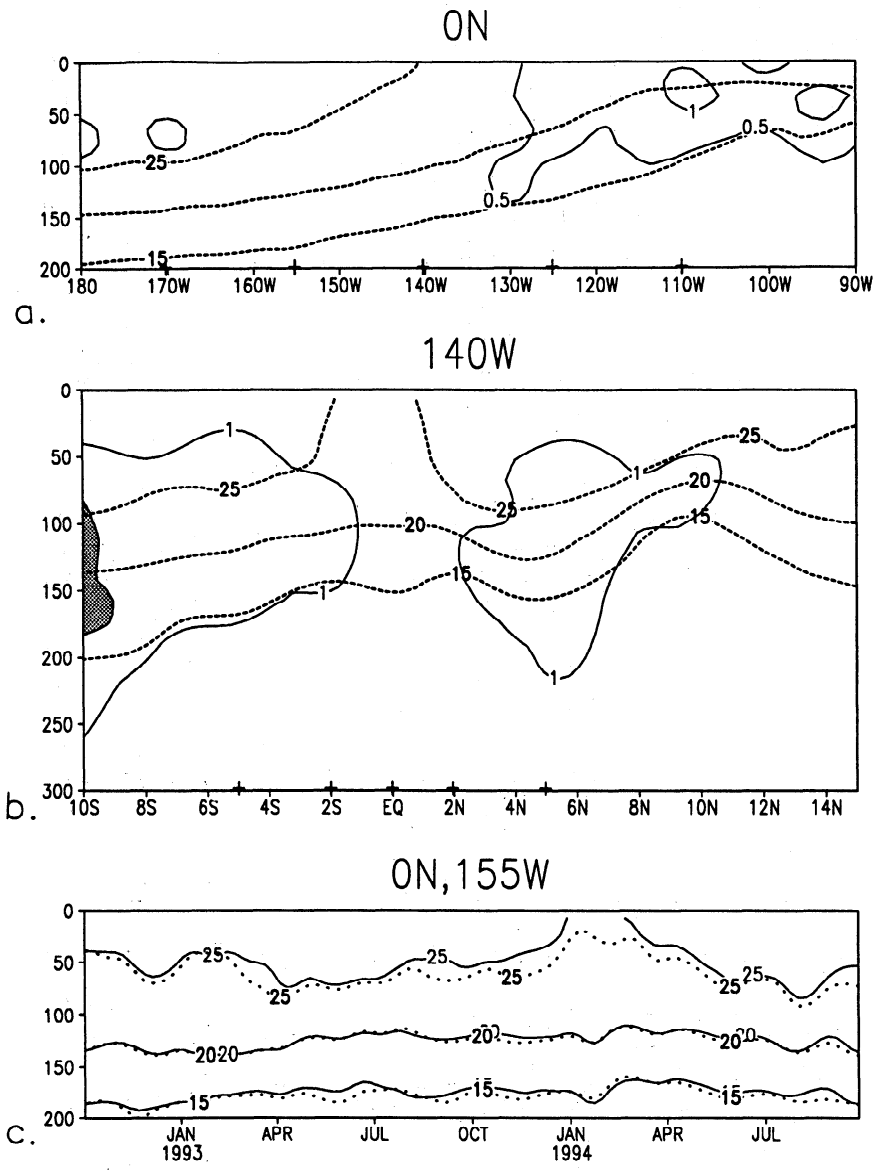

Figure 10. Relationship between observed TOGA TAO temperature and analysis temperature when observed TOGA TAO temperature is excluded from the analysis. (a) Root-mean-square temperature difference between analysis temperature and temperature from experiment 2 with longitude along equator (solid lines). Contour interval is 0.5°C. Average analysis temperature is also shown (dashed lines). Crosses along lower axis show the longitude of TOGA TAO mooring sites. (b) Root-mean-square temperature difference between Analysis temperature and temperature from experiment 2 with latitude along 140°W (solid). Contour interval is 1.0°C. Values exceeding 2°C m are shaded. Average analysis temperature is also shown (dashed). Crosses show the latitude of TOGA TAO mooring sites. (c) Time series of temperature from TOGA TAO (dotted lines) and experiment 2 (solid lines) at 0°N, 155°W.

increases by 1–2 cm to 4–5 cm. The latter represents a lower bound on the accuracy with which the analysis can reproduce observed sea level. When XBT data are withheld, the seasonal cycle is still resolved both along and off the equator. However, there is a tendency toward cooling that is not apparent when XBTs are included. Removal of the TOGA TAO thermistor data has a smaller effect on the resolution of the seasonal cycle in this analysis, with temperature differences introduced by the thermistor data that range from 0.5°C to 1.0°C at thermocline depths. This represents a lower bound on temperature error in the analysis when thermistor data are included.

Because the TAO data set is relatively more dense in time and more uniform in space than the XBT data set, it comes as somewhat of a surprise that XBTs contribute more to the analysis than do the TAO data. The explanation for this result lies in the number of degrees of freedom in the two data sets. Within the latitude band covered by the TAO array, there are typically 50% more XBT observations in each model update interval than daily averaged TAO observations. Assuming that each XBT track line is sampled at a 1° latitude spacing no more frequently than once every 2 weeks and the meridional and temporal decorrelations are 250 km and 2 weeks, we estimate that roughly 40% (100 km/250 km) of the XBT measurements are independent. In contrast only 7 (1 day/14 days) of the daily

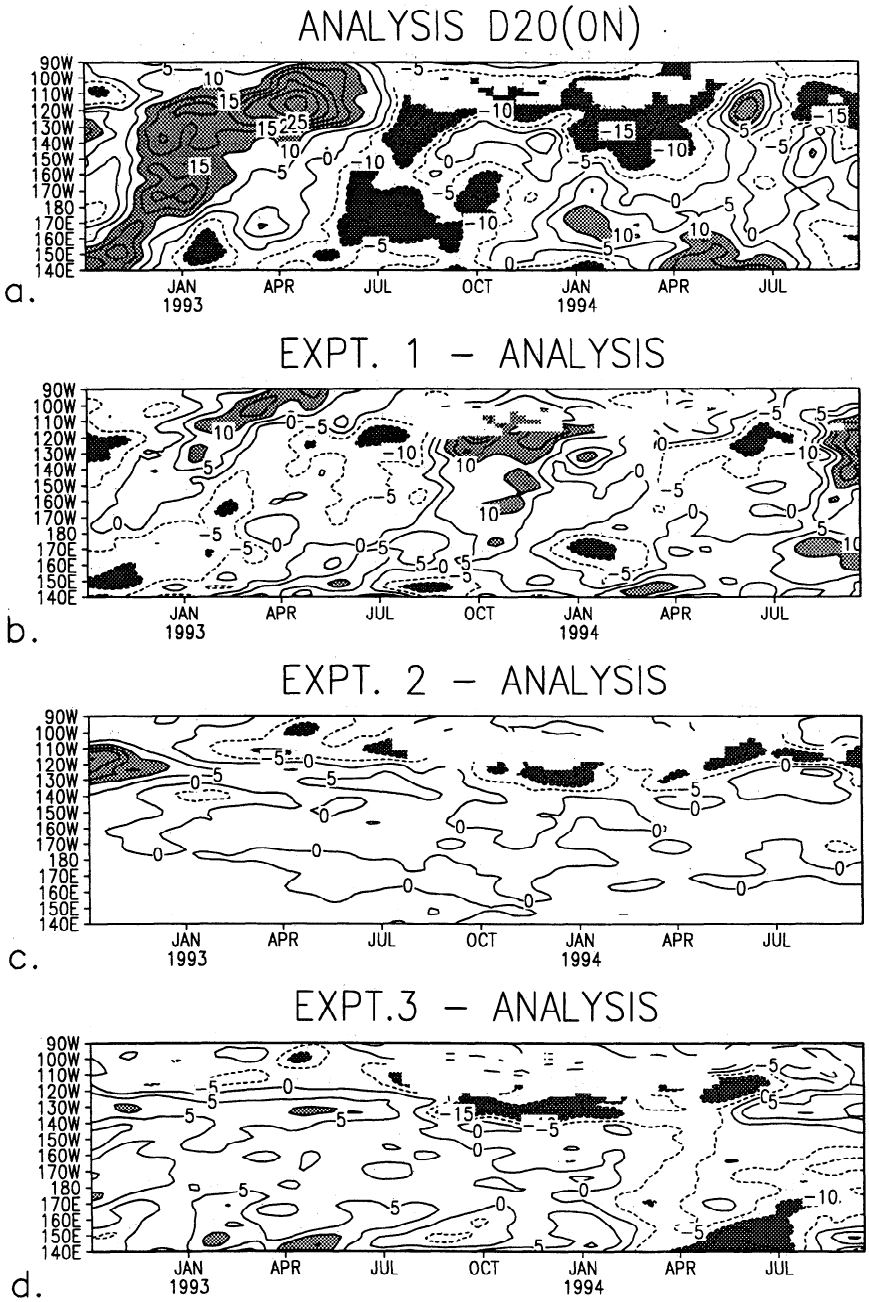

Figure 11. Variation in thermocline depth with longitude and time along equator. The thermocline is represented by the depth of the 20°C isotherm. Two-year averaged depth of the 20°C isotherm is subtracted from each case. (a) Analysis, (b) experiment 1 minus analysis, (c) experiment 2 minus analysis and (c) experiment 3 minus analysis. Vertical axis shows longitude. Areas with no contours indicate that the 20°C isotherm has approached within 7.5 m of the surface. Values exceeding ±10 m are shaded.

TAO measurements are independent. Thus although the XBTs are concentrated mainly along three shipping lanes (intersecting the equator near 150°E, 170°W, and 100°W), the greater degrees of freedom of the XBT data set appears to more than compensate for the greater accuracy and more uniform distribution of the TAO set. Because we consider only the relative impact of each data set in the context of the full suite of measurements, the altimeter data set may be compensating for whatever spatial sampling deficiencies exist in the XBT data set.

The ability of an analysis system to resolve the tropical instability waves provides a test of the resolution of intraseasonal variability. These waves are most evident along the northern edge of the equatorial cold tongue at 5°N between the date line and 110°W. They are reproduced by the model simulation in the absence of updating, but with inaccurate amplitude and phase. Our experiments show that these waves are poorly reproduced without assimilating satellite altimetry. Apparently, the spatial resolution of the XBT and the thermistor data is insufficient to resolve phenomena with spatial scales of hundreds of kilometers and timescales of weeks.

A major application of ocean analyses is in providing initial

conditions for forecasts of El Niño events. The dynamics of these events are frequently diagnosed by plotting thermocline depth along the equator as a function of time, as is shown for our experiments in Figure 11. In late 1992 an intensification of the trade winds in the western basin led to a rise of the thermocline throughout the basin. In the analysis the thermocline rise exceeded 15 m. In the absence of major components of the observing system, these estimates could be modified by 10 m. It is interesting to note, for example, the El Niño of 1992–1993 is stronger and lasts longer in this analysis than in the analysis of *Ji et al.* [1995].

We stated in the introduction that one of our goals is to provide improved initial conditions for seasonal to interannual predictions. To address this goal, a necessary next step is to carry out prediction studies with a coupled model relying on the analysis for initial conditions. A question that will need to be resolved by forecast experiments is whether the expected analysis errors are acceptably small.

These results suggest that satellite altimetry can make an important contribution to analysis of tropical ocean circulation. The future availability of altimeter data looks bright. Currently three altimeters are available, including ERS 1 and ERS 2. Geosat Follow-on should be launched in 1996. Another European altimeter will be carried on ENVISAT in 1999, and an additional French-American altimeter satellite is under consideration as part of the Earth Observing System. Thus we can anticipate high-quality altimetry for the forseeable future.

Acknowledgments. We are grateful to Mike McPhaden of NOAA PMEL for providing us with the thermistor data from the TOGA TAO array, to Bob Cheney of the NOAA Geosciences Laboratory for providing us with the T/P altimetry as well as encouragement, and to Ming Ji and David Behringer of the Coupled Modelling Group, NCEP, NOAA, for providing winds and XBT data. This work was supported by the Office of Global Programs, NOAA, under grant NA26GP0472 (J.A.C.) and NA56GP0191 (B.S.G.). Additional support was provided by NASA to J.A.C. under grant SUB JPL 958123.

References

Berry, P., and J. Marshall, Ocean modelling studies in support of altimetry, *Dyn. Atmos. Oceans*, *13*, 269–300, 1989.

Busalacchi, A. J., M. J. McPhaden, and J. Picaut, Variability in equatorial Pacific sea surface topography during the verification phase of the TOPEX/POSEIDON T/P mission, *J. Geophys. Res.*, *99*, 24,725–24,738, 1994.

Callahan, P. S., TOPEX/POSEIDON project GDR user's handbook, *JPL D-8944*, Jet Propul. Lab., Pasadena, Calif., 1992.

Carton, J. A., and E. Hackert, Data assimilation applied to the temperature and circulation in the tropical Atlantic, 1983–4, *J. Phys. Oceanogr.*, *20*, 1150–1165, 1990.

Cheney, R., L. Miller, R. Agreen, N. Doyle, and J. Lillibridge, TOPEX/POSEIDON: The 2-cm solution, *J. Geophys. Res.*, *99*, 24,555–24,563, 1994.

Daley, R., *Atmospheric Data Analysis*, 457 pp., Cambridge Univ. Press, New York, 1991.

De Mey, P., and A. R. Robinson, Assimilation of altimeter eddy fields in a limited-area quasi-geostrophic model, *J. Phys. Oceanogr.*, *17*, 2280–2293, 1987.

Derber, J., and A. Rosati, A global data assimilation system, *J. Phys. Oceanogr.*, *19*, 1333–1347, 1989.

Dey, C., and L. Morone, Evolution of the National Meteorological Center global data assimilation system: January 1982–December 1983, *Mon. Weather Rev.*, *113*, 304–318, 1985.

Giese, B. S., J. A. Carton, and L. J. Holl, Sea level variability in the eastern tropical Pacific as observed by TOPEX and Tropical Ocean-Global Atmosphere Tropical Atmosphere-Ocean Experiment, *J. Geophys. Res.*, *99*, 24,739–24,748, 1994.

Harrison, D. E., B. S. Giese, and E. S. Sarachik, Mechanisms of SST change in the equatorial waveguide during the 1982–83 ENSO, *J. Clim.*, *3*, 173–188, 1990.

Holland, W. R., and P. Malanotte-Rizzoli, Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies, *J. Phys. Oceanogr.*, *19*, 1507–1534, 1989.

Hurlburt, H. E., Dynamic transfer of simulated altimeter data into subsurface information by a numerical ocean model, *J. Geophys. Res.*, *91*, 2372–2400, 1986.

Hurlburt, H. E., D. N. Fox, and E. J. Metzger, Statistical inference of weakly correlated subthermocline fields from satellite altimeter data, *J. Geophys. Res.*, *95*, 11,375–11,409, 1990.

Ji, M., A. Leetmaa, and J. Derber, An ocean analysis system for climate studies, *Mon. Weather Rev.*, *123*, 460–481, 1995.

Legeckis, R., Long waves in the eastern equatorial ocean: A view from a geostationary satellite, *Science*, *179*, 1179–1181, 1977.

Levitus, S., and T. Boyer, *World Ocean Atlas 1994*, vol. 4, *Temperature*, *NESDIS4*, 117 pp., Natl. Environ. Satell. Data and Int. Serv., Natl. Oceanic and Atmos. Admin., Washington, D. C., 1994.

McPhaden, M. J., TOGA-TAO and the 1991–93 ENSO event, *Oceanography*, *6*, 36–44, 1993.

Mellor, G. L., and T. Ezer, A Gulf Stream model and an altimeter assimilation scheme, *J. Geophys. Res.*, *96*, 8779–8795, 1991.

Meyers, G., G. H. Phillips, N. R. Smith, and J. Sprintall, Space and time scales for optimum interpolation of temperature—Tropical Pacific Ocean, *Progr. Oceanogr.*, *28*, 189–219, 1991.

Miller, R. N., Tropical data assimilation experiments with simulated data: The impact of the Tropical Ocean, Global Atmosphere Thermal Array for the Ocean, *J. Geophys. Res.*, *95*, 11,461–11,482, 1990.

Mitchum, G., Comparison of TOPEX sea surface heights and tide gauge sea levels, *J. Geophys. Res.*, *99*, 24,541–24,554, 1994.

Ray, R., Global ocean tide models on the eve of TOPEX/POSEIDON, *IEEE Trans. Geosci. Remote Sens.*, *31*, 355–364, 1993.

Rebert, J. P., J. R. Donguy, G. Eldin, and K. Wyrtki, Relations between sea level, thermocline depth, heat content, and dynamic height in the tropical Pacific Ocean, *J. Geophys. Res.*, *90*, 11,719–11,725, 1985.

Reynolds, R. W., and T. M. Smith, A high resolution global sea surface temperature climatology, *J. Clim.*, *8*, 1571–1583, 1995.

Shaw, D., P. Lonnberg, A. Hollingsworth, and P. Unden, Data assimilation: The 1984/1985 revisions of the ECMWF mass and wind analysis, *Q. J. R. Meteorol. Soc.*, *113*, 533–566, 1987.

Wacongne, S., On the difference in strength between Atlantic and Pacific Undercurrents, *J. Phys. Oceanogr.*, *20*, 792–799, 1990.

Wagner, C. A., C.-K. Tai, and J. M. Kuhn, Improved M_2 ocean tide from TOPEX/POSEIDON and Geosat altimetry, *J. Geophys. Res.*, *99*, 24,853–24,866, 1994.

White, W. B., G. Meyers, and K. Husunuma, Space/time statistics of short-term climatic variability in the western North Pacific, *J. Geophys. Res.*, *87*, 1979–1989, 1982.

White, W. B., C.-K. Tai, and W. R. Holland, Continuous assimilation of Geosat altimetric sea level observations into a numerical synoptic ocean model of the California Current, *J. Geophys. Res.*, *95*, 3127–3148, 1990.

X. Cao and J. A. Carton, Department of Meteorology, University of Maryland, College Park, MD 20742. (e-mail: carton@atmos.umd.edu)

B. S. Giese, Department of Oceanography, Texas A&M University, College Station, TX 77843.

L. Miller, Geosciences Laboratory, NOAA, Silver Spring, MD 20910.

(Received March 21, 1995; revised February 16, 1996; accepted February 23, 1996.)