

Sea level from temperature profiles in the tropical Pacific Ocean, 1975–1982

Gilles Reverdin,¹ Alexey Kaplan, and Mark A. Cane

Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York

Abstract. We investigate how well the 1975–1992 sea level interannual variability in the tropical Pacific is captured by dynamic height from temperature profiles. For each temperature profile, a surface dynamic height relative to 300 m is estimated, assuming a constant temperature–salinity relationship. After multiplication by a latitudinally varying factor and the removal of a seasonal cycle, the dynamic height deviations fit the tide gauge sea level variability to within the sampling errors, except at a few sites near the equator west of the date line, where surface salinity variability is large. The dynamic height data are assimilated into a wind-forced linear numerical model of the sea level in the tropical Pacific, applying a Kalman filter in a space of reduced dimension. A limited number of empirical orthogonal functions of the unfiltered run (1975–1992) define the reduced space, into which the Kalman Filter covariance evolution calculation is done [Cane *et al.*, 1996]. Experiments indicate that results are better with 32 functions than with a smaller number but are not improved by retaining more functions. The resulting analyzed fields of sea level are compared to withheld dynamic height estimates from moorings, sea level data from tide gauges, and sea level analyses made with the same Kalman filter formalism applied to tide gauge measurements. The comparisons to observations suggest that the temperature profiles were usually sufficient to constrain the monthly analyzed fields to be close to the observed sea level with errors typically less than 3 cm near the equator. The comparison to tide gauge sea level reveals that this analysis is often more accurate than the analysis of tide gauge sea level data with which it shares many characteristics. Near the equator west of the date line, salinity variations are large and their neglect in estimating dynamic height has a negative impact on the analysis. The analyzed signal is underestimated in the southwest Pacific and at more than 20° off the equator. The reanalysis of the temperature data done with a primitive equation model at the National Meteorological Center (NMC) [Ji *et al.*, 1995; Enfield and Harris, 1995] does not share this problem. On the other hand, NMC reanalysis (RA4) departs more from the observations elsewhere, although more data were included than in our analysis.

1. Introduction

Accurate analyses of sea level and thermal structure will provide a better understanding of the oceanographic variability associated to El Niño/Southern Oscillation (ENSO) events. They are also required for initializing forecast experiments of short-term climatic variability [Ji *et al.*, 1994; Rosati *et al.*, 1996]. The current availability of comparatively accurate sea level data from satellite altimetry (TOPEX/POSEIDON since November 1992) provides an additional stimulus for providing long time series of sea level fields against which the newer data can be referenced [Busalacchi *et al.*, 1994]. The work presented in this note is aimed at testing the impact of temperature profiles in constraining these analyses. We will focus on analyses of sea level, because it is possible to compare the fields with tide gauge data and with their analyses [Miller and Cane, 1989]. It complements the study of Cane *et al.* [1996] where the methodology is presented and the impact of tide gauge data is investigated, and that of Miller *et al.*

[1995], which investigates the impact of a subset of the temperature profile data along major shipping lanes. Similar comparisons for the National Meteorological Center (NMC) reanalyses of temperature profiles done by a variational assimilation in an oceanic general circulation model of the equatorial Pacific are presented by Enfield and Harris [1995]. The model used here is much simpler, while the assimilation technique is more elaborated than theirs. Furthermore, instead of assimilating the vertical temperature structure, we assimilate the dynamic height. However, because sea surface steric elevation and upper thermocline isotherm displacements are typically closely correlated in the equatorial Pacific [e.g., Rebert *et al.*, 1985], the information in shallow temperature profiles reaching only 300 m is largely captured in its sea surface temperature and an estimated surface steric elevation. Also, the study of modal decomposition by Hayes *et al.* [1985] shows that much deeper vertical profiles than most available would be needed to reconstruct the vertical structure of isotherm displacements.

Rebert *et al.* [1985] demonstrate that dynamic height from temperature profiles is strongly correlated with sea level for a wide range of latitudes in the central tropical Pacific ocean. Taft and Kessler [1991] expanded this work as well as Wyrki's [1978] study to a longer period and more sites in the central Pacific Ocean. They conclude that within 15° of the equator, surface elevation estimated from temperature profiles reaching a depth of

¹ Now at Groupe de Recherche de Géodésie Spatiale, UMR CNES/CNRS 5566, Toulouse, France.

400 m correlates well with sea level and is adequate to derive geostrophic currents. Sea level and temperature content were also compared at sites in the western equatorial Pacific by *Delcroix and Gautier* [1987], who suggested that the two quantities were related but who also highlighted differences resulting from the contribution of salinity variations on the sea level. *Kessler and Taft* [1987] presented an example of the large effect of an anomalous freshening on surface dynamic height estimates in October 1982 near Christmas (1°59'N/157°29'W).

The analysis strategy here is to assimilate the data into a numerical model known for its reasonable skill in simulating ENSO variability in the equatorial Pacific [*Cane, 1984; Busalacchi and Cane, 1985*]. The model is a two-mode version of the *Cane and Patton* [1984] algorithm for solving the linear long wave approximation to the shallow water equations on the equatorial beta plane. It is less diffusive and has a higher (0.5°) meridional resolution than the version of the model used by *Miller et al.* [1995]. The model has no thermodynamics and is forced by surface wind stress. The wind anomaly fields we use, from the Florida State University (FSU) pseudo stress analysis [*Goldenberg and O'Brien, 1981*], have been smoothed and detrended [*Cane et al., 1986*], and the drag coefficient used with these monthly means has been adjusted to provide an adequate representation of the sea level variations during the 1982-1983 El Niño [*Busalacchi and Cane, 1985*]. The standard wind-driven model is run for the period 1964-1992 in a domain extending

between 28.75°N and 28.75° S and between 124°E and 80°W across the Pacific with a very simplified geometry; this will be referred to as the "unfiltered run." This run often presents less variance than tide gauge sea level, but the difference does not justify increasing the drag coefficient. An empirical orthogonal function (EOF) decomposition of the unfiltered run state variables is done, from which a reduced space can be defined by retaining a subset of EOFs.

Monthly dynamic height data are assimilated for the period 1975-1992, the same period chosen for the tide gauge assimilation using the same methodology presented by *Cane et al.* [1996]. The scheme incorporates a Kalman filter applied each month in the reduced space of a subset of EOFs. *Cane et al.* [1996] give an extensive presentation of the method, including various consistency checks of the error estimates. Discussions on the reliability and the interpretation of the error estimate are presented by *Cane et al.* [1996] and a discussion of the influence of the error model is provided by *Miller et al.* [1995].

Instead of assimilating dynamic height, the respective contributions of the two vertical modes could have been extracted from the temperature profiles. While it is in principle possible to assimilate data for the two vertical modes separately, the temperature profiles do not extend to a sufficient depth to define a second vertical mode unambiguously [*Hayes et al., 1985*]. Therefore we opt to carry the assimilation on sea level estimates. A very close alternative would have been to carry it on a model

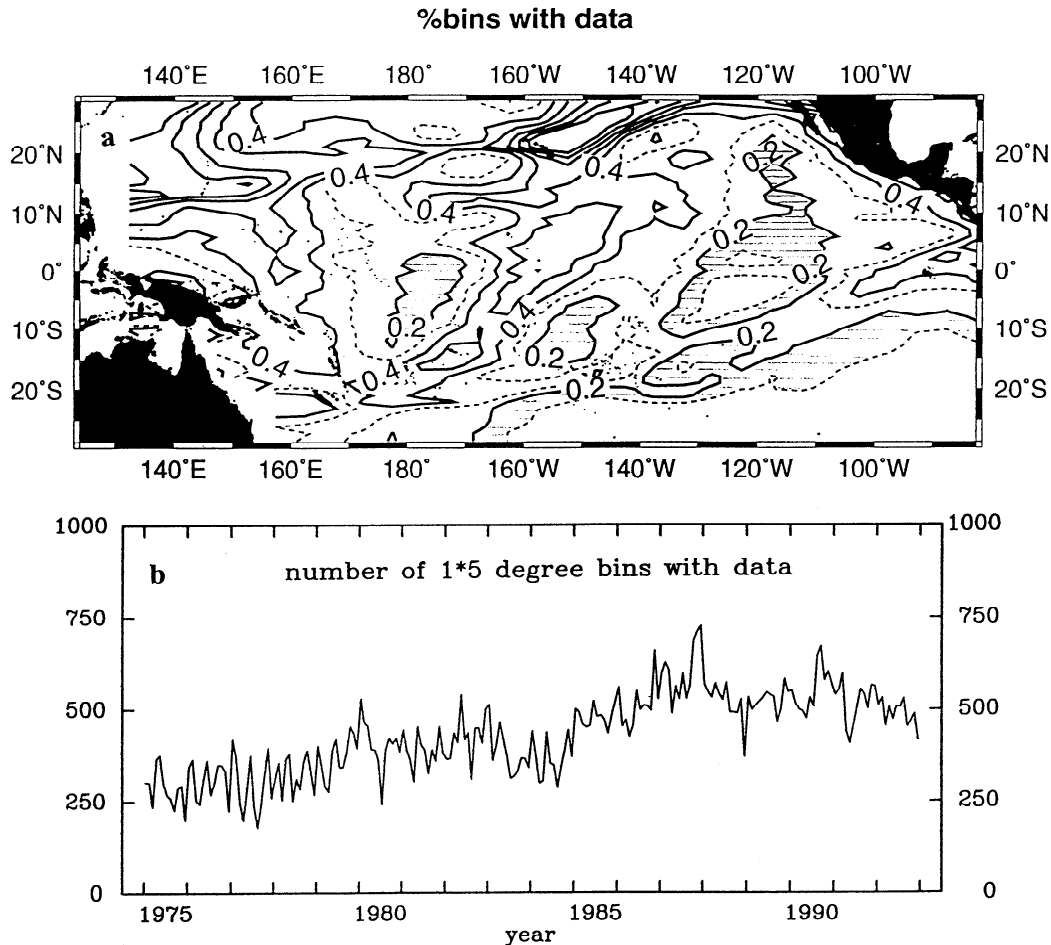


Figure 1. Data distribution: (a) spatial distribution of percentage of monthly 1°x5° bins with data within the 216 months of data (1975-1992) (contour 0.1 (10%); shaded between 0.1 and 0.2); (b) time evolution of the number of monthly 1°x5° latitude x longitude bins with data (there are 1655 bins in a given month).

dynamic height estimated from the vertical modes with the same reference level as in the data. The relative weight of the two vertical modes varies from one EOF to another, so, in principle, the assimilation can modify the relative contributions of the two vertical modes to sea level as compared to the unfiltered run. However, the two modes are correlated in the unfiltered run at the low frequencies we consider. This is especially true near the equator where the correlation coefficient usually exceeds 0.8. Hence the two modes are not well separated by the EOFs [Busalacchi and Cane, 1985], and we do not expect that assimilating sea level data will modify the ratio between the two modes substantially. Exceptions may occur if an adjustment of the meridional structure of variability is required by the data.

A noteworthy feature of the assimilation experiments with tide gauge sea level data presented by Cane *et al.* [1996] was that the results are degraded when too many EOFs are retained. Comparisons with withheld tide gauge stations suggest that overfitting the limited set of data (34 tide gauges) degrades the analysis away from the assimilated tide gauge stations. Spatial coverage is more complete for the temperature profiles, although large areas which are devoid of data remain (Figure 1a). Little information is available in large portions of the northeastern equatorial Pacific between the equator and 20° N, as well as in the southeastern Pacific. The data coverage has changed in time (Figure 1b), with more data from Japanese fishery vessels before 1980 and a higher density of data collected along lines for the Tropical Ocean and Global Atmosphere (TOGA) and World Ocean Circulation Experiment (WOCE) programs toward the end of the period (after 1985). We did not incorporate data from the Tropical Atmosphere-Ocean (TAO) surface moorings [McPhaden, 1993; Hayes *et al.*, 1991], which can therefore be used as an independent set of data against which to verify the analysis.

2. Estimation of Sea Level Deviations From Temperature Profiles

To estimate sea level from temperature profiles, we estimate dynamic height using temperature profiles and a modeled salinity from an average S(T) relationship. The dynamic height is computed relative to 300 m, and to estimate sea level height, a multiplicative factor is derived from comparisons with tide gauge sea level. It is well known that deep isopycnal vertical

displacements contribute significantly to surface sea level variability, in particular poleward of 10° off the equator [Taft and Kessler, 1991], whereas its contribution is smaller near the equator, probably of the order of 1 cm rms. Our approach is an empirical attempt to include a portion of this variability correlated to the one in the upper 300 m.

Temperature profiles incorporated in this analysis are: expendable bathythermographs (XBTs) often reaching 450 m (the larger source of data), Nansen casts often reaching 1000 m, and mechanical bathythermographs (MBTs), primarily from Japanese fisheries, reaching only 250 m. Because of these last profiles (which contribute to a large share of the total number of data before 1975) and because ultimately we also want to incorporate data from thermistance chains on the TAO surface moorings, which only reach 300 m, we have chosen a level of reference at 300 m for the computation of dynamic height. We also incorporate in the analysis profiles reaching only 250 m. For these profiles, we extend the profile downward to 300 m by comparing it to climatology and assuming similar vertical displacements below 250 m as between 150 m and 250 m. The top of the thermocline in the tropical Pacific within 30° of the equator is usually above 250 m, the occasional exceptions occurring mostly in the extreme northwest and south to southwest areas. The resulting error of extrapolating too shallow profiles is less than 1 cm rms according to tests with deeper profiles. Depth estimates of XBTs are known to be biased [Kessler, 1989], which we correct by multiplying the reported depth by 1.034 [Hanawa *et al.*, 1994].

In addition to temperature, salinity is needed in order to derive a dynamic height. Gridded fields of surface salinity are estimated from 38057 conductivity-temperature-depth (CTD) and Nansen casts, as well as gridded fields of S(T) below the surface layer down to the intermediate layers (5°C). The salinity data are too scarce to establish clearly the seasonal cycle, and we have not incorporated additional surface salinity data which have been used to derive the surface seasonal cycle [Delcroix and Hénin, 1991; Donguy, 1994]. Below the near-surface layer, we construct an average S(T) relationship by averaging individual profiles of S(T) within 1° latitude by 5° longitude boxes. Gaps are filled and the fields are smoothed, mostly in longitude, using a local function fitting algorithm [Reverdin *et al.*, 1994]. The average salinity at the surface and between the surface and the first subsurface isotherm is also estimated

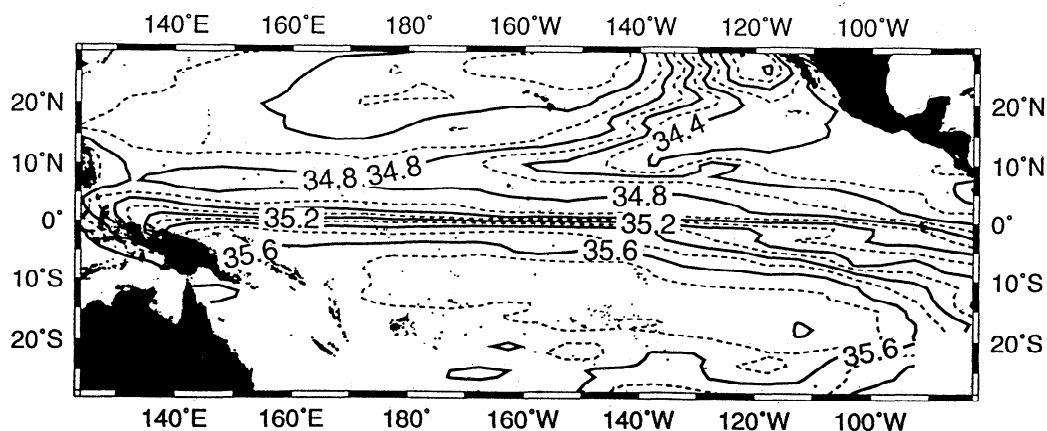


Figure 2. Average salinity at $T = 20^{\circ}\text{C}$ on the $1^{\circ}\times 5^{\circ}$ latitude \times longitude grid. Smoothing of the raw estimates was done by an objective function fitting algorithm involving at each grid point the neighboring grid points in latitude and longitude.

An example of a resulting $S(T)$ field is presented for the core of the thermocline in Figure 2. It shows the contrast of $S(T)$ across the equator between the high-salinity central waters of the subtropical South Pacific and the fresher and less ventilated waters to the north of the equator [Tsuchiya *et al.*, 1989]. This contrast extends from the vicinity of the surface down to 11° C. The tongue of low-salinity water centered at 10°N connects to the low-salinity waters off Baja California at the depth of the 20°C isotherm. At shallower surfaces, however, the 10°N tongue is also connected with the freshest water off Central America. The more saline central waters of the North Pacific, which originate north of Hawaii, are found north of 10°N (see the 35.0 contour on Figure 2).

For each temperature profile, we derive a salinity from the local $S(T)$ relation unless T is within 1° C of the surface. In this surface layer, we interpolate S linearly with depth between the $S(T)$ at the base of the layer and the surface S climatology. Then, a dynamic height is computed for the 0/300 m layer. To estimate the error associated with the assumption on salinity, we compare estimated dynamic height to dynamic height using measured salinity for the Nansen and CTD casts, as well as for the 0°N/165°E mooring for which temperature and salinity are measured by Seabird SEACATs (model SBE-16) [Sprintall and McPhaden, 1994]. According to these limited comparisons, the assumption made on salinity results in an rms error varying from 1.5 cm in the eastern Pacific and off the equator to 3.0 cm in the western equatorial Pacific (see also Taft and Kessler [1991] and Busalacchi *et al.* [1994]).

The numerical model calculates sea level deviations from the average seasonal cycle between 1975 and 1986. To assimilate the temperature profiles, we must relate 0/300 m dynamic height to sea level, and we must estimate the seasonal cycle and remove it from the data.

Observations of isotherm vertical displacements at many sites suggest that the displacements in the upper 1000 m of the water column vary almost in phase with the upper thermocline displacements, especially 5° or more away from the equator. This explains why at Honolulu, Rebert *et al.* [1985] found that the 0/300 m dynamic height, D , was correlated to sea level but that the dynamic height variability had a smaller amplitude than sea level. This suggests that some of the deep variability is in phase with the thermocline displacements. Closer to the equator, the

deviations at depth are not in phase with the upper thermocline displacements (Lukas and Firing [1985] for the seasonal cycle; Tang *et al.* [1988] for interannual deviations). However, they don't seem to contribute much to low-frequency sea level variability. To roughly take into account this dependency on latitude, we estimate sea level as $D \times \{1 + a [\int_{1000}^{300} N^2 dz / \int_{300}^0 N^2 dz]\}$, where N is the Brunt-Vaisala frequency and a is chosen as follows: $a = 0.75$ at latitudes higher than 10°, decreasing linearly to 0.1 at 5°, and $a = 0.1$ between 5°N and 5°S. This was chosen to adjust the magnitude of interannual rms variability of the monthly sea level estimates to interannual rms variability of sea level from tide gauges.

A gridded seasonal cycle of this proxy sea level is estimated by binning the individual estimates by month and 1° latitude by 5° longitude boxes. The box scales are chosen to be proportional to the decorrelation scales of the thermocline depth estimated by Meyers *et al.* [1991]; the unfiltered model simulation is highly correlated on those scales. Notice, however, that near 20°N or 20°S a large signal associated with Rossby waves at a few months period is not correlated over the box scale [Mitchum, 1994]. The average seasonal cycle retaining the annual and semiannual harmonics is estimated from these binned values between 1975 and 1986 (a similar objective mapping is used as for $S(T)$). Large errors on the average seasonal cycle, especially near the equator, result from the insufficient sampling of the interannual variability. We found that a more robust seasonal cycle could be estimated by first subtracting from the data the interannual sea level assimilation of tide gauge data [Cane *et al.*, 1996]. This seasonal cycle is then removed from the individual data to estimate interannual deviations. Errors on the estimated seasonal cycle result mainly from the data sampling but also from the neglect of any seasonal salinity variations. Individual data with large deviations which deviate strongly from other nearby data and the assimilation of tide gauge sea levels are removed (2.2%) and binned monthly deviations are estimated.

Error bars on the monthly binned deviations are presented in the Appendix with the comparisons between the binned sea level estimates and the tide gauge sea levels. These errors are estimated assuming that they originate from the sampling of high-frequency variability and the imposed salinity. These two errors are uncorrelated, and we optimistically assume that errors are uncorrelated between different bins. We assume that high-

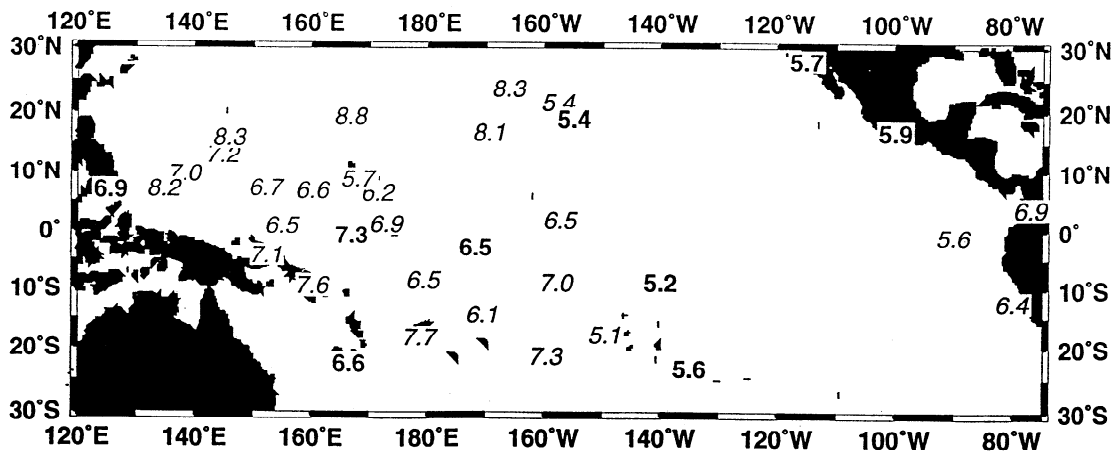


Figure 3. Rms differences (centimeters) between tide gauge sea level records in 1975–1992 and the unfiltered simulation (average seasonal cycle for 1975–1992 removed). The values are given in boldface when rms is lower than in the lower panel of Figure A1 (comparison of temperature profile-derived sea level (TSL) and tidegauge sea level (SL)).

frequency variability is a white noise with variance estimated from smoothed maps of the average variance within a month. Near the equator, the sampling of high-frequency variability or random errors contributes to errors on monthly means often exceeding 3 cm. The error resulting from the assumed constant $S(T)$ relationship is taken from a smoothed map of the variance in the difference with direct estimates of dynamic heights for Nansen and CTD casts. Central Pacific data [Taft and Kessler, 1991] suggest that this error is likely to be correlated over a few months to a few years and meridionally over hundreds of kilometers. We therefore assume that it is perfectly correlated over a bin. The linear formula used to derive sea level from 0/300 m dynamic height introduces an additional error. Its magnitude is difficult to estimate but is probably smaller than the other two and is not included in the error estimates. The comparisons in the Appendix between the temperature profile-derived sea level (TSL) and the tide gauge sea level (SL) suggest that these error estimates usually have the right magnitude. The rms differences and the correlation coefficients between TSL and SL suggest that TSL is reasonably close to SL, in particular near 10°N and 10°S

and in the eastern equatorial Pacific (see also Figure A1). It is comforting that these rms differences are often (26 out of 35) smaller than the ones between the unfiltered run and the tide gauge sea levels (Figure 3). Therefore incorporating the temperature data should improve the sea level of the model run.

3. Results of the Assimilation

The 86,616 monthly binned data with their estimated error are assimilated using a Kalman filter following the formalism discussed by Cane *et al.* [1996]. The model error is assumed to originate from random errors in the wind field whose statistical characteristics do not have spatial structure. The model error variance is estimated based on the differences between the unfiltered model runs and the data. Its variance is chosen smaller here than when compared with tide gauge data for the experiments presented by Cane *et al.* [1996]. The assumptions on the spatial homogeneity and randomness of the wind error and that data errors in different bins are uncorrelated and are certainly not valid, but their effect are difficult to evaluate. Error maps are

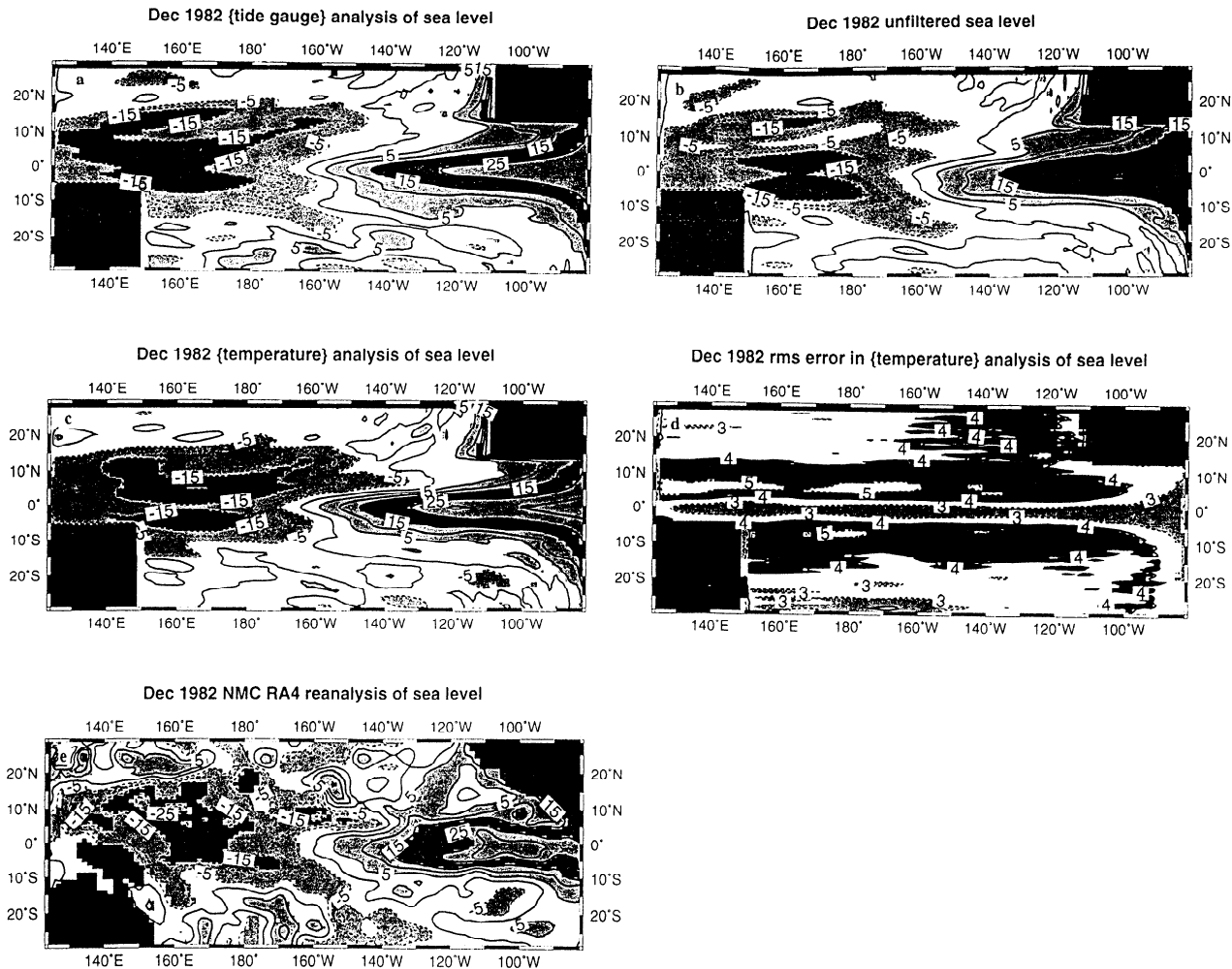


Figure 4. Anomaly fields for December 1982 (centimeters): (a) the Kalman filtered analysis of sea level from the tide gauge sea level; (b) the unfiltered run; (c) the Kalman filtered analysis of sea level from the temperature profiles (TASL) and from tide gauges, respectively; (d) the estimated rms error on the analysis TASL (this is the grid point analysis error, which is larger than the error estimated in the subspace of 32 empirical orthogonal functions (EOFs)); (e) the National Meteorological Center (NMC) reanalysis (RA4) of the temperature profiles [Enfield and Harris, 1995]. For RA4, the seasonal cycle is removed for 1982–1992; for the others, the average seasonal cycle is removed for 1975–1992.

produced which depend on the model error, although the sensitivity to it is not too large [Miller *et al.*, 1995]. The analysis is done in a space of reduced dimension whose possible states are spanned by a truncated set of empirical orthogonal functions (EOFs) of the unfiltered run. We will refer to the result of the Kalman filter as the temperature analysis of sea level (TASL). Assimilation experiments were done with a varying number of EOFs.

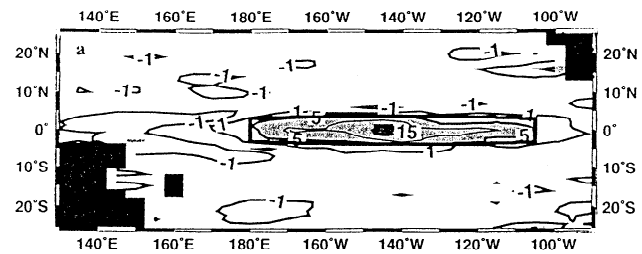
Comparisons are required to find whether the analyses and their error estimates are realistic. These analyses can be compared to the original data, to independent sea level estimates from temperature moorings, or to tide gauge data (a more direct estimate of sea level). They can also be compared to analyses of tide gauge data (GASL) using the same model and data assimilation methodology [Cane *et al.*, 1996] or to analyses of temperature profiles done at NMC (NMCSL) by variational data assimilation in a primitive equation model [Ji *et al.*, 1995; Enfield and Harris, 1995]. TASL is produced for 1975–1992. This is the same period as the GASL and most tide gauge records. However, most mooring data begin after 1987, and the overlap of TASL and NMCSL is for July 1982 to 1992.

In Figure 4, sea level fields are presented for one particular month, December 1982, at the peak of the 1982–1983 El Niño. GASL and TASL are closer to each other than to the unfiltered run, which has markedly smaller deviations in the eastern Pacific and in the western Pacific just north of the equator. The deviations in NMCSL (departures from an average seasonal cycle in 1982–1992) have comparable values to TASL but more synoptic scale structure. In particular, in NMCSL the maximum deviation near the equator in the eastern Pacific is more spread meridionally, and the three bands of high values in the western Pacific are less distinguishable (in particular, the one north of 10°N, which appears as three isolated areas). These differences are typical of the ones found in other months, as is the estimated error map of TASL with smaller values at the equator than near 10°N and 10°S.

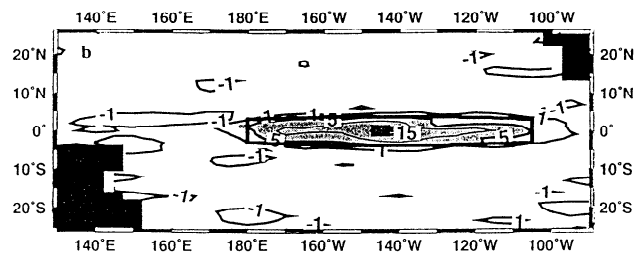
The question of how to validate these analyses will be addressed with the aim of finding which representation of sea level provides the smallest rms errors in the large scale structures active in low-frequency ocean dynamics. Comparison to the unfiltered run will not properly credit the influence of assimilated observations. Comparing the analysis to data included in the assimilation is not ideal either. The difference between analysis and data would be smallest if the analysis overfits the data, including features that are real but local, but also data noise, so that one cannot judge how the largest scales of the variability are reproduced.

Another question to be addressed is the influence of the dimension of the reduced space in which the Kalman filter formalism is applied. If the observations are sparse, the analysis would not be improved by increasing the number of EOFs beyond the few needed to capture what little information the data provide about large-scale structure. Adding more EOFs will allow the analysis to fit the data points more closely, but this will not be indicative of the quality of the large scales, which are likely to be degraded by such overfitting (this is related to the issue that the system noise be red enough, as discussed by Cane *et al.*, [1996]). If the data distribution is adequate to define the scales of the variability, one expects the analysis to improve with increasing numbers of EOFs. However, it is possible that the EOF structures extracted from the unfiltered run do not match well the observed variability, in which case the analysis would not be improved by adding EOFs. Fortunately, the comparisons for this analysis

change of variance to obs. when data withheld from box 6 (17 EOFs)



change of variance to obs. when data withheld from box 6 (32 EOFs)



change of variance to obs. when data withheld from box 6 (93 EOFs)

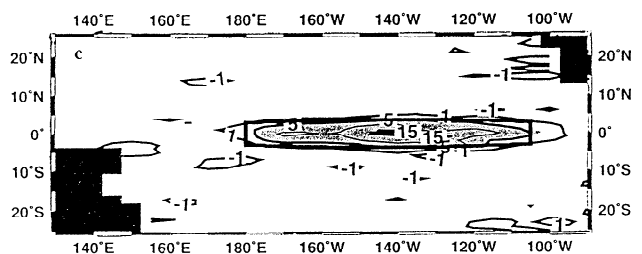


Figure 5. Changes in the variance (square centimeters) of the differences to the observations when data from area 6 are removed from the analysis (area 6 is bounded by the heavy line rectangles). The estimates on the 1°x5° grid have been smoothed by a running mean, first in longitude with weights (1/4, 1/2, 1/4) and then in latitude with weights (1/8, 6/8, 1/8). The different panels correspond to assimilation experiments with different numbers of EOFs.

suggest that this is not a major problem, except in the vicinity of the continents. Improvement will cease when the information in the data is properly represented by the EOFs. Adding additional EOFs may make matters worse by including more noise in the analysis imposing spurious structure (i.e., the comparison with data not included in the analysis will be worse). Table 1 presents comparisons between the analysis with data in some relatively well sampled areas, representative of equatorial and off-equatorial domains. It also presents the comparison of data with the 1-month forecast (the numerical simulation is forwarded 1 month after the analysis). Note that 17, 32, and 93 EOFs account for 80%, 90%, and 99%, respectively, of the variance in the unfiltered run (cf. Cane *et al.* [1996]). The comparisons for the analysis show that the rms differences with assimilated data decrease only marginally when the number of EOFs is increased beyond 32; even the difference between 17 and 32 is not large. These differences are comparable to the estimated rms error on the binned averages; that is TASL fits the data closely even with 32

Table 1. Rms Differences of the Kalman Filtered Analysis (Upper Line) and 1-Month Forecast (Lower Line) for Various Numbers of Empirical Orthogonal Functions (EOFs) With the Sea Level Observations From Temperature Profiles

Area	17 EOFs		32 EOFs		93 EOF		Rms Error
	With	Without	With	Without	With	Without	
1 17°N-22°N 128°E-150°E	9.7 9.8	9.9 9.9	9.5 9.8	9.8 9.9	9.4 9.8	9.9 10.0	8.3
2 3°N-5°N 135°E-160°E	5.8 6.7	6.1 6.8	5.5 6.7	6.0 6.8	5.1 6.5	5.8 6.6	5.6
3 5°S-7°S 155°E-180°E	5.8 6.8	6.1 6.9	5.4 6.6	5.9 6.7	5.0 6.8	5.8 6.9	5.4
4 4°N-10°N 155°W-180°W	6.2 6.7	6.6 6.8	5.9 6.5	6.4 6.6	5.6 6.6	6.4 6.8	5.6
5 10°N-15°N 155°W-180°W	5.2 5.4	5.6 5.6	5.0 5.4	5.6 5.6	4.8 5.5	5.7 5.9	5.0
6 3°N-3°S 105°W-180°W	4.5 5.8	5.4 5.9	4.3 5.8	5.3 5.9	4.1 5.7	5.1 5.8	4.8
7 5°N-10°N 95°W-120°W	5.9 6.2	6.1 6.2	5.7 6.1	6.1 6.2	5.3 6.1	6.0 6.2	5.0
8 3°S-6°S 95°W-120°W	4.2 4.5	4.3 4.5	4.2 4.6	4.3 4.6	4.1 4.6	4.3 4.6	4.7
9 3°N-5°S 135°E-180°E	4.8 5.7	5.7 6.1	4.5 5.7	5.5 6.0	4.2 5.6	5.4 5.9	5.0

The comparison is done for specific areas both when all data are assimilated and when data from that area are withheld. The last column reports the estimated average rms error of the binned data in the area.

EOFs. The 1-month forecasts show even less difference as the number of EOFs is varied.

For purposes of both validation and choosing the number of EOFs, a more revealing test is to withdraw observations from the analysis which may then provide an independent test of its accuracy (similarly to what has been done for the assimilation of tide gauge data of *Cane et al.* [1996]). For example, we withheld all observations in specified areas (examples provided in columns marked as "Without" of Table 1, Figure 5) to test whether the spatial correlation in the analyzed field imposed from the limited set of spatial structures is real and whether the error estimates are realistic. The conclusions are expected to be somewhat dependent on the size of the area from which data are withheld and on the actual distribution of data in surrounding areas. Usually, we find that removing data from an area primarily affects nearby areas and has the largest effect where data were removed. Even so, the error in the withheld area is not as large as it was for the unfiltered run. Comparisons of the 1-month forecasts confirm the effectiveness of the assimilation: for the simulated field 1 month after the assimilation step, it is often difficult to identify the difference between the run with data withheld and the run with all data used, suggesting that information from nearby areas propagates within 1 month.

Comparison of the analysis to observations in withheld areas (Table 1) suggests that 32 EOFs is a small but notable improvement over 17, while almost tripling the number to 93 brings only a slight additional improvement. It would be hard to argue for the statistical significance of any of these differences, which are at most a few millimeters in rms departures from

observations. In their assimilation of 34 tide gauge stations, *Cane et al.* [1996] found that 17 EOFs performed marginally better than retaining a higher number of EOFs. The spatial coverage of the data is better in the present study, with an average of 401 1° by 5° boxes occupied each month out of 1655. One would thus expect these data to require more structure functions to be represented properly. Still, as noted above, including too many EOFs can degrade the analysis away from the data. On the other hand, when the number of EOFs is small (17), withholding data in one area actually often improves the analysis outside of that area, as illustrated by the large areas with values less than $-1 \text{ cm}^2 \text{ s}^{-2}$ on the upper panel of Figure 5. This suggests that 17 EOFs do not capture well the structure of the signal in the data. We will therefore present in the following the analyses for 32 EOFs, which seems a minimum number required to avoid too much influence of the choice of a particular set of EOFs.

Another data set, withheld from the analysis, originates from the TAO array of temperature moorings in the equatorial band [*McPhaden*, 1993; *Hayes et al.*, 1991]. Sea level is estimated from the temperature profiles sampled at discrete levels on a mooring. First, a cubic spline fit is applied to deviations from an average profile characteristic of the area to create a continuous profile. From this profile we estimate dynamic height as was done for the other temperature profiles. Thus these dynamic heights have the same error due to the assumption on salinity as the data assimilated. The monthly binned values from the moorings are free of temporal sampling error but present an additional error resulting from the limited number of measurements levels. This error was checked based on continuous temperature profiles in

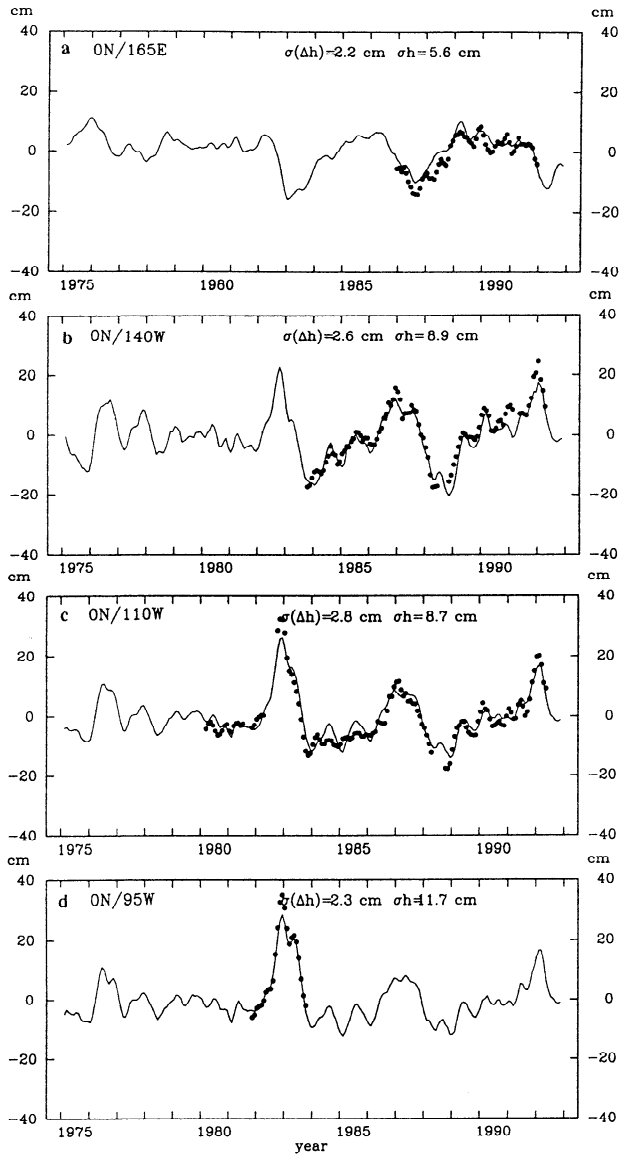


Figure 6. Comparison along the equator of the Kalman filtered analysis of the temperature sea level (solid line) to sea level from moorings (dots) not included in the analysis: (a) 165°E; (b) 140°W; (c) 110°W; (d) 95°W. The time series have been filtered by a 3-month running mean. Rms differences between the two series $\{\sigma(\Delta h)\}$ as well as the rms variability in the joint time series $\{\sigma h\}$ are indicated for each plot.

their vicinity. These estimates have virtually no systematic bias and rms errors of the order of 1 cm, which should be smaller when considering monthly binned averages. The comparisons between TASL and the moorings along the equator at 0°/165°E, 0°/140°W, 0°/110°W, as well as the shorter records at 0°/124°W and 0°/95°W, suggest quite small errors in TASL: less than 3 cm rms on the three-month running means (Figure 6). To some extent, this is because some of the well-sampled areas are in close proximity to the moorings (Figure 1a). The monthly rms differences are less than the error estimates of the analysis, which are larger than 3 cm near the equator (example on Figure 4). This is reasonable in that the error estimates include the error for the differences between a sea level and a dynamic height, which is not part of the difference in this comparison. Some of the other

shorter records are especially interesting because they are not in well-sampled areas. The comparison along 110°W with the moorings at 5°N, 2°N, 0°, 2°S, and 5°S (Figure 7) suggests that TASL captures well the meridional structure of interannual variability in the eastern Pacific in 1985–1992, with a peak in variability at the equator and smaller variability at 5°N than at 5°S. The comparison is less satisfactory along 165°E, in particular at 5°S (not shown), where the analyzed variability is less than in the observations, particularly for the episodes of low sea level during El Niño events.

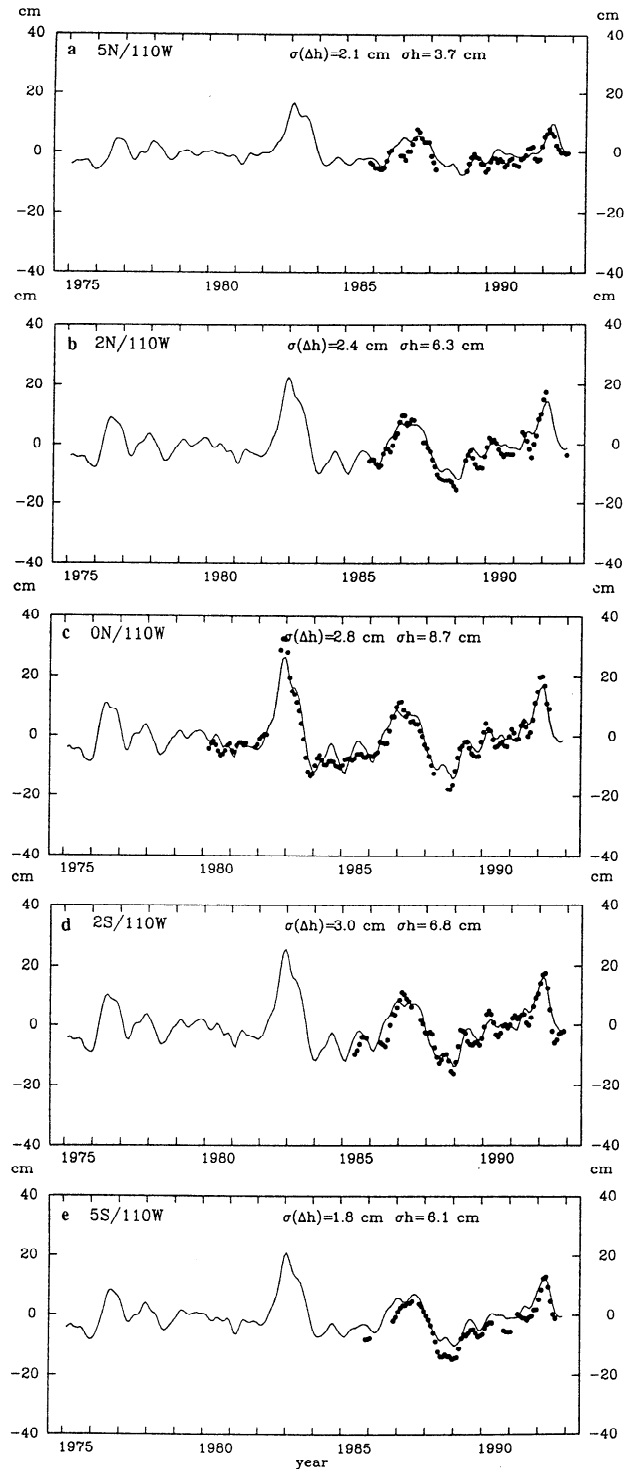


Figure 7. Same as Figure 6, but along 110°W.

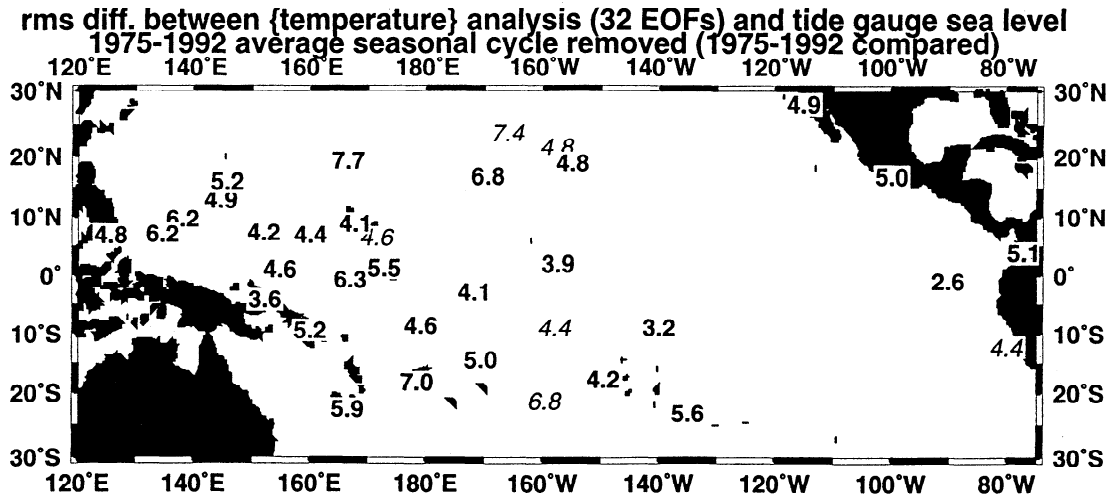


Figure 8. Rms differences (centimeters) between the filtered analysis of the temperature sea level with the tide gauge sea levels. The values are given in boldface when they are lower than in the lower panel of Figure A1 corresponding to the comparison between the sea level data from temperature profiles and the tide gauge sea level.

Though expressed in terms of sea level, the comparisons presented up to now have only addressed the ability of the fields to reproduce dynamic height estimated from temperature profiles. The remaining question is whether TASL also provides an accurate representation of sea level. This issue was also tackled for the NMC reanalyses by *Enfield and Harris* [1995]. We will therefore compare TASL to the tide gauge sea level records and also to GASL, the analysis of the tide gauge sea level records with the same model and assimilation procedure. This last comparison provides a different perspective than the direct comparison to the tide gauges, because the two sets of assimilated fields are filtered on the same sets of spatial functions and the Kalman filter of GASL has propagated spatially the information of the tide gauge records.

Since rms differences between dynamic height and tide gauge records of sea level are smaller than between the unfiltered model and these records (Figure 3), it is not surprising to find that the rms differences with the model fields decrease after applying the Kalman filter. Except for six instances, the rms differences between TASL and tide gauge records (Figure 8) are also lower than when considering the comparisons between the dynamic height data and the tide gauge sea level. This fortunate feature indicates that the analysis reduces the noise in the local data by propagating information spatially.

We will illustrate the comparison of the analysis with sea level data for a few individual stations. In this case, we adopt a longer TASL experiment ending in August 1995, and we correct the tide gauge records for atmospheric pressure variations, using an inverse barometer correction of sea level (this is not done for the other figures nor for GASL). This is expected to remove a large part of the barotropic ocean response to the atmospheric pressure variations (though this does not hold at higher frequencies [Luther, 1982]). The inverse barometer model using atmospheric sea level pressure from the Comprehensive Ocean-Atmosphere Data Set corresponds to monthly corrections of maximum amplitude of ± 4 cm within 20° of the equator. The correction is often correlated with the sea level signal. For example, higher-than-normal atmospheric pressure in the western Pacific during El Niño contributes to low sea level, but this effect is less than 10% of the observed sea level variability. The correction often

improves the comparison with the analysis, although the improvement is usually less than 0.2 cm for rms deviations. Not surprisingly, at stations for which the dynamic height data were close to the sea level records, the agreement for TASL is good. This is particularly true at Santa Cruz or Christmas (Figures 9a and 9e). More generally, it is interesting to find from the comparisons to the equatorial moorings and tide gauges that the zonal differences in the eastern Pacific are well reproduced in this analysis. For example, the second episode of high sea level of the 1982-1983 El Niño appears in early 1983 and increases relative to the first one from the central Pacific to South America, a feature which is already present to some extent in the unfiltered run [Busalacchi and Cane, 1985]. The agreement is also fairly close near 5°N - 10°N except in the extreme western Pacific. At many sites between 5°S and 15°S in the western Pacific, TASL has less variability than observed with a systematic underestimation of the low sea level events associated with El Niño (namely, Honiara on Figure 9d). As this is not a deficiency of the original dynamic height data, we conclude that the data were not sufficient to correct an extensive deficiency of the unfiltered simulation. On the other hand, the lag of 1 to 2 months at Honiara ($9.4^\circ\text{S}/160^\circ\text{E}$) of TASL with respect to sea level (the lagged correlation is 0.93 compared to 0.90 at lag 0) is also present in the dynamic height data. This might be related to real variations on a small spatial scale related to the bathymetry and the presence of islands. At latitudes close to 20°N or further north, TASL has too little rms variability compared to the tide gauge sea level or nearby dynamic height data. This effect decreases when a larger number of EOFs is included, allowing the variability in the data to be represented in the analysis.

Results are barely improved over the unfiltered run at the near-equatorial sites of the western and central Pacific, in particular at Nauru ($0^\circ 32'\text{S}/166^\circ 54'\text{E}$) and Tarawa ($1^\circ 22'\text{N}/172^\circ 56'\text{E}$) (Figures 9c and 9b). In this area, TASL is close to the dynamic height data, either from the assimilated profiles or from the unassimilated moorings, and the comparisons presented in the Appendix suggest that the difference with the tide gauge sea levels originates largely from salinity. In that respect, note that further east at Christmas ($1^\circ 59'\text{N}/157^\circ 29'\text{W}$) (Figure 9e), the time with largest differences is late 1982 when a very fresh surface water is

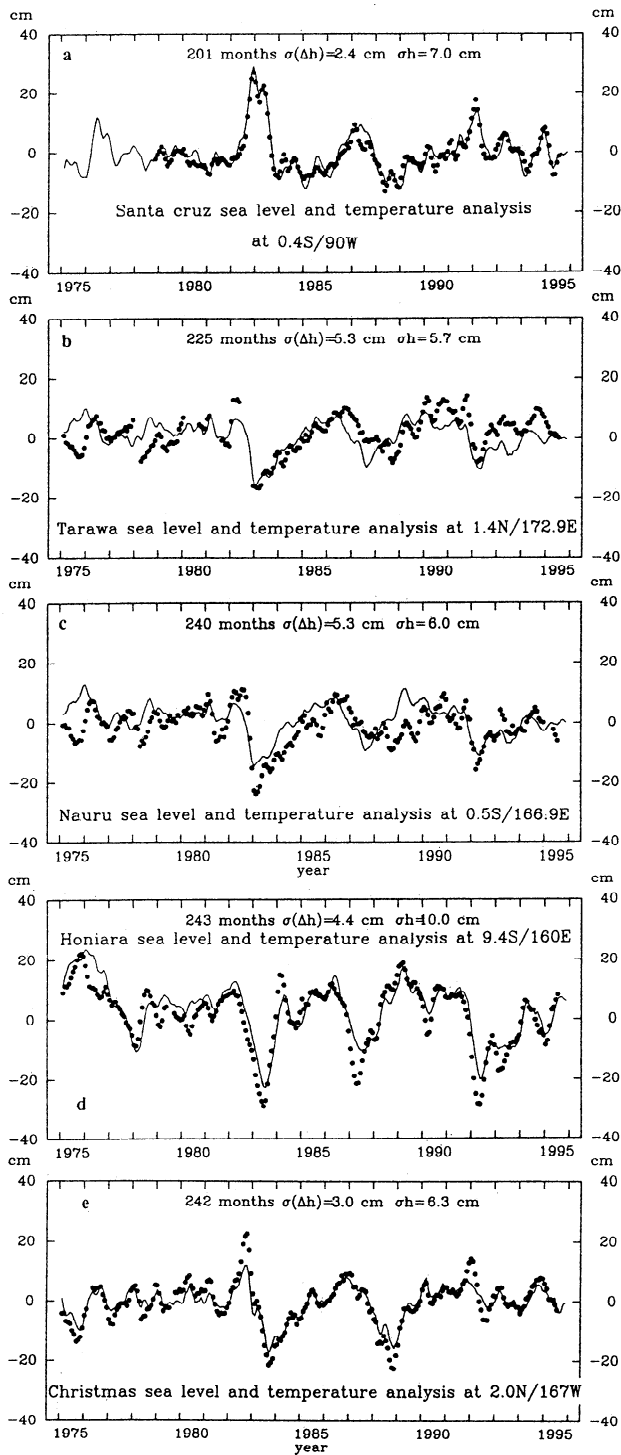


Figure 9. Comparison of the Kalman filtered analysis of the temperature sea level (solid line) to sea level from tide gauges with an inverse barometer correction of sea surface atmospheric pressure fluctuations (dots): (a) Santa Cruz; (b) Tarawa; (c) Nauru; (d) Honiara; (e) Christmas. The time series have been filtered by a 3-month running mean (rms differences between the two series $\{\sigma(\Delta h)\}$ as well as the rms variability in the joint time series $\{\sigma h\}$ are indicated for each plot).

known to have been present [see Kessler and Taft, 1987]. The same influence of surface salinity is likely to be true at Nauru and Tarawa. Despite this discrepancy resulting from the salinity effect, TASL is closer to the tide gauge sea level in the western

equatorial Pacific than the unfiltered run. Withholding profiles in this area does not improve the analysis elsewhere, which indicates that the data did not have a negative impact on the sea level analysis. This suggests little remote influence from the neglect of salinity variations. Most of the strong deviations of salinity from climatology are located very close to the sea surface. Linear model studies (Roemmich et al., 1994) of freshwater jets suggest that this near-surface trapped density anomaly resulting from surface salinity change would not project much on the two gravest baroclinic modes, the ones retained in our model.

Finally, the differences between TASL and sea level sometimes reveal trends. This is most obvious before 1985 at Yap (9.5°N/138°E) (Figure 10) but is also present to a lesser extent at Majuro and Papeete. At Yap, the trend in the tide gauge data is absent in the dynamic height from temperature data and is also absent in GASL with Yap withheld (Figure 3b of [Cane et al., 1996]). Tectonic motion is a likely candidate for explaining the differences at Yap.

Since the sea level data are used in the GASL analysis, it is not surprising that the comparisons of sea level records with TASL are not as good as with GASL. However, we often find (26 out of 34) that the comparisons between TASL and tide gauge records are better than for GASL, when the station compared was excluded from the GASL assimilation. Although in such comparisons TASL has the advantage of including data quite close to the comparison point, it nonetheless is fair to conclude that in many places the TASL sea level analysis is as good or better than GASL. The two analyses are much closer to each other than either is to the unfiltered run (Figure 11). The correlation coefficient between the two analyses is also very large, in particular near the equator, and at about 10° N near the west and east boundaries. Maximum rms difference and low correlation coefficient extends from 4°N in the western Pacific to

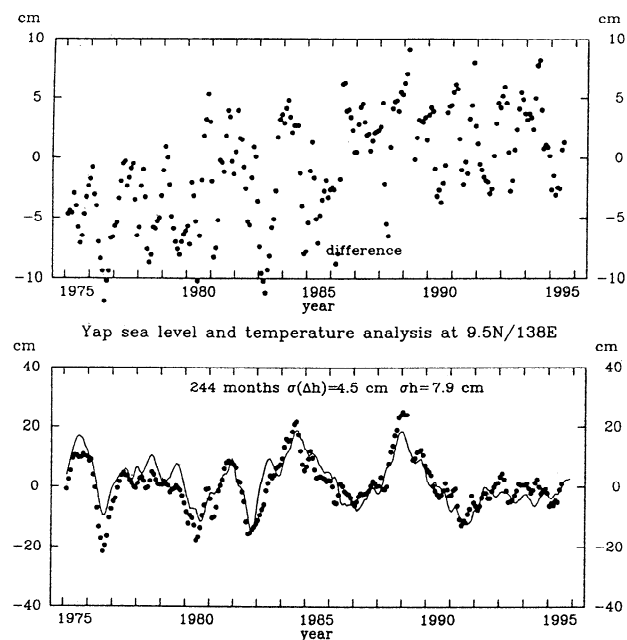


Figure 10. Comparison of the Kalman filtered analysis of the temperature sea level (solid line) to sea level from the Yap tide gauge with an inverse barometer correction of sea surface atmospheric pressure fluctuations (dots). The time series have been filtered by a 3-month running mean (rms differences between the two series as well as the rms variability in the joint time series are indicated for each plot).

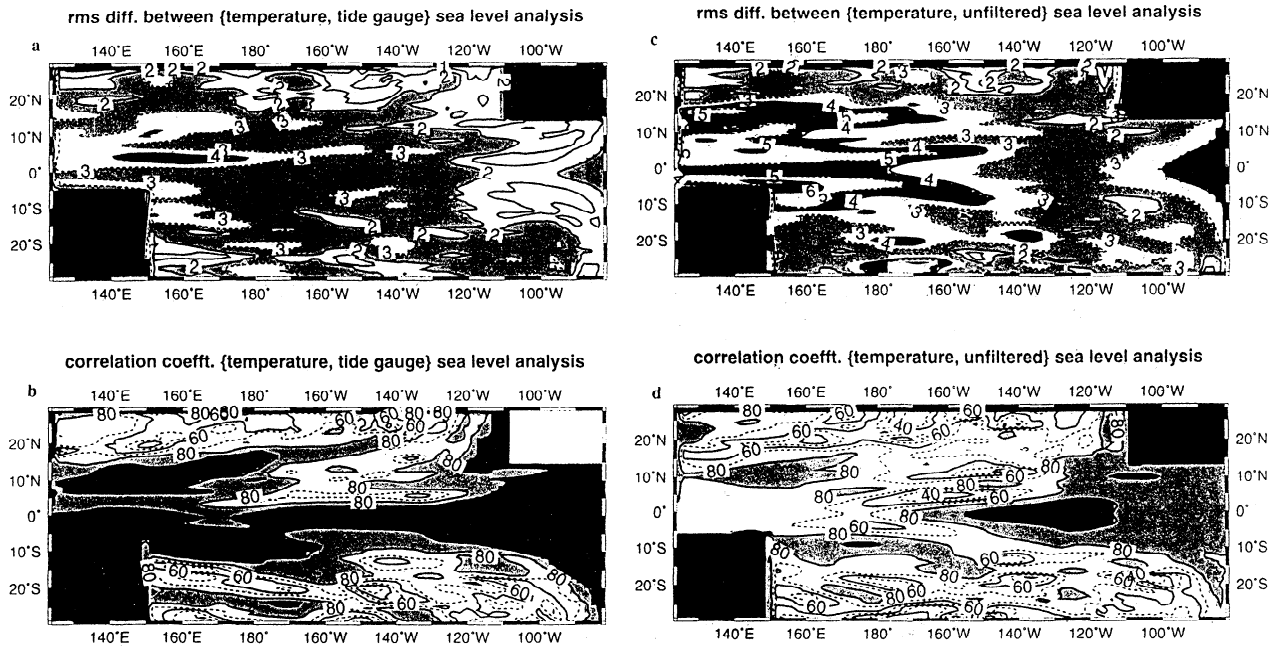


Figure 11. Maps of rms differences (centimeters) (light shading between 2 and 3 cm; dark shading between 4 and 5 cm) and correlation coefficients (percent) between various simulations in 1975 to 1992 (projection on 32 EOFs) (light shading between 80 and 90%; dark shading between 90 and 100%). The comparison between the unfiltered run and the Kalman filtered analysis of sea level from temperature profiles (TASL) is presented (Figures 11a and 11b) as well the comparison of the two Kalman filtered analysis of sea level (TASL and analyses of tidegauge data (GASL)) (Figures 11c and 11d).

6° N in the central Pacific. In the central Pacific, this is slightly north of the average ridge separating the South Equatorial Current from the North Equatorial Countercurrent [Taft and Kessler, 1991]. In this area, comparisons to the tide gauges do not clarify which field is a better representation of sea level. We suspect it is TASL, which has better data coverage there.

4. Conclusions

Sea level estimated from shallow temperature profiles (300 m) is a substitute for observed sea level data with an error of less

than 2 cm, except when there is large salinity variability unrelated to temperature variations, in particular in the western equatorial Pacific. In fact, we find that the analysis of 86,616 monthly binned data from temperature profiles (an average of 401 bins each month) does at least as good a job in reproducing observed variability as the analysis of 34 tide gauge sea level time series. The analysis is carried in a state of reduced dimension, defined by a limited set of orthogonal functions. Because we have used a Kalman filter with a model which has been shown to have some skill in reproducing low-frequency variability [Busalacchi and Cane, 1985; Cane, 1984], the Kalman filter should produce more

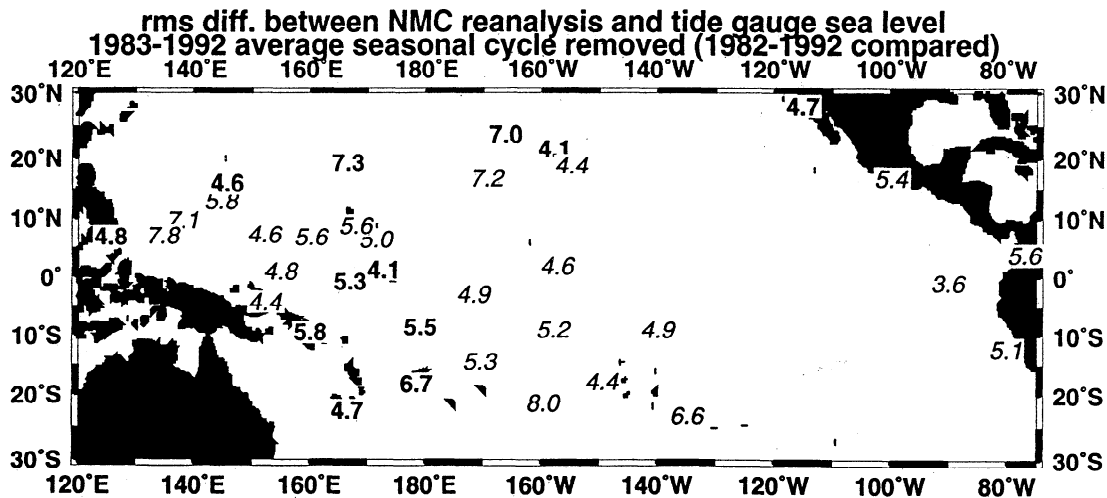


Figure 12. Rms differences (centimeters) between the NMC reanalysis (RA4) of the temperature profiles (1982-1992) with the tide gauge sea levels (in both cases, the seasonal cycle for 1982-1992 has been removed). The values are written in boldface when the rms difference is lower than in Figure 8 corresponding to the comparison between the Kalman filtered sea level from temperature profiles and the tide gauge sea level.

accurate analyzed fields than objective function fitting [e.g., *Bretherton et al.*, 1984]. Of course, this assumes that the estimate of model error is reasonable. This was discussed in detail by *Cane et al.* [1996] and is a difficult issue to tackle. We find as given by *Miller et al.* [1995] that the Kalman filter is not optimal, as the innovation time series (differences between forecast and analysis) is not white. The situation is worse for a more recent data assimilation with this model extending to 1995, for which the detrended winds forcing the model differed systematically from the observed winds after 1991, in particular in the central Pacific, resulting in a larger zonal slope of the thermocline in the unfiltered model than observed.

We found that the fields of both the tide gauge sea level analysis and the analysis presented here systematically underestimate variability between 5°S and 15°S in the western Pacific, despite the presence of numerous data. We verified that these spatial structures could have been reproduced by the set of EOFs used in the analysis. This discrepancy suggests to us that the model contains systematic errors in this part of the basin which are not well represented in the error model. The systematic errors could arise because of the specification of the models boundaries, the simplification of the dynamics, or the wind forcing applied. We believe this last hypothesis is likely to be relevant in view of the sparsity of wind observations in this region (this is discussed further by *Miller et al.* [1995]). It is also useful to consider the monthly sea level fields resulting from the assimilation of temperature profiles between July 1982 and 1992 in the NMC three-dimensional primitive model of the ocean [*Ji et al.*, 1995; *Enfield and Harris*, 1995]. The initial reanalysis presented by *Ji et al.* [1995] uses a completely different wind forcing and a different variational assimilation scheme [*Derber and Rosati*, 1993]. It presented large differences to the observations, some of which were attributed to the wind forcing. Later reanalyses RA3 and RA4 use a wind forcing close to the one we use and present a closer agreement to observations [*Enfield and Harris*, 1995]. We accessed RA4 in May 1995. In RA4, observations are analyzed with a radius of influence which is isotropic in latitude and longitude, and the analyzed fields present more synoptic scale structure than TASL (Figure 4). The observations assimilated in RA4 include the profiles we used but also the temperature mooring data. We most commonly find (23 out of 35) that the sea level in this reanalysis is not as close to the sea level from tide gauges as in our analysis (Figure 12). However, RA4 is closer to the sea level records in the southwest Pacific, primarily because the amplitude of the low sea level episodes during El Niños is more correctly represented. This difference is strong enough to influence the EOF decomposition of the sea level fields. The second EOF in the NMC reanalysis (seasonal cycle removed) has a large maximum near 8°S, which is not present in the second EOF of TASL (not shown). This suggests that the defect of our temperature analysis in the southwest Pacific can be corrected, possibly by changing the error model formulation. Part of the improvement of RA4 relative to TASL can also be related to the extra data assimilated, in particular near (0°/165°E) (this was tested in a new assimilation in which the mooring data are also assimilated).

As given by *Cane et al.* [1996] we have found that retaining a relatively small number of degrees of freedom in the Kalman filter provided a decent data assimilation with errors less than 3 cm rms near the equator. It is notable that this study involves a great deal more data. In fact, it is difficult to envision a large-scale oceanographic data assimilation problem with appreciably better space-time data coverage. Since it would be

computationally feasible to increase the size of the covariance matrix 1 or 2 orders of magnitude beyond the 32 used here, the implication is that the reduced state space version of the Kalman Filter is feasible for most (if not all) ocean data assimilation problems.

As noted by *Cane et al.* [1996], the conclusion that such a small number of EOFs is adequate may well be contingent on the flaws in our error models. That is, with a better description of both model and data errors, it is conceivable that additional information could be extracted from the observational data, requiring additional structure in the Kalman Filter error covariances. Whether this turns out to be true or not, one must conclude that obtaining better representations for the errors is the primary issue for the further development of the method.

Appendix: Validation of Dynamic Height From Temperature Records

We have presented in section 2 how dynamic height was estimated from temperature profiles extending at least to 250 m and how binned monthly sea level estimates (TSL) were computed with their error estimates. Figure A1 shows the rms differences and correlation between these noisy estimates and tide gauge records of sea level. TSL is remarkably close to SL near the Galapagos Islands and at many sites 5° to 10° off the equator in the central and western Pacific. Even for Honolulu (23° N), it is very close to the tide gauge sea level (though it is less so at Hilo on Hawaii Island, due to a poorer sampling as well as local circulation). Part of the structure is related to differences in sampling frequency at the different sites; at many sites the rms differences strongly decrease when some time smoothing is applied. However, the high rms difference near the Philippines probably results from the profiles not being close enough to the tide gauge site, aliasing the large steric slope across the intense western boundary currents. In the western Pacific, the lower correlation between 170°W and 150°E near the equator is caused by the presence of a large salinity variability which contributes significantly to sea level fluctuations.

To a large extent, the differences between TSL and the tide gauge records are coherent with the estimated errors on TSL, and we will present comparisons at a few sites to illustrate this point. On Figure A2, the line is for the tide gauge record, and the bars correspond to the estimated ± 1 standard deviation on TSL. At Santa Cruz/Baltra (Figure A2a), the correlation coefficient between TSL and tide gauge sea level is particularly large (0.96). The errors on the tide gauge record are small, as indicated by the comparison of the two sites Santa Cruz and Baltra, with rms differences less than 1 cm when the two records overlap. At this site, the statistics of the differences between TSL and the tide gauge sea level indicate that we have overestimated the rms error on TSL by close to 30%. The small set of Nansen and CTD casts in the vicinity suggests that the error resulting from the approximation on salinity is smaller than the spatially homogeneous 3 cm rms error we have chosen. The situation at Honolulu is different (Figure A2b) with error estimates which have the right magnitude. We have applied a 3-month running mean on the data to reduce high-frequency noise and random errors). A small part of the difference is also attributed to errors on the tide gauge record, with the rms differences between monthly sea level records from the two sides of the island Oahu being of the order of 1.75 cm [*Wyrki et al.*, 1988]. The comparison at Christmas (Figure A2c) illustrates a case where the rms error estimate is too small, possibly by 20%. (We again apply

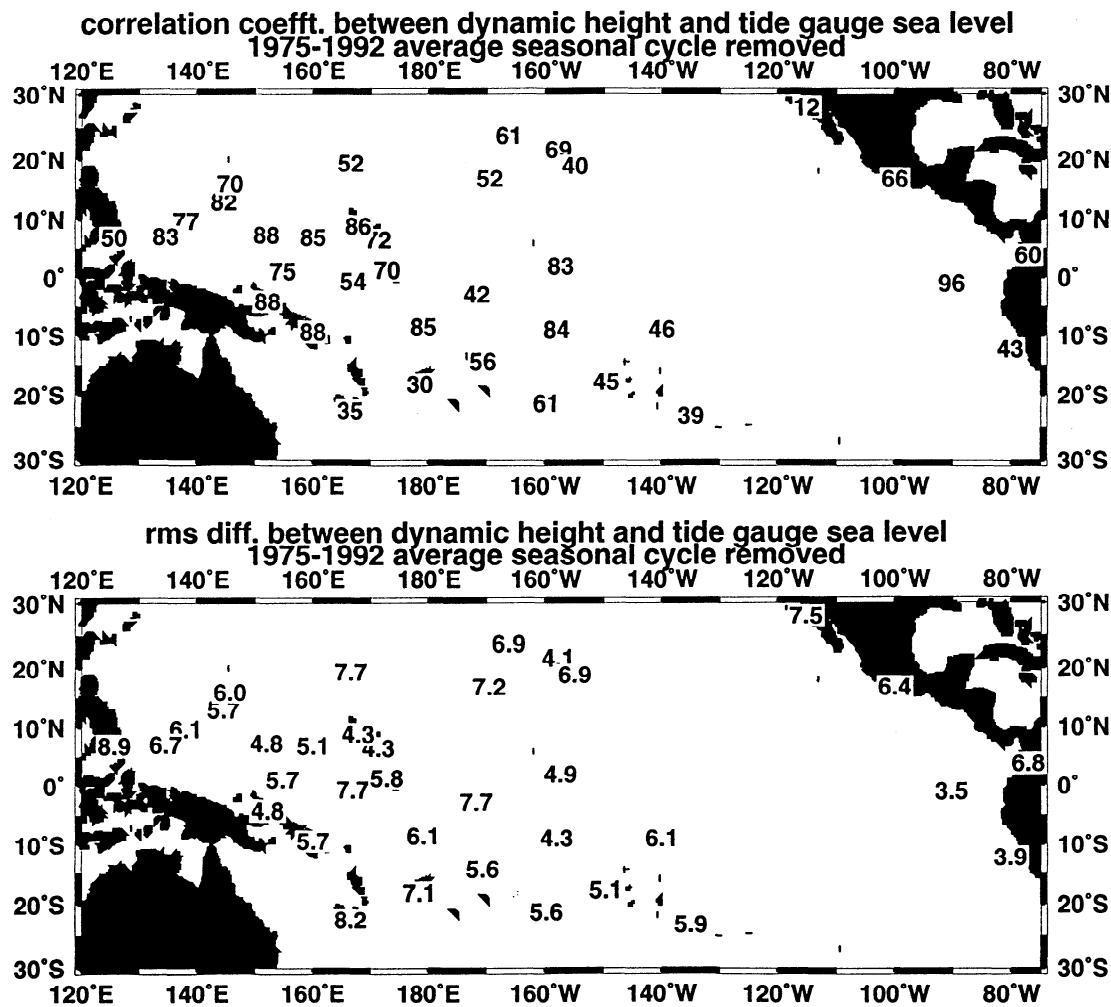


Figure A1. Comparison (correlation coefficient (percent) and rms differences (centimeters)) of TSL with tide gauge sea level records in 1975-1992. TSL is from the time series of gridded sea level from temperature profiles.

a 3-month running mean on the data). It has been suggested [Kessler and Taft, 1987] that part of the differences there originates from salinity variability. At that site, we have probably underestimated the error resulting from its neglect.

Salinity is expected to vary even more near the equator west of the date line [Delcroix and Masia, 1989; Delcroix and Hénin, 1991; Sprintall and McPhaden, 1994]. Note that in this area the comparison presented on Figure 3 is particularly poor. A good illustration of the issue is provided by the set of cruises along 165°E (27 between 1984 and early 1991) which were used in the study of Gouriou and Toole [1993] and which to a large extent are not included in the set of data analyzed. We compare these data to the tide gauges at Nauru (0.5°S/167°E), Kapingamarangi (1.1°N/155°E), and Tarawa (1.4°N/173°E). Actually, we find the agreement to be closer when instead of considering each of these latter two stations separately, we consider the linear interpolation in longitude of the two records to 1.2°N/165°E. The comparisons are done with the average of the hydrographic stations within 0.5° latitude of the site. Clearly, near Nauru (Figures A3a and A3b) the inclusion of salinity when estimating the dynamic height referred to 300 m greatly improves the agreement with the monthly tide gauge record: rms differences decrease from 6.2 to 4.6 cm. In particular, without S, the dynamic heights would be

too low in late 1987/early 1988 and too high in 1989. At this site, a large part of those differences is related to the large surface salinity change (Figure A3c). No significant improvement is found at the Kapingamarangi/Tarawa latitude (with rms in both cases close to 4.0 cm). The result at Nauru is supported by a comparison between the cruise data and the dynamic heights estimated from the daily mooring temperature profiles at 0/165°E. In this instance, the agreement with the cruise dynamic height is much closer when no salinity is included (2.1 cm rms differences) than when it is included (4.4 cm). Similar results are obtained for the comparisons with the moorings at 2°N and 2°S but are less clear at 5°N and 5°S, where interannual variability of S is much lower according to this data set. These conclusions are also supported by the short records of temperature and salinity of the 0/165°E mooring.

We have not taken into account this complicated spatial structure of surface salinity variability in our estimation of error, instead taking an average 3-cm error based on a rather limited set of comparisons of the estimated dynamic height from Nansen and CTD casts. The error of relating sea level variability to dynamic height referred to a level as shallow as 300 m (and even 250 m in some cases) should also vary spatially. We tested that along 165°E by using various reference levels. However, we found that

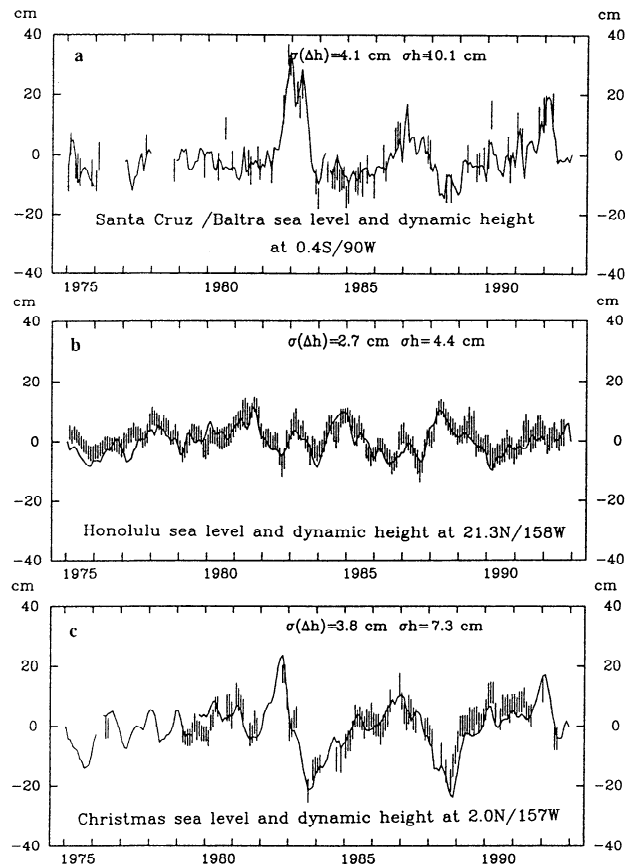


Figure A2. Comparison for a few stations of the tide gauge sea level (solid line) to sea level from temperature profiles with their error estimates (± 1 standard deviation of the error estimate). Time series in Figures A2a and A2b have been filtered by a 3-month running mean. Rms differences between the two series $\{\sigma(\Delta h)\}$ as well as the rms variability in the joint time series $\{\sigma h\}$ are indicated for each panel.

dynamic height referred to a deeper reference like 1000 decibars was a much more noisy estimate than our sea level estimated from the shallower reference, so that one could not conclude from the limited comparisons whether systematic differences in variability with the tide gauge could arise from the choice of reference level. Overall, our estimated errors might be overestimated in some areas but seem to be of roughly the right magnitude.

Acknowledgments. This research was supported by the grants from NOAA (NA16RC04320 and NA47GP0188), NASA (JPL 958123), and ONR (N00014-94-1-0134). Eric Greiner kindly read this manuscript, and Bob Miller and Tony Busalacchi provided useful advice during the course of the project. This is Lamont-Doherty contribution 5529.

References

- Bretherton, F., M.J. McPhaden, and E. Kraus, Design studies for climatological measurement of heat storage, *J. Phys. Oceanogr.*, **14**, 318–337, 1984.
- Busalacchi, A.J., and M.A. Cane, Hindcasts of sea level variations during 1982/83 El Niño, *J. Phys. Oceanogr.*, **15**, 213–221, 1985.
- Busalacchi, A.J., M.J. McPhaden, and J. Picaut, Variability in equatorial Pacific sea surface topography during the verification phase of the TOPEX/POSEIDON mission, *J. Geophys. Res.*, **99**, 24,725–24,738, 1994.

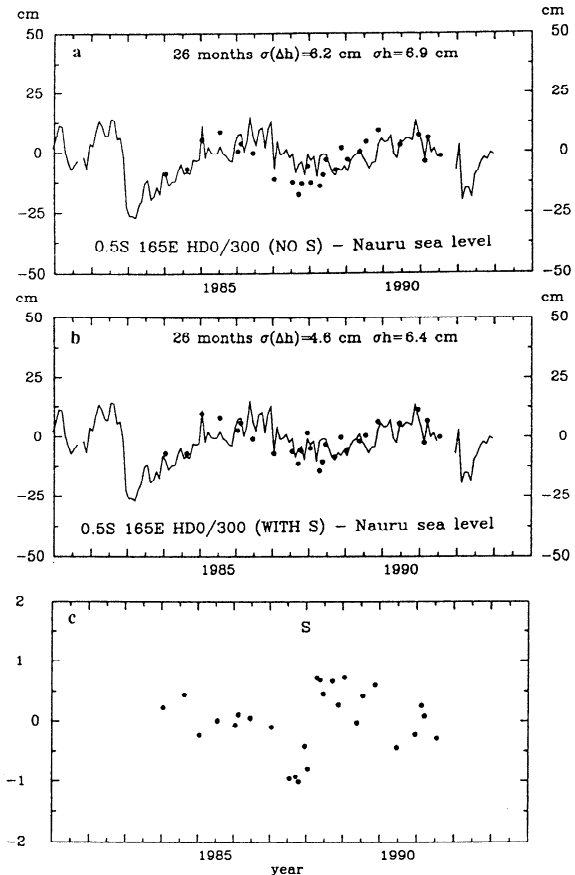


Figure A3. Comparison of cruise dynamic heights near $0.5^{\circ}\text{S}/165^{\circ}\text{E}$ with the Nauru tide gauge record: (a) dynamic height estimated from the 0- to 300-m temperature profiles; (b) dynamic height referred to 300 db estimated directly from the T and S of the stations; (c) salinity variability at this site. Rms differences between the two series $\{\sigma(\Delta h)\}$ as well as the rms variability in the joint time series $\{\sigma h\}$ are indicated for the upper panel.

- Cane, M.A., Modeling sea level during El Niño, *Nature*, **321**, 827–832, 1984.
- Cane, M.A., and R.J. Patton, A numerical model for low-frequency equatorial dynamics, *J. Phys. Oceanogr.*, **14**, 1853–1863, 1984.
- Cane, M.A., S.E. Zebiak, and S.C. Dolan, Experimental forecasts of El Niño, *Nature*, **321**, 827–832, 1986.
- Cane, M.A., A. Kaplan, R.N. Miller, B. Tang, E.C. Hackert, and A.J. Busalacchi, Mapping tropical Pacific sea level: data assimilation via a reduced state space Kalman filter, *J. Geophys. Res.*, in press, 1996.
- Delcroix, T., and C. Gautier, Estimates of heat content variations from sea level measurements in the central and western tropical Pacific from 1979 to 1985, *J. Phys. Oceanogr.*, **17**, 725–734, 1987.
- Delcroix, T., and C. Hénin, Seasonal and interannual variations of sea surface salinity in the tropical Pacific Ocean, *J. Geophys. Res.*, **96**, 22,135–22,150, 1991.
- Delcroix, T., and F. Masia, Atlas des variations de température et de salinité de surface du Pacifique tropical (1969–1988), *Rapp. Sci. Tech., Sér. sci. Mer*, **2**, 151pp., Cent. ORSTOM de Nouméa, New Caledonia, 1989.
- Derber, J.D., and A. Rosati, A global oceanic data assimilation system, *J. Phys. Oceanogr.*, **19**, 1333–1347, 1993.
- Donguy, J.R., Surface and subsurface salinity in the tropical Pacific Ocean. Relations with climate, *Prog. Oceanogr.*, **34**, 47–78, 1994.
- Enfield, D.B., and J.E. Harris, A comparative study of tropical Pacific sea surface height variability: Tide gauges versus the National Meteorological Center data-assimilating ocean general circulation model, 1982–1992, *J. Geophys. Res.*, **100**, 8661–8675, 1995.

- Goldenberg, S., and J.J. O'Brien, Time and space variability of tropical wind stress, *Mon. Weather Rev.*, *109*, 1190-1205, 1981.
- Gouriou, Y., and J. Toole, Mean circulation of the upper layers of the western equatorial Pacific Ocean, *J. Geophys. Res.*, *98*, 22,495-22,520, 1993.
- Hanawa, K., P. Rual, R. Bailey, A. Sy, and M. Szabados, Calculation of new depth equations for expandable bathythermographs using a temperature error-free method (application to Sippican /TSK T-7, T-6, and T-4 XBTs), *UNESCO Tech. Pap. Mar. Sci.*, *67*, 46 pp., UNESCO, Paris, 1994.
- Hayes, S.P., P. Ripa, and L.J. Mangum, On resolving vertical modes with observational data, *J. Geophys. Res.*, *90*, 7227-7234, 1985.
- Hayes, S.P., L.J. Mangum, J. Picaut, A. Sumi, and K. Takeuchi, TOGA-TAO: A moored array for real-time measurements in the tropical Pacific Ocean, *Bull. Am. Meteorol. Soc.*, *72*, 339-347, 1991.
- Ji, M., A. Leetmaa, and J. Derber, An ocean analysis system for seasonal to interannual climate studies, *Mon. Weather Rev.*, *123*, 460-481, 1995.
- Kessler, W.S., Observations of long Rossby waves in the northern tropical Pacific, *Tech. Memo. ERL PMEL-86*, 169 pp., Pac. Mar. Environ. Lab., Seattle, Wash., 1989.
- Kessler, W.S., and B.A. Taft, Dynamic heights and zonal geostrophic transports in the central tropical Pacific during 1979-1984, *J. Phys. Oceanogr.*, *17*, 97-122, 1987.
- Lukas, R., and E. Firing, The annual Rossby wave in the central equatorial Pacific Ocean, *J. Phys. Oceanogr.*, *15*, 55-67, 1985.
- Luther, D., Evidence of a 4-6 day barotropic, planetary oscillation of the Pacific Ocean, *J. Phys. Oceanogr.*, *12*, 644-657, 1982.
- McPhaden, M.J., TOGA-TAO and the 1991-1993 El Niño-Southern Oscillation event, *Oceanography*, *6*, 36-44, 1993.
- Meyers, G., H. Phillips, and J. Sprintall, Space and time scales for optimal interpolation of temperature - Tropical Pacific Ocean, *Prog. Oceanogr.*, *28*, 189-218, 1991.
- Miller, R.N., and M.A. Cane, A Kalman filter analysis of sea level heights in the tropical Pacific, *J. Phys. Oceanogr.*, *19*, 773-790, 1989.
- Miller, R.N., A.J. Busalacchi, and E.C. Hackert, Sea surface topography fields of the tropical Pacific from data assimilation, *J. Geophys. Res.*, *100*, 13,389-13,425, 1995.
- Mitchum, G. T., Comparison of TOPEX sea surface heights and tide gauge sea levels, *J. Geophys. Res.*, *99*, 24,541-24,553, 1994.
- Rebert, J.P., J.R. Donguy, and G. Eldin, 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.
- Reverdin, G., C. Frankignoul, E. Kestenare, and M.J. McPhaden, Seasonal variability in the surface currents of the equatorial Pacific, *J. Geophys. Res.*, *99*, 20,323-20,344, 1994.
- Roemmich, D., M. Morris, W.R. Young, and J.R. Donguy, Fresh equatorial jets, *J. Phys. Oceanogr.*, *24*, 540-558, 1994.
- Rosati, A., K. Miyakoda, and R. Gudgel, The impact of ocean initial conditions on ENSO forecasting with a coupled model, *Mon. Weather Rev.*, in press, 1996.
- Sprintall, J., and M.J. McPhaden, Surface layer variations observed in multi year time series measurements from the western equatorial Pacific, *J. Geophys. Res.*, *99*, 963-979, 1994.
- Taft, B.A., and W.S. Kessler, Variations of zonal currents in the central Pacific during 1970 to 1987: Sea level and dynamic height measurements, *J. Geophys. Res.*, *96*, 12,599-12,618, 1991.
- Tang, T.Y., R.M. Weisberg, and D. Halpern, Vertical structure of low-frequency variability in the eastern equatorial Pacific Ocean, *J. Phys. Oceanogr.*, *18*, 1009-1019, 1988.
- Tsuchiya, M., R. Lukas, R.A. Fine, E. Firing, and E. Lindstrom, Source waters of the Pacific equatorial Undercurrent, *Prog. Oceanogr.*, *23*, 101-147, 1989.
- Wyrtki, K., Monitoring the strength of equatorial currents from XBT sections and sea level, *J. Geophys. Res.*, *83*, 1935-1940, 1978.
- Wyrtki, K., B.J. Kilonsky, and S. Nakahara, The IGOSS sea level pilot project in the Pacific, *JIMAR Data Rep. 3*, 59 pp., Joint Inst. for Mar. and Atmos. Res., Univ. of Hawaii, Honolulu, 1988.

M. Cane and A. Kaplan, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964.

G. Reverdin, Groupe de Recherche de Géodésie Spatiale, UMR CNES/CNRS 5566, 18 Avenue Edouard Belin, 31401 Toulouse, France. (e-mail: reverdin@pontos.cst.cnes.fr)

(Received August 4, 1995; revised April 15, 1996; accepted May 13, 1996.)