Toward the Use of Altimetry for Operational Seasonal Forecasting

J. SEGSCHNEIDER, D. L. T. ANDERSON, AND T. N. STOCKDALE

European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, United Kingdom

(Manuscript received 11 August 1999, in final form 16 November 1999)

ABSTRACT

The TOPEX/Poseidon and *ERS-1/2* satellites have now been observing sea level anomalies for a continuous time span of more than 6 yr. These sea level observations are first compared with tide gauge data and then assimilated into an ocean model that is used to initialize coupled ocean–atmosphere forecasts with a lead time of 6 months. Ocean analyses in which altimeter data are assimilated are compared with those from a no-assimilation experiment and with analyses in which subsurface temperature observations are assimilated. Analyses with altimeter data show variations of upper-ocean heat content similar to analyses using subsurface observations, whereas the ocean model has large errors when no data are assimilated. However, obtaining good results from the assimilation of altimeter data is not straightforward: it is essential to add a good mean sea level to the observed anomalies, to filter the sea level observations appropriately, to start the analyses from realistic initial temperature and salinity fields, and to assign appropriate weights for the analyzed increments.

To assess the impact of altimeter data assimilation on the coupled system, ensemble hindcasts are initialized from ocean analyses in which either no data, subsurface temperatures, or sea level observations were assimilated. For each kind of ocean analysis, a five-member ensemble is started every 3 months from January 1993 to October 1997, adding up to 100 forecasts for each type. The predicted SST anomalies for the equatorial Pacific are intercompared between the experiments and against observations. The predicted anomalies are on average closer to observed values when forecasts are initialized from the ocean analysis using altimeter data than when initialized from the no-assimilation ocean analysis, and forecast errors appear to be only slightly larger than for forecasts initialized from ocean analyses using subsurface temperatures. However, even based on 100 coupled forecasts, the distinction between the two experiments that benefit from data assimilation is barely statistically significant. The verification should still be considered preliminary, because the period covered by the forecasts is only 5 yr, which is too short properly to sample ENSO variability. It is, nonetheless, encouraging that altimeter assimilation can improve the forecast skill to a level comparable to that obtained from using Tropical Ocean Atmosphere–expendable bathythermograph data.

1. Introduction

Coupled ocean–atmosphere interactions in the Tropics are a significant and partly predictable source of global interannual climate variability (Stockdale et al. 1998). Recent attempts to forecast the 1997–98 major El Niño event imply that the key to successful forecasting of equatorial sea surface temperature (SST) is the information contained in the subsurface density structure of the tropical oceans, in particular in the Pacific. Variations in the depth of the thermocline that change the upper-ocean heat content in the west Pacific as part of the El Niño–Southern Oscillation (ENSO) cycle can propagate eastward at subsurface levels along the equatorial waveguide. Because of the shallowing of the thermocline in the eastern Pacific, the subsurface anomalies in the west Pacific can effect SST anomalies in the eastern equatorial Pacific with a time delay of several months. For seasonal climate forecasts with lead times of up to 6 months, it is therefore important to initialize the ocean–atmosphere model not only with SSTs, but also with a realistic oceanic subsurface density structure, in particular in the Tropics.

So far the most common technique to obtain oceanic initial conditions is to force an ocean model with the recent history of surface fluxes. Surface stress fields are of particular importance but not yet of sufficient accuracy, and models have systematic errors. A common strategy to correct for forcing and model errors, also used at the European Centre for Medium-Range Weather Forecasts (ECMWF) for the present quasi-operational ocean analysis, is to assimilate in situ temperature data (Smith et al. 1991; Ji et al. 1998). In situ data, however, exist only for a limited number of points. Relatively dense observations are provided by the Tropical Atmosphere Ocean (TAO) buoy array in the equatorial Pacific, but in the Atlantic an ocean observation system is only now evolving, and few data are obtainable in

Corresponding author address: Dr. Joachim Segschneider, European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading RG2 9AX, United Kingdom.
E-mail: j.segschneider@ecmwf.int

real time for the Indian Ocean. Altimetry has the potential to improve oceanic initial conditions because it can provide information on variations in upper-ocean heat content (e.g., Chambers et al. 1998) with a data coverage in space and time that is way beyond anything obtainable from conventional observing systems. The difficulty lies in the projection of the surface signal onto the subsurface density structure.

Numerous papers have been published in the fast-evolving field of assimilation of sea level into ocean models (e.g., Stammer and Wunsch 1996; Oschlies and Willebrand 1996; Cooper and Haines 1996; Weaver and Anderson 1996; Fischer et al. 1997; Ji et al. 1998; and references therein). Much of the work, however, was done in a twin experiment environment or aimed at improvement of simulated ocean dynamics in higher latitudes. Here the main goal is to provide improved oceanic initial conditions for a global coupled ocean–atmosphere model that is used to make seasonal forecasts. Altimetry is used to correct for displacements of the thermocline, which are caused by model error or by errors in the wind forcing. The improved representation of the ocean model's upper-ocean heat content should then lead to improved forecasts made by the coupled system. The assimilation of altimeter-observed sea level anomalies in that context has been attempted by Carton et al. (1996) and Ji et al. (2000), but either no (Carton et al. 1996) or only a few forecasts (Ji et al. 1999) were started from the ocean analyses. In both of the above studies, altimeter data were assimilated together with temperature observations into limited-area models of the tropical Pacific, whereas here, either sea level *or* subsurface temperature data are assimilated into a global ocean model.

Chen et al. (1998) found that the assimilation of sea level from tide gauges improved forecasts of the 1997–98 El Niño in the Lamont prediction model. Using a more complex forecast system that consisted of a tropical Pacific Ocean model coupled to a statistical atmosphere, Fischer et al. (1997) showed that assimilation of sea level anomalies had a positive impact on forecast skill of Niño-3 SST anomalies over certain periods. However, the study was done using *pseudo* sea level data obtained from an ocean analysis at the National Meteorological Center (now known as the National Centers for Environmental Prediction) rather than from satellite. Further, only anomalies relative to the ocean model's seasonal cycle were assimilated, thus putting a relatively weak constraint on the model state and preventing corrections to the seasonal cycle of the model. Salinity was completely neglected, on the grounds that the density of seawater in the Tropics is to first order determined by its temperature. More recent studies, however, indicate that western Pacific salinity may be of importance for the genesis cycle of ENSO events (Vialard and Delecluse 1998), and interannual salinity variations can account for as much as 80 mm of the satellite-observed sea level anomaly in the western Pacific (Vossepoel et al. 1999). The ocean analysis system that is used at ECMWF to initialize seasonal forecasts consists of a global ocean model with a free surface, which carries salinity as a prognostic tracer and so should overcome some of the above shortcomings. Alves et al. (2000) assimilated *pseudo* sea level data into the ECMWF ocean model in a twin-experiment environment, and found that erroneous displacements of the thermocline could largely be corrected, but no forecasts were started from the ocean analyses. By replacing the *pseudo* sea level data with the sea level *observations* from TOPEX/Poseidon (T/P) and *ERS-1/2* and by using the coupled seasonal forecast system at ECMWF, a next step is possible. The impacts of assimilation of sea level observations observed by satellite on seasonal forecasts can be investigated in the context of a sophisticated global real-time coupled ocean–atmosphere forecast system.

However, new problems arise when assimilating observed sea level that were not addressed in Alves et al. A mean sea level has to be added to the observed anomalies, the representativeness of the data for the specific problem has to be assessed, the accuracy of the observations has to be estimated, and suitable observations have to be obtained against which the success of the assimilation can be measured. The forecasts have to span as many years as possible and therefore the high-quality sea level observations, which are available only several months behind real time, have to be combined with the fast delivery near-real-time data. Last, to estimate the impact of the assimilation on the forecast skill, a set of ensemble forecasts is needed that has several members for each forecast date and suitable methods have to be used for the evaluation of the forecasts. All these points will be addressed in this paper.

Although the analysis is global, and some results will be shown for the Atlantic and Indian Oceans, the focus of the study is on the equatorial Pacific mainly for two reasons. First, the signal-to-noise ratio of interannual SST variations is largest in that area; second, subsurface temperature observations are relatively dense in the tropical Pacific and are already improving the current ocean analysis at ECMWF. If it can be shown that the assimilation of altimeter observations works well here, then we may reasonably assume that it improves the ocean analyses in areas where only few subsurface temperature observations exist, such as in the tropical Indian and Atlantic Oceans.

The paper is organized as follows: in section 2 the observational basis on which this work is based is described and altimeter sea level is compared to tide gauge data. In section 3 the ocean model is briefly introduced and the assimilation procedure is described, followed by a discussion of the ocean-only experiments in section 4. Results of the ocean analyses are compared with observations in section 5. In section 6 the coupled forecasts are discussed and an evaluation of the use of altimetry is summarized in section 7.

2. Database for assimilation and verification

a. Sea level observations from satellite

The sea level anomalies (SLA) used in this study were computed from geophysical data records of TOPEX/Poseidon as well as from *ERS-1/2* ocean products at Collecte Localisation Satellite, Toulouse (CLS; Le Traon et al. 1995, 1998). To avoid confusion about the term ''anomalies'' in the context of altimetry, note that here sea level anomalies are not relative to an averaged seasonal cycle but are relative to the 3-yr mean over the years 1993–95, which is computed at each observation point by a repeat track analysis. The reason that sea level cannot be given as absolute value is that the geoid of the earth is still insufficiently well known, though efforts are being made to derive an accurate geoid (Tapley et al. 1996).

Several datasets of SLA are available from CLS. By combining measurements from T/P and *ERS-1/2*, a ''homogeneous-historical'' (HH) dataset is created at CLS with an estimated rms error of only ± 30 mm. Data are quality controlled and corrections are applied for tidal effects, the inverse barometer effect and long wavelength error as well as dry and wet troposphere error. The corrected sea level anomalies are available along the satellites' ground tracks, but also are further analyzed at CLS to produce gridded maps of sea level anomaly together with maps of the combined interpolation plus instrument error. The horizontal resolution of the sea level maps is $0.25° \times 0.25°$. The maps are available every 10 days from 22 October 1992, but only several months behind real time. Maps of so-called near-real-time (NRT) data, which are only 7 days behind real time, have been produced weekly since 14 January 1998.

The maps for both HH and NRT are either produced from T/P only, extending from 63°S to 63°N, or from a combination of T/P and *ERS-1/2* extending from 82°S to 82°N. The ground track of *ERS-1* did not repeat for cycles during January 1994 to March 1995; however, as the satellite was used for ice monitoring (January–March 1994) and geodetic phases (April 1994–March 1995). A repeat track analysis was not possible over that period, resulting in a long data gap for the combined T/P–ERS data. These data gaps led us to use the T/P-only product for the period January 1993–April 1998 for two reasons. First, the forecast system at ECMWF requires forecasts over as long a period as possible to allow an accurate estimate of the drift of the coupled model. Second, the forecasts need to span as long a period as possible in order to capture as many extreme states of the ENSO cycle as possible. We use the combined T/P–*ERS-2* product in a NRT analysis from May 1998 on. Time series of area-averaged sea level in the tropical Pacific (not shown) indicate that the transition between the two chosen datasets is quite smooth. A more detailed estimation of the relative accuracies of HH and NRT data is currently under way at CLS (C. Boone 1999, personal communication).

TABLE 1. Selected stations for comparison of altimeter- and tide gauge–observed sea level anomaly. WEP = western equatorial Pacific, CEP = central equatorial Pacific, EEP = eastern equatorial Pacific, SP = southern Pacific, NP = northern Pacific, I = Indian, and A = Atlantic.

Station	Latitude	Longitude	Location
Betio	1.2°N	172.5°E	WEP
Christmas Is.	1.6°N	157.3°W	CEP
Galapagos	0.3°S	90.3°W	EEP
Funafuti	8.3°S	179.1°E	SP
Johnston Is.	16.5°N	169.3°W	NP
Seychelles	4.4°S	55.3°E	I
Bermuda	32.2°N	64.4°W	A

b. Possible means of quality control

The assimilation of observations generally requires a quality control procedure to identify and flag data that are either erroneous or incompatible with the model. The altimeter observations that we use are already quality controlled at CLS and a combined mapping and instrument error is provided with the altimeter maps for each grid point. This error, however, represents to first order the distance of the respective grid point from the satellite's ground track. As pointed out later (section 3), we smooth the gridded SLAs and interpolate the data to the model grid before assimilation. Therefore we do not use the error maps provided. As an alternative, we test the suitability of tide gauge data for quality control purposes by comparing the altimeter SLA with the independent SLA from tide gauges.

We obtain daily values of SLA observed by tide gauges from the fast delivery web site provided by the University of Hawaii Sea Level Center (http://uhslc.soest.hawaii.edu/uhslc/woce.html). Currently, data on that site are available approximately two months after real time. The daily data are then averaged over 10-day intervals, which are centered at dates for which HH altimeter maps are available. Like the altimeter data, tide gauges measure sea level relative to a long-term mean, which is individual to each station. Therefore, to allow comparison with altimeter data, the mean sea level for 1993–95 is removed from each tide gauge time series separately.

Shallow water, enclosed bays, or even long-term motions of the seafloor at the location of the tide gauge are known causes that may contaminate the large-scale sea level signal (Mitchum 1994). To rule out locally dominated tide gauge stations as far as possible, we chose the subset of tide gauge stations listed in Table 1. These stations seem to best represent undisturbed open-ocean conditions and cover all ocean basins. One known potential source of deviations that remains is the 60-day tide, which is not removed from the altimeter observations (Mitchum 1994).

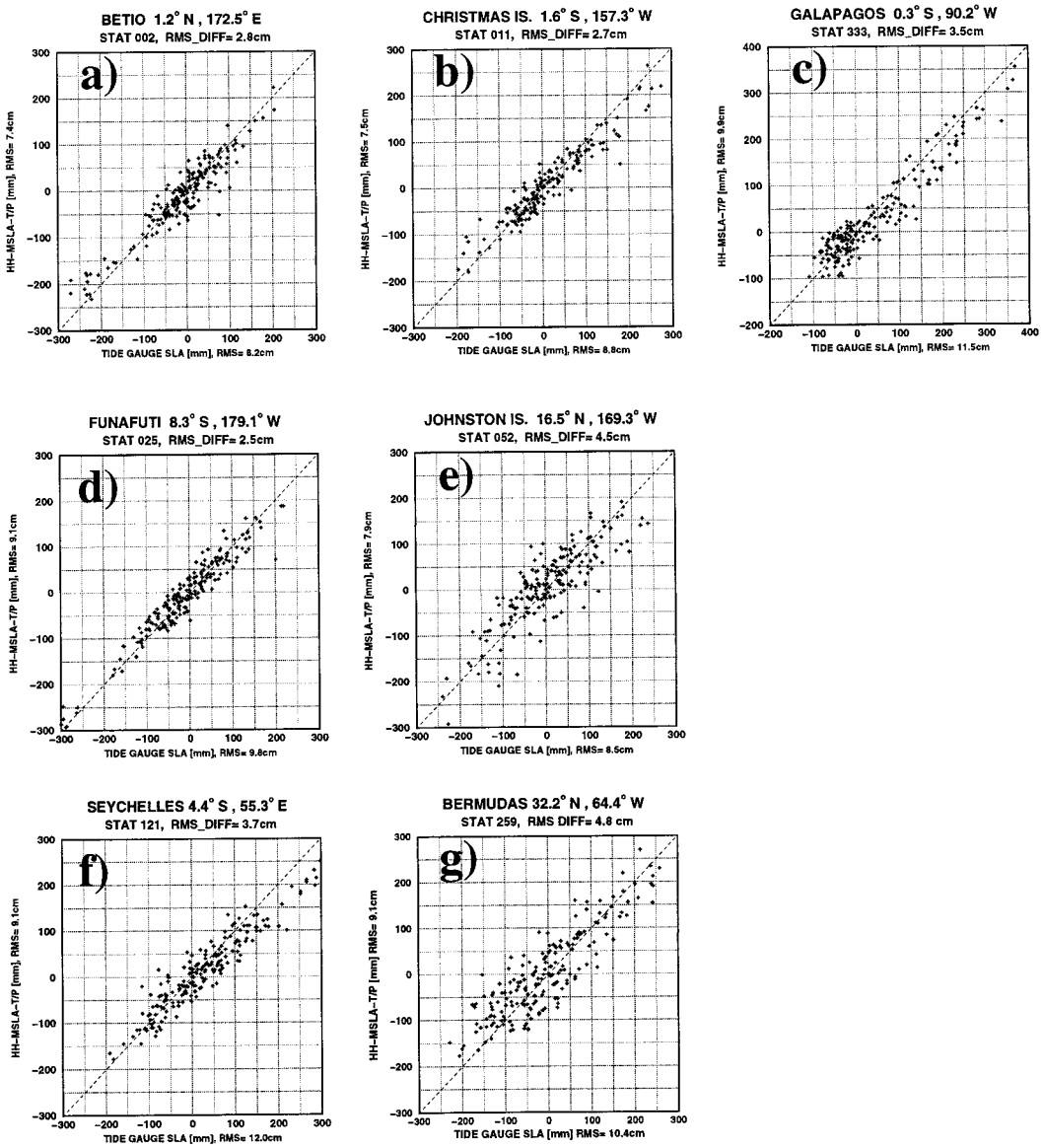


FIG. 1. Scatterplots of sea level anomalies observed at tide gauge stations vs SLA observed by TOPEX/Poseidon at the nearest grid point. Data are averaged over 10 days for the period Jan 1993–May 1998. (a) Betio, (b) Christmas Island, (c) Galapagos, (d) Funafuti, (e) Johnston Island, (f) Seychelles, and (g) Bermuda. The rms difference between tide gauge and altimeter SLA is given in the heading of each panel, the individual rms variations in the respective axis labels. Note the different origin in (c), which is introduced to capture the high anomalies caused by the 1997–98 El Niño.

Scatter diagrams of tide gauge and T/P sea level anomalies for the stations in Table 1 are shown in Fig. 1 for January 1993–May 1998. The rms variability of the individual time series of SLA is given in the axes labels, the rms difference between the two time series is shown in the subtitle. Both datasets contain randomly distributed noise, and therefore even in the absence of systematic errors the points would be scattered. Therefore one cannot expect the points to be distributed along the diagonal, but to form a relatively narrow cloud symmetric to it. This is to first order the case at the equatorial stations (Figs. 1a,b,c,f) and at 8.3°S (Fig. 1d). For the off-equatorial stations (Figs. 1e,g) the spread is larger but still almost symmetric to the diagonal (thin dashed line). Therefore the relationship between the sea level as observed by the altimeter and the tide gauges is to first order unbiased and no large systematic differences are present.

The rms differences between SLA from altimeter and tide gauges have been computed from the respective time series to quantify the deviations between the two observation platforms, but as no rms error is given with

the tide gauge data, and the rms of ±30 mm for the altimeter data is a global average, they can be used only for a relative estimate of whether systematic differences between the two observation sources are present. At stations a–c, all of them in the tropical Pacific, the rms differences are comparable to the order of the altimeter accuracy. Off the equator, the agreement is still excellent at Funafuti (d, rms = 25 mm), whereas at Johnston Island (e), 10° farther off the equator, the relation is worse (rms = 45 mm). In the Indian Ocean at the equator, the Seychelles (f, rms = 37 mm) there is good agreement between tide gauge and altimeter except when the tide gauge has large positive anomalies (larger than 200 mm), for which cases the altimeter is biased low. Bermuda is at considerably higher latitude than the other stations and the rms difference of 48 mm is the largest of the stations shown, but still not much larger than 30 mm. The overall agreement for the chosen stations suggests that tide gauge observations could serve as a means of monitoring the quality of real-time altimeter data. For the present setup of the forecasts system, however, this would require the tide gauge data to be available not later then 10 days behind real time.

c. Observations used for verification

To verify the ocean analyses, we compare the results not only to the observed sea level but also to observations of subsurface temperature. Subsurface temperature observations are derived from the Tropical Ocean and Global Atmosphere–TAO thermistor chains (McPhaden 1995) deployed across the equatorial Pacific between 8°S and 8°N, from Expendable bathythermographs (XBTs) mainly along ship routes, and also from Profiling Autonomous Lagrangian Circulation Explorer floats, which drift at subsurface levels and occasionally ascend to the surface, on their way recording temperature (T) and salinity (S) profiles. The subsurface temperature observations are routinely obtained at ECMWF from the Global Temperature and Salinity Pilot Project at the National Oceanic Data Center. The observations are assimilated into the current quasi-operational ocean analysis system at ECMWF. Furthermore, they have been used to derive dynamic height anomalies.

Dynamic height anomalies relative to the 1993–95 3-yr mean have been computed from subsurface temperatures observed at TAO moorings for the period January 1992 to September 1998 by the TAO–Pacific Marine Environmental Laboratory project office. The daily values of dynamic height were box averaged at ECMWF over 10-day intervals centered at dates for which altimeter maps are available. Dynamic height can then be compared to sea level anomalies from the maps with some restrictions. First, dynamic heights are computed using observed temperatures from the surface down to only 450-m depth, whereas the integrated temperature over the whole ocean depth is observed by satellite. Second, salinity is only very infrequently observed at TAO moorings and is therefore derived from a long-term T–S relationship. Hence, no account of interannual variations of surface or subsurface salinity is taken in the calculation of dynamic height anomalies, whereas variations in the T–S relationship contribute to the sea level from the altimeter. The first restriction is unlikely to cause problems, because for most of the tropical oceans the thermocline, around which most of the interannual temperature variability takes place, is located at a depth of less than 200 m. There is indication, however, that changes in the water mass characteristics can cause some problems when comparing sea level with dynamic height in particular in the western equatorial Pacific (Ji et al. 1999). Zonal shifts of the near-surface salinity front, caused by changes in the trade winds and the related precipitation patterns, as well as changes in the subsurface flow patterns, which advect relatively salty water from the subtropical gyre in the southern Pacific to the western equatorial Pacific at about 150-m depth, can alter the local T–S relationship.

Although it is true that the density of seawater in the upper tropical oceans is to first order determined by temperature, salinity variations in the western Pacific can be large enough to have significant impact on sea level to the extent of several centimeters. We will come back to this point in section 5a.

3. The ocean model and the altimeter assimilation procedure

The oceanic component of the coupled ocean–atmosphere forecast system at ECMWF is a modified version of the Hamburg Ocean Primitive Equation model (HOPE) developed at the Max Planck Institute for Meteorology and described in detail by Wolff et al. (1997). Sea level is computed as a prognostic variable. The model uses an Arakawa E-grid with a zonal resolution of 2.8125° at all latitudes. To allow a better representation of tropical waves, the meridional resolution is refined to ½° within 10° of the equator. Poleward of 10° the meridional grid spacing increases linearly to 2.8125° at 30° lat and poleward thereof. The uppermost 12 layers are centered at 10, 30, 51, 75, 100, 125, 150, 175, 206, 250, 313, and 425 m. The model time step is 2 h.

Although studies from as early as the beginning of this century (Helland-Hansen and Nansen 1916) imply that a proper simulation of the Gulf Stream should be important for predicting interannual climate variations over Europe, and we would like to use an eddy-resolving ocean model, the currently available computer resources do not allow the application of such a model. With the quasi-operational forecast system, 200 days of coupled integration are performed every day, and even only doubling the current ocean model resolution increases the computer time required by the ocean model by a factor of 10, thus doubling the cost of a coupled forecast if the atmospheric resolution is kept constant. Therefore at present our efforts to improve the ocean analyses are

restricted to the large-scale heat content variations in the Tropics.

To project the sea level observations on the model's subsurface density fields, we use the method developed by Cooper and Haines (1996, henceforth CH96). Vertical shifts of the *T*–*S* profiles are used to update the model's temperature and salinity fields. The magnitude of the shifts is such that the depth-averaged change of weight of the water column after assimilation compensates for the difference between simulated and observed sea level. CH96 has the advantage that, as compared to statistically derived relationships between sea level displacement and subsurface temperature changes (used in Fischer et al. 1997; Carton et al. 1996; Ji et al. 1998, 2000), *T* and *S* increments are based on the model's *T* and *S* profiles at the time and location at which the observation is assimilated. Therefore the scheme is able to capture advective changes in water mass properties, whereas statistically derived projections are not, as they are based on time-averaged variability at a fixed location. CH96 is designed to correct displacements of the thermocline, which are caused by wind or some model errors. Potential drawbacks are that it is not able to correct for steric sea level changes nor for errors in the freshwater input nor for errors in the model's water mass characteristics.

The horizontal resolution of the mapped altimeter data is much higher than the resolution of the ocean model. The mesoscale eddy activity resolved by the observations can therefore not be resolved by the ocean model. In particular at higher latitudes where the internal Rossby radius is on the order of 50 km, the effective 2° resolution of the model is much too coarse to resolve eddies. To damp out the mesoscale noise, the sea level maps are smoothed by linear regression using locally defined spatial decorrelation scales (LOESS smoother; Chelton and Schlax 1994). No smoothing in time is performed as the model's time step of 2 h is short compared to the 10-day interval between maps. The shape of the spatial smoothing takes into account a combination of the resolution of the model and of the physical decorrelation scales. In the Tropics, typical decorrelation scales for sea level on weekly to monthly timescales are on the order of 1000 km zonally and 100 km meridionally. The zonal *e*-folding scale of the smoother is set to 1160 km ($4\Delta x$) at all latitudes. The meridional length scale grows linearly from 100 km ($2\Delta y$) at the equator to 1160 km poleward of 30°. The smoother of the altimeter data thus has an elliptical shape in the Tropics and is circular in high latitudes.

The relatively large-scale smoothing outside the Tropics has the unwanted side effect of giving more weight to steric effects, which dominate the seasonal cycle of large-scale sea level variation at higher latitudes. The assimilation scheme that we use cannot correct for steric effects. However, our model is forced by daily values of precipitation minus evaporation, and we therefore should be able to capture the sea level variations due to freshwater fluxes at least partially. To damp remaining errors, less weight is given to the altimeter data at higher latitudes during the assimilation. This simpler approach is justified as the mesoscale variability present in the sea level maps cannot be resolved by the ocean model.

Two examples of SLA before and after smoothing are given in Fig. 2. SLA is shown as time series at two tide gauge locations, one station on the equator and the other station at 32°N, for the original data (solid) and the smoothed data (dotted). For the equatorial station the SLA is hardly altered at all, but for the midlatitude station, where the model resolution is poorer than at the equator, the higher-frequency temporal variations are smoothed out as a result of the spatial smoothing of the altimeter maps. The smoothed maps of the sea level anomalies are then interpolated to the grid points of the ocean model where a mean sea level is added before the data are assimilated (the choice of the mean sea level will be discussed in section 4).

Temperature and salinity increments are finally computed from the vertical shifts of the *T* and *S* profiles. However, the increments are not applied instantaneously but spread over 10 days in order to reduce any potential shock to the model and to allow a smoother geostrophic adjustment. The time-evolving background field is updated by 1/120 of the increment every 2 h over the subsequent 120 model time steps, followed by the next assimilation cycle. Thus sudden changes in the density field are avoided, but a potential drawback is that the model solution can never match the observations exactly. Furthermore, in the equatorial Pacific, baroclinic waves can travel quite fast (e.g., first mode Kelvin waves can travel 30° in 10 days). Hence, by the time the last part of an increment is applied, the errors that are being corrected for, may well have traveled several grid points farther. However, the large-scale sea level field in the equatorial Pacific varies over longer timescales than 10 days (Wyrtki 1985). Moreover, tests of the assimilation scheme in a twin experiment environment as performed by Alves et al. (2000) showed that under idealized conditions, that is, a perfect model and perfect data, the assimilation of *pseudo* sea level observations corrected for errors in the depth of the 20°C isotherm (D_{20}) over much of the tropical Pacific.

4. Ocean analyses

a. Experimental setup

Four experiments not including assimilation of altimeter data are used to guide us in the choice of a mean sea level to add to the observed sea level anomalies and later for comparison with the experiments in which altimeter data are assimilated. The set consists of one control experiment and three experiments in which subsurface temperature observations are assimilated.

The control experiment (CO-1) was started on 1 January 1985 from Levitus climatological conditions (Lev-

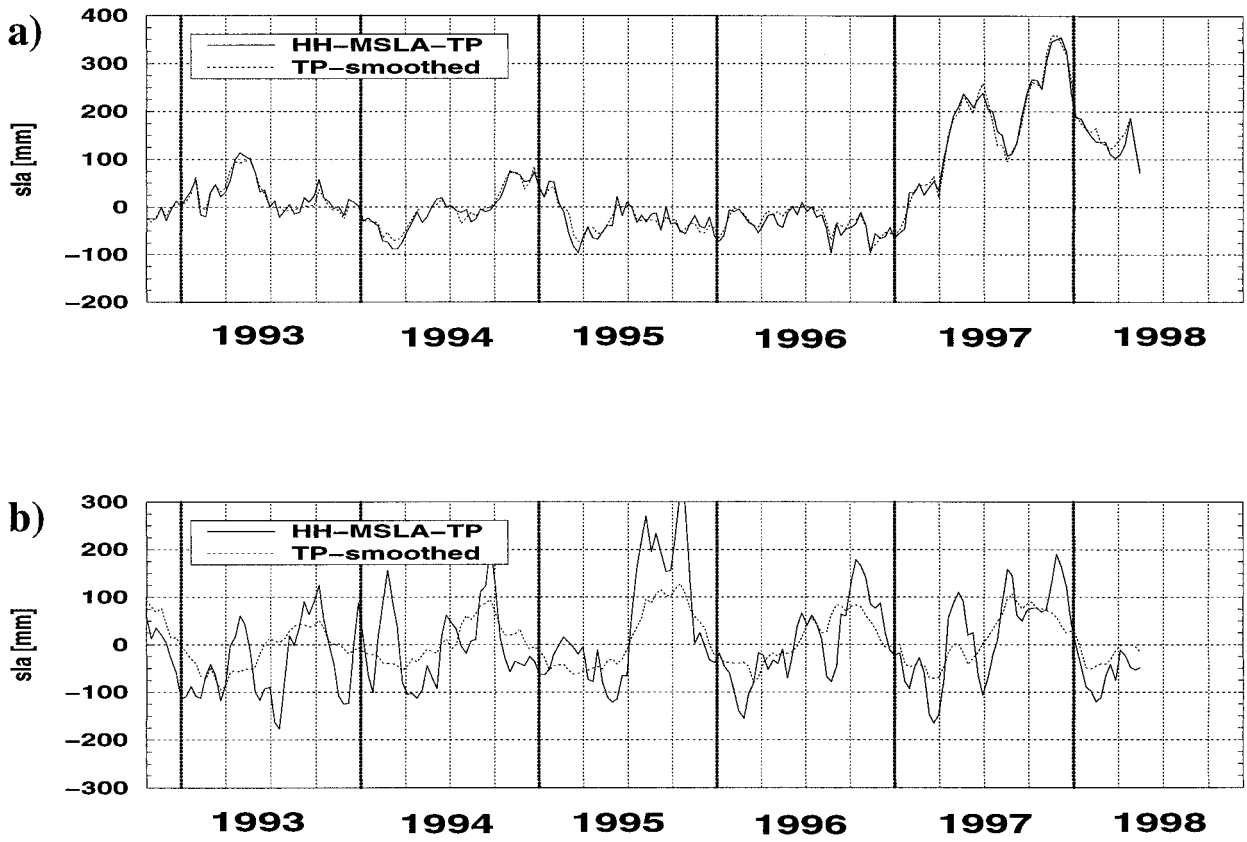


FIG. 2. Time series of unfiltered (solid) and spatially smoothed SLA (dotted) from T/P for (a) an equatorial tide gauge station (Galapagos) and (b) a midlatitude station (Bermuda, 32°N).

itus and Boyer 1994; Levitus et al. 1994). The surface forcing fields were daily values of wind stress, heat flux, and precipitation minus evaporation from the ECMWF reanalysis from 1 January 1985 until 31 December 1993, and thereafter from the ECMWF operational atmospheric analysis system until 30 September 1997. In addition to that forcing, the uppermost layer was strongly relaxed to observed weekly averages of NRT SST (Reynolds and Smith 1995) with a fast timescale of 3 days and to sea surface salinity climatology (Levitus and Boyer 1994; Levitus et al. 1994) with a timescale of 1 month.

The first experiment in which subsurface temperature observations are assimilated is denoted OI-1. It is similar to the ocean analysis that is currently run quasi-operationally at ECMWF to initialize the seasonal forecasts. OI-1 was started from the control run on 1 January 1990 but differs from it in that subsurface temperature observations are assimilated via a univariate optimum interpolation (OI) scheme (Smith et al. 1991; Smith 1995). Subsurface temperatures were bunched into 10-day windows (5 days on either side of the actual date) and an OI was performed level by level on overlapping subdomains to calculate corrections to the model state. As for the assimilation of sea level (discussed in section 3) the corrections were not applied in a single application but spread over 10 days.

The setup of the second OI experiment (OI-2) is very similar to that of OI-1: the only difference is that subsurface temperature and salinity are relaxed to climatological data with a timescale of 1 month. The relaxation is done to damp changes in water mass properties that are induced when temperature is updated but salinity is not (due to the lack of observations).

The third OI experiment (OI-3) is the strongest fit to observed temperatures: observed subsurface temperatures are assimilated every single day rather than every 10 days and the increments are spread over only 1 day. The data window, however, remains at 10 days so that observations are used several times in the analysis. Subsurface temperature and salinity are relaxed to climatology as in OI-2. The above experiments are listed in Table 2.

Experiments in which altimeter data are assimilated from 1 November 1992 until 30 September 1997 are listed in Table 3. The forcing of the model is the same as in CO-1, but the experiments differ with respect to initial conditions, the choice of mean sea level, the treatment of subsurface salinity, and the weight given to the analysed temperature and salinity increments.

b. Sensitivity of results to choice of mean sea level

As pointed out in section 2, altimeter sea level data are not given as absolute values, but are anomalies rel-

TABLE 2. Experimental setup for control run (CO-1) and TAO–XBT subsurface temperature assimilation experiments (OI-1 to OI-3). The timescale for relaxation is 3 days for surface temperatures (T_{sfc}), and 1 month for subsurface temperatures (T_z), surface salinity (S_{sfc}), and subsurface salinity (S_z).

Expt.	Assimilation (frequency)	Initialization	Relaxation
CO-1	—	Levitus @1.1.85	T_{sfc}, S_{sfc}
OI-1	TAO–XBT (10 day)	CO-1 @1.1.90	T_{sfc}, S_{sfc}
OI-2	TAO–XBT (10 day)	CO-1 @1.1.90	T_z, S_z
OI-3	TAO–XBT (10 day)	CO-1 @1.1.90	T_z, S_z

TABLE 3. Altimeter assimilation experimental setup for experiments used to investigate model-derived mean sea level, initial conditions, and relaxation to climatological subsurface data. The timescale for relaxation is 3 days for T_{sfc}, and 1 month for S_z, S_{sfc}.

Expt.	Mean	Initial conditions	Relaxation
ALT-1	CO-1	CO-1	T_{sfc}, S_{sfc}
ALT-2	OI-3	CO-1	T_{sfc}, S_{sfc}
ALT-3	OI-1	OI-1	T_{sfc}, S_{sfc}
ALT-4	OI-1	OI-2	T_{sfc}, S_z
ALT-5	OI-3	OI-3	T_{sfc}, S_z

ative to the 1993–95 mean. The question then arises as to how best to use the data in the ocean model. One choice would be to compute anomalies relative to the observed seasonal cycle from the altimeter data and to assimilate only these into the model, as is done in statistical schemes. This, however, does not seem to be an optimal strategy as then the data cannot be used to correct for errors in the model seasonal cycle. Indeed it was shown by Segschneider et al. (1999) that the seasonal cycle observed by satellite differs from the simulated seasonal cycle for all the above experiments. The best way to use the data in the CH scheme is rather to find an estimate of the mean sea level. This cannot yet be found by altimetry but must be obtained by other means. A natural choice is to use the mean sea level from a model experiment in which subsurface temperatures have been assimilated. However, as will be shown in more detail later, even a strong fit to subsurface temperature observations as in experiment OI-3 is unlikely to provide an ideal mean sea level.

Because there is no perfect choice of mean, we first try to estimate the sensitivity of the ocean analyses to mean sea level by using mean states from three different ocean analyses (CO-1, OI-1, and OI-3); the fit of the analyses to observed subsurface data will be examined later. We begin with two experiments denoted ALT-1 and ALT-2. ALT-1 uses the mean from the control experiment and ALT-2 uses the best available estimate of a mean sea level from the strong fit to observed temperatures (OI-3). For the comparison, monthly means of sea level and D_{20} were spatially averaged over the key regions of Niño-3 ($\pm 5°$, 90°–150°W) and Niño-4 ($\pm 5°$, 150°W–160°E) in the equatorial Pacific. By comparing the time series of the sea level for the two experiments, one finds that the offset of added mean sea level is maintained during the assimilation.

For variations in upper-ocean heat content, which are, to a large extent, represented by the variations of D_{20}, the sensitivity to the choice of mean sea level is less simple. From Fig. 3a it is obvious that the difference in mean sea level causes an offset of D_{20} between ALT-1 (dashed) and ALT-2 (solid) that is established after only a few assimilation steps. The variations of D_{20} in the two experiments, however, differ in the first 2 yr of assimilation before they become increasingly similar. The amplitude of the sharp rise of D_{20} that begins in late 1994 is increased from about 30 m in ALT-1 to about 40 m in ALT-2. The stronger variability of ALT-2 shows even more clearly in temperatures at 100-m depth (Fig. 3b). A drift toward lower temperatures occurs in both experiments during the period 1993–96. To investigate whether this drift is already present in the control experiment but is not corrected by the assimilation, or whether it is introduced by the assimilation, we compared the subsurface temperatures of ALT-1 and ALT-2 to those of the control experiment. From the comparison it was found that the drift of the control experiment is similar to that of ALT-2. While errors of the control experiment were not corrected in ALT-2, an even stronger drift was introduced by the assimilation in which an uncorrected mean sea level was used. At 250-m depth (Fig. 3c), the drift toward lower temperatures is stronger than in the control run in both experiments. Furthermore, caused only by the difference in mean sea level, the temperatures in ALT-1 and ALT-2 slowly diverge. At the end of the experiments, the temperature difference at 250-m depth is almost 3°C.

The drift of the subsurface temperature occurred even though we used our best estimate of mean sea level in ALT-2. Why is the drift of the control experiment not corrected? The initial conditions for ALT-2 are taken from the control run, whereas the mean sea level was from OI-3. As a result the initial conditions are not in balance with the sea level, neither in T nor S. As pointed out above, the nature of CH is such that while it can correct $T(z)$ and $S(z)$, it cannot correct for errors in $T(S)$, since water mass properties are preserved: any errors in the T–S relationship present in the initial conditions will be maintained. Therefore, in subsequent experiments, discussed below, initial conditions and mean sea level will be taken from identical or very similar experiments. This, however, turned out to be not sufficient to constrain the subsurface temperature drift.

c. *Influence of salinity*

It was then found not only that temperature drifts but also that salinity at subsurface levels was degraded in the eastern equatorial Pacific where salinity profiles are strongly nonmonotonic and the thermocline is shallow. The subsurface salinity drift was stronger in the altimeter experiments than in the control run. Although not

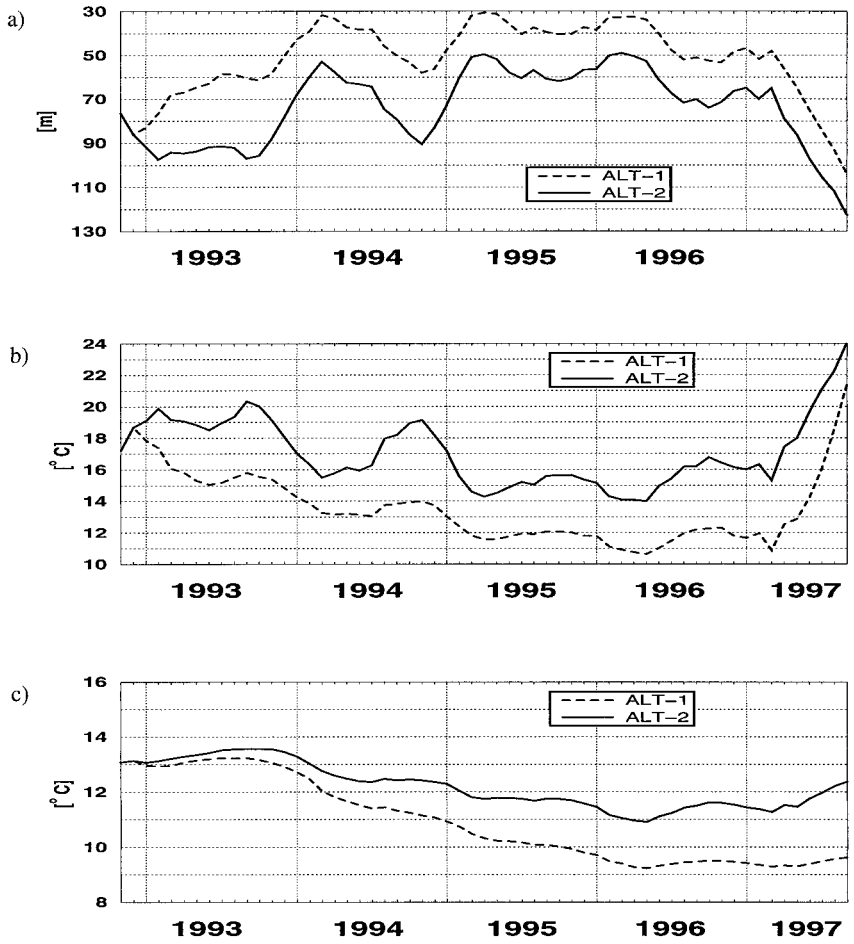


FIG. 3. Averages over the Niño-3 area (5°S–5°N, 90°–150°W) of (a) depth of the 20° isotherm, (b) temperature at 100-m depth, and (c) temperature at 250-m depth for experiment ALT-1 (dashed) and ALT-2 (solid). Curves overlap for the first month.

fully investigated yet, the drift seems to be caused by a slow erosion of vertical salinity gradients, which occurred after successive assimilation steps. Due to the limited vertical resolution of the ocean model, the salinity profile on model layers has very localized extrema, which are represented only by one model layer. For such profiles, the vertical shifts can result in smoothing of the extrema. Although splines are used in CH to minimize vertical smoothing, it seems this is not sufficient for the given model configuration and profiles. Furthermore, where the thermocline is relatively shallow, the computed vertical shifts can sometimes be larger than the mixed layer depth. Once warm water or salt is removed from the water column by the vertical shifts of the T and S profiles, it can be reintroduced only by surface relaxation or advective processes, both of which act rather slowly.

Two further experiments were designed to test if relaxation to subsurface salinity could prevent the model from drifting. ALT-3 uses the mean sea level obtained from OI-1, in which temperature data were assimilated every 10 days. ALT-4 uses the same mean sea level from experiment OI-1, but differs from ALT-3 in that subsurface salinity is relaxed to the Levitus seasonal cycle with a timescale of 1 month. Furthermore, to avoid an initial discrepancy of subsurface salinity, experiment ALT-4 is started from experiment OI-2 in which subsurface salinity is relaxed to Levitus rather than from OI-1 as used in ALT-3. In OI-1, subsurface salinity is not relaxed to Levitus, and at the beginning of 1993 differs slightly from the climatology (about 0.1 psu at 100-m depth when averaged over Niño-3). However, the tropical temperature fields and mean sea level of OI-1 and OI-2 are very similar. Therefore, the comparison of experiments ALT-3 and ALT-4 allows an assessment of the sensitivity of results to subsurface salinity relaxation.

Figure 4a shows that even when the initialization is made from similar states as in experiment ALT-3 (dashed) and ALT-4 (solid), and the same mean sea level is used, differences in D_{20} averaged over the Niño-3 region are as large as 10 m, depending on whether or

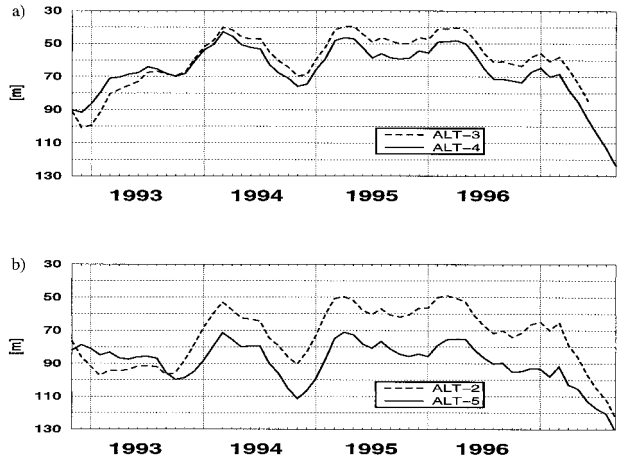


FIG. 4. Averages of D_{20} over Niño-3 from (a) experiment ALT-3 (dashed) and ALT-4 (solid) and (b) experiment ALT-2 (dashed) and ALT-5 (solid).

not subsurface salinity is prevented from drifting. A change of D_{20} on the order of 10 m corresponds to a sea level signal of about 5 cm. This is larger than the accuracy of T/P, and implies that salinity is no longer negligible in the context of altimetry. As only very few salinity observations exist, we chose to relax salinity to climatological data, although this has the unwanted side effect of damping interannual salinity variations.

d. Initial conditions and reduction of weights

The fifth experiment in which altimeter data are assimilated is denoted ALT-5. Experiment ALT-5 makes use of the same mean sea level as experiment ALT-2, but is started from the corrected subsurface conditions obtained from OI-3. It differs further from ALT-2 in that subsurface salinity is relaxed to Levitus climatology. Furthermore, reduced weights are applied to the computed temperature and salinity increments. In experiments ALT-1 to ALT-4 full weight was given to the altimeter data. In experiment ALT-5 the T and S corrections are reduced slightly close to the equator and more strongly in higher latitudes by an empirically derived factor of $0.9 \cos^2\phi$, where ϕ is latitude.

The reduced weights given to altimeter data at higher latitudes were needed to prevent the model from developing unstable density stratification in the area of the middepth salinity maxima, such as to the south of the southern Pacific subtropical gyre. In integrations giving full weight to altimeter data, the assimilation induced unrealistic T and S variations after approximately 2 yr, which then propagated into the equatorial region on a timescale of a further 2–3 yr. Consequently, Niño-3 SST at the end of the integration (1997) turned out to be erroneous. This was no longer the case when the weights of the corrections were reduced.

Comparison of experiment ALT-2 with experiment ALT-5 allows an assessment of the combined impact of improved initial conditions, relaxation of subsurface salinity, and reduced assimilation increments. Figure 4b shows that in ALT-5 (solid) D_{20} does not drift, whereas in ALT-2 (dashed), which uses the same mean sea level, the thermocline drifts to a state that is up to 25 m shallower. Relaxation to salinity in subsurface layers with a timescale of 1 month as in experiments ALT-4 and ALT-5 constrained the salinity efficiently, but it was only by additionally using the mean state and initial conditions from the strong fit to subsurface temperature observations (OI-3) that subsurface temperature was prevented from drifting (ALT-5).

5. Comparison of ocean analyses with observations

a. Interannual variability of sea level

So far no measure of reality has been given for the ocean analyses. In this section, we compare observed and simulated sea level anomalies. The comparison is restricted to the time evolution along the equator. The differences between SLA as observed by T/P and as simulated by experiments CO-1, OI-1, OI-3, and ALT-5 are shown in Fig. 5. The 3-yr mean from 1993–95 and the seasonal cycle, based on the same years, were removed from the observations and the model results before the differences were computed. For the comparison we use gridded HH data, smoothed and interpolated to the model grid as described in section 3. To allow better readability, further smoothing has been applied to the difference plots and errors of more than 20 mm are shaded. Plotted is the *absolute value* of the differences.

The largest differences between observed and simulated sea level anomalies occur in the control experiment (Fig. 5a). Errors of more than 40 mm prevail in the Pacific east of the date line, in particular during 1993, but also in late 1994 at about 115°W and at 140°W in late 1995 to early 1996. The erroneous sea level of the control experiment in the Pacific is substantially corrected by the 10-daily assimilation of subsurface temperatures (Fig. 5b). In the Atlantic, however, OI-1 performs slightly worse than the control experiment and for the Indian Ocean no overall improvement from the assimilation is visible.

In experiment OI-3 the temperatures are assimilated every day, and salinity and subsurface temperatures are relaxed to climatological data to constrain the density fields. While this approach works well in the east and central Pacific, errors remain in the western Pacific and in the Atlantic and Indian oceans (Fig. 5c). In particular in the Atlantic, where only few temperature observations are available, the relaxation to climatological temperature and salinity damps the variability imposed by the forcing fields. In the western Pacific this cannot be the reason for the sea level errors, because temperature observations are provided by the TAO moorings. Here salinity variations may play a significant role (Ji et al. 2000).

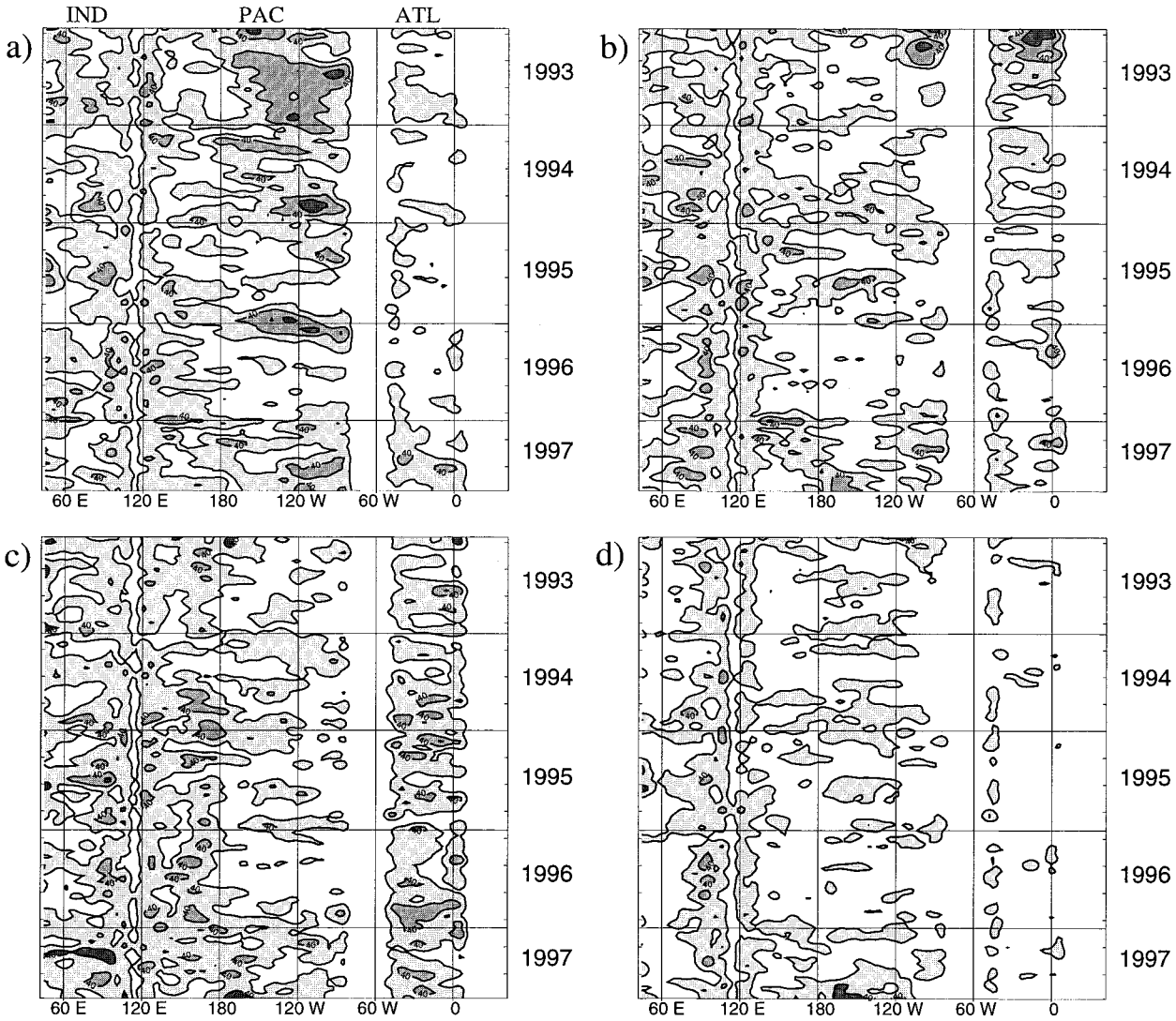


FIG. 5. Absolute difference between simulated and observed SLA along the equator for (a) control experiment CO-1, (b) 10-daily assimilation of in situ temperatures OI-1, (c) daily assimilation of in situ temperatures OI-3, and (d) altimeter experiment ALT-5. Contours are at 20, 40, and 60 mm. Differences of more than 20 mm are shaded, time runs from top to bottom.

To investigate the extent to which salinity variations can contribute to the sea level anomalies, we compared the dynamic height at TAO moorings with the sea level observed by satellite. Sea level of OI-3 and TAO dynamic height are similar because the same in situ temperature data that are assimilated in OI-3 are used for the computation of dynamic height. A Hovmöller diagram of the difference between TAO dynamic height and sea level anomalies from satellite along the equator (Fig. 6) shows that in the western equatorial Pacific the dynamic height is as much as 90 mm higher than the sea level observed by satellite during 1995 and 1996. The average difference is in the range of 30 to 60 mm, which is on the same order as the differences between OI-3 and T/P in the western equatorial Pacific. We can therefore assume that the differences between OI-3 and T/P are not caused by model errors but rather by damping out salinity variations. Figure 6 also implies that even our best estimate of mean sea level, which was obtained from experiment OI-3, is likely to be imperfect, in particular in the western Pacific.

The assimilation of the observed sea level itself (ALT-5) reduces the errors of the control run effectively (Fig. 5d), in particular in the Atlantic, but also in the eastern and western equatorial Pacific and in the western Indian. However, some smaller errors remain. For example, the erroneous sea level of the control run in the central equatorial Pacific is only partly corrected. The relatively large errors in the eastern Indian Ocean and the Indonesian Throughflow region, which are present in all four experiments, are likely to be caused by relatively strong tidal noise of

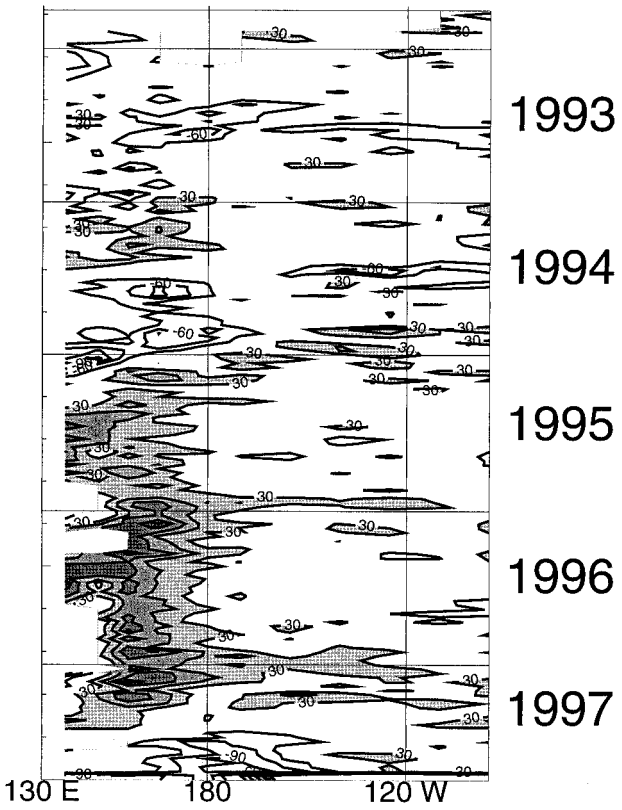


FIG. 6. Hovmöller diagram of the difference between TAO-derived dynamic height and altimeter observed sea level. Dynamic height is computed from data at 2°S, the equator, and 2°N, the altimeter SLA is from the equator. The contour interval is 30 mm. The zero contour is not shown. Values larger than +30 mm are shaded.

the altimeter data, which cannot be resolved by the ocean model.

b. Sea level from model, satellite, and tide gauges

In this section we further compare simulated and observed sea level. Model results are now compared with both altimeter and tide gauge to quantify the model agreement with two independent observation sources. The experiments are CO-1 in which no data are assimilated, OI-1 in which subsurface temperature data are used, and ALT-5 in which altimeter data are assimilated. OI-1 is chosen rather than OI-3 for consistency since in OI-1 assimilation is done every 10 days (as in experiment ALT-5) and because the setup corresponds to the present quasi-operational ocean analyses. Because the main impact of an improved ocean representation on the seasonal forecasts is expected from the tropical Pacific, we first show time series of sea level at Betio in the equatorial west Pacific and Galapagos in the eastern equatorial Pacific (Fig. 7). The appropriate 1993–95 mean sea level is removed from both the observations and model results but the seasonal cycle is not.

Before the model results are compared with altimeter and tide gauge data, time series of the two observation sources are intercompared. The two upper curves in each panel show the time series of unsmoothed altimeter SLA at the map point nearest to the station (bold, solid) and the difference between SLA from tide gauges and satellite (thin, dashed). The differences are more or less randomly distributed in time for the western Pacific station, indicating that no systematic offset of the seasonal cycles is present (Fig. 7a). At the eastern Pacific station, however, SLA from tide gauge is systematically higher than from T/P during the first quarter of each year and lower during the third quarter, thus indicating an offset of the seasonal cycles (Fig. 7b). Interannual variations agree well for most of the time, but the arrival of the Kelvin wave in the east Pacific in March 1997 is more pronounced in the tide gauge data than in the altimeter data (Fig. 7b). Indication that the tide gauge data are more realistic than T/P in March 1997 at this particular location is given from the time series of dynamic height at the TAO mooring at 95°W (not shown), suggesting that there is a problem with the mapped altimeter data at this point.

Also in Fig. 7, the differences between the SLA from satellite and the three model experiments are plotted (CO-1, OI-1, and ALT-5; lower three curves). For better readability, the difference curves are offset. Improvement of the model results from assimilation of either subsurface temperature observations or sea level is evident at the eastern station (Fig. 7b). In particular during 1993 and 1994, the control experiment (dash–dotted) disagrees with the observed sea level over longer periods. The sea level is too high during the first eight months of 1993 and too low in 1994. These errors are corrected to some degree by assimilating subsurface temperature in OI-1 (thin, solid) and even more by assimilating altimeter data (dashed). The respective rms differences between altimeter data and model experiments can be reduced from 29.6 mm for the control run to 23.8 mm by assimilation of in situ temperatures to 18.2 mm by assimilation of altimeter data.

For the west Pacific station (Betio), improvements from assimilation are small. The rms error is reduced from 26.0 mm in the control experiment to 24.6 mm in experiment OI-1 and to 24.9 mm in ALT-5. The differences between simulated and observed sea level show errors that can last for several weeks and are remarkably similar in all experiments, suggesting either errors in the altimeter, or errors in the wind stress used to force the ocean model or in the ocean model itself. The differences between observed and simulated sea level are remarkably in phase in all model experiments (i.e., also in those that have no information about sea level) but are only occasionally in phase with the differences between tide gauge and T/P. We therefore suspect that errors in the daily wind fields used for the forcing are the main reason for the deviations in sea level. However, on several occasions the differences between model and altimeter and between tide gauge and altimeter *are* in

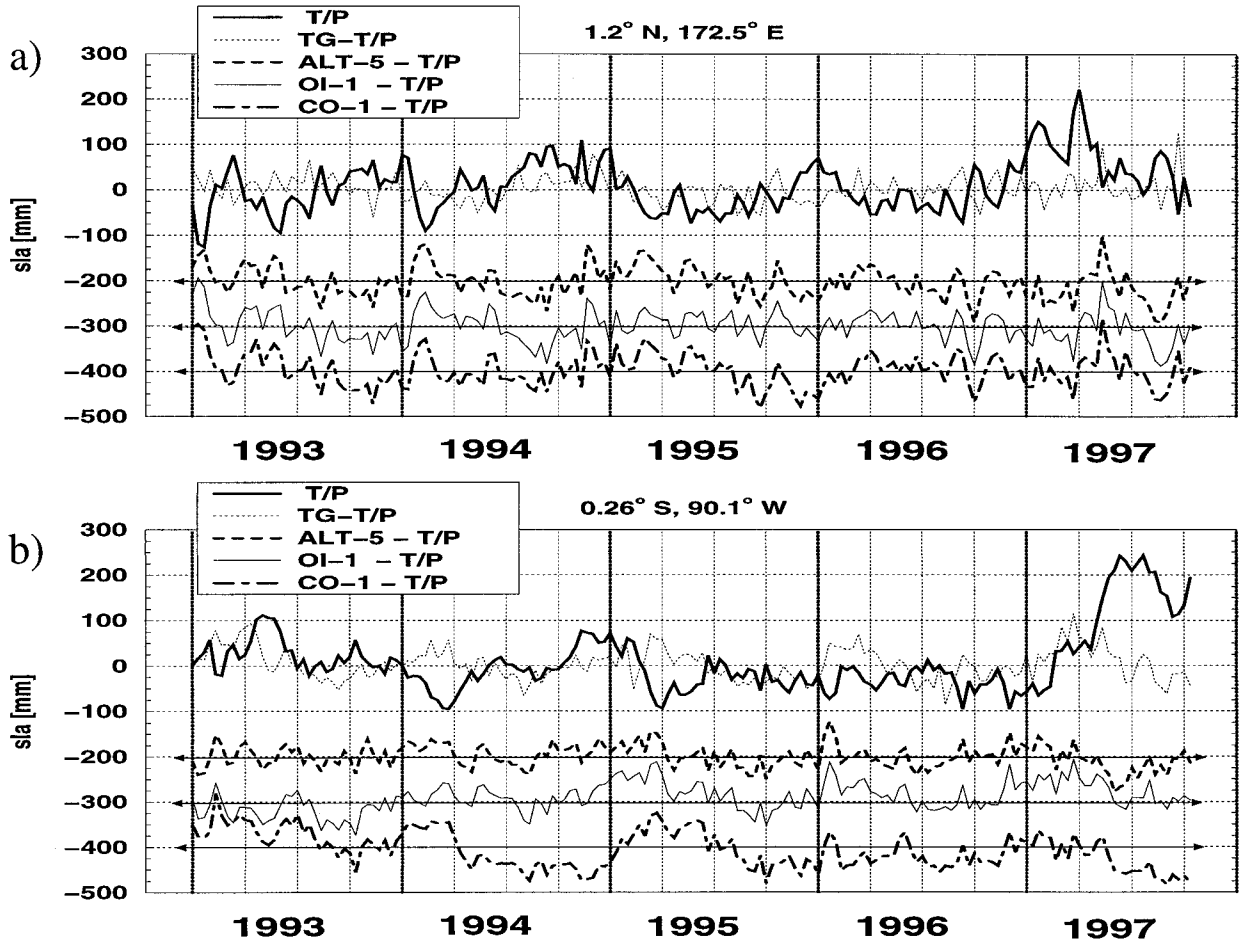


FIG. 7. Time series of SLA from T/P (bold) and the difference of tide gauge minus T/P (thin, dotted) from Jan 1993 to Sep 1997 (upper two curves) as well as the difference of model minus observed sea level (lower three curves): ALT-2 minus T/P (bold, dashed; offset −200 mm), OI-1 minus T/P (thin, solid; offset −300 mm), and CO-1 minus T/P (bold, dash-dotted; offset −400 mm). Shown are (a) a station in the equatorial west Pacific (Betio) and (b) a station in the equatorial east Pacific (Galapagos). Lines with arrow heads indicate the offset zero levels for the model-minus-observation time series. Altimeter SLA is obtained from the grid point closest to the tide gauge location.

phase (e.g., in October 1995 and May 1997 at Betio, and in March 1993 and August 1994 at Galapagos) and this suggests that model and tide gauge agree and that the altimeter data are in error for these dates.

TABLE 4. Rms differences between satellite-observed sea level (unsmoothed) and simulated sea level from experiments CO-1, OI-1, and ALT-5 (first three columns) and differences between tide gauge data and simulated sea level (last three columns) at tide gauge stations. The rms errors (mm) are computed for the period Jan 1993 to Sep 1997 from 10-daily averages.

Station	*T/P* − ALT-5	*T/P* − OI-1	*T/P* − CO-1	TG − ALT-5	TG − OI-1	TG − CO-1
Betio	24.8	24.6	26.0	25.7	26.7	23.8
Christmas Is.	21.1	21.2	28.2	28.3	28.2	33.6
Galapagos	18.2	23.8	29.6	24.1	24.1	26.9
Funafuti	21.4	30.0	31.6	24.1	31.6	32.9
Johnston Is.	45.2	55.2	55.2	51.1	60.2	60.2
Seychelles	27.7	31.8	37.5	34.7	37.0	46.7
Bermuda	46.0	63.9	61.7	54.6	74.2	72.1

We do not show time series for all tide gauge stations, but the rms-differences between model results and altimeter data (columns 1–3) and between model results and tide gauge observations (columns 4–6) for all the selected tide gauge stations (see section 2b) are listed in Table 4. Generally, the rms error is reduced by assimilation of either sea level or temperature and agreement is better with T/P than with tide gauge data. For the assimilation of altimeter data, this may not seem too surprising, but the comparison is not as direct as it might appear: first, the sea level observations were used to derive temperature and salinity corrections, not sea level corrections. Second, *smoothed* sea level observations were assimilated in ALT-5 and are compared here with the *unsmoothed* satellite-observed sea level. The sea level of experiment OI-1, which has no information about the observed sea level, neither from satellite nor from the tide gauges, also agrees better with the tide gauge sea level. For the control experiment the rms

differences between model and tide gauge are smaller than between model and altimeter at Betio and Galapagos but for the other stations the agreement of the control run is better with the altimeter observations. Overall, regardless of whether observations are assimilated or not, the model is more in balance with the altimeter observations than with the tide gauge data.

c. Replication of D_{20}

It has been shown above that sea level errors could be reduced relative to the control either by assimilation of altimeter observations or by assimilation of subsurface temperatures. But how well is the observed upper-ocean heat content simulated? To investigate whether the assimilation of altimeter data can correct for errors in the control experiment in the equatorial Pacific, time series of D_{20}, averaged over the Niño-3 and Niño-4 areas, are compared between the control experiment, in situ temperature assimilation experiment OI-3, and altimeter assimilation experiment ALT-5 in Fig. 8. Temperature data from the TAO moorings provide a relatively dense database in the Niño-3 and Niño-4 areas and are assimilated every day in experiment OI-3. Therefore D_{20} of that experiment serves as observation against which the two other experiments are compared.

Figure 8 shows that in the control run (dashed), D_{20} is too shallow in the Niño-3 and Niño-4 areas by 30–40 m. Furthermore, the temporal variations of D_{20} in the Niño-3 area are underestimated. In contrast, D_{20} in the altimeter experiment (solid) agrees rather well with that of experiment OI-3 (dotted) in Niño-3 and especially well in Niño-4. That the offset of D_{20} is corrected in the altimeter experiment is largely related to the fact that the mean sea level of experiment OI-3 is used. The corrected mean of D_{20} is not the only improvement, however. The variations of D_{20} in the altimeter experiment are in general much more similar to those of OI-3 than for the control experiment. The only exception is the first half of 1993 in the Niño-3 region. The onset of the 1997–98 El Niño event is captured in all three experiments, but the control run almost completely fails to simulate the deepening of the thermocline beginning mid 1994 in the Niño-3 area. The assimilation of altimeter data results in a much improved simulation of D_{20}, which almost matches the D_{20} of the in situ temperature assimilation experiment. In the Niño-4 region, the agreement between OI-3 and the altimeter experiment is even better (Fig. 8b). Overall, the extent to which the two experiments agree is likely to be in the range of the uncertainty of the observations, taking into account that both subsurface temperatures and sea level observations are filtered through the model and the assimilation procedures before they are compared.

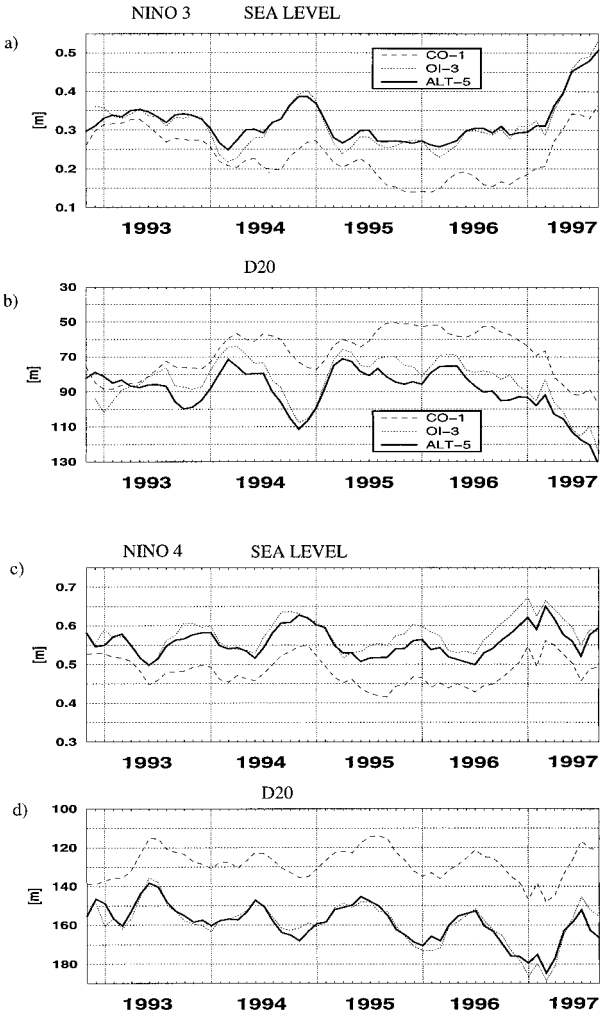


FIG. 8. Time series of (a) sea level averaged over Niño-3, (b) D_{20} (Niño-3), (c) sea level averaged over Niño-4, and (d) D_{20} (Niño-4) for experiments CO-1 (dashed), OI-3 (dotted), and ALT-5 (solid).

d. Combining HH and NRT data

In order to use the altimeter data operationally to provide oceanic initial conditions for the coupled forecasts, it is necessary to use high-quality HH data for as long a period as possible in order to be able to calculate the coupled-model drift and then to use NRT data to initialize actual forecasts. The assimilation of HH and NRT data is attempted in experiment ALT-HHNRT. The setup of the experiment is as for experiment ALT-5, that is, mean sea level from experiment OI-3 is used, subsurface salinity is relaxed to climatology, and latitude-dependent weights are given to the computed T and S increments. The experiment is integrated until the end of October 1998, thus fully capturing the 1997–98 El Niño event. The HH data are available only until May 1998; NRT data are available from 14 January 1998 on, providing an overlap of some four months. In ALT-HHNRT, we chose to assimilate HH data as long as possible until April 1998 and then to use NRT data from May 1998 on. The NRT data are available every 7 days (as opposed to 10 days for HH data) and were

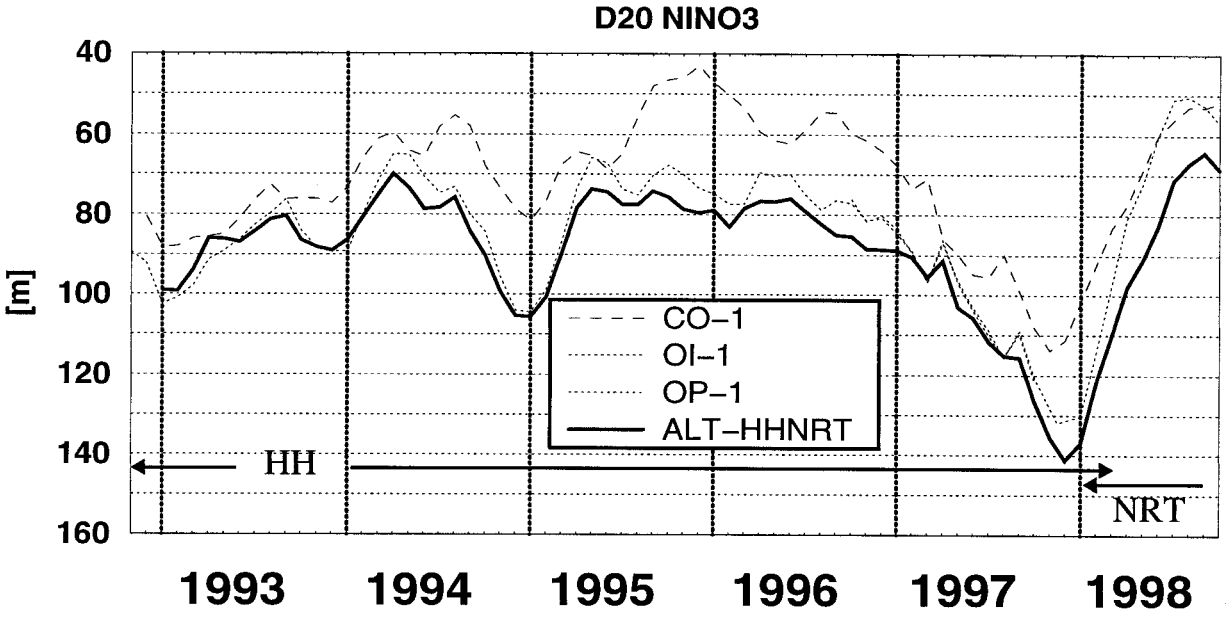


FIG. 9. Time series of simulated D_{20} averaged over Niño-3 for experiments CO-1 (dashed), OI-1 (up to Sep 1997, dotted), and OP-1 (from Apr 1996, also dotted), and ALT-HHNRT (solid). Arrows indicated the periods covered by HH data and NRT data, respectively.

interpolated in time to provide the maps at the required 10-day intervals.

From experiment ALT-HHNRT, the Niño-3 averaged time series of D_{20} is compared with the control experiment and the in situ temperature assimilation experiment OI-1 in Fig. 9. Experiment OI-1 is extended by the current quasi-operational ocean analysis (denoted OP-1) for dates after April 1996. The current quasi-operational ocean analysis has the same setup as experiment OI-1 but extends beyond September 1997. The only difference between the two experiments is that the subsurface temperature observations are manually quality controlled in experiment OI-1, whereas in the quasi-operational ocean analysis the quality control is done without manual supervision.

As seen above, the control experiment tends to simulate too shallow a thermocline, whereas the assimilation of altimeter data replicates the ''observed'' D_{20} quite closely, although the thermocline is too deep at the peak of the 1997–98 El Niño, if we take the operational analysis as the truth. The subsequent rise of the thermocline is well captured in the NRT altimeter analysis, but D_{20} is never as shallow as in OP-1 and in fact the time of minimum depth is slightly delayed. Whether the larger discrepancies, which occur during the phase in which NRT data are assimilated, are caused by lower quality of the NRT data or whether salinity effects, for example, are playing a significant role can be assessed only when the HH data for the same period have been processed, which are not available yet. The differences in D_{20} at the end of the simulation are as large as 20 m, however, and imply that further improvements for the use of altimetry are still required before it can be used operationally.

6. Coupled hindcast experiments

In this section we estimate the impact of altimetry on seasonal forecasts. Climate forecasts over 6 months are performed by coupling the HOPE ocean model to the ECMWF atmospheric numerical weather prediction model at T63 resolution by use of the Ocean–Atmosphere–Sea-Ice–Soil coupler. Because a fully coupled system is used, the coupled model drifts. To couple only the anomalies of the subsystems is one possible means of reducing such a drift, but there are serious disadvantages in anomaly coupling. So at ECMWF the two systems are fully coupled and subsequently an estimate of the drift is computed, which, if necessary, can then be subtracted from the model results a posteriori.

Forecasts are started every 3 months (i.e., on 1 January, 1 April, 1 July, and 1 October of each year) from 1.1.1993 to 1.10.1997 and integrated forward in time for 184 days. The atmospheric initial conditions are taken from the operational analysis–forecast system run at ECMWF for weather forecasting. Oceanic initial conditions are obtained from the control ocean analysis CO-1 (the coupled experiment is denoted C-CO), the subsurface temperature analysis OI-1 (coupled experiment denoted C-OI), and altimeter analysis ALT-5 (coupled experiment denoted C-ALT).

Because of the chaotic nature of the atmosphere, extensive ensembles of forecasts are required. For each forecast date we generate an ensemble by perturbing

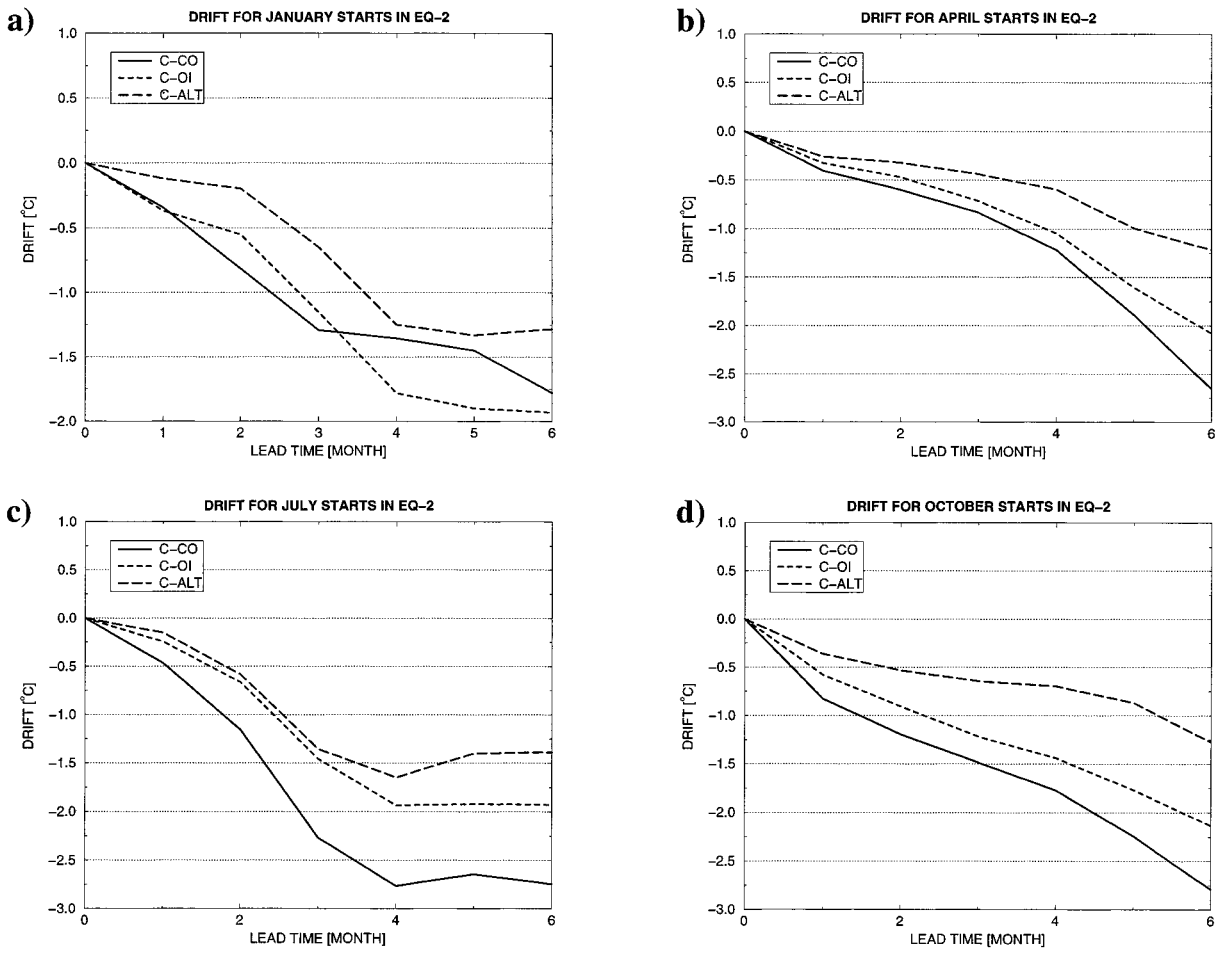


FIG. 10. Average model drift in the EQ-2 region (130°–170°W, 5°S–5°N) for forecasts started in (a) Jan, (b) Apr, (c) Jul, and (d) Oct. Shown are experiments C-CO (solid), C-OI (dashed), and C-ALT (long dashed) for lead times of 6 months.

oceanic initial conditions in the equatorial Pacific, in particular SST perturbations of 0.01°C amplitude are applied between 5°N and 5°S to the regions EQ-1 (90°–130°W) or EQ-2 (130°–170°W). The chaotic response of the atmospheric model to the applied SST changes causes a spread of the ensemble members of the forecasts. Here, each ensemble consists of five members. Since there are 20 start dates this amounts to 100 coupled forecasts for each experiment and a total of more than 150 yr of coupled model integration.

The model drift, based on the years 1993–97, was computed for each of the three coupled experiments. The SST drift is computed separately for forecasts starting in January, April, July, and October from the difference between the predicted and the observed climatologies for the 5-yr period. Because the model drift is not spatially uniform, it is estimated over subdomains in the equatorial Pacific, namely Niño-12, Niño-3, Niño-4, and EQ-1, EQ-2, EQ-3. The drift is largest for the control forecasts and smallest in the forecasts using altimeter data in five of the six regions. Only in the near-coastal region Niño-12 does C-ALT have a larger drift than the control forecasts. As an example we show the SST drift in the EQ-2 region in Fig. 10. The drift is large, frequently larger than the signal we are seeking to predict. However, although it is in general desirable to have a reduced drift, and in that sense the assimilation of altimeter data has been beneficial, there is not a simple relationship between drift and forecast skill. In experiments in which we inserted a correction to the heat flux of 15 W m^{-2} in the tropical strip the overall drift to cooler temperatures was largely removed, but there was little impact on the forecast skill. A more detailed description can be found in Stockdale (1997). As will be shown below, the reduced drift of C-ALT relative to the forecasts using in situ temperatures does not result in better forecasts.

We estimate the relative performance of the three coupled experiments by comparing predicted SST anomalies to observed values (Reynolds and Smith 1995). Although comparing only SST anomalies in the equatorial Pacific does not make full use of the information that is provided by the forecasts, the comparison is restricted to these key regions partly because the signal-to-noise

ratio is largest in the tropical Pacific and partly because improvements of the ocean analysis are likely to lead to better forecasts there.

First, we compute anomalies from the predicted SST relative to the ''predicted climatology'' and anomalies from the observed SST relative to the observed climatology. The predicted climatology is calculated separately for forecasts started in January, April, July, and October. So the climatology for January-starts covers January–July, the one for April-starts covers April–October, etc., but the values for April are different between January- and April-starts. Anomalies of SST could now be assessed relative to these model climatologies but there are some disadvantages in doing so as the climate for 1993–97 is not representative of the long-term climate. (The shorter period is biased warm as it includes a major El Niño and other warm years.) To overcome this bias we calculate the difference between the observed climatology for 1993–97 and the observed long-term climatology from the Climate Analyses Center (based on the years 1950–79) and correct the model climatologies by this amount before calculating anomalies. Observed SST anomalies against which the model forecasts are compared are computed relative to the observed climatology for 1950–79. Adjusting the model climatologies in this way makes no difference to the rms errors calculated below, but does influence the anomaly correlations, acting to increase them. The influence on the relative performance of the various experiments is small, however.

We will examine four ways in which to evaluate the coupled forecasts. There is no absolute way of intercomparing the three experiments and of validating them against data, but by giving a range of tests we show that the basic conclusions are robust. In the first method, we compare SST anomalies in the EQ-2 region and compute rms-errors and anomaly correlations. Monthly averages of observed SST anomalies (solid) and ensemble averages of predicted SST anomalies (dashed) averaged over the EQ-2 region are shown in Figs. 11a, respectively, for experiment C-CO, C-OI, and C-ALT for the period January 1993–March 1997. This figure shows that similar ocean conditions lead to similar forecasts, in that C-OI and C-ALT are more similar to each other than to the control forecasts. Recall from Fig. 8 that the variations of D_{20} were similar between the two ocean-only analyses in which either altimeter data or subsurface temperatures were assimilated, whereas D_{20} in the control experiment was different from the other two. By comparing the time evolution of predicted SST anomalies in Fig. 11 one sees that the forecasts of C-ALT and C-OI are more similar to each other than compared to the control forecasts (e.g., forecasts started at April 1993, July 1993, October 1994, January 1995, April 1996, October 1996). Further, the two experiments that benefit from data assimilation generally predict the observed SST anomalies better than the control forecasts, although the forecast started in October 1994 is an exception.

Figure 12a shows the rms error between predicted and observed SST anomalies and Fig. 12b shows the anomaly correlation coefficient. Rms error is *first* computed for each of the five forecast sets making up the individual experiments whereas for the computation of the anomaly correlation coefficient the five ensemble members for each start date are first averaged and then the anomaly correlations of the ensemble averages are computed. The vertical bars in Fig. 12a indicate two times the standard deviations of the rms error. The rms error is largest for the control forecasts (Fig. 12a). In particular for the first 3 months the error of the control forecasts grows quickly and is larger than for the persistence forecast. In experiments C-OI and C-ALT the rms errors grow more slowly than for the persistence forecast. Rms errors appear to be slightly larger in the experiment using altimeter data, but differences between the two experiments are small. We applied a Wilcoxon–Mann–Whitney test to the 100 ensemble members to investigate whether the mean of the probability density function of the absolute error was significantly shifted between C-OI and C-ALT. It turned out that the shift was only significant at the 40% level. This means that based on 100 forecasts the relative performance of experiments C-OI and C-ALT was not statistically different. By contrast, the control forecasts differ from the two other experiments with a significance greater than 99%.

The anomaly correlation coefficients (Fig. 12b) are similar between all three coupled experiments; so it is difficult to estimate the relative performance of C-ALT, even compared to the control forecasts. By eye, the anomaly correlations of the ensemble mean of experiment using altimeter data are worse than those of C-CO and C-OI for lead times of more than 2 months. We then computed the anomaly correlation coefficient for forecasts started between 1994 and 1997 instead of 1993 and 1997 on the grounds that the forecasts started in January 1993 were particularly bad in C-ALT (as shown by the ensemble mean in Fig. 11c) because the ocean model was still in a spinup phase after only 2 months of altimeter assimilation. Figure 13 shows that the anomaly correlations of C-ALT are more similar to those of C-OI when 1993 is excluded from the calculation. Using anomaly correlation coefficient as a measure, Fig. 13 also implies that the main benefit from the assimilation is restricted to the first 2 months of the forecasts, during which the anomaly correlation coefficients for C-CO is lower than for C-OI and C-ALT. For months 3–6, the anomaly correlation coefficients for the control forecasts are similar.

The second approach to evaluate the forecasts is to measure the *relative* performance. The forecasts using altimeter data are considered ''better'' if the ensemble average of the predicted SST anomaly is more than 0.2°C closer to the observed value than for C-CO or

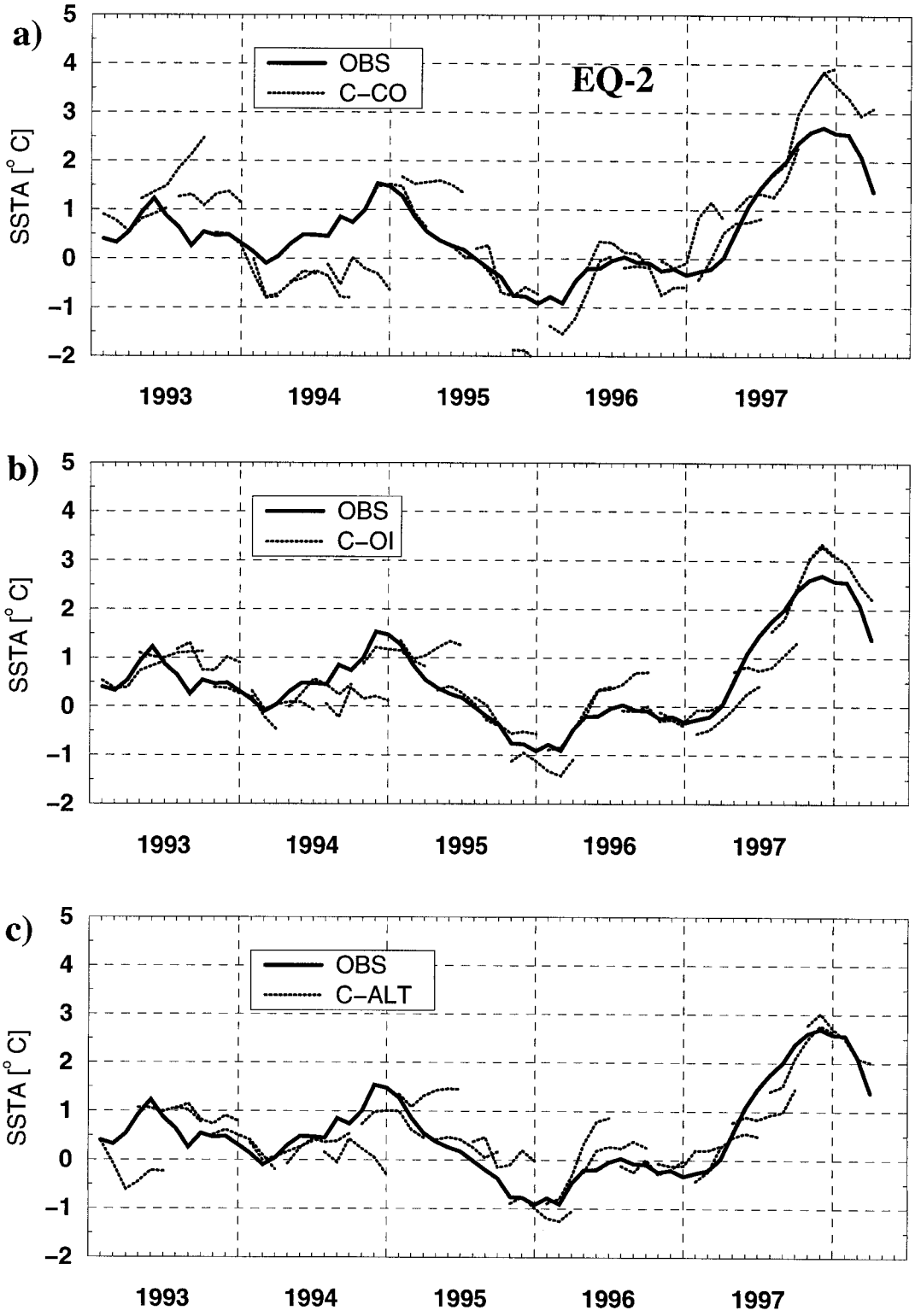


FIG. 11. Time series of observed (solid) and predicted SST anomalies (dashed) averaged over the EQ-2 region (5°S–5°N, 130°–170°W): (a) for experiment C-CO, (b) experiment C-OI, and (c) experiment C-ALT. Vertical dashed lines mark December of each year.

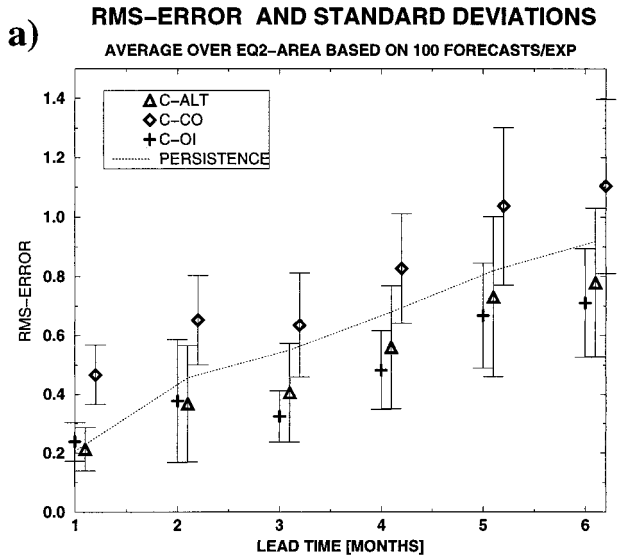


FIG. 12. (a) Rms error between predicted and observed SST anomalies as average over all ensemble members and forecast dates (1 Jan, 1 Apr, 1 Jul, and 1 Oct) in the EQ-2 region in the central equatorial Pacific (5°S–5°N, 130°–170°W), vertical bars indicate two times the standard deviation, centered at the ensemble mean value. (b) Anomaly correlation coefficient of the ensemble means for EQ-2. Shown are experiments C-CO (diamonds), C-OI (plus-signs), C-ALT (triangles), and the persistence forecast (dashed line).

C-OI, respectively. If the ensemble average of the SST anomaly is more than 0.2°C further from the observed SST anomaly than the one predicted by the two other experiments, C-ALT is considered ''worse.'' If the difference of predicted SST anomaly error is less than 0.2°C, the forecast is considered neutral. The relative performance of C-ALT is computed as the number of better minus worse forecasts. It is shown in Fig. 14a for

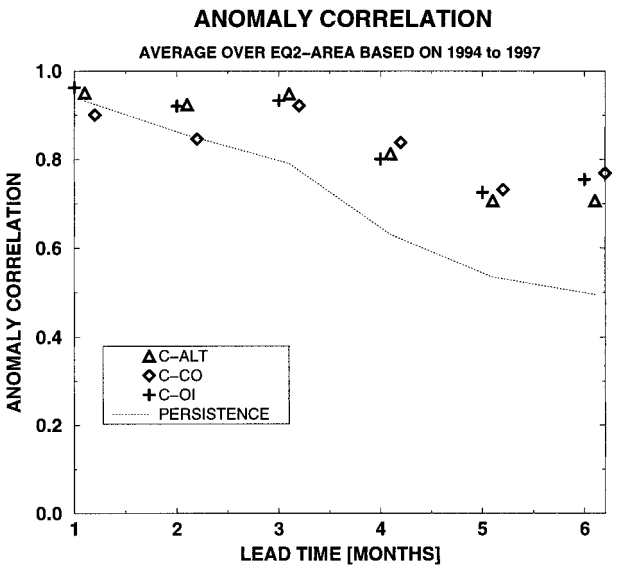


FIG. 13. Anomaly correlation coefficient as in Fig. 12b but based on the years 1994–97 only.

the regions Niño-12, Niño-3, and Niño-4, and in Fig. 14b for the regions EQ-1, EQ-2, and EQ-3. Solid bars indicate the forecasts compared to C-CO, and empty bars the forecasts compared to C-OI. A bar in the pos-

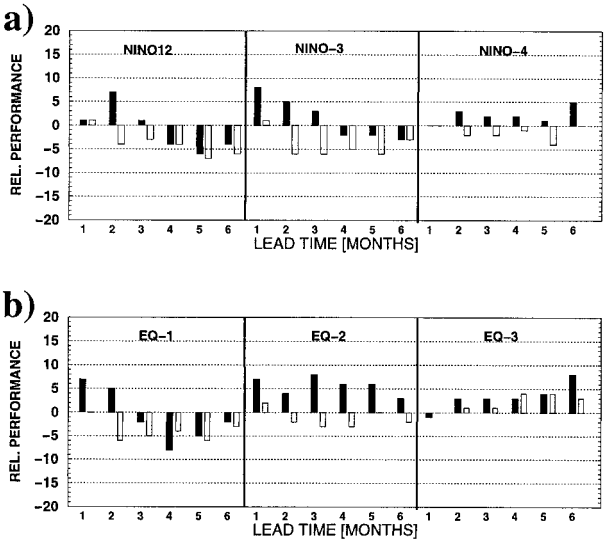


FIG. 14. Relative performance of C-ALT compared with C-CO (solid bars) and compared with C-OI (empty bars) as a function of lead time, averaged over (a) Niño-12 (10°S–0°N, 80°–90°W), Niño-3 (as all following regions 5°S–5°N, 90°–150°W) and Niño-4 (150°W–160°E), and (b) EQ-1 (90°–130°W), EQ-2 (130°–170°W), and EQ-3 (170°W–150°E). C-ALT is defined better when the predicted SST anomaly is more than 0.2°C closer to the observed SST anomaly than that of C-CO or C-OI, and defined worse when the error is larger by more than 0.2°C. Shown is the number of better minus the number of worse forecasts. Bars in the positive range mean that C-ALT performed on average better, bars in the negative range that it performed worse.

itive range means that C-ALT performs relatively better and a bar in the negative range that it has performed worse. Each bar is based on 20 forecasts.

The performance of C-ALT is quite satisfactory in the sense that forecasts are on average better than the control forecasts in Niño-4, EQ-2, and EQ-3. In Niño-3 and EQ-2, and the coastal region Niño-12 there is an improvement at shorter lead times of up to 3 months, but a deterioration for longer leads. Comparisons with the forecasts using in situ temperatures are less favorable. Only in EQ-3 are the forecasts using altimeter data more skillful than the ones that use in situ temperatures.

In a third approach we compare predicted and observed SST anomaly averaged over 3 month periods. The drift-corrected predicted SST anomalies from the individual ensemble members, averaged over the Niño-3 region and over the first 3 months of the forecasts, are shown in Fig. 15a and the average over the second 3 months is shown in Fig. 15b. Figures 15a,b give an idea of the spread of the ensemble members. Three-month averages of the ensemble means are shown in Figs. 15c,d. The solid line in Fig. 15 shows the 3-month running mean of observed SST anomalies. As described above, observed anomalies are adjusted to the 1950–79 long-term climatology and therefore do not average to zero over the period 1993–97.

In general when averaged over the first 3 months, the forecasts predict the observed SST anomalies to first order (Fig. 15a). The errors are considerably larger for 4–6 month lead times where errors are on the order of 2°C for the worst members of the control forecasts even when observed SST anomalies imply normal conditions (Fig. 15b). One striking feature is that the sharp rise of SST anomalies during 1997 is underpredicted in all experiments. The event was predicted better by the operational forecast system at ECMWF, which is very similar to C-OI, but the operational system has more ensemble members, forecasts were started every month, and the model drift is based on a longer period than for the experiments shown here. Figure 16 shows similar results but for Niño-4. Errors of the forecast SST anomalies for short lead times are much smaller than for Niño-3. For longer lead times, however, the errors are of similar magnitude as for Niño-3. As the amplitude of the SST anomalies is much smaller in the Niño-4 region, this implies a larger relative error.

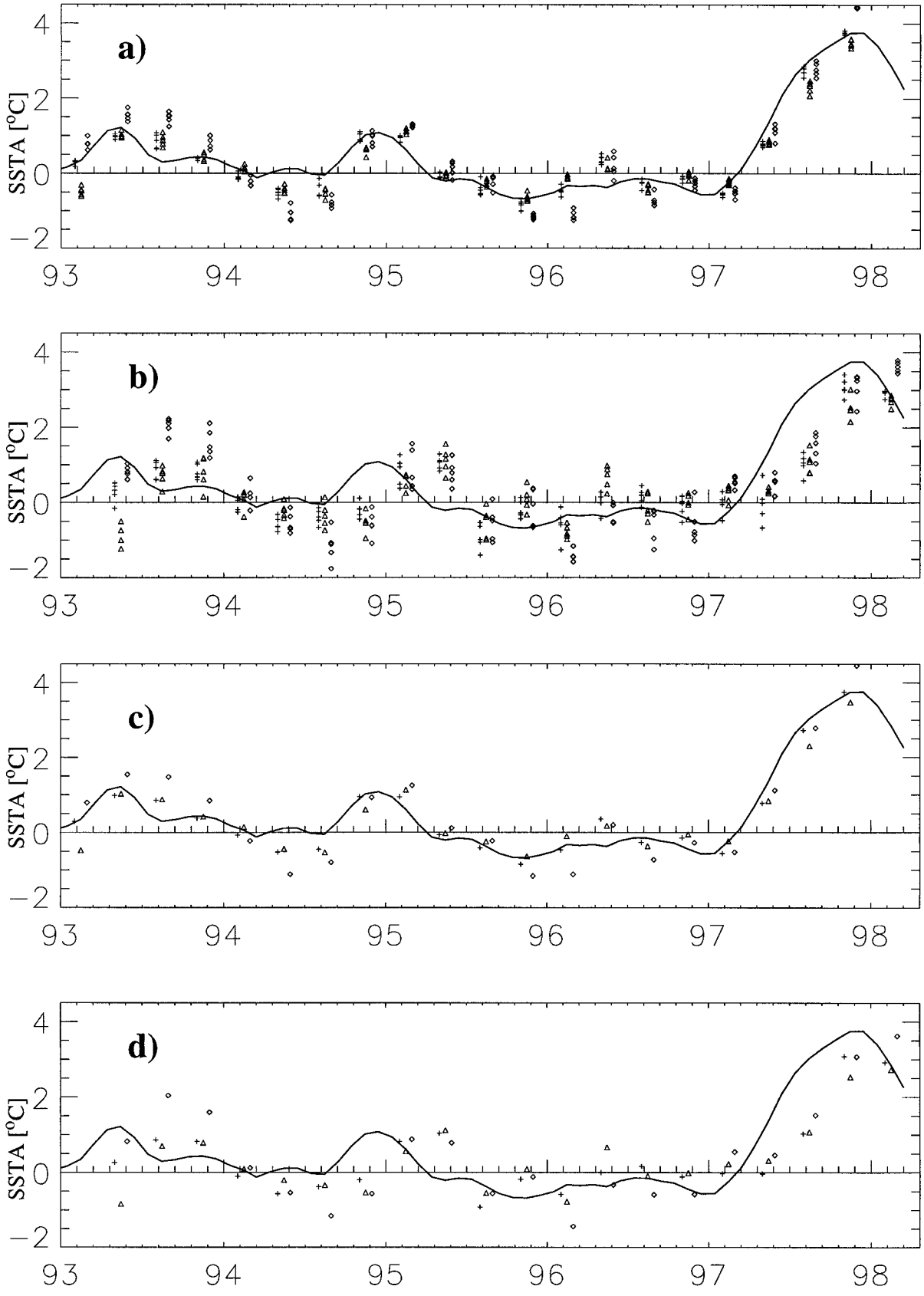
Based on the data shown in Figs. 15 and 16, we finally evaluate the forecasts using the concept of ''false alarm,'' and ''missed'' and ''hit'' events. The method takes into account whether a forecast agrees with observations when extreme events occur. First, we define an ''event'' as a 3-month average SST anomaly in excess of ±1°C in either Niño-3 or Niño-4. For the period covered, this results in a total number of 21 events (i.e., the sum over both Niño regions and lead times of 1–3 and 4–6 months). A false alarm is defined as a forecast error of more than ±1°C, if no event is observed. A hit is defined as a prediction of an observed event within 0.2°C. An event is declared missed if an observed event is underpredicted by more than 1°C.

The following numbers are based on the ensemble averages shown in Figs. 15c,d and 16c,d. The hit rate, combined from the Niño-3 and Niño-4 areas and 1–3 and 4–6-month lead time forecasts, is 8:10:5 (C-CO: C-OI:C-ALT), the false alarm rate is 12:3:2, and the missed events distribute as 3:3:3. While the hit rate of the control forecasts is relatively high, there are many false alarms. The number of false alarms is effectively reduced in both C-OI and C-ALT. Perhaps one of the larger benefits of the assimilation of observations is that it leads to a reduction of false alarms. This reduction works equally well for the assimilation of altimeter data and for the assimilation of subsurface temperatures. Missed events occur only for the 4–6 month averages of forecast Niño-3 SST anomaly. The rise of temperatures in late 1994 and during 1997 is underpredicted in all experiments, but in particular in C-ALT, which is somewhat disappointing.

However, as was pointed out above, seasonal predictions are of a probabilistic nature, and therefore the underprediction of the SST anomalies in the Niño-3 area during 1997–98 does not mean that altimetry is of no benefit to seasonal forecasting. Furthermore, our control forecasts already have information about the observed SST. It is then only the additional information at subsurface levels that potentially improves the forecast skill. One principle problem for the evaluation of the impact from altimetry is that we are confined to a limited number of ensemble members, due to the usual limitations of computer time. The second problem is the limited period for which altimeter data are provided. Although, from a historical point of view, TOPEX/Poseidon and *ERS-1/2* have been providing sea level observations for quite a long period, this period is still too short to properly evaluate the impact of the altimeter assimilation on the forecasts. Because the period of the ENSO cycle is several years, forecasts that cover several decades might be required to capture a sufficient number of extreme states of the ENSO cycle.

→

FIG. 15. Predicted SST anomalies averaged over the Niño-3 area for (a) time average over the first 3 months of the coupled forecasts and (b) average over months 4–6. Diamonds indicate SST anomalies from C-CO, plus-signs SST anomalies from experiment C-OI, and triangles indicate SST anomalies from experiment C-ALT. For better readability, the plus signs are offset by −14 days, and the diamonds by +14 days. Forecast are started every 3 months from 1 Jan 1993 to 1 Oct 1997. Ensembles for the forecasts are created by perturbing the oceanic initial conditions. The observed SST anomaly, smoothed by a 3-month running mean, is shown as solid line.



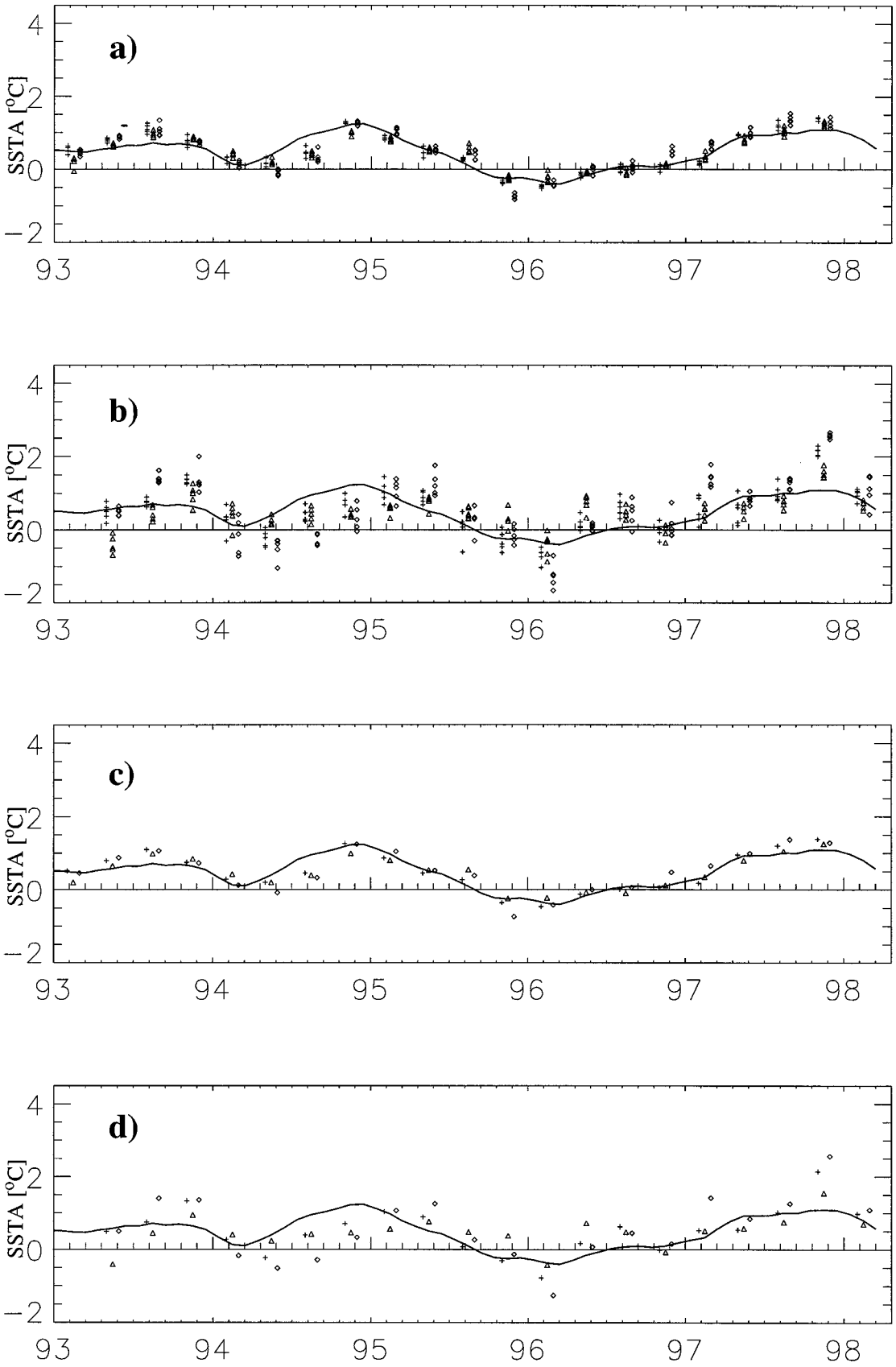


FIG. 16. As Fig. 15 but for the Niño-4 area.

7. Summary and conclusions

By assimilating sea level anomalies only, it was possible to replicate equatorial Pacific upper-ocean heat content variations of an ocean analysis in which subsurface temperature observations were assimilated, whereas the control experiment had large errors. However, the altimeter analysis turned out to be sensitive to initial conditions and the model-derived mean state of sea level that was added to the observed anomalies. Constraining the subsurface density field by relaxation to climatological subsurface salinity was required to prevent the analysis from drifting. Initial oceanic conditions and mean sea level were taken from an analysis in which subsurface temperature data were assimilated, and in this sense the analysis in which altimeter data were assimilated was not completely independent of the TAO–XBT observation network.

Ensembles of coupled forecasts were initialized from the ocean analysis in which altimeter data were assimilated, from an ocean analysis in which subsurface temperature data were assimilated, and from an ocean analysis in which neither altimeter nor subsurface temperature data were used. Based on 100 coupled forecasts for each set of experiments, we showed that differences between the forecasts started from the ocean analysis in which altimeter data were assimilated and the forecasts started from the ocean analysis using subsurface temperatures were not statistically significant, but the differences with a control set in which data assimilation was not used were highly significant at greater than 99%.

Economic considerations come into play when the quality of a forecast method has to be estimated by its financial value. For example, in order to avoid the waste of resources for unnecessary preventive measures, it may turn out to be just as important not to raise false alarms when there is no event than to predict an event correctly. In that sense the reduction of false alarms by assimilation of altimeter data is also an achievement because it results in an increased reliability of the forecasts.

In this paper we have shown how to use altimeter data to good effect, in the absence of subsurface temperatures. We are now extending our results to the assimilation of both sea level and subsurface temperature observations.

Acknowledgments. We wish to thank M. McPhaden and D. McClurg for providing the dynamic height anomalies, C. Boone for help with the altimeter data, M. Balmaseda for assistance with the ocean model, and K. Haines for fruitful discussions. O. Alves performed experiments CO-1 to OI-3, C-CO, and C-OI. This work has been supported by the European Union Environment and Climate project DUACS (ENV4-CT96-0357). The altimeter products have been produced by the CLS Space Oceanography Division as part of the European Union Environment and Climate project AGORA (ENV4-CT9560113) and DUACS (ENV4-CT96-0357). Part of the computations were carried out on a VPP-300 donated by Fujitsu Limited. The ocean model was provided by the Max-Planck-Institut für Meteorologie, Hamburg; ocean data assimilation software by the Bureau of Meteorology Research Centre, Melbourne; and coupling software by CERFACS, Toulouse.

REFERENCES

Alves, J. O. S., K. Haines, and D. L. T. Anderson, 2000: Sea level assimilation experiments in the tropical Pacific. *J. Phys. Oceanogr.,* in press.

Carton, J. A., B. S. Giese, X. Cao, and L. Miller, 1996: Impact of altimeter, thermistor and expendable bathythermograph data on retrospective analyses of the tropical Pacific Ocean. *J. Geophys. Res.,* **101,** 14 147–14 159.

Chambers, D. P., R. H. Stewart, and B. D. Tapley, 1998: Measuring heat storage changes in the equatorial Pacific: A comparison between TOPEX altimetry and Tropical Atmosphere–Ocean buoys. *J. Geophys. Res.,* **103,** 18 591–18 597.

Chelton, D. B., and M. G. Schlax, 1994: The resolution capability of an irregularly sampled dataset: With application to Geosat altimeter data. *J. Atmos. Oceanic Technol.,* **11,** 534–550.

Chen, D., M. A. Cane, S. E. Zebiak, and A. Kaplan, 1998: The impact of sea level assimilation on the Lamont model prediction of the 1997/98 El Niño. *Geophys. Res. Lett.,* **25,** 2837–2840.

Cooper, M., and K. Haines, 1996: Altimetric assimilation with water property conservation. *J. Geophys. Res.,* **101,** 1059–1077.

Fischer, M., M. Flügel, M. Ji, and M. Latif, 1997: The impact of data assimilation on ENSO simulations and predictions. *Mon. Wea Rev.,* **125,** 819–829.

Helland-Hansen, B., and F. Nansen, 1916: *Temperature Variations of the North Atlantic and in the Atmosphere* (in German). Videnskapsselskapets Skrivter. I. Mat.-Naturv. Klasse, Vol. 9, Kristiania, 341 pp.

Ji, M., D. W. Behringer, and A. Leetma, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part II: The coupled model. *Mon. Wea. Rev.,* **126,** 1022–1034.

——, R. W. Reynolds, and D. W. Behringer, 2000: Use of TOPEX/Poseidon sea level data for ocean analyses and ENSO prediction: Some early results. *J. Climate,* **13,** 216–231.

Le Traon, P. Y., P. Gaspar, F. Bouyssel, and H. Makhmara, 1995: Using TOPEX/Poseidon data to enhance *ERS-1* orbit. *J. Atmos. Oceanic Technol.,* **12,** 161–170.

——, F. Nadal, and N. Ducet, 1998: An improved mapping method of multisatellite altimeter data. *J. Atmos. Oceanic Technol.,* **15,** 522–534.

Levitus, S., and T. P. Boyer, 1994: *Temperature*. Vol. 4, *World Ocean Atlas 1994*, NOAA NESDIS, 129 pp.

——, R. Burgett, and T. P. Boyer, 1994: *Salinity*. Vol. 3, *World Ocean Atlas 1994*, NOAA NESDIS, 111 pp.

McPhaden, M. J., 1995: The Tropical Atmosphere Ocean array is completed. *Bull. Amer. Meteor. Soc.,* **76,** 739–741.

Mitchum, G., 1994: Comparison of TOPEX sea surface heights and tide gauge sea levels. *J. Geophys. Res.,* **99,** 24 541–24 553.

Oschlies, A., and J. Willebrand, 1996: Assimilation of Geosat altimeter data into an eddy-resolving primitive equation model of the North Atlantic Ocean. *J. Geophys. Res.,* **101,** 14 175–14 190.

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

Segschneider, J., J. Alves, D. L. T. Anderson, M. Balmaseda, and T. N. Stockdale, 1999: Assimilation of TOPEX/Poseidon data into a seasonal forecast system. *Phys. Chem. Earth (A),* **24,** 369–374.

Smith, N. R., 1995: SIANAL—A statistical interpolation routine.

Documentation, BMRC, Melbourne, Australia, 35 pp. [Available from BMRC, Box 1289K, Melbourne, Victoria 3001, Australia.]

——, J. E. Blomley, and G. Meyers, 1991: A univariate statistical interpolation scheme for subsurface thermal analyses in the tropical oceans. *Progress in Oceanography*, Vol. 28. Pergamon Press, 219–256.

Stammer, D., and C. Wunsch, 1996: The determination of the large-scale circulation of the Pacific Ocean from satellite altimetry using model Green's functions. *J. Geophys. Res.*, **101** (C8), 18 409–18 432.

Stockdale, T. N., 1997: Coupled ocean–atmosphere forecasts in the presence of climate drift. *Mon. Wea. Rev.*, **125,** 809–818.

——, D. L. T. Anderson, J. Alves, and M. Balmaseda, 1998: Seasonal rainfall forecasts with a coupled ocean–atmosphere model. *Nature*, **392,** 370–373.

Tapley, B. D., and Coauthors, 1996: The Joint Gravity Model 3. *J. Geophys. Res.*, **101,** 28 029–28 049.

Vialard, J., and P. Delecluse, 1998: An OGCM study for the TOGA decade. Part I: Role of salinity in the physics of the western Pacific fresh pool. *J. Phys. Oceanogr.*, **28,** 1071–1088.

Vossepoel, F. C., R. W. Reynolds, and L. Miller, 1999: Use of sea level observations to estimate salinity variability in the tropical Pacific. *J. Atmos. Oceanic Technol.*, **16,** 1401–1415.

Weaver, A. T., and D. L. T. Anderson, 1997: Variational assimilation of altimeter data in a multilayer model of the tropical Pacific Ocean. *J. Phys. Oceanogr.*, **27,** 664–682.

Wolff, J. O., E. Maier-Reimer, and S. Legutke, 1997: The Hamburg Ocean Primitive Equation Model. Deutsches Klimarechenzentrum Tech. Rep. 13, Hamburg, Germany, 98 pp. [Available from DKRZ, Bundesstr. 55, D 20146 Hamburg, Germany.]

Wyrtki, K., 1985: Water displacements in the Pacific and the genesis of El Niño cycles. *J. Geophys. Res.*, **90,** 7129–7132.