Forecast model bias correction in ocean data assimilation

Gennady A. Chepurin, James A. Carton, and Dick Dee

(submitted to Monthly Weather Review)

January 7, 2004

Department of Meteorology, University of Maryland,
College Park, Maryland, 200742 USA,
Chepurin@atmos.umd.edu

\(^1\)Data Assimilation Office, NASA/Goddard Space Flight Center,
Greenbelt, MD
Abstract

Numerical models of ocean circulation are subject to systematic errors resulting from errors in model physics, numerics, inaccurately specified initial conditions, and errors in surface forcing. In addition to a time-mean component the systematic errors include components that are time varying, that could result for example from inaccuracies in the time varying forcing. Despite their importance, most assimilation algorithms incorrectly assume that the forecast model is unbiased. In this paper we characterize the bias for a current assimilation scheme in the tropical Pacific. The characterization is used to show how relatively simple empirical bias forecast models may be used in a two-stage bias correction procedure to improve the quality of the analysis.
1. Introduction

Sequential ocean data assimilation combines a forecast produced by a general ocean circulation model with observations to construct an improved estimate of the state of the ocean. The optimality of such a procedure depends on the validity of statistical assumptions about the errors of the observations and the forecasts. One critical assumption is that the error in the forecast is random, with no systematic errors. Unfortunately, studies reveal that current ocean models do have large systematic errors. These errors, which we refer to collectively as bias, are introduced through a variety of mechanisms including inaccurate initial conditions, systematic errors in surface forcing, parameterizations of unresolved physics, or numerics, or through nonlinear processes when forced by random errors. In this paper we examine the structure of bias in a current assimilation system and explore a method designed to account for it during the data assimilation cycle.

The true bias is the difference between the forecast of the ocean state, $\omega^f$ (consisting of the gridpoint values of temperature, salinity, velocity, and pressure), the unknown true state $\omega^r$, and the random component of the forecast error $\varepsilon^f$, that is, the part of the forecast that has zero expectation:

$$\beta^l = \omega^f - \omega^r - \varepsilon^f, \quad \langle \varepsilon^f \rangle = 0.$$  (1)
Thus the true bias may be determined by taking the expectation of the difference between the state forecast and the true state. If the expectation operator is approximated by the time-mean or annual harmonic operators we obtain approximate estimates of the time-mean or annual cycle of the bias, and so forth.

State forecasts are produced using numerical ocean general circulation models. Thus, we can anticipate the characteristics of bias by reviewing ocean model simulations. Among the recent studies, Gent et al. (1998) document problems with time mean deep and intermediate water mass formation rates as well as excessive diffusion of the thermocline, a problem that seems characteristic of z-coordinate models. Smith et al. (2000) show that despite very high 1/10x1/10 degree horizontal resolution, some differences in the variability of the western boundary currents and their mean position remain. The results of Sun and Bleck (2001) indicate that problems with water mass formation/penetration remain despite use of isopycnic vertical coordinates. Seasonality in the bias is evident in the study of Tandon and Zahariev (2001) where it appears to be caused by unresolved diurnal variations in surface forcing. In the tropical Pacific sea surface temperatures in the eastern basin are generally too low and the thermocline is too high (Stockdale et al., 1998).

The problem of model bias has been widely recognized in the meteorological community, and before that in the engineering literature. As a result, a number of approaches have been developed. The most straightforward approaches to handling time-mean bias involve computing model or analysis climatologies and then introducing
correction terms into the equations of motion (e.g. Saha, 1992). To correct for rapidly changing bias in a data-rich environment, a second class of approaches has been proposed involving examination of the past few updating cycles for a systematic tendency, which is then corrected (Thiebaux and Morone, 1997; DelSole and Hou, 1999; D’Andrea and Vautard, 2000; Griffith and Nichols 1996, 2000).

This second class of approaches has the advantage that they allow the bias estimate to evolve in time as would be expected, for example, if the bias is being advected by the time changing large-scale flow field. These approaches have the disadvantage that they do not retain long-term memory and thus cannot fully account for predictable biases with longer time-scales such as those linked to the annual cycle. A third class of approaches useful in linear one-dimensional problems involves pre-whitening the errors so that their frequency spectrum resembles white noise (Kamen and Su, 1999).

In this paper we explore a fourth class of approaches we refer to as ‘two-stage estimation’, which was first introduced by Friedland (1969) (see also Mendel 1976, and Ignagni 1981). The two-stage estimation algorithm begins with the assumption that a reasonable estimate of the bias may be made prior to estimating the state of the system itself, thus allowing the estimation procedures for the bias and the state to be carried out successively. In its original form this bias was assumed to be steady. Ignagni (1990) expanded the bias model to allow for time-dependent bias (a problem considered earlier by Thacker and Lee 1972). Other authors (Mendel 1976; Zhou et al. 1993) have proposed
modifications to the two-stage estimation algorithm designed to handle nonlinear state models. Application to atmospheric data assimilation has been explored by Dee and da Silva (1998) and Dee and Todling (2000).

Here we apply the two-stage estimation algorithm of Dee and da Silva (1998) and Dee and Todling (2000) to the problem of sequential data assimilation in the ocean. We focus on the tropical Pacific sector (30S-30N) because of the strength of its interannual variability and because of its importance in climate research. Previous efforts to use the Dee and da Silva algorithm for the ocean (e.g. Carton et al. 2000a,b; Martin et al., 2002) have assumed a very simple forecast model for the bias. But as suggested above, systemic errors in ocean models are actually geographically oriented, with temporally varying structure resulting from errors in identifiable phenomena such as thermocline water mass formation or mixed layer entrainment. Some errors are persistent, others are cyclic, and still others are noncyclic but predictable. Thus our approach will need to account for each of these.

2. Bias correction algorithm

We represent the state forecast $\omega^f$ at time $t_k$ as a column vector containing the $N$ state variables. It is produced by forward integration of the ocean model $\Omega$ beginning from the analysis at time $t_{k-1}$

$$\omega^f_k = \Omega(\omega^a_{k-1})$$  

(2)
The bias forecast estimate is similarly the sum of the true bias and a zero expectation random error $\eta^f$, both column vectors of length N:

$$\beta^f = \beta + \eta, \quad \langle \eta \rangle = 0$$

(we drop the subscript k throughout except where it is needed for clarity) and is produced by the (linear) bias model $B$ beginning from the bias analysis at time $t_{k-1}$. A general linear model for the bias would look like:

$$\beta^f_k = B \beta^a_{k-1} + Z_k$$

where Dee and da Silva (1998) assume $B$ is diagonal and $Z_k = 0$. In this paper we explore models with nonzero $Z$ assuming $B = 0$.

Suppose now that at time $t_k$ we have M unbiased observations contained in a column vector $\omega^o$. This column vector may be decomposed into the true values interpolated to the observation locations plus an unbiased error $\epsilon^o$ associated with each observation

$$\omega^o = H \omega^f + \epsilon^o, \quad \langle \epsilon^o \rangle = 0 .$$
Here $H$ is the $M \times N$ interpolation matrix that maps variables specified at the model locations onto the observation locations. We would like to provide an analysis of the state as well as the bias based on a linear combination of the same set of $\omega^a$ and an unbiased forecast $\tilde{\omega}^f = \omega^f - \beta^a$. Dee and da Silva (1998) and Dee and Todling (2000) propose to compute the analyses in two stages:

$$
\beta^a = \beta^f - L \left[ \omega^o - H(\omega^f - \beta^f) \right], \tag{6a}
$$

$$
\omega^a = \tilde{\omega}^f + K \left[ \omega^o - H\tilde{\omega}^f \right] \tag{6b}
$$

using successively improved estimates of the unbiased state forecast. Thus, $\omega^a$ is an unbiased estimate of $\omega^f$. Here $K$ and $L$ are $N \times M$ gain matrices which account for the relative accuracies of the model forecast and the observations. Minimization of the mean square errors of the state and bias analyses under the assumptions that the observation error, state forecast error, and bias forecast error are mutually independent (e.g. $\langle \epsilon^f \eta^f \rangle = \langle \eta^f \epsilon^f \rangle = 0$) leads to the following equations for the gain matrices (Dee and da Silva, 1998):

$$
K = P_{\omega}^f H^T (H P_{\omega}^f H^T + R^o)^{-1} \tag{7a}
$$

$$
L = P_{\beta}^f H^T (H P_{\omega}^f H^T + H P_{\beta}^f H^T + R^o)^{-1} \tag{7b}
$$
The observation error covariance is $R^o \equiv \left\langle \epsilon^o (\epsilon^o)^T \right\rangle$.

To complete this analysis we need to prescribe the bias error covariance $P^f_\beta \equiv \left\langle \eta^f (\eta^f)^T \right\rangle$ and the unbiased forecast error covariance $P^f_\omega \equiv \left\langle \epsilon^f (\epsilon^f)^T \right\rangle$. We assume that the bias error covariance has horizontal scales similar to those of the basin and may be geographically oriented, while the unbiased forecast error has horizontal scales of a few hundred kilometers (e.g. see the analysis of Carton, et al., 2000a) and is roughly homogeneous. Following Dee and Todling (2000) the bias-corrected observation-minus-forecast differences are:

$$v = \tilde{v} + v' = \omega^0 - H(\omega^f - \beta \tilde{f}^f) = \epsilon^o - H(\epsilon^f - \eta \tilde{f}^f),$$

where we introduce $\tilde{v}$, and $v'$ to represent the basin- and small-scale components of the observation-minus-forecast differences. If we neglect the covariance between large and small scales of observed-minus-forecast differences ($\left\langle \tilde{v} v'^T \right\rangle = \left\langle v' \tilde{v}^T \right\rangle = 0$), then the bias-corrected observation-minus-forecast covariance is given by

$$\left\langle v v'^T \right\rangle + \left\langle v' v'^T \right\rangle = R^o + HP^f_\omega H^T + HP^f_\beta H^T.$$  

As a result of the assumption of the independence of basin- and small-scales we can separate (9) into two relations:
Equations (10) allow us to calculate $P^f_\omega$ and $P^f_\beta$. Based on 30 years of forecast-minus-observation differences we calculate the random forecast error covariance $P^f_\omega$ by fitting a model that decays exponentially as a function of separation in latitude, longitude, and time, and applying minimum variance estimation (see Carton, et. al. 2000a for details). The resulting spatial scales of exponential decay are smaller than 500 km. We assume that all covariance at scales greater than 500 km is due to bias.

We estimate the covariance $P^f_\beta$ of the bias forecast error by first binning the forecast-minus-observation differences into 5x5 degree bins (thus filtering out the random forecast error). Then we compute $P^f_\beta$ in this reduced space. In fact, it will be shown in the following section that $P^f_\beta$ is dominated by a few basin-scale structures and thus we will approximate it by a small number of principal components. Additional experiments not reported here show the results to be insensitive to the precise specification of the covariances in (10).

Next we turn our attention to the bias model. Consistent with the observation that $P^f_\beta$ is dominated by a few basin-scale structures we assume that the bias forecast model
and bias analysis can be represented by the product of truncated set of Empirical Orthogonal Functions $G$ and principal components $\tau$:

$$\beta' = G \cdot \tau'$$  \hspace{1cm} (11)

$$\beta^a = G \cdot \tau^a$$

where $G$ is a matrix of size $N_x Q$ containing the first $Q$ EOFs representing the spatial structures and $\tau' \beta'$ is a vector of size $Q$ containing the principal component coefficients at time $t_k$.

Since the columns of $G$ are orthonormal we can compute the bias analysis time series $\tau^a$ by multiplying both sides of (6a) by $G^T$ (dropping the time subscript $k$ as usual for convenience):

$$\tau^a = G^T \beta^a = \tau' - G^T L [\omega^o - H \hat{\omega} f]$$ \hspace{1cm} (12)

and then

$$G^T L = P^f [H G]^T \left( H G P^f [H G]^T + H G P^f [H G]^T + R^o \right)^{-1}$$ \hspace{1cm} (13)

where \( P^f = G^T P^f G \), \( P^f = G^T P^f G \).
The equations (2), (4), (6), and (7) together with the specification of the bias model (11-13) and truncated of $G$, represent our complete set of assimilation equations. In order to develop a reasonable bias model we now begin examination of the bias in a current assimilation scheme.

3. Bias

Here we examine the bias as it appears in a current ocean data assimilation system, described below. Our analysis procedure is directed toward identifying features in the bias field that have broad spatial scales and long temporal scales because of their relevance for climate estimates and because these features are more statistically stable. As a result of this examination we will propose a model of bias in Section 4.

The data assimilation analyses rely on the Simple Ocean Data Assimilation methodology of Carton et al. (2000a,b) which uses a forward model using Geophysical Fluid Dynamics Modular Ocean Model IIb (MOM-2) numerics with $1^\circ \times 1/2^\circ \times 20$-level resolution near the equator expanding to $1^\circ \times 1^\circ$ resolution in midlatitudes. All experiments span the 31-year period 1970-2000 with initial conditions provided by climatological temperature and salinity. Our analysis is limited to the last 30 years to reduce the impact of the initial conditions. Sponge layers are inserted poleward of 62S and 62N that relax temperature and salinity to their climatological monthly values. Wind stress is provided by the monthly observation-based analysis of da Silva (1994) for the years prior to 1991.
Winds for the period 1991-2001 are provided by a combination of the National Centers for Environmental Prediction monthly anomalies added to the seasonal cycle provided from the da Silva winds in order to minimize the shock introduced by the change in wind analyses.

The basic subsurface data set in the tropical Pacific consists of approximately $1.6 \times 10^6$ temperature profiles, of which two-thirds were obtained from the World Ocean Database 2001 (Boyer et al., 2002; Stephens et al., 2001) and extended by operational temperature profile observations from the National Oceanographic Data Center\NOAA temperature archive and including observations from the TAO/Triton mooring thermistor array. The profile data is concentrated along commercial shipping lanes in the far eastern and western basins. SST observations were obtained from the COADS surface marine observation set.

The analysis procedure solves equations (6)-(7) at 10-day intervals. Updating is carried out using a form of digital filtering introduced by Bloom et al. (1996). A set of five experiments is carried out listed in Table 1, differing only in the bias forecast model (4). We first discuss results from the control experiment (Expt. 1) in which there is no correction for bias and this experiment does not account basin-scale errors. Although all state variables are available, we focus our discussion on the $30 \times 36 = 1080$ 10-day fields of analysis and forecast temperature. Bias is determined by looking for systematic components of the objectively gridded differences between the model forecast and the observations. Our analysis focuses on mixed layer temperature because of its importance
in influencing the atmosphere and the depth and width of the thermocline because they reflect the accuracy with which the oceanic heat storage is represented.

We begin by examining the time-mean bias. The time-mean component dominates the bias in the mixed layer (0-45m). It explains 44% of the variance averaged geographically and in time within the mixed layer. Below the mixed layer the variance explained by the time-mean component decreases with increasing depth until the depth of the thermocline where a second maximum occurs. Along the equator the forecast mixed layer is too cold in the east by up to 0.2°C (Fig. 1 upper panel). The thermocline below this is too sharp, as indicated by the warm bias in the upper thermocline. In contrast, in the central basin the upper thermocline has a cold bias indicating that the thermocline is too shallow and broad.

It is evident in Fig. 1 that the bias has different behavior in the mixed layer and the thermocline. Because of this difference we will carry out separate analyses of these, treating the mixed layer as a slab of uniform 45m depth. The thermocline depth varies much more widely throughout the basin. Here we approximate the thermocline depth as the depth of the 20°C isotherm. It is evident from the discussion above that bias includes errors in the width as well as depth of the thermocline. We estimate the width of the thermocline as the change in depth spanned by the 22°C and 14°C isotherms.

The geographic distribution of the time-mean bias is shown in Fig. 2. Time-mean bias in the mixed layer is mainly confined to equatorial latitudes and is nearly symmetric
about the equator with a maximum negative anomaly of -0.2°C. In contrast to the mixed layer, the time-mean bias in thermocline depth extends broadly into the southern hemisphere, indicating that the thermocline is 2m too deep in the east and 4m too shallow in the central basin (again, evaluated over a 10-day update cycle), while the thermocline is 10 - 20m too wide throughout the western tropics (±10°).

Examination of the evolving bias reveals that there is a significant annual component as well. This is particularly evident in the Northern Hemisphere mixed layer between 10N – 30N where the monthly forecast-minus-observation differences are strongly negative in June (0.3C) and strongly positive in December (0.4C) even though the annual mean bias averaged over this band of latitudes is small (<|0.1C|, Fig.3). Interestingly, the distributions of forecast-minus-observation differences are also skewed, with the skewness also changing sign seasonally from –1.3 to 1.8 indicating a larger negative tail in June and a larger positive tail in December. The difference distributions also have larger tails in both directions than would be expected for a Gaussian distribution (kurtosis is 8 and 11 respectively).

We evaluate the spatial structure of the annual component by computing the annual Fourier harmonic of the forecast-minus-observation differences binned in 5° × 5° bins (binning is required to compute statistics), evaluated over the 30-year record. The amplitude and phase diagrams (Fig. 4) reveal that in the northern subtropics the maximum cold bias of up to 0.1C occurs in spring and a corresponding warm bias occurs in fall, a time of year when the mixed layer reaches its warmest (the annual cycle of bias
is roughly 25% of the annual cycle of mixed layer temperature). Along the equator the reduction in the annual component of the bias reflects a reduction in the annual cycle of mixed layer temperature (the annual cycle of mixed layer temperature along the equator is less than 0.05°C west of the date line).

The amplitude of the annual component of bias decreases with increasing depth and with decreasing latitude (partly reflecting the decrease with latitude of the annual harmonic of SST itself). In the latitudes of the North Equatorial Countercurrent (5N-15N) the phase of the annual component of bias changes giving a warm bias in spring when latent heat loss associated with intensification of the trade winds should be causing mixed layer temperatures to drop to their annual minimum.

In contrast to the mixed layer, the annual component of bias in thermocline depth is fairly uniform with latitude, ranging from 1-4 meters, with somewhat higher values in the subtropics. Along the equator where the annual component has a weak local maximum, there is evidence that the annual bias propagates slowly eastward in the western basin (propagation is evident in the alternating pattern of phase shift in Fig. 4 along the equator). The annual component of thermocline width bias is greater than 5 m throughout the western basin as well as in the subtropics. The phase of thermocline width bias is quite variable.

In addition to a time-mean cold bias and annual variations, bias at the equator also contains year-to-year variations both within the mixed layer and at thermocline depths.
(Fig. 5). Within the mixed layer the forecasts have a geographically stationary cold bias during the El Nino years, a result that explains much of the time-mean component. Within the thermocline warm and cold bias anomalies propagate slowly eastward in a way reminiscent of ENSO-induced thermocline anomalies. The anomalies indicate that the forecast thermocline underestimates the thermocline anomalies associated with both El Nino and La Nina.

In order to characterize the spatial structure of the interannual variability we conduct a principal component analysis of the three dimensional forecast-minus-observation differences. The two principal components (Fig. 6), which explain roughly 30% of the variance of the bias anomalies (after mean and annual signals have been removed), are primarily confined to the equatorial zone. Their spatial patterns are rather different from the corresponding principal components of the system state (see Chao and Philander, 1993). However, their time series, as well as the principal component time series of the system state are closely related to the Southern Oscillation Index (SOI). Indeed, the first principal component has a zero lag correlation with the SOI of 0.7. The second lags the SOI by up to one year. Together they describe the eastward propagation of bias evident in Fig. 5. The higher principal components are noisy and difficult to interpret. We suspect that much of the variance described by these higher principal components results from sampling error introduced by the sparse observing system.

4. Bias Modeling
The results presented above indicate that bias contains geographically oriented, but temporally varying components as well as a time-mean. The time-varying components are associated with the time-varying state. EOF analysis of forecast-minus-observation differences shows that two first principal components in (11) contain \( \sim 80\% \) of total bias variations and are associated with the annual cycle. The next two largest components reflect systematic errors in the way the forecast model represents ENSO. Based on these indications we propose the following bias model:

\[
\beta^f_k = \Theta + G_i \cdot \cos(\omega_A t_i) + G_i \cdot \sin(\omega_A t_i) + G_i \tau_3^f + G_i \tau_4^f
\]  

(14)

where \( \omega_A = 2\pi(1y)^{-1} \), and \( G_i \) is i-th EOF. We will explore this model through a series of assimilation experiments listed in Table 1.

In order to examine the successive impact of the components of the bias model (14) we present three additional experiments. Experiment 2 retains only the time-mean term in the bias model, Expt. 3 retains both the time-mean and the annual terms, while Expt. 4 retains all three terms (again, see Table 1).

We begin with Expt. 2 in which the bias model includes only the mean bias \( \bar{\beta} \). We find that nearly all of the time-mean bias along the equator is reduced to less than 0.1°C (compare Fig. 1 upper and lower panels). The reduction in bias is particularly noticeable in the mixed layer where the large equatorial cool bias disappears, as well as in the western basin where the warm bias weakens (compare Fig. 2). There is a weak cool
bias remaining. It remains we believe because of the spatial averaging required to map mean forecast-minus-observation differences. The mean bias in the depth of the thermocline is reduced to less than 2m throughout most of the basin, a reduction by more than a factor of two. The error in the thermocline width is reduced to 5m or less throughout most of the basin and the diffuse western equatorial thermocline evident in Fig. 2 is no longer present.

The middle panel in Fig. 1 shows the time-mean uncorrected forecast-minus-observation differences (the forecast before bias correction) along the equator for Expt. 2. It is interesting to note that the time-mean bias is not significantly reduced in comparison with the control experiment even though the initial conditions for the forecasts have been improved (compare Fig. 1 upper and middle panels). The implication from this finding is that the forecasts have some error growth that occurs very rapidly, in less than the 10-day interval between successive forecasts.

Improving the time-mean component of the bias does not significantly improve the time-varying components of the bias. Thus in order to improve the forecasts still further we consider Expt. 3 in which both the mean and annual cycle of the bias are corrected. The reduction in the monthly mean bias is dramatically illustrated by the December and June averages of forecast-minus-observation differences in the latitude band 10N-30N where the monthly mean bias is reduced by a factor of 4 - 5 (Fig. 3). Interestingly, the skewness of the histograms is also reduced by a factor of two.
The spatial pattern of the bias annual amplitude shows that the mixed layer bias is reduced to under 0.05°C almost everywhere (compare Figs. 4 and 7). The bias in thermocline depth and width is also reduced by a factor of 2 - 3. Interestingly, the phase of the annual cycle of thermocline bias for Expt. 3 resembles the phase of the control experiment indicating that we have under-corrected the bias somewhat.

In the fourth experiment we introduce interannual variability into our model of bias. We examine the results by comparing principal components of the observation-minus forecast differences for the control and Expt. 4 (compare Figs. 6 and 8). As in the control experiment, the thermocline expressions of the first two principal components for Expt. 4 have their maximum amplitude within ±10 latitude while the corresponding time series show that the forecast-minus-observation differences of the third and fourth principal components have been reduced by more than a factor of two, and thus the errors associated with the historical reproduction of ENSO events have been similarly reduced.

The reductions in bias may be compared by examining the root-mean-square (RMS) forecast-minus-observation differences for each bias model (Fig. 9). In the control experiment where the bias is neglected the RMS temperature differences are more than 0.5°C in the eastern equatorial mixed layer as well as in the subtropics between successive 10-day updates. RMS thermocline depth differences are also maximum on the equator, reaching values of 4m between successive 10-day updates.
Experiment 2 in which the time-mean bias is corrected reduces RMS mixed layer temperature differences by a factor of 2 along the equator with somewhat smaller reductions at higher latitude. Similar reductions are evident in RMS thermocline depth differences. Reducing the annual cycle of bias (Expt. 3) reduces mixed layer temperature differences in the subtropics where the annual cycle is large. In contrast, the improvement in the thermocline depth differences with this bias forecast model is limited.

The final bias forecast model we consider, used in Expt. 4, includes the two interannually varying unrotated principal components that primarily describe ENSO-related differences. Figure 9 shows that the inclusion of interannual variation in the bias forecast model has little impact on the RMS forecast-minus-observation differences in the mixed layer, while it reduces differences in thermocline depth by 30%. An additional experiment not listed in Table 1 shows these results to be insensitive to an increase in the number of principal components used.

Finally, we examine the impact of the bias analysis by comparing Expt. 4 with a fifth experiment in which the bias analysis $\beta^a$ is equal to the bias forecast for Expt. 4 (in other words rather than updating the bias estimates during the assimilation according to (6a), we simply use the bias forecast given by the model (14) fitted to the forecast-minus-observation differences of the control run). This corresponds to the use of fixed, or off-line, bias estimates for correcting the bias during assimilation. The results of this
experiment (Fig.10) show that this procedure is not as effective in reducing the RMS forecast-minus-observation differences as the full two-stage procedure.

5. Discussion

In this paper we examine the effects of bias in the forecast model on the error introduced in a 31-year long historical analysis of the physical state of the ocean, focusing on the mixed layer and thermocline depth in the tropical Pacific Ocean. We find that the biases are large and contain a variety of space and time scales including a time-mean, an annual cycle, and interannual variability linked to ENSO.

In the eastern equatorial mixed layer between successive 10-day updates root-mean-square temperature differences are more than 0.5°C. If these differences are interpreted as a heat flux error it would require a nonphysically large change in surface flux of approximately $100\, Wm^{-2}$ to correct. RMS thermocline depth differences are also maximum on the equator, but largest in the western basin, reaching RMS values of 8m between successive 10-day updates. RMS thermocline width differences are over twice as large as this.

In order to reduce the impact of bias on historical analyses we introduce a two-stage bias correction algorithm based on the ideas of Dee and da Silva (1998), modified to account for the presence of geographically oriented model bias. The efficacy of this procedure is explored in a set of three additional experiments examining successively
more complete empirical bias forecast models. We focus on the degree to which the bias-corrected forecast-minus-observation differences are reduced, a key measure to the improvement of the analysis.

The first bias forecast model we explore (Expt. 2) assumes a steady but spatially varying bias. The use of this model reduces mixed layer forecast-minus-observation differences in mixed layer temperature by a factor of 2 along the equator with somewhat smaller reductions at higher latitude. Similar reductions are evident in thermocline depth differences. Interestingly, the differences between the uncorrected forecast and observations are not strongly reduced, indicating that much of the bias develops during the first few days.

We next consider a bias forecast model that additionally includes an annual cycle (Expt. 3). The largest reductions in forecast-minus-observation differences are in the subtropical mixed layer where the annual cycle is large. In contrast, the improvement in thermocline depth differences is limited. The final bias forecast model we consider (Expt. 4) also includes the next two unrotated principal components of the forecast-minus-observation differences after the time-mean and annual cycle. Examination of these principal components shows that they primarily represent bias associated with the forecast model’s representation of ENSO. We find that relative to the results from Expt. 3, there is little reduction in the differences in the mixed layer, but a 30% reduction in the differences in thermocline depth because of an improvement in the representation of interannual variability of the thermocline depth. The cumulative effect of our most
sophisticated bias model is to reduce the corrected forecast minus observation differences by a factor of two over the control experiment.

To illustrate the impact of bias correction on the state analysis we compare the state analysis from Expt. 4 and the control experiment with independent velocity measurements. Zonal velocity on the equator at 110W in the eastern basin in the control run shows much too strong westward nearsurface currents and an Equatorial Undercurrent that is much too weak at 80m depth (Fig. 11). As a result of the introduction of the two-stage bias correction procedure, Expt. 4 shows much weaker (and thus more realistic) surface currents and a substantially stronger Equatorial Undercurrent with a realistic relaxation in the boreal summer of 1997.

The results of this study indicate that the forecast-minus-observation differences are significantly biased and that this bias reduces the effectiveness of ocean applications of data assimilation. In regions of high variability bias may explain up to half of the forecast-minus-observation differences. The two-stage correction algorithm explored here is successful in correcting much of this difference despite the simplicity of the bias forecast models considered. And these results are insensitive to the choices of assimilation parameters. The results apply directly to assimilation schemes relying on optimal interpolation or three-dimensional variational approaches. Further improvements through the use of dynamically based bias forecast models as well as reduction in the bias of the forecasts due to improvements in the models will help to reduce this problem in the future.
Acknowledgements

We are grateful to the National Science Foundation Information Technology Research program (OCE 0113148) for providing support.
References


Table caption

Table 1. Bias model experiments. Each experiment begins with the same initial conditions at January 1, 1970 and is carried out for 31 years. The bias models are described in Section 4.
Table 1. Bias model experiments. Each experiment begins with the same initial conditions at January 1, 1970 and is carried out for 31 years. The bias models are described in Section 4.

<table>
<thead>
<tr>
<th>Expt</th>
<th>Bias model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No bias model -- control</td>
</tr>
<tr>
<td>2</td>
<td>Time-mean bias: $\beta^f_k = \bar{\beta}$</td>
</tr>
<tr>
<td>3</td>
<td>Time-mean and annually varying bias: $\beta^f_k = \bar{\beta} + G_1 \cdot \cos(\omega_AT_1) + G_2 \cdot \sin(\omega_AT_1) \cdot \omega_A = 2\pi(1\text{yr})^{-1}$</td>
</tr>
<tr>
<td>4</td>
<td>Time-mean, annually varying, and interannual bias: $\beta^f_k = \bar{\beta} + G_1 \cdot \cos(\omega_AT_1) + G_2 \cdot \sin(\omega_AT_1) + G_3 \tau^f_3 + G_4 \tau^f_4$</td>
</tr>
<tr>
<td>5</td>
<td>Similar to Expt. 4 except that the bias analysis is computed off-line from the control run: $\beta^a_k = \beta^f_k$</td>
</tr>
</tbody>
</table>
Figure captions

**Figure 1.** Time-mean temperature bias in the equatorial Pacific (averaged 2.5S-2.5N) computed over the 30-year record for the control experiment (Expt. 1) with no bias correction (upper panel), and the mean bias correction experiment (Expt. 2, middle and lower panels). Middle panel shows the mean uncorrected forecast-minus-observations, while lower panel shows the bias-corrected forecast-minus-observation ($\bar{\omega}^f - \omega^o$). Depth of the 14C and 22C isotherms (obtained from the control experiment) are indicated by dashed lines. The width of these isotherms is used as a measure of thermocline width. Depth of the 20C isotherm is indicated by a bold line. Depth of this isotherm is used as a measure of thermocline depth. Shading indicates bias exceeding 0.1C. Positive values indicate the forecast has a warm bias.

**Figure 2.** Geographic structure of time-mean bias in control experiment (Expt. 1, lefthand panels) and a second experiment in which the mean bias is corrected (Expt. 2, righthand panel). Upper panel shows temperature bias in the mixed layer (0-45m). Positive values (solid contours) indicate that the mixed layer is too warm. Contour intervals are 0.025C. Middle panel shows bias in the depth of the thermocline as represented by the 20C isotherm. Positive bias indicates that the thermocline is too deep. Contour intervals are 2m. Lower panel shows bias in the width of the thermocline as indicated by errors in the relative depths of the 22C and 14C isotherms. Positive bias in
thermocline width indicates that the thermocline is too wide. Contour intervals are 5m. Units are C°, m and m.

**Figure 3.** Histogram of mixed layer temperature forecast-minus-observation differences in the latitude band 10N-30N for the months of June (dark) and December (light). Lefthand panel shows the control experiment (Expt. 1), while the righthand panel shows the experiment with both mean and annual bias correction (Expt. 3). For the control experiment the mean of the differences varies from -0.3C in June to 0.4C in December, while after correction for the annually varying bias the monthly mean differences are reduced to –0.04C and –0.1C respectively. The standard deviations are reduced by about 20%.

**Figure 4.** Amplitude (left) and phase (right) of the annual harmonic of the bias for the control experiment (Expt. 1). Upper panel shows the annual harmonic of mixed layer (0-45m) temperature bias. Middle panel shows the annual harmonic of thermocline depth bias. Lower panel shows the annual harmonic of thermocline width bias. Units for amplitude are C°, meters, and meters. Units for phase are degrees of lag relative to the calendar year.

**Figure 5.** Meridionally averaged tropical (10S-10N) bias as a function of longitude and time. Left-hand panel shows temperature at 100m. Right-hand panel shows mixed layer temperature. Contour interval is 0.5°C. Note that the bias contains a variety of scales including time-mean, annual cycle, as well as interannual variability.
**Figure 6.** Interannual variability of bias for the control experiment (Expt.1). Thermocline depth anomalies associated with the third, and fourth (upper and middle panels) three dimensional EOFs ranked in order of explained variance. Contour intervals are 0.5m. Corresponding principal component time series (lower panel). Between them, these two principal components describe 30% of the variance of the interannual variability of the bias (after mean and annual signals have been removed). The Southern Oscillation Index has a zero-lag correlation of 0.7 with the first principal component.

**Figure 7.** Amplitude and phase of the annual harmonic of the bias for the experiment in which the mean and annual cycles of bias are corrected (Expt. 3). Upper panel shows the annual harmonic of mixed layer (0-45m) temperature bias. Middle panels show the annual harmonic of thermocline depth bias. Lower panels show the annual harmonic of thermocline width bias. Contour intervals for amplitude are °C, m, and 2.5 m. Contour intervals for phase are 45° lag relative to the calendar year.

**Figure 8.** Interannual variability of bias for the experiment in which the mean, annual, and interannual variability of bias is corrected (Expt. 4). Third and fourth EOFs of thermocline depth (upper and middle panels), and principal component time series (lower panel). Contour intervals are 0.5 m.

**Figure 9.** Root-mean-square forecast-minus-observation differences for the first four experiments listed in Table 1. Lefthand panels show RMS differences in the mixed
layer, while righthand panels show RMS differences in thermocline depth. Contour intervals are 0.1C and 2m.

**Figure 10.** Root-mean-square forecast-minus-observation differences for experiments 1, 5, and 4 (control experiment in which no bias correction is performed, Expt. 5 in which the bias analysis is given by the bias forecast, and Expt. 4 in which the full updating equations (6) are used). Lefthand panels show RMS differences in mixed layer temperature, while righthand panels show RMS differences in thermocline depth. Contour intervals are 0.1C and 2m.

**Figure 11.** Monthly averaged zonal velocity at 0N, 110W at 25m and 80m depth during 1.5 years beginning April, 1996. Observations (bold), control expt. (dotted), and Expt. 4 (dashed) are shown. Dramatic changes in zonal currents associated with the 1997 El Nino are evident beginning in April 1997.
Figure 1 Time-mean temperature bias in the equatorial Pacific (averaged 2.5S-2.5N) computed over the 30-year record for the control experiment (Expt. 1) with no bias correction (upper panel), and the mean bias correction experiment (Expt. 2, middle and lower panels). Middle panel shows the mean uncorrected forecast -minus-observations, while lower panel shows the bias-corrected forecast-minus-observation ($\bar{\omega}^f - \bar{\omega}^o$). Depth of the 14C and 22C isotherms (obtained from the control experiment) are indicated by dashed lines. The width of these isotherms is used as a measure of thermocline width. Depth of the 20C isotherm is indicated by a bold line. Depth of this isotherm is used as a measure of thermocline depth. Shading indicates bias exceeding 0.1C. Positive values indicate the forecast has a warm bias.
Figure 2 Geographic structure of time-mean bias in control experiment (Expt. 1, lefthand panels) and a second experiment in which the mean bias is corrected (Expt. 2, righthand panel). Upper panel shows temperature bias in the mixed layer (0-45m). Positive values (solid contours) indicate that the mixed layer is too warm. Contour intervals are 0.025°C. Middle panel shows bias in the depth of the thermocline as represented by the 20°C isotherm. Positive bias indicates that the thermocline is too deep. Contour intervals are 2m. Lower panel shows bias in the width of the thermocline as indicated by errors in the relative depths of the 22°C and 14°C isotherms. Positive bias in thermocline width indicates that the thermocline is too wide. Contour intervals are 5m. Units are °C, m and m.
Figure 3. Histogram of mixed layer temperature forecast-minus-observation differences in the latitude band 10N-30N for the months of June (dark) and December (light). Lefthand panel shows the control experiment (Expt. 1), while the righthand panel shows the experiment with both mean and annual bias correction (Expt. 3). For the control experiment the mean of the differences varies from -0.3°C in June to 0.4°C in December, while after correction for the annually varying bias the monthly mean differences are reduced to –0.04°C and –0.1°C respectively. The standard deviations are reduced by about 20%. 
Fig. 4. Amplitude (left) and phase (right) of the annual harmonic of the bias for the control experiment (Expt. 1). Upper panel shows the annual harmonic of mixed layer (0-45m) temperature bias. Middle panel shows the annual harmonic of thermocline depth bias. Lower panel shows the annual harmonic of thermocline width bias. Units for amplitude are °C, meters, and meters. Units for phase are degrees of lag relative to the calendar year.
Fig. 5 Meridionally averaged tropical (10S-10N) bias as a function of longitude and time. Left-hand panel shows temperature at 100m. Right-hand panel shows mixed layer temperature. Contour interval is 0.5 $^\circ$C. Note that the bias contains a variety of scales including time-mean, annual cycle, as well as interannual variability.
Fig. 6 Interannual variability of bias for the control experiment (Expt.1). Thermocline depth anomalies associated with the third, and fourth (upper and middle panels) three dimensional EOFs ranked in order of explained variance. Contour intervals are 0.5m. Corresponding principal component time series (lower panel). Between them, these two principal components describe 30% of the variance of the interannual variability of the bias (after mean and annual signals have been removed). The Southern Oscillation Index has a zero-lag correlation of 0.7 with the first principal component.
Fig. 7 Amplitude and phase of the annual harmonic of the bias for the experiment in which the mean and annual cycles of bias are corrected (Expt. 3). Upper panel shows the annual harmonic of mixed layer (0-45m) temperature bias. Middle panels show the annual harmonic of thermocline depth bias. Lower panels show the annual harmonic of thermocline width bias. Contour intervals for amplitude are °C, m, and 2.5 m. Contour intervals for phase are 45° lag relative to the calendar year.
**Fig. 8** Interannual variability of bias for the experiment in which the mean, annual, and interannual variability of bias is corrected (Expt. 4). Third and fourth EOFs of thermocline depth (upper and middle panels), and principal component time series (lower panel). Contour intervals are 0.5 m.
Fig. 9 Root-mean-square forecast-minus-observation differences for the first four experiments listed in Table 1. Lefthand panels show RMS differences in the mixed layer, while righthand panels show RMS differences in thermocline depth. Contour intervals are 0.1°C and 2m.
Fig. 10 Root-mean-square forecast-minus-observation differences for experiments 1, 5, and 4 (control experiment in which no bias correction is performed, Expt. 5 in which the bias analysis is given by the bias forecast, and Expt. 4 in which the full updating equations (6) are used). Left-hand panels show RMS differences in mixed layer temperature, while right-hand panels show RMS differences in thermocline depth. Contour intervals are 0.1°C and 2 m.
Fig. 11  Monthly averaged zonal velocity at 0N, 110W at 25m and 80m depth during 1.5 years beginning April, 1996. Observations (bold), control expt. (dotted), and Expt. 4 (dashed) are shown. Dramatic changes in zonal currents associated with the 1997 El Nino are evident beginning in April 1997.