Tellus
SERIES A DYNAMIC METEOROLOGY AND OCEANOGRAPHY
PUBLISHED BY THE INTERNATIONAL METEOROLOGICAL INSTITUTE IN STOCKHOLM

Tellus (2008), 60A, 716–727

TELLUS

Estimation and correction of surface wind-stress bias in the Tropical Pacific with the Ensemble Kalman Filter

By OLWIJN LEEUWENBURGH*, *Institute for Marine and Atmospheric Research (IMAU), Utrecht University, The Netherlands*

(Manuscript submitted 24 April 2007; in final form 25 February 2008)

ABSTRACT

The assimilation of high-quality in situ data into ocean models is known to lead to imbalanced analyses and spurious circulations when the model dynamics or the forcing contain systematic errors. Use of a bias estimation and correction scheme has been shown to significantly improve the balance of the analysed states in such cases. Given the large impact of zonal wind stress bias on the tropical ocean circulation we propose to estimate directly a bias in the forcing rather than in the model state vector. The bias scheme introduced by Dee and Da Silva is modified for use with the Ensemble Kalman Filter and approximate schemes are derived to achieve computational savings. Twin experiments show that for this particular application these simplified schemes efficiently reproduce the bias. The bias in the wind stress can be completely reconstructed sequentially, also when it is slowly evolving in time, leading to well-balanced analyses and unbiased forecasts. In particular the correct recovery of the amplitude of the wind stress bias is found to be important. We additionally investigate the use and correction of a mean sea level climatology in the context of forcing bias estimation and suggest an alternative approach based on the ensemble mean sea level.

1. Introduction

A major part of the effort to improve seasonal forecasting systems has concentrated on optimizing techniques for assimilating observations into the ocean component of the system. A particular challenge in this context is retaining the information inserted into the system by improving the balance of the updated state. This includes, for example, imposing the constraints of geostrophic and hydrostatic momentum balances on the increment.

On the equator the fundamental zonal momentum balance is between the momentum input through the surface wind stress and the pressure gradient force. Near the surface, this pressure gradient is primarily associated with the sea surface slope, while at depth it is also related to the subsurface density field through the hydrostatic balance. The action of the wind is distributed vertically through downward turbulent momentum mixing, a small-scale process that has to be parametrized.

Nowadays, highly accurate observations are available that can be assimilated in ocean models to correct the density and sea level fields. Unbalanced subsurface pressure gradients and spurious circulations may result, however, due to an incorrect model representation of momentum mixing (Bell et al., 2004).

At the same time, even balanced models will tend to drift to their own climatologies when they are run forward unconstrained by observations. This preferred climatological state is, to a large extent, controlled by the boundary conditions. It is well known that the upper ocean dynamics of the tropical Pacific are highly sensitive to surface wind forcing. Several studies have addressed this issue and found significant changes in the circulation when using a range of wind stress products (e.g., Busalacchi et al., 1990; Hackert et al., 2001). Comparison of these wind stress products shows large differences, suggesting that significant errors remain in each of them (see also Fig. 1). Huddleston et al. (2004) concluded that the combination of inaccurate wind-stress forcing and relatively accurate ocean temperature observations tends to produce imbalances similar to those discussed by Bell et al. (2004), again with detrimental impacts on the model forecast dynamics.

Bell et al. (2004) showed that if the dynamic errors resulting from the assimilation of observations have a systematic component, it can be suppressed by implementation of a sequential bias correction method (Friedland, 1969; Dee and Da Silva, 1998). In ocean studies such methods have most commonly been used to estimate systematic errors in the model state (e.g., Carton et al., 2000; Martin et al., 2002) or derived quantities (Bell et al., 2004; Balmaseda et al., 2007), possibly with prescribed space or time characteristics (Chepurin et al., 2005).

*Correspondence: TNO Environment and Geosciences, PO Box 80015, NL-3508 TA, Utrecht.
e-mail: olwijn.leeuwenburgh@tno.nl
DOI: 10.1111/j.1600-0870.2008.00325.x

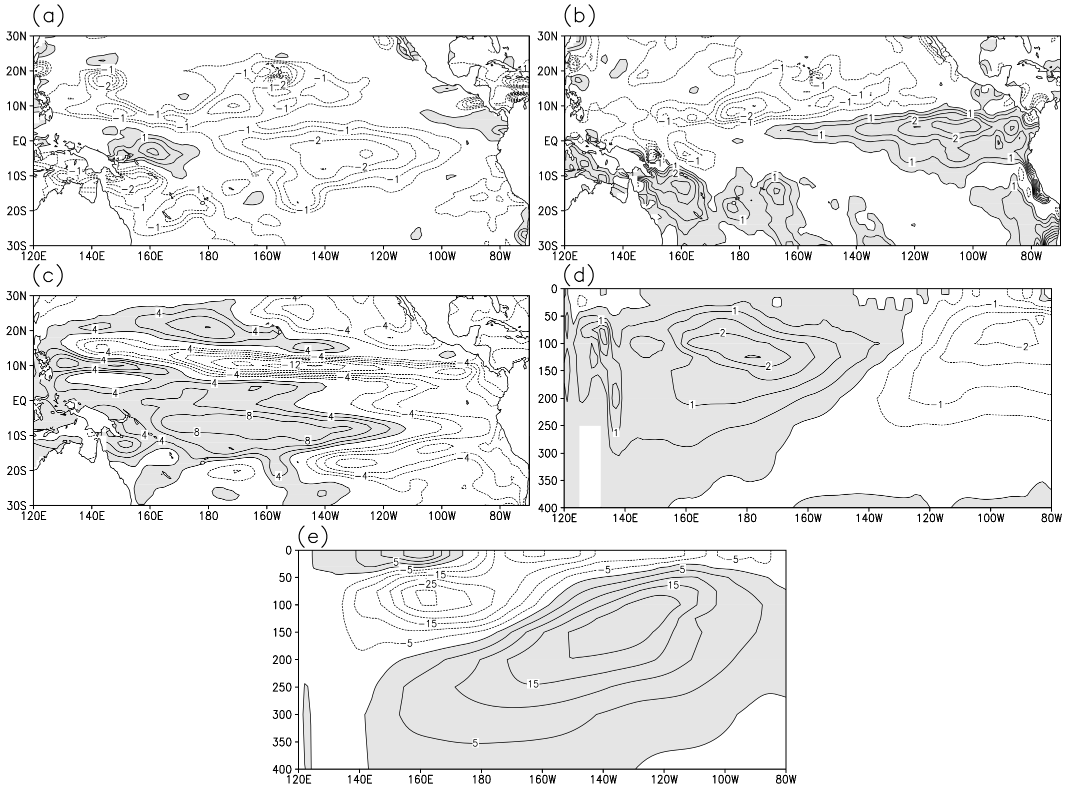

Fig. 1. Time-mean difference between ERA40 and NCEP/NCAR forced model runs for (a) zonal wind stress (10^{-2} N m^{-2}), (b) meridional wind stress (10^{-2} N m^{-2}), (c) sea level (cm), (d) temperature (K) and (e) zonal velocity (cm s^{-1}). The vertical axes in (d, e) show depth (m) on the equator.

A disadvantage of this approach is that the actual source of the systematic error is generally left uncorrected. Perhaps a better one would therefore be to directly correct the erroneous boundary conditions (e.g. the wind stress field) or parameters (e.g. the vertical momentum mixing). The feasability of this approach has previously been demonstrated for variational methods (see e.g., Yu and O'Brien, 1991; Vossepoel et al., 2004). A somewhat similar approach was followed by Bosilovich et al. (2007), who applied a sequential method to assimilate surface temperature observations into a coupled land–atmosphere model and estimate the bias in the diurnal cycle of the energy flux between the two components.

Ensemble methods, such as the Ensemble Kalman Filter (EnKF) (Evensen, 1994), also provide unique advantages that can be exploited in this context. Through the use of ensemble covariances it becomes possible to relate observed quantities to (unobserved) model state variables, but also to parameters and boundary conditions. Most importantly, the covariance evolves over time with the model state.

Keppenne et al. (2005) demonstrated that a bias correction method can be used in combination with the EnKF to account for bias in the sea surface height climatology used with altimetric measurements. In this particular application the bias information is not fed back to the model during the forecast stage.

The intent of this paper is twofold: demonstrate the ability of an ensemble-based assimilation system to correct for systematic errors associated with bias in the surface boundary conditions of an ocean model, and show that such a correction can improve the behaviour of the model during the forecast. Furthermore, we investigate the potential for reconstruction of the unbiased total sea level field through the combined use of an ensemble and multivariate observations, which would enable a more optimal use of sea level anomaly observations from satellite altimetry. An ocean general circulation model will be used to conduct twin experiments in which prescribed biases can be introduced.

Section 2 provides an overview of the ensemble filter used to combine the model with data, and of the bias correction scheme. In Section 3 the model will be introduced. The results of a series of twin experiments are described in Section 4. A discussion of the results and conclusions are given in Section 5.

2. Methods

2.1. The ensemble filter

The filter algorithm used in this study has been described in detail in previous studies (e.g., Leeuwenburgh, 2007) but is summarized here for completeness. The analysis ensemble is formed using the stochastic EnKF described by Evensen (1994) and Burgers et al. (1998). Covariance localization is implemented using a compactly supported correlation function $\mathcal{C}(\lambda, \theta)$ from Gaspari and Cohn (1999). The resulting algorithm is given by

$$\mathbf{A}^a = \mathbf{A}^f + \mathcal{C} \circ (\mathbf{P}^f \mathbf{H}^T)[\mathcal{C} \circ (\mathbf{H}\mathbf{P}^f \mathbf{H}^T) + \mathbf{R}]^{-1}(\mathbf{D} - \mathbf{H}\mathbf{A}^f)$$
$$= \mathbf{A}^f + \mathbf{K}\,(\mathbf{D} - \mathbf{H}\mathbf{A}^f), \tag{1}$$

where $\mathbf{A} = (\psi_1, \ldots, \psi_N)$ holds the ensemble of model states, $\circ$ denotes the Schur product operator, $\mathbf{R}$ is the observation error covariance matrix, and $\mathbf{H}$ is a (linear) measurement operator. $\mathbf{D} = (\mathbf{d} + \epsilon_1, \ldots, \mathbf{d} + \epsilon_N)$ is the sum of the data vector $\mathbf{d}$ and an ensemble of observation perturbation vectors $\mathbf{E} = (\epsilon_1, \ldots, \epsilon_N)$. These perturbations should be chosen such that $\langle \mathbf{E} \rangle = \mathbf{0}$ and $\langle \mathbf{E}\mathbf{E}^T \rangle = \mathbf{R}$. Applying the Schur product in observation space avoids the necessity to calculate the ensemble covariances $\mathbf{P} = \mathbf{A}'\mathbf{A}'^T/(N-1)$ explicitly in model space, which would be too costly (Houtekamer and Mitchell, 2001). Primes are used to indicate anomalies with respect to the ensemble mean and the superscripts ^{f} and ^{a} denote forecast and analysis, respectively. The filter solution is calculated independently for each grid column, where only observations within an elliptic region around the grid point are used. The half axes of the local support ellipse are $\lambda = 60°$ in the zonal and $\theta = 30°$ in the meridional direction and a scale of 500 m is used in the vertical (Keppenne and Rienecker, 2002). The horizontal scales are halved when sea level is assimilated. This so-called 'local analysis' approach is an approximation to the full problem in which all data are used but is computationally more tractable when the number of observations is very large. It has the added advantages of reducing spurious long-range correlations associated with the use of a small ensemble, and of increasing the dimension of the solution space and thereby reducing the potential for inbreeding (linear dependency between ensemble members).

2.2. Bias correction; two-step and one-step schemes

Dee and Da Silva (1998) derived a two-step online bias estimation and correction scheme with feedback for use in a sequential data assimilation system,

$$\mathbf{b}^a = \mathbf{b}^f - \mathbf{L}\,[\mathbf{d} - \mathbf{H}(\psi^f - \mathbf{b}^f)] \tag{2}$$

$$\psi^a = \psi^f - \mathbf{b}^a + \mathbf{K}\,[\mathbf{d} - \mathbf{H}(\psi^f - \mathbf{b}^a)]. \tag{3}$$

The first step is a linear update of the model bias forecast $\mathbf{b}^f$. The second step is the familiar update of the model state (analysis), but with the modification that the forecast is corrected with the latest bias estimate $\mathbf{b}^a$. Both analyses are optimal (Dee and Todling, 2000) when

$$\mathbf{L} = \mathbf{P}_b\mathbf{H}^T \left[\mathbf{H}\mathbf{P}_b\mathbf{H}^T + \mathbf{H}\mathbf{P}^f\mathbf{H}^T + \mathbf{R}\right]^{-1}, \tag{4}$$

where $\mathbf{P}_b$ is the error–covariance matrix of the bias forecast. When no dynamic model is available for the bias, persistence is commonly assumed between analysis and forecast times.

Equations (2) and (3) can be reformulated for application with an ensemble of models,

$$\mathbf{B}^a = \mathbf{B}^f - \mathbf{L}\,\left[\mathbf{D} - \mathbf{H}(\mathbf{A}^f - \mathbf{B}^f)\right] \tag{5}$$

$$\mathbf{A}^a = \mathbf{A}^f - \mathbf{B}^a + \mathbf{K}\,\left[\mathbf{D} - \mathbf{H}(\mathbf{A}^f - \mathbf{B}^a)\right], \tag{6}$$

where $\mathbf{B} = (\mathbf{b}_1, \ldots, \mathbf{b}_N)$ now contains the ensemble of model bias vectors.

Since information on the structure and magnitude of the bias is generally lacking on forehand, it has been common practise to assume that the bias error covariance is proportional to the error covariance of the unbiased model state (Dee and Todling, 2000),

$$\mathbf{P}_b = \alpha\mathbf{P}^f, \tag{7}$$

where the factor α determines the relative magnitude of the error associated with bias. If $\alpha \ll 1$, the gain matrix for the bias estimator can be approximated as

$$\mathbf{L} \approx \alpha\mathbf{K}. \tag{8}$$

When this expression is substituted in eq. (2) or (5) it can be seen that α controls the rate of change of the bias estimate. For small α, or slow rates of change, the bias forecast can be replaced with the bias estimate from the previous analysis time, and the order of the two steps can be switched (Dee, 2005).

Sometimes the model bias can be assumed to be identical for all ensemble members. This is generally the case when the same dynamic model (including parameters) and boundary conditions (up to zero-mean random perturbations) are used for each member. With all the above simplifications the final form of the two-step algorithm becomes

$$\mathbf{A}^a = \mathbf{A}^f - \mathbf{b}^f + \mathbf{K}\,[\mathbf{D} - \mathbf{H}(\mathbf{A}^f - \mathbf{b}^f)] \tag{9}$$

$$\mathbf{b}^a = \mathbf{b}^f - \alpha\mathbf{K}^b\,[\mathbf{d} - \mathbf{H}(\overline{\psi}^f - \mathbf{b}^f)], \tag{10}$$

where $\overline{\psi}^f$ is the ensemble-mean model forecast. $\mathbf{K}_b = \mathbf{K}$ if the same observations and Schur operator are used in both steps. In this case there are no significant extra computational costs associated with the bias correction step (Dee, 2005). Balmaseda et al. (2007) refer to the resulting scheme as the one-step bias correction algorithm.

The approximation (7) may no longer be justified when the spatial scales associated with the bias are substantially greater than those presumed valid for adjustment of the unbiased model state (see e.g., Keppenne et al., 2005). In order to reconstruct a bias that is dominated by spatial scales which are substantially greater than those resolved by the local observations there are two practical adaptations that we can make: (1) retain long-range correlations by reducing the degree of localization and (2) use observations that are further away from the analysis point. Both can be achieved by increasing the scales in the localizing function $\mathcal{C}(\lambda, \theta)$. To keep the computational cost limited, this can be combined with subsampling of the data set for the bias estimation step. If either the scales or the observation set are changed, the entire Kalman gain matrix $\mathbf{K}^b$ will, in principle, need to be recomputed. In this case the computational cost of evaluating eqs. (9) and (10) becomes identical to that of the two-step scheme, the only difference being the order of the two equations. However, if the bias update is based on the same observations as the analysis, some computational savings can be achieved by making the approximation discussed in the next section.

2.3. A 1.5-step bias correction scheme

A complete re-evaluation of the Kalman gain matrix for the bias estimation step would roughly double the computational cost of the one-step bias scheme. The most expensive part of the calculation is the inversion of the error covariance matrix of the innovations. Significant computational savings can be made if this term is not recomputed. The alternative scheme presented below therefore uses the same set of observations for both the analysis and bias update. For the bias step different spatial scales are only used in the localization of the covariances between the elements of the bias state and the observables (the first term of the gain matrix),

$$\mathbf{K}_b = \mathcal{C}_b \circ (\mathbf{P}^f \mathbf{H}^T)_b [\mathcal{C} \circ (\mathbf{H}\mathbf{P}^f \mathbf{H}^T) + \mathbf{R}]^{-1}. \tag{11}$$

Compared to the one-step scheme, only two additional computations need to be performed: the Schur product needs to be reapplied to the term $\mathbf{P}^f \mathbf{H}^T$ and the gain matrix needs to be reconstructed by multiplication with the inverse of the error covariance matrix of the innovations. We will refer to the resulting algorithm as the 1.5-step bias correction scheme.

2.4. Practical implementation

Rather than estimating a (proxy) bias in the model state or dynamics, we explicitly estimate a bias in the boundary conditions. The standard approach to estimating boundary conditions or parameters in filter methods is to add them to the state vector. In the current application this means that the state vector is extended with the surface wind stress field. The practical implementation is as follows.

During the forecast run the most recent bias estimate, which is identical for all ensemble members, is added to the best-guess wind stress field. At the end of the forecast run, the final wind stress field is added to the state vector and passed on to the filter algorithm. If only a bias in the wind stress is to be estimated, then, since the wind stress is not observed, eq. (9) reduces to the original EnKF analysis update eq. (1). The wind stress bias is subsequently updated by means of eq. (10), based on the model-data misfit and the ensemble cross-covariances between the observed variables and the wind stress. In experiment F (see Table 1), where sea level anomalies are assimilated, the full version of eq. (9) is used. The most recent sea level bias estimate is then used to update a mean sea level climatology which serves as a reference level for obtaining anomalies. An alternative approach to dealing with bias when assimilating sea level anomalies is investigated in experiment E and discussed further in Section 4.

Table 1. Summary of experiments. Temperature profiles are measured at TAO buoy locations. In Exp E and F additional temperature profiles from Argo floats are assimilated (see text for details). The bias parameter $\alpha = 0.25$ for all experiments except Exp G where $\alpha = 0.5$.

Exp	Data	Bias	Scheme	Scale factors
A	T	no	n.a.	n.a.
B	T	τ	2	$\beta_\lambda = \beta_\theta = 2$
C	T	τ	1	$\beta_\lambda = \beta_\theta = 1$
D	T	τ	1.5	$\beta_\lambda = \beta_\theta = 2$
E	T & SLA	τ	1.5	as in D
F	T & SLA	τ & MSSH	1.5	as in D
G	T	$\tau(t)$	1.5	as in D

In some of the experiments presented here the spatial scales for the bias estimation are modified such that $\mathcal{C}_b = \mathcal{C}(\beta_\lambda \cdot \lambda, \beta_\theta \cdot \theta)$. Note again that for the one-step scheme, $\beta_\lambda = \beta_\theta = 1$ and so $\mathcal{C}_b = \mathcal{C}$ and $\mathbf{K}_b = \mathbf{K}$.

3. The model

The model used in this study is the Max Planck Institut für Meteorologie Ocean Model (MPI-OM) (Marsland et al., 2003). The global orthogonal curvilinear grid has a spatial resolution approximating spectral truncation T42, with poles positioned over Greenland and inland of the Weddell Sea to give high resolution in the main sinking areas associated with the thermohaline circulation. Additional increase in resolution is achieved by meridional refinement (0.5°) of the grid within 10° of the equator. Horizontal discretization of the primitive equations is on the Arakawa C-grid, while the z-coordinate is discretized on 23 vertical levels. The model has a free sea surface. Surface exchanges of momentum, fresh water and heat are calculated using prescribed daily fields of surface stress, 10 m wind speed, 2 m air and dew-point temperature, short-wave radiation, cloud cover and precipitation from the NCEP/NCAR reanalysis (Kalnay et al., 1996). During the experiments a wind stress bias is simulated by adding an easterly wind patch on the equator (Fig. 2). During the ensemble runs this biased wind stress field is further perturbed by adding zonally stretched random fields which are correlated in time.

4. Experiments

4.1. Overview

The design of the twin experiments described here is based on the observation that large differences exist between available wind

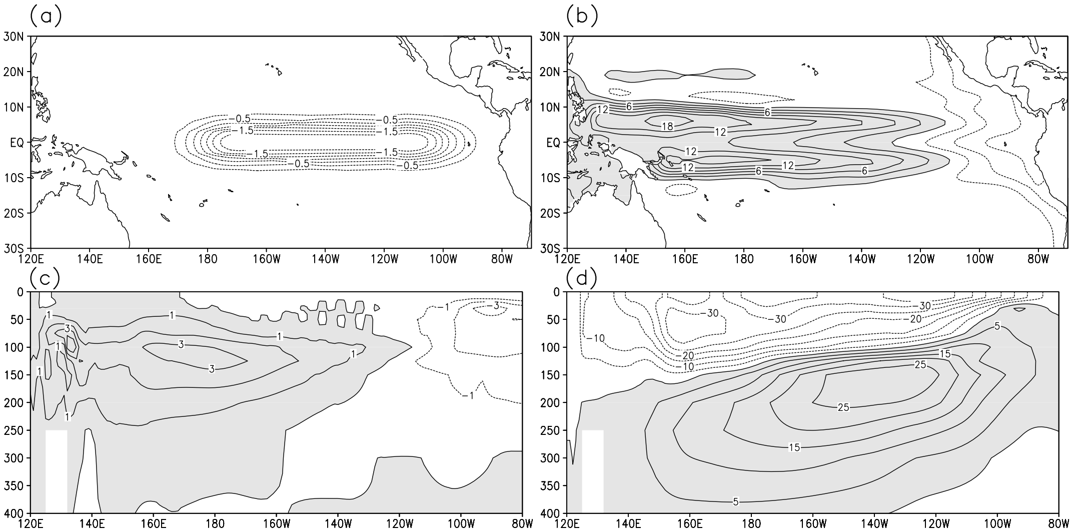

Fig. 2. Time-mean difference between control and truth for (a) zonal wind stress (10^{-2} N m^{-2}), (b) sea level (cm), (c) temperature (K) and (d) zonal velocity (cm s^{-1}). The vertical axes in (c, d) show depth (m) on the equator.

stress products, implying the presence of significant remaining errors. Figure 1a shows the 1993 annual mean difference in zonal wind stress (τ) between the ERA-40 (Uppala et al., 2005) and NCEP/NCAR (Kalnay et al., 1996) reanalyses. A large negative patch with magnitude of about 0.02 N m^{-2} can be seen occupying the central equatorial Pacific. Differences of similar magnitude are found in the meridional wind stress component (Fig. 1b). The associated difference in mean sea surface height, resulting from two model runs forced with the ERA-40 and NCEP/NCAR wind stress fields, respectively, contains large values in the entire tropical Pacific (Fig. 1c). In the ERA40-forced run more water is pushed towards the western side of the basin by the stronger easterly winds. Very large systematic differences are also found in all subsurface fields, as illustrated by temperature differences of up to 2.5 K, and zonal velocity differences up to 30 cm s^{-1}.

In order to simulate this situation twin experiments are conducted in which the 'true' ocean state is defined by a forward run of the ocean model using NCEP/NCAR forcing fields. The available wind product is assumed to differ from this true forcing by a bias in the zonal component of the tropical wind stress, that is, the westward directed wind stress is assumed to be too strong with a maximum error of 0.02 N m^{-2} on the equator (since the model grid is slightly curved near the equator, the bias also has a weak projection on the meridional component when interpolated to a latitude-longitude coordinate frame). The model run forced by this biased wind product is referred to as the 'control'. The initial conditions for both runs at the beginning of the experiment period (the first 9 months of 1993) are obtained by spinning up the model for 1 yr, starting from the same model state, but forced by the two different winds.

Mean differences between the truth and control runs for 1993 are shown in Fig. 2. The bias in the wind is seen to have affected the sea level field as well as the subsurface velocity and density distributions, the latter represented here by the temperature field. Based on the similarity between these results and those of Fig. 1, it can be concluded that it is the zonal wind stress bias which is primarily responsible for the equatorial differences between ERA40- and NCEP/NCAR-forced model runs. Since the model is the same for all runs, the only source of error in the control is the forcing bias. While a more realistic setting would of course be one where the model equations and the other components of the forcing contain (time-dependent) errors as well, the focus here will be primarily on investigating the ability of the bias estimation schemes to identify and correct the wind bias.

An ensemble of 32 model realizations is obtained by repeating the spin-up with each member forced by the control forcing plus an additional random perturbation to the zonal wind stress. The best guess model state and uncertainty therein are assumed to be given by the mean and spread of the ensemble. The perturbations are correlated over time with an e-folding scale of 60 d. Spatial scales for the perturbations are approximately 40° and 5° in the zonal and meridional grid directions.

Several experiments are performed to test the performance of the three bias estimation schemes discussed in the previous section. Exp A is an assimilation-only run where temperature is assimilated without any wind stress bias estimation. In Exp B a bias estimation step is added, according to the two-step algorithm with increased spatial scales for the bias step (results from a single bias estimation step with varying scales suggested that the wind stress bias pattern is better reproduced with increased scales, in agreement with results from earlier studies, e.g. Keppenne et al., 2005). In Exp C the one-step algorithm is used without any scale adjustment, while in Exp D, which uses the 1.5-step algorithm, the scales for bias estimation are adjusted conform eq. (11). The reader is reminded that the one-step scheme produces the same results as the two-step scheme without scale adjustment.

In Exp E and Exp F both temperature and sea level anomalies are assimilated with the 1.5-step algorithm. In Exp E only the estimated wind stress bias is corrected while the ensemble-mean sea surface height is used as a reference level to obtain sea level anomalies. In Exp F an additional sea level bias is estimated which is subsequently used to update a climatological mean sea level (MSSH). In this experiment sea level anomalies are obtained by subtracting this MSSH from the total model sea surface height (for further details see Section 4.3).

Finally, in Exp G, an annual variation of the wind stress bias is reconstructed based on temperature observations only. In all experiments temperature observations are simulated by sampling the truth at the locations of the Tropical Atmosphere Ocean (TAO) mooring sensors. In Exp E and Exp F additional Argo float temperature profiles are simulated at 50 randomly located profiles between 10° and 20° latitude. Altimetric sea level measurements are simulated by sampling the truth along TOPEX/POSEIDON ground tracks. Random errors are added to all observations. A complete set of measurements is assimilated every 10 d. The characteristics of all experiments are summarized in Table 1 for reference.

4.2. Assimilation of temperature profiles

We will first consider the results of experiments A–D in which only temperature profiles at the TAO mooring locations are assimilated. Figure 3 shows the mean analysis increment (analysis minus forecast) over the final 4 months for each experiment (the results for Exp C and D were very similar, only those for Exp D are shown). For Exp A, the assimilation-only run, the time-mean increment is found to be significantly different from zero. This suggests that the correction to the model forecast state systematically works in one direction, a clear indication of bias. Subsurface temperature above 200 m depth and east of the date line is consistently lower in the forecast than in the observations. The temperature bias in this run is associated with erroneously strong upward advection of cool subthermocline waters. These results indicate that even over a forecast period as short as 10 d, a significant readjustment of the equatorial subsurface state takes place when the model is forced with winds that are inconsistent with the assimilated temperature observations. When the temperature data is additionally used to estimate a zonal wind stress bias, as is done in Exp B, the mean amplitude of the analysis increments is visibly reduced. In particular the systematic adjustment to the vertical velocity at about 200 m depth is notably

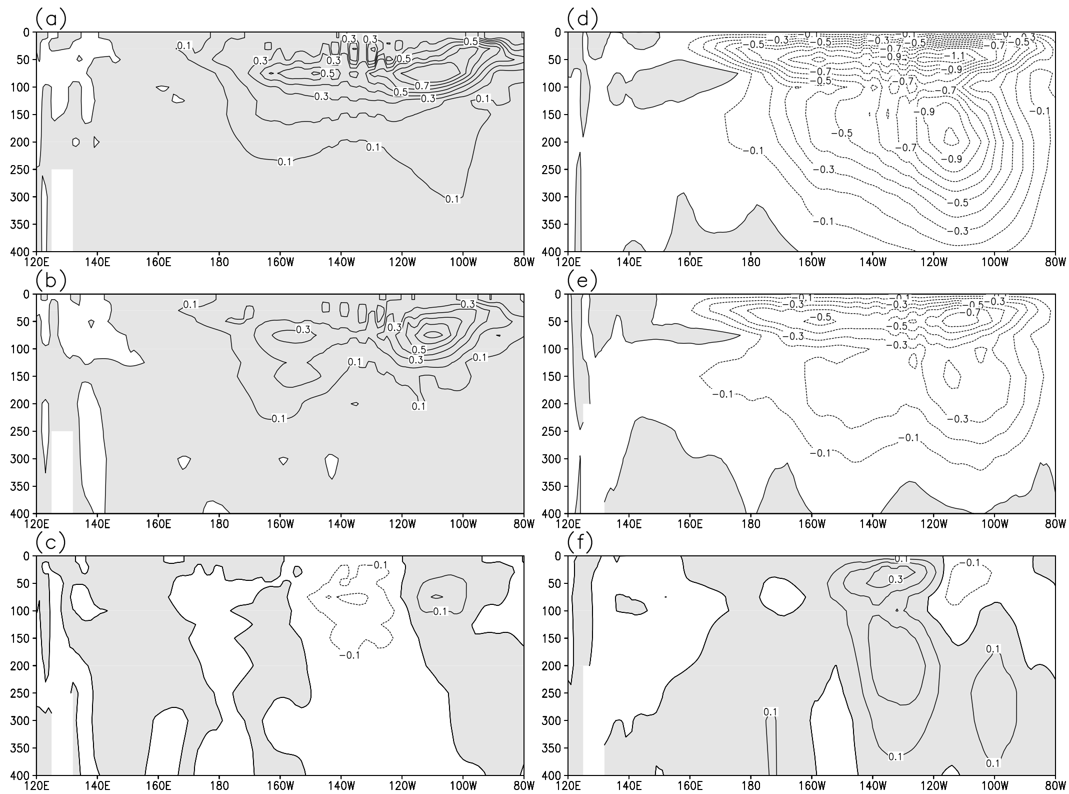

Fig. 3. Time-mean analysis increment along the equator for (left-hand side) temperature (K) and (right-hand side) vertical velocity (m d^{-1}). (a, d) ExpA: assimilation only run, (b, e) Exp B: assimilation with two-step wind stress bias correction, (c, f) Exp D: assimilation with 1.5-step wind stress bias correction. The vertical axis shows depth (m).

smaller. The third experiment shown, Exp D, appears to show an even further improved balance in the model forecast state. There is almost no indication of systematic differences between forecast and analysis.

There are thus clear indications that including a bias estimation procedure in the assimilation cycle improves the balance of the analysed state, but we are equally interested in the true forecast error. Fig. 5 shows the temperature forecast errors (forecast minus truth) for both the assimilation-only run and two of the runs with bias correction, again averaged over the final 4 months (the error for the control is reproduced in Fig. 5a for comparison). The results for Exp A (Fig. 5b) suggest that the assimilation of temperature profiles corrects a large part of the subsurface temperature bias, and furthermore that the model is able retain some of this information for at least 10 d since the errors are smaller than for the control. The error on the eastern side of the basin, however, is found to be systematically negative, consistent with the positive temperature analysis increments (see Fig. 3a). Inclusion of a bias-correction scheme leads to further reduction of these errors where the best results are obtained here for Exp D, except perhaps near the western and eastern boundaries of the domain.

A similar picture arises from Fig. 6 which shows time-mean forecast errors in the vertical velocity. It can clearly be seen that the dynamic characteristics of the forecast state will be worse than those of the control if temperature is assimilated naively, that is, without taking into account that the system is biased. Very large spurious vertical velocities appear in the forecast when assimilating subsurface temperature in this case. When the forcing bias is corrected using a bias scheme with increased spatial scales (Exp B), these erroneous velocities are reduced by approximately two-thirds. The smallest errors are obtained in Exp D, in which the spurious velocities have all but disappeared completely.

Some insight into the reason for the observed differences can perhaps be obtained from Fig. 4 which shows the mean wind stress bias estimate over the final 4 months. From a visual comparison with Fig. 2 it appears that the spatial pattern of the bias is best reproduced when increased spatial scales are used in the bias estimation step as done in Exp B and in Exp D, although the latter shows a rather strong positive patch to the north. The peak amplitude of the bias (true value 0.02 N m^{-2}), however, is clearly underestimated in Exp B where both Exps C and D produce approximately the correct equatorial amplitude. We note again that there appeared to be no significant differences between Exp C and D in the balance properties of the forecast.

Based on these results, we conclude that including a forcing bias correction step in the assimilation cycle significantly improves the dynamic balance of the analysis and accuracy of the forecast. Furthermore, we find that the experiments without or with partially adjusted spatial scales in the bias estimation step produced better balanced states here than the experiment which used a complete recalculation of the Kalman gain with increased

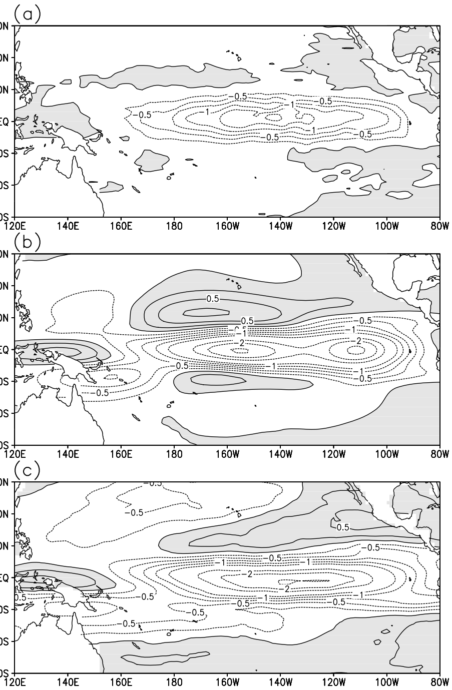

Fig. 4. Time-mean zonal wind stress bias estimate (10^{-2} N m^{-2}) for (a) Exp B, (b) Exp C and (c) Exp D.

spatial scales in the bias step. In the latter case it was that found that the amplitude of the forcing bias was underestimated. Using increased spatial scales did improve the reconstruction of the spatial pattern however. The 1.5-step bias scheme with partially adjusted scales will be used in experiments E and F, in which additional sea level measurements are assimilated, as described in the next section. As a final note, it should be remembered that the bias schemes used here are all simplified versions of the full two-step scheme introduced by Dee and Da Silva (1998) and should thus all be considered suboptimal.

4.3. Assimilation of sea level anomalies

It is known from previous studies that assimilation of sea level data can improve the subsurface model state in the tropical ocean (e.g., Fischer et al., 1997; Ji et al., 2000; Segschneider et al., 2000). It has also been demonstrated that the EnKF can be used for this purpose (e.g., Leeuwenburgh, 2005). A long-standing problem with the practical use of altimetric sea level observations, however, is the lack of an accurate geoid. In the absence of such a physically meaningful reference level, the most common practical approach has been to assimilate anomalies with respect to a time-mean value. The model time-mean sea surface height (MSSH) used to construct model anomalies is typically obtained from a potentially biased control run.

Keppenne et al. (2005) recently addressed the bias problem associated with the assimilation of sea level observations and

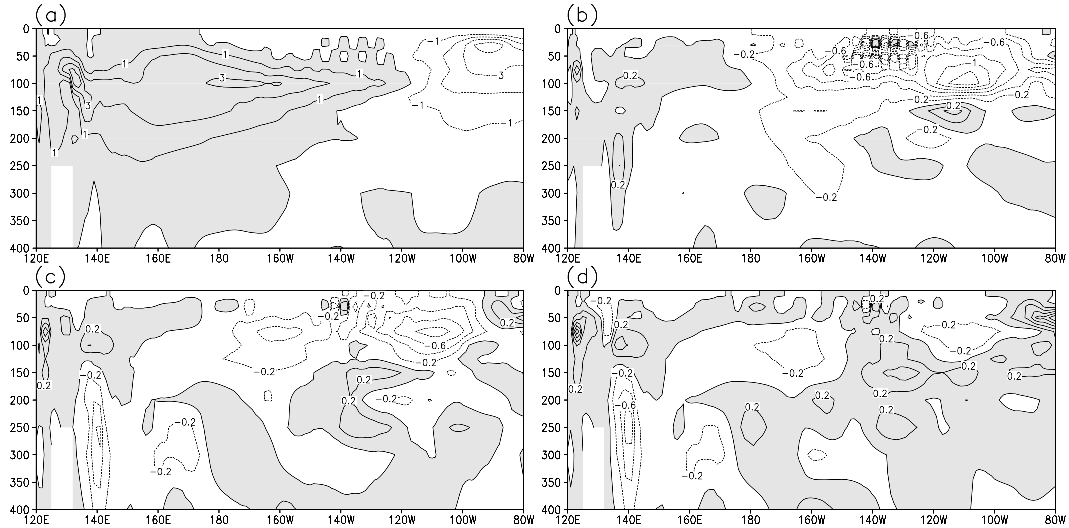

Fig. 5. Time-mean temperature forecast error (K) along the equator. (a) control, (b) Exp A, (c) Exp B and (d) Exp D. The vertical axis shows depth (m).

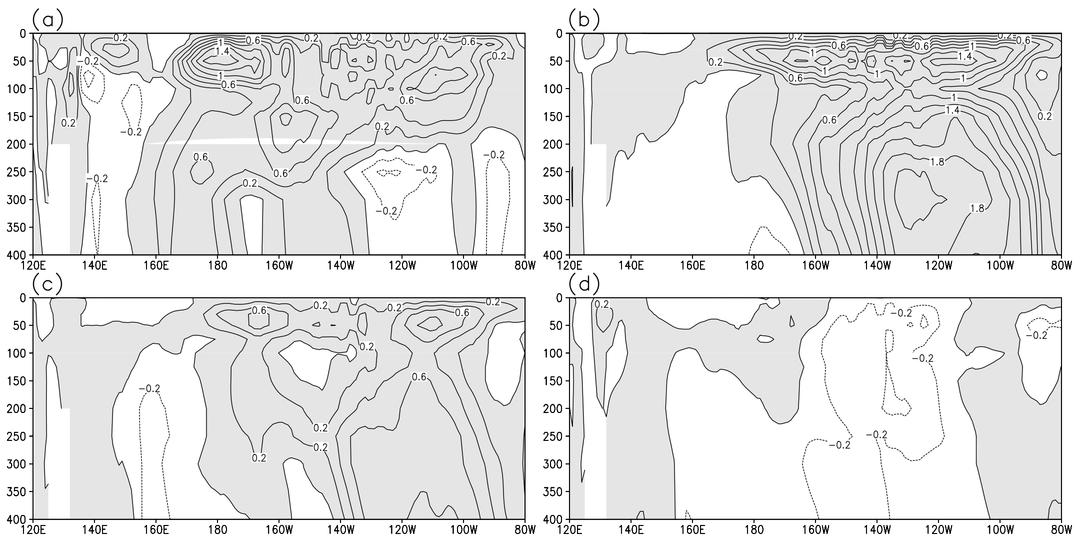

Fig. 6. Time-mean vertical velocity forecast error (m d^{-1}) along the equator. (a) control, (b) Exp A, (c) Exp B and (d) Exp D. The vertical axis shows depth (m).

found that forecast-data misfits could be reduced when the sea level innovations were used to correct the bias in the model MSSH. They followed a two-stage approach in which temperature data was assimilated first without any bias estimation. In the second stage altimetric sea level data was assimilated and used to estimate and correct the MSSH. No forcing bias was estimated. The effect of the temperature assimilation is to move the model sea level closer to the true height. The effect of the subsequent sea level assimilation is, at least initially, to move the model sea level back to it's biased position, since the observations are reconstructed with respect to a biased MSSH. This effect is expected to diminish over time if the bias estimation is converging to the true value.

The information in the temperature data is essential for estimating the MSSH bias; since the time-mean is removed from the altimetric observations, it can not be expected that time-mean biases in the model sea level can be corrected from altimetric sea level anomalies alone (some more discussion on this issue is presented in Section 5). This also suggests that the extent to which the MSSH bias can be corrected may be limited to the static sea level response to temperature change, and to the extent that this response is captured by the ensemble cross-covariances.

Figure 7a shows the sea level bias estimate obtained from Exp D (this bias was estimated simultaneously with the wind stress bias, but not used). Exp B and C produced sea level bias estimates with fairly similar amplitudes and patterns. This result

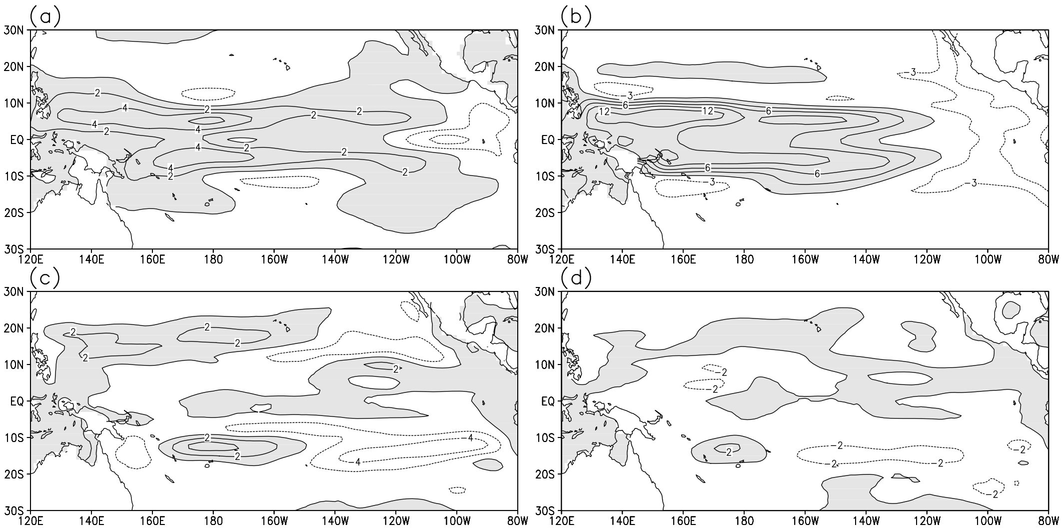

Fig. 7. (a) Time-mean sea level bias estimate (cm) from Exp D. (b) Residual error in the bias-corrected MSSH, using the bias estimate from panel (a). (c) Time and ensemble-averaged sea level forecast error for Exp D from assimilation of TAO temperature observations. (d) As (c) but with additional assimilation of Argo temperature observations between 10° and 20° latitude.

suggests that it might be possible to estimate the sea level bias associated with systematic wind-stress offsets by assimilating temperature. It can be observed, however, that even though the pattern is rather similar to that of the true bias, the amplitude is significantly smaller. The remaining fraction is most likely associated with dynamic adjustments to the wind stress that are not well captured by the temperature data. Figure 7b shows the remaining error in the corrected MSSH. While subtracting the estimated bias reduced the error somewhat, clearly a large part of the original bias remains, as can be seen by comparing with Fig. 2b. Temperature observations alone appear to be insufficient to correct the biased MSSH.

The use of an ensemble offers the possibility of a different approach. The fundamental advantage of ensemble methods is that the time-mean can be replaced by a time-evolving ensemble-mean. The applied ensemble perturbations are assumed to be mean-preserving (that is, they do not trigger non-linearities in the dynamics). If we therefore are successful in correcting the bias in the ensemble by using a bias correction scheme, we may assume that the ensemble mean will approach the true time-mean. Figure 7c shows the time-averaged ensemble-mean sea level forecast error from Exp D. Clearly this error is much smaller than that in the reconstructed MSSH. Increased error values are found in 10° wide latitude bands just south and north of the region covered by the TAO buoy array. If additional Argo-type temperature profiles are assimilated at these latitudes, these errors disappear (Fig. 7d).

The above ideas will be further investigated in two experiments. The first experiment (Exp E) uses no explicit sea level bias correction. The observations are exclusively used to estimate the wind stress bias, and model sea level anomalies are obtained by subtracting the ensemble mean. The second experiment (Exp F) is a variation on the procedure of Keppenne et al. (2005). The climatological MSSH is allowed to evolve during the assimilation cycle by using part of the sea level analysis increments resulting from temperature assimilation to update a MSSH bias estimate. Model sea level anomalies, needed for the subsequent assimilation of altimetric sea level measurements, are obtained by subtracting the updated MSSH. The wind stress bias estimate is updated after both assimilation steps. Both experiments are run over the final 4 months of the test period, continuing from the analysis and bias estimates of Exp D after month 5. It was verified that the 4 month period is long enough for the bias estimates to reach a steady value.

Figure 8 shows the mean analysis increments over the final 4 months of the experiments for both temperature and vertical velocity. It can be seen that, on average, the increments are smaller for Exp E, suggesting that this run is better balanced. In particular the values for vertical velocity in the upper central Pacific and in eastern Pacific down to 400 m depth are significantly greater for Exp F.

A similar picture arises for the forecast errors (Fig. 9), although it should be noted that also the vertical velocity errors for Exp E are larger than those for the run with temperature assimilation only. A possible explanation may be that the ensemble-mean sea level in Exp E evolves freely with time (constrained by the time scales of the forcing perturbations), while the observed anomalies were originally obtained by subtracting a time-mean value. Figure 10, finally, compares the errors in the time-mean sea surface height estimates which are obtained from both experiments. For Exp E the time-averaged, ensemble-mean, forecast error is very small. For Exp. F, on the other hand, the bias-corrected MSSH still contains very large errors.

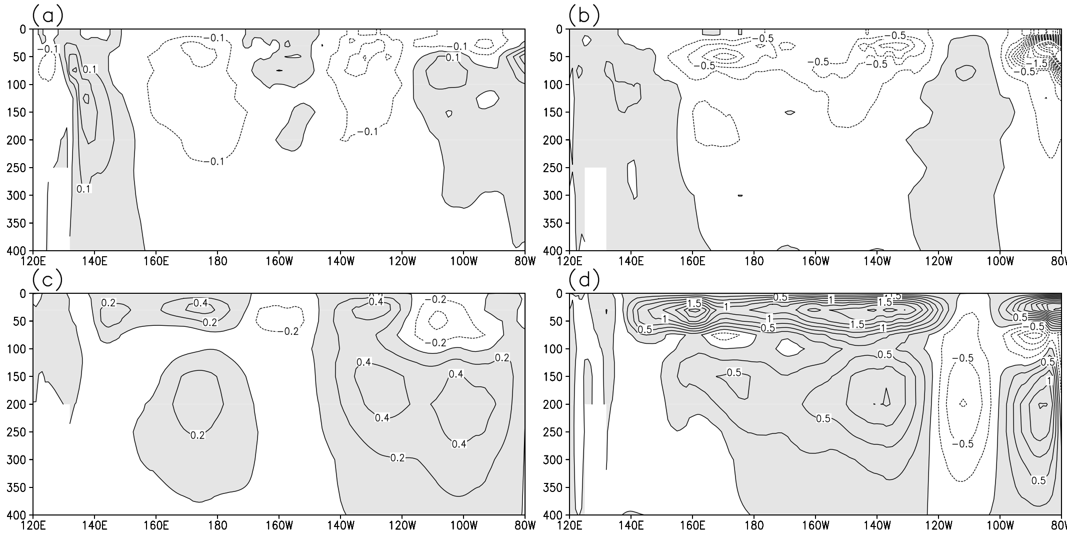

Fig. 8. Time-mean analysis increments along the equator for (top) temperature (K), and (bottom) vertical velocity (m d^{-1}). (a, c) Exp E, (b, d) Exp F. The vertical axis shows depth (m).

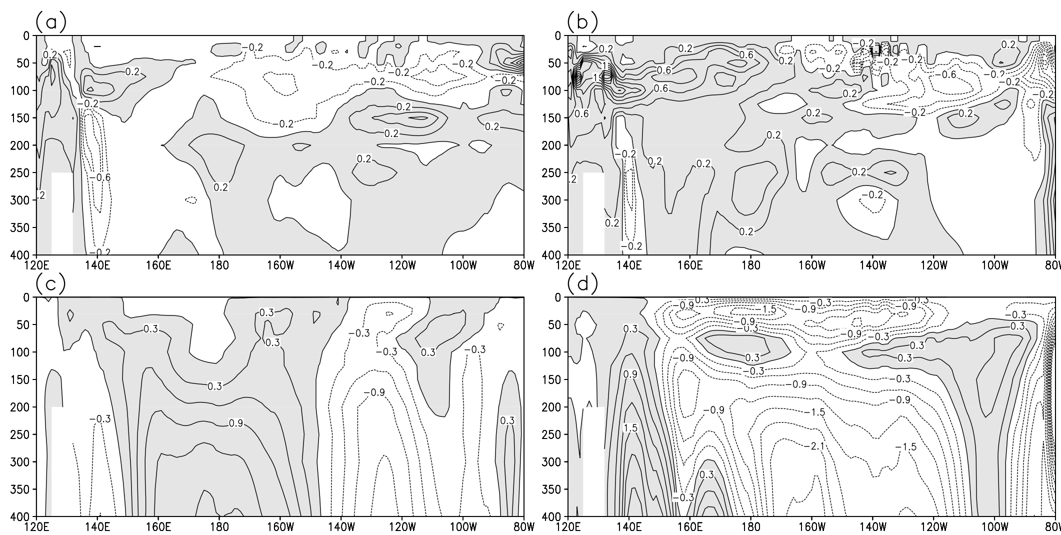

Fig. 9. Time-mean forecast errors along the equator for (top) temperature (K), and (bottom) vertical velocity (m d^{-1}). (a, c) Exp E, (b, d) Exp F. The vertical axis shows depth (m).

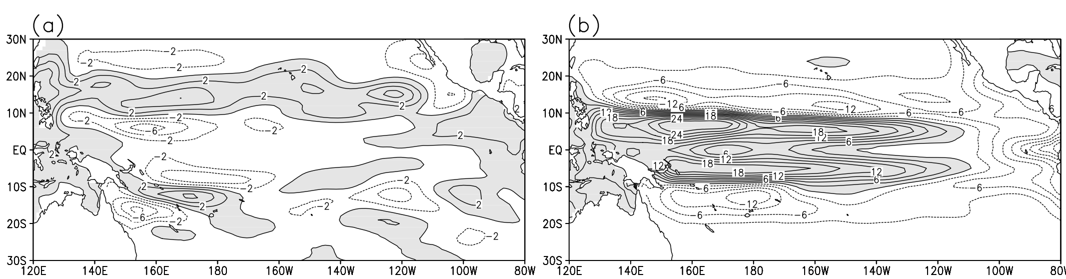

Fig. 10. (a) Time and ensemble-averaged sea level forecast error (cm) for Exp E. (b) Time-mean error in the bias-corrected MSSH (cm) from Exp F.

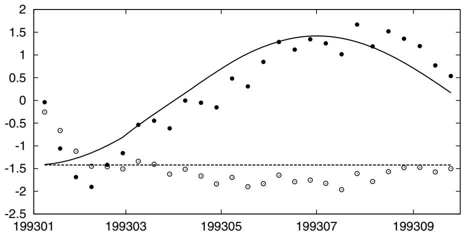

Fig. 11. Evolution of the magnitude of the bias estimate (in units of $10^{-2}\,\mathrm{N\,m^{-2}}$, vertical axis), averaged between 190° E and 250° E and 5° S and 5 ° N, between January 1993 and October 1993 (yyyymm, horizontal axis). Solid line and dots: the true and estimated bias magnitude for Exp G. Dashed line and open circles: the true and estimated bias magnitude for Exp D.

4.4. *Time-dependent forcing bias*

A final experiment is conducted in order to investigate the capability of the bias scheme to deal with slowly evolving forcing errors (Exp G). A typical example is an error in the seasonal cycle. Here the wind stress bias in the best-guess forcing of the previous sections is made to oscillate with an annual period. The spatial pattern is kept fixed in time. The 1.5-step bias scheme is again used to recover the bias from TAO temperature observations only, where now $\alpha = 0.5$. Figure 11 shows the time series of the spatially averaged amplitude of the bias estimate. It can be seen that the bias estimates follow the actual bias quite nicely, but with a slight delay of approximately 1 month. The magnitude of the bias estimate is rather good. Compared with the evolution of the bias estimate for Exp D, the variations around the true values are perhaps somewhat larger, which is most likely due to the larger value for α used in Exp G.

5. Discussion and conclusions

We have addressed the issue of bias estimation and correction in seasonal forecasting systems using an EnKF. Previous studies on bias treatment in oceanographic forecasting applications have mostly used single-model approaches in which a bias was estimated in elements of the model state vector. Here, the algorithm of Dee (2005) was adapted for use with an ensemble and further modified to estimate a bias in the boundary conditions (the surface wind stress). It was found that estimation and correction of the wind stress bias from assimilated temperature observations significantly improves the balance and accuracy of the forecast. The computationally efficient one-step bias scheme, and a modification with only partial re-computation of the Kalman gain matrix, were found to perform particularly well in this application. The balance of the analyses and accuracy of the forecasts appeared to be primarily related to the correct reconstruction of the peak wind stress bias amplitude. The experiments which used (partially) increased spatial scales were best able to reproduce the spatial pattern of the bias. Experiments E and F suggests that the combined use of a bias scheme with an ensemble filter enables better use of observations of sea level anomalies, by using the ensemble mean sea level as a reference level, rather than an (updated) climatological MSSH. The presented methodology can easily be applied with real observations to produce error-corrected surface forcing fields for ocean model studies, analogous to the work of Stammer et al. (2004) which was based on a variational method.

A bias in the model mean sea level may be reflected in systematic differences between model and observations in the position or strength of sea level anomalies associated with the position and strength of the mean currents. Only if these currents themselves can be associated with the forcing, or model parameters, and if additional observation types, such as temperature, are available, is there any hope of correcting the mean state through adjustment of forcing or parameters. Some other methods have been proposed in the past to deal with the problem of a proper reference level for sea level anomalies, in the absence of an accurate geoid. For example, Miller (1989) investigated the assimilation of differences between repeated sea level observations. More recently, two methods were shown to be able to reconstruct the mean from anomalies alone. Feron et al. (1998) showed that the statistics of altimetric anomalies (in particular the divergence of the eddy stresses) can be used to solve an equation for the mean dynamic topography starting from an initial guess based, for example, on hydrography. Van Leeuwen (1999) showed that the time-mean circulation could be computed by assimilation of anomalies with an ensemble smoother. In this application all states in the smoothing interval are estimated simultaneously using all observations taken during that interval. In both these applications information gathered over a time interval is used to correct the mean state retrospectively. Neither of these two methods is therefore suited for a sequential approach.

Some issues remain to be addressed. In particular, it remains somewhat unclear to what extent a bias in a prior MSSH estimate can be expected to be recovered from temperature profiles and sea level anomalies, both for the case where a wind stress bias is present and for the case where it is not. The alternative approach of using the ensemble-mean sea surface height as a reference level is based on the assumption that, when all biases are corrected, the ensemble-mean tends to the true time-mean. If the forcing perturbations do not contain all relevant time scales, including seasonal scales, this assumption will not be valid. While this was the case in Exp E, the impact of this error did not appear to be as serious as one might have anticipated beforehand.

It has not been the intention in this study to find optimal parameters for the bias schemes. It is likely that alternative choices for α and β, for the spatial localization scales, or for data sampling intervals will lead to improved bias reconstruction, especially under more realistic conditions, such as multiple time-dependent forcing and model errors, and a domain that includes the extra-tropics. Some of these issues have been addressed to some extent in recent work (Keppenne et al., 2005; Balmaseda et al., 2007). However, we submit that the approaches explored here provide additional advantages to those discussed in the literature.

6. Acknowledgments

The numerical computations were carried out with support from the European Centre for Medium-Range Weather Forecasts. The practical implementation of the EnKF was largely based on the package developed by Geir Evensen (NERSC). Thanks to Wilco Hazeleger (KNMI) for providing additional computational resources. I am grateful to the reviewers for their thoughtful comments on the original manuscript.

References

Balmaseda, M. A., Dee, D., Vidard, A. and Anderson, D. L. T. 2007. A multivariate treatment of bias for sequential data assimilation: application to the tropical oceans. *Q. J. R. Meteorol. Soc.* **133**, 167–179.

Bell, M. J., Martin, M. J. and Nichols, N. K. 2004. Assimilation of data into an ocean model with systematic errors near the equator. *Q. J. R. Meteorol. Soc.* **130**, 873–893.

Bosilovich, M. G., Radakovich, J. D., da Silva, A., Todling, R. and Verter, F. 2007. Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system. *J. Meteorol. Soc. Japan*, **85A**, 205–228.

Burgers, G., van Leeuwen, P. J. and Evensen, G. 1998. Analysis scheme in the Ensemble Kalman Filter,. *Mon. Wea. Rev.* **126**, 1719–1724.

Busalacchi, A. J., McPhaden, M. J., Picaut, J. and Springer, S. R. 1990. Sensitivity of wind-driven tropical Pacific Ocean simulations on seasonal and interannual time scales. *J. Mar. Syst.* **1990**, 119–154.

Carton, J. A., Chepurin, G. and Giese, B. 2000. A simple ocean data assimilation analysis of the global upper ocean 1950–95. Part I: Methodology. *J. Phys. Oceanogr.* **30**, 294–309.

Chepurin, G. A., Carton, J. A. and Dee, D. 2005. Forecast model bias correction in ocean data assimilation. *Mon. Wea. Rev.*, **133**, 1328–1342.

Dee, D. P. and Da Silva, A. K. 1990. Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* **117**, 269–295.

Dee, D. P. and Todling, R. 2000. Data assimilation in the presence of forecast bias: the GEOS moisture analysis. *Mon. Wea. Rev.* **128**, 3268–3282.

Dee, D. P. 2005. Bias and data assimilation. *Q. J. R. Meteorol. Soc.* **131**, 3323–3343.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.* **99**, 10 143–10 162.

Feron, R. C. V., De Ruijter, W. P. M. and van Leeuwen, P. J. 1998. A new method to determine the mean sea surface dynamic topography from satellite altimeter observations. *J. Geophys. Res.* **103**, 1343–1362.

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

Friedland, B. 1969. Treatment of bias in recursive filtering. *IEEE Trans. Autom. Control*, **AC-14**, 359–367.

Gaspari, G. and Cohn, S. E. 1999. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* **125**, 723–757.

Hackert, E. C., Busalacchi, A. J. and Murtugudde, R. 2001. A wind comparison study using an ocean general circulation model for the 1997–1998 El Niño. *J. Geophys. Res.* **106**, 2345–2362.

Houtekamer, P. and Mitchell, H. 2001. A sequential Ensemble Kalman Filter for atmospheric data assimilation. *Mon. Wea. Rev.* **129**, 123–137.

Huddleston, M. R., Bell, M. J., Martin, M. J. and Nichols, N. K. 2004. Assessment of wind-stress errors using bias corrected ocean data assimilation. *Q. J. R. Meteorol. Soc.* **130**, 853–871.

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

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., and co-authors, 1996. The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteorol. Soc.* **77**, 437–471.

Keppenne, C. L. and Rienecker, M. M. 2002. Initial testing of a massively parallel Ensemble Kalman Filter with the Poseidon isopycnal ocean general circulation model. *Mon. Wea. Rev.* **130**, 2951–2965.

Keppenne, C. L., Rienecker, M. M., Kurkowski, N. P. and Adamec, D. A. 2005. Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction. *Nonlin. Proc. Geo.* **14**, 1–13.

Leeuwenburgh, O., 2005. Assimilation of along-track altimeter data in the Tropical Pacific region of a global OGCM ensemble. *Q. J. R. Meteorol. Soc.* **131**, 2455–2472.

Leeuwenburgh, O., 2007. Validation of an EnKF system for OGCM initialization assimilating temperature, salinity, and sea surface height measurements. *Mon. Wea. Rev.* **135**, 125–139.

Marsland, S. H., Haak, H., Jungclaus, J. H., Latif, M. and Röske, F. 2003. The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates. *Ocean Modell.* **5**, 91–127.

Martin, M. J., Bell, M. J. and Nichols, N. K. 2002. Estimation of systematic crror in an equatorial ocean model using data assimilation. *Int. J. Numer. Methods Fluids*, **40**, 435–444.

Miller, R., 1989. Direct assimilation of altimetric differences using the Kalman Filter. *Dyn. Atmos. Oceans* **13**, 317–333.

Uppala, S. M., Kallberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., and co-authors, 2005. The ERA40 re-analysis. *Q. J. R. Meteorol. Soc.* **131**, 2961–3012.

Segschneider, J., Anderson, D. L. T. and Stockdale, T. N. 2000. Toward the use of altimetry for operational seasonal forecasting. *J. Climate* **13**, 3115–3138.

Stammer, D., Ueyoshi, K., Köl, A., Large, W. G., Josey, S. A., and co-authors. 2004. Estimating air-sea fluxes of heat, fresh water and momentum through global ocean data assimilation. *J. Geophys. Res.* **109**, C05023.

van Leeuwen, P. J. 1999. The time-mean circulation in the Agulhas region determined with the ensemble smoother. *J. Geophys. Res.* **104**, 1393–1404.

Vossepoel, F. C., Weaver, A. T., Vialard, J. and Delecluse, P. 2004. Adjustment of near-equatorial wind stress with four-dimensional variational data assimilation in a model of the Pacific Ocean. *Mon. Wea. Rev.* **132**, 2070–2083.

Yu, L. and O'Brien, J. J. 1991. Variational assimilation of the wind-stress drag coefficient and the eddy viscosity profile. *J. Phys. Oceanogr.* **21**, 709–719.