

Assimilation of along-track altimeter data in the Tropical Pacific region of a global OGCM ensemble

By O. LEEUWENBURGH*

Royal Netherlands Meteorological Institute (KNMI), De Bilt, the Netherlands

(Received XXX; revised XXX)

SUMMARY

Identical-twin experiments are performed with an ocean general circulation model ensemble to investigate the potential for correction of subsurface ocean model states through assimilation of altimetric sea level observations with the Ensemble Kalman Filter (EnKF). The EnKF provides a convenient extension to existing ensemble prediction systems. Observations are simulated for the Tropical Pacific by sampling a truth run at 10-day intervals at the TOPEX/POSEIDON along-track measurement points and adding realistic instrument and orbit errors. Ensemble spread is generated by perturbing the best-guess forcing fields. The perturbations are based on a multivariate EOF decomposition of differences between two reanalysis products. The effectiveness of the assimilation is investigated by comparison of the forecasts and analyses with a control run and with the truth. Time series of subsurface state variables along the equator show that the analyses are closer to the truth than the control in all cases, indicating a significant potential for improved ENSO forecast initialisation. A second assimilation run with an Ensemble Square-Root Filter (ESRF) shows that the analyses are very similar to those from the EnKF. However, ensemble spread in the subsurface state variables is found to be a poor proxy for the true analysis error in this experiment, in particular in the case of the ESRF. While the sea level analyses remain close to the truth, persistent offsets are introduced in the subsurface state, suggesting a role for bias correction schemes in ensemble methods.

KEYWORDS: Ensemble Kalman Filter Ensemble Square-Root Filter Assimilation Altimetry

1. INTRODUCTION

Seasonal forecasting is an emerging discipline on the boundary between weather forecasting and climate prediction. It currently aims primarily to exploit the relatively long timescales associated with the dominant low-frequency mode of variability of the world's ocean-atmosphere coupled system, ENSO (El Niño/Southern Oscillation). Theories of ENSO (e.g. Suarez and Schopf, 1988; Jin, 1997) and prediction studies (see Latif *et al.* (1998) for a review) suggest that the evolution and predictability of the coupled system in the tropics are strongly connected to the initial thermal state of the upper ocean and its subsequent evolution. Low-frequency, wind-driven variability in the tropical oceans is primarily associated with the dominant baroclinic mode (Fukumori *et al.* 1998), which describes the heaving and shoaling of the main thermocline and has a strong projection on sea level. Attempts to exploit sea level information were made in several studies on the initialisation of seasonal forecasts (Fischer *et al.* 1997; Ji *et al.* 2000; Segschneider *et al.* 2000), all of which found that the quality of the forecasts improved. In these studies sea level data was used to correct the model temperature, and in some cases, also salinity. For example, in the work of Fischer *et al.* (1997) and Ji *et al.* (2000) the sea level innovation is completely converted to a subsurface temperature correction. Vossepoel and Behringer (2000) subsequently demonstrated the potential for additional correction of the near-surface salinity field. The method developed by Cooper and Haines (1996) was used by Segschneider *et al.* (2000) and Alves *et al.* (2001) in their assimilation system to effectively lift and lower the T/S profile of the water column, leaving

* Corresponding author address: KNMI, PO Box 201, 3730 AE, De Bilt, The Netherlands. E-mail: leeuwenburgh@knmi.nl

© Royal Meteorological Society, 2004.

U and V to adjust during the forecast run. Fukumori *et al.* (1999) projected all sea level differences onto the barotropic and first baroclinic modes to correct T, S, U and V simultaneously. Cross-covariances between sea level, T, S, U and V were estimated from a model run by Borovikov *et al.* (2004) for use with a multi-variate OI scheme. Robert and Alves (2004) compared multivariate covariances estimated from a long model run with those estimated at a single instant from a model ensemble and identified significant differences associated with time-dependent adjustment processes such as wave dynamics. Keppenne and Rienecker (2003) showed how multivariate covariances associated with single temperature and salinity observations evolved over a 3 month period due to dynamical changes in a model ensemble. An assimilation method that combines all the advantages of time-evolving multivariate statistics is the Ensemble Kalman Filter (Evensen, 1994). This method furthermore provides an error estimate for the analysed state. In this paper the Ensemble Kalman Filter (EnKF) will be applied in an identical-twin experiment in which sea level is assimilated in the Tropical Pacific, with the aim of assessing the potential for improving initial conditions for ENSO forecasts. Given the potential importance of non-linear aspects for the evolution and magnitude of strong ENSO events (Jin *et al.* 2003), with the observation that only models based on the primitive equations tend to produce consistently high forecast skill across the spring barrier (van Oldenborgh *et al.* 2004), a global, fully non-linear ocean general circulation model will be used in this study.

The model will first be introduced briefly in Section 2. Section 3 subsequently gives an overview of the experiment setup. Section 4 and 5 discuss the generation of ensemble spread by forcing perturbation and the pseudo observations. The details of the EnKF algorithm implementation are explained in Section 6. The results of assimilation runs with the EnKF and with an alternative algorithm are presented in Section 7, and Section 8 summarizes the conclusions.

2. THE OCEAN MODEL

The model used in this study is the Max Planck Institut für Meteorologie Ocean Model MPI-OM version 1 (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 degrees 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 main changes in the physics with respect to the old HOPE model (Wolff *et al.* 1997) are in new parameterizations of sub-grid scale processes and details can be found in (Marsland *et al.* 2003). Surface exchanges of momentum, fresh water and heat are calculated using prescribed daily fields of surface stress, 10m wind speed, 2m air and dewpoint temperature, short-wave radiation, cloud cover and precipitation from the ERA40 and NCEP/NCAR (Kalnay *et al.* 1995) reanalyses. These fields were interpolated from the original grids to the model grid using bilinear interpolation. In all runs salinity was relaxed to Levitus climatology at all levels with a 3-year timescale. Temperature was relaxed only below the mixed layer while sea surface temperature (SST) was allowed to evolve freely to induce maximum possible spread in the ensembles.

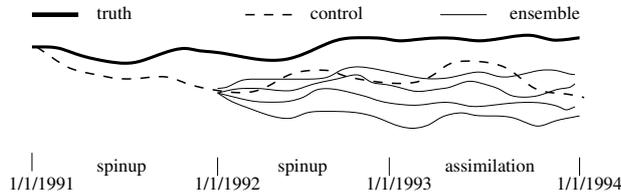


Figure 1. Schematic of the experiment setup. The ensemble members are spun off the January 1992 control state using perturbed forcing. Assimilation and validation is over 1993 only.

3. EXPERIMENT DESIGN

A twin experiment is conducted in which the true ocean state is defined by a forward run of the ocean model using NCEP/NCAR forcing fields and is therefore known exactly. ERA40 is assumed to be a best-guess estimate of the true forcing and a corresponding best-guess estimate of the ocean state is obtained by running the ocean model using ERA40 forcing fields. This run will be referred to as the 'control'. The initial conditions for both runs at the beginning of the experiment period (1993) were obtained by spinning up the model for 2 years, starting from the same model state, but forced by the two different reanalysis products (Figure 1). Plots of the true and control states show that after 2 years the two solutions have departed significantly as a result of using different forcing fields. Since the model is the same for both runs, the only sources of error in the control are the initial conditions and the forcing. These error sources are accounted for in a third run, the assimilation run. A best-guess initial condition for this run at the start of the experiment period, and the assumed uncertainty therein, will be given by the mean and spread of a model ensemble. The spread is obtained by running the ensemble forward for one year with perturbed ERA40 forcing, starting from the control. A result of this approach is that the ensemble members will be appropriately balanced both internally and with the wind field at the start of the experiment period. It is difficult to obtain spread in the deep layers from surface forcing perturbations alone, but since it is assumed that only upper ocean dynamics play a significant role in the tropical Pacific, this is not expected to cause a serious problem here. The method of perturbation of the ERA40 forcing fields will be described in the next section. The assimilation run consists of repeated 10-day forward integrations of the ensemble (the resulting mean states will be referred to as the *forecasts*), each followed by a filter step during which simulated sea level data is assimilated by the ensemble. The mean states of the resulting ensembles (the *analyses*) will be compared with the control and the truth to determine whether the assimilation has brought the model closer to the true state. An ensemble size of 40 members will be used, which is the size used by Kepenne and Rienecker (2002).

4. REPRESENTATION OF UNCERTAINTY IN THE FORCING

The statistics of the true forcing errors for this experiment could be estimated from a comparison between the ERA40 and NCEP/NCAR reanalyses over the experiment period. In reality, however, these statistics will not be available since the true forcing is unknown. The differences between ERA40 and ERA15 will therefore be used as proxies for the true forcing errors. An alternative approach

would be to assume that typical errors in ERA40 are proportional to differences between two successive 2-week averages (Bonekamp *et al.*, 2001).

There is no way of knowing if the use of ERA40-ERA15 statistics as a proxy for ERA40-NCEP/NCAR statistics is a particularly good or bad approximation of a real application where either one of these statistics is used as a proxy for those of the true forcing errors. Crucial will be whether the used error statistics are too optimistic or too pessimistic, resulting in respectively too much or too little weight given to the model forecast.

It is argued here that if the proposed approach does produce error statistics which are too optimistic, leading to analysis errors which are smaller than one might expect in real applications, this impact will be smaller than that associated with the neglect of model error in an identical-twin experiment which uses only a single model.

A 2-year record of differences between the ERA40 and ERA15 reanalyses is used to obtain a statistical description of error characteristics in zonal and meridional surface stress, air temperature, dew point temperature, and short-wave radiation. The annual and semi-annual cycles were first removed and all variability with periods shorter than 20 days was filtered out by application of a Loess smoother (Cleveland and Devlin, 1988). This latter filter step is roughly equivalent to application of a 14-day running mean (Schlax and Chelton, 1992). A combined EOF decomposition is performed on the resulting difference fields, after normalizing each variable with its standard deviation. This approach is similar to that of Robert and Alves (2004) who used EOF analysis to identify the leading orthogonal modes of intra-seasonal ERA-NCEP wind stress differences. Only differences in the 60°S - 70°N latitude band were considered, and values were gradually scaled down to zero poleward of 40°S and 50°N. The combined EOF decomposition results in individual error patterns for the different forcing variables that evolve identically over time. This allows for the representation of cross-covarying features in the variables and decreases the possibility that the perturbations of different forcing fields have opposing effects on surface temperature, which would reduce their effectiveness in increasing ensemble spread. The first EOF of the filtered record represents about 7% of the total variance. In their single 10-day ensemble run with an OGCM, Robert and Alves (2004) used a constant perturbation of wind stress. An easy method, described by Evensen (2003), has been implemented here to enforce correlation between random daily perturbations on time scales corresponding to those found in the ERA40-ERA15 time series.

5. SEA LEVEL MEASUREMENTS

The main source of sea level data are satellite altimeters, which have provided continuous measurements of global sea level since the launch of ERS-1 in 1991 and TOPEX/POSEIDON (T/P) in 1992. Altimetry measurements were simulated here by sampling the truth run along the ground tracks of the T/P satellite. After subtracting a mean model sea level, the resulting anomalies were assimilated at 10-day intervals. Since the mean sea level was identical for all runs, total sea level is effectively assimilated. This adds the possibility of correcting offsets associated with persistent bias in the forcing over the experiment period. The actual satellite tracks were subsampled at roughly 100km intervals to reduce the number of measurements to manageable quantities. All along-track samples from a 10-day

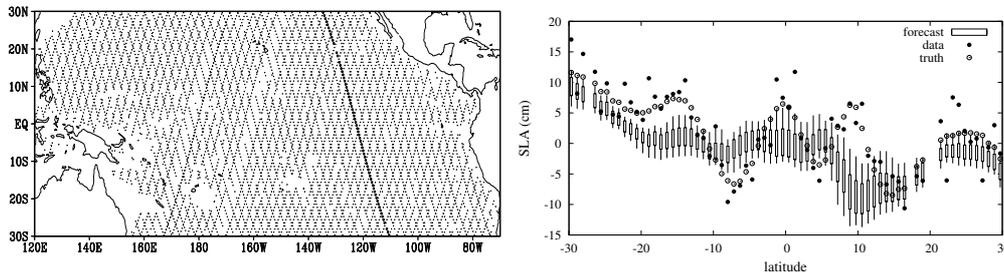


Figure 2. (left) Distribution of altimetric sea level measurements on 8 January 1993. (right) Sea level along the arc indicated with the heavy dots (see left panel) at the first assimilation step. The spread of the forecast ensemble is indicated by the mean plus and minus 1 standard deviation (boxes) and the minimum and maximum value (lines extending from each box).

T/P repeat cycle were assumed to have been measured at the assimilation time. Including ERS data and using a more realistic temporal sampling would result in a more strongly varying spatial data coverage which would make interpretation of the results more difficult. Figure 2 shows the resulting sampling pattern. Along-track data in principle have the advantage over gridded data fields that the character of the observation errors is better understood. This advantage may disappear to some extent when representation error comes into play as will be the case with real data. Two kinds of measurement error were simulated. Uncorrelated Gaussian distributed instrument noise with a standard deviation of 3.5 cm was added to each sample. A random but constant along-track bias with a standard deviation of 2 cm was also added to each arc to simulate orbit errors which are dominated by wavelengths of one cycle per revolution (Chelton *et al.* 2001). These numbers are in the range of error estimates for the T/P and ERS missions. Standard deviations of model sea level variability within a few degrees of the equator are typically between 8 and 15 cm, but may be as low as 2 to 3 cm further poleward. Figure 2 shows the true sea level, the simulated measurements, and ensemble sea level forecasts along a T/P arc at the first assimilation time. There are significant biases in the forecast at the initial assimilation time and substantial uncertainty as indicated by the large spread in the ensemble.

6. ANALYSIS ALGORITHM

The analysis ensemble is formed using the stochastic EnKF algorithm described by Evensen (1994) and Burgers *et al.* (1998). Covariance localisation is implemented following the suggestions of Gaspari and Cohn (1999). The resulting algorithm is described by

$$\begin{aligned} \mathbf{A}^a &= \mathbf{A} + \mathcal{C} \circ (\mathbf{A}' \mathbf{A}'^T \mathbf{H}'^T) [\mathcal{C} \circ (\mathbf{H} \mathbf{A}' \mathbf{A}'^T \mathbf{H}'^T) + \mathbf{R}]^{-1} (\mathbf{D} - \mathbf{H} \mathbf{A}) \\ &= \mathbf{A} + \mathbf{K} (\mathbf{D} - \mathbf{H} \mathbf{A}) \end{aligned} \quad (1)$$

where $\mathbf{A} = (\psi_1, \dots, \psi_N)$ holds the ensemble of model forecasts, $\mathcal{C} \circ$ denotes the Schur product operator, \mathbf{H} is the measurement operator, and \mathbf{D} is the ensemble of perturbed measurements. Primes indicate anomalies with respect to the ensemble mean. Both the uncorrelated instrument error and the along-track bias are accounted for in the observation error covariance matrix \mathbf{R} . Applying the Schur product in observation space avoids the necessity to calculate the

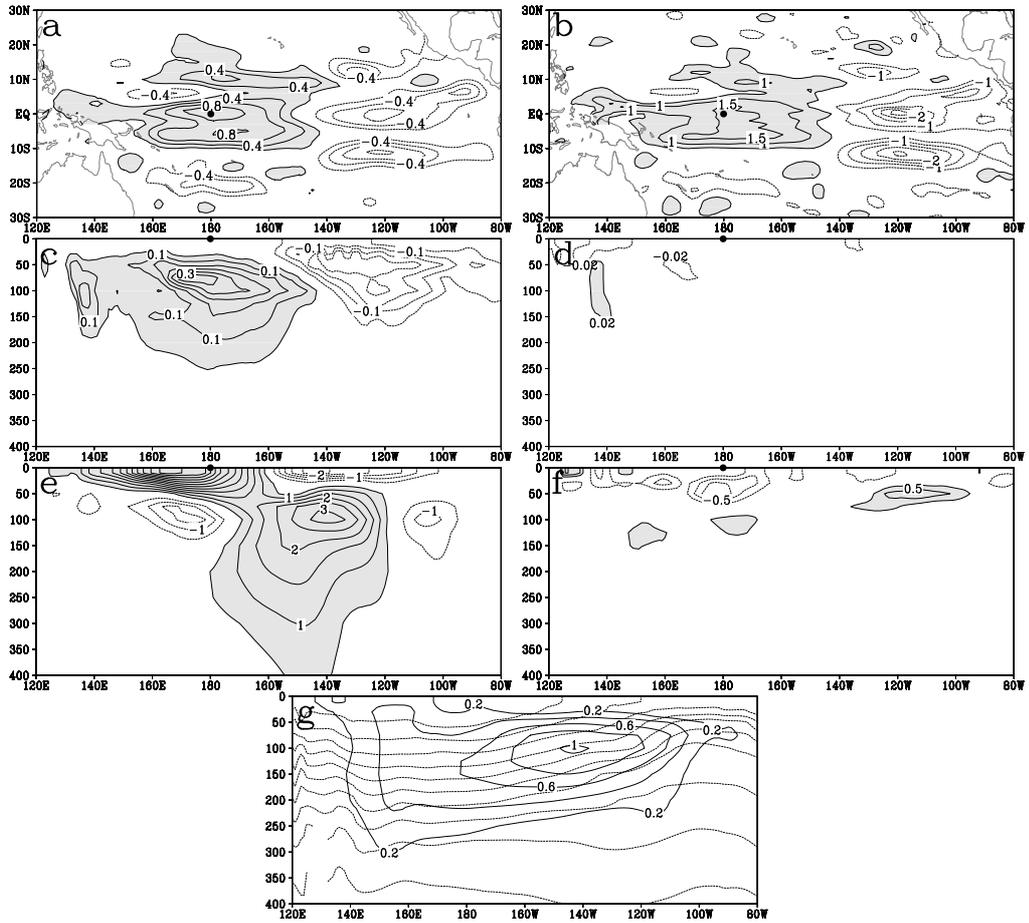


Figure 3. Analysis increments associated with a +1 cm sea level innovation located at (180°E, 0°N) on 6 June 1992 for (a) sea level (cm), (b) 20°C isotherm depth (m), (c) temperature (°C), (d) salinity (psu), (e) zonal velocity (cm/s), and (f) meridional velocity (cm/s). The zero contour is not drawn, negative contours are dashed, and values above the first positive contour are shaded. (g) Ensemble mean temperature at 2°C intervals (dashed, the 20°C contour is drawn thicker), and zonal velocity (solid lines) in m/s. Panels (c) to (g) show equatorial sections.

ensemble covariances $\mathbf{P} = \mathbf{A}'\mathbf{A}'^T / (N - 1)$ explicitly in model space, which would be too costly (Houtekamer and Mitchell, 2001). The filter solution is calculated independently for each grid point, where only observations within an elliptic region around the grid point are used. The half axes of the search ellips are 30° in the zonal and 15° in the meridional direction and a scale of 500m was used in the vertical (Keppenne and Rienecker, 2002). 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).

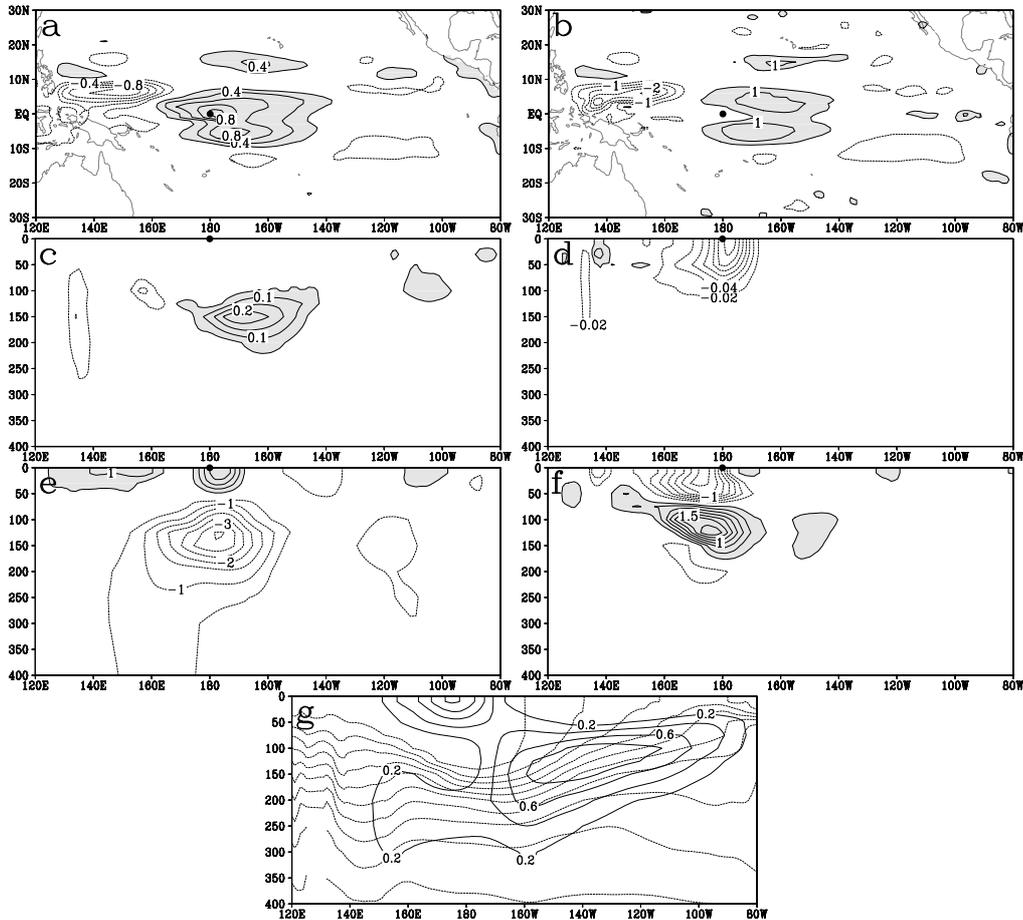


Figure 4. As in Fig. 3 but for 8 January 1993.

7. RESULTS

(a) Single observation analysis increments

Analysis increments associated with a single observation are calculated to illustrate the time-evolving character (flow-dependence) of the ensemble-based model covariances (Figs. 3 and 4). The increments presented here correspond to corrections to the mean model state obtained by assimilation of a single +1 cm sea level innovation (model-data difference) located at (180°E, 0°N). Fukumori *et al.* (1999) presented analysis increments corresponding to sea level innovations at selected locations using a time-asymptotic estimate of the error covariance matrix of the analyzed state. While their Extended Kalman Filter did account for a time-varying data distribution, it did not do so for time and space variations in errors in the forecast state estimate. The increments were further restricted to projections on the barotropic and first baroclinic modes. Instead the ensemble approach will be used here, as was done by Keppenne and Rienecker (2003) for single temperature and salinity observations. Since the innovation is a scalar with value 1, the increments correspond to Kalman gain structures captured in \mathbf{K} . The gains are computed from the full ensemble at two different times separated by

half a year. The Schur product was not applied in the calculation of the gains, and the sea level observation error variance was assumed to be zero. Comparing Figs. 3 and 4, a significantly different correction on the subsurface state would result from the same sea level innovation. A positive sea level innovation in June 1992 will result in a positive offset in the temperatures between the mixed layer and the upper thermocline around the date line, and with a negative offset in the eastern half of the basin, corresponding to a correction of the bias in the equatorial tilt of the thermocline. A similar sea level innovation in January 1993 however, can be associated with a fairly localised deepening of a warm pool (temperature $> 28^{\circ}\text{C}$) extending to approximately 160°E . This correction takes place about 75 meters deeper and 10° further east as compared to the maximum temperature change in June 1992. A clue towards the origin of the January 1993 correction can be obtained by examining the subsurface corrections in the horizontal plane. The correction in the depth of the 20° isotherm (Z20, a proxy indicator for the position of the thermocline near the equator) shows a basin-wide equatorial gradient in June 1992, whereas two distinct lobes with maximima located just north and south of the equator, reminiscent of a Rossby wave, are found in January 1993. Adjustments to salinity and zonal and meridional velocities are distinctly different at the two times as well. High sea level on the date line in June 1992, for example, is associated with zonal convergence east of 170°E in the surface mixed layer, as well as with increased strength of the Equatorial Under Current (EUC) east of the date line. In January 1993, however, this same sea level adjustment would result in an effective weakening of the EUC between 170°E and 150°W . It is obvious from these gains that different temperature and salinity corrections can be associated with the same sea level innovation. This is because the strong temperature and salinity fronts, the thermocline and the edge of the western Pacific fresh water pool, are located in different positions. Relative displacements of these fronts in the ensemble contribute most strongly to the multivariate covariances. Since it is not possible to determine from a single sea level observation whether a displacement of either the temperature or salinity front alone is the most likely cause for the model-observation misfit, both will be adjusted partly. Assimilating additional observations can help differentiate between probable causes. The changing character of the Kalman gains primarily reflects the intra-seasonal evolution of the mean state, and to a lesser extent, the time-integrated impact of the quasi-random ensemble perturbations. If, for example, the perturbations are only able to induce ensemble spread in the surface layer, and not at the thermocline depth, only variables in the surface layer will be adjusted. In the context of the discussion above, this would mean that sea level would be adjusted by a zonal displacement of the surface salinity front, rather than by a vertical displacement of the thermocline. While Figs. 3 and 4 highlight the impact of seasonal variability on model covariances, similar, or even larger differences can be expected between El Niño and La Niña states. The figures clearly show the potential benefit of time-evolving multivariate covariances and give a justification for using a relatively expensive assimilation method like the EnKF.

(b) *A single assimilation step*

A single assimilation step was performed in order to provide additional insight into the character of the corrections when many observations are available and thus into the potential for adjustment of subsurface quantities. Figure 5 compares

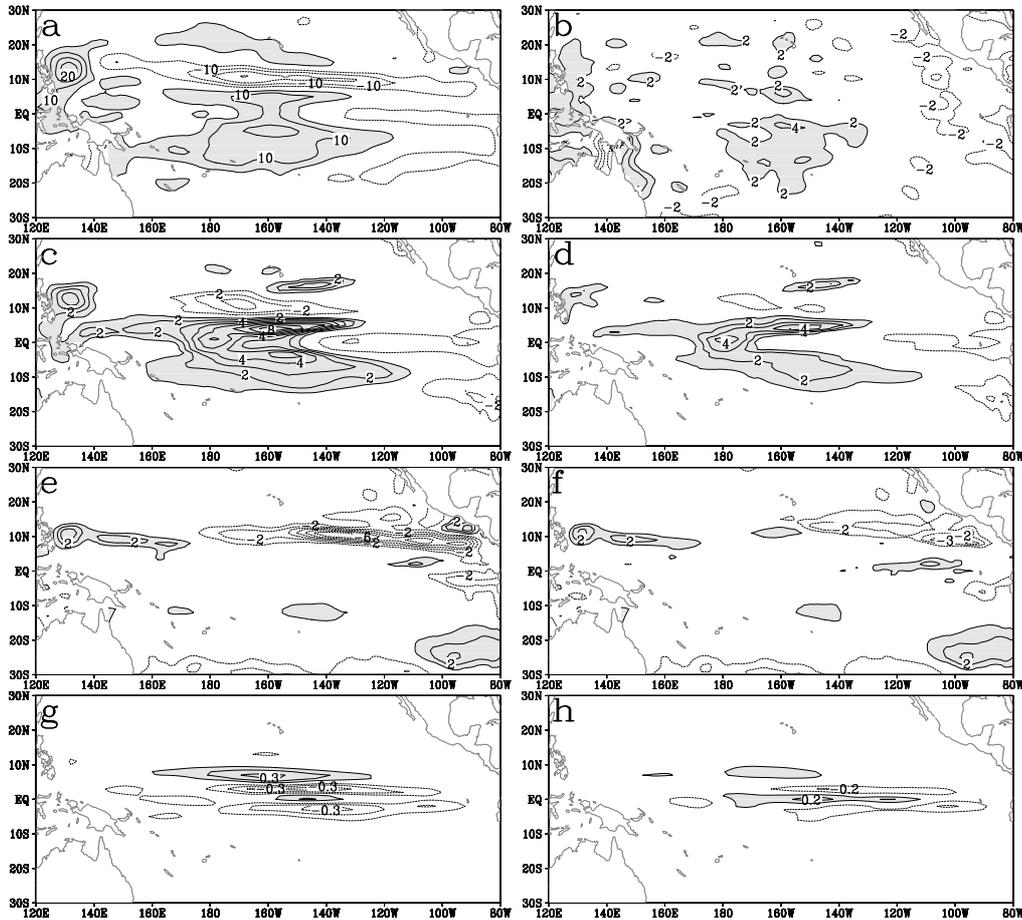


Figure 5. Errors in the forecast (left) and analysis (right) at 8 January 1993 for (a,b) sea level (cm), (c,d) temperature ($^{\circ}\text{C}$) at 150m depth, (e,f) temperature ($^{\circ}\text{C}$) at 50m depth, (g,h) zonal velocity (m/s) at 125m depth. The zero contour is not drawn, negative contours are dashed, and values above the first positive contour are shaded.

the forecast and analysis with the truth at the first assimilation step (8 January 1993). Forecast errors in sea level are as large as 10 to 15 cm in large parts of the Tropical Pacific, and correspond largely to a mean sea level bias between control and truth over the experiment period. The assimilation is very effective in reducing these errors. The sea level error in the analysis has a fairly smooth character and is typically smaller than the prescribed measurement error of 3.5 cm. Temperature errors in the forecast are as large as 6°C , both at 50m and 150m depth, and are associated with forecast errors of the corresponding sign in the sea level. For example, the large region of positive temperature error at 150m depth in the central Pacific (panel 5c) corresponds to a similarly shaped region of positive errors in the sea level forecast (Fig. 5a). The narrow band of negative errors centered at 10°N is found both in the sea level and temperature at 50m depth. These errors are primarily associated with the position of the thermocline which is typically situated around 150m depth in the central and western part of the basin and around 50m depth in the eastern part. Temperature

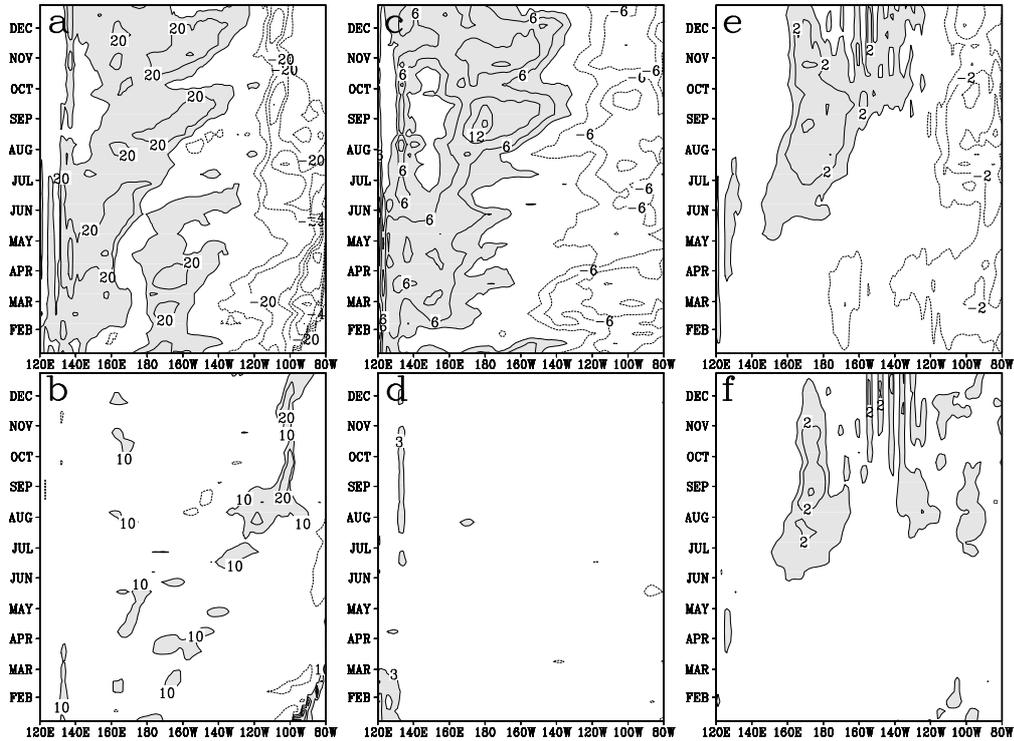


Figure 6. Time-longitude diagrams of equatorial differences between control and truth (top), and between the EnKF analysis and the truth (bottom) for (a,b) depth of the 20°C isotherm in meters, (c,d) sea level in cm, (e,f) temperature at 50m depth in °C. The zero contour is not drawn, negative contours are dashed, and values above the first positive contour are shaded.

errors are effectively reduced at both levels, corresponding to a correction of Z20, particularly within the latitude band between 15°S and 15°N.

The forecast error in the zonal current velocity at 125m depth has a banded pattern, reflecting the zonal character of the equatorial current system. The southern and northern branches of the westward flowing South Equatorial Current, as well as the eastward flowing North Equatorial Counter Current are too strong. These errors are corrected to a large extent in the analysis, although the assimilation has increased the strength of the Equatorial Under Current (EUC) too much, resulting in a positive band of analysis errors along the equator in the central part of the basin. The strength of the EUC is intimately linked to the balance between the vertical distribution of the eastward pressure force associated with the basin-wide surface slope, frictional forces and inertial accelerations (Pedlosky, 1996). The disturbance of this balance by the assimilation will be discussed further in section (d).

(c) One-year assimilation run

Figures 6 and 7 compare the analysis errors from a one-year long assimilation run with the errors in the control run. The improvement in sea level is immediately obvious but is to be expected. It is clear that sea level and Z20 are strongly related on the equator (compare Figs. 6a and c). Since a general circulation model is used here, the thermocline depth is not a model variable in the same sense as it is

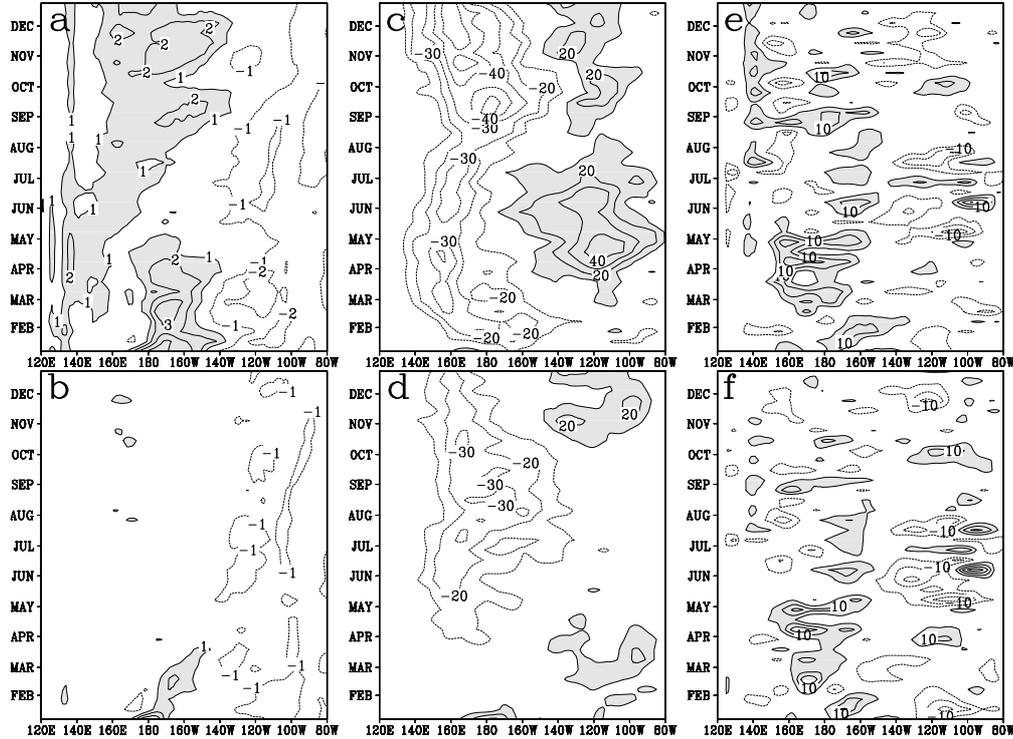


Figure 7. As in Fig. 6, but for (a,b) temperature at 150m ($^{\circ}\text{C}$), (c,d) zonal velocity at 75m (cm/s), (e,f) meridional velocity at 75m (cm/s).

in a 2-layer model and has to be determined from the subsurface temperature profile, using Z20 as a proxy. The Z20 results therefore imply that the sea level assimilation has significantly improved the subsurface temperature distribution. This improvement is primarily found in the temperature at 150m depth (Figs. 7a and b), but is less pronounced at 50m depth (Figs. 6e and f). The zonal equatorial velocity has improved most clearly between April and July in the eastern part of the basin, and all through the year in the western part. There is some indication that the zonal velocity analyses are not as good as the control in the central Pacific from June to August. The meridional velocity component has generally improved in the analysis, except perhaps east of 130°W , but the signal is too noisy to make quantitative statements based on these figures. On the whole it can be concluded that the equatorial analyses are significantly closer to the truth than the control run, and that the sea level assimilation has succeeded in correcting all subsurface ocean state variables to some extent. A more quantitative assessment of these results is presented in Fig. 10. It could be noted here that the assimilation of sea level alone can only result in limited improvement. Temperature and salinity observations can provide additional information on higher baroclinic mode structures, and can in real applications also lead to a better representation of the thermocline, which is generally overdispersive in ocean models.

It should also be determined whether the estimated uncertainties in the analyses (based on the ensemble spread) are consistent with the above results. An appropriate spread that would be consistent with all assumptions underlying

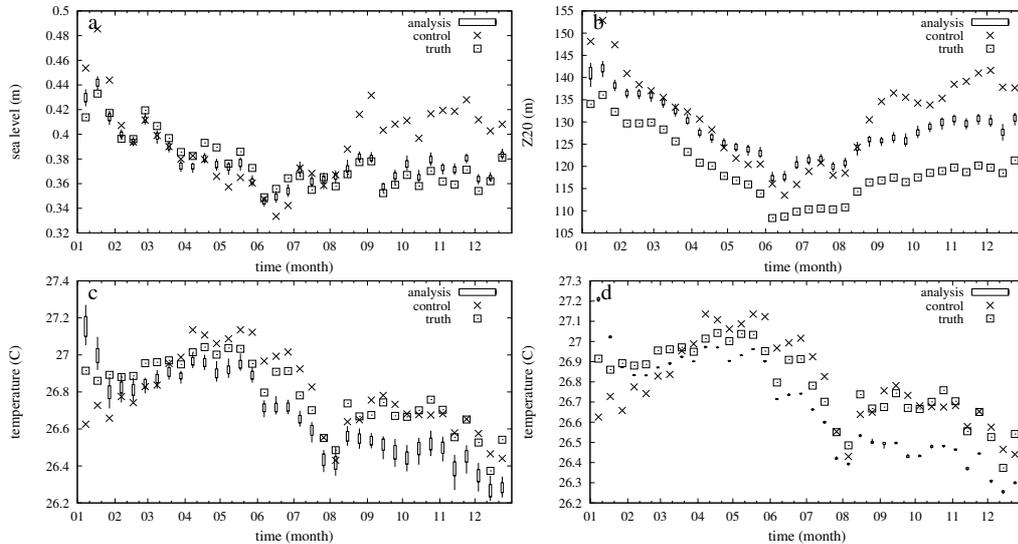


Figure 8. Time series of the truth, control, and analysis from the EnKF for Niño3.4 averages of (a) sea level, (b) Z20, (c) SST. (d) Time series of the truth, control, and analysis from the ESRF for Niño3.4 averages of SST. The range spanned by the analysis plus and minus 1 standard deviation of the ensemble spread is indicated by the boxes, minimum and maximum values are indicated by the lines extending from the boxes.

the EnKF, would be one that is comparable to the true error (the difference with the truth). Figures 8a to c show the time evolution of the true sea level, Z20 and surface temperature (SST) averaged over the Niño3.4 region (190°E - 240°E , 5°S - 5°N), as well as the corresponding values from the control and from the assimilation run, where the analysis uncertainty, as estimated from the ensemble spread, has been indicated. Sea level and Z20 are closer to the truth than the control during the entire period of the experiment, while the analysed SST appears to depart further from the truth than the control. For Z20 and SST, however, there are long periods where the true values fall completely outside the range covered by the ensemble (Gaussian statistics would allow for the truth to be outside of the one standard deviation range for approximately 30% of the time). It must be concluded that ensemble spread is not a good indicator for uncertainty during this run. This is problematic when the ensemble is used in a forecast system where one would like to have a reliable measure of uncertainty. Furthermore, underestimation of the true forecast error will unduly increase the relative weight given to the model forecast with respect to the observations and thus decrease the efficiency of the assimilation. In the next section this ensemble defect is discussed in some more detail.

(d) Ensemble spread and bias

The apparent inconsistencies between the analyses, the associated analysis ensemble spread, and the truth can in principle be ascribed to either an undervariability in the ensemble or to systematic errors (bias). The problem of maintaining representative ensemble spread has been discussed in earlier (atmospheric) studies. Houtekamer and Mitchell (1998) suggested that it is associated with the use of a single ensemble for both the calculation of the gain and the analysis error.

They found that for small ensemble sizes the ensemble spread consistently tends to underestimate the true error, and they managed to reduce this effect by using a double ensemble. Anderson and Anderson (1999) introduced an inflation factor to increase the deviation of ensemble members from the mean. The optimal value of this factor was determined by Hamill *et al.* (2001) to be a decreasing function of ensemble size. Van Leeuwen (1999) showed that small errors in the error covariance estimates, associated with the use of a finite ensemble size, introduce a tendency towards decreasing spread that is enhanced during the assimilation run. This may ultimately lead to almost complete rejection of the observations in favour of the model forecast, a situation commonly referred to as filter divergence.

Evensen (2004) suggested selecting the N dominant eigenvectors of the covariance matrix constructed from a very large ensemble to increase the rank of the initial ensemble. However, even if the initial ensemble has full rank, a nonlinear model integrator could introduce a loss of rank, or inbreeding, which will in turn lead to filter divergence. Figures 8a to c, however, show no indication of a progressively decreasing ensemble spread. Rather, the spread is already too small after the first analysis step and remains at the same level throughout the run, which would suggest that the forcing perturbations are efficient in counteracting any tendency for the spread to decrease, and that the relatively small ensemble size is not primarily responsible for the small spread. This is in agreement with Fig. (3) from Houtekamer and Mitchell (1998) which shows that increasing the ensemble size reduces the true error, but does not affect the ensemble spread very much. Note, however, that Fig. (3) of Houtekamer and Mitchell (2001) does show an increase in ensemble spread with size of similar magnitude as the true error reduction.

A second potential cause for the observed inconsistencies is the presence of bias. The persistent offsets between analyses and truth for SST and Z20 in the Niño3.4 region appear to be related to the biases observed in Figs. 6b and 6f in the eastern Pacific (note that the bias in the observed variable, sea level, is actually removed). Since the same model has been used in all runs, model bias itself can be excluded as the underlying cause. However, it is possible that the assimilation has introduced a bias in the subsurface state. It has been shown that assimilation of high-quality observations (both temperature and sea level) can offset the balance between the near-surface pressure gradient and the relatively inaccurate surface wind stress, and that this leads to a systematic error (Bell *et al.*, 2004). Unbalanced analysis increments may also induce spurious circulations which tend to erase the increments. This has led to the implementation of OI methods that produce geostrophically balanced increments on the equator (Burgers *et al.*, 2002). With the local-analysis implementation of the EnKF, there is no guarantee that balance will be maintained, and in fact, application of the Schur product has been shown to disturb linear balances (see e.g. Lorenc, 2003). Given that the bias in this experiment results from the assimilation rather than the model itself, it seems appropriate to include a bias correction scheme in the assimilation loop (Dee and Da Silva, 1998). Such schemes have so far not yet been implemented in ensemble-based assimilation systems.

(e) *Stochastic versus deterministic ensemble filters*

During recent years some studies have advocated deterministic ensemble algorithms as alternatives to the stochastic algorithm used here (see Tippet *et al.* (2003) for a review) . An advantage of these filters is that no perturbation of

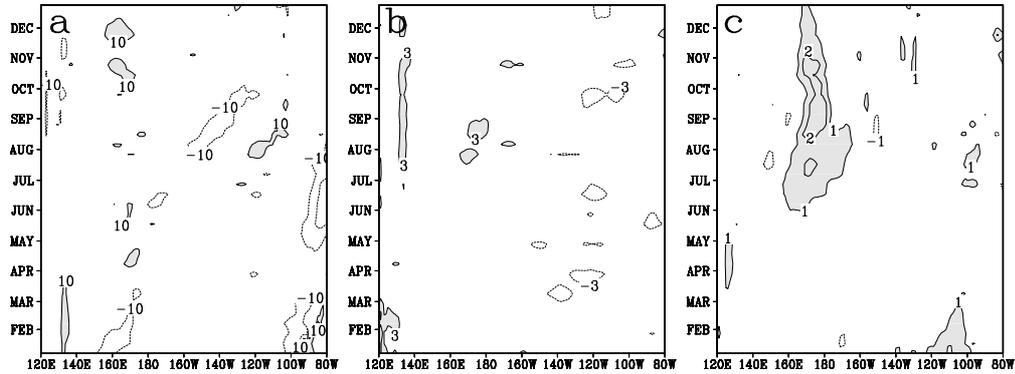


Figure 9. Time-longitude diagrams of equatorial differences between the ESRF analysis and the truth for (a) Z20 in meters, (b) sea level in cm, (c) temperature at 50m depth in °C. The zero contour is not drawn, negative contours are dashed, and values above the first positive contour are shaded.

the observations is required which avoids introducing noise into the system. An implementation of such a deterministic Ensemble Square-Root Filter (ESRF), described in Section 3 of Evensen (2004), is used here in a second assimilation run. Since no Schur product is used, the search ellips axes were halved in order to obtain an approximately equivalent effective number of observations for each local grid point analysis as for the EnKF. The resulting analysis errors for this run are shown in Fig. 9 for Z20, sea level and temperature at 50m depth, and can be compared to Fig. 6. The ESRF analyses are very close to the truth, and on some occasions better than the EnKF analyses. This can be related to the fact that the ESRF uses unperturbed observations to update the mean, while for small ensemble sizes the observation perturbations used by the EnKF will not average out to zero, which introduces noise into the system. Figure 8d shows the corresponding time series for the Niño3.4 box average of sea surface temperature. While the analyses themselves are very similar to those from the EnKF (Fig. 8c), the ensemble spread is even smaller, implying an even greater underestimation of analysis error. These results suggest that there might be an advantage to using the stochastic EnKF algorithm in that it suffers less from an underestimation of the true errors than the deterministic algorithms, and thus provides a more accurate indication of forecast uncertainty.

The statistics of sea level, Z20 and SST, averaged over four Niño boxes are summarized in Fig. 10 for both assimilation runs. Apart from the Niño3.4 SST and the Niño3 Z20 all analyses are better than the control. For the EnKF run, the analyses have on average improved sea level and SST relative to the forecast, while Z20 was only improved in the Niño4 box. The mean errors from the EnKF and ESRF are comparable, while the latter has consistently produced better Z20 analyses. The ensemble spread has in all cases been a poor indicator of the true error, with the exception of the Niño1.2 Z20 index, where the ensemble spread resulting from the ESRF is significantly smaller than that from the EnKF. The poor Z20 analyses in the Niño3 box appear to be related to relatively poor temperature analyses near the surface in the eastern Tropical Pacific, as illustrated by Figs. 6a and b and Figs. 6e and f. Low temperatures at 50m depth in the control are overcompensated in the analyses, translating into a positive bias in the Z20 analysis, as discussed in the previous section.

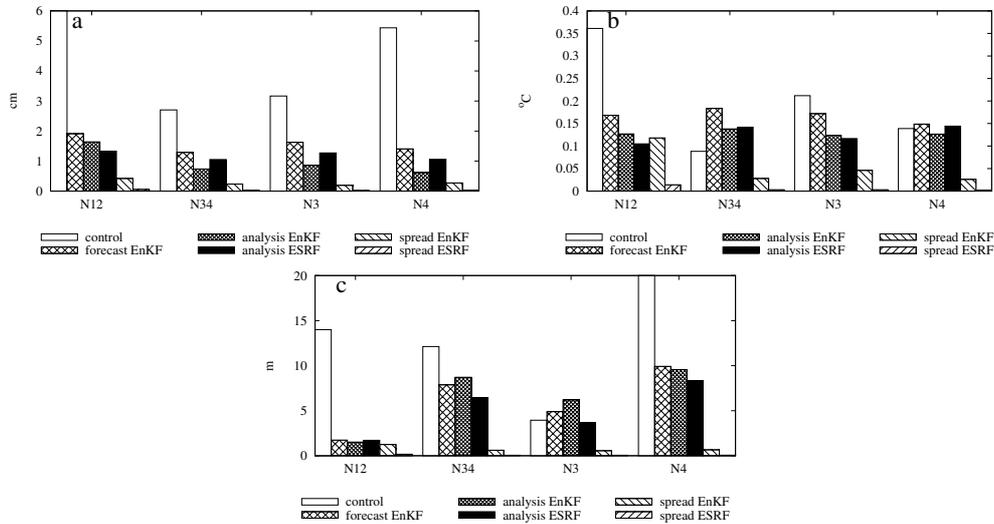


Figure 10. Mean absolute error in the control, EnKF forecast and analysis, and ESRF analysis, and the mean standard deviation of the ensemble spread from the EnKF and the ESRF analyses in 4 Niño boxes for (a) sea level, (b) SST, and (c) Z20. Niño1.2: 270°E-280°E,10°S-0°S, Niño3.4: 190°E-240°E,5°S-5°N, Niño3: 210°E-270°E,5°S-5°N, Niño4: 160°E-210°E,5°S-5°N.

8. CONCLUSIONS

The Ensemble Kalman Filter has been used to assess the potential for improvement of subsurface ocean analyses by assimilation of altimetric sea level observations in the Tropical Pacific domain of a global OGCM. An identical-twin experiment setting was used in which a truth was created by forcing the model with NCEP reanalysis fields, while the control and assimilation runs were forced with ERA40 fields. A method to represent uncertainty in the forcing was introduced that uses the spatial fields and temporal correlation scales of principal components from a combined EOF decomposition of differences between reanalysis products.

It was demonstrated that multivariate covariance patterns can differ substantially on a seasonal time scale, and it could be inferred that similar differences would be associated with ENSO events. Since the success of the assimilation depends on the prior statistics of both observations and forecast, this flow-dependence is the main advantage of the EnKF over assimilation methods which use fixed covariances.

Sea level observations were simulated along TOPEX/POSEIDON tracks, and appropriate instrument errors were added. A year long assimilation run was performed with the EnKF during which observations were assimilated every 10 days. The choices for ensemble size and data selection criteria were similar to those found in the literature. The results indicate that sea level assimilation leads to significant improvements along the equator in all subsurface fields relative to an unconstrained control run. One-year error averages based on 36 assimilation steps show reduced sea level errors in 4 Niño boxes. SST (Z20) has improved in 2 (3) of the 4 Niño boxes.

The spread of the analysis ensemble during the run was found to be a too optimistic measure of model uncertainty. At the same time, no indication was

found of ensemble collapse, implying that the forcing perturbation was sufficient to counteract any tendency towards increasing filter divergence associated with the use of a small ensemble. The introduction of imbalance by the assimilation procedure was suggested as a cause for persistent biases in the analyzed subsurface states. A second run was performed with an ensemble square-root filter. This run produced analyses of comparable quality, but with even smaller ensemble spread. Even with a very small spread though, the multivariate covariances can be determined from the model ensemble, and thus retain their flow-dependence. However, added value will result from a method that produces reliable error estimates by maintaining ensemble spread at the true error level for small ensemble sizes. Also, the presence of persistent biases in the analyses suggests that the implementation of bias correction schemes with ensemble methods must be considered whenever they are applied in the tropics.

ACKNOWLEDGEMENTS

The author would like to thank Gerrit Burgers (KNMI) for many useful discussions and suggestions, Geir Evensen and Laurent Bertino (NERSC) for the EnKF software, Helmuth Haak (MPIfM) and Noel Keenlyside (IfM) for their help with the ocean model, and Alberto Troccoli (ECMWF) for the forcing fields. The careful reading and insightful comments of one referee in particular have led to significant improvement of the manuscript. This is a contribution to the EU-funded ENACT project.

REFERENCES

- | | | |
|---|------|--|
| Alves, J. O. S., Haines, K. and Anderson, D. L. T. | 2001 | Sea level assimilation experiments in the Tropical Pacific. <i>J. Phys. Oceanogr.</i> , 31 , 305–323 |
| Anderson, J. L. and Anderson, S. L. | 1999 | A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. <i>Mon. Wea. Rev.</i> , 127 , 2741–2758 |
| Bell, M. J., Martin, M. J. and Nichols, N. K. | 2004 | Assimilation of data into an ocean model with systematic errors near the equator. <i>Q. J. R. Meteorol. Soc.</i> , 130 , 873–893 |
| Bonekamp, H., van Oldenborgh, G. J. and Burgers, G. | 2001 | Variational assimilation of TAO and XBT data in the HOPE OGCM, adjusting the surface fluxes in the tropical ocean. <i>J. Geophys. Res.</i> , C106 , 16693–16709 |
| Borovikov, A., Rienecker, M. M., Keppenne, C. L. and Johnson, G. C. | 2004 | Multivariate error covariance estimates by Monte-Carlo simulation for assimilation studies in the Pacific Ocean. <i>submitted to Mon. Wea. Rev.</i> |
| Burgers, G., van Leeuwen, P. J. and Evensen, G. | 1998 | Analysis scheme in the Ensemble Kalman Filter. <i>Mon. Wea. Rev.</i> , 126 , 1719–1724 |
| Burgers, G., Balmaseda, M. A., Vossepoel, F. C., van Oldenborgh, G. J. and van Leeuwen, P. J. | 2002 | Balanced ocean-data assimilation near the equator. <i>J. Phys. Oceanogr.</i> , 32 , 2509–2519 |
| Chelton, D. B., Ries, J. C., Haines, B. J., Fu, L.-L. and Callahan, P. S. | 2001 | ‘Satellite altimetry’. Pp. 1–131 in <i>Satellite altimetry and earth sciences</i> . Ed. L.-L. Fu and A. Cazenave, Academic Press, San Diego |
| Cleveland, W. S. and Devlin, S. J. | 1988 | Locally weighted regression; An approach to regression analysis by local fitting. <i>J. Am. Stat. Assoc.</i> ,—bf 83 596–610 |
| Cooper, M. and Haines, K. | 1996 | Altimetric assimilation with water property conservation. <i>J. Geophys. Res.</i> , 101 , 1059–1077 |
| Dee, D. P. and Da Silva, A. K. | 1998 | Data assimilation in the presence of forecast bias. <i>Q. J. R. Meteorol. Soc.</i> , 117 , 269–295 |
| Evensen, G. | 1994 | Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. <i>J. Geophys. Res.</i> , C99 , 10143–10162 |

- Evensen, G. 2003 The ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dyn.*, **53**, 343–367
- Evensen, G. 2004 Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dyn.*, **54**, 539–560
- Fischer, M., Flügel, M. and Ji, M. 1997 The impact of data assimilation on ENSO simulations and predictions. *Mon. Wea. Rev.*, **125**, 819–829
- Fukumori, I., Raghunath, R. and Fu, L.-L. 1998 The nature of global large-scale sea level variability in relation to atmospheric forcing: A modeling study. *J. Geophys. Res.*, **103**, 5493–5512
- Fukumori, I., Raghunath, R., Fu, L.-L. and Chao, Y. 1999 Assimilation of TOPEX/POSEIDON altimeter data into a global ocean circulation model: How good are the results. *J. Geophys. Res.*, **104**, 25647–25665
- Gaspari, G. and Cohn, S. E. 1999 Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.*, **125**, 723–757
- Hamill, T. M., Whitaker, J. S. and Snyder, C. 2001 Distance-dependent filtering of background error covariance estimates in an Ensemble Kalman Filter. *Mon. Wea. Rev.*, **129**, 2776–2790
- Houtekamer, P. and Mitchell, H. 1998 Data assimilation using an Ensemble Kalman Filter technique. *Mon. Wea. Rev.*, **126**, 796–811
- Houtekamer, P. and Mitchell, H. 2001 A sequential Ensemble Kalman Filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137
- 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
- Jin, F.-F. 1997 An equatorial ocean recharge paradigm for ENSO. Part I: Conceptual model. *J. Atmos. Sci.*, **54**, 811–829
- Jin, F.-F., An, S.-I., Timmerman, A. and Zhao, J. 2003 Strong El Niño events and nonlinear dynamical heating. *Geophys. Res. Lett.*, **30**, doi:10.1029/2002GL016356
- Kalnay, E. 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. and Rienecker, M. M. 2003 Assimilation of temperature into an isopycnal ocean general circulation model using a parallel ensemble Kalman filter. *J. Mar. Sys.*, **40-41**, 363–380
- Latif, M., Anderson, D., Barnett, T., Cane, M., Kleeman, R., Leetmaa, A., O'Brien, J., Rosati, A. and Schneider, E. 1998 A review of the predictability and prediction of ENSO. *J. Geophys. Res.*, **103**, 14375–14393
- van Leeuwen, P. J. 1999 Comment on ‘ Data assimilation using an Ensemble Kalman Filter technique’. *Mon. Wea. Rev.*, **127**, 1374–1379
- Lorenc, A. C. 2003 The potential of the ensemble Kalman filter for NWP - a comparison with 4D-Var. *Q. J. R. Meteorol. Soc.*, **129**, 3183–3203
- Marsland, S., 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 Modelling*, **5**, 91–127
- van Oldenborgh, G. J., Balmaseda, M. A., Ferranti, L., Stockdale, T. N. and Anderson, D. L. T. 2005 Did the ECMWF seasonal forecast model outperform a statistical model over the last 15 years? *to appear in J. Clim.*
- Pedlosky, J. 1996 Ocean circulation theory, Springer-Verlag, Berlin.
- Robert, C. and Alves, O. 2004 Tropical Pacific ocean model error covariances from Monte-Carlo simulations. *to appear as a BMRC research report*
- Schlag, M. G. and Chelton, D. B. 1992 Frequency domain diagnostics for linear smoothers. *J. Amer. Stat. Assoc.*, **87**, 1070–1081
- 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

- Suarez, M. J. and Schopf, P. S. 1988 A delayed action oscillator for ENSO. *J. Atmos. Sci.*,**45**, 3283–3287
- Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M. and Whitaker, J. S. 2003 Ensemble square root filters. *Mon. Wea. Rev.*,**131**, 1485–1490
- Vossepoel, F. C. and Behringer, D. W. 2000 Impact of sea level assimilation on salinity variability in the western equatorial Pacific. *J. Phys. Oceanogr.*,**30**, 1706–1721
- Wolff, J.-O., Maier-Reimer, E. and Legutke, S. 1997 The Hamburg Ocean Primitive Equation Model HOPE. Technical Report 13, German Climate Computer Center (DKRZ), Hamburg, Germany