Climate Dynamics (2001) 17: 687–700

S. Masina · N. Pinardi · A. Navarra

A global ocean temperature and altimeter data assimilation system for studies of climate variability

Received: 28 May 2000 / Accepted: 6 November 2000

Abstract An ocean data assimilation (ODA) system which can assimilate both temperature and altimeter observations has been applied to the global ocean and tested between January 1993–October 1996. A statistical method has been used to convert sea surface height (SSH) anomalies observations from TOPEX/POSEIDON into synthetic temperature profiles. The innovative aspect of this method is the introduction of time dependency in the correlations used to transform the altimeter observations into temperature corrections. The assimilation system is based on a univariate variational optimal interpolation scheme applied to assimilate both in situ and synthetic temperature profiles. In addition, a longer global analysis for the upper-ocean temperature starting from January 1979 and ending November 1997, has been produced to examine the skill of sea temperature assimilation with a rather simple and practical method. The temperature analysis shows encouraging improvement over a corresponding ocean simulation when compared to independent (not assimilated) temperature data both at seasonal and interannual time scales. However, the univariate data assimilation of hydrographic data does not result in an improvement of the velocity field. In fact the assimilation of sparse in situ data can introduce unrealistic spatial variability in the temperature field which affects the velocity field in a negative way. This deficiency is partially overcome when we also assimilate altimeter observations since the coverage is complete and uniform for this data. In particular, our study shows that temperature corrections due to the altimeter signal have a positive impact on the current system in the tropical Pacific.

S. Masina (✉) · A. Navarra
Istituto Nazionale di Geofisica,
Rome, Italy
E-mail: s.masina@isao.bo.cnr.it

N. Pinardi
Corso di Laurea in Scienze Ambientali,
University of Bologna, Ravenna, Italy

1 Introduction

The ocean is not observed frequently enough in space or time to allow a direct and reasonably accurate description of the large-scale oceanic state and its variability. This limitation can be overcome by applying a data assimilation system which makes use of a numerical model to dynamically interpolate information in space and time. Substantial developments of new ocean data assimilation methods have been going on since the mid-1980s, but the newer and more complex data assimilation techniques, such as Kalman-Bucy filtering (Bennett and Budgell 1987) and the adjoint technique (Le Dimet and Talagrand 1986) still appear to be impractical for application to a high-resolution global oceanic general circulation model. Another open issue in the field of data assimilation concerns the method with which different kinds of observation can be assimilated together in an efficient and appropriate way. In particular, Ji et al. (2000) and Carton et al. (2000b) showed the importance of assimilating satellite altimeter data combined with other hydrographic measurements. Ji et al. (2000) examined the impact of satellite altimeter assimilation on ENSO prediction and their analysis is limited to the tropical Pacific. Carton et al. (2000a, b) produced retrospective global analysis of the upper ocean at a rather coarse horizontal resolution mainly for diagnostic studies of climate variability.

We explore the advantages of using a univariate optimal interpolation method associated with a relatively high-resolution primitive equation global ocean circulation model. We assimilate both temperature profiles and satellite altimeter data, i.e. the most readily available global ocean measurements at the moment. The simplicity of the method and its global application gives this assimilation approach a dual purpose: first, for dynamical and climatological studies of the upper ocean, and, second, for initialization of coupled atmospheric and oceanic general circulation models (GCMs) for ENSO prediction. In particular, the assimilation of large

oceanic data sets has been shown to be important for the extension of the E1 Niño predictability limits (Miyakoda et al. 1997; Ji et al. 2000), and for climate studies at different time scales (Ji and Smith 1995; Ji et al. 1995).

This study is a partial reanalysis of the work done by Rosati et al. (1995) which shows a ten year period (from 1979 to 1988) global ocean analysis for the same model and the same temperature data assimilation scheme. We assimilate altimeter observations which are transformed into "synthetic" temperature profiles by applying the statistical method of Mellor and Ezer (1991). This work has been carried out for almost two decades (1979–1997) and has been validated by comparison with independent data sets, with particular emphasis on the tropics and the Gulf Stream region.

Details of the data sets used for assimilation and the quality control procedure will be given in Sect. 2, where we also describe the ocean model and the assimilation technique. In Sect. 3 the impact of assimilation of temperature profiles is evaluated through comparison of the ocean analyses with the model simulation and with the Levitus climatological fields. An estimate of the analysis error is made by comparison with independent data sets, together with a description of the low-frequency variability of the results. The impact of assimilation of altimeter observations is assessed in Sect. 4. Section 5 provides a summary and our conclusions.

2 The ocean data assimilation system

Part of the temperature profiles used are from XBT and CTD measurements contained in the World Ocean Data Bank-94 (Levitus and Boyer 1994). These temperature profiles covered the period from 1979 to 1991. Because of the data scarcity in the World Ocean Data Bank-94 after 1991, we switched to the GSDC Data set (see: http://www.ifremer.fr:582/sismer/program/gsdc) from the Institut Francais de Recherche pour l'Exploitation de la Mer (IFREMER) for the 1992–1997 period. In addition, the Reynolds weekly sea surface temperature (SST) analyses (Reynolds and Smith 1994) are used. The Reynolds SST fields combine in situ and satellite data analyzed using optimum interpolation on a 1° latitute longitude spatial grid.

For altimeter observations we use the TOPEX/POSEIDON product provided by the Collecte Localisation Satellites (CLS) Space Oceanography Division. In particular, the data set used in this study contains sea level anomalies relative to a 3-year mean (January 1993 to December 1995). The altimeter product is provided every 10 days on a regular grid of $1 \times 1°$ using an improved space/time objective analysis method which corrects along-track long wavelength errors (Le Traon et al. 1998).

Before being assimilated, all temperature profiles are subject to a preassimilation quality control procedure. The latter consists of preliminary quality control checks based upon a climatological and vertical stability check. The first check ensures that depths increase monotonically and the second that the temperature profile does not contain unrealistic inversions. Furthermore, observations which deviate from the climatological values by more than three times the estimate of the variability are eliminated.

The numerical model is a nearly version of the Modular Ocean Model (Cox 1984; Rosati and Miyakoda 1988) implemented for the global ocean from 78°S to 65°N. The horizontal resolution is $1 \times 1°$ everywhere except in the tropical area where the north-south resolution is increased to 1/3 of a degree. In order to show how the data assimilation affects the ocean analysis, two equivalent experiments have been run starting from the same initial conditions and applying the same atmospheric forcing. In one case the model has been run in the assimilation mode (ASSIM1 in Table 1), while in the other case the model has been simply forced without any data assimilation (SIM in Table 1). The horizontal resolution used for the two experiments is the same. The vertical resolution is higher in the forced run where the model has 18 levels unevenly spaced down to 3000 m with the first 12 levels confined in the first 250 m. This model has been used in coupled ocean-atmosphere integrations (Fischer and Navarra 2000). For the assimilation experiment 15 levels are vertically distributed over 3000 m, and the first 11 levels are in the first 250 m. The vertical level distribution was determined by the compromise between computational costs and the need to resolve the thermocline in the tropics correctly. The extra computational cost of the assimilation experiments with respect to the simulation experiments forced a slightly lower vertical resolution for the ODA experiment. The vertical diffusion and horizontal viscosity are parametrized with the Mellor-Yamada (1982) turbulence closure scheme and Smagorinsky's (1993) nonlinear viscosity, respectively. The time step is 40 min and is the same for tracers, barotropic and baroclinic velocity. The use of the same time step has been introduced to overcome the problems found by Derber and Rosati (1989) in reproducing a realistic equatorial undercurrent when different time steps are used. Unfortunately, this increases the computational cost.

The atmospheric forcing is provided by the ECMWF reanalyses fields between January 1979 to December 1993, and operational analysis fields from January 1994 to November 1997. The surface fields used are the air temperature at 2 m, the dew-point temperature at the same level, the mean sea level pressure and the zonal and meridional component of the wind at 10 m. With these variables momentum and heat fluxes are computed twice daily interactively with the model velocity fields and sea surface temperature using the method implemented by Rosati and Miyakoda (1998). The surface salinity boundary condition is relaxed to the climatological monthly mean values due to the lack of suitable hydrological data.

The assimilation scheme consists of the univariate variational optimal interpolation scheme developed by Derber and Rosati (1989), which we apply at each time step (every 40 min). The data are inserted within a time window of 30 days around the date they have been collected since we found no significant difference for the monthly mean values when we tried shorter time windows. The horizontal covariance is the same used by Behringer et al. (1998) with the only difference that the first-guess error variance has been reduced by a factor of 3 ($a_h = 0.003$ in our study) in order to consider the fact that our time step has been reduced to 1/3 of the value used by Derber and Rosati (1989). This change is similar to the overall reduction in the magnitude of the estimated first-guess error introduced by Behringer et al. (1998). For each data point, even if the first-guess error variance is small at each time step, the correction to the guess field is distributed over many time steps and therefore the effective first-guess error variance becomes similar to the observational errors. We assimilate the temperature observations only down to level 11 of the model (250 m) since the next

Table 1 List of experiments

Experiment	Time	Assimilated data	Assimilation scheme
SIM	Jan 1979–Nov 1997	Control run (no data assimilated)	No assimilation scheme
ASSIM1	Jan 1979–Nov 1997	In situ temperature	Derber and Rosati (1989)
ASSIM2	Jan 1993–Oct 1996	TOPEX/POSEIDON altimeter and in situ temperature	Mellor and Ezer (1991) and Derber and Rosati (1989)

model level is located at 467 m where the total number of temperature observations is drastically reduced. Further details of the assimilation scheme and the choice of parameters used in our study are given in Pinardi et al. (1997).

A main drawback of this assimilation scheme in its present formulation is the fact that it is univariate and therefore it does not allow correction of other model variables, for example salinity, in order to maintain the correct water mass properties. This assumption will affect the quality of our results as we will see later. The important characteristic of this assimilation scheme is that it is based upon advanced continuous assimilation algorithms, as explained by Lorenc (1986). This prevents the model adjustment processes becoming very large between intermittent data insertion steps so that monthly mean fields can be generated which contain minimal model adjustment perturbations.

3 Impact of hydrographic data assimilation

3.1 Climatological ocean analyses and statistics

We started the spin up of the model in January 1979 using the Levitus winter climatological fields as the initial conditions for temperature and salinity and zero velocities. We discarded the first two years of the analysis (1979 and 1980) to avoid any possible effect due to unbalanced initial conditions.

We first analyze the assimilation results in terms of reconstructed climatology from the 16 year data set (1981–1996) (ASSIM1 in Table 1). In the top panel of Fig. 1 we show the assimilated climatological global temperature, hereafter called the analysis, at 100 m depth. The difference between the climatological analysis and the annual mean temperature produced by Levitus and Boyer (1994) through an objective analysis technique is shown in the middle panel of Fig. 1. In an equivalent way, we compare the simulated climatology to the Levitus climatology. The comparison of the three climatologies confirms the robustness of the assimilation scheme and the improvements introduced with respect to the simulation results. The magnitude of the differences between analysis and Levitus is less than 1 °C over most of the oceans. At this depth the largest differences in the global ocean temperature are in the Gulf Stream region and in the vicinity of the North Equatorial Countercurrent trough near 10°N. In the equatorial region the difference is relatively small. Since the

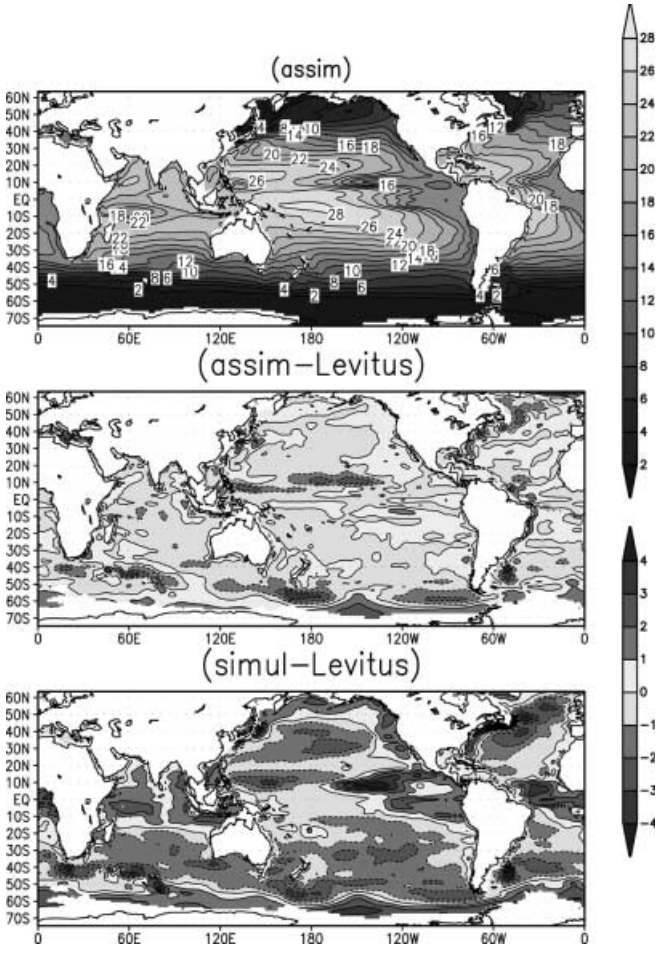

Fig. 1 Horizontal map of time mean (1981–1996) temperature at the depth of 100 m for the assimilation experiment (*top panel*). The differences between the model temperature climatology and the Levitus climatology are shown in the *middle panel* for the assimilation experiment and in the *bottom panel* for the simulation experiment

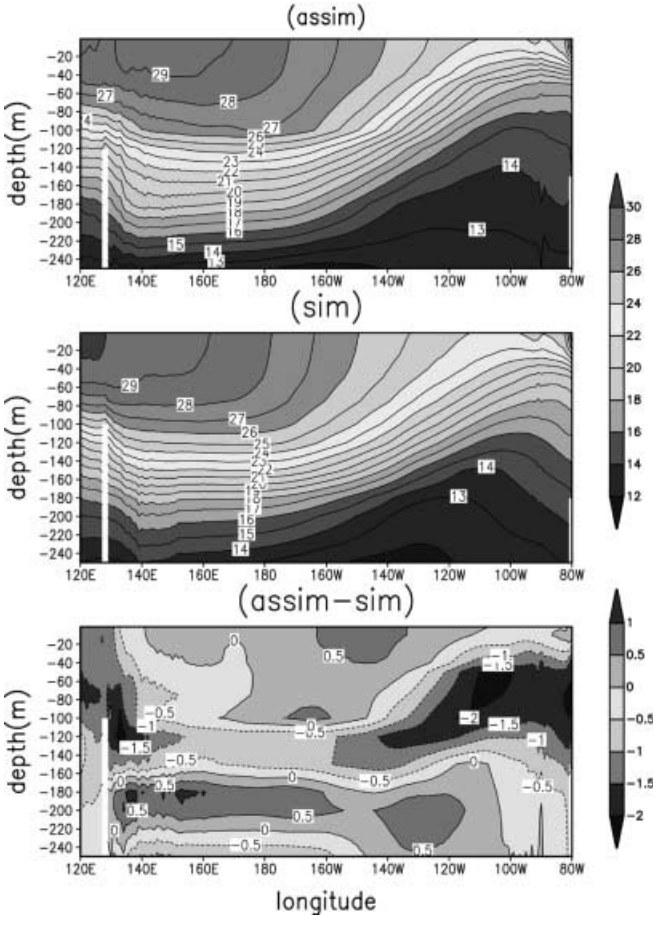

Fig. 2 Vertical section of the time mean (1981–1996) temperature along the equator in the Pacific for the assimilation (*top panel*) and the simulation (*middle panel*) experiments. The difference between the two is shown in the *bottom panel*

simulation and the analysis used the same wind forcing field, the difference in the upper ocean temperature represents the effect of data assimilation. We argue that this comparison suggests that the curl of the wind stress is too weak and therefore does not induce enough upwelling in the NECC region.

In Fig. 2 we show a temperature section at the equator across the Pacific for the assimilation experiment (top panel), the simulation run (middle panel) and the difference between the two (bottom panel). The noticeable difference between the two sections is the magnitude of the warm pool temperature signal and the structure of the subsurface thermocline, especially in the eastern Pacific. The west Pacific warm pool in the assimilated run shows a temperature decrease with respect to the simulation of about 1–2 °C in the uppermost 150 m of the water column. The thermocline structure is flat and deep in the eastern Pacific for the simulation while the assimilation shows a much sharper thermocline rise in the same region. The correspondent temperature difference between the two experiments maintains an approximate value of 1.5 °C along the region of maximum temperature gradient. The reason for the diffusive behaviour of the equatorial thermocline remains an unresolved critical question which is partly due to the difficulty in parametrizing vertical diffusion in the numerical model and the forcing inaccuracies.

A more quantitative estimate of the efficiency of the ocean assimilation method in terms of consistency with Levitus climatological fields is given by the calculation of the area-weighted correlation coefficients between the two climatologies.

In Fig. 3 we show horizontal maps of time mean (1981–1996) temperature at a depth of 75 m in the tropical Pacific for the assimilation experiment, the simulation experiment and the Levitus climatology (top panel). In the bottom panel the first two figures show the scatter plots of the model temperature data (assimilation and simulation) versus the Levitus data taken in the same area shown in the top panel. The evident spread of the scatter plot of the simulated temperatures versus the Levitus temperatures (bottom–middle panel) about an ideal slope of 45° represents the deviation of the model climatology with respect to the observed climatology. In particular, the scatter plot indicates that the model has a tendency to give temperatures warmer than the observations. From the horizontal maps we can note that this

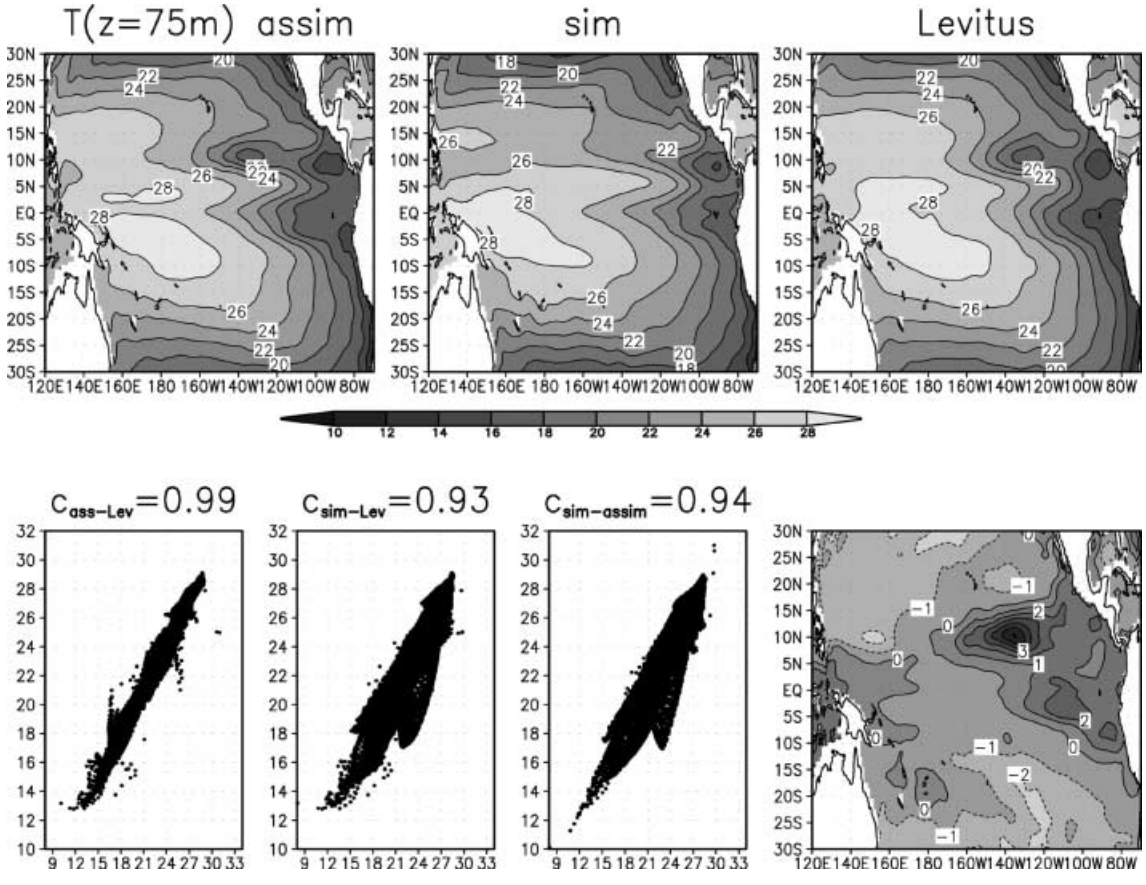

Fig. 3 Horizontal maps of time mean (1981–1996) temperature at the depth of 75 m in the tropical Pacific for the assimilation experiment, the simulation experiment and the Levitus climatology (*top panel*). In the *bottom panel* the *first three figures* show the scatter plots of the model temperature data (assimilation and simulation) and the Levitus data taken in the same area shown in the *top panel*. The correspondent correlation coefficients are written *above each plot*. The *far right figure* in the *bottom panel* shows the difference between simulation and assimilation experiments

happens along 10°N in the eastern Pacific. The scatter plot of the analysis versus the observations is usually much closer to the idealized slope as it is also shown by a correlation coefficient of 0.99 which is significantly larger than the correlation between the simulated and the observed temperatures (0.93). The correlation values are however possibly overestimated due to the fact that we correlate climatological fields.

In Fig. 4 we show horizontal maps of climatological temperature at a depth of 100 m in the Gulf Stream region from 34° to 41°N and from 80° to 60°W for the assimilation experiment, the simulation experiment and the Levitus climatology (top panel from left to right, respectively). This comparison shows that the simulation is not able to reproduce a correct Gulf Stream temperature front probably due to the coarse resolution. Not only is the amplitude of the temperature gradient much weaker than observed, but also the simulated front has a slope opposite to observed. In particular, the Gulf Stream temperature gradient simulated by the model is only 3 °C over a latitudinal extension of 5°, and is directed southward, while the corresponding observed front is twice as strong and directed northward. The analysis shows that the insertion of the observed data into the model is able to drastically sharpen the gradient of the Gulf Stream temperature front and switch its slope in the right direction east of 69°W. To the west the slope of the front is still wrong probably because of lack of data near the coasts and dynamical adjustment due to viscous processes near the coast. The bottom panels of Fig. 4 show the scatter plots of the model temperature data versus levitus data and the scatter plot of simulated versus assimilated temperatures. The scatter plots are indicative of the fact that the simulated temperatures are significantly warmer than the analysis and the observed climatology. The correspondent correlation coefficients confirm that the analysis and the Levitus climatology are highly correlated (0.94) while the simulated temperatures have a low correlation both with the observations (0.42) and the analysis (0.48).

3.2 Seasonal and interannual variability of the global upper ocean

Since one of the most likely applications of ocean analyses is the initialization of ENSO forecasts, an accurate representation of both the phase and amplitude of the propagating heat content anomalies at seasonal and interannual time scales is crucial. Therefore, we focus in this section on the seasonal and interannual upper-ocean thermal variability in the tropical oceans. The scarcity of subsurface data prevents the analysis from resolving higher frequency

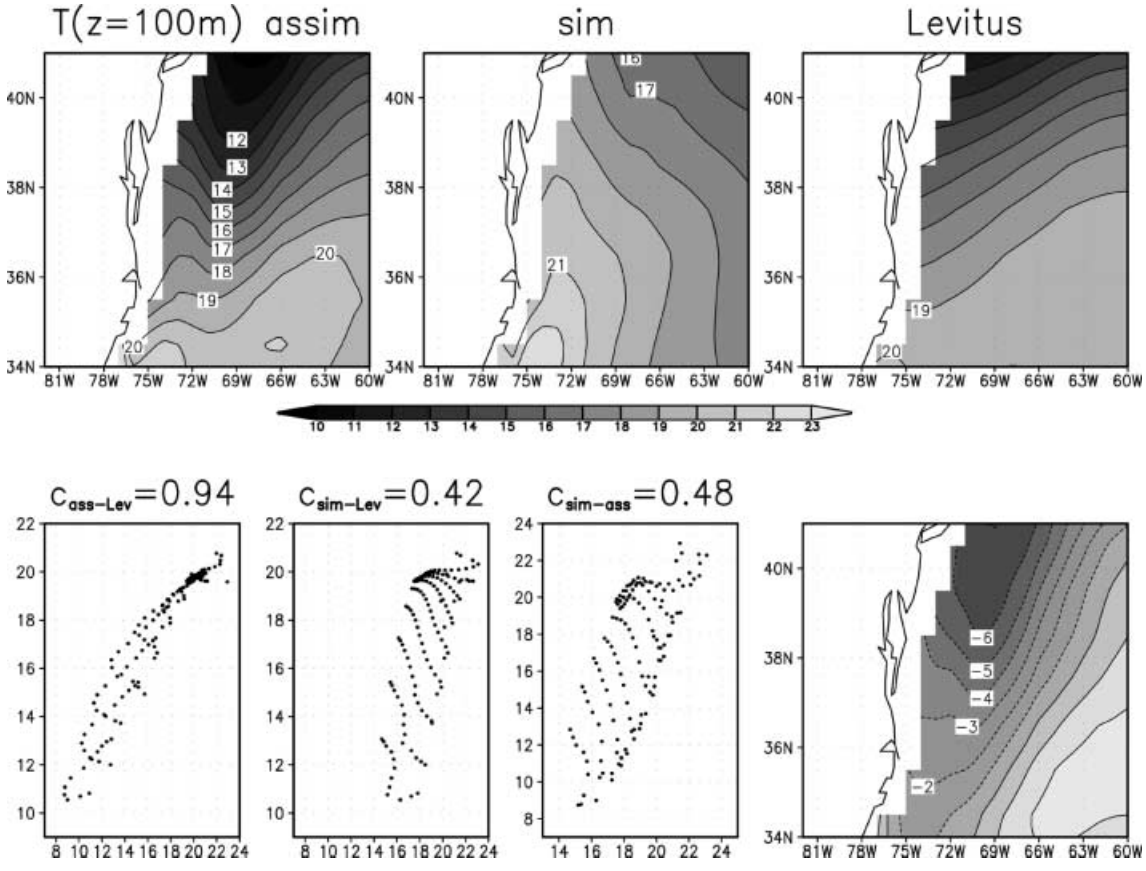

Fig. 4 As in Fig. 3 but for temperature at 100 m depth in the Gulf Stream region

variability and the limited period of the analysis does not allow resolution of decadal processes.

The comparison between the interannual variability of the monthly mean heat content anomalies averaged over a latitudinal band between 10°S and 10°N is shown in Fig. 5 for the analysis (top panel) and the simulation (bottom panel). In both diagrams the respective seasonal cycle has been removed. The comparison between the simulation and the analysis shows good correlation in the timing of the interannual variability but the amplitude of the heat content anomalies are significantly different. In particular, the western Atlantic Ocean shows differences in terms of anomaly amplitudes. In the analysis the heat content anomalies in the western Atlantic reveal a complicated pattern of alternating positive and negative anomalies which are significantly stronger than those in the eastern Atlantic, while in the simulation the amplitude of the anomalies is the same at all longitudes in the whole Atlantic. In comparison with the model simulation, the analysis shows also a stronger correlation between the ENSO signal and heat content anomalies in the Indian Ocean.

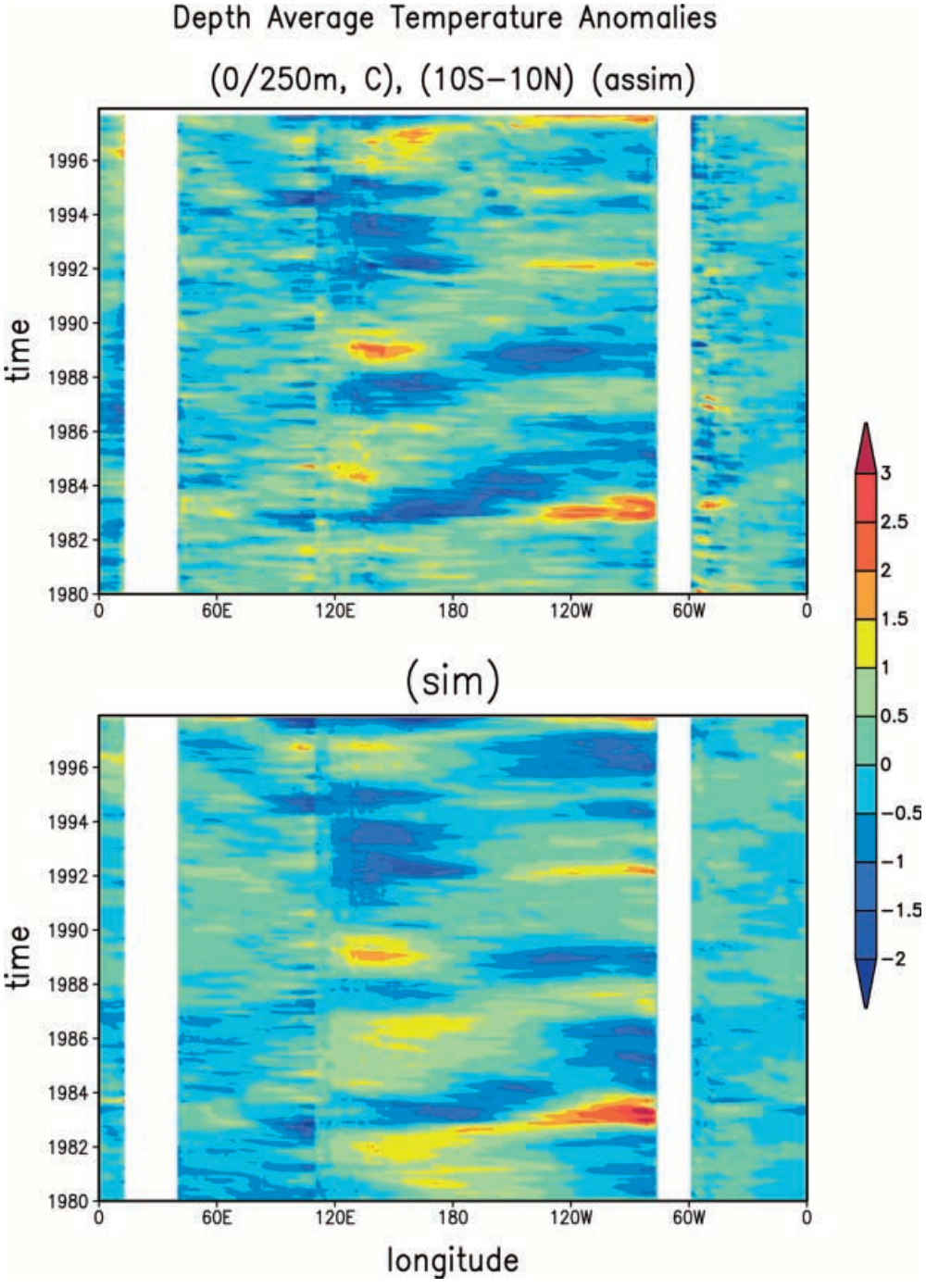

Fig. 5 Hovmöller diagrams of the monthly mean temperature anomalies averaged over the uppermost 250 m and between 10°S and 10°N. Assimilation experiment (*top panel*) and simulation experiment (*bottom panel*). The simulation and assimilation have their different seasonal cycle removed

The heat content anomalies indicate the change of ENSO characteristics from the first decade of the analysis to the second. From 1980 to 1990 two "regular" ENSO events with a period of four years are evident both in the analysis and in the simulation. In the 1990s this behaviour disappears and the negative anomaly that develops in the western Pacific from 1990 to 1994 does not propagate to the eastern Pacific where weak anomalies are present during this period. This behaviour is probably indicative of the particular character of El Niño during the years 1990–94 when it was found that the surface anomalies were strongly coupled to the mixed layer across the equatorial Pacific and a persistent warm SST anomaly in the central Pacific was already present and crucial for the initiation of the warm events (Goddard and Graham 1997). In 1996 the simulation shows a large and negative anomaly in the eastern Pacific that is weaker and less zonally extended in the analysis. The latter reproduces better the observed heat anomalies as shown in Fig. 6. The same intercomparison of Fig. 6 shows that the simulated heat content anomaly during the 1997 El Niño event is significantly weaker than the corresponding anomalies in the analysis and in the observations.

Since temperature observations from the TOGA TAO arrays are not incorporated into the assimilation and are available for most of our analysis period, they can be used as an independent data set for verification. In Fig. 7 we show time series of temperature in the upper 250 m of the ocean at the equator at 140°W from the TAO data (top panel), the assimilation experiment (middle panel) and the simulation experiment (bottom panel). The TAO profiles as well as the model solutions are monthly means. The intercomparison of the time series reveals that the insertion of the temperature data is able to sharpen the thermocline consistently with the TOGA TAO observations and to reproduce the right amplitude of the seasonal and interannual variability which the simulation generally underestimates.

The temperature assimilation also changes the velocity field. Figure 8 is the same as Fig. 7 but for the zonal velocity field in the uppermost 180 m of the ocean. According to the TAO measurements, the main fluctuations of the Equatorial Undercurrent (EUC) along the equator at 140°W is its annual cycle. In late spring or early summer the core of the EUC strengthens up to 140 cm/s and appears in shallower waters. The EUC spring intensification is followed by a deepening and considerable weakening in the next few months. The comparison of the zonal velocity time series shown in Fig. 8 indicates that the seasonal variability of the zonal velocity is reproduced by both the analysis and the simulation. However, the depth and the amplitude of the simulated EUC are generally improved when the temperature data are included into the model. In particular, the speed of the EUC is much better reproduced by the analysis than by the simulation, which has a tendency to underestimate it.

In Fig. 9 we show a comparison of zonal velocity (top panel) and temperature (bottom panel) at the same TAO location but at a single depth of 80 m. It is clear that the temperature analysis is more highly correlated with the

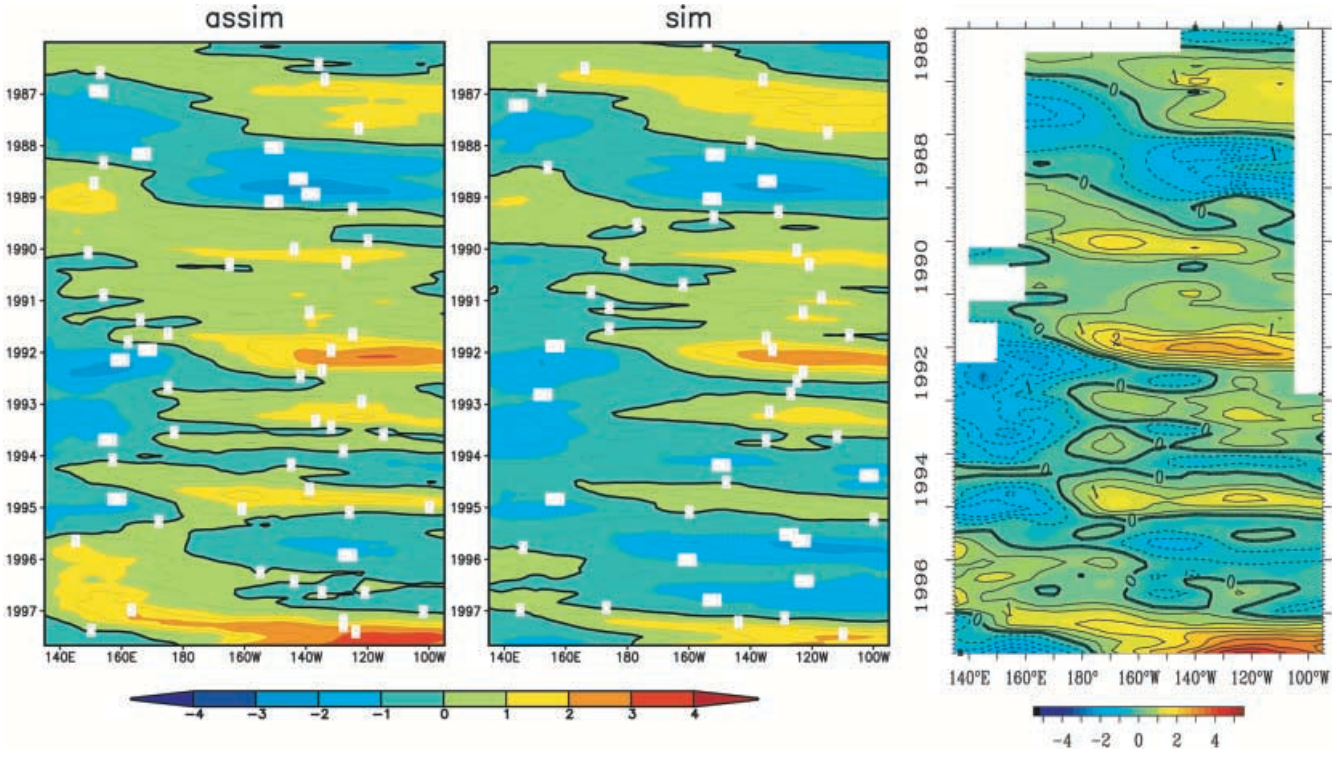

Fig. 6 Hovmöller diagrams of the monthly mean temperature anomalies (seasonal cycle removed) averaged over the uppermost 250 m in the model (first two figures) and 300 m for the TAO observations (for right figure) and between 2°S and 2°N. The TAO data have their own climatology subtracted

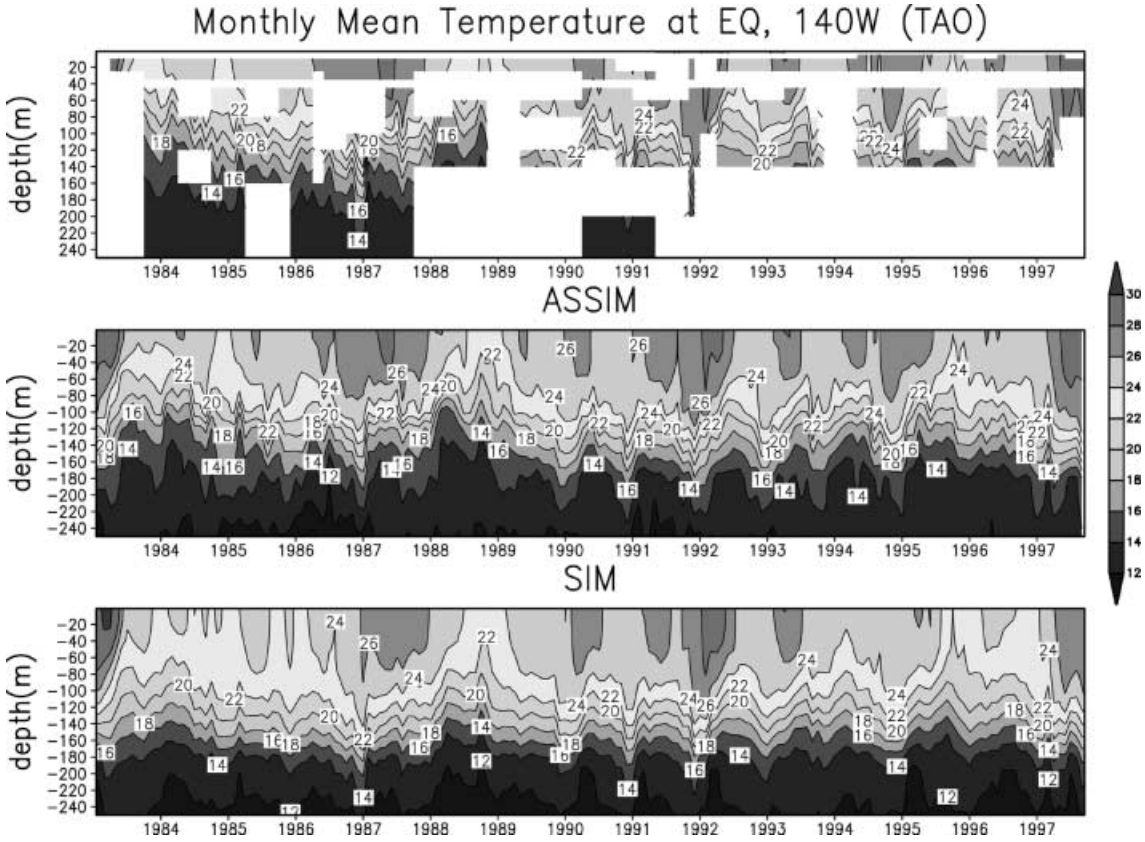

Fig. 7 Depth–time sections of monthly mean temperatures at 140° on the equator for the TOGA-TAO data (*top panel*), the assimilation experiment (*middle panel*) and the simulation experiment (*bottom panel*)

observations than the simulation but the opposite is true for the zonal velocity. From this figure it is also evident that the temperature analysis is able to reproduce larger interannual variability in a manner consistent with the observations, as can be seen in the 1987–88 ENSO event. Similar conclusions apply also to another equatorial TAO location further east in the Pacific (110°W). The comparison between the simulation and assimilation results at a depth of 45 m at this location is shown in Fig. 10. Both Figs. 9 and 10 show that the univariate temperature insertion may cause anomalous adjustment in the EUC. Other parameters in the assimilation can possibly cause this behaviour such as the data time window and the estimated background error covariance values, but in our case we believe that the problem is introduced by the univariate assumption.

4 Impact of altimeter data assimilation

4.1 The "synthetic" temperature profile method

The method implemented by Mellor and Ezer (1991) in the Gulf Stream region is applied here in order to assimilate TOPEX/POSEIDON altimeter data in the numerical model described in Sect. 2 between January 1993 and October 1996. In particular, the SSH observations are transformed into "synthetic" temperature profiles in such a way that they can be assimilated simultaneously with the hydrographic temperature profiles using the assimilation scheme described in Sect. 2. The correlations between SSH and temperature are computed from the model analysis obtained by the assimilation of the hydrographic data only. The most innovative aspect of this study with respect to the previous application and other methods is the introduction of time variability in the statistics in order to take into account the high seasonal and interannual variability of the correlations between altimeter and temperature fields.

We define a temperature anomaly, T', and a sea surface height anomaly, ssh', such as:

$$T'(x,y,z,t) = T(x,y,z,t) - T^{av}(x,y,z,t)$$
$$ssh'(x,y,t) = ssh(x,y,t) - ssh^{av}(x,y,t)$$

The terms T^{av} and ssh^{av} are seasonal means. Both T and ssh are stored three times every month as model outputs. Here both T and ssh have been computed removing a three year mean (1993–1995) derived from the analysis fields. We can then construct a synthetic temperature defined as:

$$T_{syn}(x,y,z,t) = F(x,y,z,t)ssh'_{TP}(x,y,t) + T^{av}(x,y,z,t)$$
$$F(x,y,z,t) = \frac{\overline{T'ssh'}}{\overline{ssh'^2}}$$

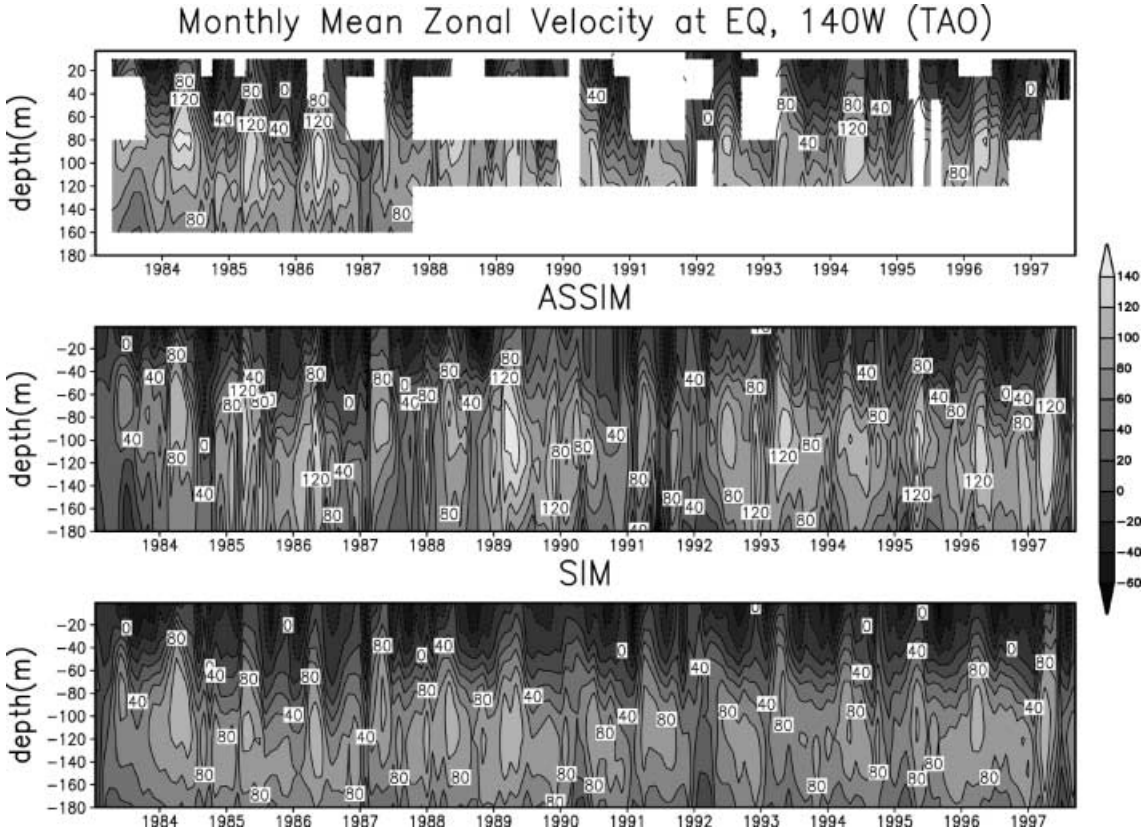

Fig. 8 As in Fig. 7 but for the monthly mean zonal velocities

where $ssh'_{TP}(x, y, t)$ is the TOPEX/POSEIDON sea surface height anomaly with respect to the 1993–1995 mean. The three year average has been subtracted from the raw data in order to subtract the geoid effect. In addition the seasonal mean has been subtracted from the raw data following the same procedure applied to the analysis fields. The overbar indicates the average of the nine model realizations available for each season and F is a correlation factor.

F is calculated seasonally using the model outputs from the assimilation run with only XBT (ASSIMI1 in Table 1). In particular, we used ten day model-averaged T and ssh fields calculated every three months. We stored the result and, in a second step, the TOPEX/POSEIDON altimeter anomalies ssh'_{TP} are "projected" into synthetic temperature profiles using F. These synthetic T profiles will be used in the experiment ASSIM2 of Table 1.

The error for the synthetic T profile is written as (Mellor and Ezer 1991):

$$\overline{T'_{syn}T'_{syn}} = F^2 \frac{(1 - C^2)}{C^2} \overline{ssh'^2_{TP}} + \epsilon_{ssh}$$

where

$$C(x, y, z, t) = \frac{\overline{T'ssh'}}{(\overline{ssh'^2}\ \overline{T'^2})^{\frac{1}{2}}}$$

is the correlation and ϵ_{ssh} is the estimated error variance of the observed ssh signal projected into temperature values.

Both the space patterns and the amplitude of the computed correlations are highly variable in time. In particular at the surface such correlations can drastically change sign from one year to the other. This suggests that methods which extrapolate density information from altimeter observations using statistics fixed in time fail to represent the natural variability of the system accurately. The correlations are in general larger at depth than at the surface indicating that the sea surface height anomalies are more correlated to the thermocline dynamics than to the mixed layer temperature changes.

4.2 Assimilation of synthetic and other in situ observations

Starting from January 1993, synthetic temperature profiles are produced every 10 days from SSH maps and then merged with the temperature observation data set. The combined data set is then assimilated with the same scheme described before for the in situ observations only and a new analysis (ASSIM2 in Table 1) has been produced. Although we do not assimilate altimeter SSH anomaly observations directly, the temperature

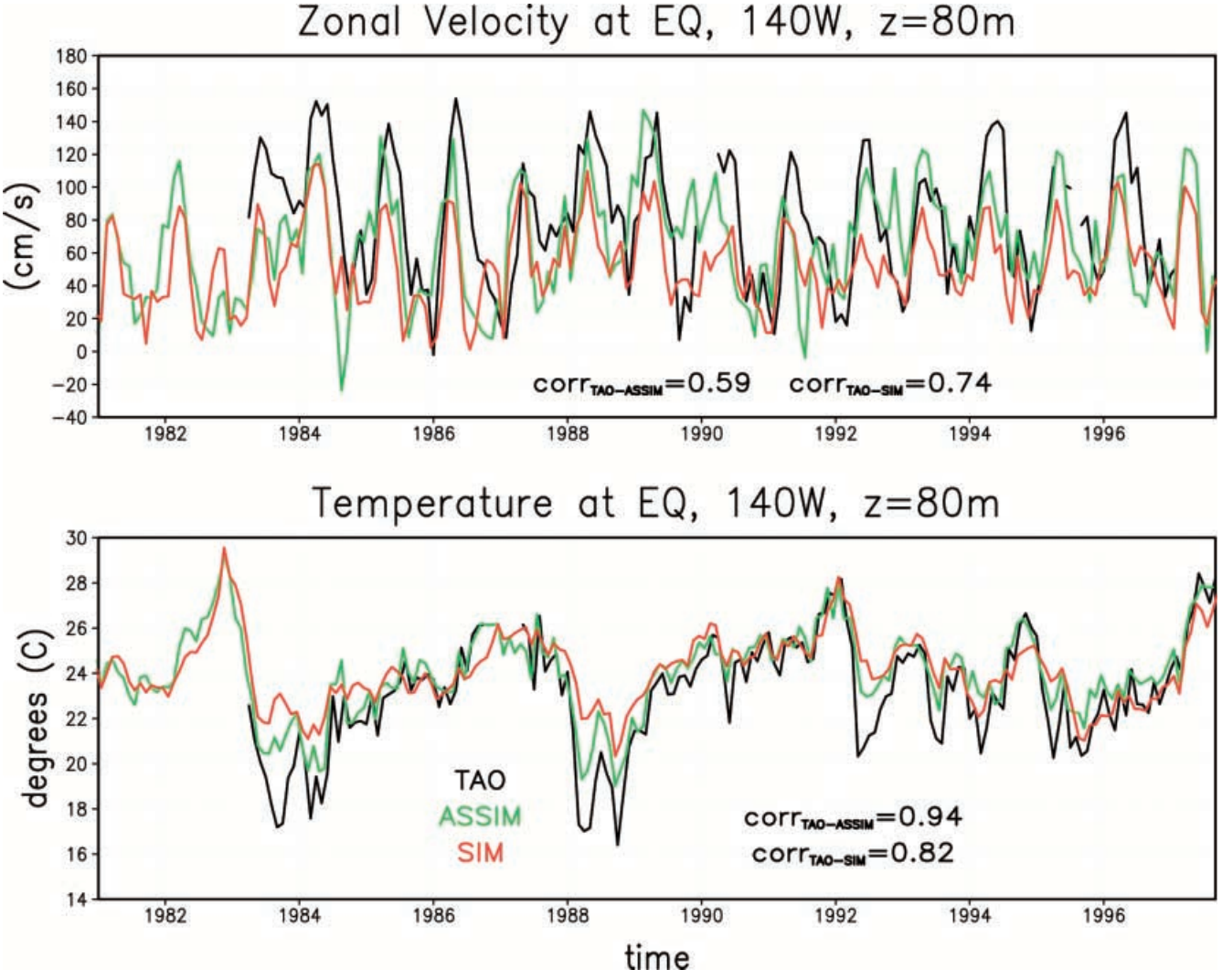

Fig. 9 *Top panel*: comparison of time series of monthly mean zonal velocities at 140°W on the equator and at the depth of 80 m for the TOGA-TAO data (*black curve*), the assimilation experiment (*green curve*) and the simulation experiment (*red curve*). *Bottom panel*: as above but for the monthly mean temperatures

corrections introduced through the statistical method have an impact on the sea surface height. Such an impact is positive in the tropical Pacific and Indian Oceans as is evident in Fig. 11 which shows the root mean square of the sea surface height anomalies differences between the TOPEX/POSEIDON data and the two analyses, respectively. The improvement of the analysis when the altimeter data are used in consistently of the order of 1 cm during the annual cycle both in the Pacific and Indian tropical oceans with peaks in the central Pacific. In contrast, in the Atlantic Ocean the impact of the altimeter assimilation is almost always negative, even if the differences between the two analyses are less than 0.5 cm. It is possible that the shallower thermocline structure of the tropical Atlantic cannot be represented well by the correlation function structure deduced from the model. Another hypothesis is that here the salinity changes are also very important and univariate data assimilation will not preserve water mass properties as well as in the Pacific. Independent subsurface temperature data should be perhaps used to compute correlation functions to be used to compute synthetic temperature profiles.

Finally, we show an intercomparison of the model simulation, the two analyses performed without and with altimeter data assimilation and an independent data set taken from the WOCE experiment. In Fig. 12 we show an example taken at a WOCE latitudinal section in the central Pacific and averaged in time from the end of June to the end of August 1993. In general, the assimilation of the hydrographic data is able to sharpen the thermocline in agreement with observations. The assimilation of XBT and CTD temperature data also introduces a spatial variability (second panel from the top) that is, of course, completely absent in the smooth model simulation (upper panel). Part of this variability is realistic, however some of the temperature structures seem to be artificially introduced by the inadequacy of the optimal interpola-

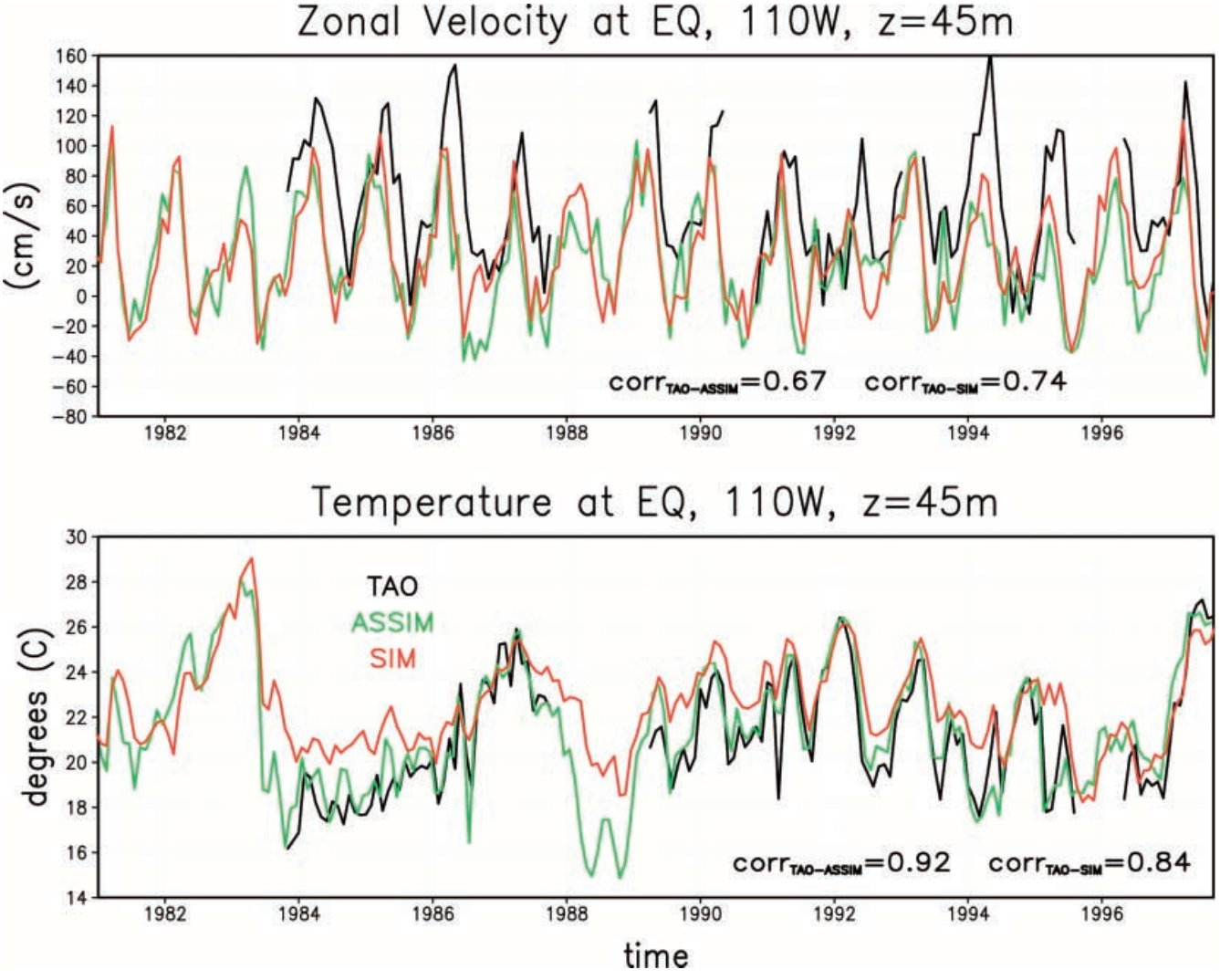

Fig. 10 As in Fig. 9 but at 110°W on the equator and at the depth of 45 m

tion in giving a smooth analysis when the data coverage is too poor. This deficiency is partially overcome when we assimilate the altimeter observation since for this data set the coverage is complete and uniform. This is why the analysis with the assimilation of both hydrographic and altimeter data (second panel from the bottom) seems to be smoother and compares better with the independent temperature section (bottom panel). In amplitude the temperature corrections introduced by the altimeter assimilation are quite small and mainly confined at the thermocline depth. However, even these small temperature corrections have a significant impact on the current system in the equatorial region. In particular, the smoother latitudinal temperature gradients produced by the assimilation of TOPEX/POSEIDON, has a beneficial effect also in the representation of the zonal velocity structure. In Fig. 13 we show the same intercomparison of Fig. 12 but for the zonal velocity. The meridional temperature gradients given by the assimilation of sparse hydrographic data produce zonal velocities (second panel from top) which are too strong (maximum EUC speed of 80 cm/s) with respect to independent data (bottom panel, maximum EUC speed of 50 cm/s) and to simulated data (top panel). The simulated zonal velocities however, are in general too weak and contain only part of the realistic spatial variability. The temperature changes introduced by the altimeter data induce changes in the zonal velocity (second panel from the bottom) in good agreement with independent data taken from the WOCE section. The speeds of the currents are reduced with respect to the speeds given by the analysis with only hydrographic data and some realistic features are also introduced. For example, note the realistic surface westward jet located between the equator and 3°S which has been clearly introduced by temperature changes induced by the altimeter signal. Similar conclusions apply to several intercomparisons made at different WOCE sections in the tropical Pacific and are confirmed by an intercomparison of time series of temperature and zonal velocity analyses with TAO observations. In Fig. 14 we

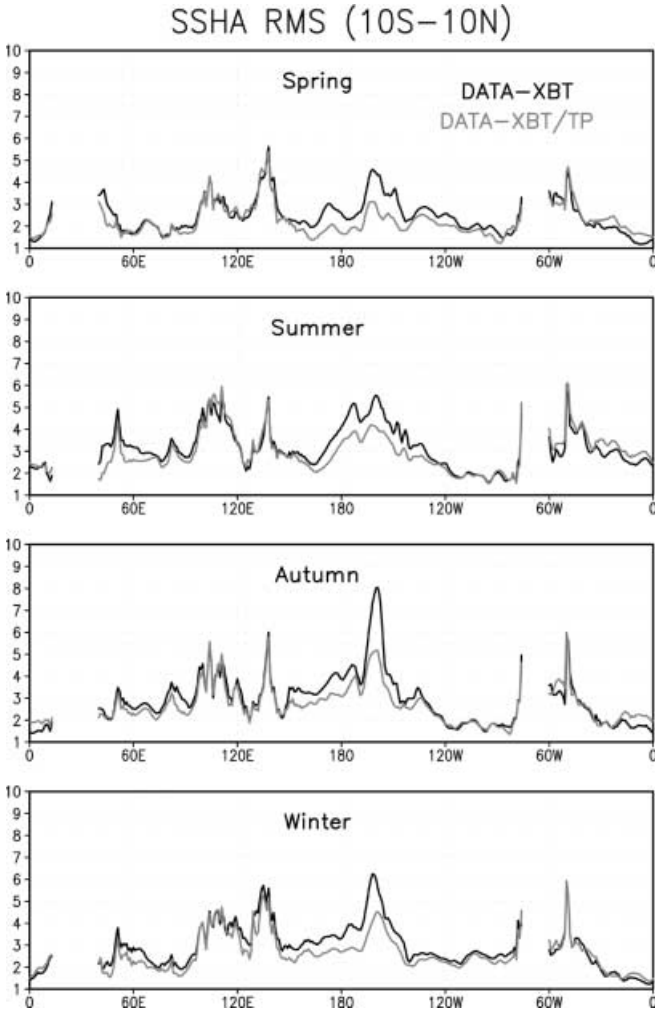

Fig. 11 A comparison between the area-averaged (10°S–10°N) seasonal rms error of sea surface height anomaly calculated by the assimilation of hydrographic data only (*black line*) and the assimilation of both hydrographic and altimeter data (*grey line*): *top* spring; *second* summer; *third* autumn and *bottom* winter

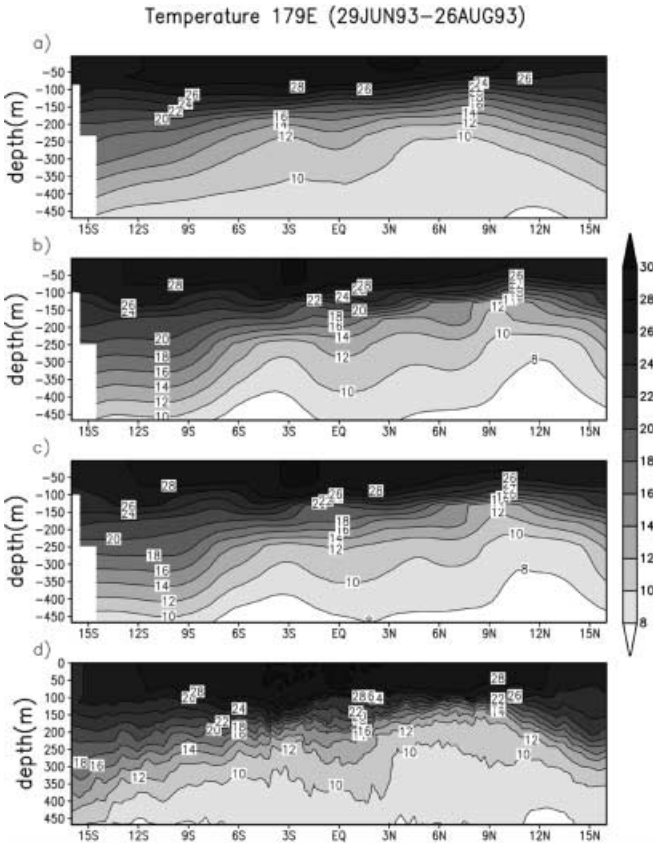

Fig. 12a–d A comparison of a depth–latitude section of temperature taken at 179°E and time-averaged between 29 June 1993 and 26 August 1993 for **a** model simulation, **b** assimilation with hydrographic data only, **c** assimilation with both hydrographic and altimeter data and **d** WOCE data

show an example of intercomparison at one TAO location (140°W) on the equator at a fixed depth of 200 m. The temperature changes introduced by the assimilation of altimeter observations are small with respect to the corrections with respect to the simulated temperatures. However, these temperature differences have a significant positive impact on the zonal velocity, which has reduced speed in the analysis obtained with the assimilation of altimeter observations with respect to the analysis with only hydrographic data assimilation. The reduced velocities are in agreement with TAO observations.

Off the tropics the coarser resolution of the model does not allow reproduction of the spatial temperature variability, though the improvement with respect to the forced simulation is still evident (Fig. 15). In particular, at middle and high latitudes our ocean analyses underestimate the speed of zonal currents in a significant way (not shown).

5 Summary and conclusions

In this study a global ocean data assimilation system, which can assimilate both in situ temperature and satellite altimeter observations, has been formulated and validated. We introduced time dependency in the statistical method used to assimilate altimeter data. This method allowed us to overcome a problem common to previous attempts to assimilate sea surface height observations, i.e. the reduced accuracy in temperature variability (Ji et al. 2000). Ji et al. (2000) found that even if the assimilation of TOPEX/POSEIDON sea level data improved the sea level variability, it also increased the temperature error during some periods of the analysis. They attributed this problem to the fact that their assimilation scheme is univariate and corrects the temperature field alone. We speculate that this weakness can also due to the fact that Ji et al. (2000) used a time-independent linear operator to transform corrections of sea surface height into temperature corrections.

Although model and forcing errors inevitably affect the quality of the analysis, the assimilation of observed temperature data and altimeter observations into the

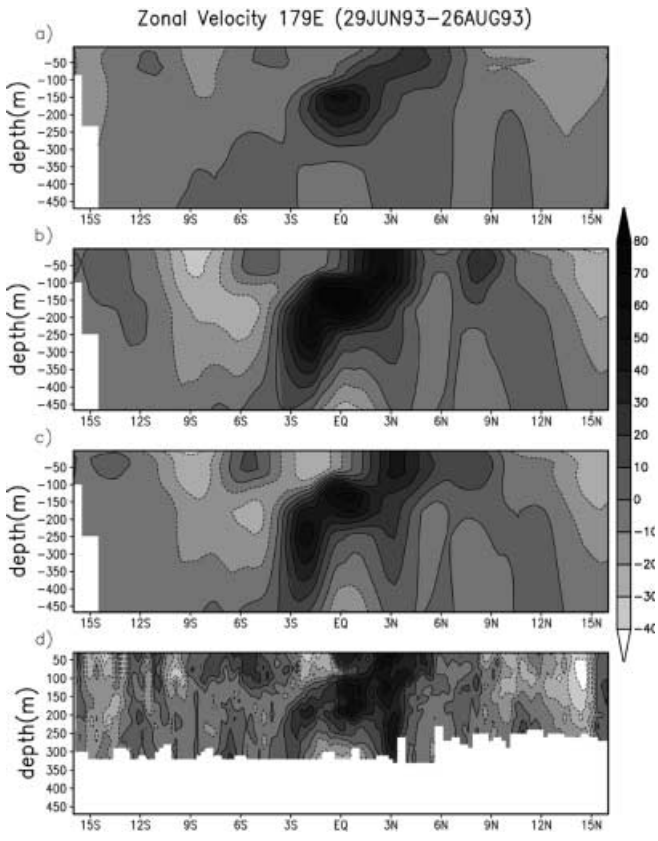

Fig. 13a–d Same as in Fig. 12 but for zonal velocity

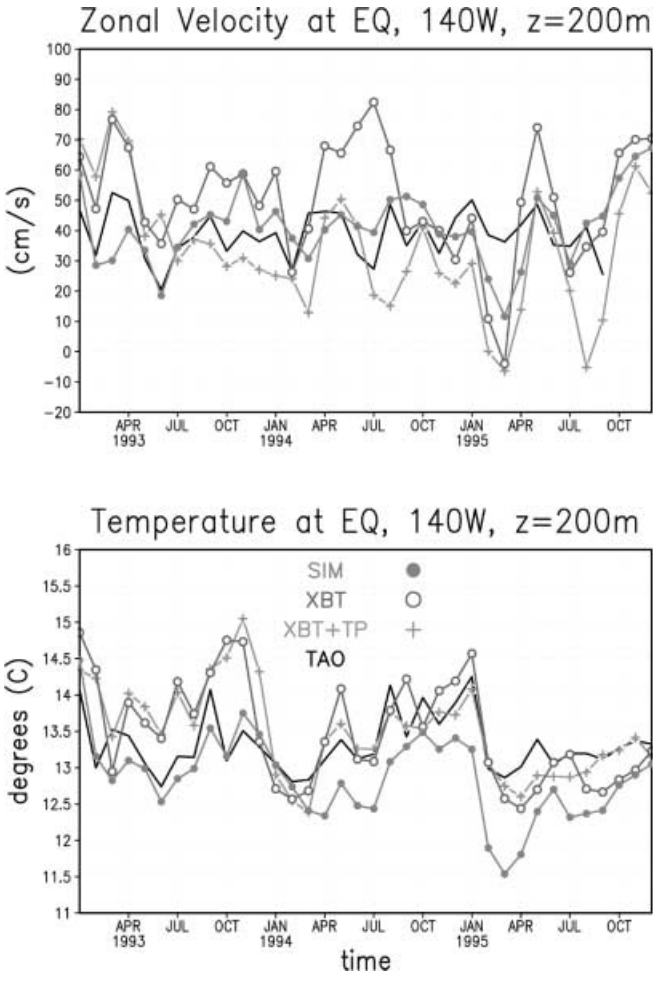

Fig. 14 *Top panel*: comparison of time series of monthly mean zonal velocities at 140°W on the equator and at the depth of 200 m for the TOGA-TAO data (*solid line*), the assimilation experiment with hydrographic data only (*open circle*), the assimilation experiment with the assimilation of both hydrographic and altimeter data (*cross*) and the simulation experiment (*solid circle*). *Bottom panel*: as above but for the monthly mean temperatures

model improves the analysis with respect to simulation estimates done with the same initial condition and atmospheric forcing. In spite of the better vertical resolution of the simulation, the coarser analyses are closer to independent observations. We have shown that the assimilation of the most comprehensive temperature profile data set into an ocean numerical model significantly reduces some of the errors of the model in the equatorial region such as the warm temperatures of the mixed layer in the western Pacific, and the deepening and diffusion of the thermocline gradient in the eastern Pacific. Away from the equator, the assimilation of temperature data improves, in particular, the vertical structure of the temperature field in the NECC region by producing a stronger upwelling than in the simulation and correcting the amplitude and the location of the climatological Gulf Stream temperature front. The comparison of temperature and zonal velocity time series in the assimilation and the simulation experiments with TAO data confirms that this method consistently improves the model estimation of the temperature both at seasonal and interannual time scales. Some of the difficulties in reproducing accurate zonal velocity fields in the tropical Pacific are particularly evident when only the hydrographic data are assimilated and have been overcome by the assimilation of altimeter data which gives a uniform and dense data coverage able to reduce some unrealistic temperature structures introduced by the assimilation of sparse CTDs. Another possibility which might explain the lack of improvement of the velocity field in some regions is the possible negative effect of the univariate assimilation on the water mass properties.

We have shown that the present ODA system improves the representation of the tropical upper ocean thermal variability both at seasonal and interannual time scales. The use of the altimeter together with the hydrographic observations shows encouraging results especially in the zonal velocity representation. One of the advantages of the statistical method used to assimilate the altimeter observations is its possible extension for obtaining also salinity corrections. Future developments concentrate on the implementation of a multivariate assimilation scheme with particular emphasis on the possibility to correct also the salinity and the velocity field.

Acknowledgements We are grateful to Mike McPhaden of NOAA PMEL for providing us with the data from the TOGA TAO array. The altimeter products have been produced by the CLS Space

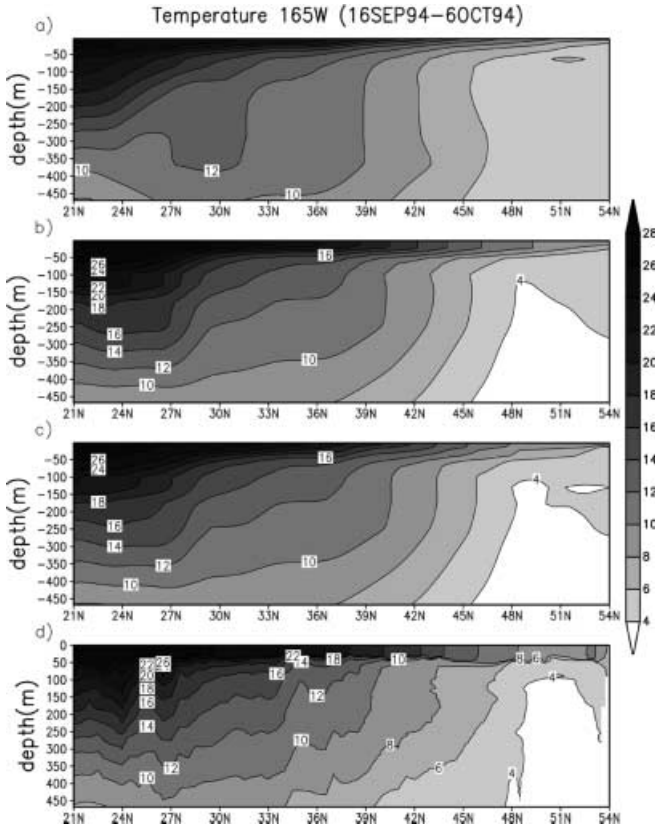

Fig. 15a–d A comparison of a depth–latitude section of temperature taken at 165°W and time-averaged between 16 September 1994 and 6 October 1994 for **a** model simulation, **b** assimilation with hydrographic data only, **c** assimilation with both hydrographic and altimeter data and **d** WOCE data

Oceanography Division as part of the European Union Environment and Climate project AGORA (Contract ENV4-CT95-0113) and DUACS (ENV4-CT96-0357) with financial support from the CEO programme (Centre for Earth Observation) and Midi-Pyrenees regional council. In particular, we wish to thank P.Y. Le Traon for the altimeter data. We also thank A. Fox for providing us with the processed WOCE data. The simulation data set has been provided by Martin Fischer. One of the authors (SM) has been supported by the Environment and Climate program, AGORA Project (Contract ENV4-CT95-0113) and the EU-CEO program, GANES Project (Contract ENV4-CT98-0744).

References

Behringer DW, Ji M, Leetmaa A (1998) An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: the ocean data assimilation system. Mon Weather Rev 126: 1013–1021

Bennett AF, Budgell WP (1987) Ocean data assimilation and the Kalman filter: spatial regularity. J Phys Oceanogr 17: 1583–1601

Carton JA, Chepurin G, Cao X (2000a) A simple ocean data assimilation analysis of the global upper ocean 1950–95. Part I: methodology. J Phys Oceanogr 30: 294–309

Carton JA, Chepurin G, Cao X (2000b) A simple ocean data assimilation analysis of the global upper ocean 1950–95. Part II: resutls. J Phys Oceanogr 30: 311–326

Cox MD (1984) A primitive equation, 3-dimensional model of the ocean. GFDL Ocean Group Tech Rep 1, pp 143

Derber J, Rosati A (1989) A global oceanic data assimilation system. J Phys Oceanogr 19: 1333–1347

Fischer M, Navarra A (2000) GIOTTO: a coupled atmosphere-ocean general circulation model, the tropics. Q J R Meteorol Soc 126: 1991–2012

Goddard L, Graham NE (1997) El Niño in the 1990s. J Geophys Res 102: 10 423–10 436

Ji M, Smith TM (1995) Ocean model response to temperature data assimilation and varying surface wind stress: intercomparisons and implications for climate forecast. Mon Weather Rev 123: 1811–1821

Ji M, Leetmaa A, Derber J (1995) An ocean analysis system for seasonal to interannual climate studies. Mon Weather Rev 123: 460–481

Ji M, Reynolds RW, Behringer DW (2000) Use of TOPEX/Poseidon sea level data for ocean analyses and ENSO prediction: some early results. J Clim 13: 216–231

Le Dimet FX, Talagrand O (1986) Variational algorithms for analysis and assimilation of meteorological observations. Tellus 38A: 97–110

Le Traon PY, Nadal F, Ducet N (1998) An improved mapping method of multisatellite altimeter data. J Atmos Oceanic Tech 15: 522–534

Levitus S, Boyer TP (1994) World Ocean Atlas 1994 vol 4: temperature. NOAA Atlas NESDIS 4, US Government Printing Office, Washington, DC, pp 117

Lorenc AC (1986) Analysis methods for numerical weather prediction. Q J R Meteorol Soc 112: 1177–1194

Mellor GL, Ezer T (1991) A Gulf Stream model and altimetry assimilation scheme. J Geophys Res 96: 8779–8795

Mellor GL, Yamada T (1982) Development of a turbulence closure model for geophysical fluid problems. Rev Geophys Space Phys 20: 851–875

Miyakoda K, Rosati A, Gudgel RG (1997) Prediction of internal climate variations. NATO-ASI series 16, Springer, Berlin Heidelberg New York, pp 125

Pinardi N, Masina S, Navarra A, Miyakoda K, Masetti E (1997) Global ocean data assimilation of temperature data: preliminary results. Operational Oceanography – The challenge for European Co-operation. Proc First Int Conf EuroGOOS, Elsevier Oceanography Series, 62, Elsevier, Amsterdam

Reynolds RW, Smith TM (1994) Improved global sea surface temperature analyses using optimum interpolation. J Clim 7: 929–948

Rosati A, Miyakoda K (1988) A general circulation model for upper ocean simulations. J Phys Oceanogr 18: 1601–1626

Rosati A, Gudgel RG, Miyakoda K (1995) Decadal analysis produced from an ocean data assimilation system. Mon Weather Rev 123: 2206–2228

Smagorinsky J (1993) Some historical remarks on the use of nonlinear viscosities. In: Galperin B, Orzag SA (eds) Large eddy simulation of complex engineering and geo-physical flows. Cambridge University Press, Cambridge, UK, pp 3–36