

Ocean Data Assimilation Using Intermittent Analyses and Continuous Model Error Correction

Bohua Huang^①, James L. Kinter III and Paul S. Schopf

Center for Ocean–Land–Atmosphere Studies, Institute of Global Environment and Society, Maryland 20705

(Received December 15, 2001; revised June 7, 2002)

ABSTRACT

A new data insertion approach is applied to the Derber and Rosati ocean data assimilation (ODA) system, a system that uses a variational scheme to analyze ocean temperature and provide ocean model corrections continuously. Utilizing the same analysis component as the original system, the new approach conducts analyses to derive model corrections intermittently at once-daily intervals. A technique similar to the Incremental Analysis Update (IAU) method of Bloom et al. is applied to incorporate the corrections into the model gradually and continuously. This approach is computationally more economical than the original.

A 13-year global ocean analysis from 1986 to 1998 is produced using this new approach and compared with an analysis based on the original one. An examination of both analyses in the tropical Pacific Ocean shows that they have qualitatively similar annual and interannual temperature variability. However, the new approach produces smoother monthly analyses. Moreover, compared to the independent observations from current meters, the new equatorial currents are significantly better than the original analyses, not only in maintaining the mean state but also in capturing the annual and interannual variations.

Key words: data assimilation, ocean general circulation model, ocean current, temperature profiles, sea surface temperature

1. Introduction

In this study, we examine two issues concerning an ocean data assimilation (ODA) system (Derber and Rosati 1989) that continuously assimilates temperature observations into an ocean general circulation model (OGCM). The first issue is how to best incorporate into the model the corrections derived from observational data. The other is how to conduct long-term analysis in a computationally efficient way. Both issues have pragmatic consequences for assimilating *in situ* observations into complex OGCMs for climate forecasting and research.

In a preliminary analysis using this ODA system, we noticed sudden and large fluctuations of the equatorial undercurrent (EU) in the Pacific when temperature observations were continuously assimilated into an OGCM (Huang and Kinter 1997). Further examinations suggest that these unrealistic features are related to a combinative effect of the repeated insertion (RI) treatment and the magnitudes of the prescribed forecast and observational error estimates.

^①E-mail: huangb@cola.iges.org

The RI technique repeatedly uses an observation in the model during a period bracketing the observational time (e.g., MacPherson 1991; Bloom et al. 1996). MacPherson (1991) demonstrated that the RI technique filters high frequency signals excited in the model by data insertion. However, the damping also continuously weakens the internal high frequency variability that provides a mechanism for model internal adjustment. This can be deleterious for ocean analysis since ocean observations are predominantly of the temperature field. Generally, ODA systems rely on model dynamics to pass assimilated information from temperature to velocity. This problem may be more serious near the equator because the geostrophic constraint disappears there. To balance the strengths and weaknesses, it is important to apply appropriate weights to the prescribed error covariance matrices for both the model and the observations in the ODA scheme (Derber and Rosati 1989). Recently, Behringer et al. (1998) reduced the magnitude of the estimated model error in the National Centers for Environment Prediction (NCEP) ODA system, and they reported improved analysis results.

The computational cost of a continuous ODA system can be quite high. For observations randomly distributed in space and time, a "station nudging" approach (MacPherson 1991) needs to solve a three-dimensional minimization problem (Derber and Rosati 1989) at every time step. This makes continuous assimilation much more expensive than intermittent methods. Moreover, there is a certain degree of redundancy in solving the problem at every time step since the time step of a model is much smaller than the RI window. Therefore, it is desirable to appropriately reduce the number of times the minimization problem is solved in the assimilation process while preserving some of the advantages of the continuous assimilation approach, such as the small initial shock and frequent error correction. Rosati et al. (1995) proposed a simplification procedure to reduce the frequency of data assimilation steps based on a computational consideration. In this study, we show that applying the incremental analysis update (IAU) technique (Bloom et al. 1996) provides benefits on both issues.

The IAU technique was originally designed for an intermittent data assimilation system in meteorology as an alternative initialization procedure to reduce analysis-induced initial shock in model prediction. The idea is to apply the correction derived from the analysis step to the model gradually as a forcing field. Bloom et al. (1996) demonstrated that this strategy efficiently suppresses gravity waves excited by the intermittent analysis insertion while not significantly affecting model internal dynamical adjustment. Carton et al. (2000a, b) implemented the IAU technique in a global intermittent ODA system for long-term analysis experiments. Apart from the usual benefit expected from this technique, they also reported a reduction in model systematic bias because of the more persistent effect from incremental forcing. During the process, both Bloom et al. (1996) and Carton et al. (2000a, b) have preserved the prediction-analysis cycle, a basic feature of intermittent systems.

In this study, we apply the IAU strategy in the opposite way to an originally continuous ODA system. Our purpose is to extend the interval between analyses while preserving the basic feature of a continuous assimilation system. First, we choose an analysis interval much smaller than the width of the RI window so that an observation can affect several adjacent analyses as in a typical continuous assimilation. Second, the ocean model is continuously driven by an IAU-type forcing derived from this analysis. Details of the IAU treatment used

here and its differences from the original procedures proposed by Bloom et al. (1996) will be discussed in the next section. A benefit of this approach is the significantly reduced computational requirement. As we will show, it also leads to an effective change of data / model contribution to the analyses.

We conduct a thirteen-year (1986–1998) analysis using this strategy. Compared with analyses using the conventional continuous assimilation, the strategy produces smoother temperature and velocity fields while still efficiently forcing the model toward observations. Moreover, the errors of the equatorial undercurrent variability appearing in the conventional analyses in the central and eastern equatorial Pacific are corrected by the new strategy. It seems that an appropriate choice of the analysis interval together with a sensible way of correcting model error can help the model's dynamical adjustments.

In the next section, we describe the ocean model, the observational data, and the assimilation method used in our study. We also present details of the new assimilation strategy, including the IAU treatment, for our experiments. Section 3 presents the results. Section 4 gives an analytical example to explain the different results from these two analysis runs. A summary and conclusions are given in section 5.

2. Ocean data assimilation system

2.1 The OGCM

The version of the GFDL Modular Ocean Model (Pacanowski et al. 1993) used here is the same as the one used in previous studies (Huang and Schneider 1995; Huang and Kinter 1997; Schneider et al. 1999). The domain is that of the world oceans between 70°S and 65°N. The minimum and maximum ocean depths are set to 300 and 4000 meters, respectively. The zonal resolution is 1.5°. The meridional resolution is 0.5° between 10°S and 10°N, gradually increasing to 1.5° at 20°S and 20°N and fixed at 1.5° poleward of 20°. There are 20 levels in the vertical direction with a constant interval of 15 meters for the top 10 levels. The time step is 1 hour. Richardson-number-dependent coefficients are chosen for the vertical mixing and diffusion of momentum, heat, and salinity (Pacanowski and Philander 1981). The horizontal viscosity and diffusivity coefficients are prescribed and set equal to a constant $2 \times 10^7 \text{ cm}^2 \text{ s}^{-1}$. The temperature and salinity near the northern and southern boundaries are relaxed to the Levitus (1982) monthly climatology.

The model was forced by the wind stress, heat flux, and salinity flux at the sea surface. The monthly mean surface wind stress was computed from the four-times-daily European Centre for Medium Range Weather Forecasts (ECMWF) 1000 hPa wind analysis, using a constant drag coefficient. The winds were interpolated from spectral resolution T106 to a spatial resolution of zonal 1.125° and Gaussian latitudes ranging from 1–1.5° spacing. The instantaneous wind stress was calculated in each grid cell and averaged within each month to form the monthly mean fields.

The heat flux is composed of radiative, sensible, and evaporative components. The radiative fluxes are divided into solar and longwave radiative parts. At each time step of the model integration, the solar radiation was prescribed by a linear interpolation of the climatological monthly means analyzed by Oberhuber (1988) from the Comprehensive Ocean Atmosphere Data Set (COADS). The solar radiation was assumed to penetrate into the upper ocean with an e-folding depth of 12 m. The other components were parameterized as functions of the model sea surface temperature (SST) and prescribed surface air temperature and

surface wind speed. The formulation for net longwave radiative flux was taken from Rosati and Miyakoda (1988). The sensible and latent heat fluxes were parameterized by bulk turbulent transfer formulae with parameters as given by Philander et al. (1987). The surface wind speed was derived from the instantaneous wind stress in an approximate manner, equivalent to assuming a constant drag coefficient of 1.4×10^{-3} . The surface (two meters) air temperature was taken from monthly mean ECMWF analyses. The model surface salinity was relaxed to the Levitus (1982) monthly climatology.

2.2 Ocean temperature observations

The observations used as input to the data assimilation system were SST and vertical temperature profiles. The SST observations were taken from the weekly global $1^\circ \times 1^\circ$ analysis of the U.S. Climate Prediction Center (CPC, Reynolds and Smith 1994).

For this study, we have assembled a comprehensive set of temperature profile observations from various sources. It is based on three data sets from the National Oceanographic Data Center (NODC): the updated World Ocean Database 1994 CD-ROM series (Levitus and Boyer 1994, referred to as WOD94 later on), the Global Temperature–Salinity Profile Program (GTSP) database (Hamilton 1994), and NODC online real-time data. Each provides observations for a time period of our analysis. More observations were added from two additional data sets. The first is from the NCEP (Ji et al. 1995) and the other from the Tropical Ocean–Global Atmosphere (TOGA) Program's Tropical Atmosphere–Ocean (TAO) moored buoy array in the tropical Pacific Ocean (Hayes et al. 1991; McPhaden et al. 1998).

The WOD94 is constructed from data in the NODC master archives and collected by the National Oceanographic Data Archaeology and Rescue and the Global Oceanographic Data Archaeology and Rescue Projects. For this study, we selected those temperature profiles that were measured during 1986–1992 and passed the NODC quality control criteria. The GTSP database contains oceanographic observations transmitted both in real-time over the Global Telecommunications System (GTS) and collected in delayed-mode since 1990, from which we have derived temperature profiles for the period of 1993–1996.

The NCEP data set contains expendable bathythermograph (XBT) measurements for the Atlantic and Pacific Oceans from near real-time GTS transmission that passed a series of error checking and correction procedures (Ji et al. 1995). While there are many duplicates in the NCEP profiles with the WOD94 and the GTSP data sets, the NCEP data are not just a subset of the other two (Huang and Kinter 1997). After eliminating apparent duplications, we include the remaining NCEP profiles for our analyses. An NCEP profile was considered a duplicate if it fell in the same $0.5^\circ \times 0.5^\circ$ square (the minimum model grid length) as a profile from either WOD94 or GTSP within one hour (the model time step). In fact, the NCEP data modestly increase the total number of profiles during 1990–1992 when the WOD94 totals were lower.

We also obtained daily mean temperature from up to 75 chain moorings of the TOGA TAO array directly from the Pacific Marine Environmental Laboratory (PMEL) FTP site. Although some of these data have been included in the data sets mentioned above, the purpose was to make sure that all TOGA TAO data were used for our analyses. Similar to the procedure discussed in the last paragraph, apparent duplications between the TOGA TAO data and the other data sets were eliminated and the remaining temperature profiles were added to the total profiles.

In summary, the temperature profiles for our analyses for 1986–1996 were mainly from either the WOD94 (1986–1992) or the GTSP (1993–1996) databases, which were supple-

mented by additional profiles from the NCEP and the TOGA TAO data sets. Starting in 1997, we have taken real-time XBT measurements directly from an NODC FTP site at the end of each month with additions from TOGA TAO. The ODA analyses at the Center for Ocean–Land–Atmosphere Studies (COLA) have been conducted on a monthly basis in near real-time. The GTSP and NCEP data were not included because they were not available in near real-time. The number of temperature profiles used for assimilation is generally around 80 000–100 000 per year, with a maximum of 120 000 per year.

For a monthly ODA analysis, all selected profiles were first interpolated linearly to model levels. Several quality control steps similar to those done at NODC were then repeated to make sure that there were no large temperature inversions and the observations were within a reasonable range of the Levitus and Boyer (1994) monthly climatology and seasonal standard deviation. This procedure provides a common criterion for profiles that went through somewhat different prior quality control procedures. It also eliminates improper features possibly created in the vertical interpolation. A topography check was then conducted to make sure that no data were on land or below the bathymetry relative to the model grid. Some of these tests were also done to the SST data assigned to the first model level. Finally, a check was done to all temperature profiles and SST together to eliminate possible duplications that existed within each original group and to make sure that a temperature profile and a separate SST observation do not appear in the same place at the same time.

2.3 Data assimilation scheme

Our experiments were conducted with a version of the Derber and Rosati scheme implemented on the OGCM at COLA (Huang and Kinter 1997). This assimilation scheme is a univariate, two-dimensional variational optimum interpolation of temperature only in which the first guess is provided by the OGCM. The observations are inserted into the ocean model integration by computing a temperature correction field derived from the optimum interpolation equation that is formulated using a variational approach. This involves minimizing a cost function with two terms that are associated with the fit of the objective analysis to the model field and the fit of the analysis to the observed field, respectively. The cost function may be written as

$$J = (T - T_f)^T E^{-1} (T - T_f) + (D(T) - T_o)^T F^{-1} (D(T) - T_o), \quad (1)$$

where T is the analyzed temperature vector and T_f is the first guess temperature vector, D is a transformation matrix (usually a spatial interpolation operator), and T_o is the vector of temperature observations. The temperature correction field ΔT is defined as the analysis minus first guess ($T - T_f$). The forecast error covariance matrix E is specified, assuming a Gaussian-like distribution of the horizontal covariance (Derber and Rosati 1989). Its e -folding spatial scale is prescribed as 4° . Assuming the observational errors are not spatially correlated, the observational error covariance matrix F is diagonal. The functional (1) is minimized using the preconditioned conjugate gradient algorithm through an iterative procedure (Derber and Rosati 1989). A maximum of three iterations is made. The model temperature is updated through the assimilation procedure only for the levels above 400 meters (upper 16 levels). If a temperature inversion occurs between the 16th and 17th levels (676 meters) at any grid point after an analysis, the temperature at the deeper level is replaced by a value 0.1°C below the temperature at the 16th level.

A 10-day RI window was used in this assimilation so that the temperature correction at

a given instant included observations within ± 5 days of that instant. This window is narrower than that used in our previous study (Huang and Kinter 1997) and some other studies (Derber and Rosati 1989; Rosati et al. 1995; Ji et al. 1995). The window width is based on the consideration that daily TOGA TAO measurements in the tropical Pacific should be adequate to resolve the propagation of equatorial waves more accurately. In principle, the error covariances of both the observations and the model should have corresponding adjustments because the resolvable phenomena are changed. However, since our original estimates of error covariance are crude, and also since no relevant statistics are available from observations, any such change will inevitably be a bit arbitrary. We produced two-year analyses using both 10-day and 30-day windows with the original covariance estimate and did not find systematic differences between them. Therefore, we decided to keep the error covariance unchanged until more solid information becomes available.

The contribution of each observation was weighted by the diagonal element of F^{-1} , that is the reciprocal of the observational error variance multiplied by a factor linearly dependent on the length of the time interval between the instant of the observation and that of the assimilation. The observational error includes systematic error, instrument error, and representation error associated with unresolved physical processes. The observational temperature error variance of temperature profile data was assigned to be 0.25°C^2 for most points. For the profiles near the northwestern Pacific and Atlantic Oceans, however, the variance was assigned to be 1°C^2 . This is because meso-scale eddies observed in these areas are not resolved by the OGCM and a single observation there is much less representative of the model's grid-scale thermal structure. The error variance of the SST data was also assigned to be 1°C^2 because the CPC analysis is not exactly an "observation".

A nominal variance for the forecast error, which determines the magnitude of the elements in matrix E , was assigned to be 0.01°C^2 . The forecast error covariance is the estimate of the error in the model state after 1 hour of integration from the last data ingest. It is partly related to the model initial error growth, but is constrained by the continuous data insertion within the 10-day window. Therefore the assumption that E is small in comparison with F is not indicative of the high quality of the model, but is a statement about the effectiveness of the repeated data insertion itself.

2.4 Intermittent analysis update (IAU)

In this subsection, we describe the modified strategy with the intermittent analysis and IAU model error correction and compare it with the original strategy. A schematic figure of these two strategies is shown in Fig. 1, which corresponds to the two analysis experiments we have conducted.

The *Control* analyses use the original ODA strategy (Fig. 1a). For this assimilation, an analysis has been done using the ODA scheme described in section 2.3 at every time step in the first twelve hours of each day. At each time step, the temperature correction ΔT is immediately inserted into the model for further integration. These instant corrections are represented with double-headed arrows in Fig. 1a. For the rest of the day, the correction is inserted at every time step with the temperature correction derived at 12 h ($\Delta T_{12\text{ h}}$). The persistent corrections are denoted with single-headed arrows in Fig. 1a. This persistent treatment was proposed by Rosati et al. (1995) for computational efficiency and has been adopted in several other studies (Ji et al. 1995; Huang and Kinter 1997).

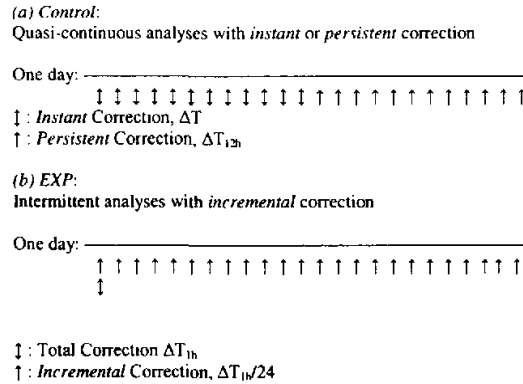


Fig. 1. Schematics of the ODA strategies for (a) the *Control* and (b) the *EXP* analyses. In *Control*, the model is forced with instant temperature corrections (ΔT , double-headed arrows) for 1–12 h each day. During the rest of the day, it is forced by persistent corrections (i.e., ΔT_{12h} , single-headed arrows, upper panel). In *EXP*, the model is forced with the incremental corrections ($\Delta T_{1h}/24$, single-headed arrows, lower panel) at each time step, where ΔT_{1h} itself is represented by a double-headed arrow in (b).

The *EXP* analyses use our modified strategy (Fig. 1b). In this case, an analysis is only conducted once daily at 1h. In Fig. 1b, the correction field (ΔT_{1h}) from this analysis is also denoted as a double-headed arrow (Fig. 1b). A constant forcing field to the OGCM during the day is derived using the IAU strategy as follows. First, ΔT_{1h} is divided by the total number of time steps in a day ($\Delta T_{1h}/24$, short arrows in Fig. 1b) and converted into a 24-h time tendency. It is then added to the model temperature field at each time step, including the one in which ΔT_{1h} is derived. These incremental corrections are denoted as single-headed arrows in Fig. 1b. (Note that the meanings of the single-headed arrows in the *Control* (Fig. 1a) and *EXP* (Fig. 1b) are different.) The ratio of computation times for solving the minimization problem between *EXP* and *Control* is 1:12.

As mentioned in the introduction, the IAU procedure adopted here is slightly different from that in Bloom et al. (1996) or Carton et al. (2000a, b). Applying the IAU to intermittent systems with a relatively large analysis interval, they follow the predictor–corrector steps in a forecast–analysis cycle. First, the instantaneous state at the mid-point of an analysis interval is forecast as the first guess for an analysis with observations within the interval. Regarding the analysis as a mean of the interval, they then go back to the starting point and carry out an integration with the IAU forcing throughout the interval (see Fig. 1 in Bloom et al. 1996). Instead, we use the IAU forcing derived from an analysis to correct forecast errors during its integration toward the next analysis time and eliminate the forecast step of the previous studies. Since the *EXP* analyses are done with data within a RI window that is much wider than the analysis interval, this change to “forward” correction should make little difference. However, it preserves a major characteristic of continuous assimilation, namely, reducing model error and bias throughout the integration.

The advantage of the IAU treatment is that it damps high frequency signals stimulated by data insertion but preserves model internal higher frequency fluctuations. Bloom et al. (1996) have analyzed the effect of different error correction methods on a linear system where

the solution can be expressed as a superposition of a "free" part determined by the initial state and an externally forced part. In comparison to the intermittent assimilation solution without initialization, the IAU-type correction damps the high frequency signals in the forced solution more strongly but does not change the free solution. Instead, a nudging-type correction, in which the model variable is relaxed to the analysis field itself at fixed rates, damps high frequency signals in both solutions. We have repeated some of the derivations to see if these characteristics are changed by our modification to a forward correction scheme, and we found that the amplitude of the complex response function is similar to the one derived by Bloom et al. (1996).

The integration for each experiment spans 13 years from 1986 to 1998 after a four-year spin-up from an ocean at rest with January climatological temperature and salinity (Levitus 1982) forced with the monthly mean climatological wind stress. Monthly mean ocean analyses and instantaneous ocean fields are saved at the end of each month for both analyses.

3. Results

Basic features of ODA analyses, such as the correction of systematic model errors in temperature, have been demonstrated in previous studies (Derber and Rosati 1989; Ji et al. 1995; Rosati et al. 1995; Huang and Kinter 1997). In this section, we examine the effects and possible improvements of the modified insertion strategy on the ODA analyses. For this purpose, we have made comparisons between the *EXP* and *Control* experiments as well as between analyses from each experiment and the observations. Some of the observations have been incorporated into the model during the assimilation process, so the comparison measures how effectively these insertion strategies force the model toward the observations. Others are valuable independent measurements of the quality of the assimilation process.

An improvement due to data insertion can be claimed only if the model error is further reduced. Therefore, we also compare the analyses with a model simulation without data assimilation (referred to as *SIM*). *SIM* is done under the same boundary and initial conditions except that the monthly climatology of the air temperature was used in the parameterized surface heat flux, as opposed to the real time monthly data, because the simulation was originally designed for another purpose (see discussions in Huang and Shukla 1996). This more diffusive surface heat flux may degrade the magnitude of simulated SST variability in the extra-tropics.

3.1 Analysis highlights

Figure 2 shows the anomalies of the upper ocean heat content (HC) along the equator in the Pacific from both analyses. The HC is defined as the averaged ocean temperature from the sea surface to 234 meters. This figure highlights the time history of the oceanic low frequency variability associated with the El Niño–Southern Oscillation (ENSO) during the period of the analysis. It features three major warm episodes in the eastern Pacific: 1986/1987, 1991/1992, and 1997/1998. A strong cold episode took place in 1988/1989 and another one with possibly larger amplitude was initiated during 1998 and was still in its developing stage at the end of the analysis period. Before the major warm events, positive HC anomalies first accumulate along the equator west of the dateline and subsequently propagate eastward. The 1991/1992 event spanned a longer lifetime than the other two, and on it were superimposed several fast eastward moving surges of warm anomalies which occurred between late 1989 and early 1991, before the major eastward shift in late 1991. After the

1991 / 1992 warm episode, low frequency oscillations seemed to have been temporarily disrupted. The eastern ocean was dominated by alternating warm / cold signals that originated in the central ocean until 1997, when a much stronger and more abrupt warm event occurred with peak HC anomaly reaching 5°C . These HC fluctuations are consistent with those described in Rosati et al. (1995), Ji et al. (1996), and Behringer et al. (1998) during the period common to this study. The development of the 1997 / 1998 warm event is similar to that described by Trenberth (1998).

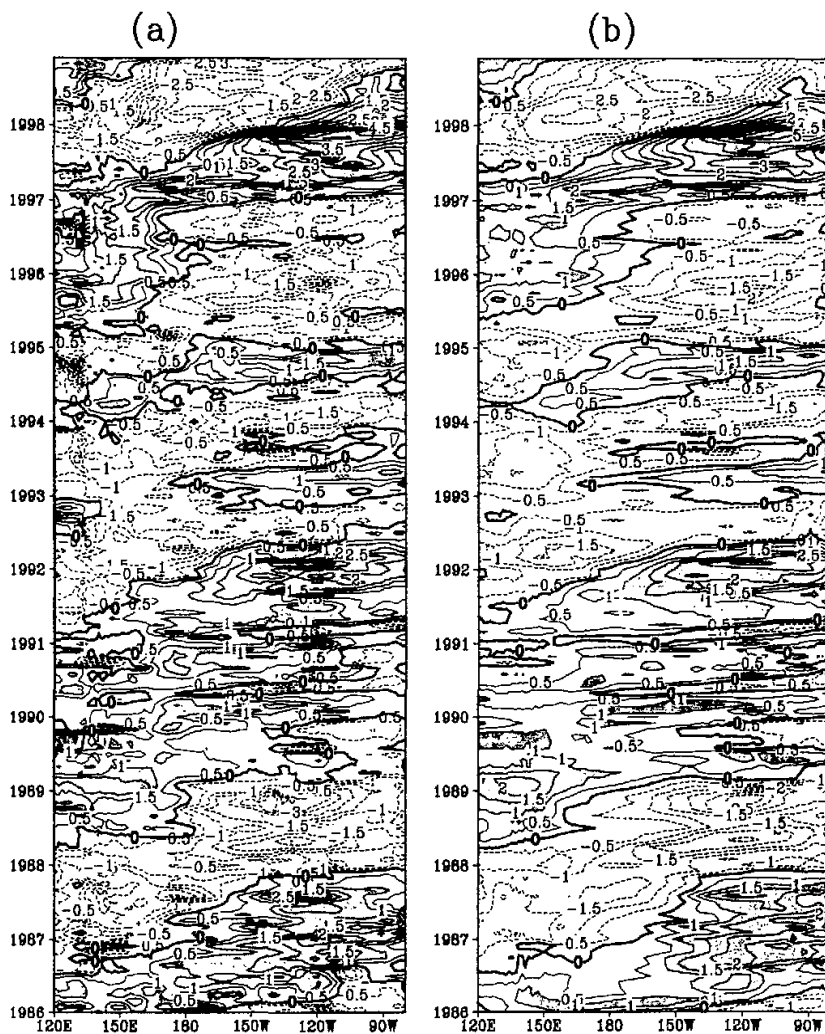


Fig. 2. Time-longitude sections of monthly HC anomalies along the equator in the Pacific Ocean during the 1986–1998 period for (a) the *Control* and (b) the *EXP* analyses. The contour interval is 0.5°C . Regions with anomalies greater than 1°C have dark shading and regions with anomalies less than -1°C have light shading.

Comparing Fig. 2a with Fig. 2b, we found that both analyses give a qualitatively similar description of the ocean low frequency variability during this period, but there are substantial differences in the higher frequency changes. The *EXP* fluctuations are generally smoother without as much small scale noise as in *Control*. The differences between the two analyses are also demonstrated in the temperature field in individual months (Fig. 3).

Figure 3 shows the longitude–depth sections of temperature along the equatorial Pacific from two chosen months. The first is January 1987 when a warm episode was close to its peak (Figs. 3a, 3c). In this month, the thermocline was quite flat along the equator. The warm pool shifted eastward with the 28°C isotherm approaching 150°W. The second month is June 1988 when a strong cold event prevailed (Figs. 3b, 3d). The situation in the equatorial Pacific, in contrast to that in January 1987, featured a very steep thermocline between east and west. The warm pool was pushed far to the west. These basic features are consistent between both analyses.

The *Control* analyses show a tighter vertical temperature gradient around the thermocline and a somewhat deeper mixed layer, especially in the warmer month. On the other hand, *EXP* appears to produce a relatively smoother spatial structure by filtering out some small-scale features and sharp changes in *Control*. One such sharp change is in June 1988 (Fig. 3d) when the thermocline (or the 20°C isotherm) deepened from the west to 180°W,

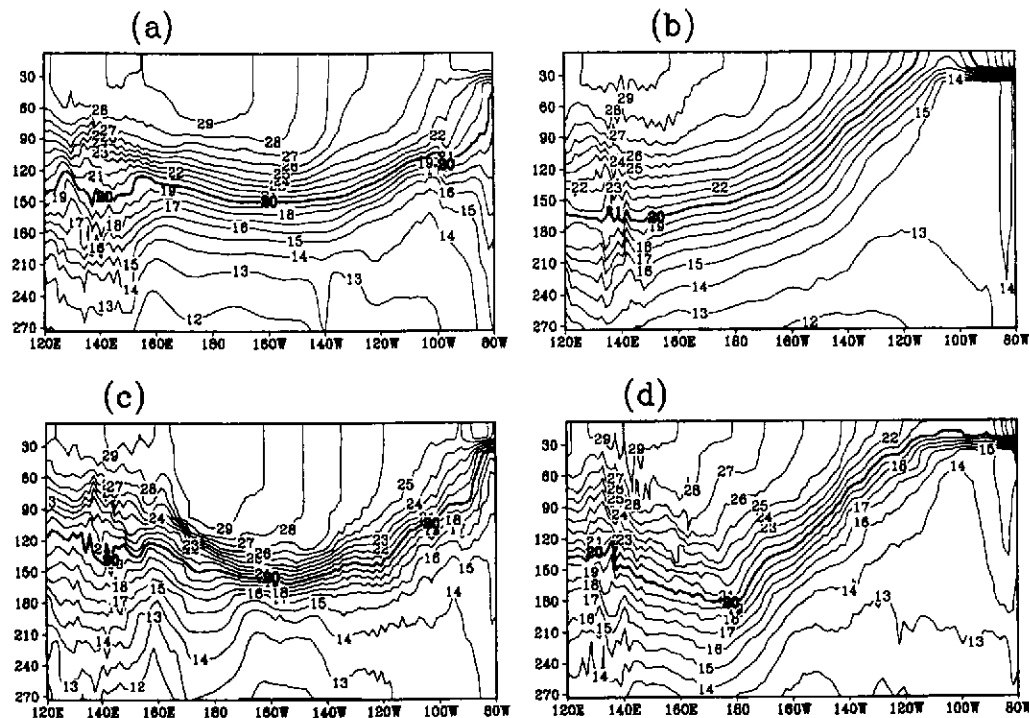


Fig. 3. Depth–longitude sections of monthly mean ocean temperature along the equator in the Pacific Ocean. The upper panels are from *EXP* with (a) for January, 1987 and (b) for June, 1988. The lower panels are from *Control* with (c) for January, 1987 and (d) for June, 1988. For each panel, the contour interval is 1°C and the thick 20°C isotherm represents the position of thermocline. The vertical axis is depth meters.

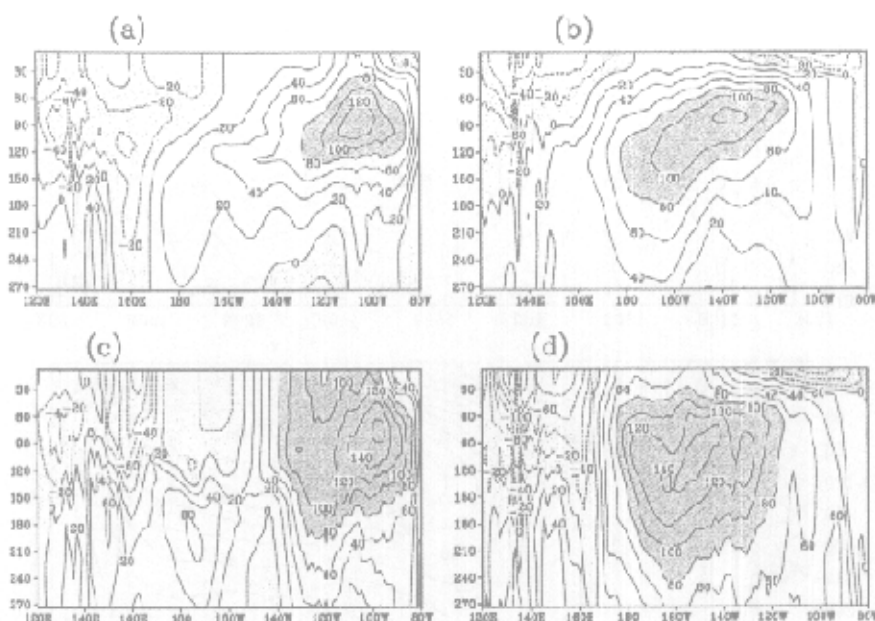


Fig. 4. Depth-longitude sections of monthly mean zonal velocity along the equator in the Pacific Ocean. The upper panels are from *EXP* with (a) for January, 1987 and (b) for June, 1988. The lower panels are from *Control* with (c) for January, 1987 and (d) for June, 1988. For each panel, the contour interval is 20 cm s^{-1} . Regions of eastward currents with speed greater than 80 cm s^{-1} have dark shading and regions of westward currents have light shading. The vertical axis is depth meters.

becoming abruptly shallower beginning at the dateline. This feature, however, is not shown in Fig. 3b, where the thermocline was actually flat in the west and started its gradual shoaling around 160°E .

Figure 4 shows the zonal velocity along the equator for these two months. In both months, the *Control* (lower panel) shows the equatorial undercurrent (EU) with broader vertical expanse and a core speed 20%–30% higher than that of *EXP*. During 1987, it also shows stronger downward mixing from the sea surface (Figs. 4a, 4c). It is interesting that the relatively small differences between *EXP* and *Control* in Fig. 3 are associated with large qualitative differences in the velocity field. Since the velocity field is also related to the time history of the mass field, the different temporal behavior of the temperature field as shown in Fig. 2 may also contribute to the current difference in Fig. 4. Moreover, one should note that the EU core in *Control* is not always stronger than in *EXP* although it is so in both months shown in Fig. 4. In fact, the EU cores of the mean currents for 1986–98 from both analyses are comparable to each other (Figs. 5a, 5b), while both are significantly stronger than that from *SIM* (Fig. 5c). This actually suggests that the EU core fluctuates more violently in *Control* than in *EXP*.

Since the velocity was not assimilated into the ocean model, its determination was totally dependent on the dynamical coupling between variables. Therefore, its quality becomes a

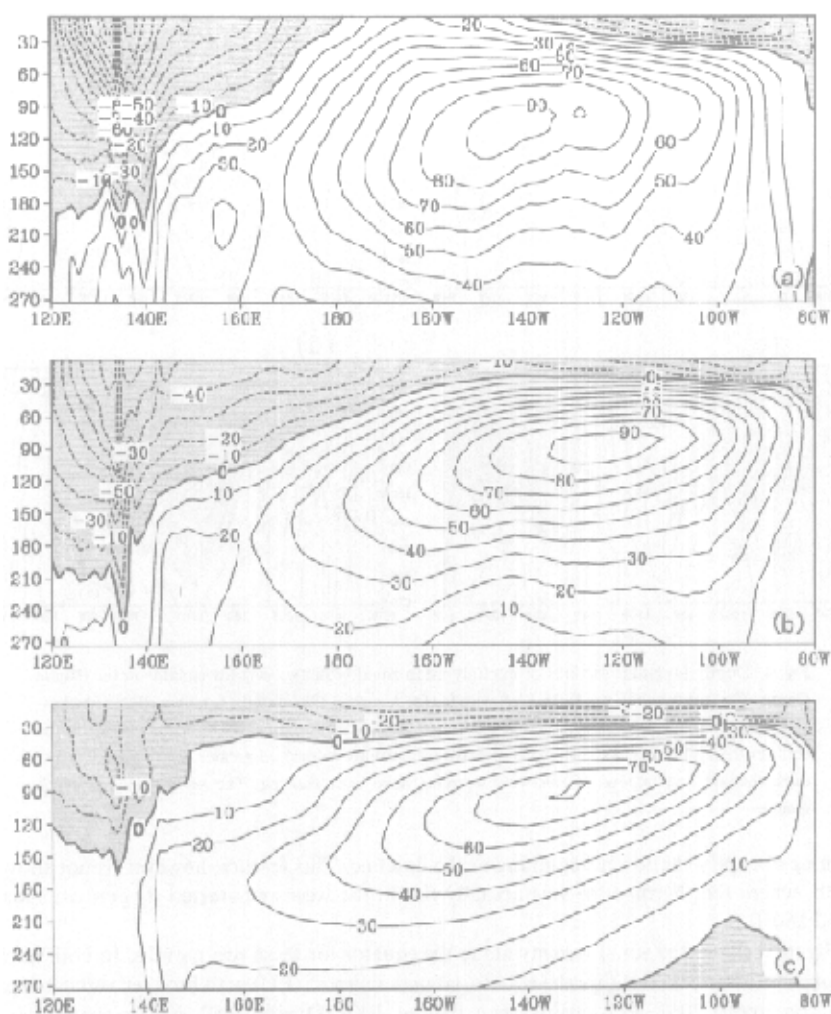


Fig. 5. Depth–longitude sections of mean zonal velocity for 1986–1998 along the equator in the Pacific Ocean from (a) *Control*, (b) *EXP*, and (c) *SIM*. For each panel, the contour interval is 10 cm s^{-1} . Regions of westward currents have light shading. The vertical axis is depth meters.

measure of how properly the dynamical and observational information is balanced in the ODA system. In the next subsection, we present evidence from a comparison with observational data to demonstrate that the equatorial currents in *EXP* do show some features that are superior to those in *Control* and *SIM*.

3.2 Comparison with current observations

The velocity observations used for this comparison are from four current meter moorings

maintained as a component of the full TOGA TAO array. Three of these mooring sites are located along the equator at 165°E, 140°W, and 110°W. They measure daily ocean currents from the surface to 250 meters together with surface winds (McPhaden and McCarty 1992; McCarty and McPhaden 1993). The observations, available from the PMEL FTP site, cover most of our analysis period with the most recent data in mid to late 1997, although there are extended periods of missing data at certain depths. The off-equatorial current meter mooring at 7°N, 140°W has a shorter span of observations from June 1988 to May 1991. For our comparison, the daily current observations are first averaged into monthly means, then vertically interpolated to appropriate model levels. None of the current measurements were assimilated in the ODA system, so the comparison represents an independent check on the assimilations.

Figure 6 shows the mean zonal currents from the analyses, the simulation, and the TAO data. These means were averaged from monthly data within the time interval of available TAO observations at each site. In Fig. 6, each panel shows the results from a mooring site, with a thick solid line for the TAO data, a thin solid line for *Control*, a short dashed line for *EXP*, and a long dashed line for *SIM*.

Compared with *SIM*, the analyses enhance the eastward motion in the upper ocean along the equator at 140°W (Fig. 6a), thus weakening the westward surface current and strengthening the EU core, which make it closer to observations. However, the EU cores from both analyses and *SIM* are all located at positions shallower than that of the observations. This may imply that the vertical mixing in the OGCM is inadequate since previous studies demonstrated that the EU core is largely determined by the balance involving the mixing processes (Yu and Schopf 1997; Behringer et al. 1998). Though having comparable core speed to that in *EXP*, the EU in *Control* is broader both above and below the core, as shown in Fig. 5.

Further to the east, the current profile at 110°W from *EXP* is considerably better than either *Control* or *SIM* (Fig. 6b). Here the eastward current in *SIM* is persistently weaker than the observed. *Control*, however, seriously distorts the shape of the profile, because it substantially increases the speed of the deeper current while keeping the EU core unchanged. The EU core for *EXP* is enhanced significantly which makes its profile very close to the observed one.

Off the equator in the central Pacific (7°N, 140°W), the TAO data features the North Equatorial Counter Current (NECC) with mean eastward currents of 30 cm s⁻¹ prevailing from the surface to near 100 m (Fig. 6c). This feature is well represented by both the analyses and the simulation. In fact, the *Control* error is smaller than the other two, while *EXP* and *SIM* are almost indistinguishable from each other.

On the equator at 165°E (Fig. 6d), the observed profile (thick solid line) is characterized by a "rest" layer of near-zero currents between the depths of 30 and 100 m. This layer separates a westward surface flow of 10 m s⁻¹ from an eastward current below. The latter peaks at a depth of 200 m, with a maximum speed of 40 m s⁻¹. This "rest" layer is well reproduced by *SIM* (long dashed line), though its surface currents are stronger and its deep currents weaker than the observed ones.

The mean currents from the analyses are degraded compared to the simulation (Fig. 6d). Both analyses show much stronger westward mean surface currents (40–50 cm s⁻¹) than the observed. Moreover, these westward currents reduce continuously to zero around the depth of 100 m, eliminating the "rest" layer. The deeper eastward currents are also broader than the observed, while their maximum speeds are weakened to 15–25 cm s⁻¹. This result is similar to the finding of Behringer et al. (1998) in comparing their ODA analyses with current profiler measurements at 165°E.

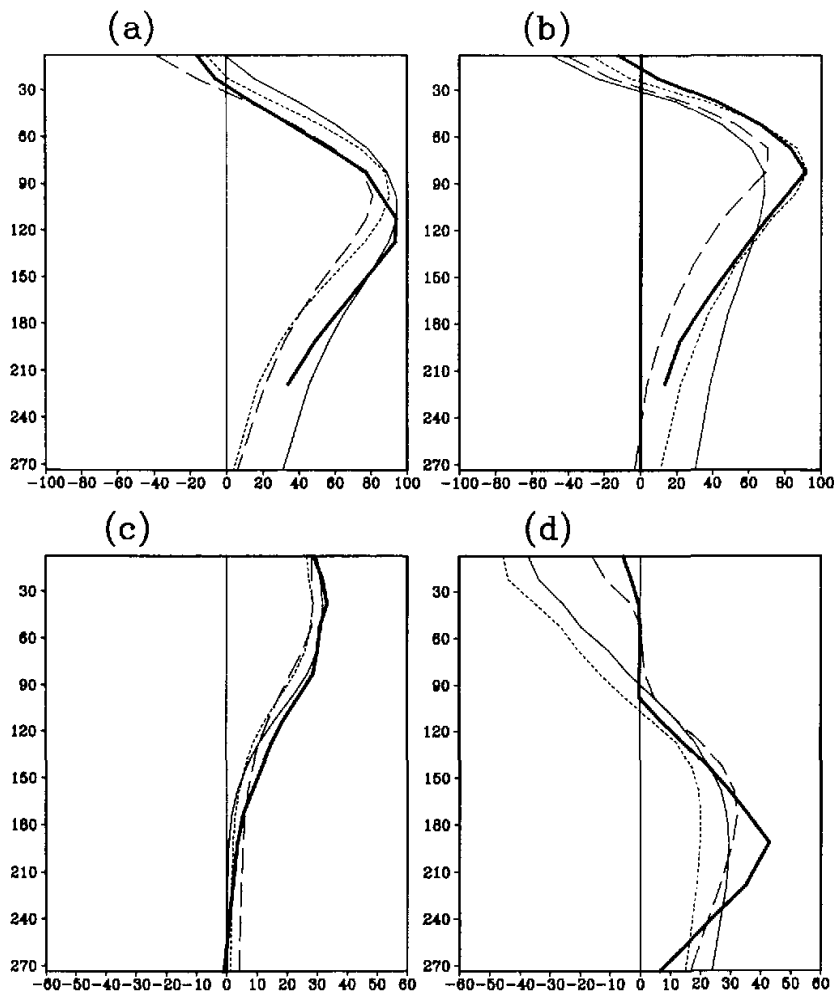


Fig. 6. Mean zonal velocity at (a) Equator, 140°W, (b) Equator, 110°W, (c) 7°N, 140°W, and (d) Equator, 165°E. For each panel, the unit of horizontal axis is cm s^{-1} . The vertical axis is depth with unit in meters. The thick solid curve is from TAO current meters, thin solid curve from *Control*, short dashed curve from *EXP*, long dashed curve from *SIM*.

The root-mean-square (RMS) errors of the zonal currents from the analyses and the simulation with respect to the TOGA TAO mooring data are listed in Table 1. The values given there are averages from all vertical levels. The table shows that *Control* generally has larger RMS errors than either *EXP* or *SIM* for all the mooring locations. *EXP* significantly reduces the errors in *Control* at 140°W and 110°W along the equator. Moreover, the *EXP* RMS errors are a little smaller than those from *SIM* at all locations except the one on the equator at 165°E, where the mean currents in both analyses are apparently problematic.

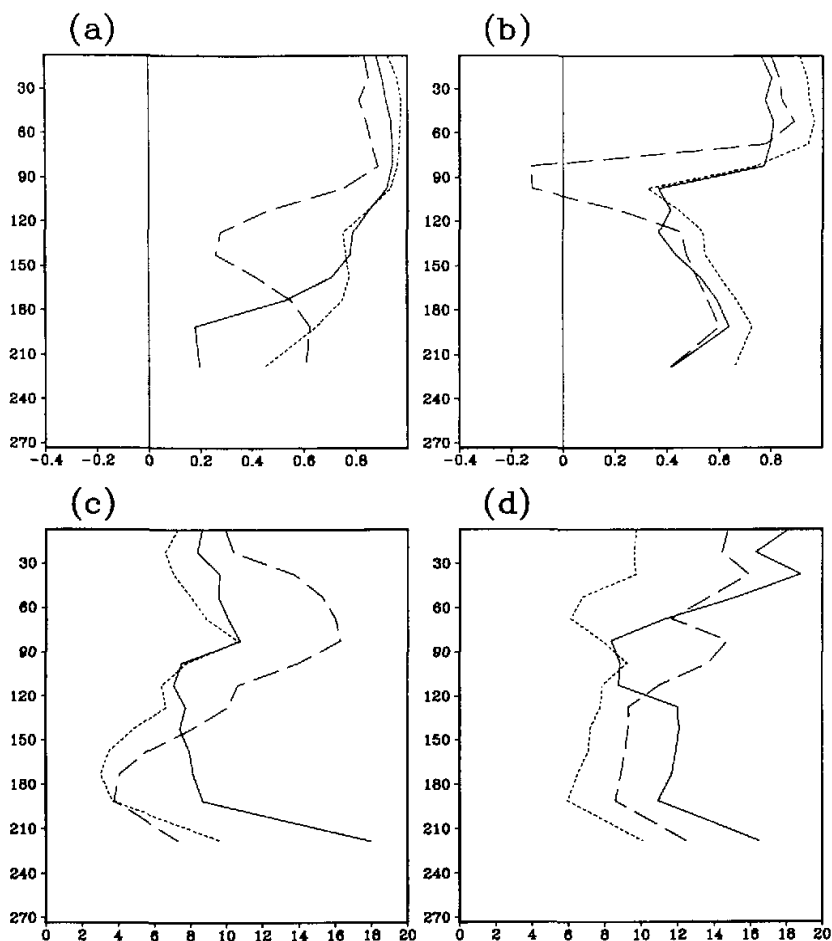


Fig. 7. Correlation coefficients and root-mean-square (RMS) errors of zonal velocity annual cycle at 140°W and 110°W along the equator with respect to TAO current meter measurements. The upper panels are correlation coefficients for (a) 140°W and (b) 110°W . The lower panels are RMS errors for (c) 140°W and (d) 110°W . For each lower panel, the unit of horizontal axis is cm s^{-1} . The vertical axis is depth with unit in meters. The solid curve is from *Control*, short dashed curve from *EXP*, long dashed curve from *SIM*.

Table 1. RMS errors of the zonal currents from *Control*, *EXP*, and *SIM* averaged for the 14 upper model levels at the TOGA TAO current meter mooring locations. Units: cm s^{-1}

Location	<i>Control</i>	<i>EXP</i>	<i>SIM</i>
0° , 140°W	40.40	26.23	27.37
0° , 110°W	38.97	24.14	30.36
7°N , 140°W	12.14	11.82	12.75
0° , 165°E	42.33	37.71	21.70

Figure 7 shows the vertical distribution of the correlation coefficients and RMS errors between the three runs and the TOGA TAO for their mean annual cycles at 140°W and 110°W. A mean annual cycle is defined as the departure of the climatological monthly mean from the annual mean. The correlation coefficients from the three runs are quite high, around 0.7–0.9 in the upper ocean. In both locations, the correlation coefficients for *EXP* are generally larger than those for *SIM* and *Control*. The RMS errors of *Control* are usually the smallest among the three at most levels. And although *Control* tends to have better annual cycle than *SIM* at 140°W, it is worse in the upper 70 m at 110°W.

As for the anomalies from the annual cycles, *Control* always has much larger RMS errors and lower correlation coefficients than either *EXP* or *SIM* at these two locations (Figs. 8a, b). The *EXP* correlation coefficients are higher than those from *Control* and *SIM* in most of the upper 120 m. The RMS errors in *EXP* are generally comparable to those from *SIM* at 140°W while slightly larger at 110°W. This is because *EXP* represents the phase of the interannual variations quite well but tends to overestimate the magnitude of the anomalies, while *SIM* underestimates them persistently. *Control* has serious errors in reproducing both the phase and the amplitudes.

Figure 9 shows time–depth sections of the zonal currents at 0°, 140°W from the observations, the analyses, and the simulation. It clearly demonstrates the improvement of current variability in *EXP* over that in *Control*. At 140°W, the major EU fluctuation from the TAO measurements is its annual cycle (Fig. 9a). During late boreal spring or early summer, the EU core usually grows in strength to above 100 cm s⁻¹ and shoals, followed by the deepening and weakening of the core over the next few months. The upward movement of the EU core causes the “surfacing” of the eastward flowing water at the surface. Interannually, the EU tends to be weaker during mature El Niño episodes (1986–1987, 1991–1992, and 1997–1998) but stronger during cold episodes and before an El Niño (early 1986, 1988–1989, 1991, 1994, 1996).

A comparison between Fig. 9a and Fig. 9b shows that most of the observed features discussed in last paragraph have been faithfully reproduced by the *EXP* analyses. Fig. 9c, however, shows a quite different picture. The current here is dominated by vertically elongated fluctuations that have not only a larger amplitude but also a different phase from both the TAO measurements and *EXP*. These large interannual fluctuations totally mask the seasonal signals even though our RMS and correlation analyses suggest that the data assimilation does improve the mean annual cycle. The currents from *SIM* are apparently too weak with less variability (Fig. 9d). The results from 110°W are similar (not shown).

At 165°W, *EXP* and *SIM* have comparable annual cycles and are both significantly better than that from *Control*. In fact, the correlations of both *EXP* and *SIM* to the observations are at 0.8 for all levels and their RMS errors smaller than 14 m s⁻¹. *Control*, however, has a minimum correlation of 0.3 around 30–60 m, together with a maximum RMS error of 20 m s⁻¹ there. For the anomalies, the correlation of *SIM* with observations is around 0.7 near the surface, gradually reducing to 0.4 around the thermocline, then increasing to 0.6 at deeper levels. Its RMS error remains smaller than 20 m s⁻¹ at all levels. In this case, the results of both *EXP* and *Control* are worse, even though the former is slightly better than the latter. In fact, *EXP* and *Control* are practically uncorrelated with the observations below 50 m. Their maximum RMS errors approach 30 m s⁻¹ for *EXP* and 40 m s⁻¹ for *Control*. The reason for these large errors in the analyses is not yet clear.

In summary, combining once–daily intermittent analyses with incremental model error correction produces equatorial currents in the central and eastern Pacific Ocean superior to

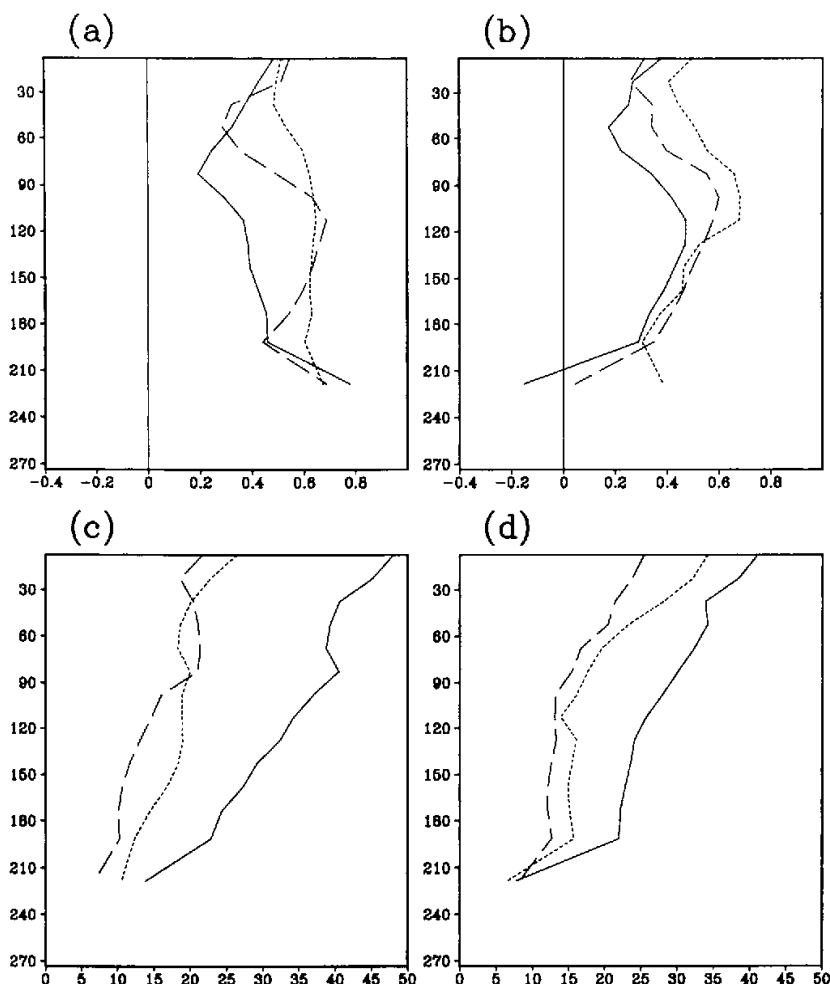


Fig. 8. Correlation coefficients and RMS errors of zonal velocity anomalies at 140°W and 110°W along the equator with respect to TAO current meter measurements. The upper panels are correlation coefficients for (a) 140°W and (b) 110°W . The lower panels are RMS errors for (c) 140°W and (d) 110°W . For each lower panel, the unit of horizontal axis is cm s^{-1} . The vertical axis is depth with unit in meters. The solid curve is from Control, short dashed curve from *EXP*, long dashed curve from *SIM*.

those from the continuous assimilation. Moreover, although inserting temperature data in both ways improves the mean annual cycle of the velocity in the central and eastern ocean, the incremental insertion seems to be able to further improve the phase of the interannual signals. This conclusion is reached based on a comparison with independent TAO mooring data not used in the assimilation process. The improvement of *EXP* over *SIM* suggests that assimilating temperature observations can properly correct ocean currents through model

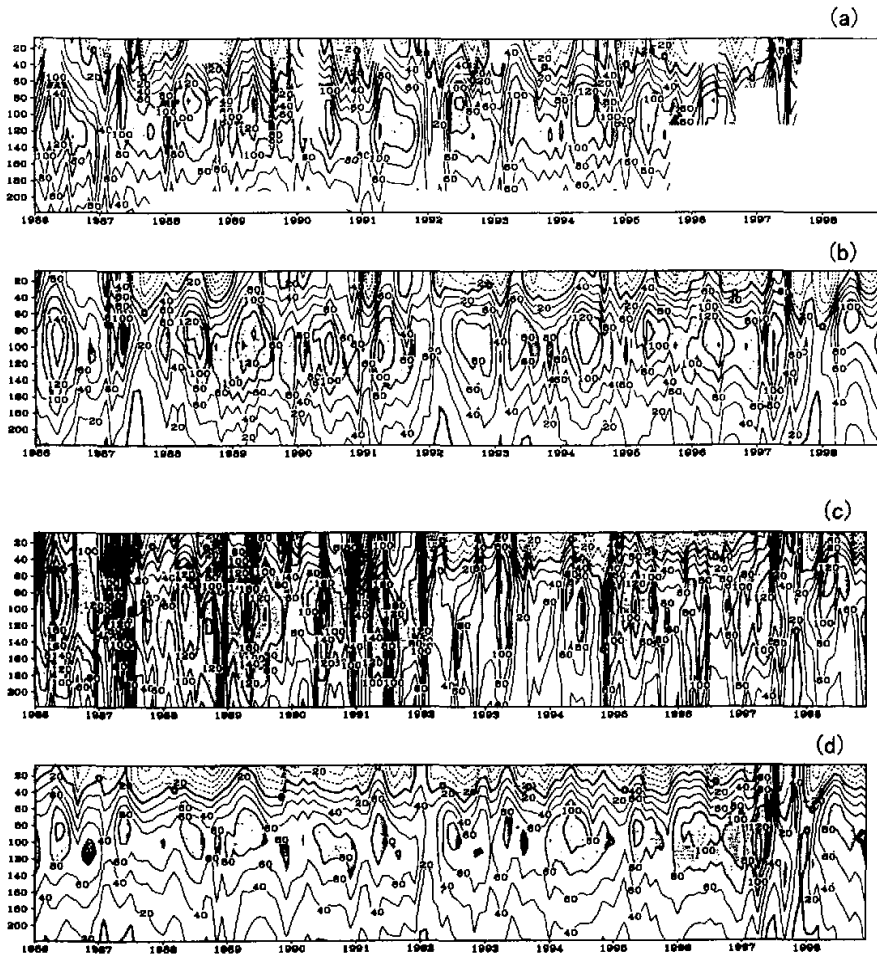


Fig. 9. Zonal velocity time-depth sections at equator, 140°W from (a) TAO current meter, (b) EXP, (c) Control, and (d) SIM. The contour interval is 20 cm s⁻¹. Areas of values greater than 80 cm s⁻¹ have dark shading. Areas of values less than -80 cm s⁻¹ have light shading.

dynamical coupling between variables. On the other hand, correcting model error incrementally rather than instantly seems to preserve the balance between the temperature and velocity.

3.3 Comparison with temperature observations

One should note that the incremental model error correction significantly reduces the contributions of temperature observations in producing the monthly analysis fields in comparison to those from the instantaneous error correction. As we will demonstrate in section 4, it is similar to reducing the relaxation coefficients in a simple nudging assimilation scheme.

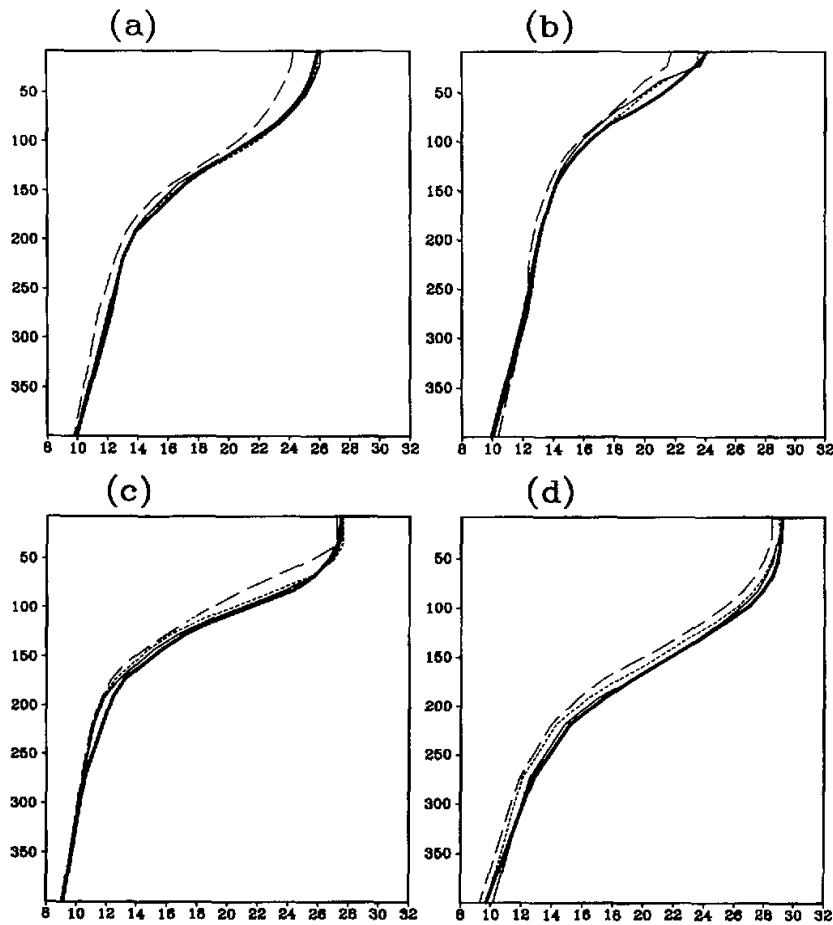


Fig. 10. Mean temperature profiles at (a) Equator, 140°W, (b) Equator, 110°W, (c) 7°N, 140°W, and (d) Equator, 165°E. For each panel, the unit of horizontal axis is °C. The vertical axis is depth with unit in meters. The thick solid curve is from TAO data, thin solid curve from Control, short dashed curve from EXP, long dashed curve from SIM.

Although this incremental scheme seems to actually improve the model equatorial currents, it is still necessary to demonstrate that the reduced relaxation is adequate for the temperature analyses. Therefore, in this subsection, we briefly summarize a comparison of the monthly temperature from the analyses with some observations that have been used during the assimilation processes. Although the observations are not independent, as was true for the current meter data, the comparison measures how effectively the present strategy can guide the model toward the observations.

Table 2 shows the spatial RMS errors of the long-term mean global SST fields (75°S–65°N) from Control, EXP, and SIM. It also gives global-averaged temporal RMS er-

rors and correlation coefficients for the mean annual cycle and the interannual anomaly. The observations used for these calculations are monthly SST analyses averaged from the CPC weekly fields as described in section 2b. Based on the global averages of the RMS errors, *EXP* reduces the model errors by 47% in the mean field, 32% in the annual cycle, and 35% in the interannual variability. These values represent a significant reduction of the model errors, even though they are lower than those from *Control* (69%, 59%, and 55%), as expected from the 1:12 reduction in the times of assimilation. The globally averaged correlation coefficient between the observed and the analyzed SST anomalies for *EXP* is 0.81, which is lower than that from *Control* (0.98) but significantly higher than that from *SIM* (0.33). For the annual cycle, the correlation coefficients from *EXP*, *Control*, and *SIM* are comparable.

Table 2. Global averaged SST RMS errors and correlation coefficients for the long-term mean, mean annual cycle, and anomaly fields from *Control*, *EXP*, and *SIM*. Units: °C for RMS

Global Average	<i>Control</i>	<i>EXP</i>	<i>SIM</i>
Mean RMS Error	0.34	0.58	1.96
Annual RMS	0.25	0.41	0.61
Anomalous RMS	0.25	0.36	0.55
Annual COR	0.99	0.92	0.94
Anomalous COR	0.98	0.81	0.33

Table 3. Mean errors as well as annual and anomalous RMS errors and correlation coefficients (COR) for the NINO3 SST index. Units: °C for RMS

NINO3	<i>Control</i>	<i>EXP</i>	<i>SIM</i>
Mean Error	−0.04	−0.29	−1.73
Annual RMS	0.06	0.10	0.24
Anomalous RMS	0.07	0.21	0.53
Annual COR	0.99	0.99	0.97
Anomalous COR	0.99	0.99	0.87

Spatially, regions with larger model errors usually benefit more from the data insertion. Table 3 gives the error measurements of the NINO3 index, which is the SST averaged within 5°S–5°N, 90°–150°W, a region of relatively large model systematic error. *EXP* and *Control* reduce the mean model error from −1.73°C to −0.29°C and −0.04°C respectively, accounting for a total reduction of 83% and 98%. On the annual and interannual scales, the reductions of the RMS error for both analyses are also about 30% higher than their corresponding global averages.

Although *EXP* exhibits generally larger total SST RMS errors than *Control*, they are still within the prescribed observational error estimate (1°C) for the ODA analysis in most regions. In fact, the global averaged RMS errors for *Control*, *EXP*, and *SIM* are respectively 0.48°C, 0.79°C, and 1.37°C. In open oceans, the RMS error for *EXP* is usually around 0.25°C to 0.5°C while that for *Control* is kept within 0.25°C. Since the observations (the CPC data) used for assimilation are continuous in both time and space, they seem to play quite a dominant role on the analyzed SST over the open ocean. Near the coast, the *EXP* and *Control* SST errors are larger because of unresolvable upwellings and other local processes. Other regions of larger errors are the Kuroshio and Gulf Stream and their extensions as well as regions near

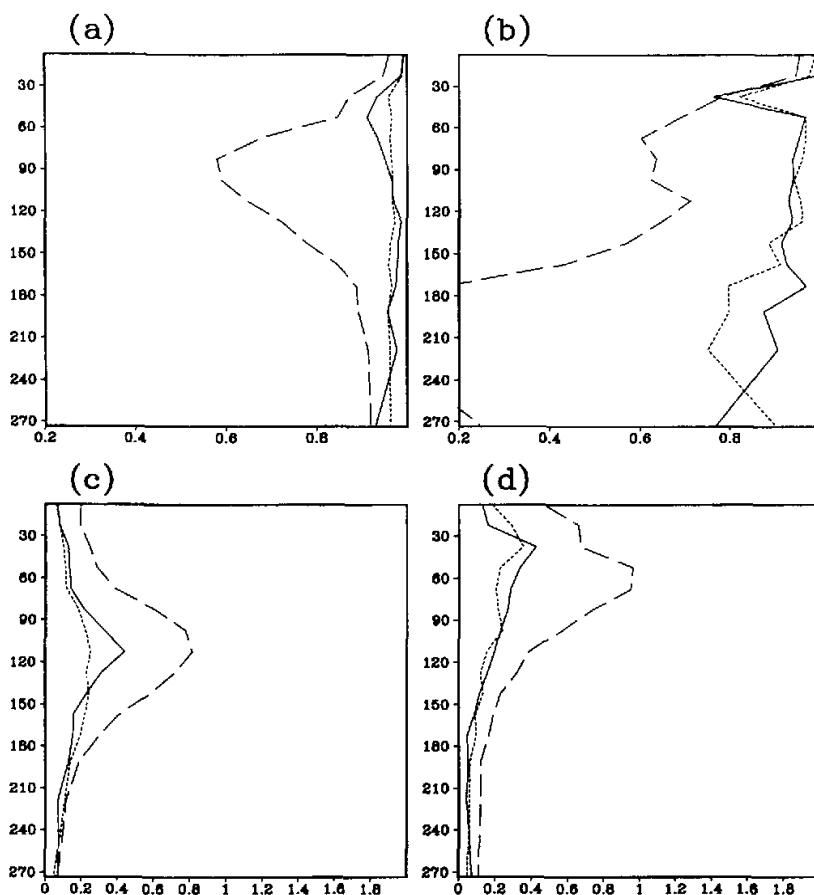


Fig. 11. Correlation coefficients and RMS errors of temperature annual cycle at 140°W and 110°W along the equator with respect to TOGA TAO mooring data. The upper panels are correlation coefficients for (a) 140°W and (b) 110°W. The lower panels are RMS errors for (c) 140°W and (d) 110°W. For each lower panel, the unit of horizontal axis is °C. The vertical axis is depth with unit in meters. The solid curve is from *Control*, short dashed curve from *EXP*, long dashed curve from *SIM*.

the southern tips of both the African and American continents. The *Control* is generally better than *EXP* there. In these regions, the oceanic meso-scale eddies are active and play a significant role in oceanic heat transport. These eddies are not resolvable by the OGCM at the resolution we have used. Therefore, the smaller *Control* SST errors in these regions are solely caused by the much stronger local relaxation of the ODA strategy.

To examine the influence of different ODA strategies on the sub-surface temperature, we use TOGA TAO temperature profiles at the same four mooring sites for the current comparison. Figure 10 shows the mean temperature profiles derived from the TOGA TAO data at

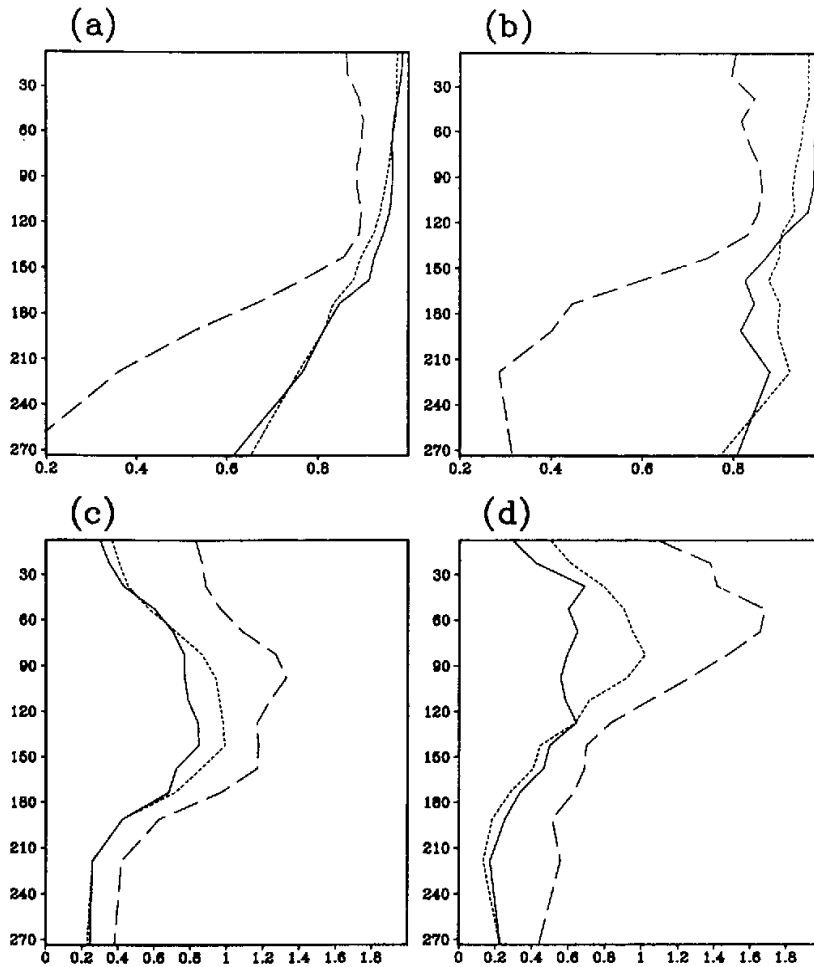


Fig. 12. Correlation coefficients and RMS errors of temperature anomalies at 140°W and 110°W along the equator with respect to TOGA TAO mooring data. The upper panels are correlation coefficients for (a) 140°W and (b) 110°W. The lower panels are RMS errors for (c) 140°W and (d) 110°W. For each lower panel, the unit of horizontal axis is °C. The vertical axis is depth with unit in meters. The solid curve is from *Control*, short dashed curve from *EXP*, long dashed curve from *SIM*.

these four locations, together with corresponding data from *Control*, *EXP*, and *SIM*. The four panels show that both *Control* and *EXP* are efficient at forcing the model profiles toward the observations. The results from *EXP* are as good as those from *Control* in most of the locations, except at 165°E along the equator, where the *EXP* errors are slightly larger than those from *Control* in the deeper ocean around 200–300 m, even so, it still shows an improvement from *SIM*.

Unlike the SST field, where the globally averaged RMS error for *EXP* is 40% higher than the *Control* error, the differences for these TOGA TAO sites are much smaller. For example, the vertically averaged total RMS errors at 140°W on the equator are 0.77°C for *Control* and 0.81°C for *EXP*. At 110°W, they are 0.66°C and 0.76°C respectively. At both sites, they account for a significant reduction from *SIM* (1.72°C and 1.69°C). This seems to suggest that *EXP* is almost as efficient as *Control* in reducing model errors. We also noticed that the RMS errors from both analyses are higher than the prescribed observational error for subsurface data (0.5°C) in the ODA system. These errors are mostly from regions around the thermocline.

As in Figs. 7 and 8, Figs. 11 and 12 present correlation coefficients and RMS errors of the temperatures at 140°W and 110°W along the equator, from *Control*, *EXP*, and *SIM* respectively for the annual cycle and interannual variability. For the annual cycle (Fig. 11), the levels of correlation and RMS error from *EXP* and *Control* are comparable to each other and they show a definite improvement over *SIM*. For the interannual variability (Fig. 12), however, the *EXP* RMS errors tend to be larger than those from *Control* around the depths of the thermocline by 0.2–0.4°C. The errors, however, are much smaller than those in *SIM*.

The systematic differences between *SIM* and TOGA TAO data below the surface are mostly associated with a more diffuse thermocline. Moreover, the time–depth sections at these four locations shows that, in comparison with the TOGA TAO data, *SIM* also has much weaker fluctuations of the thermocline depth, with some observed features being smoothed out or with delayed peaking episodes. The data assimilation is useful in enhancing the amplitude of the fluctuations and correcting their phases. In summary, the comparison to observations shows that, like the *Control* analyses, the *EXP* analyses are efficient in forcing the ocean model toward the observations even though the “rate” of relaxation is much weaker in the latter.

4. Discussion

The results of the previous section indicate that there are substantial differences between the results obtained using the old (*Control*) and new (*EXP*) insertion techniques. These differences need to be explained, since only the temporal nature of the data insertion was different in the two runs—the data and the assumed horizontal covariance functions were identical. Whether the analysis step is done once per hour or once per day should have little impact on the correction field. We believe the more fundamental difference arises from the assumptions made about the errors. Their effects on model adjustment may account for the changed state of the system.

For the following discussion, it will suffice to consider the model and observation at a single point, and to consider the state as represented by a single variable T , which is observed. The observations and model are both assumed to vary slowly in relation to the one-day analysis interval.

At the start of a day, we denote the difference between the model state estimator and the observation as T_f . (If we assume that the observation is slowly varying, we may set the observation T_o to zero). The analysis technique amounts to choosing an analysis value T that minimizes the cost function

$$J = (T - T_f)^2 E^{-1} + T^2 F^{-1}, \quad (2)$$

where E is the forecast error covariance estimate and F is the observational error covariance estimate, as defined in (1). This has the solution

$$T = \frac{T_f E^{-1}}{E^{-1} + F^{-1}}. \quad (3)$$

The increment applied to the model estimate to obtain the analysis is

$$\Delta T = T - T_f = - \frac{T_f F^{-1}}{E^{-1} + F^{-1}}. \quad (4)$$

In the *EXP* technique, we do the analysis once per day and apply

$$\delta T = \Delta T / 24. \quad (5)$$

In the *Control* technique, the analysis is done every time step for the first 12 hours, then the last increment is held steady for the next 12 hours. If we assume that the observations are unchanging over the day, and that the model does not cause any significant change in the state over one day, the repeated application of the *Control* analysis technique will result in

$$T_n = \begin{cases} T_f (1 - \varepsilon)^n, & n = 1, \dots, 12 \\ T_{12} + (n - 12)(T_{12} - T_{11}), & n = 13, \dots, 24, \end{cases} \quad (6)$$

where T_n is the analyzed state at the n -th hour of the day and

$$\varepsilon = \frac{F^{-1}}{E^{-1} + F^{-1}}. \quad (7)$$

We have assumed that the model error covariance is much smaller than the observational error covariance, with a ratio of 25 to 1. The term ε has the value $1/26$, which leads to about a 4% change in the model estimated state per analysis. Doing the analysis once per day results in this 4% change over the day, while the *Control* technique leads to a reduction in the difference between the estimate and the observation by about 60%.

We see that the *Control* technique results in a far stronger relaxation to the observations than the *EXP* technique. This stronger constraint is at the core of the difference in the model behavior. Over 24 hours, the *Control* technique results in a change

$$T_{24} = T_f (1 - \varepsilon)^{11} (1 - 13\varepsilon), \quad (8)$$

or when ε is small

$$T_{24} = T_f (1 - 24\varepsilon + 198\varepsilon^2 + O(\varepsilon^3)). \quad (9)$$

We see that for the *Control* technique to have the same effect as the *EXP* technique, it would be necessary to assume that the model error covariance is roughly 24 times smaller than that assumed for the *EXP* case. Figure 13 shows the estimate of the state after 1 day as a function of the ratio of observation error to model error for the *Control* technique, the *EXP* technique, and a modified *Control* technique. In the modified *Control* technique, we assume that the model error is 24 times smaller.

For the cases where the model error is smaller than the observational error, the modified *Control* technique provides a much closer estimate to *EXP*. Only when the model error is large (or the observational error is small) do the techniques converge by bringing the state estimate quickly to the observed state. The negative values seen in Fig. 13 for the

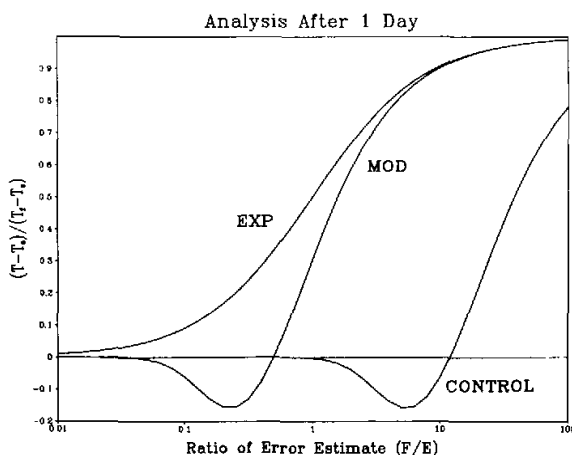


Fig. 13. Analysis state estimate at the end of 1 day as a function of the assumed ratio of observational to model error. The curve labeled MOD is the result when the *Control* technique is used with an assumed model error $1/25$ th as large.

Control technique result from overshooting the correction by applying the 12th hour correction for the second half of the day.

As a point of comparison, we note that Behringer et al. (1998) improved the NCEP ODA system by employing the *Control* technique with a much-reduced estimate of the model error covariance. The improvements in their system are similar to the ones encountered here, and indicate that the above reasoning is likely correct.

5. Summary

A new data-insertion strategy is applied to an originally continuous ODA system developed by Derber and Rosati (1989). Utilizing the same analysis component as that of the original system, the new system conducts analyses to derive model corrections at once-daily intervals, instead of at every time step. This system preserves major features of a continuous assimilation procedure due to some other characteristics of the procedure. First, each analysis is conducted with observations within a 10-day RI window, which is much larger than the once-daily analysis interval, so that variations among model corrections derived from adjacent analyses are small and continuous. Second, the correction field derived from each analysis is used as a forcing to correct the model within each day, following the Incremental Analysis Update (IAU) method of Bloom et al. (1996). In this way, the ocean model is driven continuously by a forcing derived from observations during its integration. Moreover, the IAU process incorporates the corrections into the model in a gradual manner that does not significantly disrupt the model internal dynamical development.

A 13-year global ocean analysis from 1986 to 1998 is conducted using this new ODA system. It is compared with a similar analysis based on the original Derber and Rosati system. An examination of both analyses in the tropical Pacific Ocean shows that they have qualitatively similar annual and interannual variability of the temperature field. However, the

new analysis produces smoother monthly analyses for both temperature and velocity fields and its tropical currents show significantly different features from the original analyses. Compared to the observations from TAO current meters, it is shown that the equatorial currents from the new analyses are significantly improved relative to the original analyses, not only in maintaining the mean state but also in capturing the annual and interannual variations.

Measured against some SST and sub-surface temperature data that are used in the assimilation process, the new analyses are efficient in forcing the model toward available observations, though the weight of observations is effectively reduced in producing the analyses. This is more apparent in the subsurface region where the observations are less dense and model internal processes play a more important role.

Applying both strategies analytically to a simple one-dimensional system without dynamics demonstrates that the new strategy is equivalent to a substantial reduction of the estimated model error, leading to a weaker relaxation toward observations in the assimilation process. It is a delicate issue to preserve a proper balance between the contributions from the observations and the dynamical model in producing an analysis field. If the effects of the observations are too weak, they may not be enough to correct errors in the model. However, overly strong relaxation toward observations may dominate or even disrupt the dynamical processes within the model that are important in determining variables that are not directly observed and in propagating observational information. Therefore, ideally, the short-term tendency should come mostly from the model, with the analysis applying a small correction (MacPherson 1991). For a complicated OGCM, the measure of whether or not a proper balance is maintained may be derived from the quality of its product only. The better equatorial currents in the central and eastern Pacific Ocean in *EXP* seem to suggest a more appropriate balance has been achieved.

In this sense, our results are largely consistent with those derived recently by Behringer et al. (1998). As an effort to improve the NCEP ODA system, they significantly reduced the magnitude of the estimated first-guess error, thus increasing the contribution of the model to their analyses. They have reported less noisy analyses that seem to be more favorably compared with independent observations. We achieve similar results by reducing the contributions of observations in an alternative way. One advantage of our approach, however, is that it also reduces the computational burden of the analysis procedure by reducing the number of times for solving the minimization problem iteratively. Comparing our two analyses, *EXP* reduces computational cost for the minimization part by a factor of 12 and saves two-thirds of the time for a monthly analysis run.

Our analysis concentrates on the Pacific Ocean because data are most easily available there and because our first priority is to use these analyses for ENSO studies. However, some preliminary study of the Atlantic Ocean analyses also shows similar results. The effects of this new strategy on the variability of the extra-tropical oceans have not been fully assessed, although it seems to produce a reasonable mean state there. The ultimate test will come from using these analyses to initialize coupled models for climate predictions. These issues will be addressed in further studies.

Acknowledgments. We are indebted to many people and institutions; without their help this study could not have been accomplished. We thank Dr. J. Shukla for his continuous encouragement and helpful suggestions during the course of this research. We are grateful to A. Rosati of the Geophysical Fluid Dynamics Laboratory for generously providing the GFDL data assimilation system and to N. Pinardi of Istituto per lo Metodologie Geofisiche

Ambientale for helping us implement the system at COLA. We thank S. Levitus of the National Oceanographic Data Center, A. Leetmaa, M. Ji, and R. Reynolds of the National Centers for Environmental Prediction, and M. McPhaden of the Pacific Marine Environmental Laboratory for providing the observational data sets. We also thank D. Collins of NODC for providing us monthly real-time in situ ocean measurements. This research was supported by the National Science Foundation (ATM-9321354) and the National Oceanic and Atmospheric Administration (NA46-GP0217), and computing resources were provided by the National Center for Atmospheric Research Scientific Computing Division.

REFERENCES

- Behringer, D. W., M. Ji, and A. Leetmaa, 1998: An improved coupled model for ENSO prediction and implications for ocean initialization. Part I: The ocean data assimilation system. *Mon. Wea. Rev.*, **126**, 1013–1021.
- Bloom, S. C., L. L. Takacs, A. M. daSilva, and D. Ledvina, 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.
- Carton, J. A., G. Chepurin, X. Cao, and B. Giese, 2000a: A simple ocean data assimilation analysis of the global upper ocean 1950–1995. Part I: Methodology. *J. Phys. Oceanogr.*, **30**, 294–309.
- Carton, J. A., G. Chepurin, X. Cao, 2000b: A simple ocean data assimilation analysis of the global upper ocean 1950–1995. Part II: Results. *J. Phys. Oceanogr.*, **30**, 311–326.
- Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347.
- Hamilton, D., 1994: GTSP builds an ocean temperature–salinity database. *Earth System Monitor* 4 / 4, June 1994.
- Hayes, S. P., L. J. Mangum, J. Picaut, A. Sumi, and K. Takeuchi, 1991: TOGA–TAO: A moored array for real-time measurements in the tropical Pacific Ocean. *Bull. Amer. Meteor. Soc.*, **72**, 339–347.
- Huang, B., and J. L. Kinter III, 1997: A global ocean data analysis for 1986–1992. COLA Technical Report 38, 41 pp. plus 21 figures.
- Huang, B., and J. Shukla, 1996: A comparison of two surface wind stress analyses over the tropical Atlantic during 1980–1997. *J. Climate*, **9**, 906–927.
- Huang, B., and E. K. Schneider, 1995: The response of an ocean general circulation model to surface wind stress produced by an atmospheric general circulation model. *Mon. Wea. Rev.*, **123**, 3059–3085.
- Ji, M., A. Leetmaa, and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies. *Mon. Wea. Rev.*, **123**, 460–481.
- Ji, M., A. Leetmaa, and V. E. Kousky, 1996: Coupled model predictions of ENSO during the 1980s and the 1990s at the National Centers for Environmental Prediction. *J. Climate*, **9**, 3105–3120.
- Levitus, S., 1982: Climatological Atlas of the World Ocean. NOAA Professional Paper 13 (U.S. Government Printing Office Washington, DC), 17 fiches, 173pp.
- Levitus, S., and T. P. Boyer, 1994: World Ocean Atlas 1994. Volume 4: Temperature. NOAA Atlas NESDIS 3. U.S. Department of Commerce, Washington D. C., 117pp.
- McCarty, M. E., and M. J. McPhaden, 1993: Mean Searoual cycles and interannual variations at 0°, 165°E during 1986–1992. NOAA Tech. Memo., ERL PMEL–98, 64pp.
- McPhaden, M. J., A. J. Busalacchi, R. Cheney, J. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G.T. Mitchum, P. P. Niiler, J. Picaut, R. W. Reynolds, N. Smith, and K. Takeuchi, 1998: The Tropical Ocean–Global Atmosphere observing system: A decade of progress. *J. Geophys. Res.*, **103**, 14,169–14,240.
- McPhaden, M. J., and M. E. McCarty, 1992: Mean seasonal cycles and interannual variations at 0°, 110°W and 0°, 140°W during 1980–1991. NOAA Tech. Memo., ERL PMEL–95, U.S. Dept. of Commerce, Washington, D.C., 118pp.
- MacPherson, B., 1991: Dynamic initialization by repeated insertion of data. *Quart. J. Roy. Meteor. Soc.*, **117**, 965–991.
- Oberhuber, J. M., 1988: An atlas based on the COADS data set: Budgets of heat buoyancy and turbulent kinetic energy at the surface of the global ocean. Report 15, Max–Planck–Inst. für Meteor. (Hamburg, Germany), 199pp.

- Pacanowski, R. C., K. Dixon, and A. Rosati, 1993: The GFDL Modular Ocean Model user Guide, Version 1.0. *GFDL Ocean Group Tech. Rep.*, No.2.
- Pacanowski, R. C., and S. G. H. Philander, 1981: Parameterization of vertical mixing in numerical models of tropical oceans. *J. Phys. Oceanogr.*, **11**, 1443–1451.
- Philander, S. G. H., W. J. Hurlin, and A. D. Seigel, 1987: A model of the seasonal cycle in the tropical Pacific Ocean. *J. Phys. Oceanogr.*, **117**, 1986–2002.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948.
- Rosati, A., R. Gudgel, and K. Miyakoda, 1995: Decadal analysis from an ocean data assimilation system. *Mon. Wea. Rev.*, **123**, 2206–2228.
- Rosati, A., and K. Miyakoda, 1988: A general circulation model for upper ocean circulation. *J. Phys. Oceanogr.*, **18**, 1601–1626.
- Schneider, E. K., B. Huang, Z. Zhu, D. G. DeWitt, J. L. Kinter III, B. P. Kirtman, and J. Shukla, 1999: Ocean data assimilation, initialization, and predictions of ENSO with a coupled GCM. *Mon. Wea. Rev.*, **127**, 1187–1207.
- Trenberth, K. E., 1998: Development and forecasts of the 1997 / 98 El Niño: CLIVAR scientific issues. *CLIVAR Exchanges* 3, No.2 / 3, 4–14.
- Yu, Z., and P. S. Schopf, 1997: Vertical eddy mixing in the tropical upper ocean: Its influence on zonal currents. *J. Phys. Oceanogr.*, **27**, 1447–1458.
-