A Global Oceanic Data Assimilation System

JOHN DERBER AND ANTHONY ROSATI

Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, New Jersey

(Manuscript received 26 September 1988, in final form 6 February 1989)

ABSTRACT

A global oceanic four-dimensional data assimilation system has been developed for use in initializing coupled ocean–atmosphere general circulation models and many other applications. The data assimilation system uses a high resolution global ocean model to extrapolate the information forward in time. The data inserted into the model currently consists only of conventional sea surface temperature observations and vertical temperature profiles. The data are inserted continuously into the model by updating the model's temperature solution every timestep. This update is created using a statistical interpolation routine applied to all data in a 30-day window centered on the present timestep.

Large scale features in the sea surface temperature analyses are consistent with those from independent analyses. Subsurface fields created from the assimilation are much more realistic than those produced without the insertion of data. Furthermore, information contained in the assimilation field is shown to be retained in the model solution after the assimilation procedure is terminated. The results are encouraging but further improvements can be made.

1. Introduction

A persistent problem in oceanography has been the lack of a complete and reasonably accurate description of the large-scale oceanic state. Without such an analysis, many potential studies become difficult or impossible. The applications of the analyses include the initialization of coupled atmospheric and oceanic General Circulation Models (GCMs), the verification of numerical models, dynamical and climatological studies, and the specification of horizontal boundary conditions for higher resolution limited domain models. Unfortunately, the ocean is not observed frequently enough in space or time to allow the direct analysis of the predicted variables. This limitation has been overcome in the past by assuming that the interannual variability is small and then compositing observations from many years to produce climatological analyses (e.g., Levitus 1982). If one is interested in components with strong interannual variability, however, an alternative analysis technique is necessary.

A possible solution is to use a four-dimensional data assimilation system. The basic idea of four-dimensional data assimilation is to use model equations to extrapolate information in space and time. Thus, the effective database at any particular time can be increased substantially and regions which are not observed regularly can be analyzed by use of the information extrapolated into the region by the model equations. Also, in regions where no information exists, a reasonable result (the model solution) is available. The four-dimensional data assimilation technique has been used successfully in meteorology to address a data sparcity problem similar to that found in the ocean (e.g., Bengtsson et al. 1981; Bengtsson 1975).

Four-dimensional data assimilation was first applied to oceanic problems using simple limited domain ocean models (e.g., Robinson and Leslie 1985; Rienecker et al. 1987). Recently, there has been considerable interest in applying data assimilation to larger scales using more realistic models. Clancy et al. (1988) have developed a general ocean analysis system which can be applied to larger scales but uses a simple mixed-layer model. Moore et al. (1987) have investigated the effects of inserting simulated temperature and velocity data into a one-layer reduced-gravity model and 12-level primitive equation model. Leetmaa and Ji (1989) have been producing analyses of the tropical Pacific by inserting observations into a primitive equation model. In this paper, a nearly global oceanic assimilation system using a high-resolution primitive equation model will be presented.

Several new and theoretically appealing data assimilation techniques, such as Kalman–Bucy filtering (e.g., Bennett and Budgell 1987; and Miller 1985) and the adjoint technique (e.g., LeDimet and Talagrand 1986; Thacker and Long 1987), are currently being investigated for application to the oceanic problem. These newer and more complex techniques, however, appear to be currently impractical for application to a high resolution global oceanic GCM. Thus in this study we

Corresponding author address: Dr. John Derber, NOAA/NWS, World Weather Building, Room 204, Washington D.C. 20233.

have chosen a rather straightforward technique to perform the data assimilation.

Any data assimilation system can be described in terms of three components, the dynamical model, the data and quality control procedures, and the analysis and insertion technique. The assimilating model, described in section 2, is a modification of the model used by Rosati and Miyakoda (1988, hereafter RM). This model has a horizontal resolution of 1° × 1° except between 10°N and 10°S where the latitudinal resolution is increased to ⅓°. The vertical resolution has been increased to 15 levels from the 12 levels used in RM. The surface boundary conditions have been provided using the Geophysical Fluid Dynamics Laboratory's (GFDL) FGGE IIIb 12-hourly analyses (Ploshay et al. 1983). In this study, the oceanic data (section 3) have been limited to conventional surface temperature data and vertical temperature profiles. A considerable number of these data contained erroneous values, requiring extensive quality control on the data. The resulting dataset was used in a spatial analysis routine to create correction fields to be inserted into the forecast model (section 4). This insertion was done continuously with a correction being created every timestep using data in a 30-day window around the current timestep. Results from the assimilation systems are presented in section 5 for a period within the FGGE year. In the final section, a summary of the results and a discussion of future improvements to the system will be presented.

2. The numerical model

The oceanic numerical model used in this study was developed primarily for use in a coupled ocean–atmosphere GCM to produce seasonal forecasts. Since only the upper levels of the ocean are likely to affect the atmosphere over this forecast interval, the modeling effort was concentrated on the accurate forecasting of the upper ocean given the atmospheric state. With these considerations in mind, the Bryan (1969) and Cox (1984) ocean model was modified by RM to produce an upper-ocean GCM. The major changes to the RM model for this study involved changes in the calculation of the surface fluxes, the addition of three vertical levels to allow better resolution of the thermocline (Fig. 1) and a change in the vertical mixing routine. The remainder of this section will review the RM model and discuss the changes made to the model for use in the assimilation system.

The dynamics of the ocean model are governed by the primitive equations under the hydrostatic, Boussinesq, and rigid lid approximations. These equations are approximated using a second-order energetically consistent formulation (Bryan 1969; Cox 1984). At land–ocean interfaces a "no-slip" boundary condition was used. The model equations were solved on a nearly global grid (excluding the Arctic Ocean) with realistic topography. In the horizontal, a staggered 1° × 1° latitude–longitude grid was used except between 10°N–10°S where the north–south resolution was increased to ⅓° in order to resolve finer scale equatorial structures. As mentioned earlier, the vertical resolution (Fig. 1) was increased in these experiments over those used in RM to better resolve the thermocline. This was accomplished by adding three additional levels (54 m, 100 m, and 177 m).

Because of the finite resolution of the ocean model, several subgrid scale processes must be parameterized. The most important of these processes are the horizontal and vertical mixing. The model was designed to perform vertical mixing using either constant mixing coefficients or a level 2.5 turbulence closure scheme (Mellor and Yamada 1982). Rosati and Miyakoda have shown considerable improvement of the predicted mixed layer depth using the turbulence closure scheme and thus this scheme was chosen for use in this study. For the horizontal mixing, the model is designed to incorporate either a constant mixing coefficient scheme or a nonlinear coefficient scheme based on deformation fields (Smagorinsky et al. 1965). For the assimilation

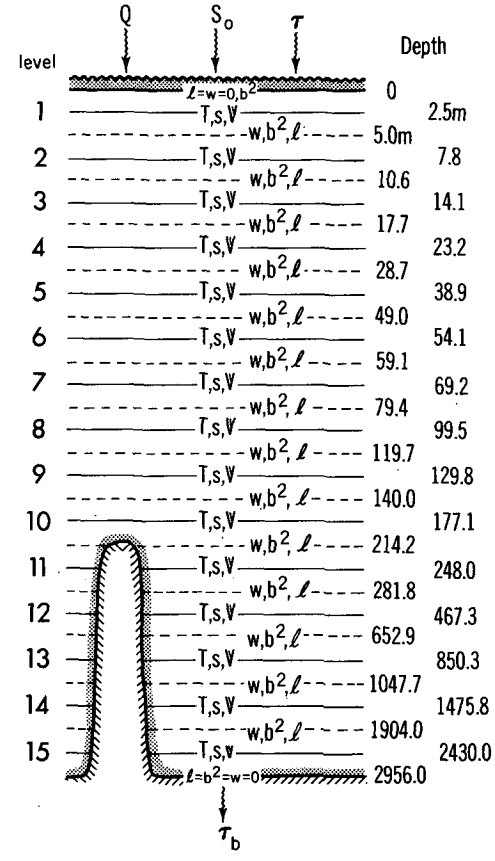

FIG. 1. A schematic diagram of the model's vertical structure; Q and S_0 represent the surface heat and salinity fluxes, τ is the wind stress, and τ_b is the bottom stress.

process, reasonably smooth fields are desirable. Using the version of the model described in RM, we were unable to produce fields of sufficient smoothness with the nonlinear scheme, and therefore it appeared necessary to use the constant coefficient scheme with a large horizontal mixing coefficient ($A_H = 10^8\ \mathrm{cm^2\ s^{-1}}$) for the assimilation. However, in an attempt to reduce model bias involving an unrealistically diffuse thermocline, a limitation on the turbulent length scale (l) in the Mellor–Yamada vertical mixing routine was incorporated. This length scale limitation substantially decreases the mixing through the thermocline by putting an upper bound on the mixing length that depends on the local static stability. The specific limitation we have incorporated is based on Galperin et al. (1988) and is given by

$$l_{\max} = 0.5b \Big/ \left(\frac{g}{\rho_0}\frac{\partial \rho}{\partial z}\right)^{1/2},$$

where b^2 is twice the turbulent kinetic energy, ρ_0 is a mean density, g is the gravitational constant and $\partial\rho/\partial z$ is the vertical density gradient. Note that this limitation is applied only when stable stratification exists.

The inclusion of this vertical mixing length limitation produced the unexpected dividend of substantially smoother horizontal fields. This allowed the use of the nonlinear horizontal viscosity scheme for the horizontal diffusion in the assimilation procedure. The specific technique for the inclusion of the vertical stability in the vertical mixing is an ongoing research topic and further changes can be expected.

The effects of ice formation and melting are parameterized in a very simple manner. The formation of ice provides a lower limit on the attainable oceanic temperatures and thus the temperatures are limited to be greater than −2°C. The surface salinity field is relaxed back to the climatological values of Levitus (1982) to parameterize the effects of ice formation and melting along with the effects of evaporation and precipitation on the salinity field. Finally, in an attempt to ensure reasonable polar oceanic temperatures, the sea surface temperatures (SST) are relaxed back to the Levitus climatology values polewards of 55°.

The heat and momentum fluxes across the ocean's surface play a substantial role in determining the state of the upper ocean. For this reason, the accurate specification of these quantities is extremely important. A portion of the heat flux is determined by the short- and longwave radiation. As in RM, the longwave radiation is empirically corrected for cloudiness and air–sea temperature differences and the shortwave radiation is allowed to penetrate through the surface layer. In this version of the model, however, a diurnal cycle has been incorporated with the hope that enhanced nighttime vertical mixing would improve the prediction of the equatorial mixed layers. Only a very slight positive impact was noted.

Several changes in the bulk aerodynamic formula for the calculation of wind stress and the latent and sensible heat fluxes have been made for this study. First, a reduction of the drag coefficient from 1.5×10^{-3} to 1.4×10^{-3} used in the calculation of the wind stress was implemented. Then, the surface wind fields used in the bulk formula were modified, as suggested by Pacanowski (1987), to reflect the motion of the underlying surface. Thus, the velocity difference between the surface wind and the surface current was used in the bulk formula. The atmospheric fields of temperatures, winds, and relative humidities used in the bulk formula were taken from the 12-hourly GFDL FGGE IIIb analyses (Ploshay et al. 1983). The changes to the surface forcings have reduced (but not eliminated) the bias involving higher than observed upwelling in the eastern Pacific noted by RM.

3. The data and quality control procedures

The data type, accuracy and distribution in space and time govern the quality of the results produced by the data assimilation system. If all predicted variables were observed perfectly and continuously in space and time there would be little need for a data assimilation system. Unfortunately, this is not the case. The observational system typically contains large regions with no observations, has observations with substantial errors and does not measure all of the forecast variables. Eventually, we would like to use as many types of observations (i.e., current measurements, ships drift data, satellite altimetry, salinity, etc.) as possible. In these initial experiments, however, we have chosen to use only conventional temperature observations. These observations were chosen primarily because of their availability, spatial coverage, and quality.

Moore et al. (1987) have presented evidence that the assimilation of density data as opposed to velocity data will be more successful globally for the level of horizontal diffusion used in this study. To update the density it would be necessary to assimilate both the temperature and salinity fields. Unfortunately, salinity observations are not currently available in sufficient quantity for assimilation. This could have important effects in the tropics and extratropics as discussed by Cooper (1988). At present little can be done except hope that the surface boundary conditions and model dynamics are sufficient to produce reasonable salinity fields.

The conventional temperature data used here consisted of SSTs from the Comprehensive Ocean–Atmosphere Data Set (COADS, Woodruff et al. 1987) and vertical temperature profiles obtained from the National Oceanic Data Center (NODC) and the U.S. Navy's Master Oceanic Observation Data Set (MOODS, Bauer 1985). The SST data were monthly averaged data within 2° × 2° latitude–longitude boxes. Within the boxes, all the observations

during the month were averaged producing a mean temperature, location, observation time, and an estimate of the variance. A typical monthly distribution (December 1979) is shown in Fig. 2a. The surface data coverage is quite good in the Northern Hemisphere and tropics. In the Southern Hemisphere extratropics, large gaps in the data still exist. The inclusion of SSTs inferred from satellite measurements may overcome this problem (Reynolds 1988).

The vertical temperature profiles consist primarily of expendable bathythermograph (XBT) measurements. However, a few measurements from Nansen casts, conductivity-temperature-depth (CTD) instruments, and salinity-temperature-depth (STD) instruments were included in the dataset. The spatial coverage of the vertical temperature profiles from the NODC First Global GARP Experiment (FGGE) dataset and the U.S. Navy's MOODS dataset was much

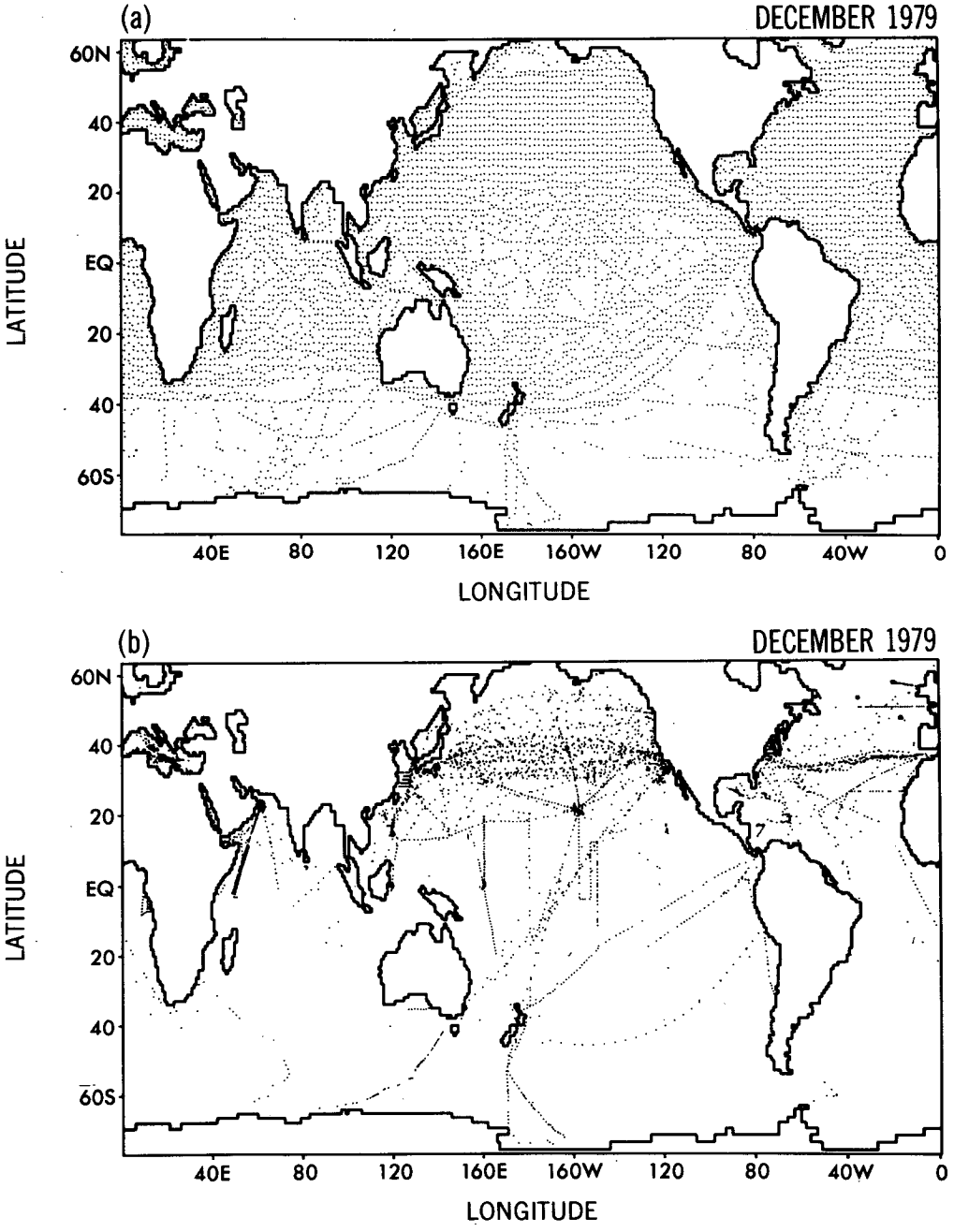

FIG. 2. Observation locations during December 1979. The COADS sea surface temperature 2° × 2° box data are shown in (a), while (b) represents the vertical temperature profile data from the MOODS and NODC datasets.

less complete than the surface coverage; however, the MOODS dataset's coverage was much better than the NODC FGGE dataset. Most of the observations contained in the NODC FGGE dataset are also in the MOODS dataset. Since the NODC FGGE dataset contained a few additional profiles in data sparse regions (i.e., along the east coast of Africa), both of the profile datasets were merged for use in the data assimilation. The locations of the December 1979 profiles for the combined datasets are shown in Fig. 2b.

The SSTs and temperature profiles contained observations with large errors and it was necessary to perform several quality control checks on the data. The initial gross checking included ensuring that the data was not over land, (as defined by the model topography), the temperature was between −5° and 35°C, the depth reported in the soundings increased monotonically, did not contain unrealistic inversions, and was not a duplicate of another profile in the dataset.

After these gross checks, the profiles were vertically interpolated to the top 11 model levels. No information was inserted below level 11 (248 m) because of a lack of data and since our interest lies primarily in the upper ocean. This interpolated data was then compared to the 1° × 1° climatological analysis produced by Levitus (1982). Within each box, an estimate of the variability of the temperature was calculated from the standard deviation of the observations in 5° boxes, as calculated by Levitus, with a minimum variability of 1°C in well-observed regions and 1.5°C in poorly observed regions. When the interpolated observation deviated from the climatological values by more than three times the estimate of the variability, the observation was eliminated from the dataset.

A final quality control procedure was performed within the data insertion process. This quality check involved the comparison of the current model solution to the observations. This final procedure will be described in more detail in the next section along with the rest of the analysis and insertion procedures.

4. Data analysis and insertion technique

The observed data were inserted into the numerical model by applying a correction to the forecast temperature field at every model timestep (every 2 hours). The correction field was created using data from 15 days to either side of the present timestep in a statistical objective analysis scheme. This 30-day window was chosen so that enough data is included for the objective analysis scheme to produce reasonable fields without including so much data that important higher frequency signals in the data are damped. By inserting the data continuously, the velocity field is allowed to adjust to the density field, and no problems resulting from an imbalance between the velocity and density fields have been noted except perhaps in the equatorial region.

The spatial objective analysis technique is based on the statistical interpolation analysis scheme of Gandin (1963). However, it is solved as an equivalent variational problem (for example, see Lorenc 1986). The variational formulation was chosen primarily because it eliminates the costly and somewhat arbitrary data selection procedure contained in more traditional statistical analysis schemes. Thus all observations are used in the analysis of every grid point. Of course, those observations that are far away from a grid point are given minuscule weight. In the variational formulation, it is also easy to account for the model grid to observation location interpolation and it also facilitates the inclusion of indirect observation of the model variables (e.g., the inclusion of surface elevation observations in a model based on the prediction of U, V, T and S).

The functional to be minimized in the variational problem (Lorenc 1986) is given by

$$I = \frac{1}{2}\mathbf{T}^{\mathrm{T}}\mathbf{E}^{-1}\mathbf{T} + \frac{1}{2}(\mathbf{D}(\mathbf{T}) - \mathbf{T}_0)^{\mathrm{T}}\mathbf{F}^{-1}(\mathbf{D}(\mathbf{T}) - \mathbf{T}_0), \tag{1}$$

where $\mathbf{T}$ is an N component vector containing the correction to the guess temperature field, $\mathbf{E}$ is an estimate of the $N \times N$ first guess error covariance matrix, $\mathbf{T}_0$ is an M component vector containing the difference between the observations and the interpolated guess temperature field, $\mathbf{D}$ is a simple bilinear interpolation operator from the grid to the observation locations and $\mathbf{F}$ is an estimate of the $M \times M$ observational error covariance matrix. The N and M denote the number of grid points and the number of observations respectively, while $(\ )^{\mathrm{T}}$ represents the transpose of a vector. The first term of the functional ($\mathbf{T}^{\mathrm{T}}\mathbf{E}^{-1}\mathbf{T}$) is a measure of the fit of the corrected temperature field to the guess field, while the second term measures the fit of the corrected temperature field to the observations. Thus, the resultant analysis is a weighted average of the guess field (which contains information from earlier periods) and the observations.

The statistics contained in the first guess error covariance matrix ($\mathbf{E}$) and the observational error covariance matrix ($\mathbf{F}$) determine the spatial structure and amplitude of the correction field. Unfortunately these statistics are not well known and thus, are now defined empirically. As these statistics become better defined, the results, shown in the next section, should also improve.

Currently, the $\mathbf{E}$ matrix is defined so that the vertical correlations are ignored and the spatial correlations are assumed to be the same for each model level. The horizontal covariances for each level are modeled by repeated applications of a Laplacian smoother. This results in a horizontal covariance between any two points given approximately by

$$ae^{-r^2/(b^2\cos\phi)}, \tag{2}$$

where $a = 0.01$, $b = 570$ km and ϕ is the latitude of the grid point. The $b^2 \cos\phi$ factor controls the filtering of the analysis scheme. The $\cos\phi$ term allows for smaller scale features in the corrections at higher latitudes. The first-guess error variance (a) is smaller than one might expect because the data is inserted continuously over a 30-day window. Away from the equator, r is defined to be equal to the distance between the two grid points. Near the equator, the east–west distance entering into the calculation of r is decreased by a factor of 2.28. This results in an anisotropic covariance function near the equator, with the function stretched parallel to the equator by a factor of 2.28. This anisotropic covariance is included to account for the well-known longer east–west correlation scales near the equator. Note that the matrix $\mathbf{E}$, rather than the matrix $\mathbf{E}^{-1}$ as used in (1), has been defined. In the solution algorithm to be described below, only the $\mathbf{E}$ matrix is necessary.

The observational error covariance matrix ($\mathbf{F}$) should be estimated not only from measurement errors but also from errors of representativeness (e.g., Lorenc 1986). For this reason, it is difficult to estimate the $\mathbf{F}$ matrix. Currently, the $\mathbf{F}$ matrix is assumed to be diagonal, i.e., the observational errors are not correlated. If the correlations were known, the inclusion of local correlations (e.g., assuming the errors are correlated in the vertical) would not be difficult; but nonlocal correlations (e.g., correlations over large spatial scales) would be more difficult to include in the solution algorithm used here.

Since $\mathbf{F}$ is a diagonal matrix, the inversion of $\mathbf{F}$ is trivial and $\mathbf{F}^{-1}$ will be directly defined. The initial estimates of the diagonal elements of $\mathbf{F}^{-1}$ are set equal to the reciprocal of an estimate of the observational error variance. This variance estimate is taken from the COADS estimate for the SSTs. For the temperature profiles the variance is set equal to $(0.25°C)^2$ except in the northwestern Pacific and Atlantic. In these regions there is a propensity to observe eddies not resolvable by this model. Thus, the assumed variance was set equal to $(1°C)^2$ in these regions. The diagonal elements are then multiplied by a time factor which increases linearly from zero to one and back to zero as the difference between the observation time and the model solution time goes from -15 days to zero to $+15$ days. The observations are given no weight when the time difference is greater than 15 days. The inclusion of the time factor allows the use of 30 days of observations in the analysis scheme, yet gives the observations closest to the present model timestep more weight. This time weighting factor contributes to the lack of balancing problems between the density and current fields. It also acts as a temporal filter reducing the effects of high frequency oscillations in the data, thus making it impossible to resolve some high frequency waves in the assimilation. These high frequency waves, however, are generally not resolved by the observations.

The final modification of the $\mathbf{F}^{-1}$ matrix occurs only when the magnitude of the difference between the guess field and the observation is greater than 5°C. For differences between 5°C and 10°C, the diagonal elements are modified by a factor given by $1/[1 + (|T_0| - 5)^2]$, where T_0 is the difference between the model solution and the observation. For differences greater than 10°C, the observations are eliminated. This modification acts as an additional quality control check, with those observations that differ from the current solution by more than 5°C given less weight.

The functional (1) is minimized using a preconditioned conjugate gradient algorithm (Gill et al. 1980; Navon and Legler 1987). This algorithm does not attempt to directly find the minimum of the functional but rather iteratively attempts to find the solution. The preconditioning in the algorithm is supplied by the $\mathbf{E}$ matrix. The use of this matrix for preconditioning allows the solution to be found without directly inverting the $\mathbf{E}$ matrix.

The key quantities in the preconditioned conjugate gradient algorithm are the derivative of the functional with respect to the correction field ($\mathbf{g}$) and $\mathbf{h} = \mathbf{E}\mathbf{g}$. For the functional used in this paper, $\mathbf{g}$ is given by

$$\mathbf{g} = \mathbf{E}^{-1}\mathbf{T} + \mathbf{D}^{\mathrm{T}}\mathbf{F}^{-1}(\mathbf{D}(\mathbf{T}) - \mathbf{T}_0). \qquad (3)$$

To allow the calculations of $\mathbf{g}$ without inverting $\mathbf{E}$, the initial guess for the correction field $\mathbf{T}^1$ is set equal to zero. Thus,

$$\mathbf{T}^1 = 0$$
$$\mathbf{g}^1 = -\mathbf{D}^{\mathrm{T}}\mathbf{F}^{-1}\mathbf{T}_0$$
$$\mathbf{h}^1 = \mathbf{E}\mathbf{g}^1.$$

Also initialized to zero are the initial search directions ($\mathbf{d}^0$ and $\mathbf{e}^0$). Then, to find the correction field ($\mathbf{T}$), the preconditioned conjugate gradient algorithm is applied. The sequence of calculations performed in each iteration is as follows:

$$\mathbf{d}^n = -\mathbf{h}^n + \beta^{n-1}\mathbf{d}^{n-1}$$
$$\mathbf{e}^n = -\mathbf{g}^n + \beta^{n-1}\mathbf{e}^{n-1}$$
$$\mathbf{f}^n = \mathbf{e}^n + \mathbf{D}^{\mathrm{T}}\mathbf{F}^{-1}\mathbf{D}\mathbf{d}^n$$
$$\alpha^n = \frac{(\mathbf{g}^n)^{\mathrm{T}}\mathbf{h}^n}{(\mathbf{d}^n)^{\mathrm{T}}\mathbf{f}^n}$$
$$\mathbf{g}^{n+1} = \mathbf{g}^n + \alpha^n\mathbf{f}^n$$
$$\mathbf{T}^{n+1} = \mathbf{T}^n + \alpha^n\mathbf{d}^n$$
$$\mathbf{h}^{n+1} = \mathbf{E}\mathbf{g}^{n+1}$$
$$\beta^{n+1} = \frac{(\mathbf{g}^{n+1})^{\mathrm{T}}\mathbf{h}^{n+1}}{(\mathbf{g}^n)^{\mathrm{T}}\mathbf{h}^n},$$

where n is the iteration counter, initially set equal to one.

Each iteration of this analysis scheme is fairly expensive, with most of the expense involved in the multiplication by the **E** matrix. Tests of the iterative algorithm revealed that most of the convergence occurred in the first three iterations. Thus, only three iterations per timestep are currently being used. This limitation does not appear to degrade substantially the results from the assimilation.

After adding the correction field to the predicted temperature field, it was noted that in regions of low vertical stability the adjustment of the temperatures could make the profile unstable. These unstable regions created adverse effects on the model solution. For this reason, a convective adjustment routine was applied to the density field as the final component of the insertion process. However, by continually creating neutrally stable layers, very large mixing coefficients were generated by the vertical mixing routine. While this is harmless for the thermal field, since it is already well mixed, it is detrimental to the momentum field. To reduce the artificial vertical mixing, the routine which predicts the vertical mixing coefficients was altered to make the vertical stability appear to be slightly enhanced. This reduced the vertical mixing during the assimilation to a level consistent with that produced when no assimilation was performed.

5. Results

The goal of oceanic four-dimensional data assimilation is to produce the best possible representation of the true oceanic state. Unfortunately, since the true oceanic state is unknown, it is very difficult to evaluate the quality of the assimilation procedure. To get some idea of the character of the assimilation technique, it will be examined in four different ways. First, the assimilation results are compared to independent analyses. This comparison suffers from two limitations. Many of the same observations were used in the independent analyses. Thus similar errors could be contained in both analyses. Also, when differences exist between the analyses, it is often difficult to determine which analysis is closer to reality. Second, the analyses from the assimilation will be compared directly back to the observations. Since the observations were used in the assimilation, this comparison is a measure of the assimilation technique's ability to force the model solution towards the data. It is also important that the model retain the information inserted into the model solution for a reasonable period of time. Thus, a model integration initialized from the assimilation was run. This solution is compared with the assimilation results and another integration initialized without any assimilation of data. Finally, the analyses from the assimilation are subjectively examined to ensure that they contain known oceanic features.

It was necessary to create a reasonable initial state to begin the assimilation procedure. This was done in the same manner as RM. The 12-level version of the model used by RM was initialized with the climatological temperatures and salinities of Levitus (1982) and integrated six years using climatological surface forcings (Hellerman and Rosenstein 1983). The results of this 12-level model integration were then vertically interpolated to the 15 vertical levels used in this study.

From these initial conditions, the assimilation was begun at 1 December 1978 with the forecast model and 12-hourly forcings described in section 2. Also begun from the same initial conditions, model and forcing was a control integration of the model in which no observational data was inserted. This integration will be referred to as the control from climatology. A third similar integration without the insertion of data was started from the assimilation solution at 1 June 1979. This integration will be referred to as the control from assimilation.

In Fig. 3, the monthly mean December 1979 global SSTs after 13 months of assimilation and the Climate Analysis Center's (CAC's) operational product (Reynolds 1988) are shown. In 1979 the operational product extended only to 40°S. Thus, the operational analyses were meshed with climatology to produce global fields. Note that while the observational datasets used to produce the operational product and the assimilation results are different, many of the same observations are used in both analyses.

Comparisons of the operational analysis and the assimilation results show few large scale differences, although the operational analysis is much smoother than the assimilation results. The operational analyses are heavily smoothed and do not contain many realistic small scale features. While not all small scale structure in the assimilated fields are real, resulting from incorrect atmospheric forcing, inadequacies in the assimilation procedure or noise in the data or forecast model, some small scale features appear more realistic than those in the operational analyses.

Figure 4 contains SST difference fields between CAC's operational product and the control from climatology (a), the control from assimilation (b) and the assimilation (c) for December 1979. The fields are shown only to 40°S, the limit of the operational analyses during this period. As mentioned earlier the only difference between the control from climatology and the control from assimilation is that, in the control from assimilation, data were assimilated between 1 December 1978 and 31 May 1979. From 1 June through 31 December 1979, the control from assimilation was continued identically to the control from climatology. The effects of the assimilation can still be seen in the SST fields in December 1979. In the tropics, the control from assimilation has cooler SSTs in the Atlantic, Indian and western Pacific. In the western northern Atlantic and Pacific oceans, large negative and positive biases in the control from climatology have been substantially reduced. As will be seen later in this

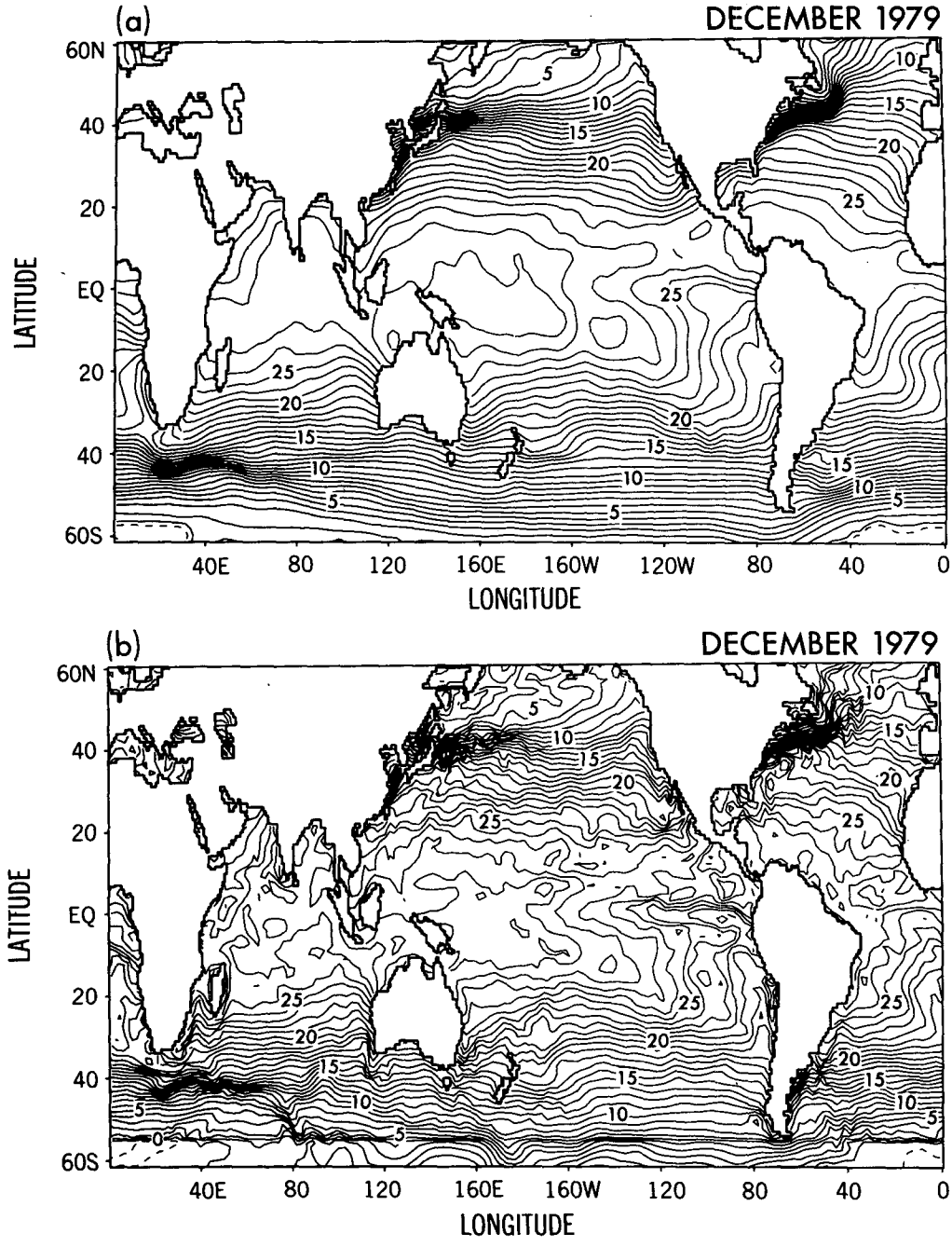

FIG. 3. Monthly mean December 1979 SST analyses from the Climate Analysis Center's operational product (a) and from the assimilation (b). Contour interval is 1°C.

section, larger differences between the two control runs are found below the surface.

Comparison of the two control runs to the analyses from the assimilation and the operational product indicate several biases due to modeling errors or improper forcing. The most prominent of these biases are the warm temperatures in the western tropical Pacific, the cool eastern tropical Pacific and the improper separation of the Gulf Stream from the east coast of North America. The Gulf Stream bias may be a result of insufficient resolution and would be difficult to reduce without improved resolution. However the tropical Pacific biases are probably not a result of resolution but rather inadequacies in the model physics or the atmospheric forcing. Large differences between the model results and the operational analysis can also be

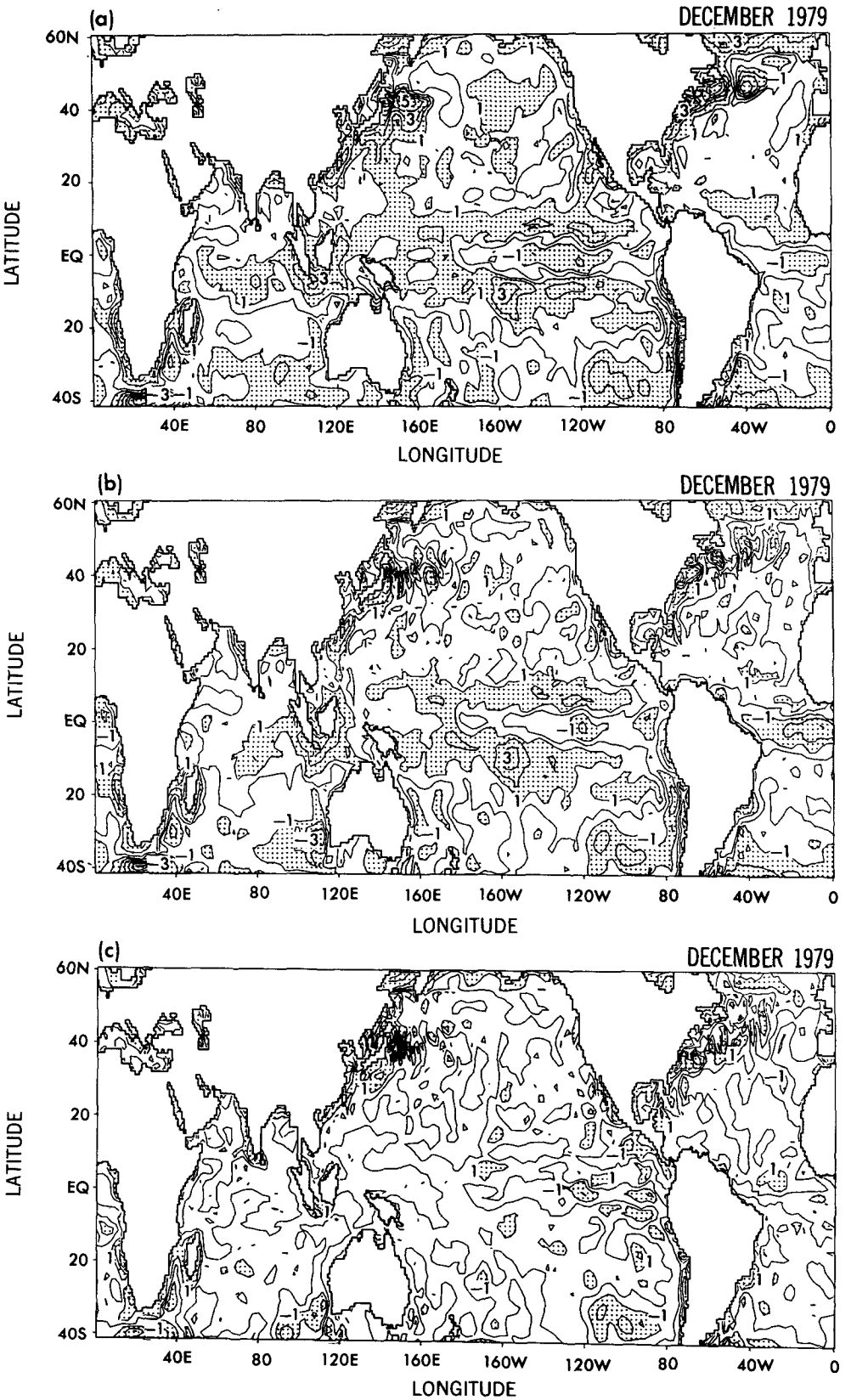

FIG. 4. Monthly mean SST differences between the Climate Analysis Center's operational product and the control from climatology (a), the control from assimilation (b) and the assimilation (c). Negative values imply the operational product is warmer than the analysis. Contour interval is 1 °C with deviations greater than 1 ° stippled.

seen near land. These differences could result from inadequate resolution, poor atmospheric forcings due to deficiencies in handling land–sea interfaces in atmospheric data assimilation and heavy smoothing of the operational analyses eliminating real features.

Some quantitative information on the analysis quality can be obtained by comparing the analyses to the observations at the observation locations. This must be done with care since the statistics are biased towards heavily observed regions, although these statistics are still instructive. Figure 5 displays the rms differences between the observations and the analyses for the 13-month assimilation interval. The statistics are divided into two regions, the tropice (10°N–10°S) and outside the tropics. In addition to comparisons to the assimilation, operational analyses (Reynolds) and the two control integrations, the observations are compared to the climatology of Levitus (1982).

As might be expected the two control runs did not perform as well as the assimilation and operational analyses. Unfortunately the control runs also do not perform as well as climatology. During this period, the signal in the surface temperature fields relative to climatology was quite small. This result implies that the model forecast may not have been as good of a guess as climatology for this time period, but as the model solution is improved through changes in the model physics and improved forcing, the guess will improve and thus the analyses will be improved. Also, the assimilation procedure has the advantage of producing consistent global current fields.

In the tropics, the control from climatology performs better than the control from assimilation. This results from an initial adjustment problem that occurs in the transition from the assimilation to no assimilation, which will be discussed later in this section.

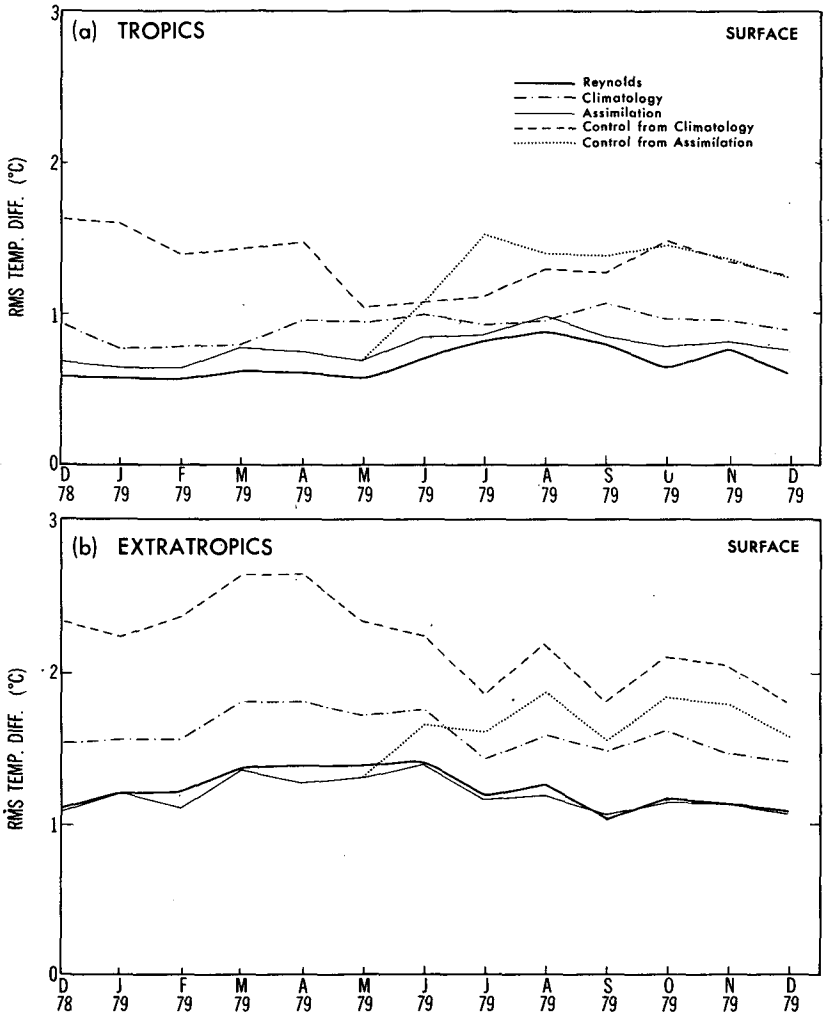

FIG. 5. Rms surface temperature differences between observations and analyses in the tropics (a) (10°N–10°S) and outside the tropics (b).

The assimilation and the operational analyses approximately fit the data equally well, with the operational analyses fitting the data slightly closer in the tropics. By changing the first guess error covariance (E), the fit at the observation locations could be improved in the assimilation, although the quality of the analyses in regions where no data exists may be degraded. As upper and lower bounds of the fit, the rms differences should not be smaller than the observational errors (the scheme should not fit observational noise) and the rms difference should not be larger than the differences from the control integrations (the assimilation scheme should improve the solution). While the assimilation solution lies between these bounds, it is obvious that in the tropics slightly better analyses could be produced, since the operational analyses fit the data better and do not contain noise. A substantial portion of the difference between the results from the assimilation and the operational analyses results from the requirement in the assimilation solution that the density field be statically stable. Because of this requirement, errors in the analyses at lower levels can force the solution away from the observations at the surface. Relaxation of this requirement can result in improved fit at the surface, but produces an overall degradation in the solution. This is not a difficulty in the operational analyses because they are two-dimensional surface analyses. Some of the errors in the assimilation also result from the model and forcing biases mentioned earlier. While the incorporation of data is able to correct much of the model and forcing errors, some of the errors still persist in the analyses from the assimilation.

The comparison of the analyses to the observations in subsurface layers generally contains similar results, except that the control from assimilation is consistently superior to the control from climatology. Figure 6 shows these statistics for a depth of 100 m. These are representative of the results at other levels except that it is the level at which the errors are largest in amplitude.

The analyses in the subsurface layers will be examined further by comparing vertical cross sections in the equatorial Pacific between the various integrations. In Fig. 7, an east–west equatorial temperature cross section is shown for the two control runs and the assimilation. The most obvious difference is the much tighter gradients in the assimilation thermocline than in the control from climatology thermocline. The control from assimilation retains much of the stronger thermocline in the west, but weakens it substantially in the east. This weakening in the east appears to be related to the previously mentioned adjustment problem which occurs in the transition between assimilation and no assimilation. In the month following the end of the assimilation process, a region of downwelling develops around 140°W with a region of upwelling east of 120°W. The downwelling region propagates eastward depressing the isotherms. The effect of the upwelling is to create cooler temperatures at the surface east of 120°W, with a rapid recovery 2–3 months later as the downwelling propagates into the region. After the downwelling feature has propagated out of the region, the upper part of the thermocline recovers to its earlier depth, while the lower part of the thermocline remains too deep, thus resulting in diffusion of the thermocline. This feature, which is not found in the control from climatology, is currently being investigated in more detail.

Finally a north–south cross section in the equatorial Pacific is examined. Figures 8 and 9 show the isotherms and zonal component of the current, respectively. Again, the most obvious difference is the much stronger thermocline in the assimilation than in the control from climatology. The weakening of the thermocline along the equator is apparent between the assimilation and the control from assimilation. Also, the ridge in the temperature field at 10°N is more intense in the assimilation than the control from assimilation. In this case, the weakening of the thermocline was gradual over several months, after which the strength stays fairly constant. Apparently the gradient is too strong for the model to retain without assistance from the data.

Considerable differences in the velocity fields can also be found between the various runs. The amplitude of the equatorial undercurrent is much stronger and more realistic in the assimilation than in either of the control integrations, but in the assimilation, the vertical extent of the undercurrent appears to be too large with the vertical structure in the control from assimilation being much more realistic. In the control from climatology, the maximum in the undercurrent is much too deep. This is probably related to the diffuse thermocline. In general, it has been noted that the equatorial undercurrent produced in the assimilation tends to be sensitive to small changes in the vertical mixing routine, even when the thermal fields appear to be similar. Off the equator, the countercurrents appear to be similar between the runs except in amplitude. As seen in Fig. 8, the north–south temperature gradient associated with the north equatorial countercurrent is strongest in the assimilation and weakest in the control from climatology. South of the equator, the opposite is true. As might be expected, the countercurrents are consistent with this pattern.

6. Summary and conclusions

A global oceanic data assimilation system has been developed. This data assimilation system was created primarily for the initialization of coupled ocean–atmosphere models for use in producing seasonal forecasts. The fields produced by the assimilation procedure have other applications as well, such as climatological and dynamical studies of the oceans and for the creation of time dependent boundary conditions for higher resolution limited domain models.

The data in this study consisted only of conventional temperature observations. For surface observations 2°

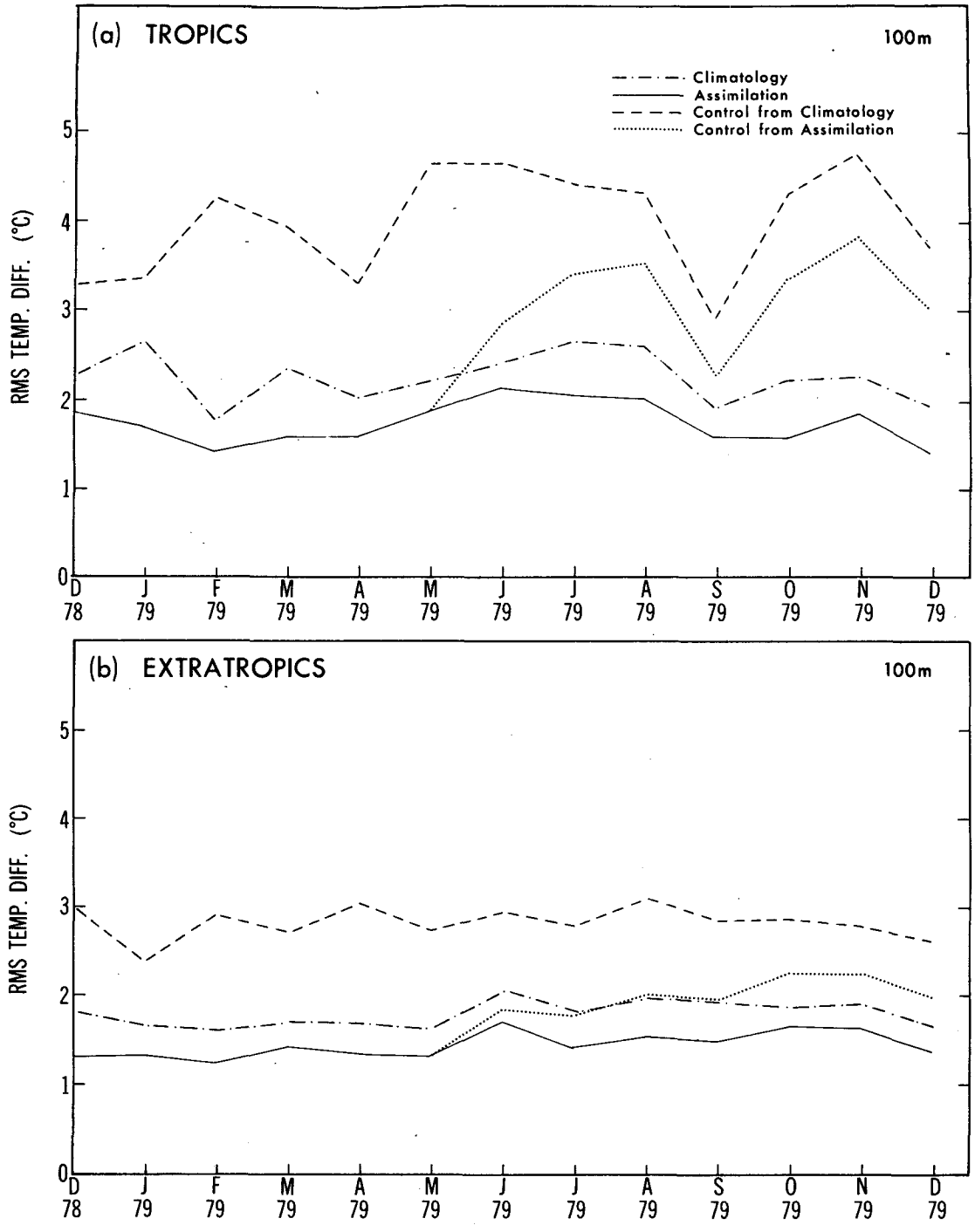

FIG. 6. As in Fig. 5 except at 100 m depth.

× 2° COADS data (Woodruff et al. 1987) were used. Vertical temperature profiles were incorporated from NODC and the U.S. Navy's MOODS dataset. While the coverage of the sea-surface temperature data was quite good during this period, the vertical temperature profile data contained large gaps in the equatorial region and in the Southern Hemisphere. In addition to problems involving inadequate data coverage, the observations also occasionally contained large errors. This required application of extensive quality control to the data.

The assimilation procedure was developed using a modified version of a global high resolution numerical model developed by RM. This model is based on the

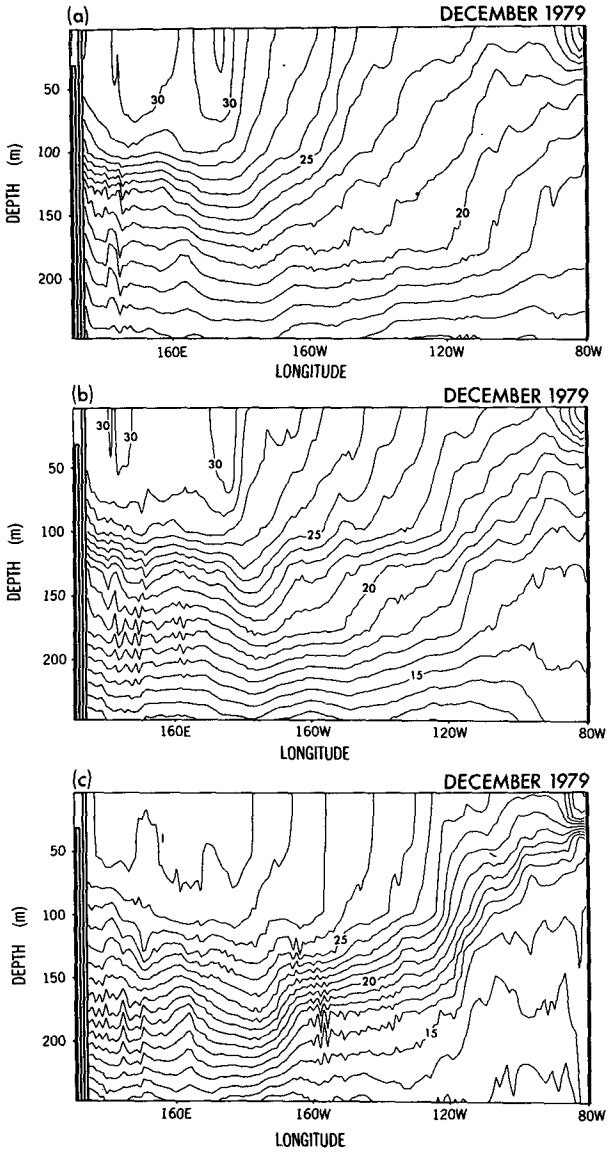

FIG. 7. East–west equatorial Pacific temperature cross-sections for the control from climatology (a), the control from assimilation (b) and the assimilation (c). Contour interval is 1°C.

primitive equations with the atmospheric forcing provided from the 12 hourly atmospheric fields created using the GFDL atmospheric data assimilation system (Ploshay et al. 1983). The data were inserted into the model using a continuous insertion technique. A temperature correction field was created every model timestep and inserted into the model solution. The temperature correction was created by applying a statistical objective analysis routine to the differences between the model solution and the data in a 30-day window around the analysis timestep.

The results from the assimilation system applied over a 13-month period are encouraging. The SST fields compare well to the operational analyses in terms of large scale features. For smaller scales, the analysis captures some features not contained in the operational analyses, but it also probably contains too much noise. Since the control forecasts contain the same level of small scale structure as the assimilation, it does not appear to be created primarily by an imbalance between the mass and momentum fields. This would be the case in an atmospheric assimilation system which generally uses an additional initialization procedure to reduce the imbalance. The rigid-lid approximation in the model, which eliminates external gravity waves, and the higher viscosity in the oceans, that damp the small-scale internal gravity waves, probably make bal-

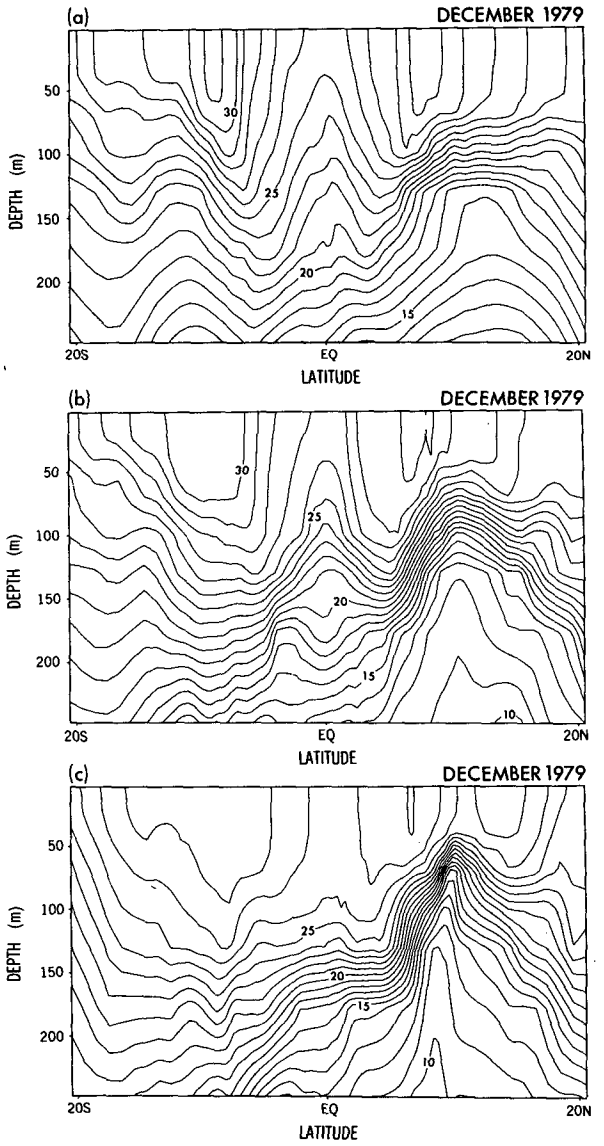

FIG. 8. North–south temperature cross sections at 149°W for the control from climatology (a), the control from assimilation (b) and the assimilation (c). Contour interval is 1°C.

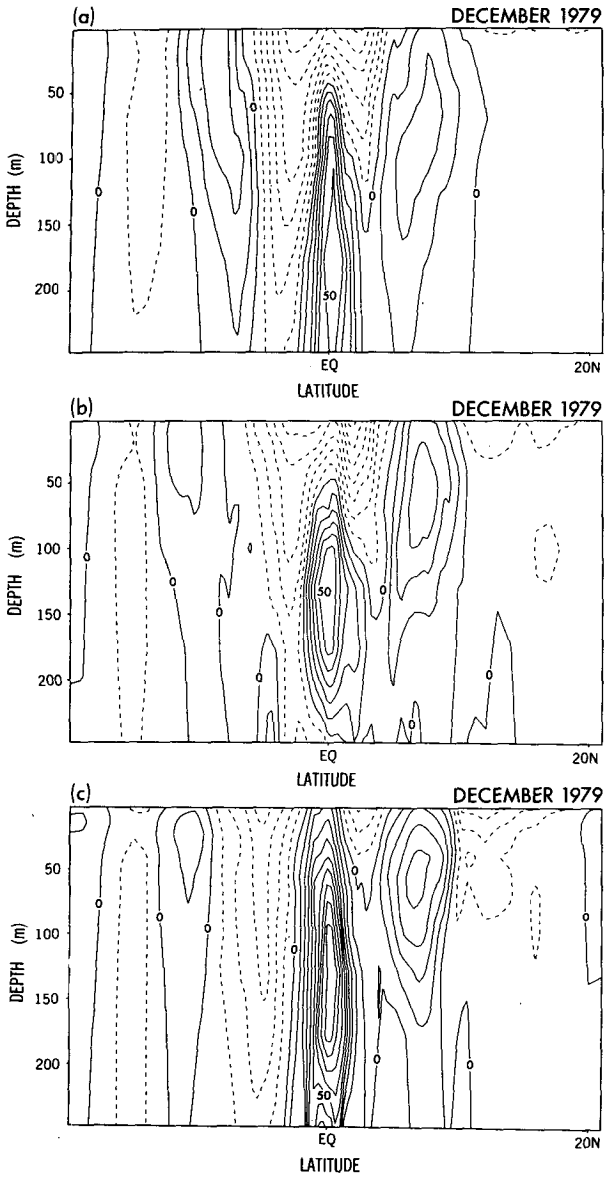

FIG. 9. As in Fig. 8 except zonal component of velocity field. Contour interval is 10 cm s^{-1}.

ancing of the mass and momentum fields unnecessary in this ocean model. At subsurface levels, the model solution is made much more realistic by the inclusion of data. In figures from the equatorial Pacific, a much stronger thermocline and deeper mixed-layers are seen in the assimilation than in the control integrations. Comparison between the control integrations reveal that the information inserted by the assimilation system is remembered by the model resulting in more realistic features, although an adjustment problem is noted in the eastern equatorial Pacific which degrades the control from assimilation solution.

While the assimilation results are encouraging, further improvements can be made to the assimilation system. The assimilation procedure can be improved by expanding the database and by improving the quality control. Since the period of this study, the number of available vertical temperature profiles has increased. The assimilation results should be substantially improved if applied to a more recent period because of this expanded database. Also, the database can be expanded by using other observation types, such as satellite derived SSTs, altimetry data, and current measurements.

Changes to the assimilation procedure itself can also enhance the results. The statistics currently in use in the statistical objective analysis scheme are somewhat ad hoc. The first-guess error statistics are obviously deficient since they only vary in the meridional direction. The ocean dynamics are very different in regions near the coast or with strongly sloping bottom topography. The model's error characteristics are quite different in these regions because of this and should be accounted for in the first guess error covariances. With improved knowledge of the model's and the data's error characteristics many of the errors in the present system can be reduced. Also, it is not clear that it is necessary to update the model solution every timestep. An intermittent insertion may produce equivalent results at less expense. This change would require some modification of the statistics and additional research to determine the optimal insertion interval.

Finally, any improvement to the numerical model solution feeds back to improve the analyses produced by the assimilation system. These improvements can be made directly to the model dynamics or physics, or can enter through the atmospheric forcing. At this point in time, the effects of improper atmospheric forcing on the assimilation are not completely known. For example, the adjustment problem during the transition between assimilation and no assimilation could be partially due to an inconsistency between the wind stress and the east–west slope of the density field inferred from the data. While further research is necessary on this particular feature, it is obvious that improved atmospheric forcings will give better oceanic simulations (Leetmaa and Ji 1989). Thus, improved atmospheric forcings will result in improved analyses from the assimilation system, but any alterations to the model will produce changes in the model's error characteristics and thus may require additional changes to the statistics in the objective analysis routine. Also, the ability to make improvements to the numerical model is enhanced by its use in a data assimilation system. The necessity of comparing the model to the data makes the model errors more apparent and thus makes their elimination more likely. In fact, several of the changes to the model from that used in RM resulted from this process.

Oceanic data assimilation is still in its infancy and many more years of research are necessary to realize

its full potential. It is not necessary, however, to incorporate all the potential enhancements to the data assimilation system to produce useful results. Useful analyses should be available from oceanic data assimilation systems in the near future.

Acknowledgments. The authors wish to thank Drs. K. Miyakoda, S. G. H. Philander and A. Leetmaa for their many useful comments during the course of the work and Drs. K. Miyakoda, L. Kantha, K. Bryan and two anonymous reviewers for their useful reviews of this paper. We would like to acknowledge S. Levitus for his invaluable assistance in providing the data and data processing routines. We would also like to thank P. Tunison and J. Connor for their assistance in preparing the figures and W. Marshall for her expert typing of the manuscript.

REFERENCES

Bauer, R. A., 1985: Functional Description Master Oceanographic Observation Data Set (MOODS). FNOC, 477 pp. [Available from Fleet Numerical Oceanographic Center, Monterey, California 93940.]

Bengtsson, L., 1975: Four-dimensional assimilation of meteorological observations. World Meteor. Org., International Council of Scientist Unions, Joint Organizing Committee, 75 pp. [GARP Publ. Ser. No. 15.]

——, M. Ghil and E. Kallen (Eds.), 1981: *Dynamic Meteorology: Data Assimilation Methods.* Springer-Verlag, 330 pp.

Bennett, A. F., and W. P. Budgell, 1987: Ocean data assimilation and the Kalman filter: Spatial regularity. *J. Phys. Oceanogr.*, **17,** 1583–1601.

Bryan, K., 1969: A numerical method for the study of the world ocean. *J. Comput. Phys.*, **4,** 347–376.

Clancy, R. M., K. D. Pollack, J. A. Cummings and P. A. Phoebus, 1988: Technical description of the optimal interpolation system (OTIS) Version 1: A model for oceanographic data assimilation. FNOC Tech. Note 422-86-02, 422 Branch, Fleet Numerical Oceanography Center, Monterey, CA.

Cooper, N. S., 1988: The effect of salinity on tropical ocean models. *J. Phys. Oceanogr.*, **18,** 697–707.

Cox, M. D., 1984: A primitive equation 3-dimensional model of the ocean. GFDL Ocean Group Tech. Rep. No. 1, 143 pp.

Galperin, B., L. H. Kantha, S. Hassid and A. Rosati, 1988: A quasi-equilibrium turbulent energy model for geophysical flows. *J. Atmos. Sci.*, **45,** 55–62.

Gandin, L. S., 1963: *Objective Analysis of Meteorological Fields,* Gidrometeorologich eskoe Izdatel 'sivo, 242 pp.

Gill, P. E., W. Murray and M. H. Wright, 1981: *Practical Optimization.* Academic Press, 401 pp.

Hellerman, S., and M. Rosenstein, 1983: Normal monthly wind stress over the world ocean with error estimates. *J. Phys. Oceanogr.*, **13,** 1093–1104.

LeDimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus,* **38A,** 97–110.

Leetmaa, A., and M. Ji, 1989: Operational hindcasting of the tropical Pacific. *Dyn. Atmos. Oceans,* in press.

Levitus, S., 1982: *Climatological Atlas of the World Ocean.* NOAA Prof. Paper, No. 13, U.S. Govt. Printing Office, 173 pp.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112,** 1177–1194.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.*, **20,** 851–875.

Miller, R. N., 1985: Toward the application of the Kalman filter to regional open ocean modeling. *J. Phys. Oceanogr.*, **16,** 72–86.

Moore, A. M., N. S. Cooper and D. L. T. Anderson, 1987: Initialization and data assimilation in models of the Indian Ocean. *J. Phys. Oceanogr.*, **17,** 1965–1977.

Navon, I. M., and D. M. Legler, 1987: Conjugate-gradient methods for large-scale minimization in meteorology. *Mon. Wea. Rev.*, **115,** 1479–1502.

Pacanowski, R., 1987: Effect of equatorial currents on surface stress. *J. Phys. Oceanogr.*, **17,** 833–838.

Ploshay, J., R. White and K. Miyakoda, 1983: FGGE level IIIb daily global analysis, Parts I, II, III, and IV. Dec. 1978–Nov. 1979, NOAA Data Rep. ERL GFDL 1, 278 pp.

Reynolds, R. W., 1988: A real-time global sea surface temperature analysis. *J. Climate,* **1,** 75–86.

Rienecker, M., C. N. K. Mooers and A. Robinson, 1987: Dynamical interpolation and forecast of the evolution of mesoscale features off Northern California. *J. Phys. Oceanogr.*, **17,** 1189–1213.

Robinson, A. R., and W. G. Leslie, 1985: Estimation and prediction of oceanic eddy fields. *Progress in Oceanography,* Vol. 14, Pergamon, 485–510.

Rosati, A., and K. Miyakoda, 1988: A GCM for upper ocean simulation. *J. Phys. Oceanogr.*, **18,** 1601–1626.

Smagorinsky, J., S. Manabe and S. L. Holloway, Jr., 1965: Numerical results from a nine-level general circulation model of the atmosphere. *Mon. Wea. Rev.*, **93,** 727–768.

Thacker, W. C., and R. B. Long, 1987: Fitting dynamics to data. *J. Geophys. Res.*, **93,** 1227–1240.

Woodruff, S. D., R. J. Slutz, R. L. Jenne and P. M. Steurer, 1987: A comprehensive ocean–atmosphere data set. *Bull. Amer. Meteor. Soc.*, **68,** 1239–1250.