

Assimilating XBT Data into HYCOM

WILLIAM CARLISLE THACKER

Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida

OLEG EVGENIEVICH ESENKOV

Cooperative Institute for Marine and Atmospheric Studies, Miami, Florida

(Manuscript received 8 December 2000, in final form 9 October 2001)

ABSTRACT

A scheme is presented for assimilating expendable bathythermographic data into HYCOM, an oceanic circulation model featuring a hybrid vertical coordinate. The scheme is fully multivariate, using observations of temperature to correct density, pressure, salinity, and momentum, in addition to temperature. Central to the scheme is the estimation of companion profiles of salinity and potential density. The potential density profiles are used to estimate the thicknesses of the model's layers, so that layer-averaged values of potential density and potential temperature can be computed. These derived data and the derived layer thicknesses are assimilated via optimal interpolation. Salinity corresponding to the corrected potential density and potential temperature fields is determined by the equation of state of seawater, and corrections to the momentum field are computed geostrophically from the corrections to the pressure field. The scheme is illustrated using data from March 1995 in the Atlantic Ocean.

1. Introduction

Given initial conditions and subsequent boundary fluxes, numerical models of oceanic circulation simulate the evolution of temperature, salinity, and currents. Unfortunately, both the initial state and the fluxes are subject to substantial uncertainty, so the simulated oceanic state can be expected to be in error. Errors in density manifest in the pressure field to adversely effect currents and, via advection, feed back to further aggravate the density distribution. Furthermore, modeling unresolved mixing processes is difficult at best, and errors in buoyancy only make matters worse. Without some way to control errors, simulations can diverge from reality. While this is a concern for all oceanographic models, this paper focuses on a particular model: the Hybrid-Coordinate Oceanographic Model (HYCOM) (Bleck 2002), which describes the ocean as a stack of fluid layers that are density-like in and below the pycnocline and pressure-like within the surface mixed layer.

Data assimilation is the process by which observations are used to correct the evolving state of the model (Bengtsson et al. 1981; Daley 1991; Malanotte-Rizzoli 1996). Most methods for assimilating data are statistical

in nature and are supported by a common formalism, even though they might differ widely in their implementation. While this formalism is often presented within a variational context, at heart it amounts to nothing more than regression: corrections are estimated as linear functions of model–data differences. The principal difference between the various methods lies in how the regression coefficients are obtained, or more precisely, how error covariances are modeled. For example, Kalman filtering (Kalman 1960; Gelb 1974; Cohn 1997) attempts to exploit the effects of dynamics on the evolution of errors, even though the initial errors, their correlations, and the rate of injection of error are quite difficult to approximate. On the other hand, optimal interpolation (Gandin 1963; Derber and Rosati 1989; Carton and Hackert 1990; Behringer et al. 1998) often assumes the error structure to be temporally stationary and spatially inhomogeneous. Although both approaches generally assume the model to be unbiased, ignoring mean errors while relying on a covariance-based formalism, Dee and da Silva (1998), Evensen et al. (1998), and Carton et al. (2000) do address the important issue of correcting for model bias.

The statistical basis for any data-assimilation scheme should derive from a careful study of model–data differences. If any biases are detected, the model should be corrected, either via the introduction of empirical compensating terms or, better, by tracking the biases back to identifiable causes that can be corrected in a

Corresponding author address: Dr. William Carlisle Thacker, NOAA/Atlantic Oceanographic and Meteorological Laboratory, 4301 Rickenbacker Causeway, Miami, FL 33149.
E-mail: thacker@aoml.noaa.gov

more physically meaningful way. Once biases have been removed, the scatter of the data about their model counterparts provides information about the error covariances needed for correcting the model state. Because of the sparsity of the data and the irregularity of the sampling, the process of inferring the error statistics is difficult. Details of this process are beyond the scope of this paper. Here, these statistics are assumed to be given, albeit with considerable uncertainty.

An important aspect of Kalman filtering is that, by accounting for the dynamically evolving uncertainties, it generates cross-field error covariances. For example, errors in the mass field evolve into errors in the momentum field. However, with no strong dynamical connection between temperature and salinity,¹ cross covariances between errors of those fields cannot be expected to evolve if initially absent, even though they might be implied by water-mass properties. Given our limited ability to quantify the model errors and their correlations, worrying about how errors should be propagated dynamically or how assimilation should impact the error statistics is not of highest priority. Our strategy is to start with a simple scheme that can exploit the bulk of the information carried by the data. However, this scheme should be fully multivariate, correcting all model fields. Later, as confidence builds in the model and in the error statistics, the assimilation scheme can be improved.

The focus here is on the assimilation of XBT profiles, as they are the most numerous in situ oceanographic observations and provide the core of the present oceanic observing system.² Some strategy is needed for correcting salinity, which is essentially unobserved. Here, each temperature profile is complemented by *estimated* companion profiles of salinity, density, and pressure, which are reduced to data for the pressure (depth) of the model's layer interfaces and for layer averages of potential temperature and potential density. These layer-specific values are the data to be assimilated. Near the surface, interface pressures should need little or no correction, because the preprocessing of the XBT profiles is designed to reflect the model's preferred layer thicknesses; here, layer properties are adjusted as for a conventional z -coordinate model. For deeper layers, density estimated from the XBT data and its model counterpart should both be close to the model's target density, so density should need little or no correction; layer thickness and temperature, on the other hand, might need substantial adjustments.

When assimilating data, it is important to keep in mind exactly what is being modeled in order to understand what information the data provide. The nu-

merical model starts as a code for solving the partial-differential equations governing oceanic dynamics and the boundary and initial conditions that specify the particular situation. Its spatial and temporal resolutions can depend on the particular configuration, as can its spatial and temporal extents, allowing the same code to be used for high-resolution, short-term problems as for low-resolution long integrations. Data, on the other hand, reflect all spatial and temporal scales and must be interpreted within the appropriate context. This is accomplished through the specification of error covariances, which determine the impact of the data. The problem is that of estimating month-to-month changes in the circulation as expressed by the XBT data acquired over the past few decades. Consequently, details such as eddies will not be represented in the model, even though they are manifest in the data. To handle the data properly is a monumental task that is beyond the scope of this work, requiring estimates of which casts lie within eddies and how representative of ambient conditions any given cast might be. Conversely, for an eddy-resolving model, it would be necessary to identify which model eddies should be eliminated as being only numerical artifacts. Nevertheless, the approach taken here should accommodate these issues, as long as attention is given to extracting the needed information from the available data.

2. Model

HYCOM has evolved from the University of Miami's Isopycnic Coordinate Oceanic Model (MICOM) (Bleck and Chassignet 1994; Bleck 1998) to allow for more flexible vertical resolution, building on the foundations laid by Bleck and Boudra (1981) and Bleck and Benjamin (1993). It preserves the original intent of using a density-like vertical coordinate to eliminate unrealistic diapycnic mixing caused by truncation errors, while allowing the low-density layers needed to characterize tropical near-surface waters to become more pressure-like at higher latitudes, where they can give additional resolution within the surface mixed layer. Each layer's potential density is maintained close to a specified target value, wherever possible, so that the oceanic stratification is well resolved. When there is no water locally that is near the target values for the upper layers, those otherwise unused layers default to a prescribed minimum thickness and characterize the local density of the near-surface water. As the ocean restratifies and water resembling some target values reappears, interfaces are adjusted so that the layers become more nearly isopycnic.

The upper panel of Fig. 1 illustrates the hybrid nature of HYCOM with results³ taken from a coarse-resolution configuration for the North Atlantic⁴ after 30 years of

¹ Because of the geostrophic relationship, correcting density can improve momentum, but nothing guarantees that correcting temperature will improve salinity or density; in fact, the opposite might occur.

² Satellite-based observations will be addressed in a subsequent paper.

³ Thanks to George Halliwell.

⁴ Figures 11 and 13 illustrate the model's Mercator grid. Resolution increases from $2^\circ \times 2^\circ$ at the equator to $1^\circ \times 1^\circ$ at 60°N .

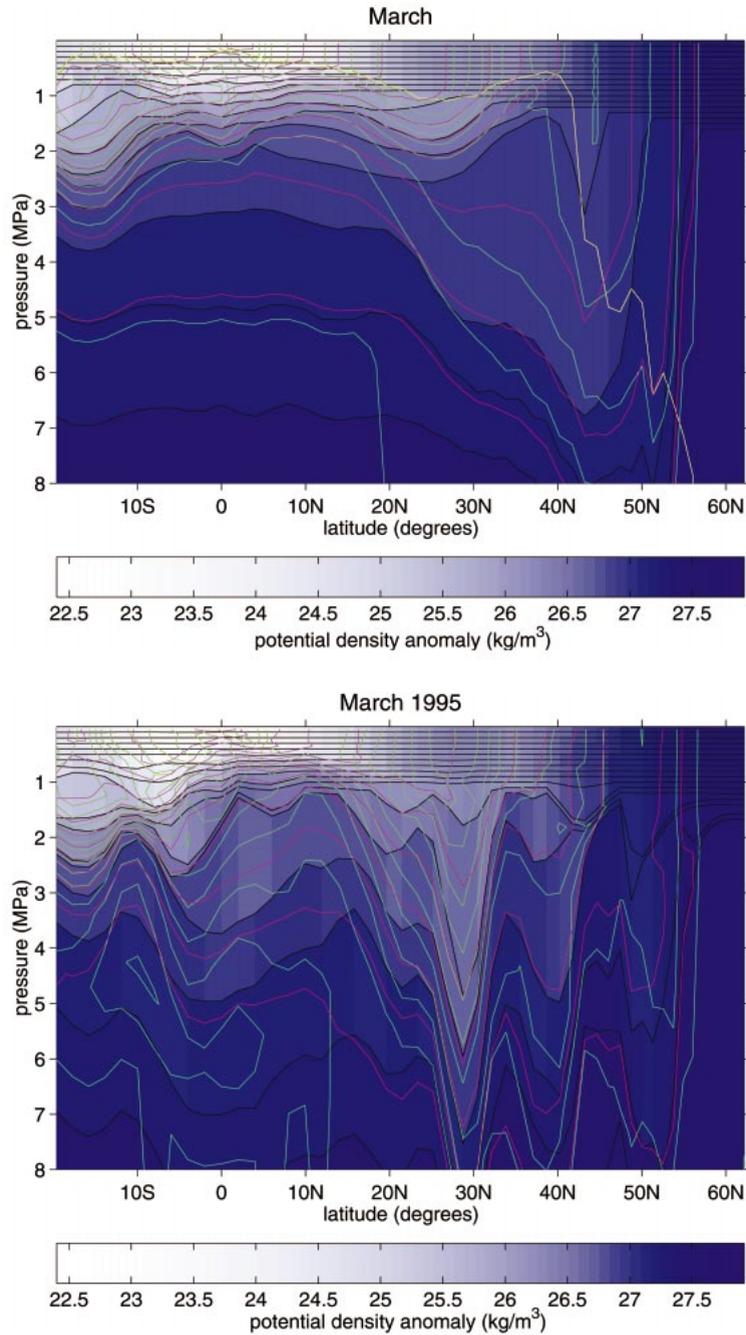


FIG. 1. Vertical section at 25°W from HYCOM simulation forced for 30 yr with climatological annual cycle of surface fluxes before (upper) and after (lower) assimilating data for Mar 1995. Shades of blue indicate variations of potential density with pressure and latitude. Pressures of layer interfaces are indicated by black curves, contours of potential temperature (2°C spacing) by magenta, contours of salinity (0.2 psu spacing) by green, and bottom of the mixed layer by yellow.

forcing by the climatological seasonal cycle of air–sea fluxes. Shown is the meridional section at 25°W for March as characterized by the 15th day of the month. The changing nature of the model’s layers can be seen by examining layer interfaces (black) and potential den-

sity anomaly⁵ (blue). Pressure is read on the vertical

⁵ Potential density anomaly is the deviation of potential density (referenced to 0 MPa) from 999.975 kg m⁻³.

axis.⁶ Contours of potential temperature⁷ (magenta) and salinity (green) have been added to emphasize their contributions to density. The depth of the mixed layer (yellow), as computed by the model (Large et al. 1994), is also shown. Note that interfaces become horizontal as they cross into the deep northern mixed layer, indicating a transition of the vertical coordinate from being density-like to being pressure-like.

To allow for the changing nature of the layers, the model transports both salinity and potential temperature. In addition, each layer follows the motion of its target water wherever it can, and the thicknesses of the layers evolve dynamically to conserve mass, while the slopes of the interfaces determine the pressure-gradient forces.

When using observations to correct the model state, at times it may be necessary to move layer interfaces and at other times to correct layer densities. The strategy taken here is to make both types of corrections, allowing the data to determine how much of each type. Correcting the mass distribution will change the pressure forces felt by the currents that advect heat and salt. While the currents might be expected to adjust quickly to the pressure corrections without unduly altering the corrected mass distribution, it is best to correct the currents as part of the assimilation process.

3. Formalism

Data assimilation can be described quite generally by a variational formalism that defines the updated model state as a compromise between a background (prior) estimate and the data (Lorenz 1986; Evensen 1994). The desired compromise is the model state vector \mathbf{x} that minimizes the quadratic objective function:⁸

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{b})^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{b}) + \frac{1}{2}(\mathbf{H}\mathbf{x} - \mathbf{d})^T \mathbf{D}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{d}). \quad (1)$$

The vector \mathbf{b} represents the prior state estimate with uncertainties characterized by the error-covariance matrix \mathbf{B} , the vector \mathbf{d} contains the observational data with errors characterized by \mathbf{D} , and the matrix \mathbf{H} transforms the state vector into a vector of model counterparts of the observations.

Here, \mathbf{b} will be taken to be the uncorrected model state, but caution is advised: this formalism implies that \mathbf{b} should be unbiased, so the model and observations should not disagree in any systematic way. The state vectors \mathbf{x} and \mathbf{b} encompass potential temperature θ , po-

tential density anomaly σ_θ , and layer thickness Δp . While salinity and velocity components could be included in the state vectors, it is simpler to diagnose their updates as described below. When the XBT data are preprocessed as described in section 4, \mathbf{d} provides layer averages of potential temperature, potential density, and layer thickness for observed grid cells. Because observations are sparse, vectors $\mathbf{H}\mathbf{x}$ and \mathbf{d} are generally much shorter than vectors \mathbf{x} and \mathbf{b} .

Differentiating J to get the condition for its minimum, solving first for $\mathbf{H}\mathbf{x}$ and then substituting, gives

$$\mathbf{x} - \mathbf{b} = \mathbf{K}(\mathbf{d} - \mathbf{H}\mathbf{b}), \quad (2a)$$

where

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{D})^{-1}. \quad (2b)$$

Each row of the matrix \mathbf{K} contains regression coefficients for updating one of the model variables as a linear combination of background–data differences. The factor $(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{D})^{-1}$ converts the differences $\mathbf{d} - \mathbf{H}\mathbf{b}$ into their data-space “fingerprint” (Hasselmann 1993; Thacker 1996) to account for redundant and uncertain information. The error-covariance⁹ matrix $\mathbf{B}\mathbf{H}^T$ provides the mechanism for correcting unobserved aspects of the model state. For example, if measurements reveal that the temperature for a particular layer and a particular grid cell should be corrected, then it is reasonable that temperatures for that layer in nearby cells have related errors and should also be corrected. Matrix \mathbf{B} effects these corrections via spatial covariances of temperature errors within that layer. If all other covariances are neglected, that is, other matrix elements of \mathbf{B} are zero, the formalism reduces to a conventional two-dimensional, univariate scheme that spreads the temperature correction in that layer over a region of influence.

This formalism can be applied at each step in a sequential assimilation procedure. The background estimate \mathbf{b} taken from a HYCOM simulation changes as the simulation evolves in time, and, in principle, its error-covariance matrix \mathbf{B} should also change. First, \mathbf{B} should reflect the reduction of uncertainty after data have been assimilated.¹⁰ Second, as the flow evolves, \mathbf{B} should also evolve (Gelb 1974; Thacker and Lewandowicz 1994). Ideally, the model could be used to propagate the uncertainties, but this could be computationally prohibitive. Third, model imperfections can be expected to increase uncertainty at each time step, further modifying \mathbf{B} ; modeling this additional error covariance is difficult. Kalman filtering, which implements these three tasks, fits nicely into this formalism.¹¹ For systems with only a few variables and a wealth of data, it might

⁶ To convert from international units to decibars, note that 100 db = 1 MPa.

⁷ Distinction between temperature and potential temperature (referenced to 0 MPa) is insignificant at these pressures.

⁸ This objective function presumes that errors in the data and those in the background state are uncorrelated.

⁹ Interpret the matrix of covariances of errors of model counterparts of the data and those of the state variables as $\mathbf{B}\mathbf{H}^T = \text{cov}(\mathbf{H}\delta\mathbf{x}, \delta\mathbf{x})$.

¹⁰ In other words, \mathbf{B} should be replaced by $(\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}$, the posterior error-covariance matrix given by the inverse of the Hessian matrix of J (Thacker 1989) before additional data are assimilated.

¹¹ Within this context, \mathbf{K} is referred to as the Kalman gain matrix.

be practical to determine the covariances of the initial state errors and the modeling errors through model–data comparisons and to update the error statistics with a Kalman filter. However, within the context of oceanographic modeling, our limited ability to estimate the initial-error and dynamical-error covariances, the unfavorable ratio of data to variables, and the Kalman filter’s computational costs suggests using less involved estimates of \mathbf{B} .

While the formalism addresses all model variables and all data via a single objective function J , simplifications are possible. For example, the objective function might be decomposed into terms corresponding to different layers:

$$J = \sum_k J_k. \quad (3)$$

This would not be the case for altimeter data, because the model counterparts of surface elevation would involve the thicknesses of all of the layers, but this could be the case for XBT data when cross-layer error-covariances are neglected. Minimizing J then amounts to minimizing each of the J_k separately, giving a solution like (2) for each layer. Neglecting cross-field error covariances allows each of the J_k to be further decomposed into terms for each field, giving solutions like (2) for each field in each layer. As each row of (2a) is a regression equation for one of the state variables, J can also be partitioned into regions (indexed by \mathbf{k}) so that each term in (3) involves only the observations that carry information about the variables in that region.

This formalism can be implemented numerically in two ways. One is to compute the corrected model state using an optimization algorithm such as conjugate–gradient descent (Gill et al. 1981) to minimize J iteratively. In this approach corrections for all state variables are computed simultaneously as the solution to a very large optimization problem. Alternatively, each row of the matrix equation (2a) can be computed independently, solving for the corrections one by one. In doing so, it is useful to break the calculation into two steps: first, solving $(\mathbf{HBH}^T + \mathbf{D})\mathbf{y} = \mathbf{d} - \mathbf{Hb}$ for the data-space fingerprint \mathbf{y} and, second, multiplying each element of \mathbf{y} by the appropriate covariance to compute the corrections for each state variable in turn. For the example discussed below, we use the second approach.

It is difficult to incorporate salinity, temperature, and density all into the variational formalism, because they are related via the equation of state of seawater (Bryden 1973; Fofonoff 1977; Brydon et al. 1999). This could be done with the use of Lagrange multipliers (Thacker and Long 1988), but it is much simpler to include only two of the three in the formalism and to use the constraint to correct the third. Because temperature is observed and density is essential for configuring the model layers, it is reasonable to work with these two, diagnosing S from the updated values for θ and σ_θ .

As is the case for salinity, there are no velocity data

to assimilate and none to gauge the magnitude of the velocity errors. Moreover, when available, velocity measurements tend to exhibit much more high-frequency variability than do measurements of temperature or salinity, rendering them difficult to assimilate. Traditionally, currents have been inferred geostrophically from the observed mass distribution. Similarly, when assimilating data, corrections to velocity can be based on dynamic balance. While such corrections can be effected via the background error-covariances (Lorenz 1981), they might also be computed diagnostically. One approach is to work with the corrected pressure field, iterating the model’s momentum equation until the uncorrected velocity field is brought into a closer balance (Bryan 1984). Another is to assimilate the data repeatedly with reduced amplitude over an interval spanning the observation time, letting the velocity field adjust to the changing pressure field (Behringer et al. 1998).

4. Preprocessing XBT data: Salinity strategy

Within the context of HYCOM, when correcting temperature, it is necessary to decide whether to move interfaces, keeping potential densities of the layers unchanged, or to correct the densities, leaving the interfaces unchanged. The decision should be based on the actual density profile. Because salinity is unobserved, it must be estimated. The simplest approach is to estimate a companion salinity profile for each observed temperature profile and to use it to estimate potential temperature and potential density profiles. Then the potential density profiles can be used to determine the depths of the model’s layer interfaces, and layer-specific data can be inferred from the profiles. This approach provides information for adjusting both pressure and density, in addition to temperature, so the question is not which to adjust but how much of each.

One strategy for estimating companion salinity profiles is to base them on the background state: the profile can be assumed to vary linearly between adjacent layer centers and to be constant in the top half of the top layer. However, because the background state is to be provided by the model, this strategy could be dangerous. The evolving salinity field is subject to substantial uncertainties in the flux of moisture through the air–sea interface, and there would be no mechanism for correcting the accumulating errors in the salinity field.

We prefer to estimate the salinity profile to be the climatological mean for the time and place of the XBT cast. Climatological data (Levitus et al. 1994; Antonov et al. 1998; Conkright et al. 1998) for specified depths can be used to estimate the profiles in a manner similar to that described in the previous paragraph. The climatology of Carnes et al. (1994), which relates salinity profiles to values of surface observables and to temperature profiles, could provide an alternative estimate. While this approach avoids perpetuating drift in the model’s salinity, it cannot recognize how the true sa-

linity differs from the climatological mean. For example, it would not be able to handle events like the great salinity anomaly (Dickenson et al. 1988).

A better approach would be to base the companion salinity profiles on temperature–salinity correlations manifest in past conductivity–temperature–depth (CTD) data (Hansen and Thacker 1999). Near the sea surface where such correlations are weak, additional observations of surface salinity would be most useful; otherwise the estimate reverts to climatology. For this strategy to be implemented, it is necessary to take on the task of examining the CTD archives to construct the empirical models for salinity profiles on a region-by-region basis. Until all regions have been modeled, a compromise would be to use such estimates where available and to use climatology everywhere else.

The companion salinity profiles provide the basis for estimating companion profiles of potential temperature and potential density. The question remains of how to transform such profiles into information about the model's layers.¹² The answer is found in the motivation for the hybrid nature of the model's layers: each layer is required to have a minimum thickness and, after that requirement is satisfied, to be as close as possible to its specified target value for potential density. Thus, each cast is processed as follows. The estimated surface density for the cast is compared to the top layer's target density to decide whether any sufficiently low-density water was observed. If not, the layer is assigned its minimum thickness, potential temperature and potential density for the first layer are computed as averages, and the question is repeated for the second layer. Once water with the target density is encountered, the remainder of the potential density profile can be partitioned so that layer averages correspond to target densities. This process gives "data" values of θ , σ_θ , and Δp (thickness) for each layer to the depth of the cast.

Four XBT casts have been chosen at random to illustrate the preprocessing at four different latitudes for March 1995. Figure 2 shows their companion salinity profiles (magenta) estimated by interpolating the climatological March salinity (Levitus et al. 1994). Depth is converted to pressure using an empirical function of latitude (Saunders and Fofonoff 1976). The measured temperature and estimated salinity are combined using an equation of state (Bryden 1973; Fofonoff 1977; Brydon et al. 1999) to obtain the estimated profiles of potential temperature (Fig. 3, magenta) and of potential density (Fig. 4, magenta). The pressure interfaces for the layers are computed from the potential density profiles in accordance with the specified minimum layer

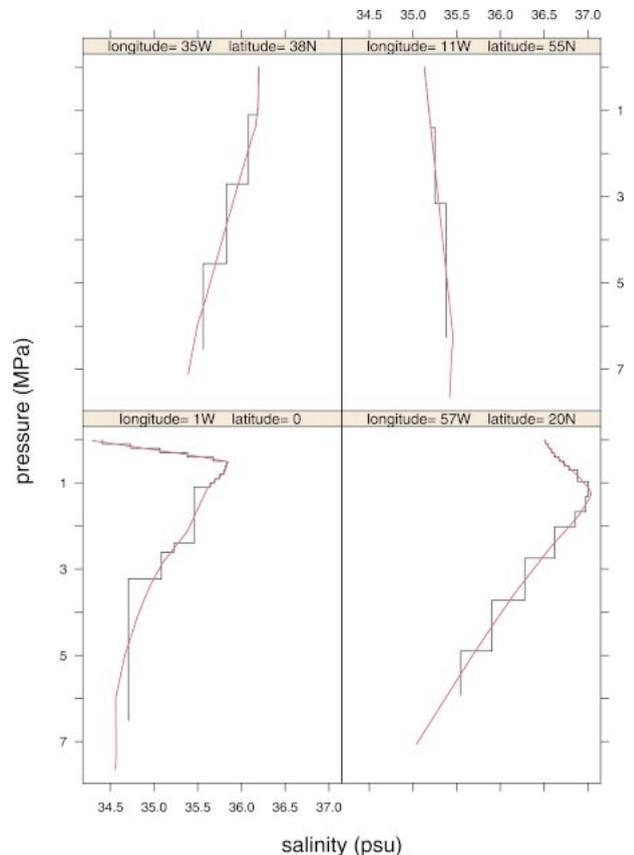


FIG. 2. Estimated salinity profiles (magenta) at locations of four XBT casts in Mar 1995 and their approximation for HYCOM as layer averages (black).

thicknesses (0.1 MPa for all layers) and target densities¹³ (Fig. 4, cyan), and layer averages for all variables are computed. The piecewise-constant profiles (Figs. 2, 3, and 4; black) illustrate both the interface pressures (horizontal segments) and layer averages (vertical segments). In Fig. 4 the lower two panels, which correspond to low latitudes, show relatively few target densities that are less than the lowest estimated density, so there are relatively few minimum-thickness layers. On the other hand, at high latitudes, where there is a deep surface mixed layer, each of the many cyan lines corresponding to unobserved low target densities is reflected as a layer with prescribed minimum thickness.

There is no direct information about the thickness of the layer in which the cast terminates. The cast might barely penetrate, providing no useful information, or it could probe well into that layer and provide useful lower bounds on the density and thickness and an es-

¹² Target densities should be chosen so that none of the estimated density profiles will indicate water that is appreciably lighter than the target value for the top layer or heavier than that for the bottom layer.

¹³ The target densities for the 22 layers are 21.20, 21.60, 22.00, 22.40, 22.90, 23.40, 24.02, 24.70, 25.28, 25.77, 26.18, 26.52, 26.80, 27.03, 27.22, 27.38, 27.52, 27.64, 27.74, 27.82, 27.88, and 27.92. Note that the lightest is not indicated in Fig. 3, nor are the three heaviest.

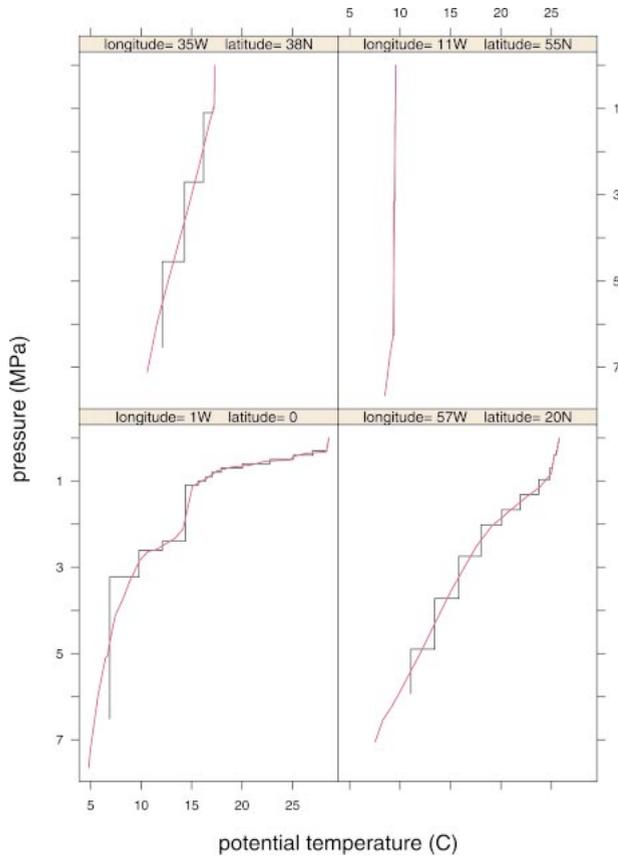


FIG. 3. Estimated θ profiles (magenta) based on observed temperature and on estimated salinity profiles from Fig. 2, and their layer averages (black).

time, perhaps biased, for the potential temperature. An expedient is simply to discard this additional information, but a strategy might be devised for exploiting it. A related issue is what should be done below the bottoms of the XBT casts.¹⁴ Here, for simplicity, no corrections are made to unobserved layer thicknesses. If there is evidence that lower-ocean corrections are needed, this might be handled by estimating extensions to the profiles. Such details are beyond the scope of this work.

Data-error variances should reflect the spatial and temporal variability that can be expected within the model's grid cells in the month of March. Information about climatological variability at standard depths (Antonov et al. 1998; Conkright et al. 1998) can be used to estimate error variance of the observed and inferred profiles at these depths. The error variance for potential temperature is estimated by the corresponding climatological variance for temperature. Error variance for potential density is taken as the square

¹⁴ The problem of what to do below the XBT soundings is similar to that of inferring subsurface conditions from satellite-based altimetric data.

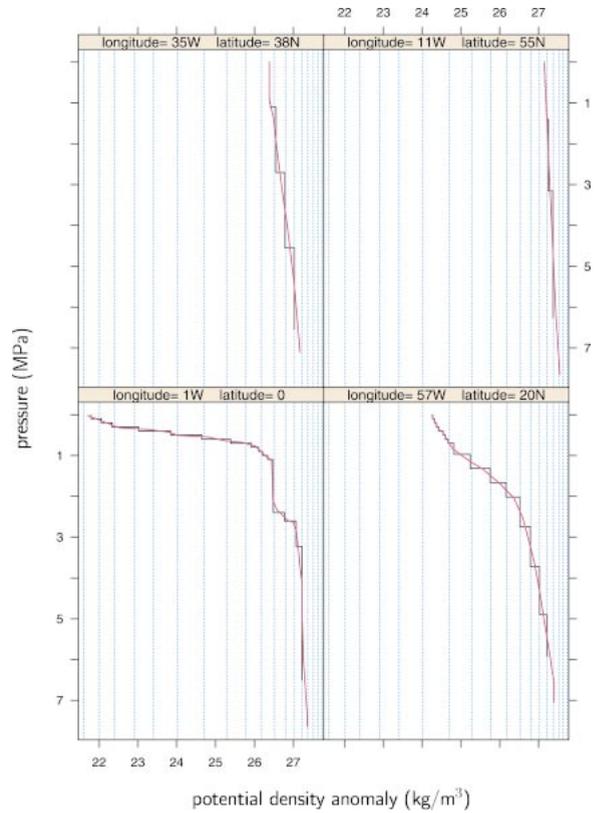


FIG. 4. Estimated σ_θ profiles (magenta) based on estimated S and θ profiles from Figs. 2 and 3 and their layer averages (black). Target values for the model's layers are indicated (cyan) for comparison.

of its standard deviation $\text{var}(\sigma_\theta) = \text{SD}^2(\sigma_\theta)$, which is estimated as¹⁵

$$\text{SD}(\sigma_\theta) = \left(\frac{\partial \sigma_\theta}{\partial \theta} \right) \text{SD}(\theta) + \left(\frac{\partial \sigma_\theta}{\partial S} \right) \text{SD}(S), \quad (4)$$

where the $\text{SD}(\theta)$ and $\text{SD}(S)$ are climatological standard deviations of temperature and salinity, respectively, and where the partial derivatives of the polynomial equation of state are evaluated using the climatological mean values. We neglect uncertainty in pressure due to fall-rate conversion and depth-to-pressure conversion,¹⁶ as errors in representing monthly mean conditions for the grid cell dominate.

Figure 5 shows the climatological variability of po-

¹⁵ If covariances between salinity and potential temperature were readily available, we could have estimated the error variance as $\text{var}(\sigma_\theta) = \text{var}(\theta)(\partial \sigma_\theta / \partial \theta)^2 + \text{var}(S)(\partial \sigma_\theta / \partial S)^2 + 2 \text{cov}(\theta, S)(\partial \sigma_\theta / \partial \theta)(\partial \sigma_\theta / \partial S)$.

¹⁶ A. Mariano (2001, personal communication) has estimated these uncertainties to be about 5%–7%.

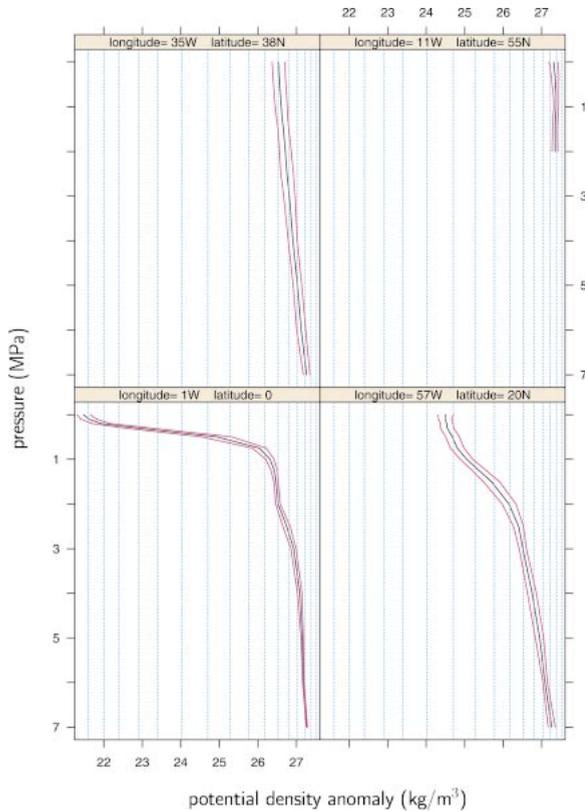


FIG. 5. Estimated uncertainties (magenta) of derived σ_θ profiles relative to profiles (black) based on climatological variability of temperature and salinity. Target values for the model's layers are indicated (cyan) for comparison.

tential density¹⁷ for the locations of the four casts of Figs. 2–4. The central black curves indicate means for March without suppressing inversions resulting from using climatological mean salinity and temperature, and the adjacent curves indicate standard deviations (4). The profiles of Fig. 4 are consistent with these estimates of climatological variability.

The task of assigning uncertainties to the layer-specific data resulting from the preprocessing is complicated by the hybrid nature of the model. When layers are assigned minimum thickness, because water with the target density is not present, the vertical coordinate is pressure-like, so uncertainties are associated with potential density rather than with pressure. On the other hand, when layers are defined by their target densities, uncertainties are associated with thickness rather than with potential density. In that case, the bounding curves

¹⁷ Comparison with Fig. 4 reveals that the climatology at the location of the northernmost of these four casts does not extend to the full depth of the sounding. However, climatology in adjacent cells suggests that the uncertainty remains constant throughout the mixed layer that extends beyond the depth of the sounding. In any case, this illustrates the need to supplement the archived $1^\circ \times 1^\circ$ climatology when estimating both salinity and uncertainties for preprocessing.

in Fig. 5 can give an idea of the uncertainty that should be assigned to thickness. In general, the uncertainty should be partitioned between the two variables to reflect the hybrid nature of the model.

The standard deviation of the error of observed potential density can be interpolated to the middepth $(p_k + p_{k+1})/2$ of layer k , as estimated from the XBT data, to give an initial approximation $SD(\sigma_k)_0$ to the error standard deviation for that layer. However, the final approximation should reflect the fact that the error standard deviation should be small when the layer's potential density is essentially the same as its target value $SD(\sigma_k)_T$:

$$SD(\sigma_k) = \min\{SD(\sigma_k)_0, \max[|(\sigma_k)_0 - (\sigma_k)_T|, SD(\sigma_k)_{\min}]\}, \quad (5)$$

if the observed thickness of layer k is greater than the minimum value specified for HYCOM. Here, we take the minimum error standard deviation $SD(\sigma_k)_{\min}$ to be 0.001 kg m^{-3} . When minimal thickness is assigned to the layer, $SD(\sigma_k)$ is taken to be $SD(\sigma_k)_0$.

Uncertainties must also be estimated for the observed layer thicknesses. Everywhere minimum thickness is assigned, the error standard deviation should be small: here, $SD(\Delta p_k)_{\min}$ is 0.001 MPa , that is, about 10 cm . However, for layers that are assigned target potential densities, the error in thickness should reflect the uncertainties illustrated in Fig. 5. The closer $SD(\sigma_k)$ is to $SD(\sigma_k)_{\min}$, the more $SD(\Delta p_k)$ should account for the variability of the potential density profile. Here, rather than trying to estimate uncertainty in layer thickness from the profiles, we assign a maximum uncertainty¹⁸ $SD(\Delta p_k)_{\max}$ taken as $0.5\Delta p_k$. Then, for layers with thickness greater than HYCOM's minimum allowed value, the error standard deviation of the layer-thickness data is

$$SD(\Delta p_k) = SD(\Delta p_k)_{\max} - [SD(\Delta p_k)_{\max} - SD(\Delta p_k)_{\min}] \times \left(\frac{SD(\sigma_k) - SD(\sigma_k)_{\min}}{SD(\sigma_k)} \right). \quad (6)$$

Uncertainties in estimates of layer thicknesses are certainly related to those of layer densities, and uncertainties in estimates of densities are related to those of temperatures. Similarly, estimating the positions of layer interfaces causes between-layer error correlations. Nevertheless, we have made no effort to account for correlated errors. Instead, we proceed as though all data errors were uncorrelated, taking \mathbf{D} to be diagonal. In the future it might be useful to model its off-diagonal elements.

A large number of XBT casts might provide information for the same grid cell. While the formalism ref-

¹⁸ What is important is the size of the uncertainty relative to the thickness error of the model state. Later we assign the background error relative to this arbitrary maximum.

erees redundancies and conflicts between these data to some extent, we have found it better to combine the data within each cell into “super obs.” For each model cell, values for layer thickness, potential temperature, and potential density from all casts are averaged. This not only reduces the size of the matrix $\mathbf{HBH}^T + \mathbf{D}$, making the computations faster, but also avoids amplification of differences between the individual casts. The super obs are treated as being located at the centers of the cells.

5. An example

Because HYCOM is still under development, our goal is only to illustrate qualitatively how the multivariate assimilation scheme works. The 30-yr test run¹⁹ illustrated by Fig. 1 provides an appropriate context. After being forced by the climatological seasonal cycle of surface fluxes, starting from zonally constant temperature and salinity fields,²⁰ the slowly changing model state can be regarded as an oceanic climatology. For example, the mid-March fields provide background estimates for any particular March. Here, data from March 1995 are used to improve the model’s background estimate so that it reflects March conditions for that particular year.²¹

a. The data

The XBT data used here are the March 1995 data assimilated²² by the National Centers for Environmental Prediction (NCEP) into their operational ocean model (Behringer et al. 1998). Figure 6 compares the 704 March 1995 XBT casts with the National Oceanographic Data Center’s 1998 March climatology (Antonov et al. 1998; Conkright et al. 1998) for four of the depths for which the climatology has been compiled. Each observed temperature has been standardized by subtracting the local climatological mean temperature and then dividing by the standard deviation. The standardized temperatures are plotted versus the latitude of the XBT cast. Horizontal lines have been added at plus and minus three standard deviations to make the identification of unlikely data easier; if the data were distributed normally with climatological means and standard deviations, the probability of data lying outside these lines would be less than 0.27%. Note that many of the suspect data are more than four standard deviations from the mean,

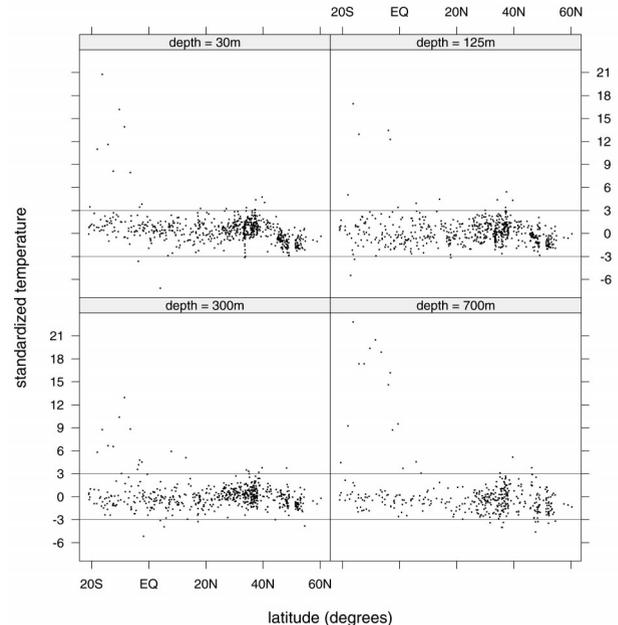


FIG. 6. Departure of observed temperatures from the Mar climatological mean relative to the local climatological standard deviation vs latitude of the XBT cast for four different depths. All observations within three standard deviations of the local climatological mean lie within the two horizontal lines.

which should occur with a probability of less than 0.0064%. Except for the outliers, the data were distributed fairly symmetrically about the climatological means, the exceptions being near the surface south of the equator and around 40°N. The 649 casts with temperatures within three standard deviations of climatology at all levels were retained as candidates for assimilation. Of these, 11 were not used, because during preprocessing the climatology considered their locations to be over land. Another five were discarded because they were in boundary cells just outside of the active model domain, and five more were discarded for having unrealizable negative values for temperature. After this screening process, there were 628 casts to be assimilated.

The spatial distribution of the 628 March 1995 XBT casts is shown in Fig. 7. When compared with that of other months for which XBT data are archived, this coverage is relatively good. To give a feeling for the relationship between the model’s resolution and the XBT sampling, Fig. 8 shows for which grid cells data are available for assimilation. Of the 1579 cells comprising the Atlantic grid, 351 are observed. Some cells have information from several casts, while there is no direct information for many other cells. Preprocessing provides 351 super obs to be assimilated, one for each of the observed cells.

b. The background state

The data-assimilation formalism assumes that the background state is unbiased. This is not the case here.

¹⁹ In this simulation, a year had 360 days and each month had 30 days.

²⁰ Zonal averages of Levitus’ 1994 estimates for January.

²¹ If the forcing would reflect the day-to-day conditions, the same data-assimilation scheme would be appropriate. However, the error covariances would have to characterize the richer level of detail, allowing assimilation to compensate for inadequate forcing and erroneous initial conditions.

²² NCEP do not assimilate observations differing by more than 5°C from their model’s background estimates (Ji et al. 1995).

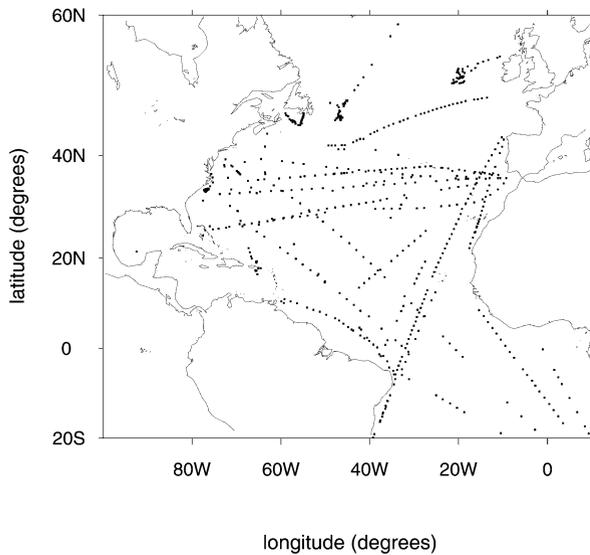


FIG. 7. Dots mark the locations of 628 XBT casts for Mar 1995.

Figure 9 shows the differences between the potential temperature²³ for layer 6, and the climatological March mean temperature for the corresponding grid cell interpolated to the middle of the layer, which is about 55 m below the sea surface everywhere in the domain. The points represent departures from climatology of the 628 casts divided by the local standard deviation for March so that unlikely values are easily recognized. Note that the model has a warm bias near 50°N associated with temperatures that are too high in the Labrador Sea. This bias might be a consequence of the model's treatment of the open northern boundary or of the air–sea heat exchange. Because the model state is from a preliminary developmental run, it is not profitable to pursue such biases further here. However, in the future it will be important to recognize and to eliminate them if at all possible. Sequential data assimilation can serve to reduce the biases by pulling the simulation toward the data, but the optimality of the formalism is compromised, and error covariances are forced to compensate for incorrect error means. It is best to fix the model's tendency to drift away from the data.

While there are no salinity data to assimilate, it is nevertheless important to judge whether the model's treatment of salinity is consistent with climatological data. Figure 10 shows a scatterplot for salinity that is analogous to that of temperature in Fig. 9. Because some of the climatological standard deviations were found to be quite small, the smallest being 0.00422 psu, to avoid exaggerating the difference between model and climatology, standardization was made by dividing the difference by the larger of 0.1 psu and the climatological value. Note that the model values are much too saline

²³ Temperature and potential temperature are essentially the same at this depth.



FIG. 8. Squares indicate the 351 grid cells with data for Mar 1995.

around 50°N and too fresh around 35°N. These large differences, like the temperature differences, suggest systematic problems with the model. Improving the treatment of the open boundaries might help in both cases, but the errors in the treatment of evaporation and precipitation are largely independent from those of surface heat flux. The temperature observations compensate for errors in heat flux, but there are no similar

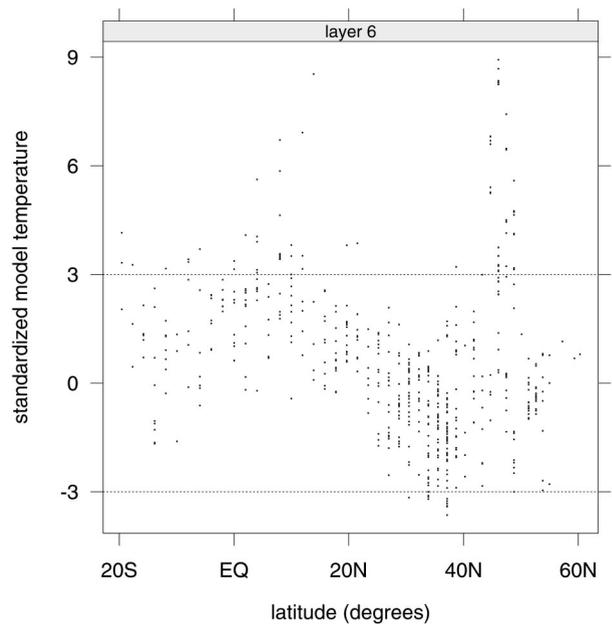


FIG. 9. Departure of model potential temperatures from the Mar climatological mean temperatures relative to the local climatological standard deviation vs latitude for layer 6. Model values within three standard deviations of the local climatological mean lie within the two horizontal lines.

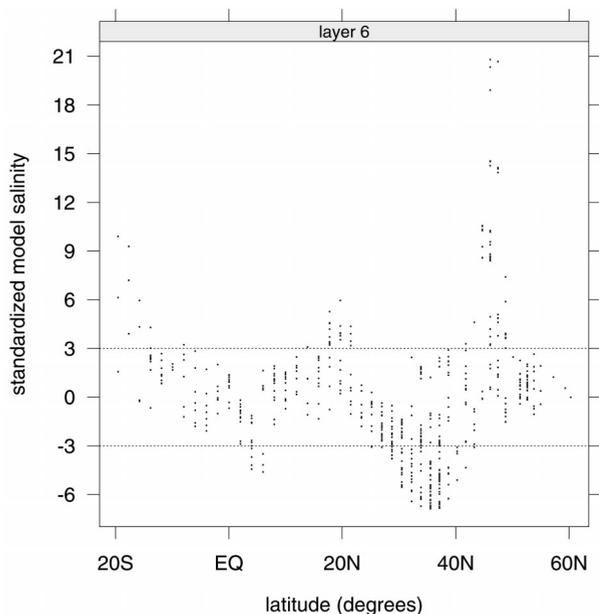


FIG. 10. Departure of model salinities from Mar climatological mean salinities divided by the local climatological standard deviation vs latitude for layer 6.

observations that compensate for errors in water flux. Using companion salinity profiles based on climatology, as suggested here, provides a reasonable basis for correction when the model tends to depart from the usual range of salinities.

c. Background error covariances

The background error correlations between potential temperatures (or potential densities or layer thicknesses) of adjacent grid cells should be appreciable. Here we take the background error covariances for all variables and all layers to be isotropic²⁴ Gaussian functions of the distance

$$B_{\mu,\nu;\xi,k} = A_{\xi,k} \exp\left\{-\frac{1}{2R^2}[(i_\mu - i_\nu)^2 + (j_\mu - j_\nu)^2]\right\}, \quad (7)$$

where i and j are, respectively, latitude and longitude of the grid cells expressed as integer indices; μ and ν indicate the pair of cells for which the covariance is computed; and R is taken to be 2, that is, the two-cell Nyquist cutoff of the computational grid; ξ and k indicate model field and layer, respectively. This horizontal parameterization is consistent with the objective of a low-resolution estimate of monthly conditions. Because HYCOM is still under development and the character of its errors remains to be explored, the factors $A_{\xi,k}$ are chosen so that for most cases the data are con-

sidered more accurate than the model, but the model is considered more accurate for the largest 20% of model–data differences. The factors were taken to be independent of layer: $(1.75^\circ\text{C})^2$ for potential temperature, $(0.5 \text{ kg m}^{-3})^2$ for potential density, and $(0.1 \text{ MPa})^2$ for layer thickness.

d. Results

The 628 XBT casts for March 1995 were preprocessed, as described in section 4, to obtain model–layer-specific data for potential temperature, potential density, and pressure thickness. These data were then assimilated to correct HYCOM’s climatologically forced March state. The corrections were computed layer by layer and field by field, as described in section 3, using the background error covariances described in section 5c to distribute model–data differences away from the observation points. Salinity was estimated from the corrected potential temperature and potential density fields using the equation of state for seawater. Velocity corrections were estimated geostrophically from the pressure corrections. Corrections for interface pressures were estimated by adding the corrected pressure thicknesses, assuming pressure at the sea surface to be zero.

The lower panel of Fig. 1 shows the results of data assimilation for the 25°W section. Note that, while the data add structure to the overly smooth model fields²⁵ in the upper panel, the hybrid nature of the model’s layers is preserved. The thicknesses of the near-surface layers change little, while changes in potential density are evident. And in the thermocline the depths and thicknesses of the layers change, while their densities remain close to the prescribed target values.

Figure 11 shows potential temperature, potential density, and thickness of layer²⁶ 6 before and after assimilation; corresponding diagnosed fields are shown in Fig. 12. As expected, thickness and depth of layer 6 are largely unchanged, except in the Tropics where the surface mixed layer is relatively shallow. On the other hand, the potential density of layer 6 is modified significantly. Note that after assimilation the water is less dense off the east coast of North America. The XBTs provide direct information about potential temperature, so changes to the model’s potential temperature field is no surprise. Near 45°N and 60°W, where the model was too warm, there is a dramatic cooling. A first impression is that this cooling should be associated with an increase in density, not the decrease that is shown here. In fact, density would have been increased, if the background

²⁴ The background error covariance is not homogeneous but rather reflects the inhomogeneities of the Mercator grid.

²⁵ The model fields were smooth, because they were computed starting from zonally averaged climatological means and were forced by seasonal climatologies of surface fluxes. The inferred changes are real to the extent that (i) the temperature data represent the monthly conditions and (ii) the March climatological salinity represents the salinity field for March 1995.

²⁶ Layers are numbered from top to bottom.

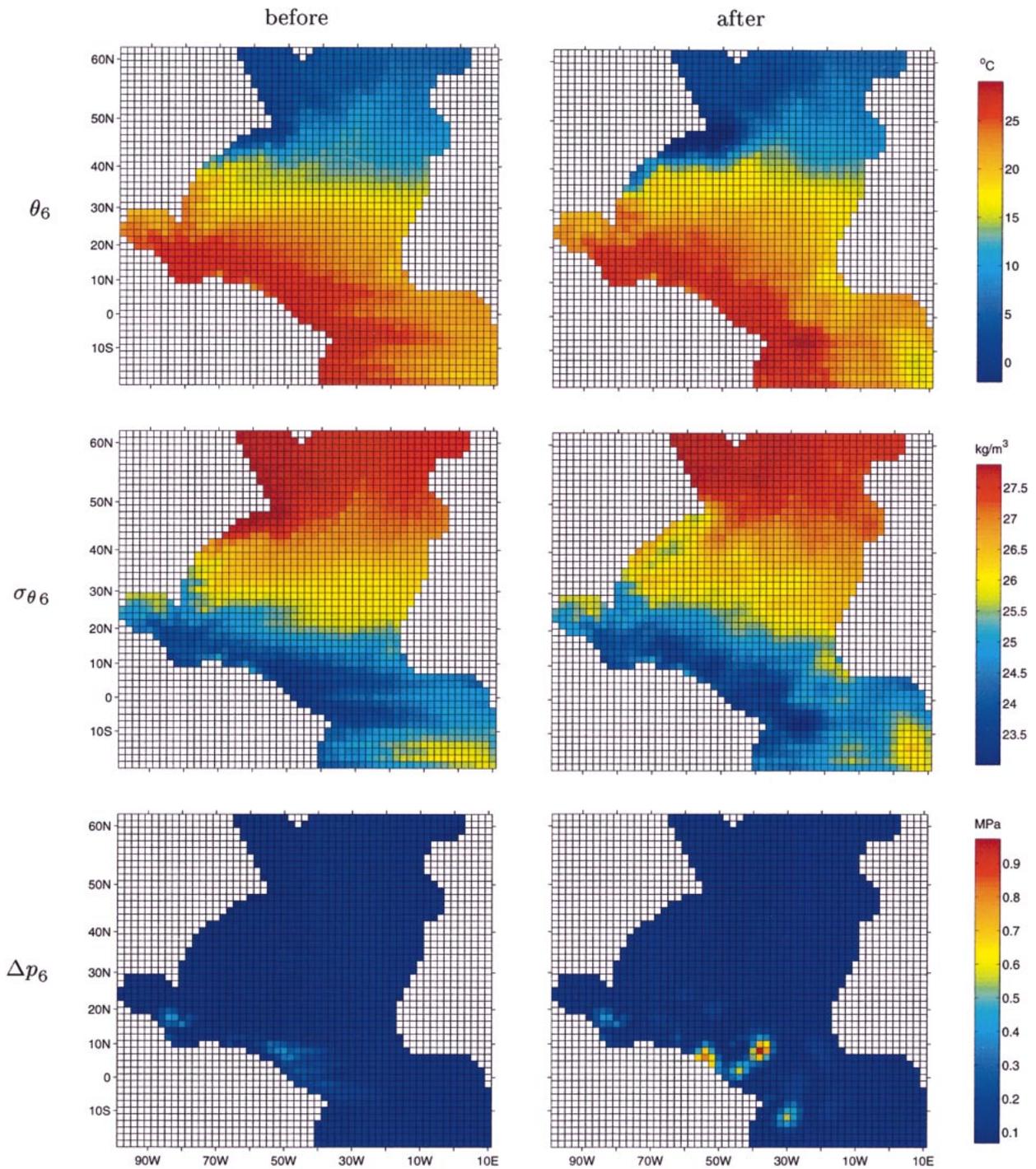


FIG. 11. Background (left) and corrected (right) potential temperature (upper), potential density (middle), and thickness (lower) for layer 6.

salinity had been used for estimating potential density,²⁷ and the “after” would have been the same as the “be-

²⁷ Because density corrections in this region depend so strongly on the salinity strategy, it would be wise to try to improve on climatological salinity using a method like that of Hansen and Thacker (1999).

fore” for salinity in Fig. 12. The actual salinity corrections reflect the differences between model and climatology shown in Fig. 10. These differences argue that preprocessing should use estimates of salinity at least as good as those from climatology.

Figures 13 and 14 show the results for layer 12. As

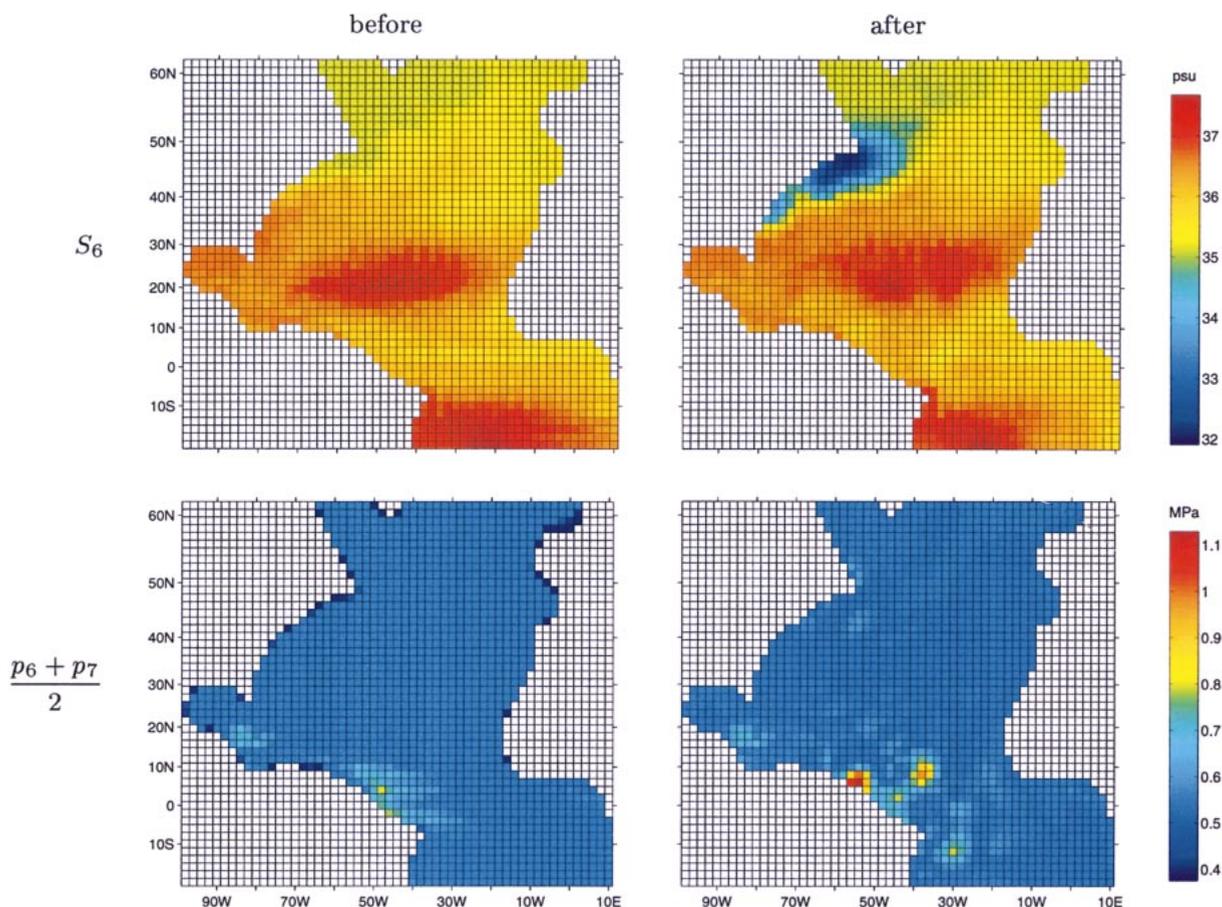


FIG. 12. Background (left) and corrected (right) salinity (upper) and pressure (lower) for layer 6.

expected, in the midlatitudes where this layer is in the pycnocline, its thickness and depth are modified by the XBT data, while its potential density is largely unchanged. The small changes of potential density in the Tropics appear to be associated with the preprocessing algorithms, which give data for layer 12 that tend to be slightly lighter than the target value. South of the Labrador Sea, corrections to potential temperature, potential density, and salinity are similar to those for layer 6, again illustrating the importance of a salinity strategy.

Because our focus is principally on adjusting the mass distribution, we implemented the simplest possible corrections to the velocity field. The corrections were determined geostrophically from the corrections to the pressure field, except near the equator where no corrections were made. Results were quite reasonable.²⁸ In subsequent work, when the data-assimilation codes are coupled to the model codes, our plan is to correct the velocity field using the model codes to bring the flow into dynamic adjustment, as discussed in section 3.

²⁸ As the figures illustrating the changes were uninteresting, they are not included.

6. Conclusions

The procedure presented here for assimilating thermal data from XBTs into HYCOM works as anticipated. It corrects depth and thickness of the model's layers, as well as potential temperature, potential density, salinity, and velocity. In accordance with the hybrid-coordinate philosophy, assimilation preserves the pressure-like nature of the vertical coordinate in the well-mixed surface layer and its density-like nature in the pycnocline. Central to its success is the preprocessing of the observed temperature profiles to estimate data pertaining to the model's layers. The most important step is estimating companion profiles of salinity, which are used in constructing potential density profiles needed for determining the depths and thicknesses of the model's layers. Once the layer-specific data have been determined, they can be assimilated using a variety of techniques. At this stage of HYCOM's development, optimal interpolation is a reasonable choice.

This approach can be generalized to accommodate other types of data. High quality coincident profiles of temperature, salinity, and density from CTD probes, while few in number, could be easily accommodated.

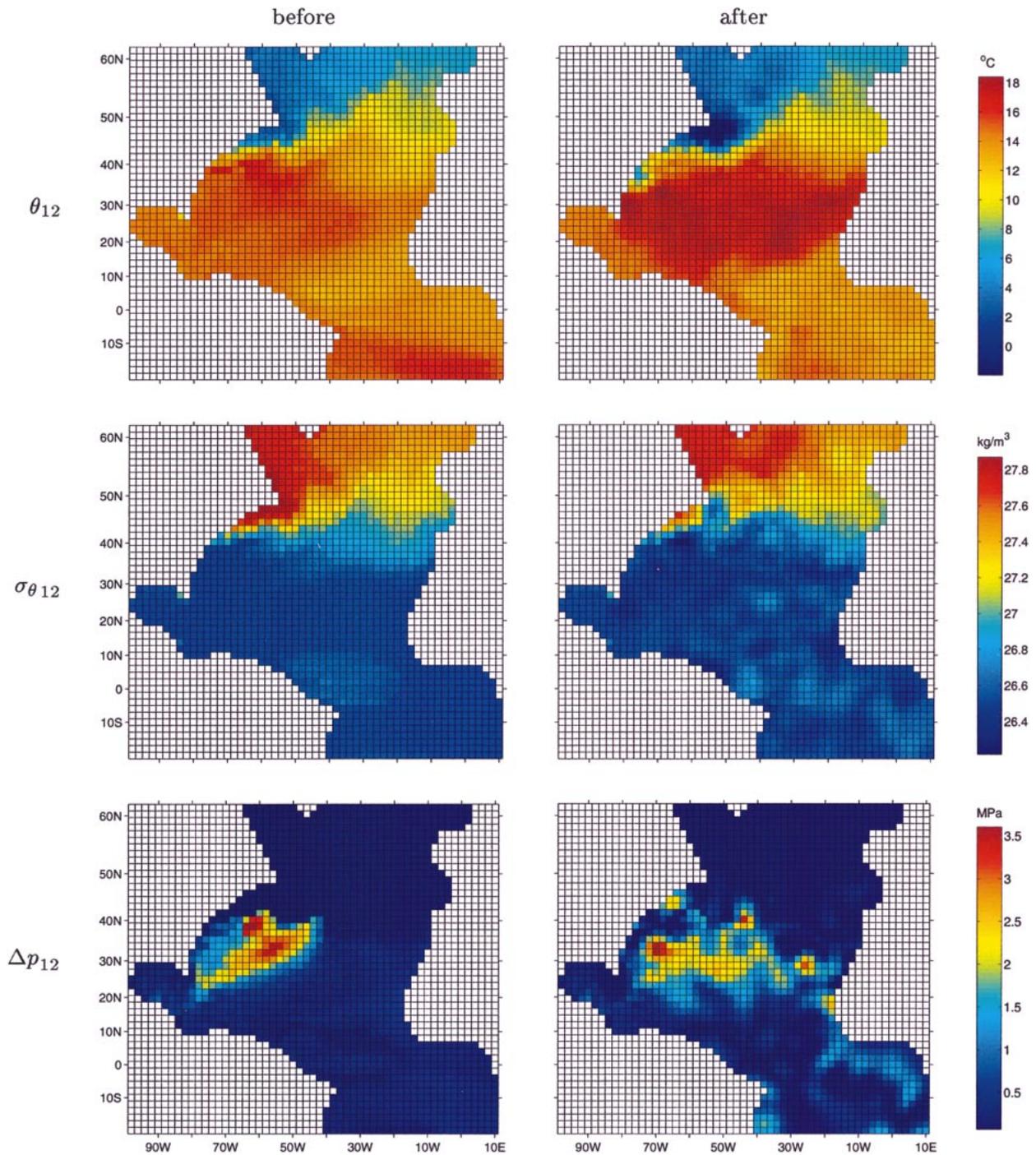


FIG. 13. Background (left) and corrected (right) potential temperature (upper), potential density (middle), and thickness (lower) for layer 12.

Profiling floats, which might provide a substantial amount of data in the near future, also fit easily into this framework. Surface temperatures observed from satellites could be treated as “short XBT casts” to provide additional information about the model’s upper layer. Altimetric data would require more attention. Be-

cause sea surface elevation reflects the thermal state of the underlying water as a function of depth, there is the question of how to infer subsurface corrections from surface discrepancies.

As optimal interpolation bases its corrections on covariances of errors of the model state, this assimilation

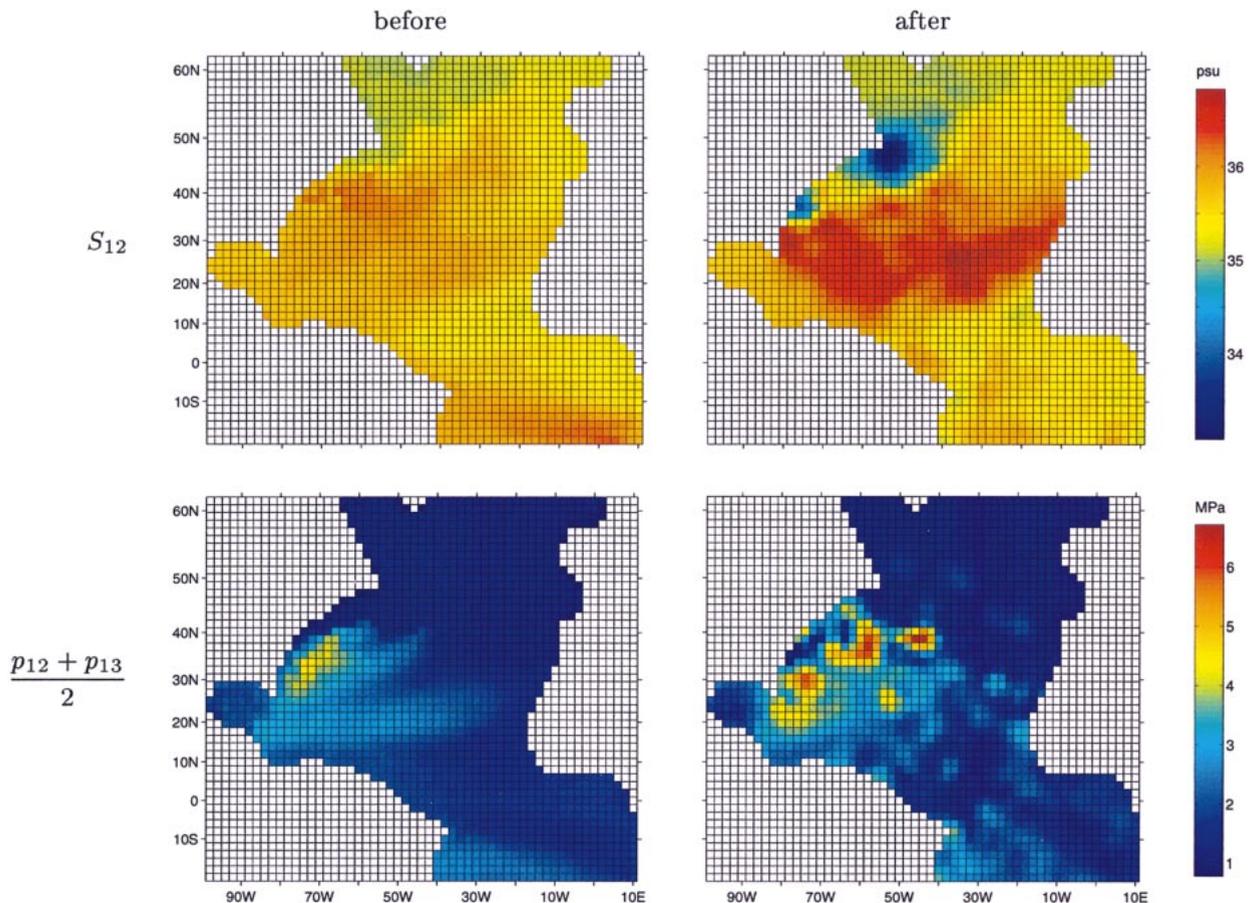


FIG. 14. Background (left) and corrected (right) salinity (upper) and pressure (lower) for layer 12.

procedure can best be improved by improving the way error covariances are modeled. In particular, more attention should be given to parameterizing the local nature of the errors. The careful model–data comparisons that are needed to model the error covariances should also promote improvements to the model to correct for systematic discrepancies.

Acknowledgments. This work was supported by the National Oceanographic Partnership Program and by the Atlantic Oceanographic and Meteorological Laboratory.

REFERENCES

- Antonov, J. I., S. Levitus, T. P. Boyer, M. E. Conkright, T. D. O'Brien, and C. Stephens, 1998: *Temperature of the Atlantic Ocean*. Vol. 1, *World Ocean Atlas 1998*, NOAA Atlas NESDIS 27, 166 pp.
- 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.
- Bengtsson, L., M. Ghil, and E. Källén, Eds., 1981: *Dynamic Meteorology: Data Assimilation Methods*. Academic Press, 330 pp.
- Bleck, R., 1998: Ocean modeling in isopycnic coordinates. *Ocean Modeling and Parameterization*, E. P. Chassignet and J. Verron, Eds., NATO Science Series, Vol. 516, Kluwer Academic, 423–448.
- , 2002: An oceanic general circulation model framed in hybrid isopycnic–Cartesian coordinates. *Ocean Modell.*, **4**, 55–88.
- , and D. B. Boudra, 1981: Initial testing of a numerical ocean circulation model using a hybrid (quasi-isopycnic) vertical coordinate. *J. Phys. Oceanogr.*, **11**, 755–770.
- , and S. G. Benjamin, 1993: Regional weather prediction with a model combining terrain-following and isentropic coordinates. Part I: Model description. *Mon. Wea. Rev.*, **121**, 1770–1785.
- , and E. P. Chassignet, 1994: Simulating the oceanic circulation with isopycnic coordinate models. *The Oceans: Physiochemical Dynamics and Resources*, S. K. Majumdar et al., Eds., Pennsylvania Academy of Science, 17–39.
- Bryan, K., 1984: Accelerating convergence to equilibrium of ocean-climate models. *J. Phys. Oceanogr.*, **14**, 666–673.
- Bryden, H. L., 1973: New polynomials for thermal expansion, adiabatic temperature gradient and potential temperature of sea water. *Deep-Sea Res.*, **20**, 401–408.
- Brydon, D., S. Sun, and R. Bleck, 1999: A new approximation of the equation of state for seawater suitable for numerical ocean models. *J. Geophys. Res.*, **104**, 1537–1540.
- Carnes, M. R., W. J. Teague, and J. L. Mitchell, 1994: Inference of subsurface thermocline structure from fields measurable by satellite. *J. Atmos. Oceanic Technol.*, **11**, 551–566.
- Carton, J. A., and E. C. Hackert, 1990: Data assimilation applied to the temperature and circulation in the tropical Atlantic, 1983–84. *J. Phys. Oceanogr.*, **20**, 1150–1165.
- , G. Chepurin, X. Cao, and B. Giese, 2000: A simple ocean data assimilation analysis of the global upper ocean 1950–95. Part I: Methodology. *J. Phys. Oceanogr.*, **30**, 294–309.

- Cohn, S. E., 1997: An introduction to estimation theory. *J. Meteor. Soc. Japan*, **75**, 257–288.
- Conkright, M., and Coauthors, 1998: World Ocean Database 1998. NOAA National Oceanographic Data Center Internal Rep. 14, 114 pp.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- Dee, D. P., and A. M. da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.*, **124**, 269–295.
- Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347.
- Dickenson, R. R., J. Meincke, S.-A. Malmberg, and A. J. Lee, 1988: The “Great Salinity Anomaly” in the northern North Atlantic 1968–1982. *Progress in Oceanography*, Vol. 20, Pergamon, 103–151.
- Evensen, G., 1994: Inverse methods and data assimilation in nonlinear ocean models. *Physica D*, **77**, 108–129.
- , D. P. Dee, and J. Schröter, 1998: Parameter estimation in dynamical models. *Ocean Modeling and Parameterization*, E. P. Chassignet and J. Verron, Eds., NATO Science Series, Vol. 516, Kulwar Academic, 337–398.
- Fofonoff, N. P., 1977: Computation of potential temperature of seawater for an arbitrary reference pressure. *Deep-Sea Res.*, **24**, 489–491.
- Gandin, L. S., 1963: *Applied Optimal Estimation*. Israel Program for Scientific Translations, 184 pp. (Translated from the Russian.)
- Gelb, A., Ed., 1974: *Objective Analysis of Meteorological Fields*. Springer-Verlag, 374 pp.
- Gill, P. E., W. Murray, and M. H. Wright, 1981: *Practical Optimization*. Academic Press, 401 pp.
- Hansen, D. V., and W. C. Thacker, 1999: On estimation of salinity profiles in the upper ocean. *J. Geophys. Res.*, **104**, 7921–7933.
- Hasselmann, K., 1993: Optimal fingerprints for the detection of time-dependent climate change. *J. Climate*, **6**, 1957–1971.
- Ji, M., A. Leetmaa, and J. Derber, 1995: An ocean analysis system for climate studies. *Mon. Wea. Rev.*, **123**, 460–481.
- Kalman, R. E., 1960: A new approach to linear filtering and prediction problems. *Trans. ASME, J. Basic Eng.*, **D82**, 35–45.
- Large, W. G., J. C. McWilliams, and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, **32**, 363–403.
- Levitus, S., R. Burgett, and T. P. Boyer, 1994: *Salinity*. Vol. 3, *World Ocean Atlas 1994*, NOAA Atlas NESDIS 3, 99 pp.
- Lorenc, A. C., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Mon. Wea. Rev.*, **109**, 701–721.
- , 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177–1194.
- Malanotte-Rizzoli, P., Ed., 1996: *Modern Approaches to Data Assimilation in Ocean Modeling*. Elsevier, 455 pp.
- Saunders, P. M., and N. P. Fofonoff, 1976: Conversion of pressure to depth in the ocean. *Deep-Sea Res.*, **23**, 109–111.
- Thacker, W. C., 1989: The role of the Hessian matrix in fitting models to measurements. *J. Geophys. Res.*, **94**, 6177–6196.
- , 1996: Climatic fingerprints, patterns, and indices. *J. Climate*, **9**, 2259–2261.
- , and R. B. Long, 1988: Fitting dynamics to data. *J. Geophys. Res.*, **93**, 1227–1240.
- , and R. Lewandowicz, 1994: Dynamics of information and uncertainty. *Tellus*, **46A**, 651–670.