

A Simple Ocean Data Assimilation Analysis of the Global Upper Ocean 1950–95. Part I: Methodology

JAMES A. CARTON, GENNADY CHEPURIN, AND XIANHE CAO

Department of Meteorology, University of Maryland at College Park, College Park, Maryland

BENJAMIN GIESE

Department of Oceanography, Texas A&M University, College Station, Texas

(Manuscript received 26 January 1998, in final form 10 February 1999)

ABSTRACT

The authors describe a 46-year global retrospective analysis of upper-ocean temperature, salinity, and currents. The analysis is an application of the Simple Ocean Data Assimilation (SODA) package. SODA uses an ocean model based on Geophysical Fluid Dynamics Laboratory MOM2 physics. Assimilated data includes temperature and salinity profiles from the *World Ocean Atlas-94* (MBT, XBT, CTD, and station data), as well as additional hydrography, sea surface temperature, and altimeter sea level.

After reviewing the basic methodology the authors present experiments to examine the impact of trends in the wind field and model forecast bias (referred to in the engineering literature as “colored noise”). The authors believe these to be the major sources of error in the retrospective analysis. With detrended winds the analysis shows a pattern of warming in the subtropics and cooling in the Tropics and at high latitudes. Model forecast bias results partly from errors in surface forcing and partly from limitations of the model. Bias is of great concern in regions of thermocline water-mass formation. In the examples discussed here, the data assimilation has the effect of increasing production of these water masses and thus reducing bias.

Additional experiments examine the relative importance of winds versus subsurface updating. These experiments show that in the Tropics both winds and subsurface updating contribute to analysis temperature, while in midlatitudes the variability results mainly from the effects of subsurface updating.

1. Introduction

Understanding the ocean’s role in climate variability requires an appreciation of the historical changes occurring in the ocean, particularly in its upper layers. In this paper we describe the Simple Ocean Data Assimilation (SODA) package and its application to the construction of a retrospective analysis of the ocean. Our goals are twofold: first, to develop methodology to explore some of the technical aspects of data assimilation and, second, to construct a long, nearly 50-year retrospective analysis of the temperature, salinity, and circulation of the upper layers of the ocean.

In developing this package we have attempted to choose simplicity over innovation where possible. The general circulation ocean model on which the analysis is based uses the Geophysical Fluid Dynamics Laboratory Modular Ocean Model 2.b code, with conven-

tional choices for mixing, etc. The constraint algorithm we apply is an extension of optimal interpolation data assimilation, an algorithm that has been widely applied in meteorological numerical weather forecasting. Our extension includes changes in the spatial dependence of the statistics and in the assumption of bias in the model forecast. Still, our approach is less sophisticated than the computationally intensive Kalman filter that actually predicts the temporal and spatial evolution of the error statistics. Optimal interpolation data assimilation bears a close similarity to variational methods as well, when the constraint being minimized is the mean square error. A comprehensive discussion of the alternatives is provided by Bennett (1990), Wunsch (1996), and Malanotte-Rizzoli (1996).

The domain of this analysis extends from 62°S to 62°N. The main datasets used to constrain the model are the hydrographic data contained in the *World Ocean Atlas 1994* (WOA-94; Levitus and Boyer 1994); additional hydrographic data, satellite, and in situ sea surface temperature (Reynolds and Smith 1994); and altimetry from the Geosat, *ERS-1*, and TOPEX/Poseidon satellites. The historical dataset is sufficiently limited that we have made no attempt to resolve midlatitude me-

Corresponding author address: Dr. James A. Carton, Department of Meteorology, University of Maryland at College Park, 2417 Computer and Space Science Building, College Park, MD 20742-2425.
E-mail: carton@metosrv2.umd.edu

noscale variability and have limited our attention to the upper 500 m of the water column. The control analysis and analysis experiments begin January 1950 and continue through December 1995. The basic analysis fields, thus, consist of 552 monthly averages of temperature, salinity, sea level, and the horizontal components of velocity. In order to simplify presentation of the results, most of the discussion in this paper focuses on heat content defined as the vertical integral of temperature 0/500 m.

In a companion study Carton et al. (2000) examine the accuracy of the SODA analysis by comparison to a variety of independent observations. The independent observations include tide-gauge sea level, altimeter sea level (when it is excluded from the updating dataset), hydrographic sections, station temperature and salinity time series, surface drifters, and moored currents. In summary, these comparisons show that the analysis explains 25%–35% of the observed tide-gauge sea level variance at longer than annual frequencies. The root-mean-square difference of observed minus analysis sea level in the Tropics (15°S–15°N) lies in the range of 3.1–4.0 cm, increasing somewhat in midlatitudes. A second study (Chepurin and Carton 1999) compares the control analysis temperature with that of six other basin-to global-scale analyses (White 1995; Levitus et al. 1994; Rosati et al. 1995; Smith 1995; Ji et al. 1995). The comparison reveals good agreement of heat content in the tropical Pacific (correlations exceeding 80%) and reasonable agreement in the North Pacific. Elsewhere the agreement is less complete (e.g., Acero-Schertzer et al. 1997).

Central to formulating a data assimilation algorithm are the assumptions made about the observation and forecast error statistics. In both the meteorological and oceanographic cases it is normally assumed that the model forecasts are unbiased [the problem of bias in meteorological models is described in Thiébaux and Morone (1990)]. In reality the presence of slowly evolving errors in the forecast is, we believe, the major source of error for the climate problem. There are several causes of this forecast bias: insufficient resolution, inadequate modeling of unresolved physics, and biases in surface forcing fields of wind, heat, and freshwater. An important example of the effects of unresolved physics is the inability of this and most coarse-resolution models to simulate properly the separation of the Gulf Stream. To understand the impact of neglecting forecast bias we examine the change in the analysis that happens when this assumption is eliminated (expt 1 in Table 1). Variability enters the analysis either through surface forcing or through the updating procedure. The relative importance of these two sources is examined, as well as the forecast model's contribution to the analysis in a series of four experiments. Additional experiments examine the impact of subsurface salinity and assumptions about the relationship between temperature and salinity.

TABLE 1. Ocean analysis experiments presented in the text. Each experiment covers the period 1950–96.

Experiment	Description
Control analysis	Basic analysis with detrended winds
1	Basic analysis except assuming significant forecast error bias
2	Simulation with no subsurface updating
3	Basic analysis with climatological monthly winds
4	Basic analysis with complete winds
5	Basic analysis except with salinity updating with observed salinity, but without <i>T/S</i> error covariance
6	Basic analysis except without <i>T/S</i> error covariance
7	Basic analysis except replacing model with clim. monthly temperature

2. Data

In the period before the mid-1980s the main datasets we have to provide constraints on the subsurface density field are profile measurements of temperature from mechanical bathythermograph (MBT), expendable bathythermograph (XBT), conductivity–temperature–depth (CTD), measurements from thermistors and reversing thermometers and salinity from CTD and station measurements. Since late 1986 satellite altimetry has added an additional powerful constraint on column-averaged density. However, even during this later period, subsurface measurements are required in order to determine the vertical distribution of density variations implied by the altimetry. The sources of these data are described below.

a. Hydrography

We began with the datasets contained in *WOA-94*. From *WOA-94* we extracted temperature and salinity data records obtained by four different instrument types: MBTs, XBTs, CTDs, and salinity–temperature–depth probes (combined with CTDs in Fig. 1), and station data. The temporal distribution of the datasets is shown in Fig. 1a. Until the late 1960s MBTs provide the majority of upper-ocean temperature profiles, while ocean stations (Nansen and Niskin bottle measurements of salinity and reversing thermometer measurements of temperature) provide the majority of salinity and temperature measurements below 200 m. After 1968 XBTs replace MBTs as the major dataset, while the number of station observations begins to fall off. The change to XBTs means a substantial increase in data between 200 and 450 m as well. The spatial coverage of the subsurface temperature data is presented in Fig. 1b. The observations are mainly limited to the Northern Hemisphere and are concentrated along commercial shipping routes.

The error characteristics of the instruments vary widely. The MBT provides a direct measurement of pressure. The XBT instruments, on the other hand, infer depth from drop rate and thus are subject to random and systematic errors that increase with depth (the random error

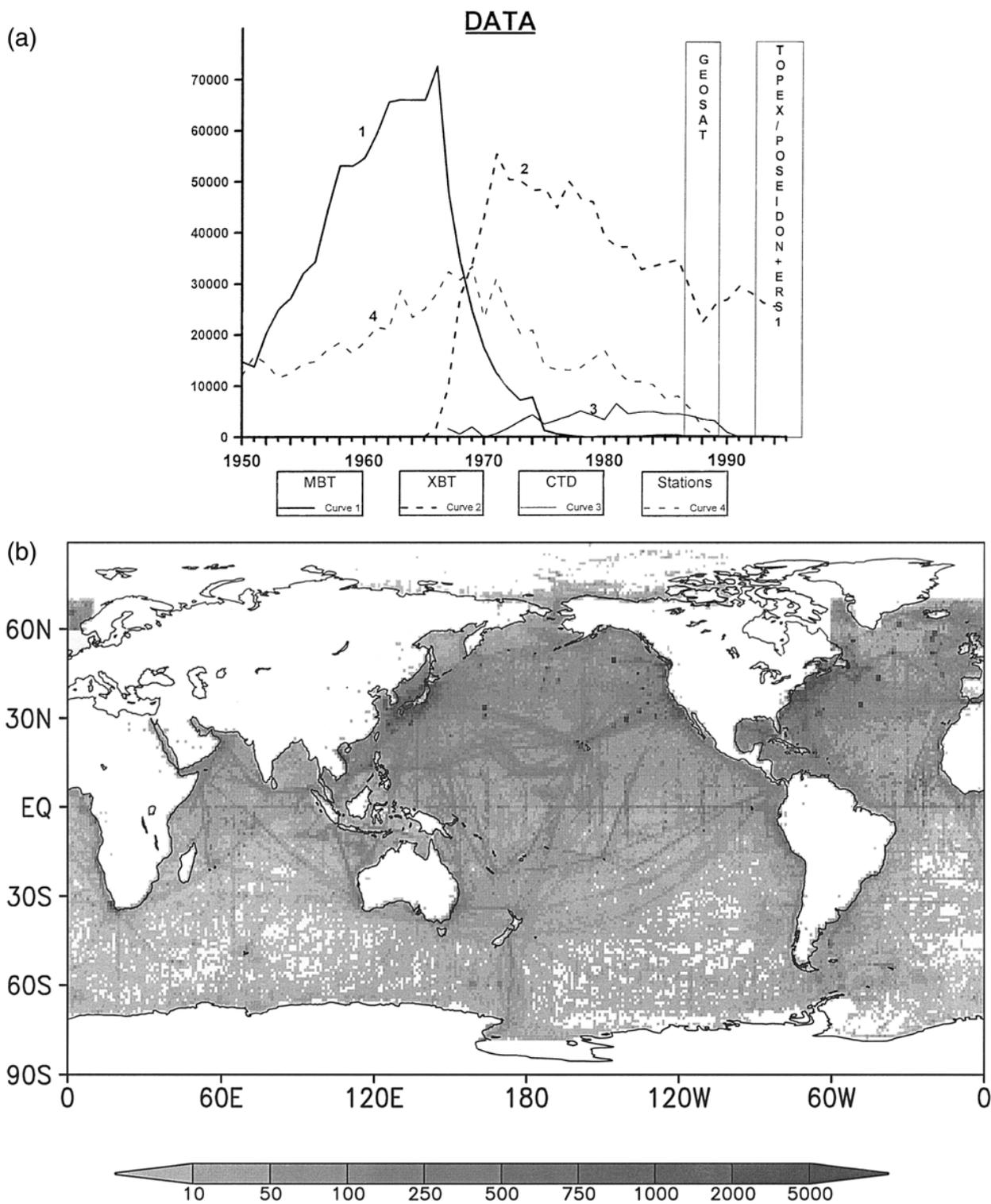


FIG. 1. Distribution of (a) combined subsurface temperature observations at 75 m with time. The dates of the recent altimeter satellites are indicated. Distribution of (b) the combined subsurface temperature observations with latitude and longitude. Shading indicates density.

is generally assumed to be 1%–2% of depth). Systematic errors in the drop rate of XBTs depend on the type, and even the batch, of the XBT. Here we have applied the empirical corrections of Hanawa et al. (1994) to try to limit this source of bias. CTDs have temperature measurement errors sufficiently smaller than errors associated with unresolved physical processes that they may be neglected. The same may be true of the XBT and MBT measurement errors (S. Levitus 1998, personal communication).

The data contained in *WOA-94* is limited mainly to the years prior to 1991. We have supplemented this with data from the XBT archive of the National Oceanographic Data Center (Searle 1992) and the Coupled Modeling Project of NOAA's National Centers for Environmental Prediction (NCEP). We have also included temperature data produced by the Tropical Ocean Global Atmosphere Tropical Atmosphere–Ocean (TOGA–TAO) moored thermistor array. Additional data is provided by a variety of research programs, including, for example, the Soviet SECTIONS tropical program and the Western Tropical Atlantic Experiment. Data from the World Ocean Circulation Experiment hydrographic surveys has been withheld to provide independent comparison (Carton et al. 2000).

With such a wide variety of datasets, quality control becomes extremely important. Much of the combined dataset was checked prior to entry into *WOA-94*. We found it useful to complete our own additional checks for duplicate reports and errors in the recorded position and time of observations. We also check each profile for static stability and for the extent of its deviation from climatology, including the relationship between temperature and salinity. Observations differing from climatology by more than four standard deviations were assumed to be in error and thus were excluded. Altogether these data checks eliminated approximately 10% of the data. Of the criteria we applied, the most restrictive were the comparison to climatology and the tests for static stability. We believe the most serious problems with the in situ observations result from mistakes in station location reports.

The analysis is carried out at each of the model levels between the surface and 444 m. The profile data is interpolated onto these model levels using a quadratic interpolation procedure (profiles were eliminated if the vertical resolution was insufficient). The main reason for this limitation is the lack of data at deeper model levels. After vertical interpolation the temperature observations were binned into $1^\circ \times 1^\circ \times 5$ day bins. The resulting *superobservations* form the basic temperature and salinity datasets used in this study.

b. Altimetry

The altimetry used in this study is the Pathfinder Project version 2.1 (Koblinsky 1997, personal communication) and has been obtained from the Pathfinder Web

TABLE 2. Satellite altimetry used in this study.

Satellite	Days/repeat cycle	Revolutions/ repeat cycle	Span of data
Geosat ERM			
<i>ERS-1</i> (35-day repeat)	17.05	244	11/8/86–9/30/89
	35.00	501	4/14/92–12/20/93
TOPEX/Poseidon	9.92	127	9/30/92–11/28/96

site. The altimetry comes from four different instruments flying on Geosat, *ERS-1*, and the TOPEX/Poseidon platforms. The orbital characteristics of the altimeters are described in Table 2. Each instrument and platform requires independent calibration procedures, most of which were carried out as part of the Pathfinder processing. We have added the usual corrections for geophysical effects and then averaged sea level estimates into 1° latitude segments. The once-per-revolution harmonic has been removed from the Geosat altimetry because of contamination by planetary-scale error. No additional filtering or interpolation has been performed.

3. Analysis methodology

a. Assimilation methodology

In this section we briefly describe our estimation algorithm and associated statistical models. The theoretical basis for optimal interpolation is described in Daley (1991), while extensions to account for bias in the forecast model are described in Dee and da Silva (1997, submitted to *Quart. J. Roy. Meteor. Soc.*). We begin by assuming that a forecast \mathbf{w}_k^f exists at time t_k . Here \mathbf{w}_k^f is a state vector containing all model variables at each grid point; \mathbf{w}_k^f is provided by integration of a model based on initial conditions at time t_{k-1} . We assume \mathbf{w}_k^f differs from the “true” state \mathbf{w}_k due to the presence of model bias \mathbf{g}_k^f and random Gaussian distributed forecast error \mathbf{e}_k^f :

$$\mathbf{w}_k^f = (\mathbf{w}_k^f - \mathbf{g}_k^f) - \mathbf{e}_k^f. \quad (1)$$

The quantity in parentheses is referred to as the bias-corrected forecast: $\tilde{\mathbf{w}}_k^f \equiv \mathbf{w}_k^f - \mathbf{g}_k^f$. Throughout this discussion we use a tilde to indicate a bias-corrected variable. All observations (of temperature, salinity, sea level, etc.) are collected into an observation vector \mathbf{w}_k^o of length p_k at time t_k . We define the observation error \mathbf{e}_k^o to be the difference between the observed and true state interpolated to the observation location $\mathbf{w}_k^o - \mathbf{h}_k(\mathbf{w}_k^f)$; \mathbf{e}_k^o includes measurement error, error of representativeness, and error due to unresolved physics, and \mathbf{h}_k is the interpolation operator that maps the analysis variable type into the observation variable type at the observation location and time. We immediately approximate this equation as

$$\mathbf{e}_k^o \equiv \mathbf{w}_k^o - \mathbf{H}_k \mathbf{w}_k^f,$$

where

$$\mathbf{H}_k = \left. \frac{\partial \mathbf{h}_k}{\partial \mathbf{w}} \right|_{\mathbf{w}_k^f}.$$

Now we summarize the analysis equations. Minimizing the mean square error leads to the interpolation equation for the analysis

$$\mathbf{w}_k^o = \tilde{\mathbf{w}}_k^f + \mathbf{K}_k[\mathbf{w}_k^o - \mathbf{H}_k \tilde{\mathbf{w}}_k^f]. \quad (2)$$

The forecast bias is assumed to be slowly evolving with time. We can anticipate, for example, that there will be a substantial seasonal dependence to the bias because of errors in the formation of mixed layers. The certainty with which we know the bias deteriorates even more rapidly with time. To account for these time dependencies we will use as our first guess of the bias at t_k a fraction of the bias at the previous update time (in this study, 10 days earlier):

$$\mathbf{g}_k = \mu \mathbf{g}_{k-1}. \quad (3)$$

After examining the results of a few experiments we have chosen a value of $\mu = 1/2$, making the first guess of the bias decay with a timescale of roughly 20 days. We begin at t_o by assuming zero bias. The bias is updated using equations similar to the analysis equation

$$\tilde{\mathbf{g}}_k = \mathbf{g}_k - \mathbf{L}_k[\mathbf{w}_k^o - \mathbf{H}_k \tilde{\mathbf{w}}_k^f]. \quad (4)$$

The gain matrices \mathbf{K}_k and \mathbf{L}_k depend on the forecast, observation, and bias error covariances

$$\mathbf{P}_k^f = \langle \mathbf{e}_k^f (\mathbf{e}_k^f)^T \rangle; \quad \mathbf{P}_k^o = \langle \mathbf{e}_k^o (\mathbf{e}_k^o)^T \rangle;$$

$$\mathbf{P}_k^b = \langle (\mathbf{g}_k^f - \langle \mathbf{g}_k^f \rangle) (\mathbf{g}_k^f - \langle \mathbf{g}_k^f \rangle)^T \rangle,$$

as

$$\mathbf{K}_k = (1 - \alpha) \mathbf{P}_k^f \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k]^{-1} \quad (5a)$$

$$\mathbf{L}_k = \alpha \mathbf{P}_k^b \mathbf{H}_k^T [\mathbf{H}_k \mathbf{P}_k^b \mathbf{H}_k^T + \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k]^{-1}. \quad (5b)$$

The parameter α determines what part of the total error is random and what part is due to bias. A value of $\alpha = 0$ corresponds to the assumption that all error is random, whereas a value of 1 corresponds to the assumption that all error is due to bias. For the sake of computational efficiency (5a,b) are currently solved locally in a series of patches spanning 25 horizontal grid points each. The global solution is then constructed by assembling the patches.

In the control analysis we choose the usual assumption of no bias and set $\alpha = 0$. The dependence on this assumption is examined in experiment 1. In this experiment we assume $\alpha = 0.7$ following the recommendation of Dee and da Silva (1997, submitted to *Quart. J. Roy. Meteor. Soc.*) based on studies with a shallow-water model. The result of assuming the existence of bias is for the analysis fields to be shifted in regions of high data coverage to correct for forecast bias. This result is illustrated in Fig. 2 at a point in the central North Pacific (49°N, 184°W). Here the bias correction

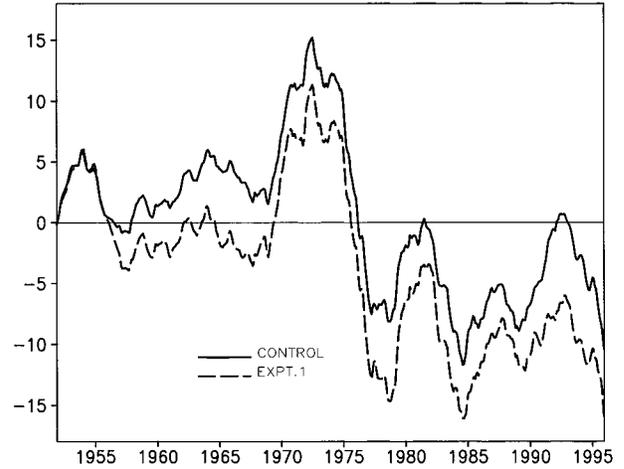


FIG. 2. Influence of bias on analysis. Sea level from control analysis ($\alpha = 0$) and expt 1 ($\alpha = 0.7$) in the central North Pacific (49°N, 184°W). The average difference is 3.9 cm.

algorithm has the effect of lowering the analysis sea level by about 2–4 cm. The drop in sea level is accompanied by an elevation of the thermocline by up to 20 m.

The effect on temperature and salinity accuracy of the bias correction algorithm can be examined by comparing the control analysis and experiment 1 with observations at station S in the northwest Atlantic (32°N, 64°W). At this station the impact of the bias correction at 444 m (within the thermocline) is to reduce the mean temperature error averaged over 1954–95 from 0.16° to 0.04°C, and to reduce the total root-mean-square error from 0.41° to 0.35°C. Further discussion of the results at S is presented in Carton et al. (2000). The impact of the bias correction algorithm is even larger in the high data regions of the western North Pacific and Atlantic.

Implementation of the estimation algorithm becomes primarily a matter of choosing error covariances properly. In optimal interpolation data assimilation the unbiased forecast error is assumed to be time stationary and homogeneous, while observation error is assumed to be uncorrelated and forecast bias error is assumed to be zero. The Kalman filter, in contrast, introduces predictive equations for the forecast error covariances. Here we choose a middle ground by allowing the forecast error covariance to vary with latitude and depth, but requiring it to be time stationary. We retain the assumption that the observation errors are uncorrelated, but allow our estimate of the forecast bias to evolve according to the equations above.

We will assume that the unbiased covariance of two forecast error variables $\tau(x_i, y_i, z_i)$ and $\eta(x_j, y_j, z_j)$ has an exponential form whose weights vary with their average latitude and depth (y_o, z_o), as well as their separation ($\Delta x, \Delta y, \Delta z$):

$$P^f(\Delta x, \Delta y, \Delta z, y_o, z_o) = \Lambda^{(\tau, \eta)} \exp \left\langle - \left[\left| \frac{\Delta x}{C_x(y_o, z_o)} \right| + \left| \frac{\Delta y}{C_y(y_o, z_o)} \right| + \left| \frac{\Delta z}{C_z(y_o, z_o)} \right| + \left| \frac{\Delta t}{C_t} \right| \right] \right\rangle, \quad (6)$$

where x and y are zonal and meridional distance. The horizontal covariance structure functions C_x and C_y were estimated at each depth as follows: The observed minus forecast differences were formed by subtracting climatological monthly temperature from the observed temperature. These differences are frequently referred to as the innovations. The domain was then divided into tropical (20°S–20°N) and midlatitude (poleward of 20°) regions. Within each region, each month, and at each vertical level, all possible data pairs $\mathbf{v}_k, \mathbf{v}_l$ were formed. The sets of data pairs were then segregated by zonal and meridional distance for which C_x was computed from pairs separated meridionally by at most 25 km,

and C_y was computed from pairs separated zonally by at most 25 km.

The zonal and meridional lagged autocorrelation functions for temperature error at 75 m in the tropical and midlatitude domains are shown in Figs. 3a–d. In addition to the observed correlations for each basin, the panels include a correlation model. Assuming that the bias has been estimated correctly, the error at zero lag provides an estimate of the normalized observation error covariance \mathbf{P}_k^o . These values are typically around 0.5. The spatial scales of the correlations all remain in the range of 250–375 km at this depth, with the largest lags in the zonal correlation in the Tropics. These scales are

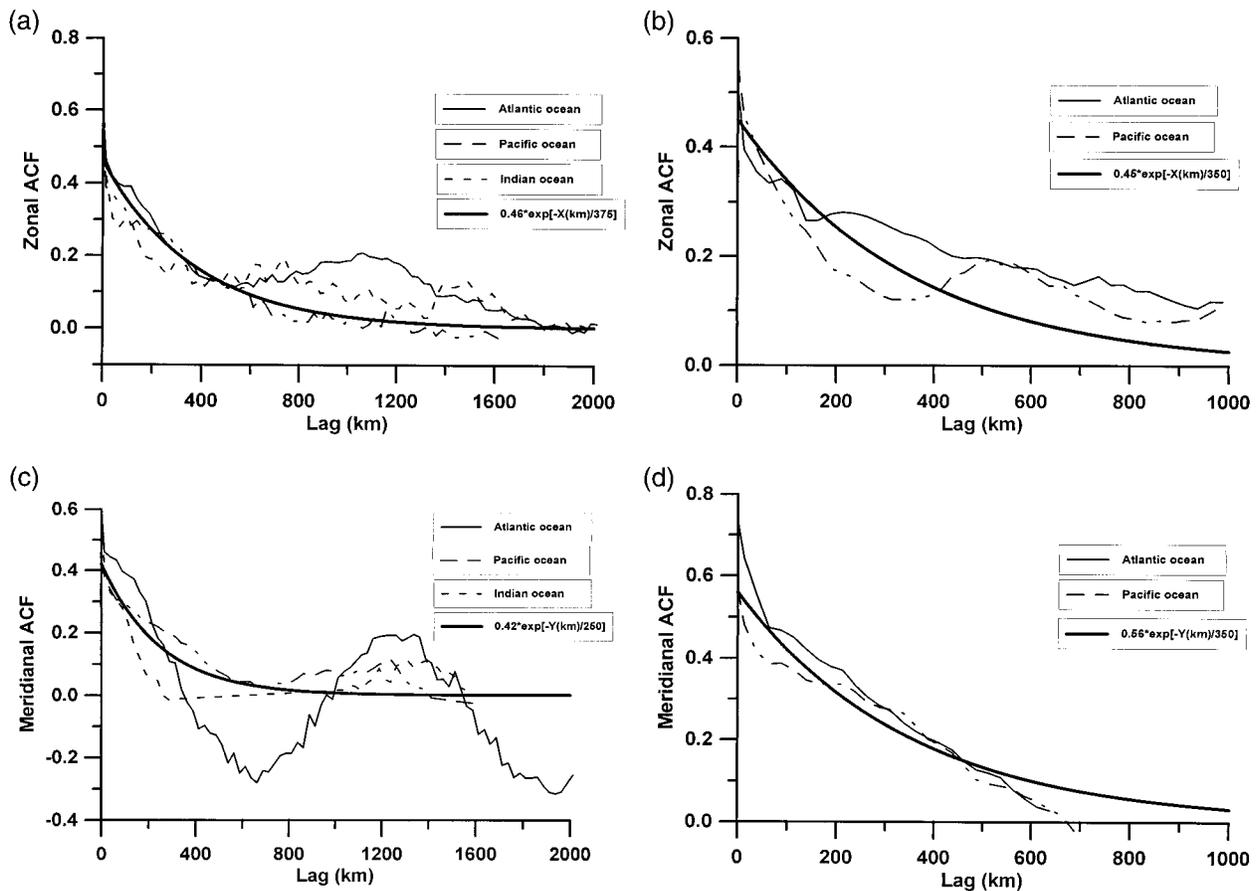


FIG. 3. Covariance of temperature observation error with zonal and meridional lag for three ocean basins. (a) Zonal tropical, (b) zonal midlatitude, (c) meridional tropical, and (d) meridional midlatitude. The vertical axis is $(^{\circ}\text{C})^2$. Solid curves show model covariance. Vertical line at zero lag indicates the size of the analysis variance.

similar to those proposed by Meyers et al. (1991) based on XBT data in the Pacific.

A striking feature of the meridional autocorrelation in the Tropics is a strong oscillatory pattern with a wavelength of 1200 km in the tropical Atlantic (Fig. 3c). We have examined the cause of this pattern and found it to be the result of the ridge–trough system of thermocline undulations associated with the zonal equatorial currents. When one current is altered, the others are affected. The Pacific, which also maintains a ridge–trough system, is not so tightly coupled.

After fitting exponential curves to each of the spatial covariances, we construct an analytic form for the dependence of C_x and C_y on depth and latitude as

$$C_x = \left(450 + \frac{375 - 450}{50^\circ} |\phi| \right) \left(\frac{1200 - z \left(1 - \frac{|\phi|}{50^\circ} \right)}{1200} \right) \text{ km}, \quad (7a)$$

$$C_y = \left(250 + \frac{375 - 250}{50^\circ} |\phi| \right) \left(\frac{1200 - z \left(1 - \frac{|\phi|}{50^\circ} \right)}{1200} \right) \text{ km}. \quad (7b)$$

The timescale of the autocovariance of observation-minus-forecast differences was similarly estimated to be $C_1 = 30$ days. The vertical covariance function of unbiased forecast error is particularly inhomogeneous. This quantity is very large in the mixed layer and of the scale of the pycnocline at pycnocline depths.

After completing the analysis we would like to go back and check the consistency of our assumptions. We have two options to consider: Daley (1992) has suggested examining the temporal correlation of the observation-minus-forecast differences. Improvements in the analysis system should be reflected in reductions in the temporal (and spatial) scales of correlation. Alternatively, Hollingsworth and Lonnberg (1989) have proposed examination of the spatial structure of observation-minus-analysis differences. In principle, these latter differences should be uncorrelated. Any spatial correlation in these differences represents information that has not been extracted by the analysis. Here we follow Hollingsworth and Lonnberg and examine the observation-minus-analysis differences. The 75-m temperature differences were binned into small 25×25 km² bins in the same regions as used in deriving (7). The variance of observed temperature in the tropical Atlantic is 1.6 (°C)² (Fig. 4). The covariance at lags between 25 and 50 km is less than 0.2 (°C)².

We now discuss the structure of the covariance $\Lambda^{(\tau, \eta)}$ in (7). In the simple case where the variables τ and η represent forecast temperature errors at two locations, $\Lambda^{(\tau, \eta)}$ is simply the product of the unbiased forecast error standard deviations at two locations (σ_i^f, σ_j^f). We antic-

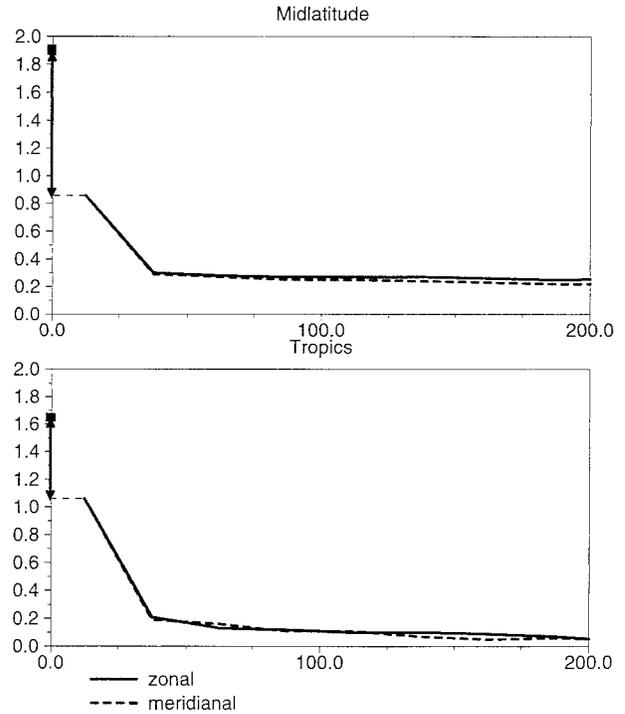


FIG. 4. Covariance of 75-m observed minus analysis temperature differences with latitude and longitude computed in the North and tropical Atlantic. The covariance at zero lag reflects observation error.

ipate a close relationship between temperature and salinity errors at pycnocline depths in many parts of the ocean [the importance of the temperature–salinity relationship to the assimilation problem is discussed in Cooper and Haines (1996)]. However, the relationship varies with location because of the varying distribution of water masses. Here we use our CTD dataset to construct a lookup table for the temperature–salinity error covariance as a function of x_o, y_o, z_o . Among the difficulties encountered in constructing this table is the necessity of excluding the high latitude regions of the northern ocean where the relationship between temperature and salinity error is multiply valued.

One of the most extensive datasets used in our assimilation analysis is altimeter-derived sea level. Here we follow the reasoning of Carton et al. (1996). We anticipate that sea level and density forecast errors are negatively correlated because of the tendency of the ocean to compensate for a rise of sea level with a depression of the pycnocline. This observation together with the assumption that the variations in sea level are “small” leads to the conclusion that for sea level error and temperature error $\Lambda^{(\tau, \eta)}$ should be approximately proportional to the mean vertical temperature gradient $[-\gamma(x_o, y_o, z_o)\partial T/\partial z]$. In this formula γ is an empirical nondimensional constant that determines the amplitude of isopycnal depth displacement error, implied by a given sea level error. Based on comparison of the TOGA–TAO thermistor data and TOPEX/Poseidon altimetry

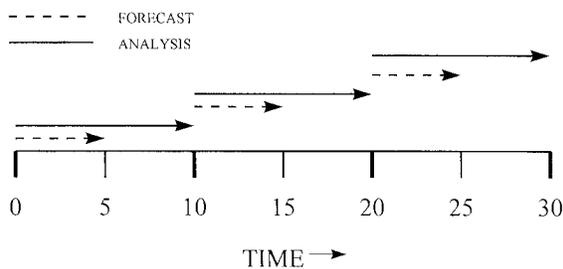


FIG. 5. Illustration of the incremental update analysis procedure of Bloom (1996). Preliminary 5-day forecasts are indicated with dashed lines. The first forecast begins on day 0. At day 5 forecast errors are estimated. A second 10-day forecast is carried out starting from day 0 (solid line). In this forecast the mass fields are continuously corrected for the estimated forecast error. The analysis at day 10 provides initial conditions for the next 5-day preliminary forecast. The procedure repeats itself every 10 days.

Carton et al. show that γ should lie in a range of 100–200 in the tropical Pacific (a +1 cm sea level forecast error implies a -1 to -2 m isopycnal depth error). For the current study we have recalculated $\gamma(x_o, y_o, z_o)$ based on the global historical sea level and isopycnal depth variations obtained from a preliminary assimilation analysis. Similarly, we have recalculated the relationship between forecast salinity error and sea level error. Neglecting this relationship would contribute 1–2 cm to the observed sea level error.

Because detailed structure of the earth's geoid is uncertain, the sea level estimates from Geosat and those of later overlapping missions (*ERS-1* and *TOPEX/Poseidon*) each differ from "true sea level" by an unknown time stationary, but spatially varying, mean. In this study we compensate for these two unknown means by removing time-mean sea level from the datasets and replacing them with time-mean sea level produced by an experimental analysis in which altimetry was not assimilated (computed for the same period as each mission). In this way, the effect of the altimetry in the time-mean sea level is minimized, while still contributing to the trend and to shorter timescale variability.

The procedure for updating temperature errors in the mixed layer is somewhat different than for subsurface errors. The reason for this distinction is the availability of a more extensive dataset and that the mixed layer has broader spatial scales and stronger vertical coherence (Carton and Hackert 1990). At the update time step the forecast depth of the mixed layer is the depth at which forecast density has increased from its surface value by $\Delta\sigma_T = 0.03$. This value was chosen as the result of a series of experiments and depends on the model resolution and mixing parameterization. The sea surface temperature of Reynolds and Smith (1994), resulting from objectively combining satellite, shipboard, and buoy sea surface temperature observations, is used as an estimate of mixed layer temperature. Sea surface salinity currently is relaxed to its climatological seasonal

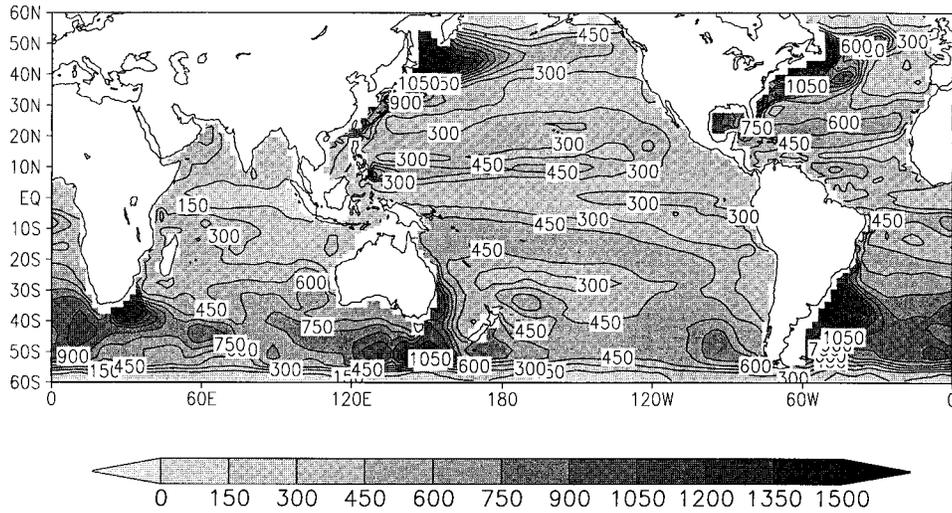
value in this layer (temperature and salinity errors are assumed to be uncorrelated within the mixed layer).

Time stepping in the analysis is carried out using the incremental update analysis method of Bloom (1996). In this approach, illustrated in Fig. 5, we begin with a preliminary 5-day forecast (dashed line). At day 5 the forecast error is estimated using the procedure outlined above. Then we go back and carry out a 10-day integration starting from day 0, correcting the mass fields for the estimated forecast error (solid line). This corrected integration is our analysis. The analysis at day 10 provides initial conditions for the next 5-day preliminary forecast. The forecast error is computed on day 15, then we go back and carry out a new 10-day analysis beginning on day 10 and the cycle repeats itself. This predictor–corrector time-stepping method has two advantages. The first is that it acts like a continuous assimilation method by keeping the mass and momentum fields in balance, thus suppressing gravity waves otherwise expected to be produced by the updating procedure. The second is that it reduces the impact of bias on the analysis by 50% because the forecast is continually corrected for the estimated forecast error, which includes an estimate of the remaining bias. The cost of this procedure is a 50% increase in the integration time of the model forecast over that of a simple 10-day intermittent analysis.

The assimilation procedure relies on the forecast of an ocean model based on the Geophysical Fluid Dynamics Laboratory Modular Ocean Model 2.2 software. The model horizontal resolution is $2.5^\circ \times 0.5^\circ$ in the Tropics, expanding to a uniform $2.5^\circ \times 1.5^\circ$ resolution at midlatitude. The basin domain extends from 62°S to 62°N for a total of 146×96 horizontal grid points. At the polar boundaries the temperature and salinity fields relaxed to Levitus and Boyer (1994) climatology. We make no attempt to model cryospheric or deep-water formation processes explicitly. A weak 5-yr relaxation of the global temperature and salinity fields is included in order to reduce weak forecast bias in deep water masses. Bottom topography is included. The model has 20 levels in the vertical, with 15-m resolution in the upper 150 m. This resolution was chosen to be as high as possible while still allowing experiments to be carried out on workstations. Horizontal friction and diffusion is included with a constant value of $6 \times 10^7 \text{ cm s}^{-2}$. Vertical friction and diffusion are Richardson number dependent with a maximum value of 3 cm s^{-2} . A 50-year analysis with a 1-h time step currently requires three weeks on a Digital alpha workstation with a 333-MHz CPU.

For the period 1950–1992 the winds are provided by an historical analysis of Comprehensive Ocean–Atmosphere Data Set (COADS) surface wind observations by da Silva et al. (1994). We have removed a linear trend from both components of wind at each grid point. The impact of removing the trend is examined below.

IMPACT OF OBSERVATIONS



IMPACT OF MODEL

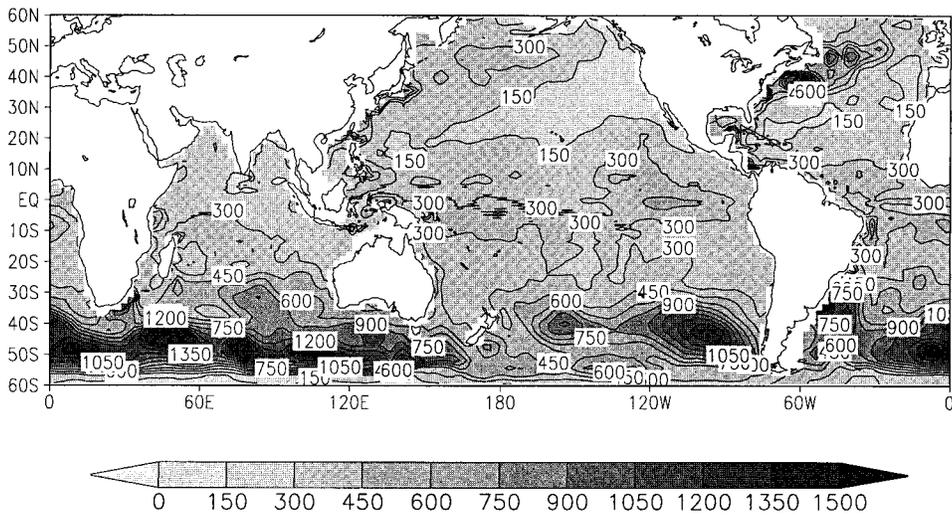


FIG. 6. Relative impact of wind, model, and subsurface temperature variability on the 0/500 m heat content variability from its climatological seasonal cycle. Units are $^{\circ}\text{C m}$. Upper panel shows the root-mean-square difference between the control analysis heat content and heat content from expt 2 in which no data updating is carried out. The difference reveals the impact that subsurface observations have on the analysis temperature. Middle panel shows the root-mean-square difference between the control analysis heat content and heat content from expt 7 in which the model is replaced by climatological monthly temperature. The difference in this panel shows the impact of the model (including winds). Lower panel shows the root-mean-square difference between the control analysis heat content and heat content from expt 3 in which the model is retained, but winds are replaced with their climatological monthly values. The difference in this panel shows the specific impact of historical winds.

In order to allow us to extend the analysis past 1992, we have added to the wind record using National Center for Environmental Prediction (NCEP) winds for 1992–95. The mean strength of the COADS and NCEP winds are different. In order to reduce changes in the analysis due to changes in the wind field we have corrected the

NCEP winds to the mean of the COADS winds based on comparing the mean winds of the two historical analyses during a 2-yr period of overlap. Surface heat flux is effectively determined by surface temperature observations. Sea surface salinity is relaxed to climatological monthly salinity in this study.

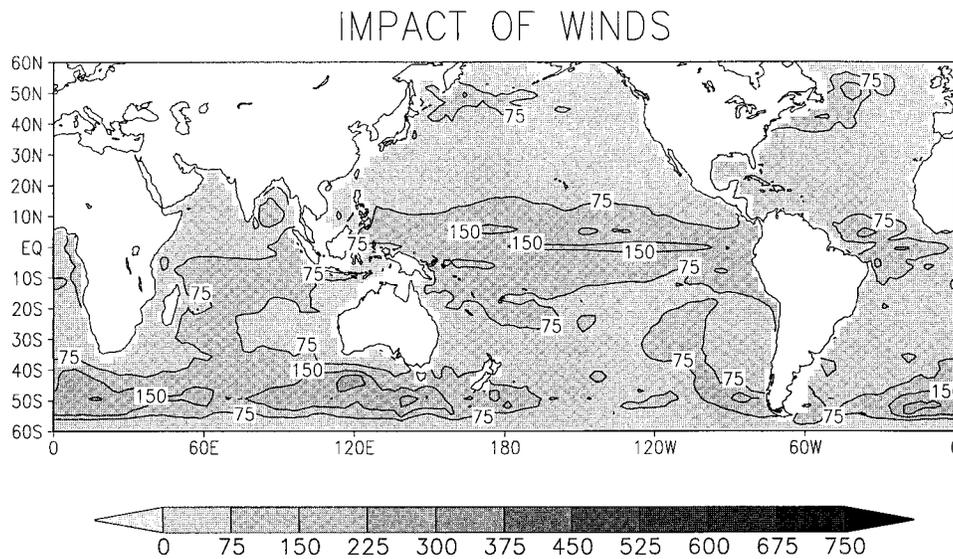


FIG. 6. (Continued)

b. Data, model, and wind impact

The first issue we address is the relative importance of winds versus subsurface updating. Variability is introduced into the analysis either through inaccuracies in the forecast model and its associated initial/boundary conditions or through the updating procedure and its associated suite of observations. In experiment 2 (see Table 1) the control analysis is repeated except without any subsurface updating (a simulation). In this experiment interannual and decadal variability is due primarily to variability in the historical wind field and SST. Thus, the difference between experiment 2 heat content and heat content from the control analysis indicates the impact of the subsurface observations (Fig. 6, upper panel). The subsurface observations have an impact throughout the basin approaching 300°C m in most places. In the region of the western boundary current the impact is even larger. The large corrections here are associated with correcting the position and scale of the boundary current.

In experiment 7 (see Table 1) the control analysis is repeated except now using climatological monthly temperature as a first guess (a statistical objective analysis). In this experiment interannual and decadal variability is due to surface and subsurface temperature updating. Thus, the difference between experiment 7 heat content and heat content from the control analysis indicates the impact of the model, including the interannually varying winds (Fig. 6, middle panel). The impact of the model exceeds 300°C m in the Tropics and exceeds 500°C m in the region of the Gulf Stream and the Southern Ocean. We believe that the impact of the model in the Southern Ocean is due to the lack of subsurface observations here so that much of the interannual variability is introduced through the action of the winds.

Finally, in experiment 3 the control analysis is re-

peated except that the winds are replaced with their climatological monthly average. In this experiment, which may be thought of as a subset of the situation examined in experiment 5, interannual and decadal variability is due to variability in the updating datasets, which may have remote influence due to the generation of waves. Thus, the difference between experiment 3 and the control analysis heat content is solely the result of wind variations (Fig. 6, lower panel). The contribution of the winds is generally less than 100°C m except in the Tropics and the Antarctic Circumpolar Current. Comparison of the results shown in the three panels of Fig. 6 suggests the subsurface data have most of the impact on the analysis of the subsurface thermal field on interannual and longer timescales. In the Tropics, though, winds do contribute significantly.

In the Tropics the model plays an additional role due to the presence of tropical wave dynamics. In Fig. 7, we compare heat content anomalies along the equator in the Pacific with a similar picture from experiment 7 in which climatological monthly temperature is used as a first guess. Prior to 1980 the no-model case has much reduced variability. In contrast, the control case shows a succession of six Niño events prior to that date. After 1980 the two analyses more closely resemble each other, although the no-model case is still noisier than the control case.

We next consider the source of the long-term trend in heat content. Recent studies by Clarke and Lebedev (1996) have shown that the trend in the historical wind field is not matched by a trend in surface pressure, and thus is likely to be spurious. Our response to this result has been to remove the long-term trend from the wind field in most of our experiments. We examine the impact of removing the wind trend by comparing the trend in heat content from experiment 4 in which the wind field

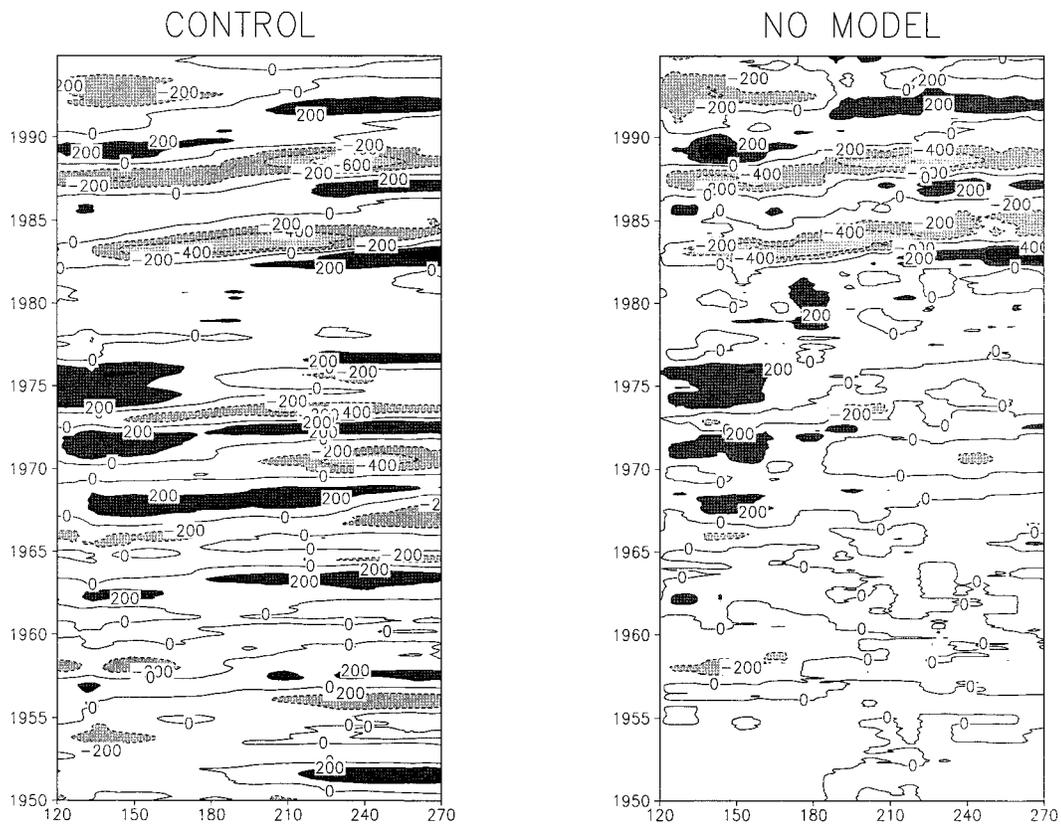


FIG. 7. Heat content anomaly along the equator in the Pacific with time showing the influence of the model on the analysis. Left-hand panel shows the control analysis, while right-hand panel shows expt 5 in which the model has been replaced with climatological monthly temperature. Note the reduced variability in the no model case prior to 1980.

has not been detrended with the trend in heat content from the control analysis (Fig. 8). The linear trend of heat content in the control analysis is small, generally less than $3^{\circ}\text{C m (200 d)}^{-1}$ with the largest trends at high latitudes. In the subtropics heat content has been rising by a more modest $1\text{--}2^{\circ}\text{C m (200 d)}^{-1}$, while in the Tropics the trend has been toward cooling at a rate of $-1^{\circ}\text{C m (200 d)}^{-1}$. When the wind field trend is retained (experiment 4) the heat content trend in the Tropics increases by a factor of 2 or more with strong cooling in the eastern Tropics and warming in the west as the result of the strengthening of the trade winds with time.

STATION S TEMPERATURE AND SALINITY

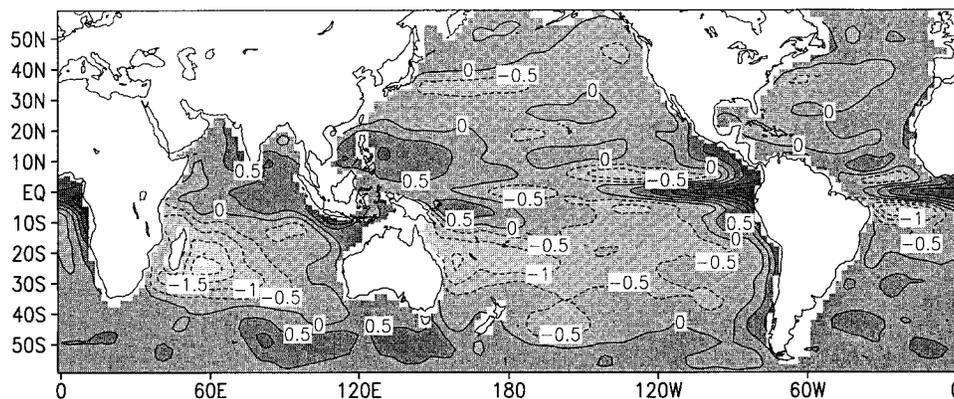
Thirteen permanent ocean weather stations were established in the North Atlantic and Pacific following World War II by the United Nations (Dinsmore 1996). Most of the stations maintained by the United States were terminated by 1977 so that their record lengths are short. Others are too far north to be of interest to us. However, station S, in the western side of the Atlantic subtropical gyre (32°N , 64°W), is well suited for use in examining the properties of the analysis (WHOI and BBSR 1988). In order to make the comparison at this

station independent the station S hydrographic data, a combination of bottle data and CTDs (including the BATS data), has been specifically excluded from the control analysis and the experiments.

The observed mixed layer at S is shallow, less than 25 m deep, throughout the summer. By late winter this layer deepens to 200 m or more with mixed layer temperatures decreasing to 20°C (Michaels and Knap 1996). Within this mixed layer temperature and salinity are largely independent. Figures 9a,b (upper panels) show that from year to year the mixed layer temperature and salinity at this location vary independently of each other. Joyce and Robbins (1996) observe that mixed layer salinity has a longer timescale than temperature. They conclude that salinity is more closely coupled to the slowly varying properties of the water mass below the mixed layer, while temperature is more closely linked to surface meteorology.

Much of our interest in the time series at station S is in the properties of this water mass. The distinctive subtropical-mode water mass is characterized by a uniform temperature with depth, or thermocline, of nearly 18°C . It is present in a layer approximately 200 m thick throughout the Sargasso Sea but is formed in a more limited region to the northeast of Bermuda in late winter.

TREND DUE TO WINDS



TREND DUE TO DATA AND MODEL

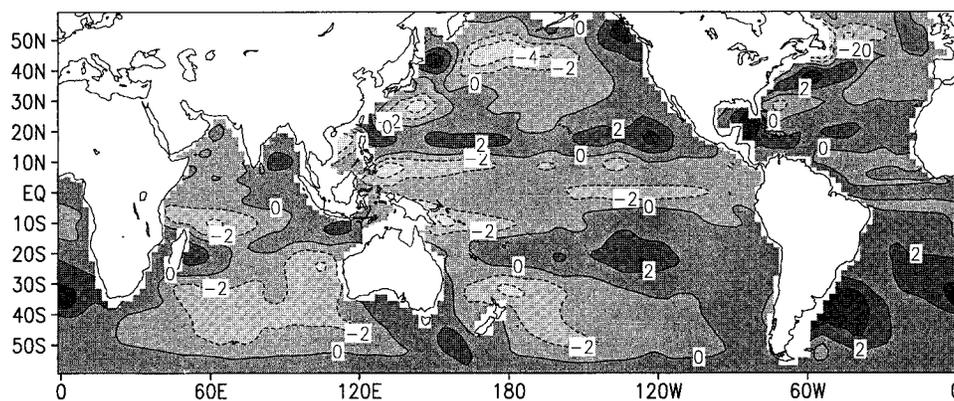


FIG. 8. Contribution of trends in the winds to trends in 0/500 m heat content. Upper panel shows the linear trend of heat content at each point computed using the full 46-yr record for expt 4 in which the wind field was not detrended. Striking changes occur along the equator. In the Atlantic and Pacific, the intensification of the trade winds with time leads to a gradual intensification of the zonal gradient of heat storage. Lower panel shows a similar map of heat content trend for the control analysis for which a linear trend was removed from the winds. Units are $^{\circ}\text{C m (200 d)}^{-1}$.

Changes in the thickness of the 18°C thermocline at Bermuda are thought to indicate changes in the rate of water mass formation.

Talley and Raymer (1983) have examined the early part of the record at S from this perspective. They argue that 18°C Water formation ceased during 1972–75 leading to lower temperature and salinity (this is apparent in the depth of the 18°C isotherm in Fig. 9a, upper panel, and the 36.5 psu contour in Fig. 9b, upper panel). The spatial structure of this event in the early 1970s was examined by Levitus (1989), who showed that the changes extended throughout the subtropical gyre. A second, less dramatic change in water mass formation apparently occurred in 1984/85. Below the layer containing subtropical mode water are the thermocline and halocline whose 50 m vertical (or 300 km horizontal) excursions are highly correlated with sea level (Roemich 1990).

Cooper and Haines (1996) have recently pointed out

the need to maintain consistency between temperature and salinity. In the remainder of Fig. 9 we explore this issue by presenting vertical cross sections of temperature and salinity from the control analysis and two experiments. In experiment 6 the analysis is carried out without any constraint on the salinity field, except a weak relaxation to climatology. Although the reconstruction of the temperature field is good, the analysis is unable to reproduce the observed subsurface salinity maximum. Indeed, the upper water column becomes too fresh by up to 0.5 psu. In the control analysis where salinity errors are updated based on observed temperature errors the water column is no longer too fresh. In addition, the salinity anomalies of 1972–47 and 1984–58 are now evident below the mixed layer, but are much weaker than observed. The mixed layer has no salinity anomalies since temperature and salinity errors are not correlated in the mixed layer. In experiment 5 when direct observations of salinity are assimilated (except

OCEAN STATION "S" TEMPERATURE

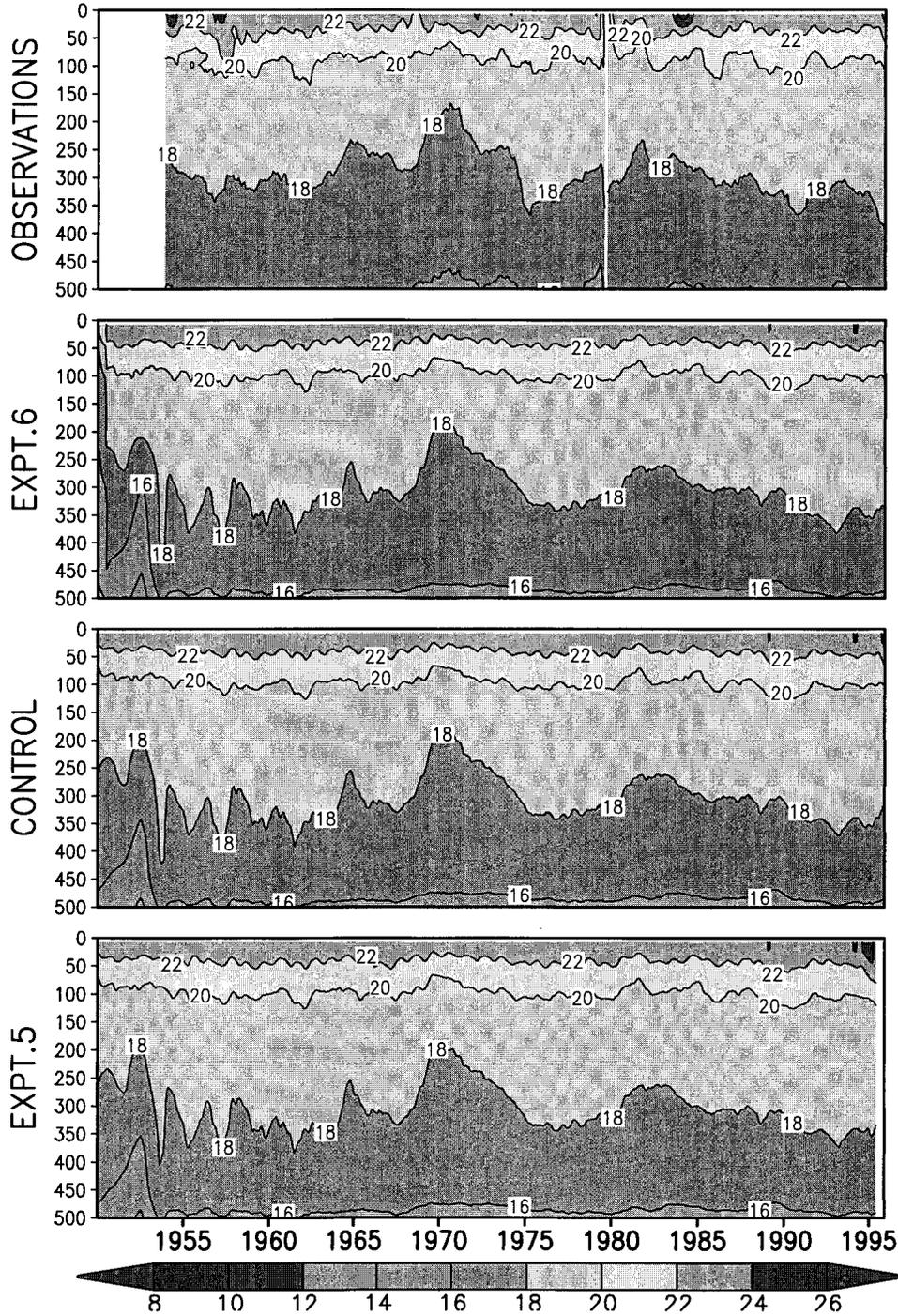


FIG. 9. Observed and analysis temperature and salinity at station S (32°N, 64°W). Upper panel shows observations. Second panel shows expt 6 analysis without any salinity updating. Third panel shows control analysis, which includes a covariance model for the temperature and salinity errors, but without using salinity observations. Bottom panel shows expt 5 analysis that includes salinity observations (except those at S).

OCEAN STATION "S" SALINITY

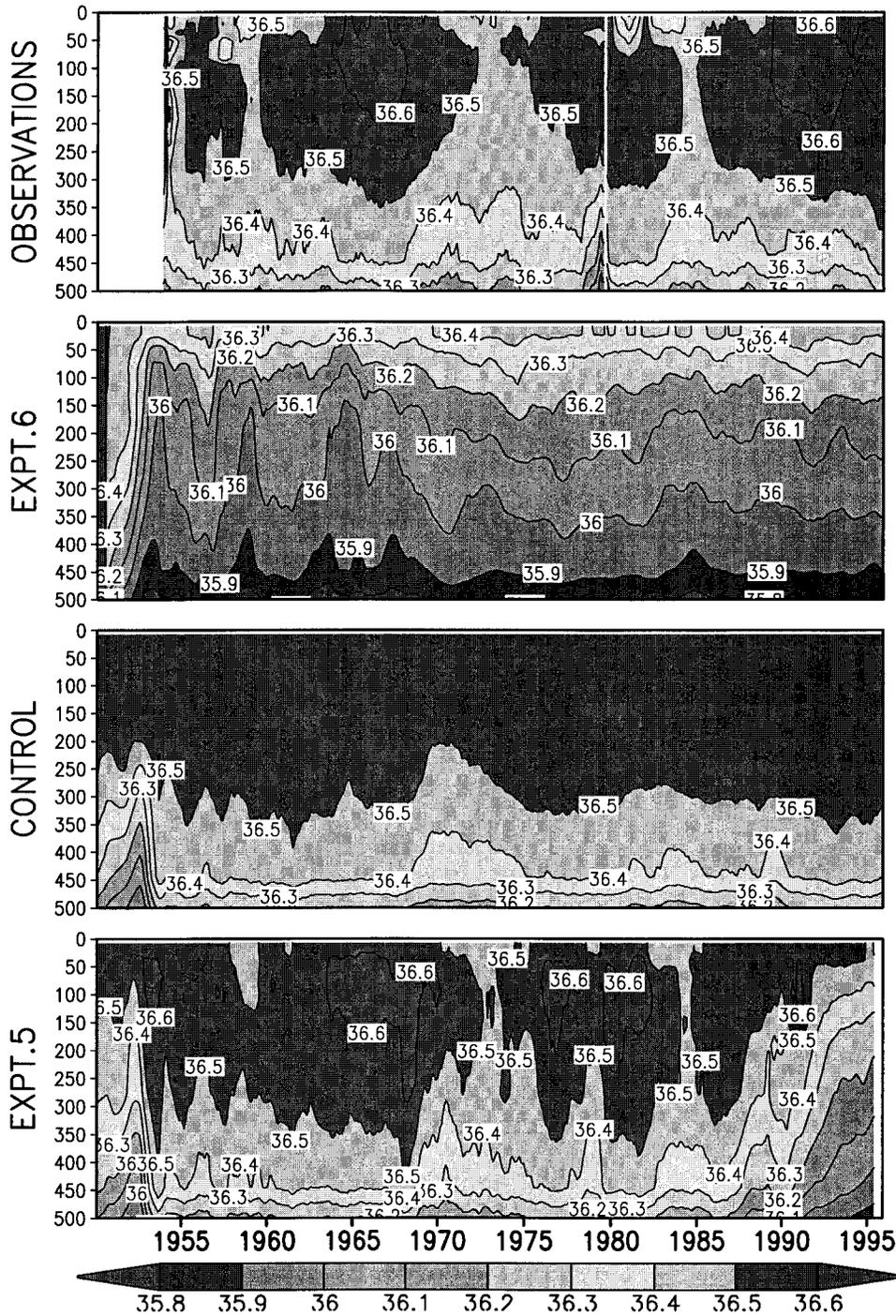


FIG. 9. (Continued)

the observations at S) the salinity becomes most realistic of all. The main errors are before 1953, due to initialization transients, and after 1990 (the last date of salinity measurements in the WOA-94 dataset).

In all three of the analyses shown in Fig. 9, the tem-

perature field at S is well represented. However, in the absence of subsurface temperature updating (expt 2) the 18°C thermocline disappears by 1960. The disappearance of the 18°C Water indicates that the model is unable to form subtropical water properly in the absence of data

assimilation. A number of possible causes of this problem have been suggested to the authors. These possible causes include limited model resolution, insufficient mixed layer salinity variability, and insufficient surface cooling due to the absence of strong cold air outbreaks in the historical surface forcing fields. Insufficient 18°C Water production in the forecast model is a problem that needs further exploration.

4. Conclusions

Here we present a description of the Simple Ocean Data Assimilation. SODA contains all components of an analysis system that we can anticipate in the future, including an ocean general circulation model; models of the observation and forecast error; the basic hydrographic, altimeter, and SST datasets; and a constraint algorithm. Particular attention is focused on the problem of bias in the model forecast and its effects on the analysis. We then use SODA to construct a retrospective analysis of temperature, salinity, and current in the upper layers of the ocean globally during the past five decades.

The constraint algorithm is based on optimal interpolation data assimilation. This technique uses a numerical forecast model to provide a first guess of the mass and momentum field at each analysis time. Statistical estimates of the forecast error are computed a priori and are assumed to be steady. These assumptions are responsible for the computational efficiency of our approach. Observations of temperature, salinity, and sea level are used to estimate the error in the first guess of the mass field as well as to estimate forecast model bias. The error estimates are then used to update the mass and momentum fields with a time-continuous updating approach that ensures that the mass and momentum fields remain close to geostrophic equilibrium.

We begin our exploration of the results by examining the relative impact of subsurface observations, winds, and model physics on the control analysis. By comparing the control analysis results with results from experiments without one or another of these factors we are able to identify regions in which each have substantial effect. Along the equator heat content variability results from both historical winds and historical direct observations. However, their impact is not distributed uniformly with time. Prior to 1980, before the expansion of the ship of opportunity XBT lines, moored thermistor array, and altimetry, the winds and model are clearly the most important factors in producing an analysis of heat. By 1985 the thermal field along the equator is well resolved by direct observations.

The situation in the midlatitude ocean is somewhat different. There, we find that the model driven by historical wind stress is inadequate to reproduce the historical record of heat content variations. One possible explanation is that the historical wind stress does not adequately resolve smaller-scale variations. Because the curl of wind stress drives the circulation in the interior

ocean, the missing smaller-scale variations may lead to large errors in the basin-scale circulation. Direct salinity observations as well as the relationships between temperature and salinity errors also need to be taken into account in the Tropics and midlatitude. They are necessary to the estimation of the salinity field and, in particular, the subduction of pycnocline waters.

Throughout the ocean the forecast model is shown to have substantial bias due to errors in the mixed layers and in the production rates of subtropical and tropical water masses. Among the causes of these errors is lack of knowledge of the historical surface freshwater budget, errors in specification of mixed layer dynamics, and in resolution of important topographic features controlling the exchange of water between basins. Model inadequacies are certainly responsible for errors in the region of the western boundary currents. Some of these problems may be addressed by shifting to an isopycnal coordinate model. It is also possible that improvements in the data assimilation methodology, such as a shift to adjoint or streamline assimilation techniques, may improve the analyses. We will be exploring these possibilities in the near future.

Acknowledgments. We are grateful to a number of people who have given us access to their datasets. Mark Swenson and Zengxi Zhou of the Atlantic Oceanographic Marine Laboratory/NOAA have provided the monthly averaged surface drifters and Sydney Levitus and Robert Cheney and their colleagues at the National Oceanographic Data Center/NOAA have given us access to the hydrographic and altimeter data. We have benefited from the datasets collected by principal investigators of the Tropical Ocean Global Atmosphere and World Oceanographic Circulation Experiments. The IOC/GODAR project has played an important role in increasing the historical hydrographic dataset. We thank Dick Dee and Arlindo da Silva of the Data Assimilation Office/NASA for many discussions on the subject of data assimilation. Finally, we want to express our gratitude for support from the Office of Global Programs/NOAA under Grants NA66GP0269 to JAC and NA76GP0559 to BSG, and the National Science Foundation under Grant OCE9416894 to JAC.

REFERENCES

- Acerro-Schertzer, C. E., D. V. Hansen, and M. Swenson, 1997: Evaluation and diagnosis of surface currents in the National Centers for Environmental Prediction's ocean analyses. *J. Geophys. Res.*, **102**, 21 037–21 048.
- Bennett, A. F., 1990: *Inverse Methods in Physical Oceanography*. Cambridge University Press, 346 pp.
- Bloom, S. C., 1996: Data assimilation using incremental analysis updates. *Mon. Wea. Rev.*, **124**, 1256–1271.
- Carton, J. A., and E. C. Hackert, 1990: Assimilation analysis of tropical Atlantic circulation during 1983–84. *J. Phys. Oceanogr.*, **20**, 1150–1165.
- , B. S. Giese, X. Cao, and L. Miller, 1996: Impact of altimeter, thermistor, and expendable bathythermograph data on retro-

- spective analyses of the tropical Pacific Ocean. *J. Geophys. Res.*, **101**, 14 147–14 159.
- , G. Chepurin, and X. Cao, 2000: A simple ocean data assimilation analysis of the global upper ocean 1950–95. Part II: Results. *J. Phys. Oceanogr.*, **30**, 311–326.
- Chepurin, G. A., and J. A. Carton, 1999: Comparison of retrospective analyses of the global ocean heat content. *Dyn. Atmos. Oceans*, **29**, 119–145.
- Clarke, A. J., and A. Lebedev, 1996: Long-term changes in the equatorial Pacific trade wind systems. *J. Climate*, **9**, 1020–1028.
- Cooper, A., and K. Haines, 1996: Data assimilation with water property conservation. *J. Geophys. Res.*, **101**, 1059–1077.
- Daley, R., 1991: *Atmospheric Data Analysis*. Cambridge University Press, 457 pp.
- , 1992: The lagged innovation covariance: A performance diagnostic for atmospheric data assimilation. *Mon. Wea. Rev.*, **120**, 178–196.
- da Silva, A. M., C. C. Young, and S. Levitus, 1994: *Atlas of Surface Marine Data 1994, Vol. 1: Algorithms and Procedures*. NOAA Atlas NESDIS 6, U.S. Department of Commerce, NOAA, NESDIS, 83 pp.
- Dinsmore, R. P., 1996: Alpha, Bravo, Charlie. *Oceanus*, **39**, 9–10.
- Hanawa, K., P. Rual, R. Bailey, A. Sy, and M. Szabados, 1994: Calculation of new depth equations for expendable bathythermographs using a temperature-error-free method (application to Sippican/TSK T-7, T-6 and T-4 XBTs). Intergovernmental Oceanographic Commission 42, UNESCO, 46 pp. [Available from NODC/NOAA, Silver Spring, MD 20910.]
- Hollingsworth, A., and P. Lonnberg, 1989: The verification of objective analyses: Diagnostics of analysis system performance. *Meteor. Atmos. Phys.*, **40**, 3–27.
- Hsiung, J., R. E. Newell, and T. Houghtby, 1989: The annual cycle of oceanic heat storage and oceanic meridional heat transport. *Quart. J. Roy. Meteor. Soc.*, **115**, 1–28.
- Ji, M., A. Leetmaa, and J. Derber, 1995: An ocean analysis system for seasonal to interannual climate studies. *Mon. Wea. Rev.*, **123**, 460–481.
- Joyce, T., and P. Robbins, 1996: The long-term hydrographic record at Bermuda. *J. Climate*, **9**, 3121–3131.
- Levitus, S., 1989: Interpentadal variability of salinity in the upper 150 m of the North Atlantic Ocean versus 1955–1959. *J. Geophys. Res.*, **94**, 16 126–16 131.
- , and T. Boyer, 1994: *World Ocean Atlas 1994, Vol. 4: Temperature*. NESDIS Atlas series, NOAA, Washington, DC, 117 pp.
- , —, and J. Antonov, 1994: *World Ocean Atlas, 1994, Vol. 5: Interannual Variability of Upper Ocean Thermal Structure*. NESDIS Atlas series, NOAA, Washington, DC, 176 pp.
- Malanotte-Rizzoli, P., 1996: *Modern Approaches to Data Assimilation in Ocean Modeling*. Elsevier, 455 pp.
- Meyers, G., H. Phillips, N. Smith, and J. Sprintall, 1991: Space and timescales for optimal interpolation of temperature—Tropical Pacific Ocean. *Progress in Oceanography*, Vol. 28, Pergamon, 189–218.
- Michaels, A. F., and A. H. Knap, 1996: Overview of the U.S. JGOFS Bermuda Atlantic time-series study and the Hydrostation S program. *Deep-Sea Res.*, **43**, 157–198.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analysis using optimum interpolation. *J. Climate*, **7**, 929–948.
- Roemmich, D., 1990: Sea level and the thermal variability of the ocean. *Sea-Level Change*, Geophysics Study Committee, Commission of Physical Sciences, Mathematics and Resources, National Research Council, and National Academy Press, 208–217.
- , 1995: Sea level and variability of the oceans. *Glaciers, Ice Sheets, and Sea Level: Effect of a CO₂-Induced Climatic Change*, DOE/ER/GO235-1, Dept. of Energy, 104–115.
- Rosati, A., R. Gudgel, and K. Miyakoda, 1995: Decadal analysis produced from an ocean assimilation system. *Mon. Wea. Rev.*, **123**, 2206–2228.
- Searle, B., 1992: *Proceedings of the Ocean Climate Data Workshops*, Goddard Space Flight Center, 97–108. [Available from NODC/NOAA, Silver Spring, MD 20910.]
- Smith, N. R., 1995: An improved system for tropical ocean subsurface temperature analysis. *J. Atmos. Oceanic Technol.*, **12**, 850–870.
- Thiebaut, H. J., and L. L. Morone, 1990: short-term systematic errors in global forecasts: Their estimation and removal. *Tellus*, **42A**, 209–229.
- White, W. B., 1995: Design of a global observing system for basin-scale upper ocean temperature anomalies. *Progress in Oceanography*, Vol. 36, Pergamon Press, 169–217.
- WHOI and BBSR, 1988: *Station “S” off Bermuda Physical Measurements, 1954–84*. WHOI and BBSR, 189 pp.
- Wunsch, C., 1996: *The Ocean Circulation Inverse Problem*. Cambridge University Press, 442 pp.

