An Inversion–Assimilation Approach Using Hydrographic Data in a Coarse-Resolution Ocean Model

ADRIAN HINES*

Department of Mathematics, Keele University, Keele, Staffordshire, United Kingdom

PETER D. KILLWORTH

Southampton Oceanography Centre, Southampton, United Kingdom

(Manuscript received 27 March 2000, in final form 9 January 2001)

ABSTRACT

Attempts to estimate the state of the ocean usually involve one of two approaches: either an assimilation of data (typically altimetric surface height) is performed or an inversion is carried out according to some minimization scheme. The former case normally retains some version of the time-dependent equations of motion; the latter is usually steady. Data sources are frequently not ideal for either approach, usually being spatially and temporally confined (e.g., from an oceanographic cruise). This raises particular difficulties for inversions, whose physics seldom includes much beyond the geostrophic balance. In this paper the authors examine an approach midway between the two, examining several questions. (i) What is the impact of data assimilated continuously to a steady state on regions outside the data sources? (ii) Can remote data improve the long-term mean of a model whose natural response is not close to climatology? (iii) Can an eddy-free model assimilate data containing eddies?

The authors employ an inversion using a simple North Atlantic model, which permits no eddies, but contains better dynamics than geostrophy (the frictional planetary geostrophic equations), and an assimilative scheme rather simpler than those normally employed, almost equivalent to direct data insertion, run to a steady state. The data used are real subsurface data, which do contain eddies, from World Ocean Circulation Experiment cruises in the northern North Atlantic. The presence of noise in these data is found to cause no numerical difficulties, and the authors show that the impact of even one vertical profile can strongly modify the water mass properties of the solution far from the data region through a combination of wave propagation, advection, and diffusion. Because the model can be run for very long times, the region of impact is thus somewhat wider than would occur for assimilations over short intervals, such as a year.

1. Introduction

Historically, observations of the ocean have always been less abundant than meteorological observations of the atmosphere. Despite recent efforts to improve the data coverage of the oceans, the number and frequency of observations still lag far behind that of meteorology. Oceanographers frequently still rely on hydrographic data with sparse spatial coverage of the oceans to infer the mean large-scale circulation using ''inverse models.''

These inverse models rely on geostrophy in the form of the thermal wind equations relating the vertical derivatives of velocity to the horizontal derivatives of density. Such calculations require the determination of reference velocities to compute the circulation (Wunsch 1978; Schott and Stommel 1978; Killworth 1986; Cunningham 2000). To determine these velocities, one often uses the distributions of additional properties, the application of conservation properties, or additional dynamical equations (Tziperman 1988). The choice of additional constraints is clearly somewhat arbitrary. Schlitzer (1993) takes an alternative viewpoint and demands that mass, heat, and salt budgets are satisfied exactly, and applies geostrophy as a constraint on the solution. Comprehensive reviews of inverse models are contained in Bennett (1992) and Wunsch (1996).

In the search for a more accurate description of the oceans, oceanographers have looked toward combining numerical models (with more complete dynamics than the simple geostrophic balance) with oceanographic data to produce a picture of the ocean circulation that is consistent with both our beliefs about the dynamics and our observations. This has led to the development of oceanographic data assimilation techniques, many of

* Current affiliation: Ocean Applications, Met Office, Bracknell, Berkshire, United Kingdom.

Corresponding author address: Dr. Adrian Hines, Ocean Applications, Met Office, Room 254A, London Road, Bracknell, Berkshire RG12 2SZ, United Kingdom.
E-mail: ahines@meto.gov.uk.

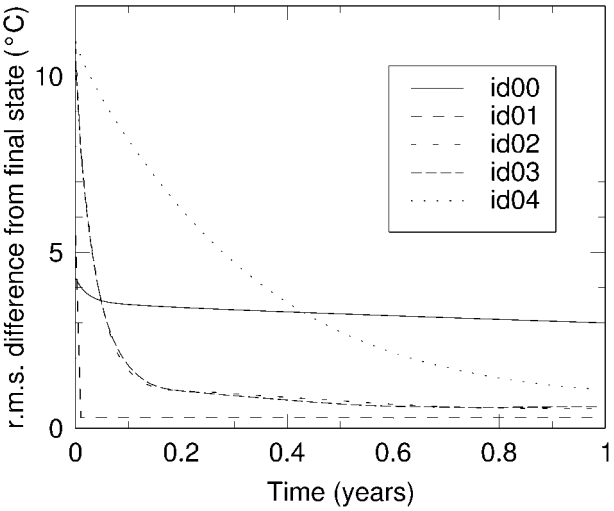


FIG. 1. Convergence of T to the final state, showing runs id00, id01, id02, id03, and id04.

which are adapted from techniques used in meteorology. These techniques range from the simplest ''nudging'' techniques—to be used here—through techniques such as optimal interpolation (widely used in meteorology), to the statistically optimal techniques of the Kalman filter (Ghil et al. 1981) and the adjoint method (Le Dimet and Talagrand 1986).

Optimal methods formally carry with them error statistics, though these suffer from two distinct difficulties. First, they are almost impossible to compute for realistic problems; the model covariance matrix usually has to be estimated (e.g., Derber and Rosati 1989). In addition, the adjoint model, for example, requires the Hessian matrix to be computed; as Marotzke and Wunsch (1993) wryly note, ''the computational load has so far deterred anyone from actually carrying out the recipe.'' Second, error estimates rely entirely upon the model errors being both known and having a Gaussian probability distribution. What little we know about model errors suggests that it is highly likely they will be systematic, and even the nonbiassed component will almost certainly have a skewed distribution. Thus the error estimates arising from these models may be at best far from optimal, and at worst, illusory. In addition, little is known about the response of most models when combined with an optimal assimilation scheme that changes the model physics (i.e., methods other than the strong constraint adjoint; cf. Woodgate and Killworth 1997).

While optimal methods are prohibitively expensive, practical schemes have been developed to overcome their computational drawbacks (Miller 1986; Evensen 1994). However, a trade-off between the sophistication of the assimilation scheme and the complexity of the numerical model inevitably occurs, with optimal schemes being applied to simplified models and suboptimal schemes proving more practical with complex models (Ghil and Malanotte-Rizzoli 1991). Nonoptimal

TABLE 1. Identical twin runs.

No.	Description	Initial T	Initial S (psu)	α
id00	Control run	Linear	35	None
id01	Assimilation at all points	Linear	35	$0.5/dt$
id02	Assimilation on $i = 7, 13$; $j = 7, 13$ only	15°C	30	$0.5/dt$
id03	As id02, but with reduced α	15°C	30	$0.01/dt$
id04	As id02, but with reduced α	15°C	30	$0.001/dt$
id05	Assimilation at one point only	Linear	35	$0.5/dt$

schemes, in particular the nudging technique, have therefore been given much attention in the context of assimilating data in more sophisticated numerical ocean models, even though they convey no error information. Due to the limited data available, in the application of any of the assimilation techniques the numerical model acts as a dynamical extrapolator/interpolator between regions where data are available and those which are unobserved. Hence, a scheme that is practical for use with a model that best represents our knowledge of the ocean dynamics is desirable.

Malanotte-Rizzoli and Young (1995) consider the assimilation of localized nonsynoptic hydrographic datasets in a primitive equation model of the Gulf Stream system using a simple (nudging) assimilation scheme and compare their results to those obtained by assimilating a global dataset and both datasets simultaneously.[1] They find that while the localized data are effective around the regions of observation, over the full model domain they are less so; the combination of global and local observations provides the most realistic results. To what extent their results are specific to the region under consideration, and depend upon the location of the local datasets, is uncertain. However, regularly available global data await autonomous instrument networks, so that for the near future at least the effects of local data need to be considered.

Assimilation techniques of any kind can work well only if the model in which they are embedded runs naturally close to the data; that is, a ''good'' model is needed. This requirement is frequently in contradiction to computer resources available. At present, modelers have two choices. Coarse-resolution models can be run to a steady state (necessary for inversions of the mean state) but yield results somewhat far from climatology. Fine-resolution models, however, are capable of reproducing the observed ocean state (though subject to drift) but need spinup times longer than can be afforded.

This joint requirement of a long-term integration and a model running close to reality remains one of the major stumbling blocks in eliciting a good mean ocean inversion. Further difficulties can occur when the mod-

[1] Fukumori and Malanotte-Rizzoli (1995) used a simplified Kalman filter (though still with unknown model errors), finding that this generally outperformed nudging.

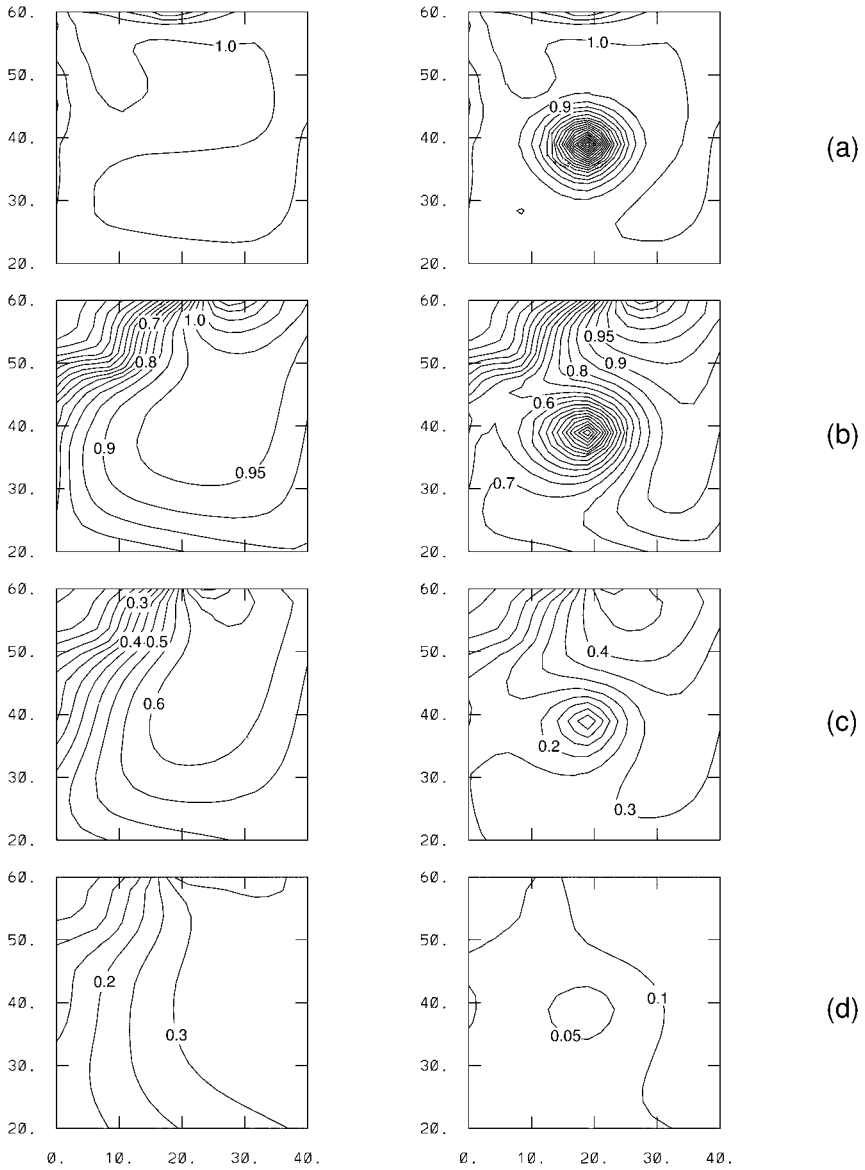


FIG. 2. Plots of $|T_{\text{model}} - T_{\text{final}}|/|T_{\text{initial}} - T_{\text{final}}|$ on $z = 2000$ m, for the control run (left panel) and run id05 (right panel) at (a) $t = 1.2$ months, (b) $t = 6$ months, (c) $t = 2$ yr, and (d) $t = 5$ yr.

els used do not reproduce reality well. If an assimilative method that changes the model is employed, and the data are only available in localized areas, then the changed model state relaxes back to its natural state over a period of time that depends on spatial resolution and mixing parameterizations (Li et al. 2000, unpublished manuscript). Use of an adjoint approach in an inaccurate model means that the model trajectory cannot run close to observations unless other aspects of the model (e.g., forcing) are changed. Initial experience with a coarse model (D. Stammer 2000, personal communication) shows that early state estimates assimilating the World Ocean Circulation Experiment (WOCE) data still lie far from climatology.

If the mean state is required, the loss of information from assimilated data can be avoided by constantly assimilating the same data to a steady state. If data over an entire basin were employed for this exercise, it would not be surprising if realistic temperature–salinity (T–S) structures were reproduced; there would, furthermore, be little information both on the impact of the assimilation and on how the circulation was changed. However, if data were only available in limited areas, but still assimilated to a steady state, this would permit an examination of the impact of the data by seeing where the model was moved toward climatology *outside* the region of data.

This study, then, examines the impact of small

TABLE 2. Rms differences from Levitus climatology.

Depth range (m)	Control		Assimilation	
	Temperature (°C)	Salinity (psu)	Temperature (°C)	Salinity (psu)
Above 100	1.5	0.35	1.4	0.36
100–500	5.8	0.60	3.9	0.51
500–1000	6.2	0.68	2.4	0.44
Below 1000	2.6	0.18	0.4	0.09

amounts of deep T–S data on ocean state estimates, using a simple nudging approach, which can be considered as a cross between simple assimilative and inverse approaches. Such an inversion clearly cannot be expected to recover the entire basin water mass structure accurately, but it does permit some analysis of the impact of continuous assimilation of the same data on remote areas of an ocean basin. (There are few studies on *how* assimilation modifies water masses and circulation in the literature; most research has concentrated on success or failure of the methods.)

Our study uses a model with highly simplified physics (frictional planetary geostrophy). This model, unlike pure geostrophy as normally employed in inversions, can permit simplified western boundary currents and the inclusion of the equator. The model does not support gravity waves, thus avoiding initialization problems (though Malanotte-Rizzoli et al. 1989 found no shocking to occur when direct insertion was employed). Such simplified models have been used previously (Tziperman et al. 1992; Marotzke and Wunsch 1993), although as simplified versions of general circulation models rather than specifically written codes. However, the data used here are near-synoptic and localized, unlike the smoothed gridded data used in those studies, and *do* contain both eddies and gravity waves, which the model does not permit. Indeed, one of our aims is to find out whether reduced-physics models can assimilate noisy data.

The assimilation approach is considerably simpler than the usual statistical approaches, as used by Malanotte-Rizzoli and Young (1995), and is almost equivalent to direct insertion of data from hydrographic sections in a localized region of the northwest Atlantic. The immediate disadvantage of this approach is the apparent lack of error information (a feature of many simple assimilation approaches). As noted above, however, until we have good estimates of *model* errors, the error estimates produced by more sophisticated approaches are unlikely to be meaningful. (Indeed, Marotzke and Wunsch 1993 posed the model as a strong constraint, thus removing the possibility of model errors in their analysis.) Conversely, the advantage of a simple assimilative approach is that it becomes much easier to see which elements of the model physics produced the responses observed (although, unfortunately, we do not possess an adjoint for the model used).

We carry out a series of idealized identical twin experiments. The sensitivity of the inversion to the nudging coefficient, and to the amount of data available, is investigated. Information assimilated in a localized region may be expected to influence the entire domain through advection and the propagation of Rossby waves, and we investigate the spreading of information away from a localized region. Experiments in a realistic domain are completed with strong nudging with a view to constraining the model to be close to the observations in the region that data are available. We discuss how effective is the combination of the model and the (restricted) data for reproducing the climatological fields of the North Atlantic.

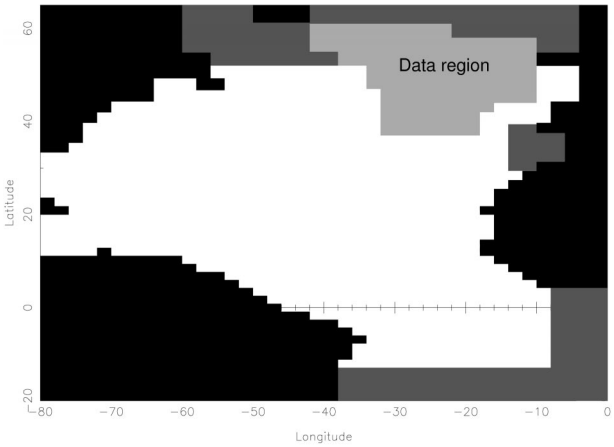


FIG. 3. Model geometry (black), relaxation zones (dark gray), and the region in which data are available from the CONVEX and Vivaldi cruises (light gray).

2. The numerical model

There are two available frictional planetary geostrophic numerical models; those of Edwards (1996), and Samelson and Vallis (1997). The requirement that the model handle varying coastline and topography restricts the choice to Edwards (1996), which has also been used to study the stability of the thermohaline circulation by Edwards et al. (1998). The governing equations are the thermoline equations, with the addition of linear drag in the horizontal momentum equations to permit western boundary currents. Convection and both horizontal and vertical diffusion of heat and salt are included. Full details of the physics and the numerical scheme used are described in Edwards (1996).

Conditions of no normal flow and no diffusive flux are applied at the solid boundaries. The latter condition is applied through the mathematical device of specifying horizontal diffusivities that vanish at the boundaries, thus avoiding the necessity of raising the order of the equations in order to accommodate the no diffusive flux

conditions. A full discussion of this condition is contained in Edwards (1996).[2]

The surface heat flux is applied through a pseudo-Haney forcing term (Haney 1971), which relaxes the temperature in the top model level to the Levitus (1982) annual climatological sea surface temperature. A surface salt flux is also prescribed, derived from the surface evaporation minus precipitation ($E - P$) field. The salt flux is zonally uniform, with the form of a single cosine structure in latitude between values of $-S_0$ at the southern boundary and S_0 at the northern; the amplitude S_0 is equivalent to an $E - P$ field with a maximum magnitude of 1 m yr^{-1} at 35 psu at the northern boundary.

The inversion approach

As indicated above, the inversion is performed using a simple assimilation scheme in the frictional planetary geostrophic system, using the nudging technique first introduced to meteorology by Anthes (1974), and later applied to oceanographic data assimilation by Verron and Holland (1989) and Holland and Malanotte-Rizzoli (1989). The general formulation for assimilating observations from N stations of a prognostic variable f of the model is given by the equation

$$\frac{\partial f}{\partial t} = \text{rhs} - \sum_{n=1}^{N} G(\varepsilon_n, \delta t_n, \delta r_n)(f - f_n^{\text{obs}}). \qquad (1)$$

Here, the rhs represents all other terms in the evolution equation for f, and f_n^{obs} is the nth observed value of f for $n = 1, \ldots, N$. The relaxation function G is an inverse timescale, which is in general a function of the standard deviation of the nth observation ε_n; the separation of the time of the observation from the model evaluation time δt_n; and the distance of the observation point from the grid point δr_n. In general, G is also connected with the time-stepping scheme of the model.

We choose the simplest formulation, following Anthes (1974), where

$$G = \begin{cases} \text{constant} > 0 & \text{for } \delta r_n = 0, \\ 0 & \text{otherwise.} \end{cases} \qquad (2)$$

This nudging technique has proved successful in oceanographic applications, particularly in the context of reproducing deep flows by the assimilation of altimetric data (Holland and Malanotte-Rizzoli 1989; Haines 1991). A major advantage of this nudging scheme is simplicity, which is particularly advantageous when incorporating the scheme into an existing ocean model. The drawback, however, is that a rigorous assessment of errors and consistency is not possible when reconstructing an unknown ocean flow.

The simple assimilation of hydrographic data into the model described in this section requires the addition of nudging terms to the equations for both T and S. These terms take the form $-\alpha (T - T^{\text{obs}})$ and $-\alpha (S - S^{\text{obs}})$, respectively, where α is the constant nudging coefficient that determines the relative weighting given to the model dynamics and the observational data. Since our aim is to reconstruct the climatological fields as reflected in the hydrographic data, we choose α so as to constrain the model to be close to the observed data. Note that $\alpha = 1/\Delta t$ (where Δt is the time step) corresponds to neglecting the model state in favor of the data in all but the advective and diffusive terms, and hence we will usually choose α to be close to $1/\Delta t$. This choice is almost equivalent to direct insertion of the ''observed'' data in the model. We show below that weaker relaxation (e.g., of the size employed by Malanotte-Rizzoli and Young 1995) gives slower convergence but otherwise has similar properties.

3. Identical twin experiments

In order to assess the approach, we perform some ''identical twin'' experiments, in which artificial data generated from the model itself are used in the inversion (Daley 1991). The model is spun up to a near–steady state, then data from this final state are assimilated in subsequent runs, and the convergence toward the final state is monitored. Since the model dynamics and the artificial data are necessarily in agreement, identical twin experiments are inevitably optimistic. They do, however, give an indication of the feasibility of the method and the dependence of the scheme on factors such as the distribution of the data. These latter sensitivity tests could be equally effectively carried out with an adjoint of the frictional geostrophic model, but we do not possess this (and given the instant vertical mixing implied by the convection scheme, an adjoint would be difficult to produce).

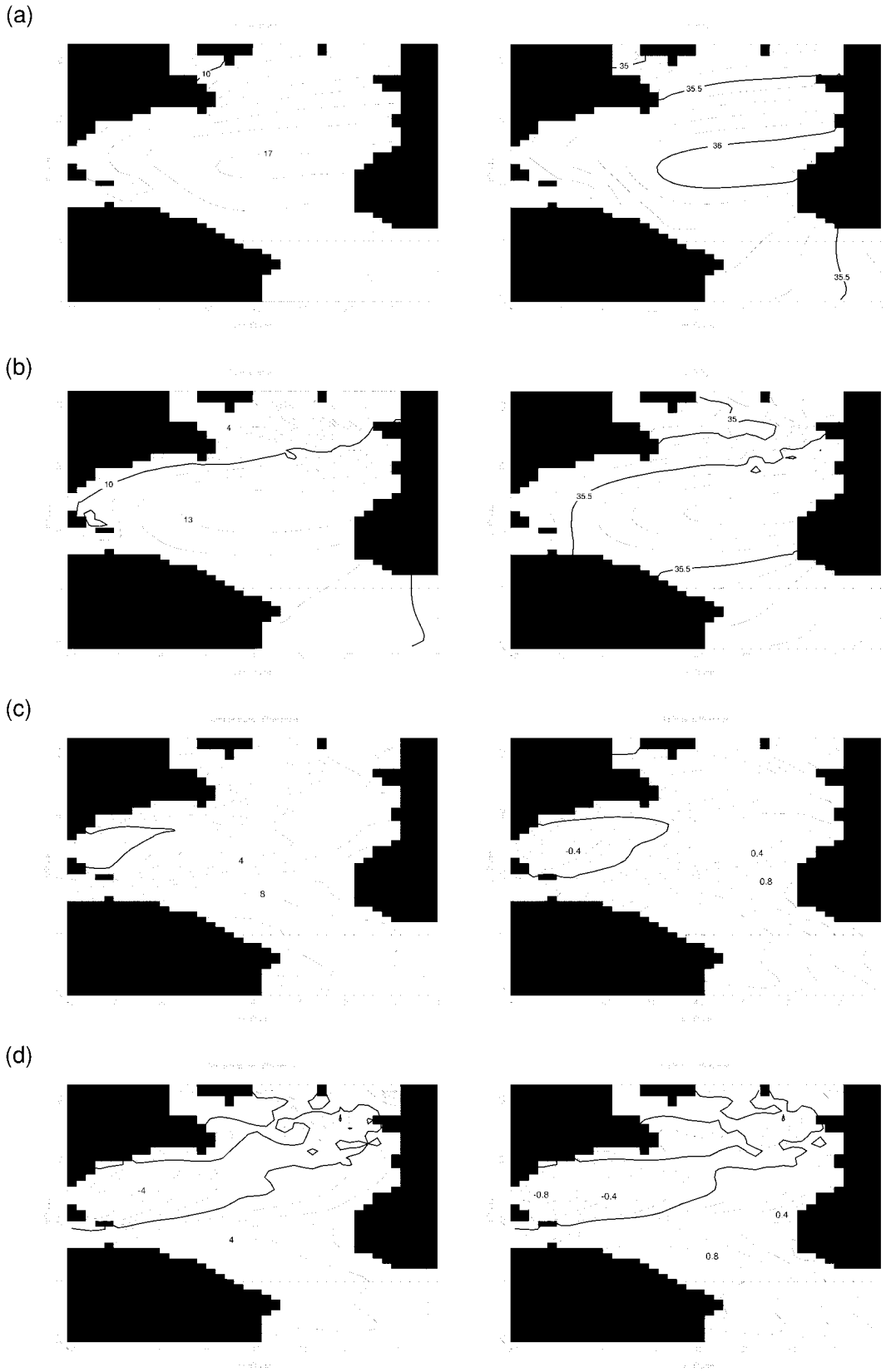
a. Model configuration

The model described in section 2 was configured to run in a 40°-wide sector ocean, extending from $\theta_s = 20$ to $\theta_n = 60$°N. The model grid contains 20×20 horizontal grid points, equally spaced in longitude and in $\sin\theta$, where θ is latitude. Note that with this choice of grid, all grid boxes on each vertical level have equal volume. Sixteen equally spaced vertical levels span the ocean depth of 4000 m.

The equation of state used is the United Nations Educational, Scientific, and Cultural Organization formula, which gives a close approximation to the ocean density in kilograms per meter cubed (Gill 1982); for the identical twin runs, a polynomial approximation was used, and the exact formula for the later more realistic calculations.

The surface wind stress is entirely zonal ($\tau^s = 0$), and has the form

[2] Samelson and Vallis (1997) handle the issue differently; providing the model has sufficiently coarse resolution, tests with idealized ocean basins show little difference between the codes.



$$\tau^{\phi} = \tau_0 \cos\left[\frac{2\pi(\theta - \theta_s)}{\theta_n - \theta_s}\right] \quad (3)$$

for latitudes $\theta_s \leq \theta \leq \theta_n$, where the amplitude τ_0 is 0.07 N m^{-2} as used by Edwards (1996). Other parameters are $\lambda = (1.25 \text{ days})^{-1}$, $\kappa_H = 3185$ m^2 s^{-1}, $\kappa_V = 2 \times 10^{-4}$ m^2 s^{-1}; the relaxation time for the Haney forcing at the surface is 61 days.

The model is initialized with a uniform salinity of 35 psu throughout the ocean, and with a stable linear stratification for T, with $T = 10$°C at the surface and $T = 5$°C at the ocean floor. An alternative initial condition in which the linear stratification of T is replaced by a uniform distribution with $T = 15$°C throughout the ocean was used for some runs.

b. Experiments and results

The artificial data for use in the identical twin experiments were generated by running the model from rest for 100 yr, then storing the temperature and salinity at all grid points. Runs were then performed to investigate the effect of assimilating (i) using different subsets of the data, (ii) using data containing artificial noise, (iii) using different initial conditions, and (iv) using different values of the nudging coefficient. Another run was completed to investigate the mechanisms by which information travels around the ocean. A summary of the runs performed is shown in Table 1. Since all the runs carried out were relatively short, and hence computation time was not an issue, a small time step of 5.0×10^{-4} yr (4.4 h) was used for all runs.

To assess the impact of the assimilation, the difference between the model temperature and salinity and those of the final state were monitored throughout the integrations, and these differences were quantified by calculating the unweighted rms difference between model and steady state over all the grid points.

For comparison, the convergence of the control run was measured by rerunning the model in the identical configuration to that from which the data was taken, and monitoring the convergence to its final state. Hence the rms difference for the control run decreases monotonically, reaching zero after 100 yr.

The results obtained were similar for both the temperature and salinity fields, and hence we focus attention on the convergence of the model temperature to the final state from which the data are taken. Figure 1 shows the rms difference in the temperature field from the final state with and without assimilation. The assimilated data are taken from the final state from the control run (run id00), which takes 100 yr to reach this state. Random noise is added to the data being assimilated to simulate the noise occurring in real datasets. This noise is generated by multiplying the data value by a random factor between 0.8 and 1.2. In run id01, the final temperature and salinity data are assimilated continuously at every grid point with a strong nudging. Clearly, the convergence toward the final state is extremely rapid when the data are assimilated, with the model approaching the final state in a matter of days. However, the model settles into a final state slightly different to the original final state. The data and the model dynamics are not exactly consistent due to the addition of noise to the data, and furthermore the state toward which the assimilation is nudging the model is not in exact agreement with the original final state, so the existence of a small residual difference in final states is not surprising. In an equivalent run without noise added to the data, the assimilation converges at the same rate but no residual differences occur.

The effect of assimilating a subset of the data, and using different initial conditions, is also shown in Fig. 1. The impact of assimilation using only even indices (not shown) was also extremely rapid, converging in a little over a month. Run id02 assimilates data on only two north–south and two east–west transects, at one-third and two-thirds across the basin, from an alternative initial state. After a year, the temperature field has not quite reached the final state; despite the fact that only a small subset of the data is being used, convergence is still considerably more rapid than in the control run. Again, a small residual difference from the final state remains due to the noise added to the data; such differences are not seen in an equivalent run using data without noise. Interpretation of convergence rates must be made with care, as Lee (2001) has shown that tracer fields in turbulence naturally tend toward control solutions on the Rhines–Young timescale. The calculations here differ in several respects from her runs (the flow is neither specified nor turbulent, and there is active insertion of data), but nonetheless the timescale $(L/U)(\kappa_H/UL)^{1/3}$, about 0.6 yr, would appear only slightly too large from Fig. 1.

Experiment id02 was repeated using different values of the nudging parameter α, also displayed in Fig. 1. Reducing the nudging coefficient by a factor of 10 (id03) clearly has little effect on the convergence of the solution to the final state; essentially the nudging remains direct insertion. A further reduction to 1/100 times the original value (id04) greatly slows convergence, although still considerably quicker than in the control run. In this case, with relatively weak nudging toward noisy data on two north–south and two east–west sections

←

FIG. 4. Temperature and salinity at $z = 530$ m for (a) control run, (b) assimilation run, (c) control run − Levitus climatology, and (d) assimilation run − Levitus climatology. Contour intervals are 1°C and 0.1 psu. In (a) and (b), firm lines mark 10°C and 35, 35.5, 36 psu. In (c) and (d), the firm line marks the zero contour, and dashed contours are negative.

alone, starting from different initial conditions, significant improvement is seen in the convergence of the model towards the final state over the control run.

To investigate the mechanisms by which assimilation in a localized region influences the solution in the whole domain, an additional run, id05, was performed in which assimilation was confined to a single vertical profile, located in the center of the domain. This has conceptual similarities to the use of representers (Bennett 1992), though without the accompanying statistical advantages: in both cases, the idea is to examine the impact on the ocean physics of the act of assimilation.

Figure 2 shows plots of the distance of the model temperature from the final temperature as a proportion of the distance of the initial model temperature from the final temperature at various times throughout the integration for the control run (left panel) and for run id05 (right panel). This proportion (plotted at depth $z = 2000$ m), therefore, has a uniform value of 1 at all points at the beginning of the integration, and decreases to zero as the model converges to the final state.

Figure 2a shows the fields shortly after the integration has begun ($t = 1.2$ months). Both the control run and the assimilation run show a rapid adjustment near the boundary of the domain, while the signal of the data is evident in the assimilation run. This signal is asymmetric about the data point, extending farther westward due to the propagation of long Rossby waves. After 3 months (not shown) the rapid adjustment near the boundary in both runs has continued, with the most rapid change occurring in the northwest corner and spreading outward from this region. Information from the data point has reached the western boundary, suggesting a propagation velocity of around 0.2 m s^{-1}. There is also evidence of a signal from the data point traveling anticlockwise around the southern boundary, and subsequently up the eastern boundary.

In Fig. 2b ($t = 6$ months), this signal has reached the northeast corner of the domain. The northwest region, in contrast, is unaffected by the signal from the data point. Indeed, there is little evidence of any effect on the western boundary to the north of 50°. The more rapid progress of the signal around the boundary than in the interior may be evidence of a boundary wave passing around the domain, while the small amount of progress of the signal to the north is likely to be due to diffusion.

After 1 yr (not shown), the signal has reached around the boundary to the north, while the interior is now everywhere affected by the assimilation, with information appearing to have leaked from the boundary into the interior. In addition, advection may be playing a role in spreading the information around the interior. Only the northwest corner of the domain remains insensitive to the signal from the data. After approximately 2 yr (Fig. 2c) the signal begins to have a significant impact on this region.

Finally, after 5 yr (Fig. 2d) the entire domain feels significant effects due to the introduction of the data. Clearly, the spreading of information from the data around the domain depends upon Rossby wave propagation, advection, diffusion, and boundary waves. The boundary waves are of numerical origin, arising from the discretization of the governing equations near a boundary, with speeds typically 0.2–0.4 m s^{-1} (Killworth 1985). Both propagation speeds for Rossby waves and advective speeds are typically 0.01–0.05 m s^{-1} in the North Atlantic (Killworth et al. 1997), and diffusion would take around 150 yr to cross the basin, implying a typical velocity of $O(10^{-4})$ m s^{-1}. (The linear drag is equivalent to a diffusion of the Rossby wave propagation, with a coefficient about half the applied κ_H.) All of these processes, with the addition of topographic Rossby waves in an ocean with topography, help to ensure that the data from a single profile is able to rapidly influence the entire domain.

4. Data assimilation experiments

We now apply the model and assimilation scheme described in section 2 to the simulation of the North Atlantic circulation.

a. Model configuration

1) MODEL DOMAIN AND GRID

The model domain extends over the North Atlantic from 20°S to 65°N, and from 0° to 80°W. A grid of 40 × 42 points in the horizontal, and 16 vertical levels was used. The grid points are equally spaced in longitude, with a resolution of 2°, and equally spaced in the sine of latitude, giving a resolution ranging from 1.8° in the south to 3.5° at the north, and an average resolution of 2°. The vertical levels are logarithmically spaced, with the depth (m) of the kth density level from the surface given by

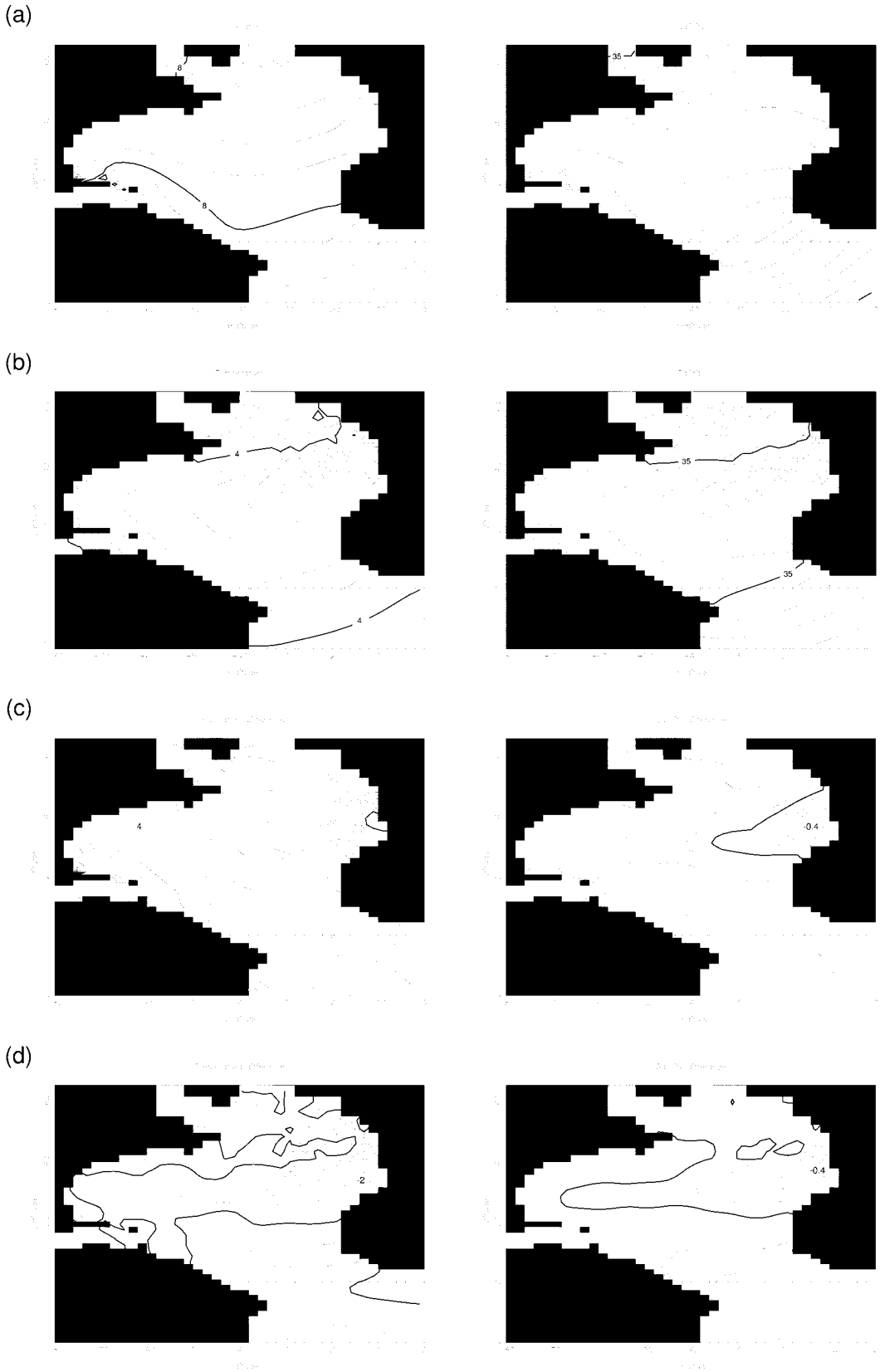
$$d_k = 550\ [11^{(2k-1)/2N_k} - 1], \tag{4}$$

where N_k is the number of vertical levels. The vertical extent of the bottom layer is 10 times that of the top layer.

The topography was obtained from the Earth Topography Five Minute Grid (ETOPO5) dataset; the depth

→

FIG. 5. Temperature and salinity on $z = 1420$ m for (a) control run, (b) assimilation run, (c) control run − Levitus climatology, and (d) assimilation run − Levitus climatology. Contour intervals in (a) and (b) are 0.25°C and 0.025 psu. In (a) and (b), firm lines mark 4° and 8°C and 35, 35.5, 36 psu. In (c) and (d), the contour intervals are 0.25°C and 0.1 psu. The firm line marks the zero contour.



at each point was the median of the ETOPO5 data falling within that box, converted to the nearest vertical level depth. Channels of width less than three grid points were manually removed, while attempting to preserve the coastline as closely as possible. Any islands occurring in the interpolated topography were lowered to below the topmost grid point.

2) SURFACE FORCING

The formulation of the surface heat flux is the same as that used for the identical twin runs described in section 3. Runs were also completed using the alternative formulation of Barnier et al. (1994). However, perhaps because the model is rather coarse, the results obtained were relatively insensitive to the surface forcing used. The surface salt flux condition described in section 2 is replaced by a relaxation of the model sea surface salinity toward the climatological values of Levitus (1982).

We consider forcing the model using two different sets of wind stress data. First, the monthly analyses of Hellerman and Rosenstein (1983) have been averaged to give an annual mean wind stress. The second set of wind stress data uses European Centre for Medium-Range Forecasts (ECMWF) monthly mean wind stress data, produced from an analysis of data from 1980–89, averaged to give an annual mean.

3) INITIALIZATION

The model has been initialized using the Levitus (1982) annual climatological temperature and salinity data. Initialization shocks, often problematic in numerical model initialization, are avoided here for two reasons. First, the numerical model does not permit gravity waves, which can be generated during initialization in more complete dynamical models. Second, since the runs to be undertaken are relatively short, and hence computation time is not a major restriction, we are able to choose a time step short enough to avoid initialization problems associated with nudging the model strongly at the beginning of the integration.

4) RELAXATION ZONES

Initial tests with the model showed that its steady state possessed poor water mass structure. As in any single-basin model experiments, one can either accept that the lack of dense water forcing at high latitudes will induce unrealistic thermohaline structure, or add in buffer zones to emulate the dense water production. Since we wish to minimize the degree to which the model is forced toward the data, it is necessary to have a model that demonstrates some realism. Accordingly we chose to have regions where temperature and salinity were restored to Levitus (1982) values, with an inverse timescale varying linearly over four grid points from 0.5 day^{-1} at the boundary to 3.5 day^{-1} at the innermost of the four grid points. The location of these regions is shown in Fig. 3, which also serves to show the model geometry. These regions include the northern and southern boundaries (to account for interbasin flows) and the Mediterranean outflow, without which the North Atlantic salinity will be incorrect. Such zones are common in basin models.

b. The data

The data used in the experiments come from two separate datasets, both from surveys of the North Atlantic carried out in 1991. Data from the CONVEX expedition (Pollard et al. 1996) cover the subpolar northeast Atlantic, around 54°–60°N and 0°–40°W, with observations at 95 locations, typically at 50-km spacing. The second dataset, from the Vivaldi study (Read and Gould 1992), is from 37 more widely spaced locations, typically 200 km apart, and covers the region around 35°–54°N and 10°–40°W. Both datasets contain full-depth profiles of temperature and salinity. The region in which data are available is also shown in Fig. 3.

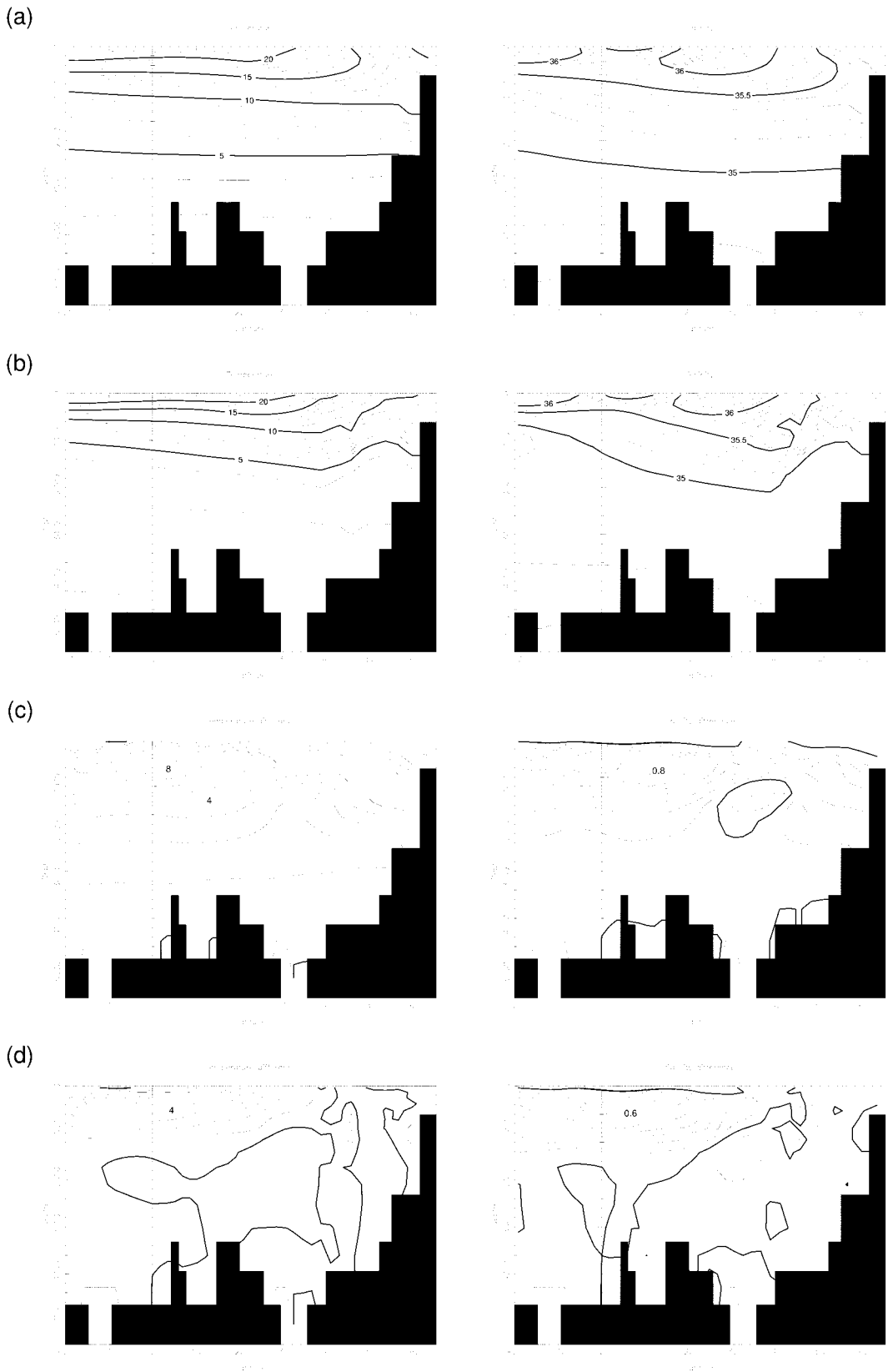
Since the model grid is relatively coarse, and we therefore have a large number of data values falling within each grid box, we generate a set of data at each model grid point from the available data. To obtain data values at model grid points we first obtain data values from each profile at each model level by linear interpolation. Then at each level the median value of the profiles falling in a grid box is taken to be the data value for that grid box. Where a grid box with data lies within the relaxation zones described in section 4a(4), the relaxation to Levitus is turned off. This avoids direct competition between the assimilation and the relaxation at any individual grid point.

c. Experiments and results

We performed 50-yr integrations, since Döscher et al. (1994) show that 10–15 yr are required merely for the meridional overturning to come to quasi equilibrium. Assimilation times of this length are not yet possible for more complete models, so that the deep T–S struc-

→

FIG. 6. Meridional sections of temperature and salinity on 20°W for (a) control run, (b) assimilation run, (c) control run − Levitus climatology, and (d) assimilation run − Levitus climatology. Contour intervals are 1°C for all temperature plots. Salinity contour interval 0.025 psu for (a) and (b), and 0.1 psu for (c) and (d). The firm lines mark 5°, 10°, 15°, and 20°C and 35, 35.5, 36 psu in (a) and (b). In (c) and (d) the firm line marks the zero contour and dashed contours are negative.



ture in such models is unlikely to be completely assimilated. (This is not a problem for an adjoint model, but coarse models run far from climatology, as previously noted.) We tried three combinations of forcing fields. First, control and assimilation experiments were carried out using the Haney relaxation conditions on T and S together with the Hellerman and Rosenstein (1983) wind stresses. Similar experiments were performed using the ECMWF wind stresses. These second experiments were repeated using the Barnier et al. (1994) formulation for the surface heat flux. In all runs, the remaining parameters were as in the identical twin experiments.

Due to the coarse resolution of the model configuration, the results obtained proved to be relatively insensitive to the surface forcing used, with the rms difference in temperature and salinity between any runs (both control or both assimilation) being less than 0.18°C and 0.03 psu, respectively. In view of this insensitivity, we subsequently focus attention on the results from the first pair of runs, using the Hellerman and Rosenstein winds.

The net horizontal circulation differs naturally between control and assimilation; the range of values of the barotropic streamfunction were, respectively, −0.2 to 8 Sv and −18 to 14 Sv. Both values are small in comparison to those predicted by finer-resolution models (e.g., Semtner and Chervin 1992), which are able to resolve features such as the Gulf Stream and its associated large transport. In both runs the subtropical gyre is well-developed; the subpolar gyre is somewhat indistinct in the control, but much stronger than the subtropical gyre in the assimilation.

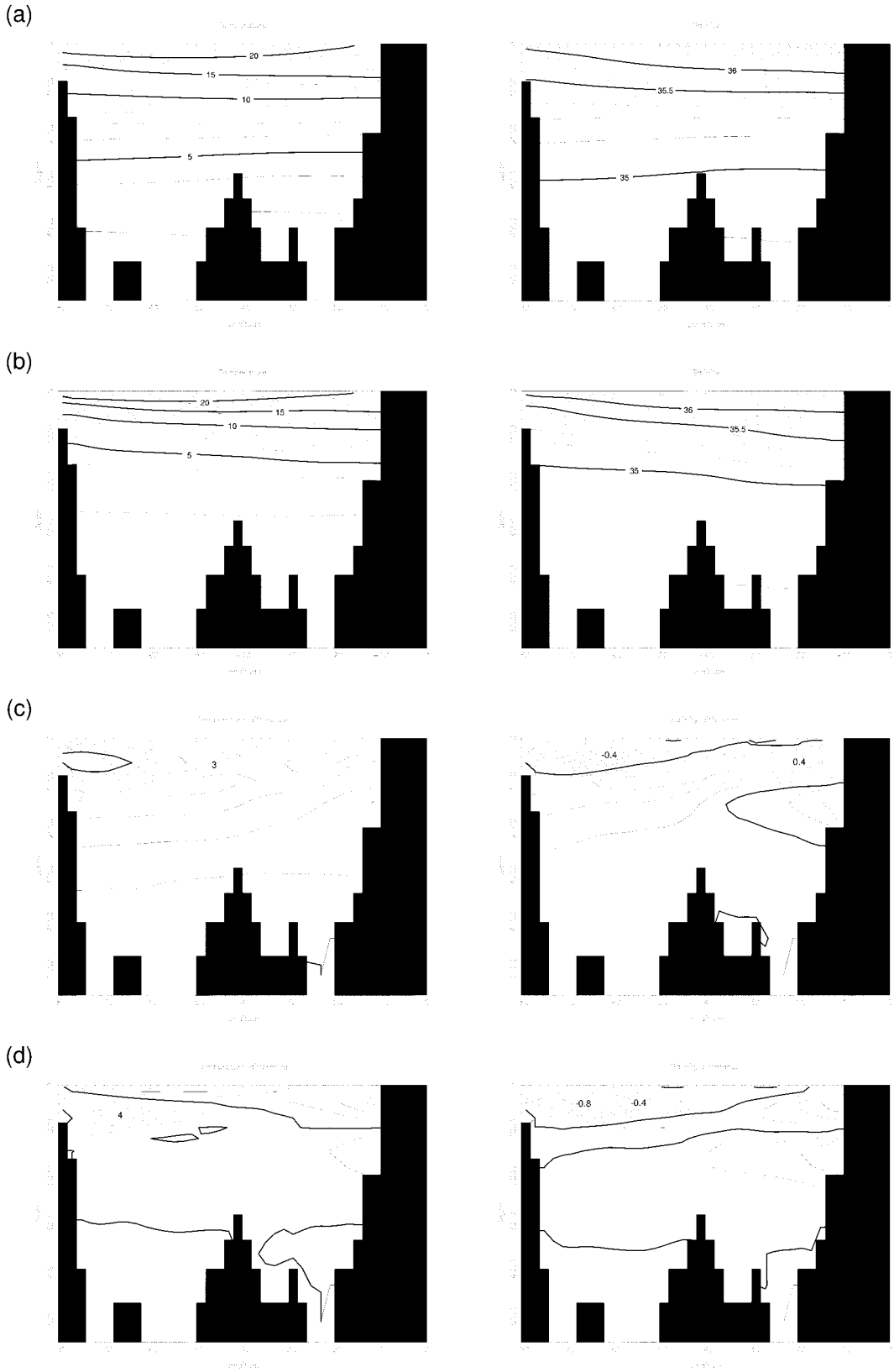
We make the assumption that similarity to climatology is a reasonable measure for the success or failure of the inversion. We have noted that coarse-resolution ocean and climate models are not particularly successful at reproducing climatology, frequently possessing too weak a vertical heat and salt gradient produced by too much vertical mixing. The control run is shown below to share this feature. Were we assimilating data from the entire basin, this measure would be questionable (one should, after all, move closer toward something that is being assimilated). However, in our case the model is generating most of the information, with noisy, temporally restricted data confined to a small region, and we will be interested in the effect of these confined data, where they have an impact, and whether they improve the inversion. Nonetheless, if the Levitus smoothing and whatever smoothing is implied by our model are on different space scales, it is unclear that comparison with climatology is a perfect measure.

Near-surface temperatures and salinities (not shown) are constrained by the surface relaxation condition to be close to the Levitus climatology. Figures 4a,b show temperature and salinity, and Figs. 4c,d show and temperature and salinity differences from the Levitus fields on $z = 530$ m. Both the control and assimilation runs fail to reproduce the structure of the Levitus temperature and salinity at this depth. The model temperatures have a maximum value in the center of the domain, while the Levitus temperature has a maximum located in the far west of the domain; the model salinities have a maximum on the east coast, and the Levitus salinity in the far west. This is emphasized by the differences from the Levitus fields, which show that both control and assimilation runs have cold, fresh errors in the far west. The sharp gradients of temperature and salinity poleward and equatorward of the maxima in the west are not reproduced. The model resolution precludes the development of relatively narrow features such as pronounced western boundary currents, and the associated sharp temperature and salinity gradients, and hence the failure to reproduce the slightly sharper gradients seen in the Levitus climatology is not surprising. Comparison of the errors relative to Levitus for the control and assimilation runs shows that the assimilation has reduced the errors over much of the domain through a large-scale cooling and freshening. However, errors in the far west have been increased by the assimilation. The control run is generally (and erroneously) warmer and more saline throughout, and hence comes closer to the maxima in the west than the assimilation run. The reduction in errors due to assimilation near and to the south of the equator are large, despite being induced by data insertion in the north.

Quantitatively, there are significant differences between the temperature and salinity values of the control run and those of the Levitus data. This is, of course, a common problem with coarse-resolution ocean models, which lack the more complete dynamics necessary to reproduce observations accurately (including sharp gradients, although the Levitus data are smoothed). Errors are partly kept in check by the relaxation zones; preliminary calculations without such zones showed temperature differences at high latitudes between control and Levitus of up to 6°C. Despite the zones, at high latitudes the control run is several degrees too warm, and the salinity about 0.3 psu too salty. In contrast, the assimilation run shows realistic values for both temperature and salinity at high latitudes, where the northward deflection of temperature and salinity contours is also reproduced. Over much of the North Atlantic, the

→

FIG. 7. Zonal sections of temperature and salinity on 30°N for (a) control run, (b) assimilation run, (c) control run − Levitus climatology, and (d) assimilation run − Levitus climatology. Contour intervals are 1°C for all temperature plots. Salinity contour interval 0.025 psu for (a) and (b), and 0.1 psu for (c) and (d). The firm lines mark 5°, 10°, 15°, and 20°C and 35, 35.5, 36 psu in (a) and (b). In (c) and (d) the firm line marks the zero contour and dashed contours are negative.



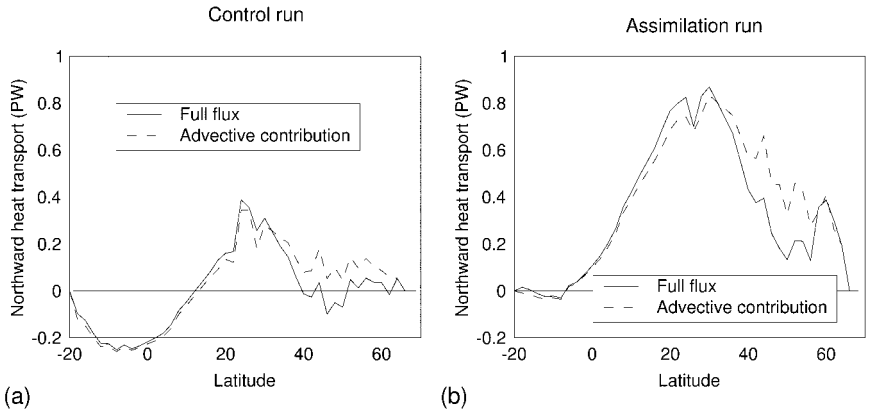


FIG. 8. Northward heat transport (PW) for (a) control run and (b) assimilation run. The solid line shows the transport, and the dashed line the advective contribution, which dominates in most regions (except high latitudes in the assimilation case).

assimilated results are not much further from climatology than even the 1/2° results of Semtner and Chervin (1988), which are frequently of order 1°–2° too warm or cold in areas where climatology would be an acceptable solution, for example, far from frontal and other active regions. In the equatorial region, however, both the control run and the assimilation run are considerably warmer and saltier than the climatology. (The linear drag addition to geostrophy does not give a good representation of equatorial dynamics.)

The failure of the control run to reproduce the structure of the Levitus fields is again evident on $z = 1420$ m (Fig. 5). In the Levitus temperature and salinity the Mediterranean outflow is evident as a tongue of warm, salty water extending from the eastern coastal boundary, throughout the mid-Atlantic. This structure is absent from the fields from the control run (Fig. 5a). The assimilated run, however, shows a strong tongue, although with distinctly too weak a salinity signal. This is remarkable, since both control and assimilation possess identical relaxation zones near Gibraltar.[3] The deficiencies in the representation of the outflow are clear in the differences from the Levitus fields. The control run (Fig. 5c) has temperatures that exceed even the temperature maximum within the outflow, and hence has large warm errors, which increase away from the outflow. The assimilation run (Fig. 5d), in contrast, does not represent the maximum temperature in the outflow, but shows smaller cold errors decreasing away from the outflow. Although neither run captures the salinity maximum, errors in the assimilation run decrease away from the outflow, while those for the control run decrease, then increase again as the fairly uniform salinity field fails to capture any of the structure of the outflow.

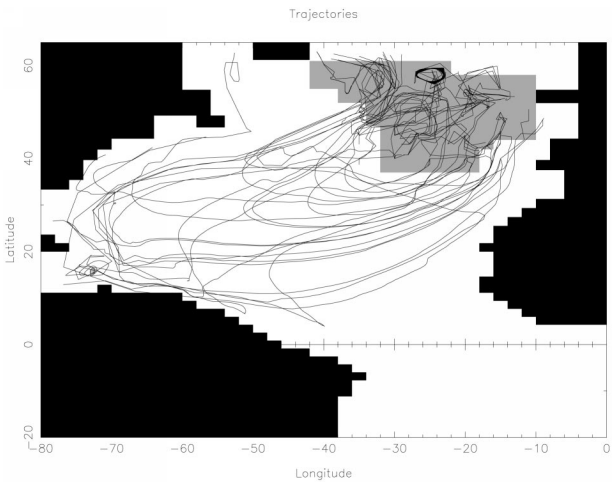


FIG. 9. Fifty-year trajectories projected onto a horizontal plane, integrated away from the assimilation region (gray box) using the model assimilation run velocity field.

Away from the Mediterranean outflow region, the control run is again too warm and too saline throughout, and fails entirely to reproduce either the qualitative or quantitative structure of the Levitus fields. In contrast, the errors in the assimilation run are significantly smaller, with reasonable values of temperature and salinity. Both temperature and salinity show the westward decrease in magnitude, as evident in the Levitus fields, albeit with far smaller gradients.

Figures 6a,b show meridional sections of temperature and salinity, and Figs. 6c,d show temperature and salinity differences from Levitus on 20°W, where the domain extends to the southern boundary. The control run (Fig. 6a) shows poor agreement with the Levitus fields everywhere, with temperature and salinity errors as large as 9°C and 1 psu, respectively, in upper layers at low latitudes (Fig. 6c). These errors, which extend down to the ocean floor, are due to the thermocline being far

[3] Note that such a relaxation zone demonstrably does not necessarily produce correct water masses; relaxation can induce a pressure contrast that induces flow around, rather than through, the relaxation zone, inducing a failure to create correct water masses.

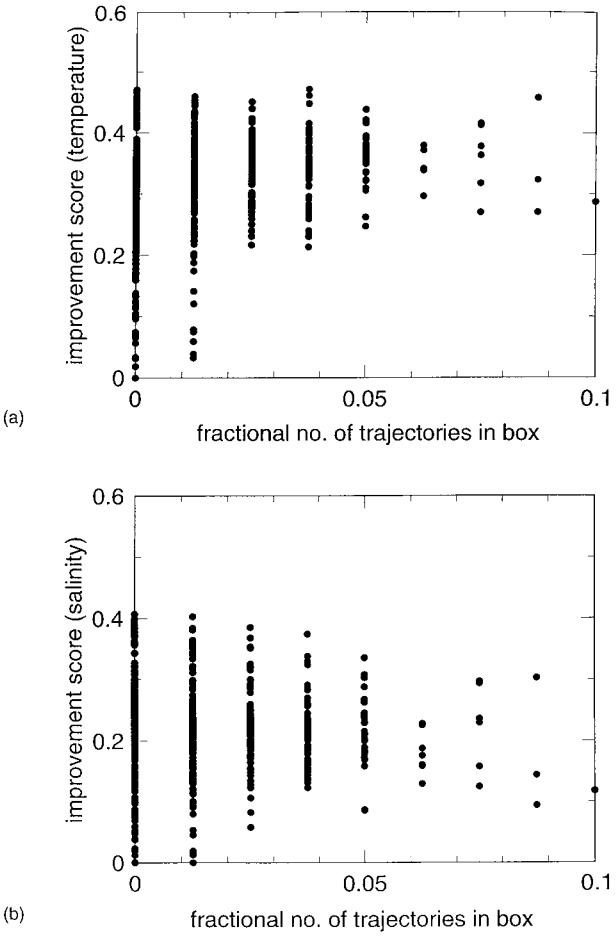


FIG. 10. Scatterplot of improvement produced by assimilation against fractional number of trajectories passing through a horizontal grid box; trajectories within the data assimilation region are not shown.

too deep and diffuse. Errors increase toward the south, where the tightening of thermocline in equatorial regions seen in the Levitus fields is not reproduced.

The assimilation run (Fig. 6b) shows better agreement at most latitudes, with a much shallower thermocline and halocline; the cooling and freshening impact of the CONVEX data at depth has a clear effect throughout the basin (even at the southern boundary, perhaps due to coastal wave propagation). Errors are reduced throughout (Fig. 6d), particularly at depth, and dramatically so at high latitudes. However, despite the tightening of gradients in the thermocline due to the assimilation, large errors remain at thermocline depths where the vertical resolution of the model precludes the representation of the sharp gradients present in the Levitus fields.

Zonal sections of temperature and salinity, and temperature and salinity differences from Levitus, through the northern half of the domain (30°N; shown in Fig. 7), again show the improvements caused by the inversion. The assimilation run (Fig. 7b) shows sharper subsurface gradients of temperature and salinity and reproduces the shallowing of the deep isotherms and isohalines to the west present in the Levitus fields. The control run (Fig. 7a) does not reproduce the rapid variation with depth in the top 1500 m, and again exhibits deep temperature and salinity fields that are too warm and too salty at all longitudes. Similar contrasts are seen farther north. The error fields (Figs. 7c,d) highlight the improvement in the assimilation run, which shows smaller errors throughout much of the section. Again, the large errors in the control run arise from the poor representation of the thermocline and halocline. The main exception to this is in the west, between around 500- and 2000-m depth. In the Levitus fields, the near-surface isotherms and isohalines deepen toward the west. This is not reproduced by either run, leading to errors in both runs in the far west. Since the isotherms and isohalines in the control run are deeper throughout, errors in the far west are smaller than in the assimilation run.

Clearly, the assimilation of the data offers significant improvement in the ability of the model to produce a state of motion in which the temperature and salinity are consistent with climatology. (Equally clearly, the fit to climatology remains poor in many areas; as noted, coarse-resolution models do not naturally match climatology, and we only assimilate data from a small region into the model.) The improvement can be quantified by considering the rms difference between the model state and the Levitus climatology, averaged over all the ocean grid points. For the control run, the rms differences in temperature and salinity are 4.2°C and 0.42 psu, respectively; for the assimilation run they are 2.1°C and 0.30 psu, respectively. It is clear from Figs. 4–7 that the results vary significantly with depth, hence rms differences from Levitus have been computed for different level ranges, and are presented in Table 2. Root-mean-square differences in the surface layer are not greatly affected by the assimilation due to the surface relaxation to Levitus fields. The largest improvements due to the assimilation occur around thermocline depths, particularly in the 500–1000-m-depth range, where the assimilation produces significant tightening of vertical gradients and a corresponding reduction in rms temperature and salinity errors of 3.8°C and 0.23 psu, respectively. This improvement in the representation of the thermocline and halocline also impacts the rms errors in deeper level ranges, where the rms temperature and salinity errors of the control run exceed those of the assimilation run by 2.2°C and 0.10 psu, respectively.

Figure 8 shows the northward heat transport for the control and assimilation runs. The northward heat fluxes increase toward the center of the domain, with maximum values at around 20°N, of 0.4 and 0.85 PW, respectively. Heat fluxes from coarse-resolution models tend to be small compared with the estimates from global inversions and data (the flux at 24°N is believed to be about 1.2 PW; MacDonald and Wunsch 1996). Even

finer-resolution models are known to struggle to reach the observational estimate (Böning et al. 1995), with fluxes rather dependent upon north and south boundary conditions as well as eddy parameterizations in most North Atlantic models. (The effect of the southern boundary is clear in Fig. 8.) The effect of assimilation of the CONVEX data, as confined to the northern end of the basin, has thus been to lift the small heat flux natural to coarse-resolution models to more, but not completely, realistic levels. This has been achieved by an increase of the top-to-bottom temperature contrast. The assimilated flux agrees reasonably with results from other models (e.g., Semtner and Chervin 1992; Chassignet et al. 1996) and from inverse models (e.g., Wunsch 1984). Because there are considerable diabatic effects due to data insertion and relaxation, the northward fluxes are not necessarily related well to surface fluxes; however, as the dashed lines show, the majority of the heat flux is advective despite the coarse-resolution diffusion used.

To investigate the geographical spread of the influence of the data insertions, the trajectories from the final state velocity field were used to determine to what extent advection is able to carry information to regions of the ocean remote from the area of insertion.[4]

The projection onto a horizontal plane of the trajectories from 80 starting points within the data assimilation area (dashed box), integrated for the length of the model integration, are shown in Fig. 9. Trajectories extend throughout the mid-Atlantic, suggesting that advection plays an important role in the effectiveness of the assimilation. No trajectories reach near or south of the equator although, as observed in the identical twin experiments in section 3, advection is not the only process by which data can traverse the domain. It is likely that the propagation, and topographic steering, of Rossby waves may also play a significant role in transmitting information around the domain.

The effects of advection can be partly quantified. We define an improvement score due to assimilation as

$$\max\left(1 - \left|\frac{\text{climatology} - \text{assimilation}}{\text{climatology} - \text{control}}\right|, 0\right)$$

over the region of the ocean outside the data area.[5] These scores, averaged over a fluid column for temperature and salinity, reach values above 0.4 for salinity and almost 0.5 for temperature. The majority of the basin has temperature scores above 0.2 and scores around 0.2 for salinity, even in the South Atlantic. These scores are plotted against the fractional number of trajectories that pass through any given grid box (Fig. 10). (Although we cannot quantify the effects of wave propagation, the consideration of points only lying outside the data region means that modifications to propagation behavior due to data insertion have been removed.) When no trajectory reaches an area, there is still wide scatter in the scores (suggesting again the role of waves). However, for areas that trajectories reach, there is clear improvement in the temperature scores as the number of trajectories increases; the improvement is also visible in salinity, but less clearly.

The improvement over much of the domain due to the assimilation of data confined to a relatively localized area raises the question of how much data are required to offer significant improvement. To investigate this question further, an assimilation run was carried out, assimilating a single temperature and salinity profile. This yields a very strong subpolar barotropic gyre (44 Sv), suggesting that single gridpoint changes induce strong (and not well balanced) bottom pressure torque effects on the vorticity field. The predicted T and S are, intuitively, somewhat intermediate between the control and assimilation runs, showing better quantitative agreement with the Levitus fields than the control run, but not matching the improvement of the assimilation run. (The net ocean stratification is already dramatically improved even by one data point insertion, however.) The improvement scores are about 0.1 smaller for both temperature and salinity. We conclude that even a single profile has a clear impact on the inversion, in agreement with the coarse-grid identical twin experiments.

5. Summary and conclusions

Ocean inversion–assimilation experiments have been carried out using the frictional geostrophic ocean model of Edwards (1996) and Edwards et al. (1998), together with a simple nudging (nearly direct data insertion) scheme, with the aim both of providing a description of the ocean state consistent with model dynamics and the assimilated data. The data are entirely within the subpolar gyre, and the experiments have investigated how wide the effect of data can spread in a nongeostrophic model. The model used is thus somewhat superior to simple geostrophic balance used in most inversions, while the assimilative approach is less sophisticated than most in general use. The advantage of the approach is that the model can be run to a steady state with finite computational resources, which permits the effects of data insertion to spread far wider and to have a stronger impact than in assimilations of limited lengths (e.g., a year).

The scheme is tested in a series of identical twin runs, performed in an idealized flat bottom ocean, in which temperature and salinity data taken from the final state

[4] The trajectories shown here are not therefore a true representation of the advection of information during the spinup of the model. Indeed, during the early stages of a model integration, initialization produces large velocities as the model adjusts, so it is uncertain how information is advected during these early stages.

[5] There are grid points in which the control is very close to climatology, which would otherwise generate large negative values for the score; hence the cutoff at zero. Only 6% of the temperature values were worsened by assimilation, but as many as 20% of the salinity values.

of the model control run are assimilated, and the convergence to the final state is monitored. The assimilation offers a significant increase in the rate at which the model converges to the final state. This result is robust to changes in the value of the nudging parameter, to initialization further from the final state, and to the use of only a subset of the data, and the use of data containing artificial noise produces comparable results. Indeed, using only two zonal and two meridional data sections, with the data containing noise, and initializing the model further from the final state results in convergence to the final state in little over a year, compared to 100 yr for the control run. These results are consistent with previous successful application of nudging techniques (e.g., Haines et al. 1993).

The mechanisms by which the data influences the surrounding ocean are examined by assimilating at one data point only, and monitoring the model convergence. Results support the idea that Rossby wave propagation, advection, diffusion, and boundary processes all play a role in transmitting data around the domain.

The model is configured for a realistic North Atlantic topography and coastline, and data assimilation–inversion experiments are performed using hydrographic data from the CONVEX and Vivaldi expeditions, both confined to the northeast corner of the Atlantic. The temperature and salinity fields arising from the control run and the assimilation run are compared to the annual mean climatologies of Levitus (1982). The fields in both runs show good agreement near to the sea surface due to the surface boundary conditions. However, at greater depths there is much closer agreement between the assimilation fields and the climatological fields over the entire domain, with the improvement decreasing with distance from the region of data insertion. In the deep ocean, the control run is significantly too warm and too saline, features which are largely removed over the entire basin by the small area where data are provided. The assimilation produces an average improvement of 2°C (rms) per grid point.

The results are sensitive to the amount of data assimilated, and this is confirmed by assimilating a single data profile. In this case the results are intermediate compared to those obtained in the control run and those obtained assimilating all the data. However, a single data profile still offers considerable improvement over the control run, though it results in a much smaller reduction in the average rms difference from the Levitus temperature. Hence, considerable improvements in the inversion results can be achieved by essentially fixing the temperature and salinity depth profile at a single point, rather than any more complicated approach.

The northward heat transport of the assimilation run shows reasonable agreement with values computed from other numerical models (e.g., Semtner and Chervin 1992) and from inverse models (Wunsch 1996), and is significantly better than that of the control run, having double the value around 24°N.

The assimilation of noisy, synoptic, localized hydrographic data clearly has a dramatic impact on the ability of the coarse, frictional geostrophic model considered here to reproduce the climatological fields of the North Atlantic, although our choice of spatially confined data deliberately precluded an accurate reproduction over the entire ocean. As well as providing a useful means for obtaining an inverse estimate of the ocean state, the results presented here suggest that including hydrographic data to constrain the subsurface flow may drastically improve the assimilation of satellite altimetric or other data; this is in agreement with the findings of Malanotte-Rizzoli et al. (1989). In addition, the effect of assimilating local data are not just significant for the region in which the data are given, but may have an impact over the whole ocean basin through direct advection as well as wave propagation. This occurs because the model can be run long enough for information to spread over the whole domain, unlike more complicated, eddy-permitting models.

Acknowledgments. This work was partly carried out under a WOCE research studentship. We thank Andrew Willmott for help and encouragement.

REFERENCES

Anthes, R. A., 1974: Data assimilation and initialization of hurricane prediction models. *J. Atmos. Sci.*, **31,** 702–719.

Barnier, B., L. Siefridt, and P. Marchesiello, 1994: Thermal forcing for a global ocean circulation model from a three year climatology of ECMWF analyses. *J. Mar. Syst.*, **6,** 363–380.

Bennett, A. F., 1992: *Inverse Methods in Physical Oceanography.* Cambridge University Press, 346 pp.

Böning, C. W., W. R. Holland, F. O. Bryan, G. Danabasoglu, and J. C. McWilliams, 1995: An overlooked problem in model simulations of the thermohaline circulation and heat transport in the Atlantic Ocean. *J. Climate,* **8,** 515–523.

Chassignet, E. P., L. T. Smith, and R. Bleck, 1996: A model comparison: Numerical simulations of the north and equatorial Atlantic oceanic circulation in depth and isopycnic coordinates. *J. Phys. Oceanogr.,* **26,** 1849–1867.

Cunningham, S. A., 2000: Circulation and volume flux of the North Atlantic using synoptic hydrographic data in a Bernoulli inverse. *J. Mar. Res.*, **58,** 1–35.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

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

Döscher, R., C. W. Böning, and P. Herrmann, 1994: Response of circulation and heat transport in the North Atlantic to changes in thermohaline forcing in northern latitudes: A model study. *J. Phys. Oceanogr.,* **24,** 2306–2320.

Edwards, N. R., 1996: Unsteady similarity solutions and oscillating ocean gyres. *J. Mar. Res.*, **54,** 793–826.

——, A. J. Willmott, and P. D. Killworth, 1998: On the role of topography and wind stress on the stability of the thermohaline circulation. *J. Phys. Oceanogr.,* **28,** 756–778.

Evensen, G., 1994: Sequential data assimilation with a non-linear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99,** 10 143–10 162.

Fukumori, I., and P. Malanotte-Rizzoli, 1995: An approximate Kalman filter for ocean data assimilation: An example with an idealized Gulf Stream model. *J. Geophys. Res.*, **100,** 6777–6793.

Ghil, M., and P. Malanotte-Rizzoli, 1991: Data assimilation in me-

teorology and oceanography. *Advances in Geophysics,* Vol. 33, Academic Press, 141–266.
——, S. Cohn, J. Tavantzis, K. Bube, and E. Isaacson, 1981: Application of estimation theory to numerical weather prediction. *Dynamic Meteorology: Data Assimilation Methods,* L. Bengtson, M. Ghil, and E. Kallen, Eds., Springer-Verlag, 139–224.
Gill, A. E., 1982: *Atmosphere–Ocean Dynamics.* Academic Press, 662 pp.
Haines, K., 1991: A direct method for assimilation of sea surface height data into ocean models with adjustments to the deep circulation. *J. Phys. Oceanogr.,* **21,** 843–868.
——, P. Malanotte-Rizzoli, R. E. Young, and W. R. Holland, 1993: A comparison of two methods for the assimilation of altimeter data into a shallow water model. *Dyn. Atmos. Oceans,* **17,** 89–133.
Haney, R. L., 1971: Surface thermal boundary condition for ocean circulation models. *J. Phys. Oceanogr.,* **1,** 241–248.
Hellerman, S., and M. Rosenstein, 1983: Mean monthly wind stress over the World Ocean with error estimates. *J. Phys. Oceanogr.,* **13,** 1093–1104.
Holland, W. R., and P. Malanotte-Rizzoli, 1989: Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies. *J. Phys. Oceanogr.,* **19,** 1507–1534.
Killworth, P. D., 1985: A two-level wind and buoyancy driven thermocline model. *J. Phys. Oceanogr.,* **15,** 1414–1432.
——, 1986: A Bernoulli method for determining the ocean circulation. *J. Phys. Oceanogr.,* **16,** 2031–2051.
——, D. B. Chelton, and R. A. De Szoeke, 1997: The speed of observed and theoretical long extratropical planetary waves. *J. Phys. Oceanogr.,* **27,** 1946–1966.
Le Dimet, F. X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus,* **38A,** 97–110.
Lee, M.-M., 2001: Shear dispersion and error decay in idealized twin experiments. *Mon. Wea. Rev.,* in press.
Levitus, S., 1982: *Climatological Atlas of the World Ocean.* NOAA Professional Paper 13, 173 pp.
MacDonald, A. M., and C. Wunsch, 1996: An estimate of global ocean circulation and heat fluxes. *Nature,* **382,** 436–439.
Malanotte-Rizzoli, P., and R. E. Young, 1995: Assimilation of global versus local data sets into a regional model of the Gulf Stream system. 1. Data effectiveness. *J. Geophys. Res.,* **100,** 24 773–24 796.
——, ——, and D. B. Haidvogel, 1989: Initialization and data assimilation experiments with a primitive equation model. *Dyn. Atmos. Oceans,* **13,** 349–378.
Marotzke, J., and C. Wunsch, 1993: Finding the steady state of a general circulation model through data assimilation: Application to the North Atlantic Ocean. *J. Geophys. Res.,* **98,** 20 149–20 167.
Miller, R. N., 1986: Toward the application of the Kalman filter to regional open ocean modeling. *J. Phys. Oceanogr.,* **16,** 72–86.
Pollard, R. T., M. J. Griffiths, S. A. Cunningham, J. F. Read, F. F. Pérez, and A. F. Ríos, 1996: Vivaldi 1991—A study of the formation, circulation and ventilation of Eastern North Atlantic Central Water. *Progress in Oceanography,* Vol. 37, Pergamon, 167–192.
Read, J. F., and W. J. Gould, 1992: Cooling and freshening of the subpolar North Atlantic Ocean since the 1960s. *Nature,* **360,** 55–57.
Samelson, R. M., and G. K. Vallis, 1997: A simple friction and diffusion scheme for planetary geostrophic basin modes. *J. Phys. Oceanogr.,* **27,** 186–194.
Schlitzer, R., 1993: Determining the mean, large-scale circulation of the Atlantic with the adjoint method. *J. Phys. Oceanogr.,* **23,** 1935–1952.
Schott, F., and H. Stommel, 1978: Beta spirals and absolute velocities in different oceans. *Deep-Sea Res.,* **25,** 961–1010.
Semtner, A. J., and R. M. Chervin, 1988: A simulation of the global ocean circulation with resolved eddies. *J. Geophys. Res.,* **93,** 15 502–15 522.
——, and ——, 1992: Ocean general circulation from a global eddy-resolving model. *J. Geophys. Res.,* **97,** 5493–5550.
Tziperman, E., 1988: Calculating the time-mean oceanic general circulation and mixing coefficients from hydrographic data. *J. Phys. Oceanogr.,* **18,** 519–525.
——, W. C. Thacker, R. B. Long, and S.-M. Hwang, 1992: Oceanic data analysis using a general circulation model. Part I: Simulations. *J. Phys. Oceanogr.,* **22,** 1434–1457.
Verron, J., and W. R. Holland, 1989: Impacts de données d'altimitrie satellitaire sur les simulations numeriques des circulations generales oceaniques aux latitudes moyennes (in French). *Ann. Geophys.,* **7,** 31–46.
Woodgate, R. A., and P. D. Killworth, 1997: The effects of assimilation on the physics of an ocean model. Part I: Theoretical model and barotropic results. *J. Atmos. Oceanic Technol.,* **14,** 897–909.
Wunsch, C., 1978: The North Atlantic general circulation west of 50°W determined by inverse methods. *Rev. Geophys. Space Phys.,* **16,** 583–620.
——, 1984: An eclectic Atlantic Ocean circulation model. Part I: The meridional flux of heat. *J. Phys. Oceanogr.,* **14,** 1712–1733.
——, 1996: *The Ocean Circulation Inverse Problem.* Cambridge University Press, 442 pp.