Oceanic Data Analysis Using a General Circulation Model. Part II: A North Atlantic Model*

ELI TZIPERMAN

Department of Environmental Sciences and Energy Research, The Weizmann Institute of Science, Rehovot, Israel

WILLIAM CARLISLE THACKER AND ROBERT BRYAN LONG

Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida

SHOW-MING HWANG

Cooperative Institute for Marine and Atmospheric Sciences, Miami, Florida

STEPHEN R. RINTOUL

CSIRO Division of Oceanography, Hobart, Tasmania, Australia

(Manuscript received 5 December 1990, in final form 29 January 1992)

ABSTRACT

A general circulation model and North Atlantic climatological data of temperature, salinity, wind stress, evaporation minus precipitation, and air–sea heat fluxes are used to examine the possibility of solving inverse problems using a full-scale numerical GCM and real oceanographic data, combined through an optimization approach.

In this study several solutions for the model inputs and the structure of the cost function as a function of the model inputs are examined to demonstrate two of the main difficulties confronting such large-scale nonlinear inverse problems (about 30 000 unknowns and a similar number of constraints for the problem examined here). The first is the possible existence of local minima of the cost function, which prevents convergence of the optimization to the global minimum representing the desired optimal solution for the model inputs. The second difficulty, which seems the dominant one for many of the problems examined in this part as well as in Part I, is the ill conditioning of the inverse problem. Simple model equations are used to analyze the conditioning of the optimization problem and to analyze the role of both dissipation and waves in the model dynamics in conditioning the problem. The analysis suggests what might be an improved formulation of the cost function resulting in better conditioning of the problem.

The relation between the optimization approach and the robust diagnostic method of Sarmiento and Bryan is explicitly demonstrated, and the solution obtained by combining the two methods is used to examine the performance of the GCM used here for the North Atlantic Ocean.

1. Introduction

This is the second part of a work exploring the possibility of posing meaningful and solvable inverse problems using sophisticated general circulation models (GCMs) of the oceanic circulation. For this purpose, the inverse problems are posed as optimization problems, minimizing a cost function that measures the distance between model solution and data and the distance of the model solution from a steady state. In Part I (Tziperman et al. 1992), we used simulated data with and without added noise to carefully analyze this problem under controlled conditions where the true solution to the optimization problem is precisely known. While the solution method performed quite well with no added noise, difficulties arose when simulated noise was added to the model-generated data. When applying the methodology of Part I to real data, the noise level in the data is expected to be quite high, and furthermore, there is no guarantee that the model and data are consistent even within the assumed error bars. We explained in Part I that, when the noise level added to the simulated data is further increased, simulated noisy data do not offer any advantages over real oceanographic data for exploring the performance of the optimization. On the other hand, there is much to be

* Contribution No. 30, Department of Environmental Sciences and Energy Research, The Weizmann Institute of Science.

Corresponding author address: Dr. Eli Tziperman, The Weizmann Institute of Science, Department of Environmental Sciences, 76100 Rehovot, Israel.

learned by using real oceanographic data with the same methodology. In this part of the work, North Atlantic data are therefore used to further analyze the difficulties found with noisy simulated data, and in particular, the problem of conditioning of the inverse problem is discussed and analyzed in detail.

The North Atlantic Ocean was chosen for this study because it is perhaps the most observed and studied of the world's oceans, and therefore serves as an adequate test-bed for new methodologies. Recent numerical simulations of the North Atlantic Ocean have reached a fairly high resolution and level of sophistication. Yet, these models must specify the poorly known surface forcing fields and interior mixing parameters, while they cannot use the available data directly in order to improve estimates of the unknown model inputs. Inverse models have also been used to study the North Atlantic general circulation, using mass, heat, and salt balances in order to calculate an absolute velocity field from hydrographic data. These models were often formulated as coarse-resolution box models (Wunsch 1978, 1984; Schlitzer 1988; Rintoul 1988), or alternatively used local non–mass conserving dynamics (Olbers et al. 1985). More recently, a nonlinear inverse model was used by Mercier (1989) to calculate the general circulation of the western North Atlantic, as well as to obtain an optimal estimate for the density field. While inverse models can, in principle, calculate unknown model parameters from the available data, their resolution is often significantly lower than that of GCMs used to study the North Atlantic Ocean, and their dynamics are simpler.

The optimization approach used in Part I has the potential of combining the advantages offered by the inverse models with the superior resolution and dynamics of the numerical GCMs. To further explore this potential, the methodology and GCM presented in Part I will be used to analyze North Atlantic climatological temperature and salinity data and surface forcing (i.e., heat and freshwater fluxes and wind stress) data. Different inverse problems are posed and the ability of the so-called adjoint method to solve them is examined. As this is the first application of the adjoint method to a complex GCM and real oceanographic data, our purpose here is not to obtain a definite estimate of the North Atlantic circulation, but to demonstrate what can be done using the method and to examine the difficulties that may arise. We did not, therefore, try to use the "best" data for the North Atlantic, but rather chose to use readily available and conveniently gridded climatological datasets for both the hydrography and the surface fluxes. Similarly, whenever possible, we chose simplicity rather than sophistication in the model formulation, for example, in choosing the mixing parameterization. Because this is the first time we are able to attempt the calculation of all model inputs including the hydrography and surface forcing fields, interesting and previously unencountered difficulties arise, and various possible explanations are suggested here for these difficulties. We believe that both the specific results of the calculations presented here and the discussion of the difficulties encountered, as well as the possible extensions to this work, should prove useful and serve as an important step toward a more definite calculation of the North Atlantic general circulation.

This part has three main objectives. The first is to examine the difficulties encountered in Part I with noisy simulated data by using real data that is both noisy and possibly inconsistent with the model used to analyze it. Our second objective is to analyze the model performance by confronting it with real data, using both the optimization approach and the robust diagnostic approach of Sarmiento and Bryan (1982). The relation between the two approaches is, therefore, explicitly demonstrated and discussed. Finally, we analyze in detail the problem of conditioning of the optimization problem involving the fit of data to a steady model. The role of both dissipation and waves in the model dynamics in conditioning the steady optimization problem is demonstrated and discussed. The problem of conditioning seems to be the source of many of the difficulties found here and in Part I, and the careful analysis here should help in future efforts to apply the optimization approach to oceanic general circulation models.

In the following sections, we briefly describe the model and data used for this study (section 2), formulate the optimization problem to be solved (section 3), and describe the results of some North Atlantic inversions (section 4). Then, (section 5) we analyze the conditioning of the optimization using some simplified model equations (section 5), in order to better understand the difficulties demonstrated here in terms of the physics of the model, and to point at possible remedies. Finally, we present our concluding remarks in section 6.

2. Model and data

The GCM was described in detail in Part I. For modeling the North Atlantic Ocean, the horizontal resolution was chosen to be 2° latitude and 3° longitude. The model extends from 9.5°N to 59.5°N, spanning the whole width of the North Atlantic Ocean. There are 12 levels in the vertical so that temperature, salinity, and the horizontal velocities are evaluated at the depths 25, 81.5, 159, 275, 451.5, 714.5, 1088.5, 1589, 2216.5, 2956, 3783 and 4672.5 m.

The temperature and salinity data are taken from the Levitus (1982) atlas, wind data are taken from Hellerman and Rosenstein (1983), the data of Schmitt et al. (1989) are used for evaporation minus precipitation, and the air–sea heat flux data are those of Esbensen and Kushnir (1981). The hydrographic data were calculated on the model grid by subsampling the

original data in the horizontal direction and using simple linear interpolation in depth. The surface forcing data were simply subsampled on the model's grid. Figure 1 shows the wind, air–sea heat fluxes (W m^{-2}), and evaporation minus precipitation (cm yr^{-1}) data used for this study.

3. Optimization problem

The two basic assumptions of this study are that the yearly averaged North Atlantic circulation is at a steady state represented by the annually averaged climatological data, and that this steady state is consistent with our model equations. The steady-state assumption is commonly made in inverse calculations of the oceanic circulation, but it still requires some attention here. The oceanic circulation is characterized by many time scales from seconds–hours for the small-scale turbulence and the internal wave field, through the days–weeks time scale of mesoscale features and the seasonal time scale of the upper ocean, to interannual variability of the large-scale ocean circulation. Our model is of the large-scale oceanic circulation, where the small-scale and mesoscale features are parameterized as subgrid mixing processes, and their temporal variability may therefore be ignored. The very slow interannual variability is taken into account in our model formulation by allowing for errors in the steady-state model equations, as described below. By using annually averaged climatological data, we hope to filter out all seasonally varying components of the circulation. The large-scale oceanic circulation is probably dominated by its steady components at midlatitude regions, which is our justification for using a steady model to describe the circulation. There are, clearly, remaining questions concerning the missing seasonal variability of the upper ocean and whether this variability may be ignored. There are also some parts of the World Ocean where time variability may be an essential part of the dynamics. Such are the western boundary regions, for example, where one wonders whether there is a consistent steady solution for the narrow Gulf Stream or whether the meandering of the current is essential to its existence as a narrow swift current. Our approach in this work is rather practical; we simply try using a steady model, and if inconsistencies arise, an attempt is made to determine whether they result from the ignored time-varying components of the circulation. As will be seen below, there are quite a few interesting and important problems that can and need to be addressed—even in the simplified framework of a steady model.

The purpose of this study is to calculate the model inputs that are consistent with both the available data and the model equations. These inputs include the temperature (T) and salinity (S) fields, the surface forcing by wind stress ($\tau = [\tau^x, \tau^y]$) and by air–sea fluxes of heat ($\mathscr{H}$) and fresh water ($E - P$), and the horizontal and background vertical tracer mixing coefficients (K_h and K_{v0}, see Part I for details of the vertical-mixing parameterization). The available climatological data and model equations are therefore combined into a single cost function that is a function of all model inputs and that measures both the deviation of the model inputs from the data and the deviation from the steady-state solution of our model equations. Let us now denote the data for the various inputs by a hat ($\hat{}$), and the temperature and salinity fields calculated by the optimization and used as the initial conditions for the model run by $T_{ijk}^{n=0}$ and $S_{ijk}^{n=0}$. The cost function to be minimized in the optimization can then be written as

$$
\begin{aligned}
J(T, S, \mathscr{H}, E-P, \tau, K_{v0}, K_h) &= \sum_{i,j,k} \left[\bar{W}_{ijk}^{T} \left(\frac{\partial T_{ijk}}{\partial t} \right)^2 + \bar{W}_{ijk}^{2} \left(\frac{\partial S_{ijk}}{\partial t} \right)^2 \right] \\
&+ \sum_{i,j,k} [W_{ijk}^{T} (\hat{T}_{ijk} - T_{ijk}^{n=0})^2 + W_{ijk}^{S} (\hat{S}_{ijk} - S_{ijk}^{n=0})^2] \\
&+ \sum_{i,j} [W_{i,j}^{\tau} ((\hat{\tau}_{ij}^{x} - \tau_{ij}^{x})^2 + (\hat{\tau}_{ij}^{y} - \tau_{ij}^{y})^2)] \\
&+ \sum_{i,j} [W_{i,j}^{\mathscr{H}} (\hat{\mathscr{H}}_{ij} - \mathscr{H}_{ij})^2 + W_{i,j}^{E-P} ([\hat{E} - \hat{P}]_{ij} \\
&\qquad - [E - P]_{ij})^2]. \quad (1)
\end{aligned}
$$

The temporal derivatives appearing in (1) are calculated by stepping the model a single time step from the initial conditions, $T_{ijk}^{n=0}$ to $T_{ijk}^{n=1}$, and letting $\partial T_{ijk}/\partial t = (T_{ijk}^{n=1} - T_{ijk}^{n=0})/\Delta t$. By minimizing the cost function (1), we obtain an optimal estimate for the model inputs that is as close to the observations as allowed by the noise in the data, and is also as consistent as possible with a steady-state solution of our model equations. The procedure used to minimize the cost function uses the adjoint model of our GCM and the conjugate-gradient (C-G) minimization algorithm and was described in Part I. Note that in addition to the optimal estimates for all model inputs, a 3D velocity field is calculated that is consistent with the model equations as well as with the various types of available data.

a. Weights

The weights (W_{ijk}^{T}, etc.) in the cost function (1) determine the relative importance of the various constraints used in the cost function. They play a major role in determining the results of the calculation and should therefore be chosen carefully. A more complete statement of the cost function should involve weighting by the inverse error covariance matrix for the observations (Thacker 1989), but the large number of terms in the cost function makes this formulation impractical for our purposes. We therefore resort to using the above form of weighting that is meant to be equivalent to weighting the various terms of the cost function by the diagonal terms of the inverse covariance matrix. In principle, when using diagonal weighting, each term in the cost function is weighted by the inverse of its

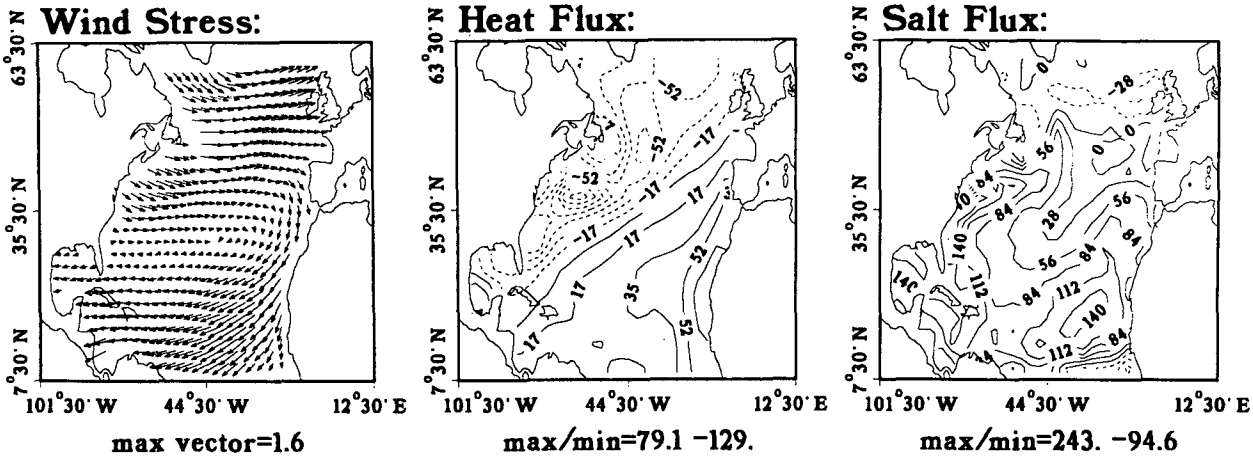

FIG. 1. Surface-forcing data used for this study: wind stress (Hellerman and Rosenstein 1983), air–sea heat flux (Esbensen and Kushnir 1981), and evaporation minus precipitation (Schmitt et al. 1989).

squared error, but there are some further considerations. Beginning with the weights for the hydrographic data, note that each temperature data point T_{ijk}, for example, represents in the model a volume element of the appropriate model grid box, $dx_i\,dy_j dz_k$. Data representing larger volumes should be given more weight in the cost function, and vice versa. Otherwise, arbitrarily choosing the model grid to be denser at some depth or location will result in an artificially high weight given to the temperature data at that location. The errors in the temperature and salinity data should also be considered, of course, when calculating the weights. Because there is no error information for the Levitus hydrographic data used here, we use rough order-of-magnitude error estimates, $\epsilon_k(T)$ and $\epsilon_k(S)$, specified as a function of depth only, and decreasing exponentially with depth as plotted in Fig. 2. These error profiles are used to calculate the weights for the hydrographic data that are proportional to the grid-box volume element over the error squared, for example,

$$W^{T}_{ijk} \propto \frac{dx_i\,dy_j dz_k}{[\epsilon_k(T)]^2}.$$

Similarly, the weights for the surface forcing data are chosen to be proportional to the area element at the grid location where the surface data is given, divided by the squared error in the surface flux data,

$$W^{\mathscr{F}}_{ij} \propto \frac{dx_i\,dy_j}{[\epsilon(\mathscr{F})]^2}.$$

The weighting by the volume and area elements takes care of the relative importance of different temperature data among themselves or different surface forcing data among themselves. It is still necessary to determine the relative importance of, say, a single surface heat-flux data point versus a single temperature data point (or equivalently, the relative importance of the whole hydrographic data versus the whole surface heat-flux data).

In order to calculate the relative weights of different types of data in the cost function, we need to specify a measure of the intrinsic amount of information in each data point. For this purpose, consider the correlation scales for the various data types. Let the horizontal and vertical correlation distances of the hydro-

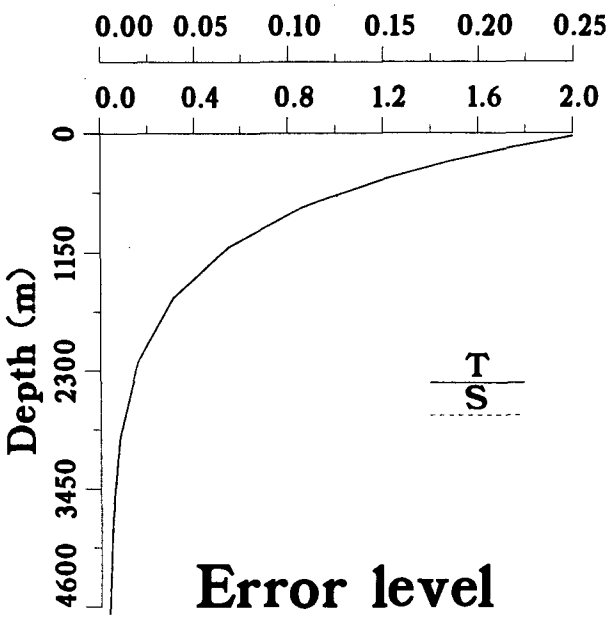

FIG. 2. The specified-error profiles for the temperature and salinity data [$\epsilon_k(T)$ and $\epsilon_k(S)$], used to calculate the weights (2, 3) for the hydrographic data terms in the cost function (1). (The expressions for the two profiles are $\epsilon_k(T) = 1.97 \times \exp[-(z_k - z_1)/800.] + 0.03$, and $\epsilon_k(S) = 0.246 \times \exp[-(z_k - z_1)/800.] + 0.004$. Note that the two profiles overlap in the figure.)

graphic data be L_H^{ocean} and L_V^{ocean}, respectively. A single temperature measurement roughly represents an ocean volume of $V_{\text{corr}} = [L_H^{\text{ocean}}]^2 \times L_V^{\text{ocean}}$. Similarly, given that the correlation distance for the atmospheric surface-forcing data (wind, air–sea fluxes) is L^{Atmos}, each surface forcing data point represents an area of $A_{\text{corr}} = [L^{\text{atmos}}]^2$. This provides a way of prescribing the relative importance of different data types. Normalizing the weights for the hydrographic data by the correlation volume for this data, and the surface forcing weights by the corresponding correlation area, we obtain a weight that measures the intrinsic value of each data point relative both to other data of its kind, and to data of other types,

$$W^{T}_{ijk} = \frac{dx_i\, dy_j dz_k}{V_{\text{corr}} \times [\epsilon_k(T)]^2}, \quad W^{S}_{ijk} = \frac{dx_i\, dy_j dz_k}{V_{\text{corr}} \times [\epsilon_k(S)]^2},$$

$$W^{\mathscr{H}}_{ij} = \frac{dx_i\, dy_j}{A_{\text{corr}} \times [\epsilon(\mathscr{H})]^2},$$

$$W^{E-P}_{ij} = \frac{dx_i\, dy_j}{A_{\text{corr}} \times [\epsilon(E-P)]^2},$$

$$W^{\tau}_{ij} = \frac{dx_i\, dy_j}{A_{\text{corr}} \times [\epsilon(\tau)]^2}. \qquad (2)$$

Estimates for the error in climatological wind and air–sea flux data such as used here vary, with a conservative estimate being about 20% (Schmitt et al. 1989). After some experimentation with the weights, we have specified a uniform error calculated as 15% of the average absolute value of the data for each of the surface forcing fields. These are the values used for $\epsilon(\mathscr{H}, E-P, \tau)$ in (2). The oceanic correlation distances used in (2) were specified to be $L_H^{\text{ocean}} = 400$ km, $L_V^{\text{ocean}} = 400$ m, and the correlation distance for the atmospheric data was taken as $L^{\text{atmos}} = 1500$ km.

The choice of weights for the terms in the cost function requiring the minimization of the temporal derivatives ($\bar{W}^{T}_{ijk}$) poses yet another difficulty. We want the weights for these steady penalty terms to be proportional, as before, to the inverse squared error in the corresponding terms of the cost function. But it is not obvious how the errors in these terms should be calculated. The deviation of the inverse solution from a steady state comes from two independent sources. The first is the long-term temporal variability of the oceanic general circulation, such as the large-scale changes in the temperature and salinity of the North Atlantic Ocean between the two 5-year periods, 1955–59 and 1970–74, described by Levitus (1989). The second source of errors in the steady-penalty terms is the discrepancy, however small, between the model equations and the actual physics of the oceanic circulation. Both the model errors and the actual deviation of the ocean from a steady state would result in an inability of the inverse calculation to fit a perfectly steady model to the climatological data, and the temporal derivatives appearing in the cost function would therefore not vanish.

A set of weights for the steady penalties to be used in the optimizations presented here, believed to be a reasonable estimate for these weights, is defined as follows. Assume that the errors $\epsilon_k(T, S)$ specified for the temperature and salinity data are also an estimate for the variability of the ocean during a 15-year time period. This is, very roughly, the order of magnitude of the temporal variation of the North Atlantic temperature and salinity fields found by Levitus (1989) over a time period of 15 years. The error in the steady-penalty terms $(\partial T/\partial t)^2$ is therefore $[\epsilon_k(T)/(15\text{ years})]^2$. The weights for these penalty terms are the inverse squared of these errors, weighted by the appropriate grid-box volume element, and correlation volumes,

$$\bar{W}^{T}_{ijk} = \frac{dx_i\, dy_j dz_k}{V_{\text{corr}} \times [\epsilon_k(T)/(15\text{ years})]^2},$$

$$\bar{W}^{S}_{ijk} = \frac{dx_i\, dy_j dz_k}{V_{\text{corr}} \times [\epsilon_k(S)/(15\text{ years})]^2}. \qquad (3)$$

b. Open boundaries

The portion of the North Atlantic Ocean modeled here is bounded by several open boundaries (Fig. 1): the southern boundary at 9.5°N, the northern boundary at 59.5°N, and the Straits of Gibraltar. The regions near these open boundaries require some special attention. Ideally, the open boundary conditions (in particular, velocities at the open boundary), should be treated as unknowns, and solved for as part of the optimization. But such a formulation is not trivial to implement, and we have chosen a different, simpler approach that may still be expected to give good results.

At the open boundaries we specify a vertical wall at which the regular set of closed boundary conditions is applied (no-slip condition for the velocities, and no flux for the temperature and salinity). In the region near this artificial wall the model dynamics is not valid, and the model is not expected to give realistic results for the velocity field as well as for other model outputs. This region is therefore treated as a limiting case of the interior in which we impose the data constraints appearing in the cost function (1) but not the dynamic constraints requiring the model to be at a steady state based on the model equations. Consequently, the weights of the dynamic constraints, $\bar{W}^{T,S}_{ijk}$, are set to zero in the regions near the open boundaries. These regions were chosen to be the three grid points nearest the northern and southern boundaries, and a small region of four latitude grid points by two longitude grid points near the Strait of Gibraltar.

This approach to handling open boundaries is similar to the use of "buffer" or "sponge" regions in numerical models: a closed boundary condition is specified, and the model equations are modified in the region near

the boundary to include a term restoring the temperature and salinity to climatological data for temperature and salinity. This amounts, as in our formulation, to using the data near the open boundaries while not enforcing the model physics there.

4. Results of the North Atlantic data analysis

The specific inverse problems examined using the model and data described above will now be described. In Part I, we concluded the set of inverse calculations using simulated data at a point where it was clear that, while the optimization was fairly successful for noise-free simulated data, there were difficulties with the convergence to the optimal solution for various model inputs when noise was added to the simulated observations. Two possible explanations were suggested for these difficulties. The first was the problem of ill conditioning of the optimization, resulting in possible stalling of the optimization and in noise amplification (i.e., low-level noise in the data may result in a very noisy solution for the model inputs). The second difficulty was the possible existence of local minima of the cost function to which the optimization may converge rather than to the global minimum representing the desired optimal solution. In Part I, we anticipated the problems to be even more severe when using real data that is noisier and less consistent with the model than the simulated data.

In the experiments shown below, the difficulty due to ill conditioning of the optimization (section 4a) is explicitly demonstrated. A case of multiple minima of the cost function (section 4b) is then presented. Finally, the relation between the optimization approach and the robust diagnostic approach of Sarmiento and Bryan (1982) discussed in Part I is explicitly demonstrated (section 4c). A robust diagnostic calculation is used to obtain an initial guess for the model inputs, and this guess is used as the starting point for the optimization. The calculation involving the robust diagnostic approach also serves to test the GCM used here with the real North Atlantic Ocean basin and data. In section 5, the conditioning problem will be analyzed using some simplified model equations, in order to better understand the difficulties demonstrated here, and to point out some possible remedies.

a. Difficulties due to ill conditioning

The conditioning of the optimization is closely related to the problem of uniqueness of the optimization solution for the various model inputs. If the cost function does not change in a direction in parameter space that corresponds to some model inputs, then the optimization cannot prefer any of the possible solutions along this direction, and the corresponding parameters are not resolved. This flatness of the cost surface in some directions may be the result of the lack of sufficient data for calculating the model parameters as was found, for example, by Tziperman and Thacker (1989). They found that the solution for their simple QG model inputs may not be unique under certain circumstances. For their simple model it was possible to show that the nonuniqueness indeed resulted from the ill posedness of the problem: there were more unknowns to be solved for than there were constraints. Such a situation may be anticipated by examining the resolution information calculated from the Hessian matrix (Tziperman and Thacker 1989), but that is much too large to be calculated explicitly for the present problem. In order to examine the conditioning of the optimization problem for our problem, we use the fact that ill conditioning results in the cost surface being much flatter in some directions in parameter space than in others (Thacker 1989). To explicitly demonstrate such a case, we therefore simply investigate the structure of the cost surface in the neighborhood of two solutions obtained by the optimization.

To obtain the first solution (run I1 in Table 1), all model inputs were calculated simultaneously. The calculated inputs included the temperature and salinity fields, the surface forcing by wind stress and by heat and freshwater fluxes, and the two mixing coefficients. The initial guess for all model inputs was taken to be the data for the hydrography and surface forcings: 1 $cm^2\ s^{-1}$ for the background vertical-mixing coefficient

TABLE 1. Indications for the ill conditioning of the optimization: two different solutions that cannot be distinguished by the optimization. The two runs are described in section 4a. The J_{beg} is the initial value of the cost function, before the optimization and based on the data values as first guess for all fields; J_{end} is the cost function at the end of the optimization. The columns marked $\Delta(\ )$ give the contribution of various terms in the cost function (5). For example, for the temperature field, the term in the cost function measuring the deviation from the data is

$$\Delta T = \sum_{i,j,k} [W^T_{ijk}(\hat{T}_{ijk} - T_{ijk})^2].$$

Note that the contributions of the steady penalties (column marked $\Delta T_t + \Delta S_t$) are by far larger than any of the data misfit terms, indicating that these solutions may not represent the desired balance between deviation from a steady-state solution and deviation from the data.

Run	J_{beg}	J_{end}	ΔT	ΔS	$\Delta T_t + \Delta S_t$	$\Delta\tau^x$	$\Delta\tau^y$	$\Delta\mathscr{H}$	$\Delta E - P$	K_h, K_{v0}	Figures
I1	2079.1	40.96	4.78e− 2	7.28e− 2	31.8	3.3	3.5	0.73	1.5	.242E+6, .1379	3a,b
I2	2079.1	33.17	0.12	0.14	32.2	8.66e− 2	9.92e− 2	0.17	0.31	.369E+6, .332	3a,b

and 10^7 cm^2 s^{-1} for the horizontal-mixing coefficient. The optimization was continued for more than 4000 conjugate-gradient iterations until no more progress toward a lower value of the cost function was observed. The second solution (run I2 in Table 1) was obtained by calculating all of the above model inputs, but in a slightly different way: first the values of the surface-forcing parameters were fixed to the data values, and only the optimal hydrography minimizing the cost function was calculated. After a significant reduction of the cost function was achieved (but minimum still not reached), we proceeded to calculate all model inputs simultaneously, restarting the optimization from the hydrography calculated during the first half of this optimization, and from the data values for the surface forcing fields. Again, a total of about 4000 iterations was needed to reach the final solution.

The results of the two optimizations are summarized in Table 1 and in Figs. 3a and 3b. Although both solutions seemed to represent minimum points of the cost function, and although the total reduction of the cost function is roughly the same for both runs, they correspond to very different locations in parameter space. In the first run (I1), where all parameters were calculated simultaneously from the beginning of the run, the final surface forcings deviate significantly from the data, while the temperature and salinity are hardly modified. In the second run (I2), where before calculating all parameters simultaneously, only the temperature and salinity were calculated, the final solution is characterized by a smaller modification of the surface forcings and a larger modification of the temperature and salinity fields.

At this stage, it is not yet clear whether the non-uniqueness of the optimization solution is due to the stalling of the optimization due to ill conditioning, or whether two separate local minima have been found. In order to try to resolve this issue, we calculated the cost function along a section across parameter space, passing through the points I1 and I2 of Table 1. The curve marked "section A" in Fig. 3b shows the cost function calculated on 60 equally spaced locations along the line passing through the two minima,

$$\mathbf{X}_i = \mathbf{X}_{\min 1} + \text{step}_i * (\mathbf{X}_{\min 2} - \mathbf{X}_{\min 1}); \quad \text{step}_i = -3, \cdots, 3, \quad (4)$$

where $\mathbf{X}$ is a vector of all model inputs calculated by the optimization and $\mathbf{X}_{\min 1}$, $\mathbf{X}_{\min 2}$ correspond to the two solutions (I1, I2). Note that the two solutions are

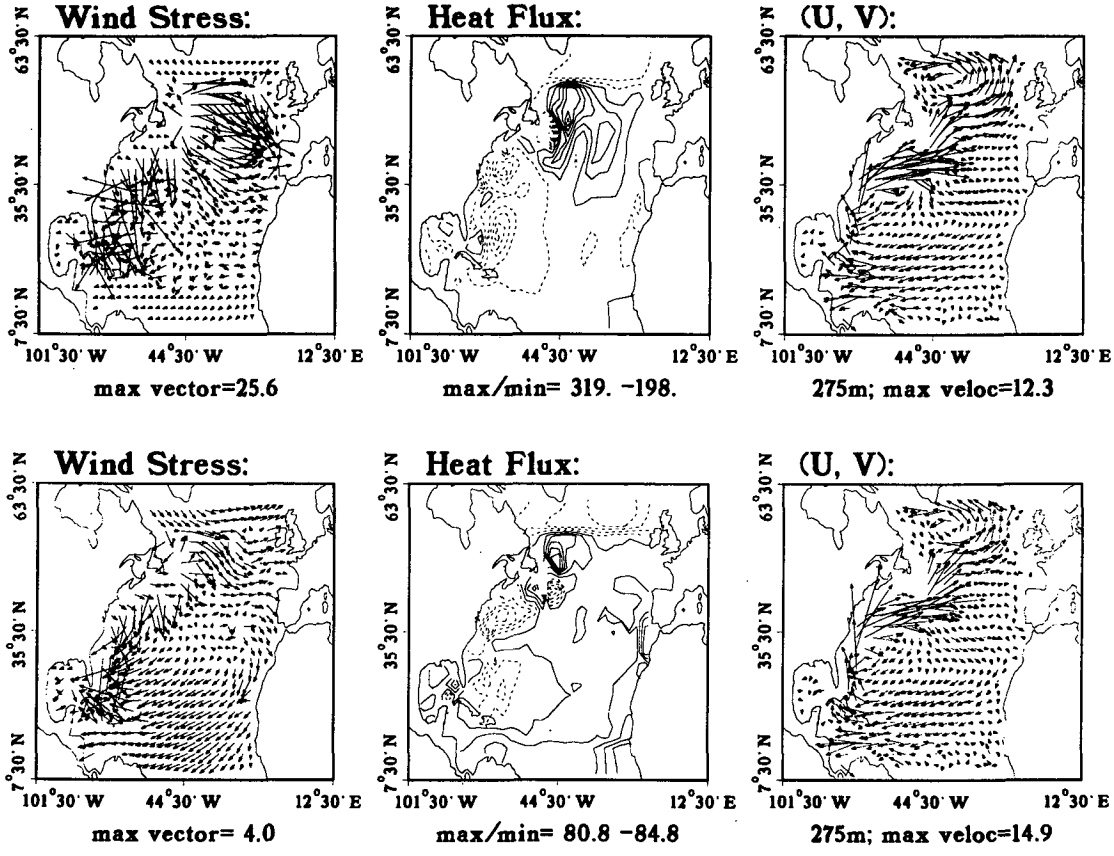

FIG. 3a. The upper and lower panels present the results of runs I1 and I2, respectively (see Table 1 and subsection 4a). From left to right: the surface wind stress, the air–sea heat flux, and the horizontal circulation at a depth of 275 m.

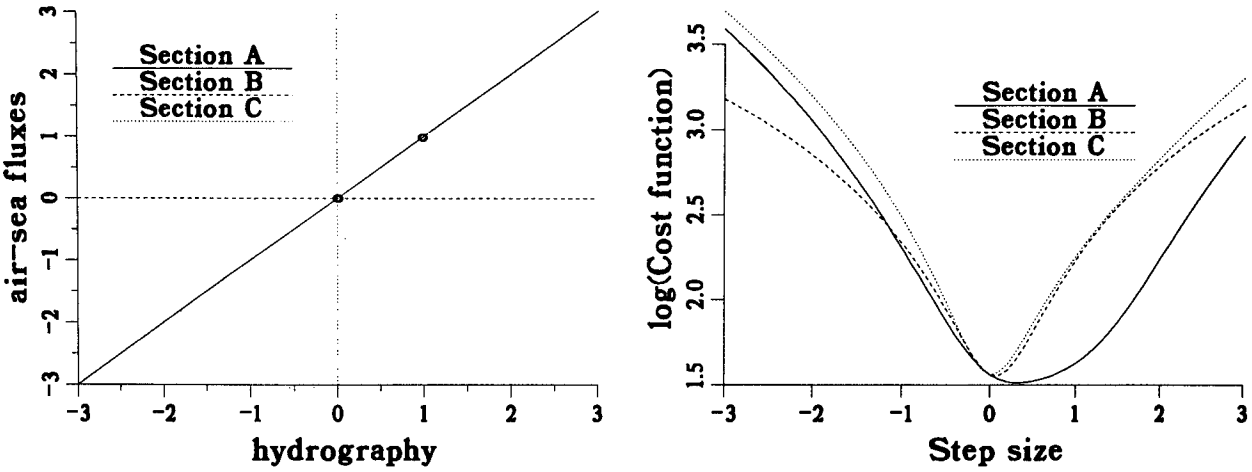

FIG. 3b. Three sections of the cost function across parameter space. Right panel: value of the cost function along the three sections. Left panel: schematic drawing of the directions of the three sections. The two solutions I1 and I2 (Table 1) are schematically denoted by two circles in this figure, at locations (0, 0) and (1, 1). See section 4a for details.

recovered for $step_i$ equal to 0 and 1. Looking at the cost function along section A, it becomes obvious that the two solutions of Table 1 are not two local minima, but both lie in a "valley" in parameter space along which the cost function hardly changes. Such a valley means that the model is unable to resolve linear combinations of parameters lying along this line, and all solutions along it are acceptable by the optimization. Note that section A in Fig. 3b also extends beyond the two solutions of Table 1 in both directions and that there the cost function eventually increases so that the valley in parameter space is of a limited extent. Carefully examining the value of the cost function along this section, we find that the minimum value of the cost function along this section does not lie at any of the solutions found by the optimization, but somewhere in the middle. Clearly, the flatness of the cost surface in this direction had caused the optimization to stall and prevented convergence to the actual minimum.

In order to demonstrate that the cost function is not necessarily flat in other directions in parameter space, the cost function was calculated along two additional

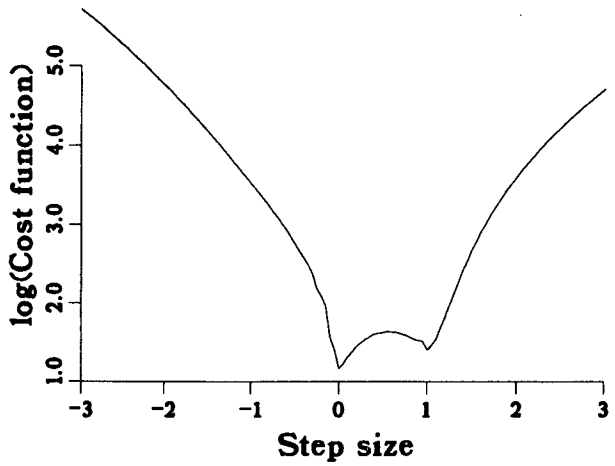

FIG. 3c. A section in parameter space indicating the possible existence of two local minimum points (solutions M1 and M2; see Table 2; see section 4b).

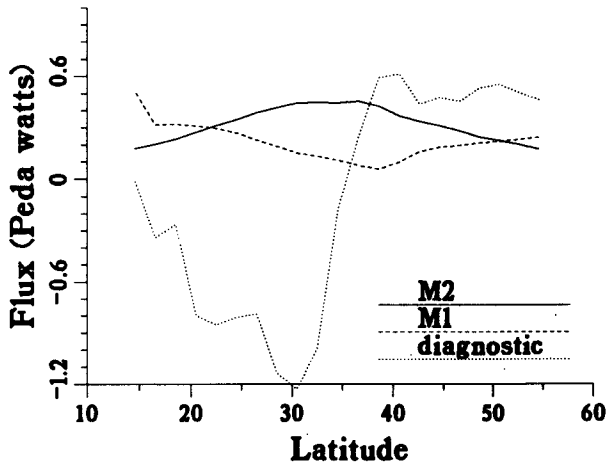

FIG. 3d. The meridional heat flux for the two solutions corresponding to the two minimum points shown in (c) (runs M1 and M2 in Table 2). Note that, although the two minimum points seem quite close to each other in terms of the reduction of the cost function relative to the initial guess (see Table 2 and section 4b), the meridional heat flux is quite different for the two solutions. Also shown: the meridional heat flux calculated diagnostically from the Levitus data, with no optimization involved.

sections in parameter space (marked B and C in Fig. 3b). Both of these sections pass through one of the solutions of Table 1, but in different directions than that of section A. The left panel in Fig. 3b schematically shows the directions of the three sections presented in the right panel. Let us divide the tens of thousands of parameters calculated by the optimization into two groups: one includes the hydrography, and the other the surface forcing by air–sea fluxes and wind. These two groups are denoted by the two axes of the left panel of Fig. 3b. The two solutions of Table 1 are schematically denoted by two circles in this figure, at locations (0, 0) and (1, 1). As can be seen in this figure, the section marked A passes through both of these solutions. The curve marked section B was calculated along the projection of section A on the "hydrography" axis. This means that at $\text{step}_i = 0, 1$ the hydrography part of the parameter space is correspondingly equal to that of the two minimum points of Table 1, while the air–sea fluxes part of the solution is equal to that of the first minimum point of Table 1 along the entire section B. The curve marked section C is calculated along the projection of section A along the air–sea fluxes axis. The curves of the cost function along the two additional sections passing through the first minimum point, but in different directions, clearly demonstrate that the cost surface is not flat in all directions as it is in the direction of section A, so that this section indeed represents a special direction in parameter space along which the optimization is unable to resolve the model parameters.

The flatness of the cost surface between the two minima may be explained as follows. The optimization can minimize the deviations from a steady state at each surface grid point either by changing the temperature and salinity fields near the surface, or by modifying the surface fluxes so that they balance any deviation from steady state in the temperature and salinity equations. The data constraints and the choice for the different weights should determine the precise compromise between modifying the surface properties (temperature and salinity) and modifying the surface forcing fields. Considering the relative magnitudes of the different terms of the cost function (Table 1), note that for the solution with $J = 33$, all data penalties (temperature, salinity, wind stress, and freshwater fluxes) together amount to less than one. This implies, of course, that the steady penalties are O(32), and the imbalance between the data and steady penalties seems to indicate that the solution does not represent a balance of dynamics (steady penalties) and data, as is intended in the optimization approach taken here. The optimization did not, consequently, use the data in order to separately resolve the hydrography and fluxes, resulting in the observed flatness of the cost function between the two solutions. Furthermore, the ill conditioning expressed in the cost function being flat between the two solutions had prevented further progress of the optimization. Note also the large number of iterations needed for the conjugate-gradient optimization to converge to the minimum of the cost function (about 4000 iterations). Ill conditioning of the optimization problem and the resulting flat structure of the cost function in some directions is known to cause a very slow convergence to the minimum point of the cost function (Gill et al. 1981; Tziperman and Thacker 1989). It will be seen that there are, in fact, solutions with lower value for the cost function; so the solutions found here (in fact, the whole valley found in parameter space) do not represent the optimal solution.

Because the solutions found here are not the optimal ones corresponding to the global minimum of the cost function, the results of these experiments cannot be used to deduce whether the data is sufficient to determine all surface forcings. We still have chosen to present these results because they demonstrate an interesting and potentially important situation where a flat structure of the cost function in some direction results in ill conditioning of the optimization.

b. Local minima

In order to examine the possible existence of local minima of the cost function, several optimizations were made in which the only control variables calculated are the temperature, salinity, and mixing coefficients. The initial guess for all model inputs was taken to be the data for the hydrography and surface forcings: 1 $\text{cm}^2\,\text{s}^{-1}$ for the background vertical-mixing coefficient, and $10^7\ \text{cm}^2\,\text{s}^{-1}$ for the horizontal-mixing coefficient. The initial cost value before the optimization was $J = 2079$. We begin by minimizing the cost function using the C-G algorithm, until converging to a minimum point (run M1 in Table 2). The minimum value obtained is $J = 26$. The reduction in the value of the gradient from the initial guess for the parameters to the minimum point is of about five orders of magnitude. At this stage, the optimization seemed unable to further reduce the cost, so we *assume* the solution represents a minimum of the cost function. Starting from the optimization solution, the model is then stepped forward in time for 2 years, resulting in the location in parameter space characterized in the entry MF in Table 2. The forward run results in an increase of the cost function to $J = 43$, as time stepping the model does not necessarily minimize the cost function. Finally, restarting the C-G optimization from the temperature and salinity at the end of the forward run, the optimization converges to a second minimum, $J = 16$, given by the entry M2 in Table 2.

As before, we want to determine whether the nonuniqueness of the optimization solution is due to insufficient data or due to the nonlinearity of the cost function resulting in two separate local minima calculated by the optimization. Figure 3c shows the cost function along the line (section) in parameter space connecting the two solutions M1 and M2 shown in

TABLE 2. Calculating the optimal hydrography and mixing coefficients. The first three entries show indication for the existence of multiple local minima in parameter space when calculating only hydrography (T, S) and mixing coefficients (see the text in section 4b for explanation). The final two entries are the result of the robust diagnostic calculation described in section 4c (entry robust), and the optimization that follows it (Z). The J_{end} is the cost function at the end of the specific calculation, whether an optimization (entries M1, M2, and Z), a forward model run (MF), or a robust diagnostic calculation (robust); J_{beg} is the value of the cost function at the beginning of the optimization. See Table 1 for explanation of the entries in this table, where in particular the columns marked ΔT_t and ΔS_t are the contributions to the cost function resulting from the steady penalties for temperature and salinity.

Run	J_{beg}	J_{end}	ΔT	ΔS	ΔT_t	ΔS_t	K_h, K_{v0}	Comments	Figures
M1	2079.1	25.9	0.40	0.32	18.0	6.9	0.38×10^7, 0.3	1st minimum	3c,d
MF	25.9	42.8	0.79	0.59	22.1	20.6	0.38×10^7, 0.3	Forward run	3c,d
M2	42.8	15.8	0.77	0.65	7.0	7.4	0.38×10^7, 0.3	2d minimum	3c,d
Robust	2079.1	9.47	1.9	2.0	2.6	2.9	1.00×10^7, 1.0	Robust diagnostics	7
Z	9.47	7.84	1.9	2.0	1.7	2.2	1.00×10^7, 1.0	Optimization	4–8

Table 2. The structure of cost surface along this section suggests that the two solutions are two local minima separated by a small maximum where the cost is equal to 43.

The section through the parameter space shown in Fig. 3c clearly shows that the two minima are distinct along the straight line connecting them. They are not part of a valley of the cost surface, directly linking the two solutions in parameter space, as is expected when data is insufficient to resolve the parameters. It is still possible, of course, that the two minima are connected but not along this straight line; we have no way of verifying this possibility due to the large number of parameters and therefore the large dimension of the parameter space. The fact that solution M2 represents a lower value of the cost function seems a mere coincidence. Because of the ill conditioning (or the existence of local minima), the final solution in these runs depends on the initial guess for the unknowns. It seems that the short forward run preceding calculation M2 resulted in an initial guess that enabled finding a solution with a lower cost function. But none of these solutions is optimal, as will be seen later.

Examining Table 2, however, it is striking how similar the amount of cost reduction is for the two solutions. In one case (M1) the reduction from the initial guess using data values for the temperature and salinity is $2079 - 26 = 2053$, while in the second (M2) it is $2079 - 16 = 2063$. One is immediately tempted to ask whether the difference between the two solutions is significant at all or is within the noise level. To answer this question, Fig. 3d shows the total (advective and diffusive) zonally integrated meridional heat flux carried by the ocean as calculated for the two minimum points, as well as for the initial guess based on the Levitus data values (the results based on the Levitus data are discussed in more detail below). It is quite obvious (and somewhat unfortunate) that the seemingly small difference between the two minimum points translates into a significant difference between the corresponding optimization solutions. Clearly, one should invest efforts in getting as close as possible to the absolute minimum of the cost function to obtain the truly optimal solution. But does one of these solutions represent the global minimum and therefore the desired optimal solution for the model inputs? Entries M1 and M2 in Table 2 indicate that there is still an imbalance of steady and data penalties, suggesting that these solutions still do not represent the desired balance of steadiness and deviation from the data. To find if there is in fact a better solution, a solution with a lower cost function is obtained in section 4c by starting the optimization with a robust diagnostic calculation (Sarmiento and Bryan 1982). This provides an opportunity to demonstrate the link between the optimization and the robust diagnostic method, as well as to discuss the relevance of our model to the North Atlantic Ocean.

c. *The relation between the robust diagnostic and optimization approaches*

Our first objective now is to try to find out if a solution exists that corresponds to a lower value of the cost function than the ones found in the previous subsections and that represents a balance of steady penalties and data penalties in the cost function. For this purpose we restrict itself the calculation of optimal estimates to the temperature and salinity fields and to the mixing coefficients, and set all other model inputs (wind stress, heat and freshwater fluxes) to the data values—not treating them as unknowns.

In Part I, it was explained that in the limit where only the hydrography (T, S) is calculated by the optimization, the calculation is roughly equivalent to a robust diagnostic calculation. It seems natural, therefore, to begin the optimization with a robust diagnostic run of the GCM used for the optimization. The robust diagnostic method of Sarmiento and Bryan (1982) is based on adding a term of the form $\gamma(\hat{T} - T)$ to the temperature equation, restoring the temperature to the data values (and similarly for the salinity). The restoring coefficient, γ, has units of $(\text{time})^{-1}$ and was chosen by Sarmiento and Bryan to be a function of depth, representing a time scale of between 50 days for shallow levels and 250 days for the deep water.

A reasonable first guess for the value of γ that results

in a significant reduction of the cost function can be calculated from the weights of the different terms in the cost function using the following heuristic argument. Consider the contribution to the cost function due to the data and steady penalties for the temperature, written as follows,

$$J(T) = \sum_{i,j,k} [\bar{W}^T_{ijk}(\mathbf{u}\nabla T - K_h\nabla_H^2 T - K_v T_{zz})^2 + W^T_{ijk}(\hat{T} - T)^2]. \quad (5)$$

At the optimal solution we expect the total contribution of the steady penalties to be of the same order as that of the data penalties. Translating this global requirement into a local condition (admittedly with no rigorous justification, but hopefully resulting in a better understanding of the relation between robust diagnostic calculations and the optimization approach), we have

$$(\mathbf{u}\nabla T - K_h\nabla_H^2 T - K_v T_{zz})^2 \approx \left[\frac{W^T_{ijk}}{\bar{W}^T_{ijk}}\right](\hat{T} - T)^2. \quad (6)$$

Now, at steady state, the robust diagnostic equation for the temperature satisfies

$$(\mathbf{u}\nabla T - K_h\nabla_H^2 T - K_v T_{zz}) = \gamma(\hat{T} - T), \quad (7)$$

which clearly suggests a relation between the weights used in the cost function, and the robust diagnostic coefficient γ that should result in a significant reduction of the cost function,

$$\gamma = \left[\frac{W^T_{ijk}}{\bar{W}^T_{ijk}}\right]^{1/2}. \quad (8)$$

Although in general, γ may be a function of location, for our choice of weights, this relation simply suggests a constant value of $\gamma = 1/(15 \text{ years})$. We have experimented with various choices of γ around this suggested value of $\gamma = 1/(15 \text{ years})$, and finally settled on a constant value of $(9.6 \text{ years})^{-1}$, which seemed to result in the most significant reduction to the cost function. Note that this is a much smaller γ than used by Sarmiento and Bryan. The larger the time scale used to define γ, the less weight put on the data, and the closer the final solution to the steady state of the model equations without the robust diagnostic term. The time needed to achieve a steady solution is also related to the choice of γ. Normally a time that is 3 to 5 times γ^{-1} suffices, which implies about 50 years of integration. We have integrated the model for more than 500 years to ensure achieving a steady solution.

The results of the robust diagnostic run are summarized under the entry "robust" in Table 2. The value of the cost after the robust diagnostic run was 9.5, which is a significantly lower value than we were able to obtain with the optimizations in the previous section. It is interesting to note that the optimization is given more degrees of freedom because the fluxes, or at least the mixing coefficients, are allowed to vary in addition to the hydrography. As a result, the optimal value of the cost function for all of these cases is expected to be lower than that of the robust diagnostic run, where only the temperature and salinity are allowed to vary from the data values. The failure of the optimization to find such solutions is another indication of the poor conditioning of the problem; this is analyzed in detail in the next section.

The robust diagnostic solution was taken as the starting point for an optimization that further reduced the cost from 9.5 to 7.8. Additional details are given under entry Z in Table 2. Unlike the minimum points found in the optimizations described previously, in which the steady-penalty terms dominated the cost, the minimum found by first using the robust diagnostic method represents a balance of steady and data penalties that are of the same order of magnitude (e.g., compare ΔT and ΔT_t in Table 2). This hopefully indicates that the solution is indeed balancing data and model dynamics, as intended in the formulation of the cost function. Experience in the previous sections with the complicated structure of the cost surface and the poor conditioning of the problem, however, suggests that there is still no guarantee that the minimum found here is indeed the desired global minimum of the cost function.

Trying to calculate surface fluxes in addition to hydrography by restarting the optimization from the solution of run Z, it was found that the fluxes were hardly modified by the optimization and the cost not reduced much. This may indicate one of two possibilities: One is that the solution found in run Z, which is a minimum with respect to the hydrography that was varied in the optimization, is also very close to a local minimum point with respect to the other model parameters (surface fluxes) although these were kept constant during the optimization. The other possibility, of course, is that the optimization is stalling due to bad conditioning near the solution of the robust diagnostic run, and therefore is unable to use the additional degrees of freedom (surface fluxes) to further reduce the cost.

The solution of the optimization that was started from the robust diagnostic solution seems reasonable, and it is worthwhile to examine it now in some detail. The detailed examination of the solution and residual fields is used to evaluate the performance of the model in the realistic setting of the North Atlantic Ocean. Therefore, various aspects of the solution are now discussed, beginning with the flow field and then the meridional circulation, meridional heat flux, the mixing coefficients, and the residual fields.

1) THE GENERAL CIRCULATION

The velocity field calculated by the optimization is given in Fig. 4 for every second model level. The circulation is composed of the classical subtropical and subpolar gyres in the upper levels, while the deeper

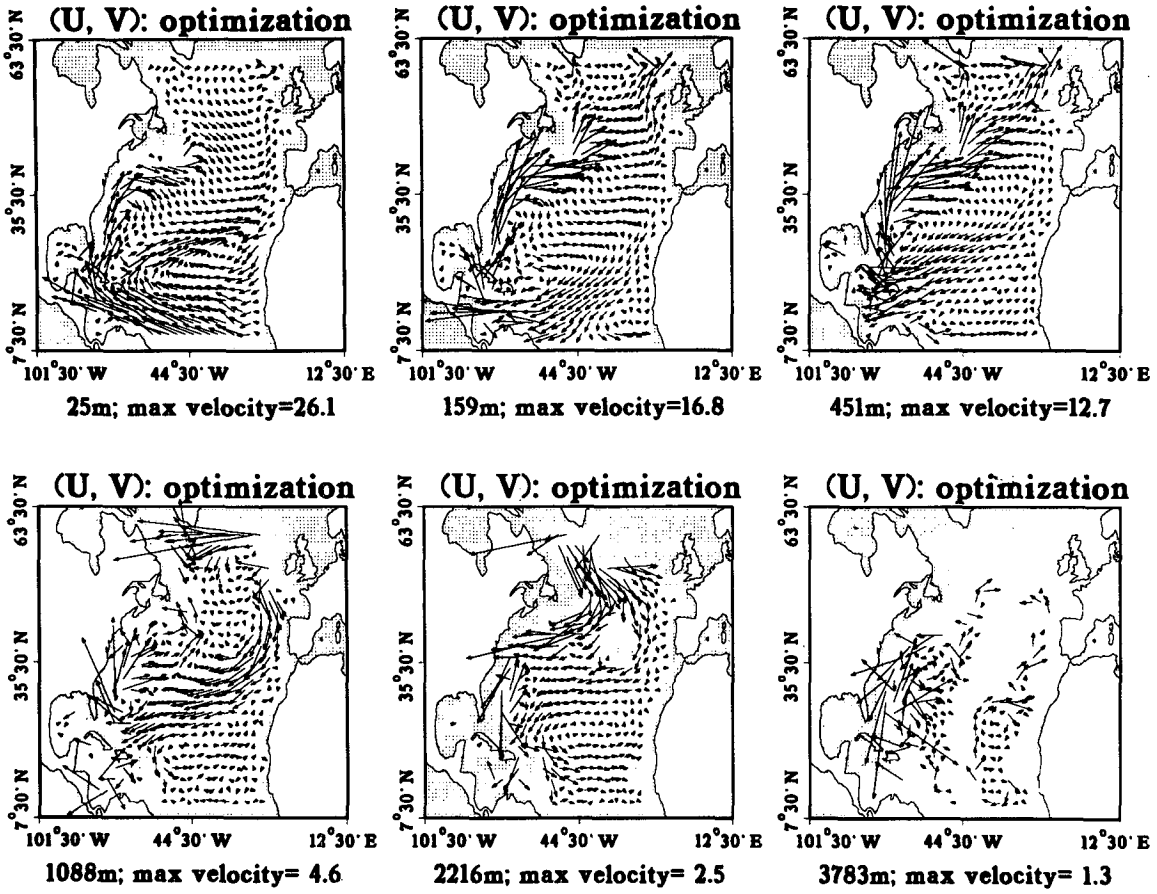

FIG. 4. The calculation of the optimal hydrography and general circulation of the North Atlantic Ocean (run Z; see Table 2 and section 4c). The horizontal velocity field at every second model level, as calculated by the optimization. The dotted regions denote land areas as specified in the model topography.

levels (2216 and 1589 m, not shown in the figure) show a southward-flowing western boundary current, and a southward interior flow. The circulation in the deepest levels, such as 3783 m (and 4672 m, not shown), is quite noisy and not much can be said about it. The circulation near the open boundaries in the northern and southern buffer regions is, of course, strongly biased by the presence of closed boundaries in the model at these locations. The velocity vectors in these regions (three grid points from both the northern and southern boundaries) should therefore be ignored. The same note of caution holds for the meridional circulation pictures (Figs. 5a, b). The portions of the meridional circulation in the buffer regions (e.g., the very strong upwelling at the southern boundary, about 8°N) are an artifact of the method used to treat the open boundary conditions and should be ignored.

2) THE ZONALLY AVERAGED MERIDIONAL CIRCULATION CELL

One of the important features of the adjoint method (being a nonlinear optimization method) is its ability to calculate optimal estimates for the temperature and salinity fields. Because the momentum equations in our model are diagnostic, the optimization modifies the velocity field by making the necessary adjustments to the temperature and salinity fields. These adjustments are of major importance to the results of the calculation and in particular to the meridional circulation. Figure 5a shows the streamfunction of the zonally averaged circulation, calculated from the original Levitus data as follows. The diagnostic momentum model equations were used to calculate a three-dimensional velocity field from the hydrographic data, and this velocity field was then zonally averaged and used to calculate the shown meridional circulation streamfunction. Note that no optimization is involved in this calculation. Figure 5b is again the streamfunction for the zonally averaged meridional circulation, calculated this time from the optimal solution for the temperature and salinity calculated by the adjoint method in run Z.

The meridional cell for the North Atlantic Ocean is believed to be composed of a northward surface flow and a returning southward deep flow, connected by a

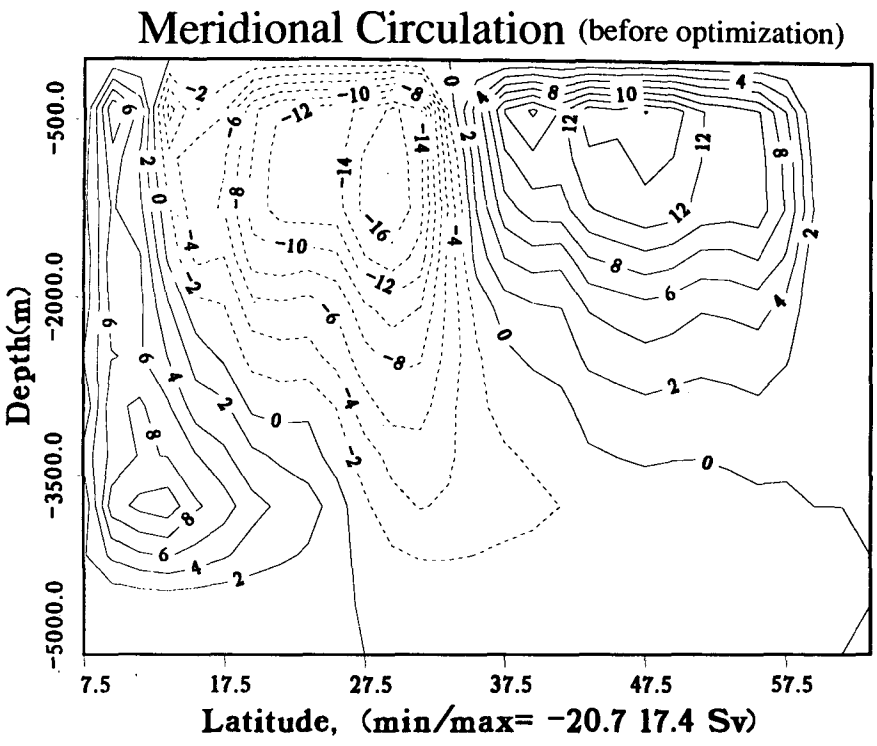

FIG. 5a. The meridional circulation calculated diagnostically from the Levitus data, before the calculation of the optimal hydrography (see Table 2 and section 4c).

sinking in a limited area in the north, and a broad upwelling over the whole North Atlantic Ocean. Such a meridional cell results in a net northward meridional heat flux carried by the ocean (see Wunsch 1984 for previous estimates of the meridional heat flux carried by the North Atlantic Ocean). Now, the meridional streamfunction calculated from the velocity field before the optimization (Fig. 5a) and based on the original Levitus hydrography, is quite far from the above picture. There is, in fact, a strong opposite cell between the latitudes 12°N and 35°N (dash contours in Fig. 5a), with a very strong upwelling signal at about 35°N.

This seemingly nonphysical meridional circulation seems a result of the inability of our data to resolve parts of the Gulf Stream recirculation region near the western boundary. The interior wind-driven Sverdrup transport of the North Atlantic subtropical gyre is normally assumed to be only about 30 Sv (Sv $\equiv 10^6$ m^3 s^{-1}), while the transport of the Gulf Stream before it separates from the North American continent is estimated to be more than 100 Sv. Much of this large transport is due to water recirculating in a relatively small area near the western boundary. This recirculation is composed of the narrow Gulf Stream flowing along the western boundary, the eastward flow of the separated current, and the westward returning recirculation water. The Levitus climatological data was heavily smoothed to remove small-scale noise, a process that in combination with the subsampling on our fairly coarse resolution grid, effectively removed the signal of the narrow Gulf Stream along the North American continent at about 30°N. The hydrographic data does contain the signal of the large transport of the separated Gulf Stream and some of the westward returning recirculation water; these can be seen in the velocity calculated diagnostically from the original data by the model (Fig. 6, right panel). Because our coarse-resolution data cannot resolve portions of the Gulf Stream, the large horizontal transport of the separated Gulf Stream seems in the diagnostic calculation to come out of the western boundary. The flow from the western boundary implies a strong upwelling at that region, supplying the water for the large transport of the separated Gulf Stream. This strong upwelling is the reason for the nonphysical meridional circulation calculated from the original data, and in particular for the opposite circulation cell seen in Fig. 5a. The recirculation region is characterized by very large deviations from steady state in the temperature and salinity equations evaluated from the original data, clearly indicating that the strong upwelling found in the diagnostic calculation based on the original data is neither correct nor consistent with the model's physics.

The optimization tries to reduce the deviations from steady state by modifying the temperature and salinity fields in order to eliminate the upwelling, and to supply the separated Gulf Stream water from south of the separation point rather than from the deeper levels. The

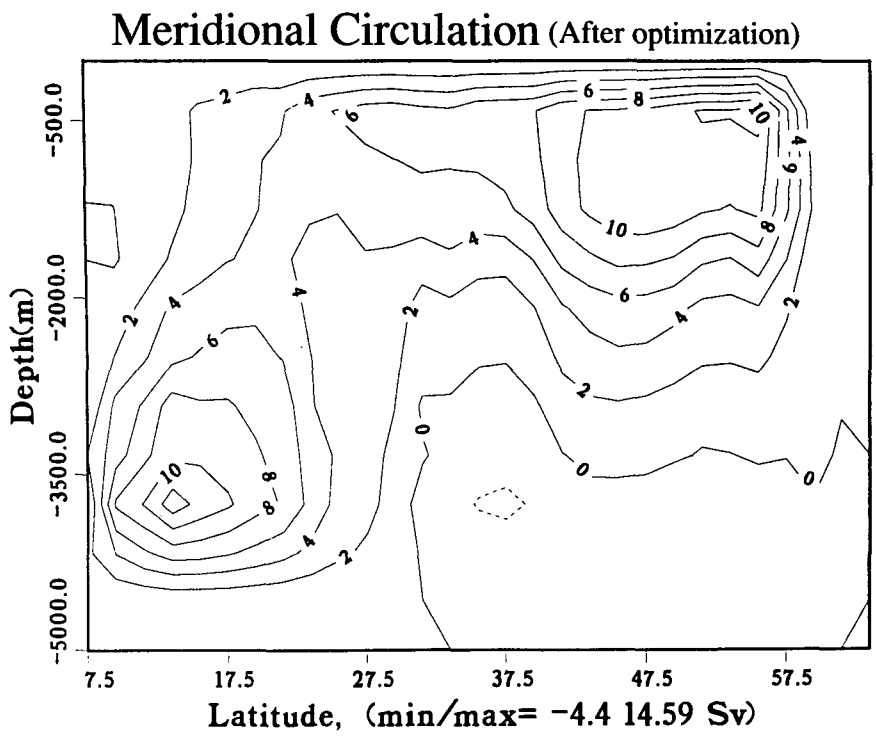

FIG. 5b. The meridional circulation calculated from the solution of optimization run Z (see Table 2 and section 4c).

main differences between the horizontal circulation before and after the optimization are indeed in the western boundary recirculation region: The optimal circulation contains the part of the Gulf Stream south of the Gulf Stream separation point that is not resolved by the original data. Figure 6 shows the velocity at the second model level before and after the optimization, and one can clearly see the northward velocity vectors along the western boundary north of the Straits of Florida that can be seen in the optimization solution (left frame of Fig. 6), but not in the velocity calculated diagnostically from the original data (right frame). In addition, the westward-returning part of the recirculation region, in the small region south of the separated Gulf Stream, was partially resolved by the original data but was removed by the optimization perhaps because it was not consistent with our simple linear momentum equations. The meridional circulation after the optimization is shown in Fig. 5b, and is much closer to what we expect to see.

Sarmiento and Bryan (1982) describe similar difficulties with purely diagnostic calculations, mentioning also previous works using diagnostic models. They show that the discrepancy between their rough bottom topography and the smoothed hydrography results in a highly cellular structure of the meridional circulation. These difficulties were overcome in their robust diagnostic calculation, which modified the density field near the bottom to be consistent with the model topography. In our case, the bottom topography is fairly smooth, and the diagnostic meridional cell not as noisy. But our diagnostic calculation still suffers major difficulties, which we attribute to the lack of horizontal resolution and oversmoothing of the hydrographic data in the Gulf Stream area.

3) THE MERIDIONAL HEAT FLUX

This force is shown in Fig. 7 both before and after the optimization. The figure shows the total (advective and diffusive) oceanic northward heat transport calculated in various ways: diagnostically from the original Levitus data; from the optimal hydrography obtained by optimization Z; and from the solution of the robust diagnostic calculation. As could be expected from examining the meridional cell, and from previous diagnostic models (Sarmiento and Bryan 1982), the meridional heat flux calculated by the model from the original Levitus hydrographic data (dash line) is quite opposite to what is expected: there is a southward heat flux in the latitude band corresponding to the opposite (anticlockwise) meridional circulation cell. The meridional heat flux calculated from the optimal hydrography is far closer to what we expect to find: a northward heat flux carried by the North Atlantic Ocean. Our meridional heat flux is similar to the free thermocline case of Sarmiento and Bryan (1982), although somewhat weaker. This similarity is probably due to the use of a small restoring coefficient γ in the robust

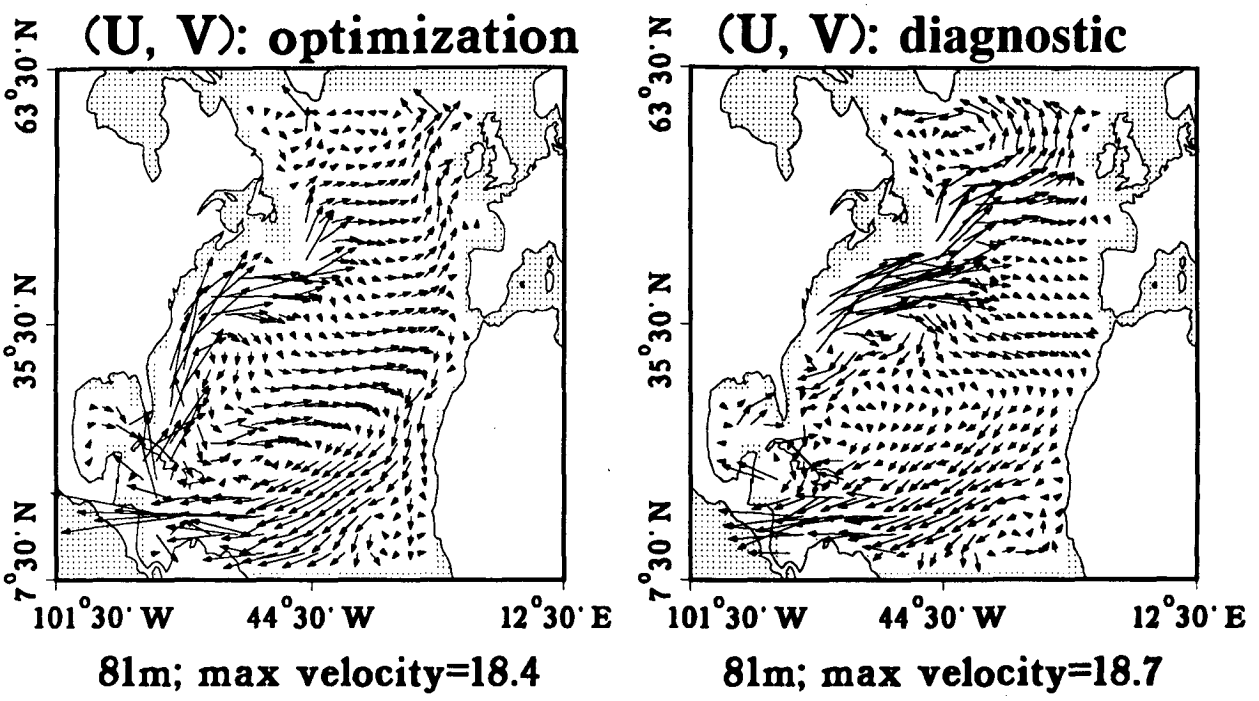

FIG. 6. The horizontal velocity at the second model level (81 m), as calculated diagnostically from the original hydrographic data (right), and as calculated by the optimization from the optimal temperature and salinity fields (left).

diagnostic run initializing the optimization, which allows the temperature and salinity to significantly deviate from the data values in a manner similar to the free thermocline case where the thermocline region was modeled with $\gamma = 0$.

An additional possibility that is, in fact, quite natural in an inverse calculation (but is not possible in regular numerical models or robust diagnostic calculations) is to constrain the meridional heat flux to be equal to what we think it ought to be by adding the appropriate constraint to the cost function, weighted by the corresponding error estimate. This is only one example of integral constraints that may be used as part of the optimization approach. Integral constraints may be useful also in improving the open boundary treatment in our model. Using buffer regions to handle open boundaries in numerical models seems to result in a too-weak meridional cell. In an inverse calculation, the buffer region (see section 3) may be supplemented with an integral constraint requiring the meridional cell closed in the buffer region near the boundary to be of a certain magnitude. Such integral constraints have proved useful in previous inverse calculations and should be implemented in inversions using a full GCM as in the present work, to take full advantage of the power of the optimization approach.

4) MIXING COEFFICIENTS

In all of the optimizations presented here, the horizontal and background vertical-mixing coefficients for the temperature and salinity equations are calculated as part of the unknown model inputs. Although the constant coefficient parameterization used here for the horizontal mixing is quite simplified, there are still some interesting points about the results we obtain. Ideally, in a steady model such as ours with no active mesoscale eddies, one would like to use tensor diffusivities (Redi 1982) that properly separate long- and

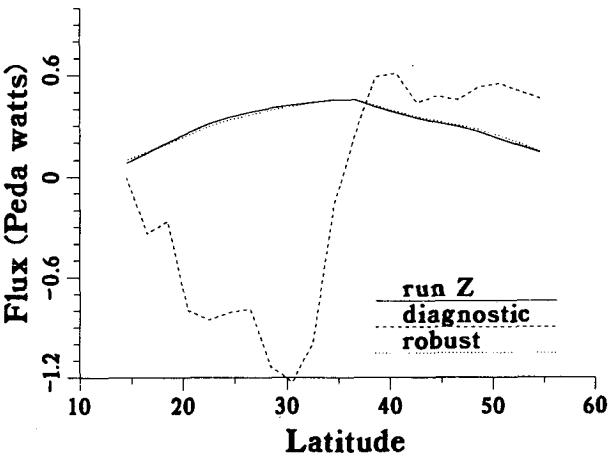

FIG. 7. The meridional heat flux, as calculated diagnostically from the Levitus data (dash), from the solution of optimization run Z (full line), using the robust diagnostic calculation preceding optimization Z (dots). See Table 2 and section 4c.

cross-isopycnal mixing in the ocean. The constant horizontal-mixing coefficient used here was, until recently, the standard parameterization in many coarse-resolution numerical studies of the large-scale ocean circulation. It is interesting, therefore, to use these results to examine how well such a parameterization fits the available data. Note that in the present calculation the mixing coefficients are not calculated locally at every horizontal grid point, but rather, a single mixing parameter is fit to the entire data. This is possible because our calculation is not local as in some beta spiral or similar inverse models (Olbers et al. 1985; Tziperman and Hecht 1988). A local calculation of the mixing coefficients may be useful for determining the present state of mixing in various oceanic regions, but cannot be used to find what mixing coefficient or mixing parameterization best fits the data globally. In addition, because the mixing is, locally, a small second-order effect, a local calculation of the mixing coefficients often fails to resolve the mixing coefficients from the noise in the data, and the coefficients are, more often than not, found to be indistinguishable from zero (Olbers et al. 1985; Tziperman and Hecht 1988), and also often negative unless explicitly constrained to be positive.

Our approach enables a global fit of a single parameterization to the whole hydrographic data, therefore avoiding the problems due to the small, local effect of the mixing. As the mixing processes (in particular, the vertical mixing) are most important in setting the global balances in the ocean and maintaining the vertical stratification in the ocean, a global fit of a mixing parameterization to the data may enable one to resolve the coefficients above the noise level in the data. A very encouraging indication that the present approach may indeed be able to successfully resolve the mixing coefficients is that the mixing coefficients calculated in all the optimizations using our model were positive, probably because the results are not dominated by the noise level as they seem to be in local calculations.

The horizontal-mixing coefficient used in coarse-resolution numerical models is intended as a crude representation of the mixing by the mesoscale eddies, and the value for this coefficient in such models is of the order of 10^7 cm^2 s^{-1}. The values calculated for this coefficient in the optimizations of sections 4a and 4b are smaller (0.2×10^6 to 0.4×10^7 cm^2 s^{-1}, see Tables 1 and 2), although the initial guess for this coefficient was always taken to be 10^7 cm^2 s^{-1}. The variation in the value of the coefficients is too large to allow any definite conclusions but may suggest that the simple mixing parameterization using the constant horizontal-mixing coefficient is simply not consistent with the hydrographic data, and the optimization therefore tries to make the coefficient as small as possible. This inconsistency is probably a result of the fact that horizontal mixing produces a much too large cross-isopycnal mixing in the presence of tilted isopycnals, and the tensor diffusivities mentioned above are thought to be a possible solution to this problem. It would be interesting to try to calculate a long-isopycnal mixing coefficient using the present approach in combination with a more sophisticated tensor diffusivity mixing parameterization, and to see whether such an improved parameterization results in a larger value for the long-isopycnal mixing coefficient. When starting the optimization with a robust diagnostic calculation (run Z), the optimization did not change the value of the mixing coefficients from their initial guess. That is probably because the robust diagnostic run strongly modified the temperature and salinity fields to be consistent with the coefficients used in that run (i.e., 1 and 10^7 cm^2 s^{-1}). The optimization was then unable to change both the hydrography and coefficients to further reduce the cost, although there may be a better solution with different mixing coefficients and a different hydrography.

As for the vertical-mixing coefficient, estimates based on observations vary between 0.1 and 3.0 cm^2 s^{-1}. Presumably, the value of this coefficient should not depend too strongly on the model resolution, as the small-scale mixing processes responsible for the vertical mixing cannot be resolved by even high-resolution GCMs and must be parameterized in any case. This leads to the hope that the value calculated here for the vertical-mixing coefficient can be used in other studies of the North Atlantic general circulation using numerical GCMs. The values found in the calculations presented here (Tables 1 and 2) tend to be in the lower part of the range of previous estimates (0.1 to 0.3 cm^2 s^{-1} for the background vertical mixing K_{v0}; in all optimization runs using the standard weights in the cost function, except for run Z, see above discussion of horizontal mixing).

5) ADJOINT METHOD VERSUS ROBUST DIAGNOSTIC

It has already been noted (Table 2 and Fig. 7) that the optimization solution of run Z does not represent a significant improvement (in terms of the cost reduction) over the robust diagnostic run used to obtain the starting point for the optimization. (It should be noted, though, that this is also the case because we have tuned the restoring coefficient, γ, in the robust diagnostic model, so that it would produce the smallest possible cost function at steady state.) It is no surprise, therefore, that the above results for the meridional circulation cell, and the difference between the diagnostic and optimal solutions, resemble very much those of the robust diagnostic model of Sarmiento and Bryan (1982). This is also the case for the improvement of the meridional heat flux: our results are not very different from those of the robust diagnostic calculations. Robust diagnostic calculations cannot be used to obtain an improved estimate of the surface fluxes, but neither could the adjoint method do this at this stage. The similarity of our results and those of the robust diagnostic calculations implies therefore that the potential benefits of the op-

timization approach using the adjoint method to calculate the gradient of the cost function have not been fully exploited here, mostly because of the difficulties due to ill conditioning.

6) RESIDUALS: IS THE MODEL CONSISTENT WITH THE DATA?

Figures 8a–f show the temperature and salinity data, the optimal solution for the temperature and salinity fields, as well as the residuals of the data and steady-penalty terms in the cost function, all given for three of the model levels, representing the surface, middepth, and deep waters. The residuals measuring the misfit to the data are $(\hat{T}_{ijk} - T_{ijk}^{n=0})$, and in order to help in physically interpreting the residuals of the steady-state penalties, they were multiplied by 15 years, ($[\partial T_{ijk}/\partial t]$ $\times$ 15 years). This gives a rough order-of-magnitude estimate for the change in temperature expected to develop if the model for 15 model years, starting from the optimization solution for the model inputs. Ideally, this change should be of the same order of magnitude as the temporal changes found by Levitus (1989) for the North Atlantic Ocean. Similarly, the deviation from the data should be of the same order of magnitude as the assumed errors given in Fig. 2.

Before examining the spatial structure of the residuals, we note that for the model and data to be consistent within the assumed error levels, we expect the total contribution of each of the terms in the cost function (1), as they are given in Table 2, to be less than or equal to one. The actual values of these terms are around two, indicating that on the average, the deviation of the steady and data penalties from their allowed

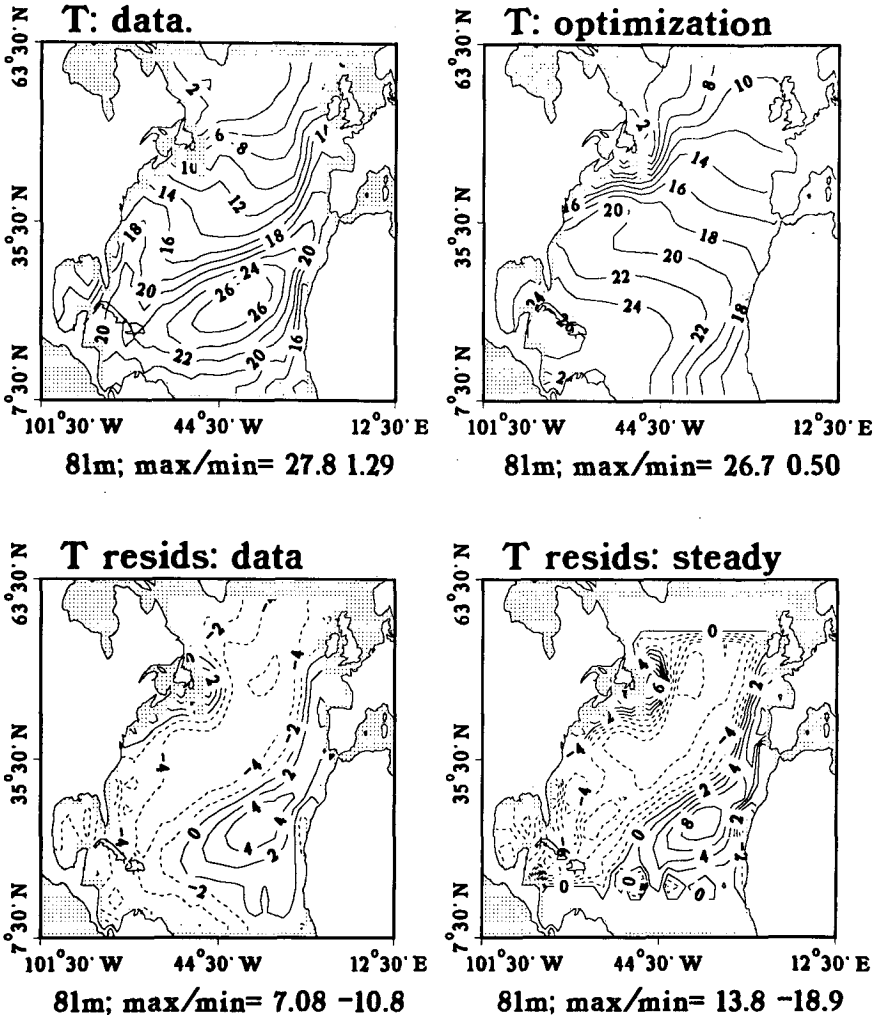

FIG. 8a. Upper left: the temperature data; upper right: the temperature calculated by optimization Z; lower left: the data residuals (difference between the data and optimal temperature); lower right: the deviation of the temperature equation from steady state (that is, the temporal derivative multiplied by 15 years; see text for details). All shown for level 2 of the model. Contour interval (CI) for all plots is 2°C.

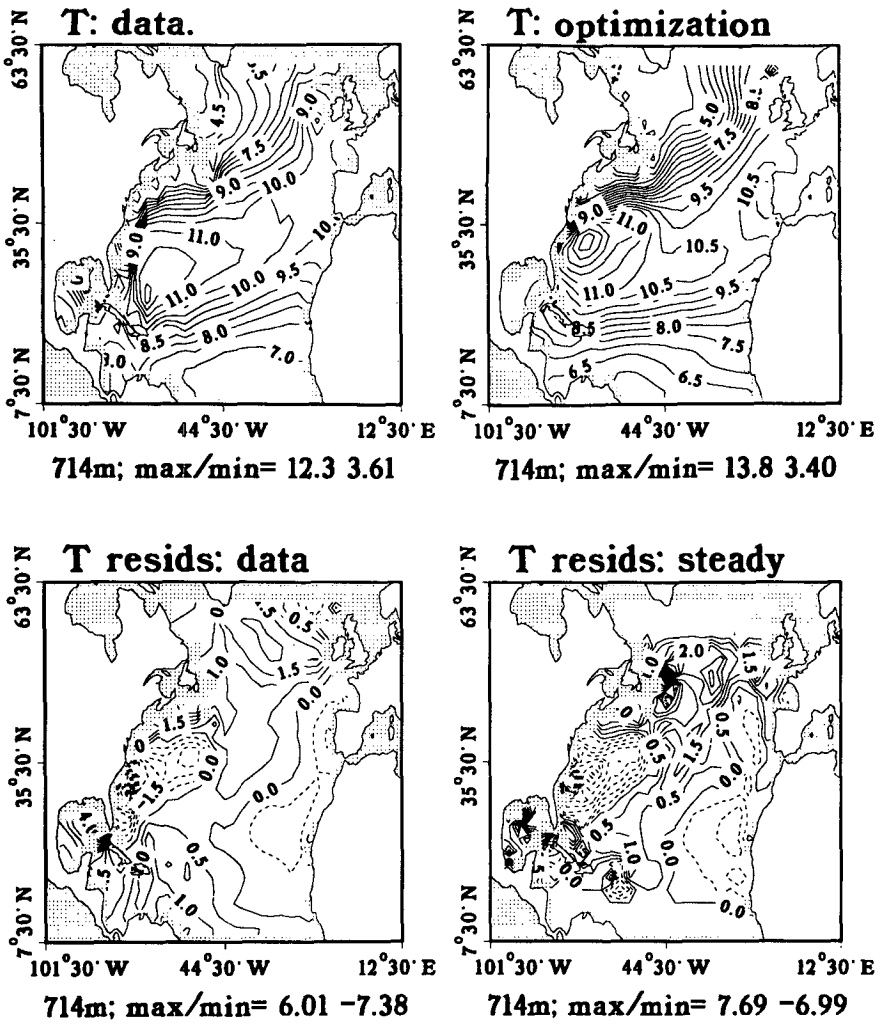

FIG. 8b. As in (a), for level 6 of the model: CI is 0.5°C.

values is somewhat too large, so that model and data cannot be expected to be fully consistent. More insight into the question of model and data consistency is obtained by examining the structure of the residuals. Looking at Figs. 8a–f, one immediately notices the similarity in the structure of the steady and data residuals. That is simply because our final solution is not very different from the solution of the robust diagnostic run used to initialize the optimization. In the steady-state robust diagnostic solution, the deviation from the data (multiplied by γ) is locally equal to the rest of the terms in the temperature and salinity equations (7).

When examining the residuals at the three plotted levels (81, 714, and 2216 m), it is useful to note that the assumed noise level for both steady and data residuals is (Fig. 2) 1.86°, 0.86°, and 0.16°C for the temperature residuals at the corresponding depths, and 0.23, 0.11, and 0.02 ppt for the salinity.

Beginning with the results at 81 m (Figs. 8a,d), we see unacceptably large temperature residuals of both types in two main regions: off the African coast (steady residuals of up to 8°C, compared to the allowed noise level of 1.86°) and in the northern part of the western boundary. In the later region, the salinity residuals are also large, and both the optimal temperature and salinity are characterized by sharper gradients than the data fields, somewhat similar to the results of Sarmiento and Bryan. At this level the model and data can probably be declared inconsistent. The reason may be problems with the heat fluxes that strongly influence the solution at this shallow depth, as well as a variety of other reasons such as vertical-mixing parameterization and mixed-layer dynamics that may be oversimplified in the model. Most probably, however, the reason is that our model is driven by surface fluxes, rather than by specified sea surface temperature and salinity. Oceanic GCMs are known to produce unacceptable results when driven by observed fluxes, and the model used here is probably not an exception. At the next

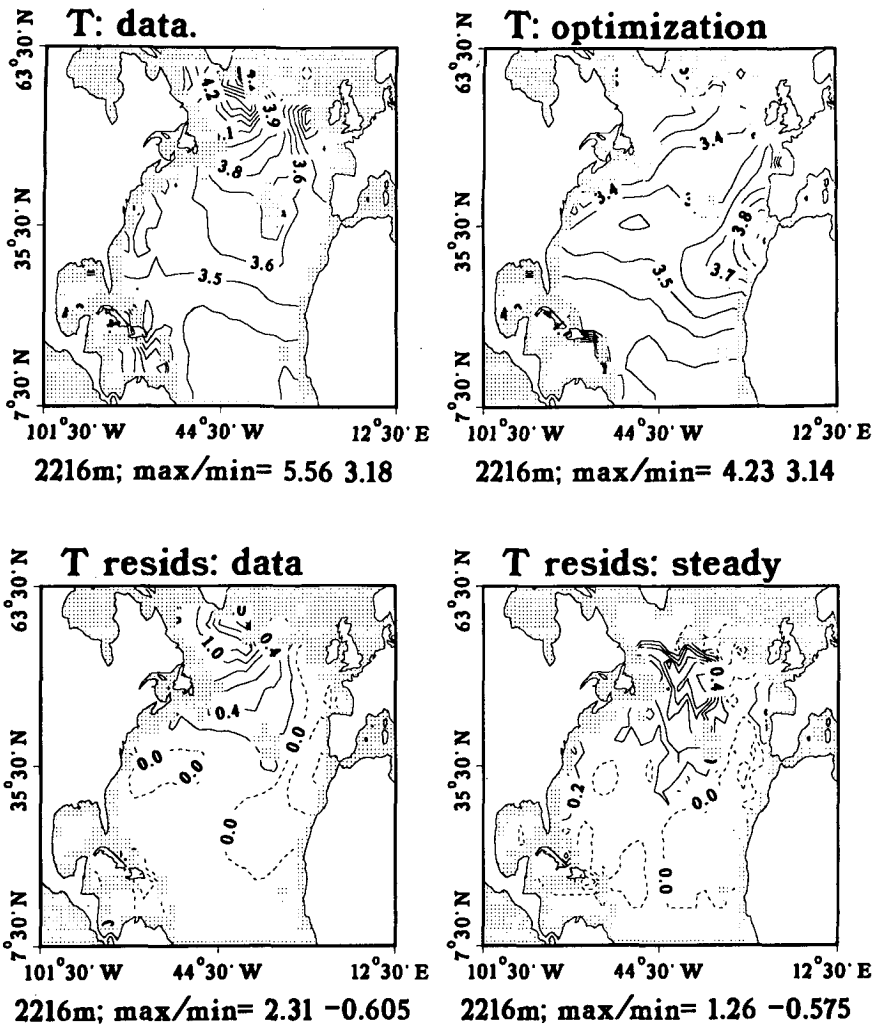

FIG. 8c. As in (a), for level 10 of the model: CI for the plots of the data, optimal solution, data misfit, and steady residuals are, correspondingly, 0.1°, 0.1°, 0.2°, and 0.2°C.

level shown, 714 m (Figs. 8b,e), the residuals are large especially in the western boundary region. This may be due to either the oversmoothing and insufficient resolution of the data in this region or the simplified linear dynamics of the model momentum balances, which may be too simplified for this region. Finally, in the deepest level shown, 2216 m (Figs. 8c, f), the assumed noise level for the temperature and salinity is 0.16°C and 0.02 ppt, correspondingly. The main differences between the optimal fields and the data are in two regions: the Mediterranean outflow, and more so in the northwestern part of the basin. In both cases, and in particular in the northern part of the basin, we tend to suspect the poor parameterization used in the present model for convective overturning, as well as the use of flux boundary conditions mentioned above. The convection parameterization is based on using a relatively large, yet perhaps not large enough, vertical-mixing coefficient in regions of unstable density stratification. This is done to avoid the problems, discovered using simulated data, in converging to the optimal solution when a stronger and less smooth convective-overturning parameterization was used. Deep convection (from the surface to 2 km deep) in the solution shown here occurs in the western part of the northern buffer region only, with shallower convection (to 200–400 m) in restricted areas of the northern part of the basin. Too weak vertical mixing due to convection may result in the less sharp temperature and salinity gradients of the optimal solution in the west northern corner of the basin, with the resulting large residuals there. As we do not calculate surface fluxes in this optimization, convection is not a major concern, but in the future, the problem of which deep convection parameterization to use can clearly be expected to be one of the important ones for optimization studies using GCMs.

Overall, the residuals have quite large values in large areas, notably in the western boundary regions. The spatial structure of the residuals is nonrandom. This

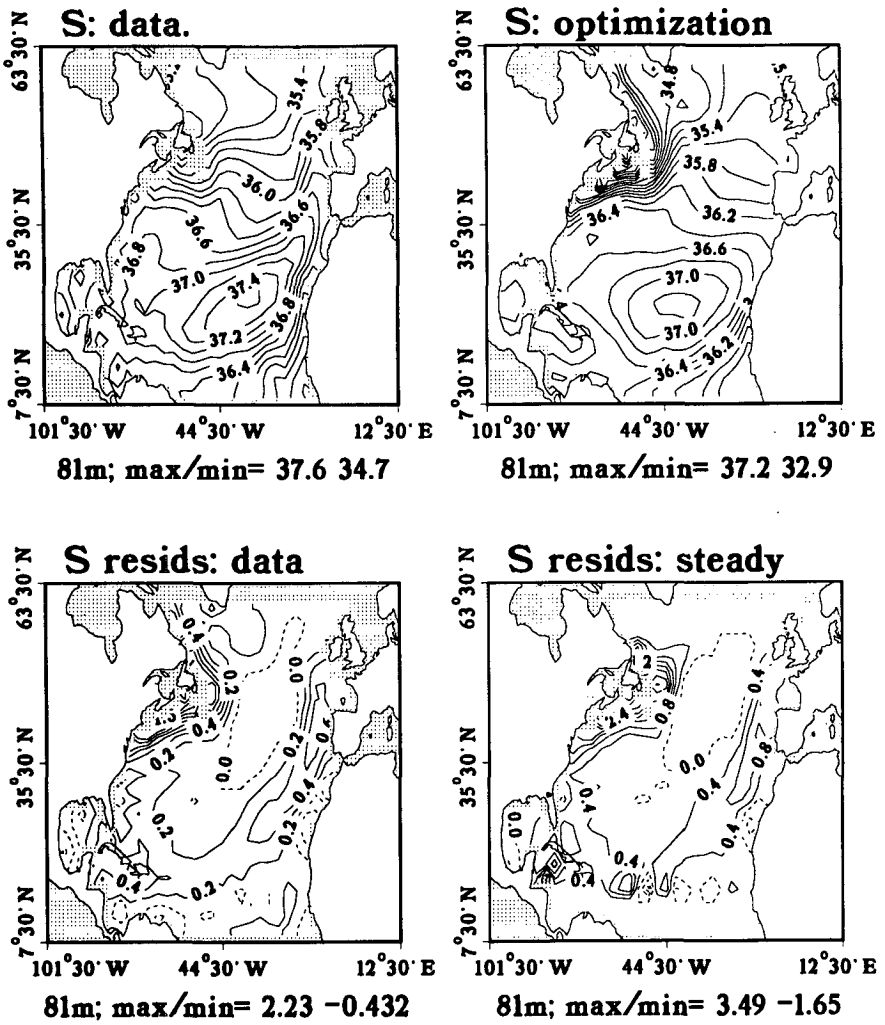

FIG. 8d. As in (a), but for salinity, at level 2 of the model: CI are 0.2, 0.2, 0.2, and 0.4 ppt.

may be expected with the smoothed data we use, but on the other hand, the correlation of the regions of large residuals with oceanographic features (western boundary currents, deep convection areas, Mediterranean tongue) indicates that what we see are inconsistencies between the model and the data, rather than just smoothed random errors in the data. The model-to-data fit can be improved by using both a better model (with nonlinear-momentum equations) and a higher resolution and perhaps more carefully analyzed hydrographic data (especially in the western boundary region) that may be able to resolve the Gulf Stream more fully. We conclude, as did Sarmiento and Bryan in their robust diagnostic study, that the model used to analyze the data must be accurate for the results to be useful.

7) ERROR ESTIMATES

The absence of formal error estimates for the calculated model parameters and inputs is somewhat of a deficiency of the present calculation. Although this information may be obtained from the Hessian matrix [matrix of second derivatives of the cost function with respect to the model inputs; see Thacker (1989) or Tziperman and Thacker (1989)], this matrix is quite difficult to handle for large-scale optimization problems such as the present one. In fact, the ability of the adjoint method to calculate the solution separately from the error information is what enabled us to approach the difficult problem of analyzing the North Atlantic general circulation using a fairly complex GCM.

Although the large size of the Hessian matrix will probably prevent the calculation of formal error bars for the solution obtained by the adjoint method in the near future, our results serve to demonstrate how some information concerning these errors can still be obtained. For example, when the minimum of the cost function is not a sharp minimum, but has the form of a valley in parameter space, as we have found in one of our optimizations, then the calculated parameters may not be uniquely determined by the data used (al-

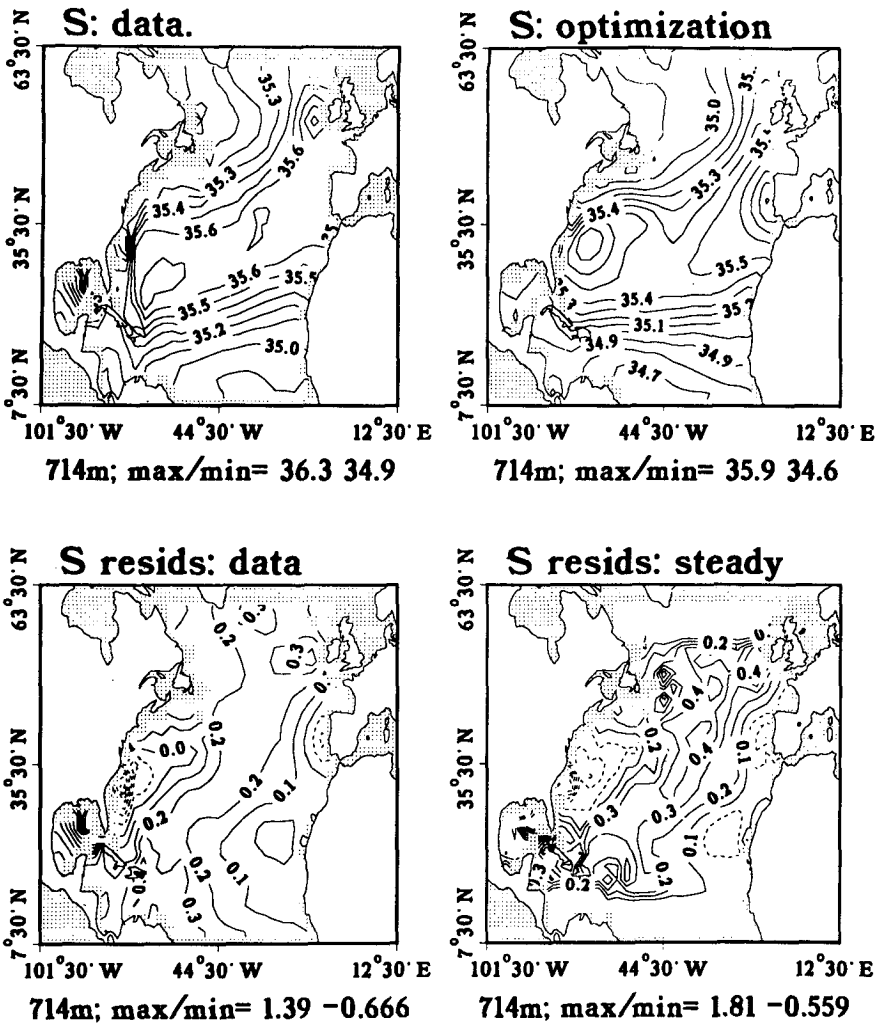

FIG. 8e. As in (d), for level 6 of the model: CI is 0.1 ppt.

though, as explained in section 4a, that was not the case here). This implies that they are in the null space of the resolution matrix that may be calculated from the Hessian, and their error bars may be expected to be quite large. The experiments of section 4a serve as an example on how the structure of the cost function can be explored without calculating the prohibitively large Hessian matrix.

We now proceed to analyze in detail the conditioning of optimization problems involving fitting a steady model to oceanographic data.

5. Analysis of the conditioning of optimization problems using a steady-state model

At this point, we can summarize what has been learned from the above calculations as follows: The optimization problem involving the fit of a steady model to oceanographic data using the cost function given by (1) seems poorly conditioned. In trying to solve the problem using the optimization approach alone, we have obtained solutions that were often not acceptable and were unable to find the optimal solution. When analyzing the difficulties due to the poor conditioning, it is quite clear that the data penalties in the cost function are not to blame. These are simply quadratic in the unknown temperature, salinity, and surface fields, and an optimization problem involving these terms only can be solved in a few iterations at most. The main difficulty seems to lie in the steady-penalty terms.

In order to gain some understanding of why the problem of poor conditioning of the steady penalties occurs and perhaps also how it may be overcome, several simple problems are considered in which we will examine the conditioning of various forms of the steady penalties using highly simplified model equations. For this purpose, we ignore the questions related to the calculation of boundary conditions and the use of data and concentrate on the calculation of the steady so-

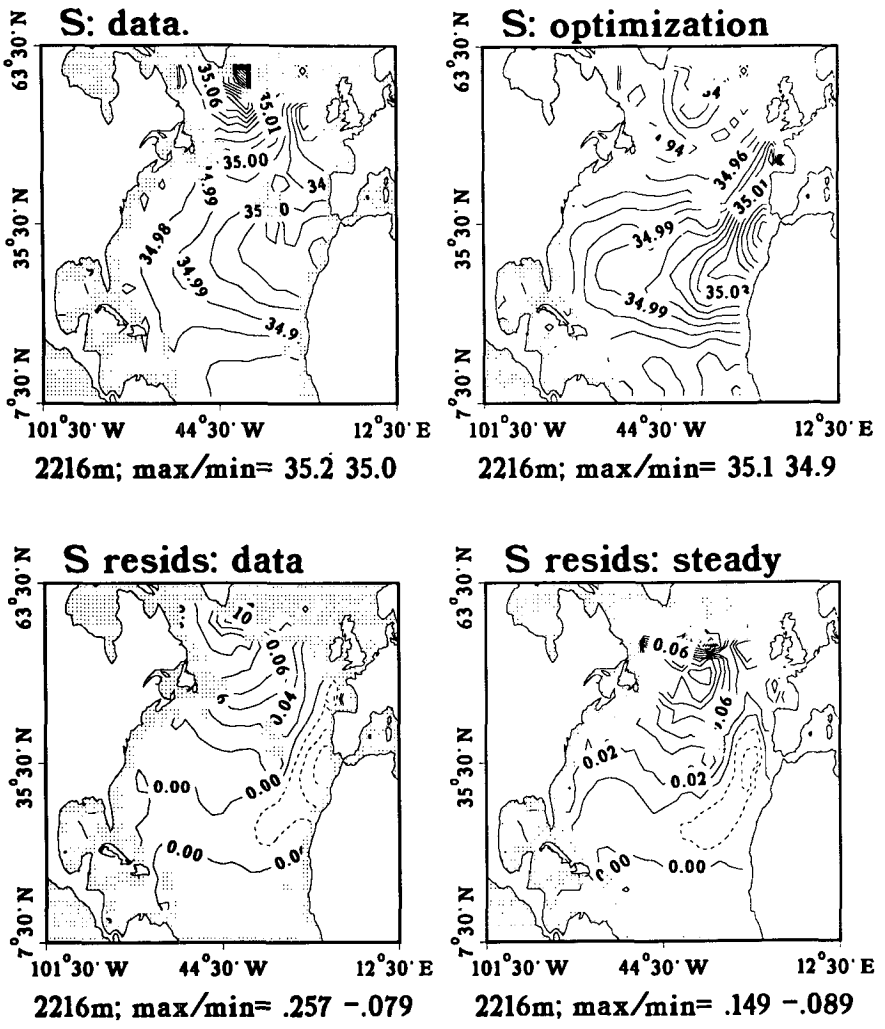

FIG. 8f. As in (d), for level 10 of the model: CI are 0.01, 0.01, 0.02, and 0.02 ppt.

lution using the optimization approach. The analysis here is, in a way, an extension of the work of Tziperman et al. (1992), who calculated the steady-state solution to a QG model using the optimization approach. Marotzke (1992) has used the same model used here, penalizing the deviation from steady state using terms of the form $(T_{\text{final}} - T_{\text{initial}})^2$ as we do here, but with the important difference that the initial and final temperatures are separated by longer time integrations rather than a single time step as we do here. He suggested that some of the problems found here may be overcome this way. In the analysis that follows we examine the conditioning both for the one time step approach, for Marotzke's approach, and also for alternative forms of the cost function based on longer integrations. In doing that we try to derive both a more optimal form of the cost function and the minimum integration time per iteration necessary to obtain good conditioning.

These objectives are achieved by considering the effects of both dissipation and wave motions on the conditioning of the optimization. First, a word of explanation on why it is useful to consider the waves present in the model equations. We assume that the optimization is started with an initial guess for the temperature and salinity that is near the steady-state solution (a reasonable first guess would simply be equal to the data). This initial state may be considered a perturbation about the steady solution for the model equations, and we expand this perturbation in linear modes corresponding to the model waves. The purpose of the steady penalties in the cost function is to eliminate these waves from the initial conditions, and to drive the initial conditions toward the steady-state model solution. In the following discussion, we argue that by using in the cost-function steady penalties that are the sum of the squared differences between the initial conditions and all (or at least some) of the following time steps, the conditioning of the problem may be expected to improve. The analysis suggests that using only the difference over one time step, or only the difference

between the final and initial states for more than one time step, may be expected to be less effective.

Consider first the 1D diffusion equation,

$$T_t = \kappa T_{xx}, \tag{9}$$

and assume initially that we pose the optimization problem based on taking a single time step with a finite-difference model based on this equation. The cost function will be

$$J = \frac{1}{2} \sum_i [T_i^1 - T_i^0]^2, \tag{10}$$

where $T_i^n = T(x_i, n\Delta t)$, $0 < x_i < L$, and

$$T_i^1 = T_i^0 + \kappa \Delta t (T_{i+1}^0 - 2T_i^0 + T_{i-1}^0)/(\Delta x)^2. \tag{11}$$

Using (11), the cost function can be written in matrix form as

$$J = \frac{1}{2} \| (I + \kappa \Delta t \partial_{xx}) T^0 - T^0 \|^2 = \frac{1}{2} \| (\kappa \Delta t \partial_{xx}) T^0 \|^2,$$

where I is the unit matrix and ∂_{xx} represents the operator (matrix) of double differentiating with respect to x. Taking the second derivative of the cost with respect to the initial conditions T_0, we find that the Hessian matrix is simply

$$H = (\kappa \Delta t \partial_{xx})^2.$$

The eigenfunctions (eigenvectors in the discrete matrix formulations) are simply $\exp(ikx)$. For the continuous problem, the eigenvalues of the Hessian matrix are therefore simply $\Lambda_k = (-\kappa \Delta t k^2)^2$. Using the finite-difference operator, the eigenvalues are

$$\Lambda_k = \left(-2\kappa \frac{\Delta t}{(\Delta x)^2} [1 - \cos(k\Delta x)] \right)^2.$$

The largest wavenumber represented in the model is $\pi/\Delta x$, and the smallest is π/L. The condition number for this problem is therefore simply

$$C = \frac{\max(\Lambda_k)}{\min(\Lambda_k)} = \left(\frac{2}{1 - \cos(\pi \Delta x/L)} \right)^2 \approx \left(\frac{2L}{\pi \Delta x} \right)^4.$$

For the optimization problem to be well conditioned, this number should be as close to order one as possible, which will not be the case for large L and small Δx.

Assume, next, that the model is integrated for more than one time step, as suggested by Marotzke (1992), and use for the steady penalties the form he suggested,

$$J = \sum_i [T_i^N - T_i^0]^2. \tag{12}$$

Writing now the temperature at the Nth time step in terms of the initial conditions, we obtain

$$T^N = (I + \kappa \Delta t \partial_{xx})^N T^0,$$

which results in the following expression for the Hessian

$$H = [(I + \kappa \Delta t \partial_{xx})^N - I]^2.$$

The Hessian eigenvalues for this case are $\Lambda_k = [(1 - \kappa \Delta t k^2)^N - 1]^2$ for the continuous operators, and

$$\Lambda_k = \left[\left(1 - 2\kappa \frac{\Delta t}{(\Delta x)^2} [1 - \cos(k\Delta x)] \right)^N - 1 \right]^2$$

for the finite-difference operators. The condition number is therefore

$$C = \frac{\max(\Lambda_k)}{\min(\Lambda_k)} = \left(\frac{\left(1 - \frac{4\kappa \Delta t}{\Delta x^2} \right)^N - 1}{\left[1 - \frac{2\kappa \Delta t}{\Delta x^2} \left(1 - \cos \frac{\pi \Delta x}{L} \right) \right]^N - 1} \right)^2. \tag{13}$$

Noticing now that the term $2\kappa \Delta t/(\Delta x)^2$ must be smaller than one due to the Courant stability condition, we see that the condition number approaches one as the number of time steps N increases. We conclude that, for this problem, increasing the number of time steps and formulating the steady penalties as in (12) should indeed improve the convergence rate of the optimization. The physical reason for this improvement is that by extending the integration time, the longer waves had the opportunity to feel the dissipation by the diffusion term and therefore they contributed to the change between the initial and final states and thus to the cost function. When stepping the model a single step or a few time steps, only the small scales were able to feel the dissipation, resulting in the mismatch between the way short and long waves are affected by the dissipation. This mismatch between the way waves of large and small wavenumbers are handled will result in a similar mismatch between their corresponding eigenvalues, and therefore in poor conditioning. But would this conclusion also hold when dissipation is not the dominant physical process? In order to answer this question, consider the following simple problem.

Suppose our model consists now of the wave equation obtained from the advective part of the temperature equation,

$$T_t + uT_x = 0. \tag{14}$$

Consider the optimization problem of trying to calculate the steady state using the cost function based on integrating the model for a time τ,

$$J = \sum_i [T(x_i, \tau) - T(x_i, 0)]^2. \tag{15}$$

Assume that the initial conditions (first guess for the steady state) include a single wave component, so that we can write

$$T(x, 0) = \cos(kx).$$

The solution to our model equation is then simply $T(x, \tau) = \cos[k(x - u\tau)]$, so that the cost function is

$$J = \sum_i [\cos[k(x_i - u\tau)] - \cos(kx_i)]^2. \tag{16}$$

Clearly, if the wavenumber k satisfies the condition $ku\tau = 2n\pi$ for some integer n, it would not contribute to the cost function and therefore cannot be eliminated from the initial conditions in order to obtain the steady-state solution using the optimization approach. In the full problem, where many waves are present in the initial conditions, there will always be wave components that will not contribute to the cost function formulated as in (7) (even waves that only approximately satisfy $ku\tau = 2n\pi$ will hardly affect the cost function). Such insensitivity of the cost function to components of the initial conditions clearly means poor conditioning. In order for all waves to affect the cost function, the integration time τ needs to be long enough for dissipation to affect all waves, which may imply fairly long integration times for the longer waves. Let us examine a different, and perhaps better, formulation of the cost function, which may get around this problem and shorten the integration time needed to obtain good conditioning.

In order for a wave of a given period to have a significant effect on the cost function, it is best to include in the cost function a term that is squared in the difference between the initial conditions and the temperature at a time when the deviation from the initial conditions is significant. Sometime within a time equal to a quarter of the period of the wave, the change from the initial conditions will be at least half the amplitude of the wave (may be up to the full amplitude, depending on the phase of the wave at the initial time). A reasonable way to make sure the wave motions affect the cost function could therefore use a cost function of the form:

$$J = \sum_{i=1}^{L/\Delta x} \sum_{n=1}^{N} w_n [T(x_i, n\Delta t) - T(x_i, 0)]^2, \quad (17)$$

where w_n are weighting coefficients, and the integration time $N\Delta t$ is at least equal to a quarter of the wave period. [Note that (17) is a special case of the general steady penalties for this 1D problem that can be written as

$$J = \sum_{i>j=1}^{L/\Delta x} \sum_{n>m,n'>m'}^{N} [T(x_i, n\Delta t) - T(x_i, m\Delta t)] \times w_{ijnmn'm'} [T(x_j, n'\Delta t) - T(x_j, m'\Delta t)],$$

where $w_{ijnmn'm'}$ is a positive-definite weighting matrix.] In order for all waves to be constrained by this cost function, the number of time steps (N) should be chosen such that the integration period is larger than a quarter of the period for the slowest waves in the model. This does not bring into consideration, however, the effects of the dissipation in the model equations, which we will do now by combining the two simple problems discussed here into one with both waves and dissipation.

Consider the model equation,

$$T_t + uT_x = \kappa T_{xx}. \quad (18)$$

As above, we are trying to find the way to precondition the optimization problem involved in calculating the steady solution of this model equation. Let the initial guess for the steady state be composed of various wave components,

$$T(x, 0) = \sum_k \hat{T}_k \cos(kx). \quad (19)$$

Then the solution after a time τ is

$$T(x, \tau) = \sum_k \hat{T}_k \cos[k(x - u\tau)] \exp[-\kappa k^2 t]. \quad (20)$$

Now, in order for the cost function (17) to be sensitive to a given Fourier component, that component needs to vary significantly from the initial time to some time $\tau_n = n\Delta t$, so that it would contribute to the corresponding term $[T(x, \tau_n) - T(x, 0)]^2$ in the cost function. Such a change from initial conditions could happen in two ways. For the long waves this would most effectively happen through the difference in temperature between the initial conditions and a time within a quarter of the wave period $\tau_{1/4} = (2\pi/\sigma_r)/4$, where σ_r is the real part of the frequency. These long waves need very long integration times in order to feel the effect of dissipation, which would imply very long integration times per iteration of the optimization, and therefore a significantly higher computational cost for the entire minimization process if we rely only on the dissipation effects. The quarter-period time may be much shorter than the dissipation time for the long waves, hopefully resulting in a more effective conditioning.

The shorter waves, on the other hand, are more strongly affected by the dissipation through the exponential term in (20). The significant difference between the initial and final conditions is most effectively achieved for these waves by integrating the model for a time that is equal to or larger than the dissipative time scale for these waves, $\tau_d = (\kappa k^2)^{-1}$, which should not be large, because the wavenumber k for these waves is quite large.

To summarize, the integration time should be chosen so that the cost function is sensitive to all wavenumbers. If the integration time τ is larger than the decay time for some short wave, then all terms in the cost function for which $n\Delta t > \tau_d$ will be sensitive to this wave component. Otherwise, for longer waves, at least one of the terms in the cost function will be sensitive to the phase change of the wave through the quarter period constraint.

Having (hopefully) obtained some intuitive feeling for how a good conditioning of a problem involving both dissipation and waves can be effectively achieved, we now consider the more realistic oceanic case, where the main waves carrying information around are the Rossby waves. In a GCM, where dissipation is present

both in the momentum and in the temperature and salinity equations, the Rossby waves are dissipated by both horizontal and vertical mixing processes, so that the dispersion relation for the first mode baroclinic waves (higher modes will be considered below) can be written as

$$\sigma = \sigma_r + i\sigma_i$$

$$= \frac{-\beta k}{k^2 + l^2 + L_R^{-2}} - i[\kappa_h(k^2 + l^2) + \kappa_v m^2], \quad (21)$$

where κ_h and κ_v are the horizontal and vertical mixing in the momentum equations (mixing in the temperature and salinity equations will have a similar effect).

Our purpose is to find the minimal integration time of the model that will provide good conditioning using the cost function of the form (17). Let us divide now the wavenumber range into two main subranges (which may overlap, as we will see below): the long-wave range in which preconditioning is most effectively achieved by constraining waves over a quarter-period time scale, and the short-wave subrange in which conditioning is best achieved through the dissipative effects acting over the diffusion time scale.

As a first step toward estimating the optimal integration time for the model, let us assume that the model is integrated for a time τ and examine which wavenumbers are properly conditioned by such an integration. Consider first the long waves. For these, the relevant part of the dispersion relation is the real part, which should yield a relation between the wavenumbers of the well-determined waves and the integration time. The well-determined waves are those whose quarter-period time is equal to or smaller than the integration time. This may be translated into a relation between their wavenumbers and the integration time,

$$\tau \geqslant \tau_{1/4} = \frac{1}{4}\frac{2\pi}{\sigma_r} = \frac{\pi}{2}\frac{k^2 + l^2 + L_R^{-2}}{\beta k}. \quad (22)$$

Let, for simplicity, $l = k$, and we now wish to extract the wavenumbers of the well-determined waves in terms of the quarter-period time. Because of the shape of the dispersion relation for the Rossby waves, there are two wavenumbers that correspond to a given frequency (or, equivalently, to a given quarter period). One would be in the very long-wave part of the spectrum, and the other toward the shorter waves. Denote these wavenumbers by $k_<$, which is the wavenumber smaller than the inverse Rossby radius L_R^{-1}, and $k_>$ that is larger than L_R^{-1} (Fig. 9). Using the appropriate approximate limits of (22) for short and long waves, we obtain the two wavenumbers bounding the wavenumbers of the well-determined waves:

$$k_> = \frac{\tau\beta}{\pi}; \quad k_< = \frac{\pi}{2\beta\tau L_R^2}. \quad (23a,b)$$

Waves with wavenumbers that fall between $k_>$ and $k_<$ will have quarter periods that are smaller than our integration time τ and should therefore be well determined.

Next, consider the short waves, to which the cost function may be sensitive due to the effects of dissipation. For the dissipation to have a significant effect on a wavenumber k, the integration time must be equal to or larger than the diffusion time,

$$\tau \geqslant \tau_d = \sigma_i = [k_h(k^2 + l^2)]^{-1}. \quad (24)$$

Assuming again $l = k$, we obtain an expression for the longest wave that is affected by the dissipation during an integration time τ. Denoting this wavenumber by k_d, we have

$$k_d = [2\kappa_h\tau]^{-1/2}. \quad (25)$$

Note that we use only the horizontal diffusion term for estimating the diffusion time scale, because for the first baroclinic mode horizontal mixing is expected to be more effective than vertical mixing. Vertical mixing will be included when discussing the higher modes below.

Now, for the least-squares optimization problem of calculating the steady state to be well conditioned, all waves must be affected by either the quarter-period constraint, or by the dissipation effects. Consider Fig. 9, where the different wavenumbers are plotted with the dispersion relation for the Rossby waves. When the integration time τ is small, there are two ranges of well-determined waves (stippled horizontal axis in Fig. 9), but there may be two gaps of poorly determined waves. The first gap is for wavenumbers $2\pi L^{-1} < k < k_<$ at the long-wave part (L is the basin size, and $2\pi L^{-1}$ is

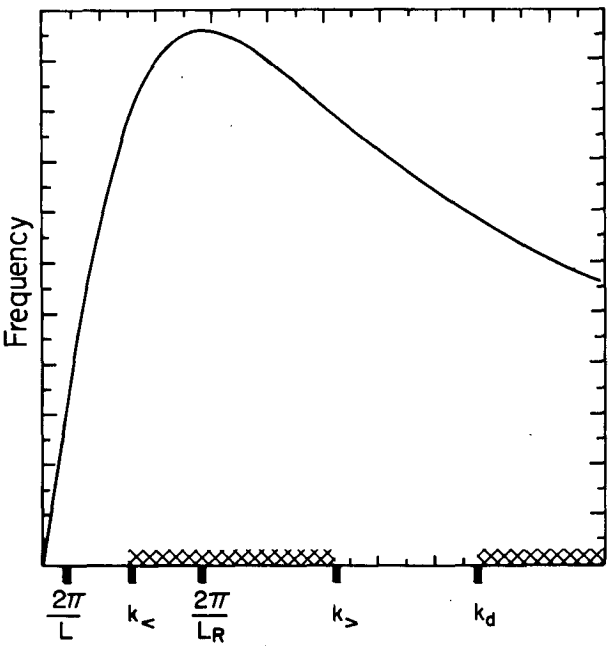

FIG. 9. A schematic drawing of the Rossby wave dispersion relation, indicating the ranges of well-determined waves for a given integration time used to calculate the cost function (17).

therefore the longest wave represented in the model). The second gap is for waves that satisfy $k_> < k < k_d$, that is, waves between those that are affected by the quarter-period constraint and those affected by dissipation.

For these gaps to be eliminated we need to first extend the integration time τ so that $k_> = k_d$. Using (23a) and (25), we obtain that the condition for closing the gap for the shorter waves is

$$\tau \geqslant \left[\frac{\pi^2}{2\kappa_h \beta^2}\right]^{1/3}. \tag{26}$$

Substituting $\kappa_h = 10^7$ cm^2 s^{-1} and $\beta = 2 \times 10^{-13}$ cm^{-1} s^{-1}, we have $\tau \approx 2 \times 10^6$ s, that is, about one month. But substituting this integration time into (23b) using $L_R = 50$ km, we find that the longest wavelength to be affected by such an integration time is only about 400 km. To extend this to a 4000-km basin, we need the integration time to be 5 to 10 times longer, which implies

$$\tau \approx 1 \text{ year}. \tag{27}$$

Next, we need to consider the effects of the barotropic and higher-mode baroclinic waves. Barotropic waves pose no problem. They are characterized by short periods and therefore are well determined by the quarter-period constraints in the cost function for short integration times, well within the times required for the first-mode baroclinic waves. Although the barotropic and first-mode baroclinic waves already account for much of the oceanic structure and variability, it is still necessary to consider the higher-mode baroclinic waves. For the latter, vertical mixing is expected to play an important role, because of the smaller vertical scale in the upper ocean. Taking, for example, 300 m as the vertical scale of the second baroclinic mode in the upper ocean, we find that the diffusion time for it due to the vertical dissipation in the momentum equations is

$$\tau_d(\text{vertical mixing}) = (\kappa_v m^2)^{-1}.$$

Using a minimal value for the vertical mixing $\kappa_v = 1$ cm^2 s^{-1} (coarse-resolution GCMs use values between 20 and 50 cm^2 s^{-1}) and m $= 2\pi/300$ m, we find that $\tau_d = 2 \times 10^7$ sec, that is, less than a year. In the deeper ocean the higher baroclinic modes have larger vertical scales, but on the other hand, the stratification there is well represented by the barotropic and first baroclinic modes, so we expect the convergence of the optimization in the deep ocean to still be reasonably good.

The significance of the above estimate for the optimal integration time (19) is that, by using the proper form of the cost function, constraining the large scales by the quarter-period constraint and the small scales by the dissipation effects, the integration time per iteration of the optimization may be significantly shorter than when using only the dissipation to constrain the optimization toward a steady state. This may enable the analysis of higher-resolution models for which integration time and storage of the history of the time integration (required for the adjoint model) are crucially limiting factors. These suggestions for a reformulation of the steady penalties clearly need to be implemented and used within an inverse calculation using a GCM to thoroughly examine their usefulness, and work is under way to do that.

6. Conclusions

We have used a GCM and North Atlantic climatological data of temperature, salinity, wind stress, evaporation minus precipitation, and air–sea heat fluxes to examine the possibility of solving inverse problems using a full-scale numerical GCM and real oceanographic data. These inverse problems are aimed at calculating model inputs, such as surface forcing fields, mixing parameters, and the temperature and salinity fields in a manner consistent with both the available data and the model dynamics. The resolution and complexity of the model used here are superior to what has been possible to attempt in previous inverse studies, thanks to the use of the optimization approach based on the adjoint method for calculating the gradient of the cost function. As this is the first application of the method to a complex GCM and real oceanographic data, this work was mostly intended to examine and present the possibilities and difficulties that are encountered when using a GCM within an inverse calculation, using the well-studied North Atlantic as a test case. In Part I of this work, we have used simulated data and found that while the minimization of the cost function proceeded quite well with model-generated data with no added noise, problems arose when simulated noise was added to the data. In this part, we proceeded to examine the performance of the method in more realistic circumstances using real data.

By examining several solutions for the model inputs, and by examining the structure of the cost function as a function of the model inputs, we have tried to demonstrate two of the main difficulties confronting such large-scale nonlinear inverse problems (about 30 000 unknowns and a similar number of constraints for the problem examined here). The first is the possible existence of local minima of the cost function preventing the convergence of the optimization to the global minimum representing the desired optimal solution for the model inputs. The second difficulty, which seems the dominant one for many of the problems examined in this part as well as in Part I, is the ill conditioning of the inverse problem. The ill conditioning is manifested in the cost function being flat in some directions in parameter space, possibly resulting in the stalling of the optimization and an inability to converge to the optimal solution. Such ill conditioning may be the result of insufficient data to calculate some of the model parameters, as well as the result of an inappropriate formulation of the inverse problem.

In order to better understand the seemingly bad conditioning of the optimization problems attempted here, we have used simple model equations to analyze the conditioning of the optimization and analyze the role of both dissipation and waves in the model dynamics in conditioning the problem. This analysis helps in understanding the difficulties encountered in this study, as well as to suggest some possible remedies. In a recent work using the same model and method used here, Marotzke (1991) has suggested that the constraints enforcing steadiness of the model temperature equation, for example, should be formulated as the sum of squares of the final temperature minus the initial temperature, as done here, but rather than one time step separating them, he used much longer integration times, up to 50 years. Motivated by this suggestion, we have examined the conditioning of steady penalties based on the one time-step approach, of penalties based on longer integrations, and also of different forms of the steady penalties in the cost function. The analysis suggests that formulating the constraints that the model be steady in a somewhat different way, by taking advantage of the model waves, may improve conditioning while requiring shorter integration times, of the order of one to a few years. The suggested form of the cost function penalties requiring the steadiness of the solution for the temperature equation, for example, involves the sum of the squares of the difference between the initial temperature and the temperature at several different times between the initial and final states.

The optimization approach is closely related to the robust diagnostic method of Sarmiento and Bryan (1982) when only temperature and salinity are varied in the optimization. The link between the two methods was explicitly demonstrated, and the solution obtained by combining them was used to examine the performance of the GCM used here for the North Atlantic Ocean. A relation was derived here between the weights in the cost function used in the optimization approach and the restoring coefficient of the robust diagnostic method. Using this relation, the robust diagnostic approach would make a useful method for obtaining the initial guess for starting an optimization based on the adjoint method.

Clearly, the new possibilities offered by the optimization approach and the adjoint method also reveal many new difficulties not arising in the context of simpler inverse models. The intrinsic nonlinearity of the calculation, and the new capabilities such as the possibility to calculate surface forcings and other model inputs, all give rise to many new and previously unknown obstacles that must be overcome. It is important to realize that practically all the difficulties reported here are related to the way the optimization problem is formulated and have nothing to do with the adjoint method for calculating the gradient of the cost function. The adjoint model used here, based on that developed by Long et al. (1989), was able to accurately and efficiently calculate the gradient of the cost function based on a complex GCM. The remaining challenges all have to do with the way the inverse problem based on the GCM is formulated (i.e., the specific form used for the cost function). Hopefully, the results reported here bring us closer to an understanding of how such inverse problems should be formulated. Considering, however, the short experience in oceanography with such complex optimization problems (as compared with the vast experience with numerical models, for example), the progress made so far is encouraging. We have tried in this work to gain some understanding of the reasons for the difficulties encountered and have tried to narrow the range of possible reasons for the difficulties. There seem to be good reasons to expect that many of the difficulties still encountered here will be corrected in future studies, helping to fulfill the method's potential in bridging the gap between numerical modeling and observational oceanography, using both models and data to form a unified picture of the oceanic general circulation.

Acknowledgments. We wish to thank Carl Wunsch and Charmaine King for helping with the initial data analysis. This research was supported by Grant No. 89-00408 from the United States—Israel Binational Science Foundation.

REFERENCES

Esbensen, S. K., and Y. Kushnir, 1981: The heat budget of the global ocean: An atlas based on estimates from surface marine observations. Oregon State University Climate Research Institute Rep. No. 29.

Gill, P. E., W. Murray, and M. H. Wright, 1981: *Practical Optimization.* Springer-Verlag, 377 pp.

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

Levitus, S., 1982: Climatological atlas of the world ocean. NOAA Tech. Paper, 173 pp.

——, 1989: Interpentadal variability of temperature and salinity at intermediate depth of the North Atlantic Ocean, 1970–1974 versus 1955–1959. *J. Geophys. Res.*, **94**(C5), 6091–6131.

Long, R. B., S.-M. Hwang, and W. C. Thacker, 1989: The finite-difference equations defining the GFDL–GCM and its adjoint. Unpublished report, Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida.

Marotzke, J., 1992: The role of integration time in determining a steady state through data assimilation. *J. Phys. Oceanogr.*, **22**, 1434–1457.

Mercier, H., 1989: A study of the time averaged circulation in the western North Atlantic by simultaneous nonlinear inversion of hydrographic and current meter data. *Deep-Sea Res.*, **36**(2), 297–313.

Olbers, D. J., M. Wenzel, and J. Willebrand, 1985: The inference of North Atlantic circulation Patterns from climatological hydrographic data. *Rev. Geophys.*, **23**, 313–356.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr.*, **12**, 1154–1158.

Rintoul, S. R., 1988: Mass heat and nutrient fluxes in the Atlantic Ocean determined by inverse methods. Ph.D. thesis, Massachusetts Institute of Technology–Woods Hole Oceanographic Institute. Joint Program in Oceanography.

Sarmiento, J. L., and K. Bryan, 1982: An ocean transport model for the North Atlantic. *J. Geophys. Res.*, **87,** 394–408.

Schlitzer, R., 1988: Modeling the nutrient and carbon cycles of the North Atlantic 1. circulation, mixing, coefficients, and heat fluxes. *J. Geophys. Res.*, **93**(C9), 10 699–10 723.

Schmitt, R. W., P. S. Bogden, and C. E. Dorman, 1989: Evaporation minus precipitation and density fluxes for the North Atlantic. *J. Phys. Oceanogr.*, **19,** 1208–1221.

Thacker, W. C., 1989: The role of the Hessian matrix in fitting models to measurements. *J. Geophys. Res.*, **94,** 6177–6196.

Tziperman, E., and A. Hecht, 1988: Circulation in the eastern Levantine Basin determined by inverse methods. *J. Phys. Oceanogr.*, **18,** 506–518.

——, and W. C. Thacker, 1989: An Optimal Control/Adjoint equations approach to studying the oceanic general circulation. *J. Phys. Oceanogr.*, **19,** 1471–1485.

——, ——, and K. Bryan, 1992: Computing the steady state oceanic circulation using an optimization approach. *Dyn. Atmos. Oceans,* **16,** 379–403.

——, ——, R. B. Long, and S.-M. Hwang, 1992: Oceanic data analysis using a general circulation model. Part I: Simulations. *J. Phys. Oceanogr.*, **22,**

Wunsch, C., 1978: The general circulation of the North Atlantic west of 50°W determined from inverse methods. *Rev. Geophys.*, **16,** 583–620.

——, 1984: An eclectic Atlantic Ocean Ocean circulation model. Part I: The meridional flux of heat. *J. Phys. Oceanogr.*, **14,** 1712–1733.