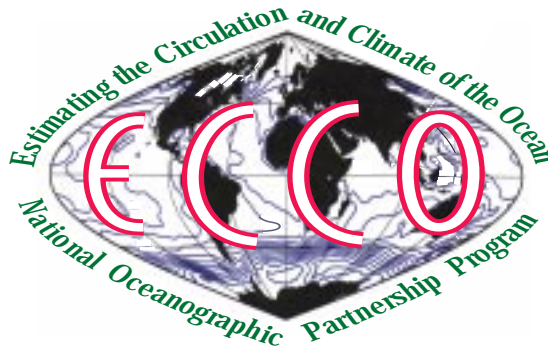


*The ECCO Report Series*¹

Estimating the North Atlantic Circulation With Nesting and Open Boundary Conditions Using an Adjoint Model.

Nadia Ayoub², D. Stammer³ and C. Wunsch⁴,



Report Number 10

October, 2001.

¹The ECCO Project is funded through a grant from the National Oceanographic Partnership Program (NOPP).

²Laboratoire d'Etudes en Geophysique et Oceanographie Spatiales/ Centre National de la Recherche Scientifique, Toulouse, FRANCE

³Copies of this Report are available at www.ecco-group.org or from Detlef Stammer, Scripps Institution of Oceanography, La Jolla CA 92093-0230, USA, ph.: (858) 822-3376; fax: (858) 534-4464; e-mail: dstammer@ucsd.edu

⁴Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Abstract

A $1^\circ \times 1^\circ$ resolution version of the MIT-GCM in the North Atlantic is used to test whether open-boundary conditions can be constrained by observations inside the domain interior using an adjoint method. The model is run over the year 1993 and with a simplified vertical mixing physics. This feasibility study is performed as a preliminary step towards an experiment over 6 years and with a version of the MIT-GCM including a full surface boundary model. The data used to constrain the model consist in monthly SST fields from Reynolds and Smith (1994), monthly climatological θ, S fields of Levitus et al. (1994) and TOPEX/POSEIDON altimetry. For this latter, we use separately the mean sea surface height and the anomalies with respect to the mean. The adjoint model is built using an automatic differentiation software: the TAMC of Giering and Kaminski (1998). The method aims at bringing the model's trajectory to consistency with the data, by adjusting the initial θ, S fields, the four components of the atmospheric forcing fields (heat and fresh water fluxes and the zonal and meridional wind stress components) as well as the θ, S and horizontal velocity fields prescribed at the open boundaries. For the surface and open-boundary conditions the adjustments are computed every 10 days and every month respectively. An originality of the work is the so-called 'nested approach', that is the use of optimized fields from a global, coarser resolution model for the open-boundary conditions and for the prior estimates of the adjustments on the surface conditions. A solution is obtained after 75 iterations. This study shows that significant changes can be obtained on the open-boundary conditions, and that a general improvement on the circulation is achieved in the constrained solution. This is particularly true in the Gulf Stream and equatorial regions. Changes at the open boundaries are difficult to interpret. They are mainly characterized by a large temporal variability and by the presence of small scales. Locally, large adjustments are found close to the bottom and are probably unrealistic. There, the method probably tends to compensate for some of the model's deficiencies by computing large corrections on the open-boundary conditions.

1 Introduction

Knowledge of the time-varying ocean circulation, of the associated meridional heat and mass transports variability, and subsequently of ocean/atmosphere interactions, are fundamental to understanding of climate evolution. But the turbulent nature of oceanic flows and the forcing of the circulation by many simultaneous factors—leading to ocean responses with a large range of spatial and temporal scales—make very complex both the interpretation of observations and the simulation of the variability with numerical models. Despite an increasing volume of observations and significant improvements in models, many problems remain.

Neither observations nor models provide a complete or accurate depiction of the ocean. Numerical models produce descriptions of the time variable ocean physical state encompassing a broad range of scales. Their realism however, is limited by model errors owing to the missing physics, the discretization in space and time, the lack of adequate resolution, and uncertainties in the boundary and initial conditions. Data remain comparatively sparse and always contain a variety of observational errors.

A picture constructed by appropriate combination of both model and data will produce a better understanding than what emerges from either alone. This present study is part of a broader attempt (Stammer et al., 2001a,b) to produce quantitatively useful understanding of the ocean circulation by state estimation methods (“assimilation”). But the global focus of this work has thus far precluded obtaining adequate spatial resolution in the general circulation model (GCM) being used. By restricting the computations to the North Atlantic, we can in principle run models with much higher resolution than is possible globally today, thus addressing one (but not all) of the major issues of oceanic state estimation. Regional models however, inevitably raise fundamental issues connected with “open-boundaries” and whose solution is far from obvious.

The methodology and overall goals are described in Stammer et al. (2001a,b) for the global problem. Our ultimate goal is to obtain an estimate of the North Atlantic circulation over the years 1992-97 at high resolution. As a step towards that goal, the focus here is on specific questions regarding the open-boundary problem. Because of its preliminary

nature, the computations described here are confined to a one year experiment in 1993. Use of a single year greatly reduces the computational load, while still raising all of the technical issues of the longer interval. The immediate goal is not to present a state-of-the-art estimate of the North Atlantic circulation; instead, we use a simplified, although still complex, version of the MIT-GCM to investigate

(1) whether a regional open-boundary model can be constrained by many different data types, using an estimation method employing Lagrange multipliers (the “adjoint method”), and

(2) by using the results of the global, lower-resolution, simulation as a first-guess for the open boundary conditions and the interior solution to demonstrate the use of a multi-resolution approach.

2 Model Description

2.1 Basic Structure

We use the MIT primitive equation model (Marshall et al., 1997), hereafter referred to as the MIT-GCM. The model as run uses the incompressible Navier-Stokes equations, within the Boussinesq and hydrostatic approximations. This version has a semi-implicit free surface formulation and does not yet include an ice-model. Atmospheric forcing consists of daily heat and fresh water fluxes, and twice-daily zonal and meridional wind stress components from the NCEP reanalysis. No restoring terms to climatological temperature and salinity fields are used at the surface. The model is integrated on a $1^\circ \times 1^\circ$ horizontal grid and 23 vertical levels, with the first six levels in the upper 100 meters. Horizontal mixing of momentum is represented by a biharmonic term with a coefficient varying as the cosine of latitude to account for the decreasing distance between meridians. Horizontal dissipation of temperature, θ , and salinity, S , occurs through Laplacian terms with constant coefficients. Because of complications inherent in the computation of the adjoint model in the presence of complete surface boundary/mixed-layer physics, the MIT-GCM is run here with constant vertical mixing coefficients for θ , S and momentum through the whole water column. Convective adjustment is permitted once per day to reduce grav-

itational instabilities. Besides the relatively coarse grid resolution, the lack of a surface boundary layer model probably constitutes the main limitation of the present work.

The model is integrated from 1 January 1992 to 31 December 1992, starting from rest, and the Levitus et al. (1994) temperature and salinity climatology for January. Data from 1993 are then employed to obtain a constrained-model estimate of the ocean state.

2.2 Open-Boundary Conditions

The domain of integration extends from 100°W to 15°E and from 25°S to 80°N (Fig.1). Communication with the remaining part of the global ocean exists via open boundary conditions: with the South Atlantic at 25°S , with the Arctic Ocean through the Fram Strait, and with the Barents Sea at 15°E between Spitsbergen and Scandinavia. The Mediterranean basin is not included in the model domain, so that the Gibraltar Strait is another open boundary. The Labrador Sea is closed at its northern limit. At this closed boundary, θ and S are relaxed toward the Levitus et al. (1994) climatological seasonal fields over the water column and over 10 adjacent horizontal grid points inside the domain interior.

Our open-boundary (OB) scheme is inspired by the method of C. Herbaut (Herbaut, pers. comm., 1998). A similar open-boundary scheme is used by Zhang and Marotzke (1999). These authors study the steady circulation of the Indian Ocean using the MIT-GCM and the adjoint method; their OB conditions are optimized as well. However, both their numerical implementation of the OB scheme and the context of their optimization problem are different from ours, leading our respective studies to address distinct issues. In particular, the time-dependent circulation makes our problem more complicate. We also constrain the free-surface model with TOPEX/POSEIDON sea level data whereas Zhang and Marotzke (1999) used a rigid-lid and no altimetry.

This OB method is relatively simple, and relies on the advection of θ , S and horizontal velocity (U, V) from prescribed values at the boundaries. These latter are specified using the results of the longer-duration global estimate at coarser resolution. The method was employed primarily for its simplicity, and is regarded as an interim procedure.

The OB scheme allows the inward and outward advection of θ, S, U, V in the direction

perpendicular to the boundaries. In the momentum equation, velocity is restored to prescribed values at the two closest points to the boundaries. The restoring coefficients are $(1 \text{ day})^{-1}$ at the closest point to the boundary and $(5 \text{ days})^{-1}$ at the furthest. There is, in principle, no need to distinguish the cases of inflow/outflow when the prescribed fields at the boundaries are constrained by observations in the domain, because the observations contain the signature of incoming and outgoing flows (Vogeler and Schröter, 1999). This OB scheme requires the use of a strong vertical viscosity ($10^{-2} \text{ m}^2/\text{s}$); presumed to be acting as a damping factor that allows the adjustment of the ocean state close to the boundaries to the conditions imposed at the boundaries. This relatively high viscosity can lead however to excessive mixing in the deep and intermediate layers of the domain interior.

To avoid a failure of volume conservation (recall this is a Boussinesq approximation model), we impose a net zero transport across each of the boundaries. One expects the TOPEX/POSEIDON sea surface height (SSH) constraints to be sufficient to force a realistic net transport in the basin. But at mid-latitudes, a flow of 8 Sv across a meridional section with a uniform depth of 4000 m corresponds to a SSH difference of only 2 cm along the section, and errors in the geoid employed with the altimetric data exceed this value, requiring the extra constraint.

3 Constraint Data

3.1 Description of the Data Sets

The model is constrained by TOPEX/POSEIDON altimetric data, climatological monthly hydrographic fields, and sea surface temperature (SST) observations. The conventional geophysical corrections, including tides and inverse barometer effect, have been applied to the TOPEX/POSEIDON data; with measurements in areas shallower than 1000 m having been eliminated. The absolute sea surface heights (SSH) are determined using the EGM96 geoid model (Lemoine et al., 1998). They are projected onto the model grid by averaging the along-track data over one day and $1^\circ \times 1^\circ$ boxes. The mean surface over 1993 is computed and SSH anomalies (SSHA) are then deduced by subtracting the mean

surface from the SSH. Both the mean and the SSHA data are used to constrain the model, but handled separately.

Hydrographic observations are represented by a gridded monthly climatology of θ and S (Levitus et al., 1994) and monthly average SST fields (Reynolds and Smith, 1994). Instantaneous hydrography and SST values could be used if the error covariance matrices (see the next section) were available.

3.2 Model/Observation Comparison

Before carrying out any assimilation, one must check the fundamental model/data consistency. (One should not average apples and oranges.) Figure 2 compares the observed and simulated mean SSH, averaged over the year 1993. The TOPEX/POSEIDON mean SSH is computed relative to the geoid model EGM96. Structures at scales higher than 2000 km are determined with a high precision, whereas smaller scales suffer from large inaccuracies. This structure is well illustrated in the Gulf Stream area where the observed mean SSH does not depict the sharp sea surface slope associated with the current. The fields in Fig. 2 are in a general good qualitative agreement, showing similarities in both the shape and amplitude of the large-scale structures. Simulated subtropical gyres are however weaker than in the data and display a smaller extension into the interior of the basins. Significant differences are also observed in the tropical basin. They may reflect a poor representation of the density field, especially in the east, where the modeled surface temperatures are too high and do not exhibit the seasonal appearance of the eastern equatorial upwelling.

Fig. 2

As a consequence of the lack of complete surface boundary layer physics in the model, the heat storage in the upper 150 m is relatively weak in the tropics and subtropics while the tropical SST can reach abnormally high values. In subpolar regions, the cooling is too strong and leads to unrealistic temperature values and convective patterns. Note also that the lack of ice formation prevents the subsurface waters from being ‘insulated’ from the atmospheric cooling. The excessively low temperature values create a large inconsistency between the model SST fields and the NCEP heat flux fields which contain an ‘ice signal’ characterized by very weak fluxes. In future experiments, the formulation of atmospheric

forcing a modified formulation of surface fluxes should prevent this form of inconsistency.

4 The Estimation Method

4.1 The Adjoint Method

Under the assumption that model-data differences are the result of errors in identifiable model parameters (including boundary and initial conditions, mixing coefficients, etc.), one can attempt to estimate values of these parameters which would bring the model into agreement, within error estimates, with the observations. Because both model and data are discrete, finite dimensional fields, the problem reduces to one of least-squares, constrained by the model.

We have chosen to use Lagrange multipliers in the so-called adjoint method as described in Wunsch (1996). Details of the application are given in Stammer et al. (2001a). One first defines a cost or objective function, J , measuring the model misfit to the data. The model is appended to J using the Lagrange multipliers, to form a new objective function J' . One seeks the stationary values of J' with respect to the ocean state, the adjustable parameters (“control variables”) and the Lagrange multipliers, $\boldsymbol{\mu}(t)$. Setting the respective partial derivatives of J' to zero results in the so-called normal equations which must be solved. In practice, the equations are solved implicitly (not written out) and iteratively, employing the adjoint compiler of Giering and Kaminski (1998) and a descent code of Gilbert and Lemaréchal (1989). A complete description of the use of the compiler with the MIT-GCM and the interpretation of the adjoint solution can be found in Marotzke et al.(1999).

4.2 The cost function

Let $\mathbf{x}(t)$ be the oceanic state vector (defined as exactly sufficient information—model velocities, pressure, temperature and salinity defined on the grid points, sufficient to take one time step forward if the boundary conditions are provided; see Wunsch, 1996). The

objective function used here is of the form:

$$\begin{aligned}
J = & \sum_{t_i}^{t_f} (\mathbf{u}(t) - \mathbf{u}^b(t))^T (\mathbf{Q}(t))^{-1} (\mathbf{u} - \mathbf{u}^b(t)) \\
& + \sum_{t_i}^{t_f} (\mathbf{E}(t)\mathbf{x}(t) - \mathbf{y}(t))^T (\mathbf{R}(t))^{-1} (\mathbf{E}(t)\mathbf{x}(t) - \mathbf{y}(t))
\end{aligned} \tag{1}$$

where $\mathbf{x}(t)$, $\mathbf{y}(t)$, $\mathbf{u}^b(t)$ are respectively the state vector, the observations and the first guess of the control variable $\mathbf{u}(t)$ at time t . \mathbf{E} is the observation matrix while \mathbf{Q} and \mathbf{R} are weight matrices. The summation is taken over the full time span of the data, $t_i \leq t \leq t_f$, here from 1 January to 31 December 1993. A more detailed description of J is found in the Appendix.

J has two types of element. Some terms represent penalties on changes to the control variables, used to modify the model trajectory through phase space to bring it closer to the observations; the second type of term represents the model-data difference that one is attempting to reduce. The computation trades decreasing the model-data difference terms for the expense of increasing the magnitudes of the variations in control terms. Here, control variables include initial conditions for temperature and salinity (θ_0, S_0), the four atmospheric forcing components (heat and fresh water fluxes, zonal and meridional wind stress components) for which an adjustment is computed every 10 days, and the monthly fields prescribed at the open boundaries (temperature, salinity and horizontal velocity, $\theta_{ob}, S_{ob}, U_{ob}, V_{ob}$) for which an adjustment is computed every month. The prior estimates are defined as follows: θ_0 and S_0 are the temperature and salinity fields for Dec. 31st 1992; H_Q, H_F, τ_x, τ_y are the NCEP reanalyses and $\theta_{ob}, S_{ob}, U_{ob}, V_{ob}$, the outputs of the global run (see Section 2-2).

4.3 Weight Matrices

As with any estimation method, specification of the weight matrices, \mathbf{Q} , \mathbf{R} , largely determines the solution. In principle, the \mathbf{Q} , \mathbf{R} represent all of the *a priori* information about the structure and amplitude of data errors and of errors in the background fields. ‘Data

error' here includes not only the instrumental and related noise, but also any physical properties (e.g., internal waves, mesoscale eddies) that the model cannot reproduce.

The weight matrices are often simply identified as the inverse error covariance matrices. Then the solution in the near-linear case becomes also the minimum variance statistical estimate (Wunsch, 1996). Normally, there exists only limited knowledge of the observational errors, but estimation theory implies using whatever information is available. Here the errors are supposed constant in time, with diagonal covariance matrices. The diagonal elements were chosen, initially, to be proportional to the θ, S error profiles provided by Levitus et al. (1994) (Fig. 3). At each boundary, a constant value is arbitrarily chosen for the uncertainty of U_{ob}, V_{ob} . This value has been changed throughout the iterations to help control the rate of descent. For the last 15 iterations, we took an error of 2 cm/s in the upper 185 m and of 1 cm/s below. For the atmospheric terms, the diagonal \mathbf{Q} elements are inversely proportional to the field variance during 1993. Comparisons between NCEP and ERS-1 wind products provided by CERSAT have been used to determine the proportionality coefficient for the wind stress error matrix. The geoid error variance provided by Lemoine et al. (1998) is used for \mathbf{R}_{geoid} . This choice relies on the assumption that the altimetric measurement errors do not have any significant mean component, and consequently that all the discrepancies between the simulated and observed mean SSH are attributed to uncertainties in the geoid estimate. For the SSH variability, the variance of the uncertainty of the observation together with the observability error is estimated to be 1/4 of the data point wise variance. Omitting the off-diagonal terms, as we have only an approximate knowledge of them, is consistent with the results of estimation theory which suggests that the consequence will be a greater uncertainty in the solution than is possible if the missing information could be provided, but without biasing the result.

Fig. 3

5 Strategy of Optimization

A major purpose of these calculations is to investigate the effectiveness of using the results of a coarse resolution model as an initial estimate of the regional, higher resolution solution, as opposed to initializing, e.g. from the flow implied by the climatology alone.

A first step towards this “nested approach” is made by prescribing the fields determined from the global model at the open boundaries. The next step is to apply to the control variables the corrections $d\mathbf{u}_G$ deduced from the global optimization.

This approach has been tested performing a set of experiments described in Table 1; Fig. 4 shows the model/data misfits for the different cases. We first start (iteration zero, written I0) with $d\mathbf{u} = 0$; after 24 iterations, we obtain the so-called ‘C24’ control run. The results in Fig. 4 are normalized by the values reached in C24. Comparing the results in I0 and C24 confirms that the method succeeds in reducing the model/data misfits, especially for the θ/S fields. For the sea surface height, the reduction is less significant. As explained later, the model is mainly constrained by the climatology during these first iterations. Further in the iterative process, we downweight the climatology influence and let the model be more constrained by altimetric observations. We then perform two experiments (N1 and N2) with $d\mathbf{u} = 0$ for the initial and open-boundary conditions and with $d\mathbf{u} = d\mathbf{u}_G$ for the heat and fresh water fluxes. For the wind stress, we take: $d\mathbf{u} = d\mathbf{u}_G$ in N1 and $d\mathbf{u} = 0$ in N2. After a single model integration, the misfits for θ, S are significantly reduced with respect to I0, which is especially obvious for the SST, in both N1 and N2. The SST misfits in N1 are even lower than in C24 (about 75% of the value in C24). This result suggests that, in the extreme case where the model is constrained by SST observations only, using the outputs of the global model is more efficient than performing 24 iterations. The model/data differences in terms of SSH anomalies are slightly increased in N1 with respect to I0, whereas they are unchanged in N2. We suspect that the corrections to the NCEP wind fields computed in the global case are not consistent with the higher resolution Atlantic configuration, and lead to the degradation of the SSH fields in N1.

A last experiment has been performed starting from iteration 23 of the control run (C23) and applying $d\mathbf{u} = d\mathbf{u}_{C23}$ for the initial and open-boundary conditions and $d\mathbf{u} = d\mathbf{u}_G$ for the atmospheric fields (including the wind). The only difference between this integration (hereafter called I24) and C24 is thus the origin of the adjustment $d\mathbf{u}$ to the atmospheric fluxes: from iteration C23 for C24 and from the global run for I24. The

Table 1
Fig. 4

model/data misfits in I24 are little changed with respect to C24 for θ, S (Fig. 4) but are reduced by half for the SST. In terms of SSH anomalies, they are about 10% higher in I24 than in C24, which is probably due to the wind adjustments as suggested above by the results of N1 and N2.

For the optimization described in this note, we have chosen to pursue the descent of J from I24 to keep the good performance obtained on the SST fields. Later in the iterative process, we increase the weights on the TOPEX/POSEIDON data constraint to force a faster descent of the cost function relative to the SSH anomalies.

In summary, the global model outputs as first guess for the regional optimization allows us to save many iterations for this latter thus significantly reducing the computational burden. This is a very encouraging result showing the benefit of the nested approach and the potential use of low to medium resolution simulations on large domains for high-resolution regional zoom where the optimization is expected to be very expensive in terms of computational costs. At the same time it shows the benefit of a down-scaling approach also on global scale - starting with coarse resolution and increasing resolution of the optimization problem once the larger scales were adjusted.

6 Statistical Assessment of the Solution

6.1 Objective Function Evolution

Figure 5 shows the evolution of J and its individual elements J_i with iteration number, and also the normalized value $J/(N + M)$ where N and M are the number of observations and of control variables respectively. In a fully consistent solution, J_i should be order 1, and the individual terms should have a frequency function approximating a χ^2 distribution. More general a posteriori test criteria exist (Wunsch, 1996; Talagrand, 1999), but our aim here is mainly exploratory, and we will not pursue definitive statistical tests.

Iterative minimization is stopped when J ceases to diminish and the distribution of the J_i is acceptable. In the present case, intermediate minima (found along the way) tend to be local rather than global, as the non-linearity of the system makes the local topology of J extremely complex. Descent methods are easily confused by the complex

shape of J which is a function of problem parameters. Following our own and others experience with non-linear optimization methods, we have in the course of 75 iterations, five times restarted the calculation after modifying the local terrain by changing the weight matrices \mathbf{Q}, \mathbf{R} within reasonable limits of our own uncertainty of their values. The changes of weight appear as jumps in the cost function evolution.

The cost function decreases significantly initially (Fig. 5). Between iterations 24 and 75, $J/(M + N)$ is reduced from 1.3 to about 0.4 (thus perhaps becoming too small). Further improvement is then slight, and the calculation was halted at iteration 75. Individual terms J_i associated with the altimetry, surface temperature and climatology vary considerably (as they should). Misfits between the model and the climatology are strongly reduced in the first iterations. Beyond iteration 36, the climatology misfit terms were downweighted, and they no longer decreased. We suspect however, that they are actually somewhat overfit. After a weight increase at iteration 24 and the use of heat flux corrections from the global run as part of the nested experiment, the misfit to the SST estimates decreases and converges quickly—resulting in J_i terms of order 1 beyond iteration 44. This result is encouraging because the unconstrained simulation was particularly poor in its representation of SST.

Misfit terms to altimetric anomalies do not show a monotonic decrease with iteration. The overall rate of decrease accelerates after the point, at iteration 36, when the climatology was downweighted, probably indicating an inconsistency between the two data types. At the end of the calculation, the relatively poor fits suggest that some TOPEX-POSEIDON data errors, in particular those reflecting the difference of spatial scales between the along-track SLA signals and the model resolution, have been underestimated. The mean SSH contributes only a small fraction to the total value of J , consistent with the large errors in the geoid estimate at small scales.

Changes in the initial conditions are greatest during the first iterations; after iteration 50, further decrease is slight. At iteration 75, $J_i/(M + N)$ is close to 0.3 and 0.15 for the temperature and salinity respectively. Rapid changes, initiated at iteration 30, cease at iteration 36 when the hydrographic climatology is downweighted.

6.2 Major Misfits of the Forward Model Alone

In the unconstrained forward run, there are some obvious model/data misfits. In particular, the use of constant vertical mixing parameters leads to systematic discrepancies between the modeled and observed temperature and salinity fields at all depths. The most drastic of these are in SST; without a proper surface mixed layer model, the simulation cannot adequately transfer heat downward from the sea surface. Misfits of more than 10°C are observed in summer in the subpolar regions, which may have severe consequences for convection during the next winter.

In the optimized calculation, SST has been artificially forced to consistency with the data by imposing a very high weight on the surface temperature boundary conditions north of 50°N . Model/data misfits are efficiently reduced (Fig. 5), especially in the western tropical Atlantic (boxes 6, 8, 9 and 12) and north of 50°N where the summertime excessive high SST are cooled down and the misfits remain lower than $\sim 2^{\circ}\text{C}$ (Fig. 6). Assuming that the variance of the residual SST misfits can be estimated by their temporal variance over the integration period, we have computed the ratio of this variance to the a priori error variance at each grid point. In an area roughly defined between 50°N and 70°N and west of 60°W and in the western tropical region, the ratio is close to 1, indicating that the solution is consistent with the a priori errors in the SST boundary condition. North of 70°N , the ratio remains higher than 2.5, despite the considerable reduction of the misfits at iteration 75. In the rest of the basin, the ratio is much lower than 1, suggesting an overestimate of the a priori SST error. In the attempt to bring the model to consistency with the observed SST while the mixing physics are incomplete, we take the risk of generating unrealistic stratification. We have compared the upper layer temperature content in the constrained and unconstrained models and in the climatology. Except in the eastern equatorial Atlantic and in the Gulf Stream area (box 24), the thermal content in the upper 260 m in the constrained model is not significantly changed with respect to the unconstrained one in mid and low latitudes. In high latitudes, the constrained model becomes closer to the observations on average. These corrections might eventually have a strong impact on the deep water mass formation rates, but the present simulation is too

Fig. 6

short to observe it.

We have also computed the ratio of the variance of the model/data misfits for the SSH anomalies by the a priori error variance. This ratio is globally decreased in the whole basin. North of 40°N , it remains close to 4, mainly because of the low variability of modeled SSH compared to the observations (see Section 3-2). The ratio is significantly decreased in the equatorial and southern regions and in an area lying roughly between 40°W - 20°W and 10°N - 20°N . There, ratios locally higher than 7-8 are reduced to less than 3. As described later, such large changes affect the mean circulation there, down to ~ 1200 m, linking the tropics to the northern subtropical branches in the constrained model.

The variance of the model/data misfits in temperature and salinity is well below the a priori error variance all over the basin. Because this misfit is mostly due to our deliberate downweighting of the constraints relative to the climatology, we will not discuss it further.

6.3 Adjustments to the Control Variables

Control variables are modified to bring the forward model trajectory through state space into closer agreement with observations. Changes in the initial and surface boundary conditions are reported in section 7 along with the description of the optimized state. Figures 7 and 8 show the mean adjustments to the θ, S initial profiles, and the atmospheric fields respectively. A striking characteristic of these corrections is the presence of small scales. The only constraint able to induce such scales is the daily altimetric constraint, because the climatological and Reynolds SST fields are relatively smooth. The adjoint method is therefore able to transfer the small-scale variability from its signature in SSH to wind, temperature and salinity fields. Real atmospheric and oceanic fields do contain variability at such scales. The question arising here is whether a solution containing small scale features with this medium-resolution model configuration is acceptable? In this particular context, we consider that as long as the medium and large scales are well simulated (see Section 7) and the level of small-scale variability is comparable to that contained in the altimetric observations (see Section 7-1), the solution is acceptable. This issue should be explored in future experiments, by prescribing non-diagonal weighting matrices for the data constraints and the controls.

Fig. 7

Fig. 8

We focus on the role of the controls at the open boundaries. Discussion of the changes is greatly complicated by the suspected regional inhomogeneities in the hydrographic climatology so that apparently significant deviations of the model results from the climatology may only represent a failure to adequately describe the error fields of the latter.

Adjustments at the open boundaries are characterized by large changes at small scales (e.g. between two adjacent grid points) and from month to month. The ability of the system to estimate boundary values is limited by the time required for information to propagate to or from the interior to the boundary. Thus, changes in the OB initial estimates are very weak in the first and last two months of the optimization—as the information has had inadequate time to connect with the interior fields. In general, adjustments in the deep layers are as large as at the surface. The mean changes over the year and their standard deviation are shown on Figs. 9 and 10. We have also computed for our analysis the ratio between the mean (in absolute value) adjustment and the prescribed error (not shown).

Fig. 9
Fig. 10

Changes in the OB conditions results in strong vertical velocities at the first points inside the integration domain (except near Gibraltar). At the southern boundary, vertical velocities are less than 1×10^{-5} m/s in the first 1500 m and are then intensified at depth (up to 24×10^{-5} m/s). Close to 78°N , they reach 10^{-4} m/s below 500 m. The largest values are observed close to the boundary at 15°E below 115 m, with a maximum downward velocity of 124×10^{-5} m/s. The combined effect of the strong viscosity and of the large vertical velocity makes the interpretation of the changes at the boundaries a bit tricky. In particular, some changes in the temperature and salinity fields (e.g. at the northern and eastern boundaries) may lead to unrealistic fields, whereas the first points inside the integration domain and adjacent to the boundaries are characterized by visually realistic profiles, globally closer to the climatology than are the initial estimates.

Southern boundary.

At iteration 75, the mean model-data misfits are significantly decreased at the first grid points inside the domain, as illustrated in Fig. 11. When diagnosed as monthly averages in the boxes of Fig. 2, they suggest that both θ, S are relatively well adjusted in the first ~ 200 m. In the central part of the basin (boxes 2 and 5), the salinity field is

Fig. 11

modified between 1000 and 2500 m and departs more from the climatology than in the non-constrained run.

The changes in the meridional velocity are believed to have a significant impact on the meridional overturning streamfunction by inducing an inflow of cold water at the bottom which then upwells at 20°S below 3700 m (see Table 2). Having water penetrate primarily in the eastern basin is not consistent with most of the literature which places Antarctic Bottom Water primarily west of the Mid-Atlantic Ridge. If this feature should persist in longer runs, it could be corrected by adding a preventive constraint. Table 2

Northern and eastern boundaries

The Nordic Seas are the most challenging region for open-boundary estimation because the low resolution relative to the deformation radii leads to inaccurate representation of the circulation, of the formation of deep waters in the central cyclonic gyre, and of the overflow of these dense waters through the Greenland-Iceland-Shetlands straits. The lack of an adequate ice model is also critical, among other problems.

The changes to the prescribed velocity and θ, S fields at the northern and eastern boundaries are larger than at 24.5°S. At the northern boundary, they are significant in three main regions. The first one is located on the continental shelf off the Greenland coast. In the first 50 meters, the East Greenland Current (EGC) is accelerated (+ 2cm/s in average), carrying southward a cooler (-2°C) and less saline (down to -0.9 in salinity) water mass. Between 50 and ~300 m, these corrections have reversed sign and display weaker amplitude. The second region is in the central part of the section (-5°W / +5°W) at depths between 50 and ~500 m. There, the meridional velocity is accelerated northward (+ 2 cm/s), the salinity increased (+0.2) and the temperature diminished (-1.5°C). Changes on opposite signs are created on θ and S below 600 m; the adjustment on the velocity changes sign at ~1000 m with a northward acceleration at the bottom (~+ 2cm/s). These adjustments display a large time variability in particular at the surface (standard deviations of 8°C and 1.5 for the changes in θ and S) and below 600 m (3 cm/s for the meridional velocity). Note that in the EGC area, the adjustments to S are characterized by a seasonal signal with a decrease of the salinity in winter and an increase in summer. The ratios between the mean corrections and the a priori error suggest that

the corrections are too large at the surface for θ and at depth for S and the velocity.

Changes in the zonal velocity at the eastern boundary show large values at the surface (+4 cm/s in average) and below 200 m, and correspond mainly to an increased outflow. Close to the bottom, the maximum mean change reaches 14 cm/s, with a standard deviation over the year higher than 8 cm/s. At the first relaxation point inside the domain, the meridional velocity display significant changes (> 2 cm/s) at depth, with a northward acceleration in the south and a southward one in the north. Adjustments to S are larger than 0.5, usually an increase, except at the northernmost part of the section. Changes in θ are maximal (3.7°C) at the surface and along the Norwegian continental slope. Characteristic changes are on small scales, without any apparent structure in the nearby regional circulation.

In terms of model-data misfits, there is a global improvement, but some misfits increase. When averaged in the boxes of Fig. 2, the misfits suggest a better representation of θ, S at the surface and subsurface in the EGC region, of the temperature field in the first 1500 m in the east (box 35) and above 400 m along the Norwegian coast (box 33) but also a quasi-systematic degradation of the salinity representation below 1500-2000 m in the whole region. Interestingly, although the adjustments can make the θ, S profiles at the boundary unrealistic (e.g. with an unstable stratification on θ), the fields in the immediate vicinity of the boundaries can be closer to the climatology. The large differences between the profiles at the boundaries and at the first grid points inside the domain are likely due to the strong vertical mixing and to the effect of horizontal advection.

The main change in the surface circulation in the Nordic Seas is the formation of a cyclonic gyre in the Norwegian basin (Fig. 12). Such a gyre is represented in the scheme of Poulain et al. (1996) and may be real. The EGC is accelerated and extended eastward. The maximum modeled velocity is about 8 cm/s whereas observations (Woodgate et al., 1999; Foldvik et al., 1988) indicate amplitudes as large as 14 cm/s. The intensification is limited to a few points close to the boundary. It is likely that with a longer integration period, the information at the northern boundary would lead to an improvement downstream, further south. However, the lack of resolution in the model would likely continue to prevent it from producing a wholly realistic estimate.

Fig. 12

Strait of Gibraltar

The circulation near Gibraltar consists of a relatively fresh and warm layer of Atlantic waters flowing into the Mediterranean Sea overlying an outflow of dense Mediterranean Water (MW). The mixing of the MW as it passes the strait's sill at 320 m deep and its spreading into the Atlantic down to a depth of ~ 1000 m proceeds from a complex dynamics that we do not expect to resolve. The Strait has been artificially increased to two grid points, leading to an opening of 175 km in the model bathymetry, much wider than in reality (12 km at the narrowest section).

Adjustments to the prescribed fields at the Straits are maximum in the bottom three layers and at the surface (S only). On the annual average, the deep changes correspond to an intensification of the inflow at the northern point of the Strait, carrying into the Atlantic a less saline and warmer water mass. The inflow at the southern point is instead weakened; the asymmetry between the changes at the northern and southern points of the strait is well marked as well in θ and S . The deep and surface adjustments display a large temporal variability: this latter reaches, for example, ~ 3 cm/s at the bottom for the zonal velocity.

The prescribed flow field at the boundary, obtained from the global output, is characterized by an inflow in the first layer (and an outflow in subsurface) throughout the year, whereas one would expect an outflow in all the surface layers. The adjustment tends to weaken the surface inflow in summertime but intensify it in wintertime. It is very likely that the first guess is so far from a realistic state that the method cannot lead to any improvement.

East of $\sim 20^\circ\text{W}$ and between 30 and 45°N , the changes between the annual mean velocity in constrained and unconstrained model are weak (amplitude lower than 2 cm/s) but significant. They are mainly located off the Portuguese coasts and in an area roughly comprised between 15 and 20°W and 33 - 37°N . In this area, they consist of an intensification of the Azores Current in subsurface flow down to ~ 300 m, and in a northward acceleration below 600 m in the constrained model. Off Portugal, the flow is intensified southward in the first 200 m and northward at intermediate depths (400-900 m). Below 900 m, the solution displays a reversed flow as compared to the unconstrained circulation

along the Portuguese coasts, with a weak southward current.

In the Gulf of Cadiz, the misfits between the model and the climatological θ/S fields are not necessarily reduced. However, as already noted for the situation in the Nordic Seas, at greater distances from Gibraltar, but still in areas possibly influenced by the open-boundary conditions, the model θ, S fields move locally closer to the climatology. This adjustment occurs along the Portuguese coast in the salinity field between the surface and ~ 1200 m, and in temperature between ~ 120 and 1500 m (below 1500 m no significant differences are observed between constrained and unconstrained solutions). The changes are closely linked to those on the circulation above described. It is unclear though, to what extent they are due to the changes of the general circulation in the Northeast Atlantic and more specifically to those at the open-boundary conditions at Gibraltar.

7 The Optimized State

7.1 Small scale Variability

One of the most obvious changes between I0 and I75 is the introduction in I75 of small-scale variability, as illustrated on the maps of the SSH standard deviation (Fig. 13). Although the simulated variability is still lower than in the observed one, its spatial structures are in better agreement with the observations in I75 than in I0. At the very beginning of the simulation, the mean kinetic energy (KE) is 8 times higher in I75 than in I0; it decreases quickly and, by the end of January it is only 2 times higher. From April, we can consider KE as stabilized around a mean level; this latter is still higher than in I0 partly because of an enhancement of the small scale variability. The relatively high variability during the first months can be explained by initialization processes and by the 'introduction' of small scales through the initial θ, S fields (see Section 6-3). The variability decrease in time indicates that the model has neither the resolution nor the physics to maintain the scales contained in the initial conditions.

7.2 Meridional Overturning Cell (MOC)

In both I0 and I75 simulations, the meridional streamfunction is underestimated, as evidenced by the low values of the Florida Strait transport (10.1 ± 0.9 Sv in I0 and 9.1 ± 1.6 Sv in I75) compared to the estimates found in the literature (30 ± 2 Sv as reported by Rintoul and Wunsch, 1991). The MOC underestimation is mainly due to the relative poor resolution and mixing physics (in particular to the high vertical viscosity). The amount of deep/intermediate water masses north of 40°N is unchanged between I0 and I75 (about 14 Sv) and is in a very good agreement with the 15 ± 2 Sv of NADW production estimated by Ganachaud and Wunsch (2000). However, the 14 Sv cell observed at $50\text{-}30^\circ\text{N}$ and 1000-1500 m in I0 is reduced in I75. Following Böning et al. (1995), we suggest this weakening is due to the intensification of upwelling along the US coasts between 100 m and 2500 m.

Large changes between I0 and I75 are observed close to the southern boundary where the northward transport in the upper 1000 m is reduced from ~ 16 Sv in I0 to 10 Sv in I75. Below 4000 m, an inflow of 2-4 Sv is observed in I75, resulting from large deep adjustments on the meridional velocity at the open boundary (see Section 6-3). Connections between the tropics and the northern subtropical gyre differ also significantly. The I75 solution is characterized by an extended vertical cell (~ 16 Sv) between 20°N and 10°S associated to a significant deepening of the southward flow between 30 and 10°N and 2000-3500 m and the development of intense zonal flow at 20°N and at the equator in intermediate layers (~ 1200 m). The extent to which these patterns are realistic is not clear; neither is the change on the control variables that leads to them.

7.3 Meridional Heat Transport

Monthly temperature and velocity fields are used to compute the meridional heat transport (Fig. 14). Changes between I0 and I75 consist in a reduction and increase of the transports respectively south and north of the equator. South of 5°S , the transport in I75 is lower than in I0 by ~ 0.2 PW and departs more from the box model calculations of MacDonald (1998) and Ganachaud and Wunsch (2000). At 23°S , however, it is in a very good

Fig. 14

agreement with the result of Macdonald (1998). Between 5 and 25°N and at about 45°N, the transport increases of ~ 0.1 PW between I0 and I75; at 13°N, it reaches a maximum value of 1.07 PW in I75. This improvement is likely due to a more realistic surface thermal field in I75. At 7°N and 48°N, the I75 values fall within the range of the Ganachaud and Wunsch (2000) and Macdonald (1998) uncertainties respectively. The heat transports in both I0 and I75 are still weaker than the estimates from observations—partly because of the weakness of the simulated MOC.

7.4 Western Subtropical Gyre and Gulf Stream

The biggest change between I0 and I75 is observed in the Gulf Stream (hereafter GS) area and in the Newfoundland Basin. South of 40°N, the Gulf Stream is locally accelerated at the surface in I75 with respect to I0 (Fig. 15). Its maximum velocity (~ 50 cm/s) is still lower than values reported in the literature (e.g. Macdonald, 1998, finds a maximum velocity of 1.77 m/s). Below 100 m, the flow is weaker in I75 than in I0 and its vertical extension slightly reduced. The GS separation point from the coast is located more to the south in I75 than in I0. At 38°N for example, 5 Sv only are flowing northward in I75, while 15 Sv are still flowing alongshore in I0. Also obvious in Fig. 15 is the strengthening down to ~ 800 m of the GS extension to the east: at 72°W and 50°W, the core of the mean flow transports eastward respectively 34.9 Sv and 34 Sv in I75 vs. 29.5 Sv and 28.6 Sv in I0. Despite some improvements in I75 these transports remain much lower than those reported in the literature as for example, by Joyce et al. (1986): 125 ± 6 Sv at 72°W.

A new connection pattern between the subpolar and subtropical gyres is observed in I75, with, in particular, the development of a southward alongshore current between ~ 50 and 40°N, down to ~ 100 m and throughout the year. It is fed by shallow waters of the Labrador Sea and carries relatively cold and fresh water ($\theta < 12\text{-}15^\circ\text{C}$, $S < 35$) southward, enhancing the density front with the warm and salty subtropical waters ($\theta \sim 20^\circ\text{C}$, $S \sim 36$) of the GS. As illustrated on Fig. 16 and 17, water masses in the slope water area (box 24) in I75 have θ/S characteristics much closer to the climatology. The circulation changes lead to differences in the mean SSH between I75 and I0 as high as 20 cm there (Fig.

Fig. 15

Fig. 16

Fig. 17

18). Since T/P mean SSH does not provide accurate information on the short-spatial scale sea level slope linked to the Gulf Stream (see Section 3-2), we deduce that the mean surface there is constrained by hydrographic data. Consequently, discrepancies between the model and T/P mean SSH increase locally between I0 and I75. Fig. 18

The adjustment to the initial temperature consists in a cooling of 1 to 3°C in the slope water area from the surface down to 600 m. The temperature is increased by 1°C on average in the whole subtropical gyre and by 1 to 3°C in the intergyre zone (i.e. centered at 40°W - 50°N), in the upper 300 m. Changes to the initial salinity exhibit similar patterns with an increase of salt in the gyre, enhanced in the intergyre region (~ 0.3), and with a decrease of 0.4 to 1 in the first ~ 100 m in the slope water area. North of 55°N, the temperature changes in surface reach more than 10°C locally in the Nordic Seas and south of Greenland, and remain on average higher than 4°C. Such adjustments are linked to our choice of a strong weight on the SST constraint. They lead to unrealistic surface initial conditions in temperature in this area, but because the model/data misfits are then reduced during the integration period, we retain this solution. Adjustments to the heat and fresh water fluxes (Fig. 8ab) are consistent with those to the initial surface θ and S . They undergo a significant seasonal variability. On average, they tend to correct the location of the thermal and salinity front associated with the mean path of the GS. The heat loss along the GS path and intergyre region is reduced by 40 Wm^{-2} on average, whereas it is increased by 20 to 30 Wm^{-2} in the slope water area. Changes to the E-P balance indicate a flux of fresh water in the slope water area (0.8 cm/yr on average) and an increase of the evaporation rate in the Newfoundland basin (up to 0.3 cm/yr). The annual mean adjustments to the wind stress reach locally 0.02 Nm^{-2} for both components. Following Ezer and Mellor (1992), we suggest that the presence of an alongshore flow of slope water in the I75 simulation is responsible for the southward shift of the GS separation from the coast. We interpret the overall changes to the control variables as the following: the formation of slope water is induced by the changes on the initial θ/S conditions, while its density structure is maintained by the heat and fresh water atmospheric fluxes.

7.5 Equatorial and Tropical Regions

As a consequence of the poor mixing physics, the simulated temperature profile departs significantly from the climatology with maximum (negative) misfits at ~ 50 m (-1°C to -4°C in both I0 and I75). In I75, the thermocline is shallower than in I0 in the eastern basin, which results in the steepening of its zonal slope and to a more realistic representation of this slope. Because subsurface waters suffer from a heat deficit, the raising of the thermocline leads to an increase of model/data misfits in the eastern basin around 50 m, from -1°C in I0 to -3°C (box 12). In the western and central regions (boxes 7-11), the model is slightly cooled between ~ 100 and 250 m, leading there to a better agreement with the climatology. A major improvement in I75 is the appearance of a seasonal cycle in the upper 200 m all along the equator. As in the climatology, the seasonal cycle is stronger in the east, with the development of upwelling from April to September (Fig. 19), but it is not always in a good agreement with the data (wrong phase or wrong amplitude). At 5°S , the optimization leads to a seasonal signal below 30 m and east of 20°W that is not observed in the Levitus climatology. The appearance in I75 of the upwelling results in a drastic decrease of the model/data SST misfits in the eastern basin (boxes 9-12). For example, some local misfits reaching $13\text{-}14^\circ\text{C}$ in fall are reduced to 6°C . Off the African coast, at about 10°S , the SST is cooled by $\sim 5^\circ\text{C}$ with respect to I0 throughout the year.

Fig. 19

The discrepancies between the modeled and observed SSH anomalies are significantly reduced from I0 to I75 over the whole tropical area and to the largest extent, in the eastern basin between June and August. Fig. 20 compares the SSH anomalies fields in March and August from I0, I75 and T/P. The seasonal cycle in I75 is in good agreement with the observed one, with, in particular, a sea level decrease (more than 10 cm) between March and August in the eastern basin. This signal is consistent with the development of the equatorial upwelling discussed earlier. Mean sea level has also dropped by ~ 5 cm in this area (Fig. 18), bringing the model mean SSH closer in I75 to the T/P observations than in I0.

Fig. 20

The initial θ field is strongly cooled (3°C on average) in the eastern basin in the upper 65 m. Adjustments to the with a westward extension and a slight intensification

from May to September. They are partly responsible for the decrease of the SST in the eastern basin. Changes to the wind stress result in an intensification of the trades in the 5°N - 5°S latitude band, from April to August. As a consequence, the Ekman divergence along the equator is expected to be amplified, which can explain the development in the spring-summer period of the equatorial upwelling in I75.

The representation of the low salinity cores associated to the African and South American rivers discharge is significantly improved in I75 (e.g. in the Amazon area). Because river discharges have a shallow structure (< 50 m in the climatology), their signal in the simulation can be forced by the atmospheric fresh water flux. In the mean adjustment to NCEP fluxes, there is indeed an input of fresh water of more than 40 cm/yr (Fig. 8). In future studies, the runoff will be included in the freshwater budget. In the western basin, the model/data misfits for salinity reach a local maximum between 100 and 200 m. At I75, the misfits are reduced, and almost set to zero (box 10). This change in salinity appears related to a slight improvement to the subsurface North Brazilian Current (NBC) simulation in I75 (see below).

Currents

The model resolution is much lower than one would expect necessary to reproduce the equatorial circulation in detail. During the first four months, an intense zonal flow develops along the equator from the surface to the bottom. Between the surface and ~ 450 m, it appears as an eastward jet, where one would expect the northern branch of the westward-going South Equatorial Current. This jet is present in I0 and is intensified in I75 with a maximum amplitude of ~ 30 cm/s in the first 65 m. Below 450 m, the I75 flow shows a wavelike vertical structure. We suggest that this zonal circulation results from initialization processes. Between April and June, the westward surface flow is strongly accelerated between I0 and I75, as an indication of the intensification of the northern and southern branches of the SEC, owing to the strengthening of the trades in I75 in early summer (Fig. 21a,d). From September to November, the North Equatorial Counter Current is intensified eastward between I0 and I75, as is the circulation west of 20°W at $\sim 1^{\circ}\text{N}$ (Fig. 21b,c,e,f), probably as the signature of the strengthening of the Guinean Current (although this latter is expected with a more limited westward extension) result-

Fig. 21

ing from the weakening of the trades in the Gulf of Guinea. From August, a subsurface eastward flow appears between 65 and 300 m (Fig. 21f). We identify it as the Equatorial Undercurrent (EUC) although its vertical extension is too deep and its amplitude very weak (less than 15 cm/s) compared to expected values (e.g., Blanke and Delecluse, 1993). Besides the poor representation of the thermocline depth, the use of a large vertical viscosity coefficient prevents the maintenance of large vertical velocity shears. The EUC is slightly accelerated between I0 and I75 east of 25°W, as a result from the steepening of the thermocline zonal slope mentioned earlier.

In I75, the NBC is diminished with respect to I0 south of the equator, with a maximum decrease at $\sim 5^\circ\text{S}$ below 100 m (~ 20 cm/s). North of the equator, the main change consists in the subsurface intensification from May of the northward flow along the Brazilian coast. As a result, the path of the NBC from the equator to the Caribbean is continuous in I75, while, in I0, the NBC retroflects completely between 5 and 10°N. At 44°W, the westward velocity reaches its maximum (38 cm/s on annual average) at the surface, while Schott et al. (1998) observe amplitudes as large as 1 m/s at ~ 100 m. At 44°W and between the surface and 1000 m, these authors find a westward transport of 34.6 Sv. The subsurface flow is enhanced in I75 with respect to I0, which leads to a more realistic value of the annual transport as compared with the estimate of Schott et al. (1998): 21.3 ± 5.3 Sv at I75 vs. 18 ± 5 Sv at I0.

8 Comparisons with Independent Data Sets

8.1 Comparison with XBT Data

XBT profiles over 1993 are used to test the reality of the simulation of the subsurface thermal field and its variability over 1993. For the comparison, we use 5-day averages of the model temperature field and interpolate them to the XBT data points. The model/data misfits are then averaged over five vertical layers (surface, 10-55 m, 55-135 m, 135-510 m and 510-1750 m) and four latitudinal bands. Figure 22 shows the global distribution of Fig. 22 the profiles over 1993.

In terms of SST, we observe a net improvement in I75 with respect to I0. The standard

deviation of the model/data differences is reduced by more than 40% at surface between I0 and I75. In the eastern equatorial region, large surface misfits remains in I75 due to the poor representation of the upwelling; they are, however reduced from 10-15°C (in absolute value) in I0 to $\sim 5^\circ\text{C}$ in I75. In the 10-135 m depth ranges, the misfits are also significantly decreased, confirming that applying a strong constraint to the model SST does not have any negative impact on the subsurface thermal field, except in the eastern equatorial region. There, the raising of the thermocline in I75 leads to an enhanced subsurface heat deficit (see Section 7-5) and thus to an increase of the model/data differences. North of 55°N , the improvement in I75 is particularly remarkable in summertime where misfits ranging from 2 to 15°C in the first 135 m in I0 are reduced to values lower than 4°C on average. In box 24, where large changes on the circulation are observed between I0 and I75, the model/XBT data misfits are strongly reduced between I0 and I75, from the surface to the deeper layer (Fig. 23). Below 135 m, the changes between I0 and I75 are relatively weak. In the tropics and mid-latitude the misfits are globally reduced in I75, whereas they increase slightly below 510 m between 50 and 60°N (boxes 27-28) and south of 15°S . There, the changes may result from the deep adjustments at the open boundaries as discussed in section 6-3.

Fig. 23

8.2 Comparison with WOCE Sections

The simulated θ/S fields are compared to hydrographic data taken along two zonal sections at 4.3°S (section A7) and at 7.3°N (section A6), in January and February 1993 during the WOCE CITHER-1 cruise (Arhan et al., 1998). The largest differences between the observed and modeled temperature are observed in the first 100 m of section A7, with values up to 3.4°C in I0 (Table 3). The misfits are reduced in I75, but remain significant, because they are mostly due to the deficient vertical mixing (see Section 7-5). Differences between the climatological and A7 temperature fields exhibit however the same kind of pattern, which suggests that the climatology is not fully appropriate to constraint the thermocline at the equator. In deeper layers, the model/data differences reach relatively large values close to the Mid-Atlantic Ridge (for θ only) at 2000-2500 m depth, in the

Table 3

western basin below 3500 and near the bottom in the eastern basin (for both θ and S). Similar misfits are observed between the climatology and WOCE fields in the western basin below 3500 m. The situations in the western and eastern basins are thus different. In the west, the climatology is probably not an appropriate constraint (at least, for a simulation over 1993). In the east, close to the bottom at 10°W (Mid-Atlantic Ridge) and at 5°E , the θ/S fields depart more from the climatology and WOCE data in I75 than in I0. Since at these depth and time of the year, the fields are mainly influenced by the initial conditions, we suggest that the changes to the initial θ/S fields are not correct there.

Along section A6, the model/data SST misfits are significantly decreased between I0 and I75 (see Table 3). Misfits in temperature are also reduced over the whole water column, although differences up to 1°C appear close to the bottom in I75. Differences at these depths are partly due to the poor representation of the mixing and pathway of the North Atlantic Deep Water and Antarctic Bottom Water through the abyssal plains and fracture zones, because of the coarse vertical and horizontal resolution of our model. However, such misfits are observed also between the climatology and the WOCE data in the Guiana basin ($40\text{-}45^\circ\text{W}$) in both θ and S . Again, the climatology should probably not be applied as a constraint there.

The most striking feature on section A6 is the presence of large differences (positive and negative) in θ and S between the model and the WOCE fields, and between the climatology and WOCE data. They are located in the depth range 50-200 m, between the coast and 35°W . Arhan et al. (1998) detect oscillations on the A6 section and suggest they are the signature of baroclinic Rossby waves. These waves are not present in the climatology, but the model/data misfits relative to the waves seem to be slightly reduced between I0 and I75. Larnicol (1998) compares the dynamic height as computed from CITHER-1 data to T/P SSH. He shows that the signals are consistent, and show westward propagation, whose characteristics are close to those of the 1st baroclinic mode (propagation speed of 12 cm/s with a period of ~ 50 days). We clearly observe westward propagation in I75 between 6 and 9°N , with the strongest amplitude during the first 100 days and west of

35°W (Fig. 24). A very rough estimate of the velocity is ~ 17 cm/s at 7.5°N, which is consistent with the value of Larnicol (1998). We thus suggest that the wave signature contained in T/P SSH anomalies usefully constrains the model. As a result, the I75 simulation reproduces some wavelike signals with consistent physical characteristics, leading to a reduction of the misfits with the WOCE data. Fig. 24

9 Conclusions

A preliminary study of full, nearly rigorous, data assimilation based upon Lagrange multipliers has succeeded in the presence of open-boundary conditions. In the solution obtained after 75 iterations, significant changes in the circulation and in the θ/S distribution are found in the vicinity of the open boundaries. Although the one year of integration is too short to obtain definitive results about the ultimate constrained model state, nothing emerges to suggest any fundamental difficulties. Sometimes boundary values stemming from the optimization do not appear completely realistic, while the nearby interior circulation does appear improved compared to the unconstrained model. We suspect that to some degree, deficiencies in the model—primarily lack of adequate resolution—are being compensated by boundary control terms forcing the model closer to the observations than would otherwise be possible. This situation appears to be most likely in the Fram and Gibraltar Straits and near the rough topography of the Mid-Atlantic Ridge at the southern boundary.

The main improvements characterizing the constrained solution are in the Gulf Stream and in the equatorial regions. The constraints applied by the SST and the climatological θ, S profiles lead to an adjustment of the Gulf Stream thermohaline front and consequently to the mean circulation there. In the tropics, the model seems to be mainly constrained by T/P SSH anomalies, with improvements to the representation of the seasonal variability and the equatorial upwelling in the eastern basin. Because the run is only 1-year long, and because of the type of data constraint we apply, the improvements to the circulation are mainly located in the upper layers. There, comparisons with independent data sets (XBT profiles and WOCE equatorial sections) suggest that the optimized solution has become more realistic.

The tendency of the constrained solution to generate small spatial scales at the boundaries is probably the result of using diagonal a priori covariance matrices, \mathbf{Q} , \mathbf{R} , for the unknown controls. It likely indicates that the solution is overfitting the small scales contained in TOPEX/POSEIDON data. For TOPEX/POSEIDON, the data might usefully be pre-processed by removing the short along-track scales. For atmospheric fields, the use of simple correlation functions and radii would be a useful interim improvement. Changes in the open-boundary conditions exhibit a relatively large temporal variability, particularly near the bottom. Such variability could well be the result of inconsistencies between the high-frequency atmospheric forcing and the adjustment on the OB fields at monthly scales and is again an issue that can be dealt with by filtering and pre-processing.

In general, interpretation of the changes at the open-boundaries is not straightforward. In particular, it is not always obvious how to determine the specific impact of the adjustment at the boundary conditions on the observed circulation changes, nor to identify the specific constraint that led to such adjustments. A logical continuation of this work is the exploration of the components of the cost function gradients at the open boundaries for different kinds of data constraints.

10 Appendix Cost Function

The detailed expression of the cost function is:

$$\begin{aligned}
J &= (\bar{\eta} - \bar{\eta}_{TP})^T (\mathbf{R}_{geoid})^{-1} (\bar{\eta} - \bar{\eta}_{TP}) \\
&+ \sum_1^{365} ((\eta' - \eta'_{TP})^T (\mathbf{R}_{TP})^{-1} (\eta' - \eta'_{TP})) \\
&+ \sum_1^{12} ((\bar{\theta} - \bar{\theta}_L)^T (\mathbf{R}_{\theta_L})^{-1} (\bar{\theta} - \bar{\theta}_L) + (\bar{S} - \bar{S}_L)^T (\mathbf{R}_{S_L})^{-1} (\bar{S} - \bar{S}_L)) \\
&+ \sum_1^{12} ((\overline{SST} - \overline{SST}_R)^T (\mathbf{R}_{SST})^{-1} (\overline{SST} - \overline{SST}_R)) \\
&+ (\delta\theta_0)^T (\mathbf{Q}_{\theta_0})^{-1} (\delta\theta_0) + (\delta\mathbf{S}_0)^T (\mathbf{Q}_{S_0})^{-1} (\delta\mathbf{S}_0) \\
&+ \sum_1^{36} ((\delta\tau_x)^T (\mathbf{Q}_{\tau_x})^{-1} (\delta\tau_x) + (\delta\tau_y)^T (\mathbf{Q}_{\tau_y})^{-1} (\delta\tau_y)) \\
&+ \sum_1^{36} ((\delta\mathbf{H}_Q)^T (\mathbf{Q}_{H_Q})^{-1} (\delta\mathbf{H}_Q) + (\delta\mathbf{H}_F)^T (\mathbf{Q}_{H_F})^{-1} (\delta\mathbf{H}_F)) \\
&+ \sum_1^{12} ((\delta\theta_{ob})^T (\mathbf{Q}_{\theta_{ob}})^{-1} (\delta\theta_{ob}) + (\delta\mathbf{S}_{ob})^T (\mathbf{Q}_{S_{ob}})^{-1} (\delta\mathbf{S}_{ob})) \\
&+ \sum_1^{12} ((\delta\mathbf{U}_{ob})^T (\mathbf{Q}_{U_{ob}})^{-1} (\delta\mathbf{U}_{ob}) + (\delta\mathbf{V}_{ob})^T (\mathbf{Q}_{V_{ob}})^{-1} (\delta\mathbf{V}_{ob}))
\end{aligned}$$

θ_0 and \mathbf{S}_0 stand for the initial temperature and salinity conditions respectively; \mathbf{H}_Q , \mathbf{H}_F , τ_x and τ_y for the atmospheric heat and fresh water fluxes and for the wind stress zonal and meridional components respectively; θ_{ob} , \mathbf{S}_{ob} , \mathbf{U}_{ob} and \mathbf{V}_{ob} for the temperature, salinity and horizontal velocity fields prescribed at the open boundaries.

\overline{SST} , $\bar{\theta}$ and \bar{S} refer to the monthly averages of the simulated SST and θ, S fields whereas \overline{SST}_R , $\bar{\theta}_L$ and \bar{S}_L to the Reynolds and climatology monthly fields. $\bar{\eta}$ and η' indicate respectively the mean SSH and the daily SSH anomalies from the model; $\bar{\eta}_{TP}$

and η'_{TP} are their altimetric counterpart.

Acknowledgments. We gratefully acknowledge the help of C. Herbaut and R. Giering with the open-boundary conditions and the use of the TAMC respectively. Special thanks are addressed to P. Heimbach for very interesting discussions. The authors thank K. Zhang and the team of J. Marshall for helpful discussions about the MIT-GCM and the open-boundary problems, as well as C. King for TOPEX/POSEIDON data processing. The ERS1 wind products have been provided by CERSAT. The work has been supported under the National Ocean Partnership Program (NOPP) by grants from NASA, the NSF and the Office of Naval Research and by prior support of NASA and NSF. The computational support from the National Partnership for Computational Infrastructure (NPACI) is acknowledged.

References

- [1] Arhan M., H. Mercier, B. Boulès and Y. Gouriou, 1998: Hydrographic sections across the Atlantic at 7°30 N and 4°30 S, *Deep Sea Res. I*, 45, 829-872.
- [2] Blanke B. and P. Delecluse, 1993: Variability of the Tropical Atlantic Ocean simulated by a general circulation model with two different mixed-layer physics, *J. Phys. Oceanogr.*, 23, 1363- 1388.
- [3] Böning C.W., W.R. Holland, F.O. Bryan, G. Danabasoglu and J.C. McWilliams, 1995: An overlooked problem in model simulations of the thermohaline circulation and heat transport in the Atlantic Ocean, *J. Climate*, 8, 515-523.
- [4] Ezer T. and G.L. Mellor, 1992: A numerical study of the variability and the separation of the Gulf Stream, induced by surface atmospheric forcing and lateral boundary flows, *J. Phys. Oceanogr.*, 22, 660-682.
- [5] Foldvik A., K. Aagaard and T. Torresen, 1988: On the velocity field of the east Greenland Current, *Deep Sea Res.*, 35, 1335-1354.
- [6] Ganachaud A. and C. Wunsch, 2000: Oceanic meridional overturning circulation, mixing, bottom water formation rates and heat transport, *Nature*, 408, 453-456.
- [7] Giering R. and T. Kaminski, 1998: Recipes for adjoint code construction, *Trans. Math. Software*, 24, 437-474.
- [8] Gilbert J.C. and C. Lemaréchal, 1989: Some numerical experiments with variable-storage quasi-Newton algorithms, *Mathematical Programming*, 45, 407-435.
- [9] Joyce T.M., R.S. Pickart and R.C. Millard, 1997: Long-term hydrographic changes at 52 and 66°W in the North Atlantic Subtropical gyre and Caribbean, *Deep Sea Res.*, II, 46, 245-278.
- [10] Large W.G., J.C. McWilliams and S.C. Doney, 1994: Oceanic vertical mixing: a review and a model with a non local boundary layer parameterization, *Rev. Geophys.*, 32, 363-403.

- [11] Larnicol G., 1998: Analyse de la variabilité de l'Océan Atlantique à partir des données altimétriques TOPEX/POSEIDON et d'un modèle inverse non linéaire, Thèse de doctorat de l'Univ. de Brest, 185 pp.
- [12] Lemoine F. and 17 others, 1998: The development of the NASA GSFC and NIMA Joint Geopotential Model EGM96, NASA/TP-1998-206861 report, 575 pages.
- [13] Levitus S., R. Burgett and T. Boyer, World Ocean Atlas 1994, vol.3, Salinity and vol. 4 Temperature, NOAA Atlas NESDIS 3-4, US Dept. of Comm., Washington DC.
- [14] Macdonald A., 1998: The global ocean circulation: a hydrographic estimate and regional analysis, *Prog. Oceanogr.*, 41, 281-382.
- [15] Macdonald A. and C. Wunsch, 1996: An estimate of global ocean circulation and heat fluxes, *Nature*, 382, 436-439.
- [16] Marotzke J., R. Giering, Q.K. Zhang, D. Stammer, C. Hill and T. Lee, 1999: Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity, *J. Geophys. Res.*, 104, 29529-29548.
- [17] Marshall J., A. Adcroft, C. Hill, L. Perelman and C. Heisey, 1997: A finite-volume incompressible Navier-Stokes model for studies of the ocean on parallel computers, *J. Geophys. Res.*, 102, 5753- 5766.
- [18] Poulain J-M., A. Warn-Varnas and P.P. Niiler, 1996: Near surface circulation of the Nordic Seas as measured by Lagrangian drifters, *J. Geophys. Res.*, 101, 18237-18258.
- [19] Rintoul S. and C. Wunsch, 1991: Mass, heat, oxygen and nutrient fluxes and budgets in the North Atlantic ocean, *Deep Sea Res.*, 38, S355-S377.
- [20] Reynolds R.W. and T.M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation, *J. of Climate*, 7, 929-948.
- [21] Schott F.A., J. Fisher and L. Stramma, 1998: Transports and pathways of the upper-layer circulation in the western Tropical Atlantic, *J. Phys. Oceanogr.*, 28, 1904-1928.

- [22] Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C.N. Hill, and J. Marshall, 2001a, The global ocean circulation during 1992 –1997, estimated from ocean observations and a general circulation model, *J. Geophys. Res.*, in press.
- [23] Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C.N. Hill, and J. Marshall, 2001b: Volume, Heat and Freshwater Transports of the Global Ocean Circulation 1992 –1997, Estimated from a General Circulation Model Constrained by WOCE Data, submitted for publications.
- [24] Talagrand O., 1998: A posteriori evaluation and verification of analysis and assimilation algorithms, in Proceedings of workshop on Diagnosis of data assimilation systems, ECMWF, Reading, England, November 1998.
- [25] Vogeler A. and J. Schröter, 1999: Fitting a regional ocean model with adjustable open boundaries to TOPEX/POSEIDON data, *J. Geophys. Res.*, 104, 20,789-20,799.
- [26] Woodgate R., E. Fahrbach and G. Rohardt, 1999: Structure and transports of the East Greenland Current at 75°N from moored currentmeters, *J. Geophys. Res.*, 104, 18,059-18,072.
- [27] Wunsch C, 1996: The Ocean Circulation Inverse Problem, Cambridge Univ. Press, 442 pages.
- [28] Zhang K.Q. and J. Marotzke, 1999: The importance of open-boundary estimation for an Indian Ocean GCM-data synthesis, *J. Mar. Res.*, 57, 305-334.

Table 1. Values taken for the prior estimates of the control variables in the five experiments described in Section 5. 'C23' refers to the iteration 23 of the control run, 'G' refers to the outputs of the global simulation of Stammer et al. (2000a, 2001). '0' means that the prior estimate is zero (i.e. there is no prior estimate).

Prior estimates for the control variables					
	<i>Initial Conditions</i>	<i>OB Conditions</i>	<i>Heat Flux</i>	<i>Fresh Water Flux</i>	<i>Wind Stress</i>
<i>I0</i>	0	0	0	0	0
<i>C24</i>	C23	C23	C23	C23	C23
<i>I24</i>	C23	C23	G	G	G
<i>N1</i>	0	0	G	G	G
<i>N2</i>	0	0	G	G	0

Table 2. Meridional transport at 20°S below 3700 m: annual mean and standard deviation (STD) over 1993 in Sverdrups. Positive (resp. negative) values indicate northward (resp. southward) transports. 'Western' and 'eastern' refer to basins respectively west and east of the Mid-Atlantic Ridge (at about 15°W).

	I0		I75	
	<i>Mean</i>	<i>STD</i>	<i>Mean</i>	<i>STD</i>
<i>Western basin</i>	-0.4	0.9	-1.2	4.6
<i>Eastern basin</i>	-2.7	0.6	5.5	6.6

Table 3. Mean differences (in absolute value) between simulated and observed temperature and salinity at 7.5°N in February 1993 (section A6) and at 4.5°S in January 1993 (section A7). Differences between Levitus and A6-A7 data are given in the last row of each array ('L'). Numbers in bold indicate cases where the differences decreases between I0 and I75. The western tip of section A6 is tilted with respect to a line along 7.5°N; therefore, points west of 50°W are excluded for the comparison.

Temperature (°C)

	A6				A7			
	<i>0-20m</i>	<i>20-100m</i>	<i>100-985m</i>	<i>> 985m</i>	<i>0-20m</i>	<i>20-100m</i>	<i>100-985m</i>	<i>> 985m</i>
<i>I0</i>	2	1.5	1.03	0.25	3.48	3.02	0.5	0.26
<i>I75</i>	1.23	1.21	0.81	0.22	2.41	2.49	0.57	0.25
<i>L</i>	0.7	1.24	0.7	0.25	1.25	0.99	0.42	0.27

Salinity

	A6				A7			
	<i>0-20m</i>	<i>20-100m</i>	<i>100-985m</i>	<i>> 985m</i>	<i>0-20m</i>	<i>20-100m</i>	<i>100-985m</i>	<i>> 985m</i>
<i>I0</i>	0.22	0.16	0.13	0.012	0.28	0.21	0.064	0.009
<i>I75</i>	0.3	0.14	0.099	0.017	0.29	0.13	0.057	0.017
<i>L</i>	0.25	0.13	0.08	0.015	0.31	0.14	0.069	0.009

Figure 1: North and Tropical Atlantic domain of integration. Division of the domain into boxes for the diagnostics of the constrained run.

Figure 2: Mean Sea Surface Height as observed by TOPEX/POSEIDON (a) and as simulated by the unconstrained model (b). In both cases, the spatial average of the mean SSH has been subtracted at each grid point.

Figure 3: Error profiles on the climatology provided by Levitus et al. (1994). (The a priori error profiles used to weight the climatology constraint in the cost function are proportional to the profiles shown in this figure).

Figure 4: Model/data misfits over the whole integration period and the entire basin, for the different experiments described in table 1. The misfits are normalized by their values in the experiment C24 (iteration 24 of the control run).

Figure 5: Evolution as function of iterations of the total cost function J a) and its different terms J_i : constraint by θ/S Levitus fields b), constraint by T/P SSH anomalies c), constraint by Reynolds SST fields d), by T/P mean SSH e), changes to the control variables, i.e. initial θ/S conditions f), heat and fresh water atmospheric fluxes g), wind stress h), open-boundary conditions on T/S i) and on velocity j). J and J_i are normalized as explained in section 6-1.

Figure 6: Misfits between the model and Reynolds SST as a function of months: averages in boxes 7, 8, 9 and 12 (equatorial region) and boxes 24, 27, 28, 34 (north of 40°N), for the non constrained (I0 – thin line) and the constrained (I75 – thick line) simulations.

Figure 7: Adjustments to the initial profiles of temperature (top) and salinity (bottom), in the first 500 m and in four latitude bands (thick line). The thin line shows the a priori error on the initial profiles.

Figure 8: Annual mean adjustment to the heat flux (a), fresh water flux (b), zonal (c) and meridional (d) components of the wind stress at iteration 75. Negative values are in gray. Thick line indicates 0 contour line. Contour interval is 10 Wm^{-2} from -55 Wm^{-2} to 45 Wm^{-2} for heat flux, and 0.1 m/yr from -2 to 2 m/yr for fresh water flux. Contours for wind stress are: $-3 \cdot 10^{-2}$, $-1 \cdot 10^{-2}$, $-5 \cdot 10^{-3}$, $5 \cdot 10^{-3}$, $1 \cdot 10^{-2}$, $3 \cdot 10^{-2} \text{ Nm}^{-2}$.

Figure 9: Annual mean and standard deviation of the adjustments to the temperature and salinity fields in I75, prescribed at the open boundaries (a) at 24.5°S and 79.5°N and (b) at 15.5°E and 6.5°W . Units are $^{\circ}\text{C}$ for the temperature.

Fig. 9, continued

Figure 10: Annual mean and standard deviation of the adjustments to the velocity fields in I75 prescribed at the open boundaries: (a) meridional velocity at 24.5°S and 79.5°N and zonal velocity to which the model is restored at 23.5°S and 78.5°N ; (b) zonal velocity at 15.5°E and 6.5°W and meridional velocity to which the model is restored at 14.5°E and 7.5°W . Units are cm/s.

Fig. 10, continued

Figure 11: Annual mean (in absolute value) of the model/data misfits in terms of temperature and salinity in the immediate vicinity of the open boundaries: (a) at 23.5°S for θ and $I0$, (b) at 78.5°N for θ and $I0$, (c) at 14.5°E for θ and $I0$, (d) at 7.5°W for θ and $I0$; (e), (f), (g) and (h): same as (a), (b), (c), (d) but for $I75$; (i), (j), (k), (l), (m), (n), (o) and (p): same but for S . Units are $^{\circ}\text{C}$ for θ .

Fig. 11, continued

Figure 12: Annual mean of the velocity fields at 30 m for I0 (a) and I75 (b) in the Nordic Seas.

Figure 13: Standard deviation of the SSH anomalies as seen by TOPEX/POSEIDON (a) and as estimated in the I0 (b) and I75 (c) simulations. Units are in cm.

Figure 14: Meridional heat transport for I0 (blue) and I75 (red) as a function of latitude. Curves indicate annual means, error bars the standard deviations around the mean. Values estimated by Macdonald (1998) and Ganachaud and Wunsch (2000) are shown in green and magenta respectively, together with their uncertainties.

Figure 15: Annual mean of the horizontal velocity averaged over the first 100 m for I0 (a) and I75 (b), and between 100 and 710 m depth for I0 (c) and I75 (d). Colors indicate the horizontal velocity amplitude in cm/s.

Figure 16: Temperature/salinity diagrams in boxes 24 (top) and 35 (bottom) as monthly averages. Thick light line corresponds to Levitus data, thin line to I0 and thick dark line to I75.

Figure 17: Mean temperature for December 1993 at 65°W as a function of depth and latitude, for I0 (a), I75 (b) and Levitus data (c).

Figure 18: Difference between the mean SSH in I75 and I0 (positive values indicate a higher SSH in I75). Thick line indicates 0 contour line. Contour interval is 5 cm. Units are in cm.

Figure 19: Temperature profiles in the first 200 m as a function of time in a point located at 10°W at the equator: a) non constrained simulation (I0), b) constrained simulation (I75) and c) climatology.

Figure 20: Differences between August and March SSH anomalies fields for a) non-constrained simulation I0, b) constrained simulation I75, c) TOPEX/POSEIDON data. Units are in cm.

Figure 21: Mean velocity fields with the amplitude in colors: a) mean over the first 35 m and from April to June for I0, b) same as a) but mean over Sep.-Oct.-Nov., c) same as b) but mean over the depth range 35-260 m, d) same as a) for I75, e) same as b) for I75, f) same as c) for I75. Units are in ms-1.

Figure 22: Distribution of the XBT data over the year 1993.

Figure 23: Model/XBT data misfits (in absolute value) as a function of months averaged over box 24 and for the four layers.

Figure 24: Sea surface height anomalies as a function of time and longitude at 7°N for a) non- constrained simulation I0, b) constrained simulation I75, c) TOPEX/POSEIDON data. Units are in cm.