

A 4D-variational approach applied to an eddy-permitting North Atlantic configuration: Synthetic and real data assimilation of altimeter observations

Bruno Ferron^{a,*}

^a Laboratoire de Physique des Océans, UMR 6523 CNRS-IFREMER-IRD-UBO, IFREMER, B.P. 70, 29280 Plouzané, France

* Corresponding author : Bruno Ferron, Tel.: +33 298 224 566; fax: +33 298 224 496, email address : bruno.ferron@ifremer.fr

Abstract:

The increasing number of oceanic observations calls for the use of synthetic methods to provide consistent analyses of the oceanic variability that will support a better understanding of the underlying mechanisms. In this study, a 1/3° eddy-permitting model of the North Atlantic (from 20°S to 70°N) is combined with a 4D-variational method to estimate the oceanic state from altimeter observations. This resolution allows a better extraction of the physical content of altimeter data since the model spatial scales are more consistent with the data than coarser assimilation exercises because of a lower error in model representativity. Several strategies for the assimilation window are tested through twin experiments carried out under the following conditions: different window lengths and either a quasi-static (also known as progressive) variational assimilation with progressive extension of the window, or a simpler direct method without prior assimilation. From our set of experiments, the most efficient strategy is the use of both a simple direct assimilation method and a 90-day window. The assimilation of synthetic altimeter data constrains the model-temperature, -salinity and -velocity fields mainly over the first 1300 m where the error is the largest. Improvements occur not only in quiescent regions, but also in more energetic meso-scale regimes. Despite the existence of model- and surface forcing-errors as well as large errors in the first guess, the assimilation of real altimeter data proves to be consistent with our twin experiments. Indeed, the analyses show a better detachment of the Gulf Stream, weaker regional biases and more accurate positions for meso-scale structures. Independent hydrographic data (Argo floats and CTD cruises) and transports estimates along the OVIDE 2002 cruise show an improvement of the analysed oceanic state with respect to the assimilation-free case though water mass properties are still incorrectly represented. After assimilation, the North Atlantic heat transport in the model is in good agreement with independent estimates based on hydrographic data.

Highlights

► Sea surface height is assimilated (4D-var) in an eddy-permitting configuration. ► Synthetic and real data assimilation show consistent behaviours. ► A time window of 90 days is efficient to constrain the primitive equation model. ► Stratification is improved down to 1300 m over the whole domain. ► Meridional overturning and heat transport are consistent with independent estimates.

Keywords: Ocean modelling; Data assimilation; Variational method; Meso-scale; Sea surface height; Heat transport; Meridional overturning

1 Introduction

Observation of the oceans with either satellites or in-situ-platforms (floats, gliders, moorings, ships) provides one with a growing number of data sets. Each of them covers a specific part of the spatial and temporal spectra of the ocean dynamics. Producing coherent syntheses of such pieces of observations to better describe and understand the puzzle of the ocean variability is a challenging task of primary importance. Among the available techniques, the four dimensional variational (4D-var) assimilation (Le Dimet and Talagrand, 1986; Thacker and Long 1988) has the advantage of producing analyses of various types of observations that are consistent with the dynamics of a general circulation model (GCM) over a given period. This study was aimed at developing an assimilation system whose dynamics remains consistent with observations from seasonal to inter-annual scales. Our main goal is to diagnose and quantify the seasonal to inter-annual variability of the circulation and heat transport of the North Atlantic Ocean to facilitate further interpretation of the variability seen in “synoptic” repeated hydrographic cruises by providing basin-scale and time-averaged oceanic analyses. We expect that these analyses will enable us to quantify the seasonal representativity of a quasi-synoptic hydrographic survey, or to which extent the variability of a regular annual survey is, indeed representative of the inter-annual variability.

The 4D-var assimilation of “instantaneous” (i.e. not climatological) oceanic observations was successfully applied to ocean GCMs at low horizontal resolution (1° or coarser) with realistic surface forcings and bathymetry (e.g. Stammer et al. 2002, 2003; Ferron and Marotzke 2003; Weaver et al. 2003; Vialard et al. 2003; Masuda et al. 2003; Forget 2005; Forget et al. 2008). As such low resolution models are quasi-linear over 1-year timescales; they permit one to use long assimilation periods. However, they have also a large representativity error due to unresolved processes, especially in the extra-tropics where meso-scale energy is significant. That limits the physical content liable to be extracted from an observation by the assimilation process. This is particularly true for sea surface height (SSH) observations since a large part

of the SSH is composed of meso-scale structures. It is then legitimate to seek for resolution enhancement so as to get a model dynamics, with a lower representativity error and fitting better with the real ocean dynamics and its observations. Given the spatial scales resolved by SSH data and the Argo network of profiling floats, an eddy-permitting resolution is a good compromise between observed scales and the numerical cost of the method.

Increasing the resolution is, however, not straightforward for the assimilation problem whose aim is to reduce the misfits between a model and observations (the so-called cost function minimisation). The more non-linear the model dynamics is, the higher the probability is to get, in the cost function, multiple minima leading to a poor convergence of the assimilation process (Schröter et al. 1993; Tanguay et al. 1995; Morrow and de Mey 1995; Pires et al. 1996). This convergence is achieved iteratively through use of cost function gradients given by the linear model adjoint. When huge and localised sensitivity regions appear in these gradients, they prevent a significant minimisation of the cost function (see Hoteit et al. 2005 for an illustration, detailed explanations and references). Several methods have been proposed in the literature to avoid packed extrema in the gradients and to find the global minimum of the cost function in (moderately) non-linear models. They mainly consist in:

- adding a smoothness term in the cost function (Schröter et al. 1993).
- using a quasi-static variational assimilation (QSVA, Pires et al. 1996) also called progressive assimilation (Blum et al. 1998) that keeps track of the absolute minimum of the cost function. The assimilation is done in N steps over which the assimilation window is progressively extended. Over the first step, the model initial conditions are optimised on a short assimilation window from the time, T_0 , to the time, T_1 . In the next step, the optimised initial conditions are used as a first guess to assimilate observations from T_0 to T_2 with $T_2 > T_1$. Then, in the last step, N , the data are assimilated from T_0 to T_N , with $T_N > T_{N-1}$, to estimate the initial conditions previously optimised during the $N-1$ step.

- using an incremental 4D-var in which the assimilation is done at low resolution (forward and adjoint models) around a trajectory of a higher non-linear model, which is frequently updated (Laroche and Gauthier 1998).
- using an assimilation of statistical moments to improve the climatological state of eddy-resolving GCM. In this method, statistical moments are derived from a high resolution GCM, compared to the equivalent-mean observations and assimilated by using a low resolution adjoint version of the same GCM (Köhl and Willebrand 2002; 2003).
- using a multi-scale method: a regional coarse assimilation defines the best first guess for the eddy-resolving assimilation adjoint (Gebbie et al. 2006).
- increasing viscosity and diffusivity coefficients in the adjoint model (Köhl and Willebrand 2002; Hoteit et al. 2005; Mazloff 2008).

Despite the lack of comparison between these methods, we assumed that using higher viscosity/diffusivity coefficients in the adjoint model should be equivalent to using an adjoint with a coarser horizontal resolution. For instance, in the study by Hoteit et al. (2005), increasing the viscosity and diffusivity parameters, in the adjoint, by a factor 30 led to typical coefficients used in models of 1° in horizontal resolution. A clear benefit of using high viscosity and diffusivity coefficients in the adjoint lies in the efficiency to damp large values that, sometimes, tend to develop during the temporal integration of the adjoint. Longer assimilation periods may then be reached. However, at each time step, the cost function gradients are smoothed everywhere in the model domain by high viscosity and diffusivity coefficients. For long assimilation periods, this excessive smoothing may remove a part of the signal contained in the observations on the smallest scales resolved by the forward model; but this effect was not quantified, here. To avoid undesirable smoothing effects, but at the expense of a shortening of the assimilation period, the QSVA method was tested in this study with change in neither the adjoint model resolution nor viscosity/diffusivity. QSVA techniques were successfully tested with the Lorenz system (Pires et al. 1996)

and with a low-resolution quasi-geostrophic model (Blum et al. 1998).

In this study to get a model solution consistent with the assimilated SSH data, several experiments were conducted at an eddy-permitting resolution and with different assimilation periods. As the QSVA method has never been tested for such a high dimensional system, it appeared to us worthwhile to do. Thus, QSVA is compared, here, to other simpler strategies through the conduct of twin experiments. Improvements in the tracer and velocity fields brought by the assimilation of SSH in the twin experiments are quantified. The results are compared to those from previous studies. Twin experiments are ideal to define an assimilation system since the true model ocean state history reconstructed by the assimilation process is known. However, they cannot address the whole complexity of real data experiments. In particular, results from real data assimilation can be worse due to the presence of errors in the model and surface forcing, approximations in error covariance matrices or large errors in the first guess. In the twin experiments conducted by Lea et al. (2006) in a 4D-var exercise carried out with a 9-km horizontal resolution GCM of the Irminger Basin, application to real data assimilation was missing. However, in their conclusion these authors felt confident that the assimilation of real data would be successful for their regional domain with moderate eddy kinetic energy and also for larger domains. In the last part of the study reported here the assimilation of real SSH observations is tested on the North Atlantic, and the results are compared to those from twin experiments and to independent hydrographic observations and transport estimates.

Section 2 of the paper presents the main characteristics of the assimilation system. Section 3 focuses on the twin experiments, whereas Section 4 is devoted to real data assimilation. The results of the study are discussed in Section 5.

2 The assimilation system

2.1 - Forward model

The primitive equations were solved with the OPA general circulation model (Madec et al. 1998). The configuration covered the North Atlantic from 20°S to 70°N, and from 96°W to 20°E. The horizontal resolution was $1/3^\circ \times 1/3^\circ \cos(\text{latitude})$ (longitude \times latitude). It was decreasing from 33 km at the Equator to 12 km close to 70°N. There were 43 levels along the vertical with vertical resolution decreasing from 10 m close to the surface to 500 m at depth. The bathymetry was based on the dataset (version 9.1) by Smith and Sandwell (1997) with a careful definition of the passages where key overflows were present. Horizontal mixing was parametrised by a bi-harmonic operator. Coefficients for viscosity and diffusivity were equal to $\cos^3(\text{latitude}) \times 2.5 \times 10^{11} \text{ m}^2 \text{ s}^{-1}$. The turbulent kinetic energy scheme by Blanke and Delecluse (1993) provided vertical eddy viscosity and diffusivity with background values of $1 \times 10^{-5} \text{ m}^2 \text{ s}^{-1}$ (diffusivity) and $1 \times 10^{-4} \text{ m}^2 \text{ s}^{-1}$ (viscosity). A free slip boundary condition and a quadratic friction with a bottom drag coefficient of 1×10^{-3} were used. The model was forced with daily mean surface fluxes from NCEP re-analysis 2 (Kamitsu et al. 2002). The domain was closed with restoring towards the seasonal climatology of Reynaud et al. (1998) on the northern and southern Atlantic boundaries as well as on the eastern Mediterranean boundary. A sea surface restoring condition towards the weekly SST fields from Reynolds and Smith (1994) was also used.

2.2 - Assimilation characteristics

The adjoint model (Weaver et al., 2003) was derived from the tangent linear model of the forward model. No turbulent kinetic energy scheme is present in the adjoint model. On the other hand, eddy viscosity and diffusivity were recorded in the forward run in addition to the model state variables needed for the

backward integration of the adjoint model. These recorded variables were daily averaged. The control variables were composed of the initial conditions for tracers, velocities and free surface (a prognostic variable in this version of OPA, Roulet and Madec 2000). Since at this eddy-permitting resolution, the ability to use long assimilation periods was unlikely, no attempt was made to control the surface forcing fields, the closed boundaries or any other internal model parameter. Restoring towards the weekly SST was maintained in the assimilation to compensate the absence of control of the surface heat fluxes.

The cost function, J , minimised by the assimilation process is such that:

$$J[x(t_0)] = 1/2[x(t_0) - x^b(t_0)]^T B^{-1}[x(t_0) - x^b(t_0)] + 1/2 \sum_{i=1}^n [y_i - y_i^o]^T R^{-1}[y_i - y_i^o]$$

where the first term on the right hand-side is the background cost, J_b , the second term is the observation cost, J_o , x is the vector of ocean state variables, and t_0 is the time at the beginning of the assimilation window. The background state, x^b , is our first guess; J_b that quantifies the changes in the initial ocean state is scaled by the covariance matrix of the background errors, B . In the assimilation window $[t_0; t_n]$, J_o sums the misfits between the observations, y_i^o , at the time, t_i , and their model counterpart, $y_i = Hx(t_i)$, where H is the observation operator of the model ocean state at the time, t_i . J_o is scaled by the covariance matrix of the observation errors, R . Thus the assimilation is expected to modify $x(t_0)$ so as to produce an optimised model trajectory, i.e. a time-varying ocean state that fits the time-distributed observations within their error bar all along the assimilation window. Since the assimilation period should be as long as possible to take advantage of the dynamical propagation by the adjoint model, the assimilation uses the full non-linear model, \mathcal{M} , as the forward model. The cost function is then non-quadratic since in the observation cost function y_i reads: $y_i = Hx(t_i) = H\mathcal{M}[x(t_{i-1})]$. Several minima may then appear depending on the importance of the non-linearities as outlined in the introduction.

Ideally, B should contain error covariances of the first-guess initial conditions, which depend on t_0 . Since

B cannot be quantified precisely, it is assumed to be diagonal for simplicity. One should note that, despite this choice, the assimilation-produced changes in the ocean state at t_0 present correlations between both grid cells and state variables. Indeed, one strength of the 4D-variational assimilation lies in the dynamics of the adjoint model that propagates the information contained in the observations backward in time. This provides dynamically correlated modifications of the control variables. Local inconsistencies may appear between nearby grid cells. In order to damp them, an auto-adjoint filter is applied to the cost function gradient with respect to the control variables, i.e. the adjoint variables at the initial time, t_0 . This filter, a weighted average of a grid-point with its four neighbours, is equivalent to applying a viscosity/diffusivity parameter 30 times the value of the forward model for only 9 hours of model adjoint integration. Its effect remains spatially localised compared to using an enhanced viscosity/diffusivity parameter at every time step of the adjoint model time integration (e.g. Hoteit et al. 2005). Twin and real data assimilation experiments use a different B matrix, and more details about it (B) are given in the corresponding Sections (Sections 3 and 4).

This study is focused on the assimilation of only SSH. The Aviso data centre provides SSH observations as along-track or gridded data. In the twin experiments, the assimilation of along-track SSH, which is a point-wise piece of information in time and space, produces occasionally, at high latitudes, large values in the cost function gradient. This may cause problems of convergence in the assimilation iterations. Such spikes come from neither dynamical inconsistencies between the assimilated data and the model since these data are generated by the same model configuration, nor assimilation of erroneous data given that the data and their errors are controlled in the twin experiments. North of 56°N , the gradient of the cost function exhibits the signature of the assimilated tracks, which is no longer the case further south. Because the adjoint dynamics is unable to smooth the assimilated information contained in the tracks, small scale structures may persist and occasionally grow up at high latitudes. The use of error covariances in B would help to smooth the spikes once they are developed. However, it seems more appropriate to

prevent their development by constraining the dynamical scales in the adjoint run. An ad hoc method would be to significantly increase the horizontal diffusion at high latitudes. Another method, used here, is to assimilate gridded observations that force the adjoint model consistently over nearly the whole domain. Thus, our real SSH assimilation uses the $1/3^\circ \times 1/3^\circ$ horizontal resolution maps of absolute dynamic topography (MADTs) in delayed time produced by Ssalto/Duacs and distributed by Aviso. Along-track sea level anomalies from satellites are mapped via an optimal interpolation with realistic correlation functions and then merged with the mean dynamic topography by Rio et al. (2005) to build MADTs. MADTs are available every 7 days and are assimilated as an average of the dynamic topography over that period. Since the interpolation between tracks is not controlled by the assimilation system as expected under ideal conditions, errors on AVISO MADTs have to be carefully quantified.

The MADT errors used in the twin and real data experiments consist of i) the error on the mean dynamic topography provided by Rio et al. (2005; Fig. 1a) and ii) the error on the mapped sea level anomalies (Fig. 1b). The latter is estimated through comparison of sea level anomaly maps (based on data from Jason-1 and Envisat satellites) against along-track sea level anomalies for the year 2004. When these tracks come from Topex/Poseidon, the maps and tracks are independent; on the other hand, they are dependent when they are issued from Jason-1. Over the year 2004, Topex/Poseidon was on an orbit halfway between different Jason-1 tracks (tandem mission). The rms of the differences in sea level anomalies between maps and tracks shows that the largest values are found in regions of high meso-scale energy such as the Gulf Stream extension and the Gulf of Mexico (Fig. 1b). Large differences are also found in a few coastal regions associated with residual errors in tidal corrections. The comparison between maps and tracks highlights an upper bound of the mapping error in the case of Topex/Poseidon tracks and a lower bound for Jason-1 tracks, which explains the banded structure on Fig. 1b. Except for some coastal regions, the error on the anomalies is far greater than the one on the mean sea level (Fig. 1ab). In the assimilation process, the error covariance matrix on the MADT data is diagonal and composed of both maps. One should note that, in shallow regions, the error on MADTs may be underestimated since

the usual corrections made mainly for the open ocean are not necessarily appropriate to coastal regions (e.g. Madsen et al. 2007, Bouffard et al. 2008).

3 Twin experiments

The performed experiments were of two types: they consisted of i) twin experiments dedicated to the assimilation of synthetic observations and ii) real data assimilation experiments. Twin experiments have the advantage of providing extensive comparisons between the optimised state and the true model state. Their results can be used as reference to gain further insight into the more complex assimilation of real data.

For twin experiments, a 10-year-long spin-up was constructed: it started at rest from Reynaud et al. (1998) seasonal climatology on January 1, 1995. Our reference “true” state was the model trajectory of the year 2001. The model free surface of this reference state was averaged over 10-day periods to create maps of synthetic SSH. At the start of the assimilation process (January 1, 2001), the initial conditions for temperature, salinity, velocities and free surface were changed to those of the model state on January 1, 2002. Statistics from these perturbations were used to compute the diagonal of the background error matrix, B . For instance, those corresponding to the temperature field were dependent upon the model depth. At a given model depth, the corresponding element in B was calculated as the squared inverse rms of the temperature differences between the first guess (January 1, 2002) and the true state (January 1, 2001). The same method was applied to the other fields (salinity, velocity, free surface). The largest differences in the vertical profiles of these rms are found in the upper ocean where the oceanic variability is the most marked (Fig. 2). The assimilation problem under study is: when starting from the perturbed initial conditions, what is the quality of the retrieval of the reference time-varying state variables in the case where only SSH maps are assimilated? One should note that, during the assimilation, the surface

forcing in use is that of the year 2001. As opposed to the real data experiments of section 4, model- and surface forcing-errors are missing in the twin experiments. If the output of a different model with an eddy-resolving resolution had been used as a “true state”, the testing of the assimilation scheme would have been more severe; it would be worth encouraging such twin experiments.

In order to get closer to reality, synthetic SSH maps are blended with an error that takes into account the errors on both the mean sea level (Rio et al. 2005) and the sea level anomalies (Fig. 1). A latitude-dependent correlation length varying from 180 km at the Equator to 90 km north of 60°N is added to the error on sea level anomalies as diagnosed from the comparisons between satellite tracks and AVISO maps. Figure 3 gives an example of error added to a synthetic map. This error is composed of a systematic bias representing the error on the mean sea level plus a time-evolving noise for the error on the sea level anomalies. The highest error amplitude (60 cm in the example of Fig. 3) is always located in regions of high meso-scale energy.

3.1 – Testing the assimilation period

In the study by Pires et al. (1996), it was proposed to use an assimilation period proportional to the doubling time of growth of small errors. Given a typical small initial error (e.g. 10% of the difference between two model states separated by a year), the error doubling time, t_{nl} in our configuration is about 20 days for the velocity fields and 30 days for the tracer fields. This also corresponds to the period of validity of the tangent linear model. The cost function is then expected to remain quadratic over periods of nearly 30 days. Hoteit et al. (2005) reported a linear behaviour for about 30 days with a 1/3° model of the tropical Pacific with usual viscosity/diffusivity values. This is also consistent with a t_{nl} of about 20 days reported by Brasseur et al. (1996) for a 4-layer quasi-geostrophic model of the North Atlantic with a horizontal grid mesh of 1/6°.

In order to find the best strategy to assimilate observations over periods longer than t_{ni} , we compared three approaches based on twin experiments (Table 1). The first one, Dir30, consists of 3 successive 30-day periods; the optimised state at the end of a period is employed as a first guess for the next period. The second one, the QSVA techniques, uses the first optimised 30-day period of Dir30: the optimised initial conditions of the first 30-day period are used as a first guess and the assimilation period is extended from 30 to 90 days. In the third approach, Dir90, the observations are directly assimilated over a 90-day period without prior assimilation. Three and 9 MADTs are assimilated over a 30-day window and a 90-day one, respectively. A given cycle is stopped when the decrease of the cost function, J , between two successive iterations is less than 2%. Reaching this minimum for the cost function requires, on average, 20 iterations for a 30-day window, and 25 iterations for a 90-day window. Thus, cycling three 30-day windows to get an optimised trajectory over 90 days saves 20% of CPU time compared to a 90-day window. With a unique NEC-SX8 processor, 25 iterations with a 90-day window take 10^5 CPU seconds. The storage needed during the experiment is directly proportional to the window length (14.5 Go for the 90-day cycle, 3-fold less for the 30-day cycle). For the 3 approaches, residuals are consistent with observation error variances, and changes in the initial state fall within the background error variance. The results presented in the next paragraphs were obtained for an assimilation period covering the winter season (January to March); one should note that starting the assimilation in summer led to similar results.

An advantage of twin experiments comes from the possibility of measuring misfits between the model and true observations (perfect MADTs) while the assimilation process is guided by the model-noisy MADT misfits. Without assimilation, the model-true MADT misfits reach an average rms of about 6 cm over the whole domain (Fig. 4, black line). The average misfit between the model and the noisy MADTs is, on average, 60% greater than the observational error before assimilation.

For a given 30-day cycle, the model-true MADT misfits exhibit a maximum for the first assimilated map and a decrease for the next two assimilated maps (Fig. 4, blue line). For the last 2 cycles of 30 days, this maximum is close to 3.8 cm and associated with model-noisy MADT misfits, which are, on average, 7% greater than the observational error. For the following two maps, model-true MADT misfits are less than 3 cm and associated with model-noisy MADT misfits, which are 5 to 12% less than the observational error. For a given window, the background constraint represented by the term, J_b in the total cost function, J , is null for the first iteration of the assimilation, and then increases with the number of iterations. Concerning J_b , it accounts for 35% of the total cost function for the last iteration of the first window. For the next two windows, this percentage is decreasing to 25% thanks to the improvement of the first guess by the previous assimilation window. Hence, the assimilation is mostly constrained by the observational term, J_o .

The QSVA method starts from a first guess equal to the optimised initial state of the first 30-day window and extends this window to 90 days. The model trajectory to be optimised is then a forecast from the optimised initial state of the first 30-day window (Fig. 4, black circle-dotted line). This forecast clearly shows a rapidly growing error after the third map, which means that the initial state optimised by the first 30-day cycle has only been partly reconstructed by the assimilation of the first 3 MADTs. In other words, the pieces of information provided by the assimilation of 3 MADTs are insufficient to constrain the model trajectory on time-scales exceeding 30 days. From days 20 to 70, the model-true MADT misfits produced by the QSVA method are either equal to or less than those from Dir30 approach; the opposite is found for the last 20 days (Fig. 4, red line). The 6 last assimilated MADTs are associated with model-noisy MADT misfits that are, on average, 7% greater than the observational error.

The interest of QSVA and Dir90 is that both produce a continuous solution over 90 days. The use of a 90-day window without any prior assimilation (Fig. 4, green line Dir90) shows the lowest model-true MADT misfits after day 40. The results are, however, close to the QSVA approach. The 6 last assimilated MADTs are associated with model-noisy MADT misfits that are on average 4% greater than

the observational error. Hence, non-linearities within a 90-day window are weak enough to permit the assimilation to converge and to extract information from the assimilated maps, which optimises the model trajectory. It is worth noting that the succession of several 30-day-long periods produces slightly smaller model-noisy MADT misfits, but slightly larger model-true MADT misfits than the use of a 90-day-long window. This suggests that the noise present in observations can be partly filtered out by using a longer assimilation period because of a better dynamical propagation of the information in the adjoint run. Conversely, the optimised initial state of a 30-day-long period is more contaminated by the observational noise, which leads to a less accurate model trajectory. For the experiment Dir90, observations are again the main constraint of the assimilation since, at the end of the iterations, the background cost function, J_b , represents only 11% of the total cost function, J . This percentage is lower for the 90-day window than for the 30-day window since the former enables the collection of 3-fold more observations.

The presence of a maximum model misfit at the beginning of an assimilation cycle is a common feature of 4D-variational assimilation when initial conditions are modified by the assimilation scheme (e.g. Hoteit et al. 2005, Titaud et al. 2010). The first guess of the second 30-day cycle comes from the optimised ocean state at the end of the first 30-day cycle. Then, it is modified by the assimilation during the second cycle. The new optimised initial conditions generated by the assimilation are not a balanced state of the model equations. Dynamical adjustments between the modified stratification and the modified velocity field take place at the beginning of the assimilation window. They create larger model misfits for the observations at the beginning of the assimilation window. Ideally, the search for new initial conditions should be made in the phase space of the dynamically balanced states, which are unknown. Dynamical adjustments at the beginning of a window may be reduced by using either an appropriate error covariance matrix, B , or a weak constrain formulation (Gauthier and Thépaut, 2001). To avoid the adjustment period in model state diagnostics, a more pragmatic solution is to make the first assimilation window overlap the next one by a few days. The model diagnostics for the overlap period are then determined from the window that ends

with the overlap.

As opposed to real data assimilation, twin experiments also enable one to measure to which extent the true state is retrieved over the whole domain for the not directly assimilated variables (velocity, temperature, salinity). A comparison of the temperature analyses issued from the QSVA method and the 90-day window without any prior assimilation (Dir90) shows that the QSVA method performs better than Dir90 at the beginning of the assimilation, but becomes worse than Dir90 from fortnight 3 to 6 (Fig. 5). This behaviour is found for all state variables and at every depth. One should note the significant improvement in the analysed temperature from fortnight 1 to fortnight 6 with a 30% decrease of the maximum rms error below the mixed layer.

The fact that the QSVA techniques do not give the best results when the assimilation time period is extended from 1 to 3 months (about $3 t_{nl}$) likely comes from the high dimensional system of our GCM. In their assessment of QSVA techniques, Swanson and Vautard (1998) already noticed that the longest assimilation period usable decreased from $8 t_{nl}$ (Pires et al. 1996) for the low-order Lorenz system (1963) to $3-4 t_{nl}$ with a 3-layer quasi-geostrophic model of the atmosphere. It is, however, worth noting that, with a 3-month-long assimilation period, convergence towards observations is unaffected by the existing nonlinearities.

After 90 days, Dir90 produces slightly better results than Dir30 and QSVA. Moreover, contrary to Dir30, Dir90 produce a continuous analysis over 3 months. These findings led us to base our following twin and real-data experiments on only the direct assimilation approach (i.e. QSVA is no longer used) with 90-day-long assimilation cycles, unless otherwise stated.

3.2 - Carrying on the assimilation over 1 year with synthetic data

The previous section showed that the first 90-day cycle not only decreased significantly the misfit between the model and the true observations, but also improved the model state variables for which no data from direct observation were assimilated. The next step is to determine to which extent the model state can be improved when several windows are carried on up to the saturation of the quality of analyses induced by the presence of errors in MADTs. Moreover, it is worth quantifying whether the assimilation is as efficient in regions of high meso-scale energy as it can be in more quiescent regions. This led us to conduct an analysis over 1 year from 4 cycles of 90-day windows. In this exercise, the optimized model state at the end of a window is the first guess of the next cycle.

Figure 6 highlights that the assimilation of 36 noisy MADTs generates a marked reduction of the model-true state misfit from 5.5 – 6 cm in the assimilation-free case to nearly 2 cm over the last 80 days of this 1-year-long exercise (Fig. 6). During that period, residuals are consistent with the error on MADTs. Each new assimilation cycle brings the model free surface closer to true MADTs. The gain between two successive windows is the weakest for the last two 90-day cycles, which indicates that the assimilation of new observations hardly brings the model free surface closer to the true MADTs. There is no additional enhancement in the quality of the model free surface analyses when the assimilation of noisy MADTs is prolonged up to 18 months (not shown on Fig 6). The impact of the error on MADTs can be quantified on condition to assimilate perfect MADTs instead of noisy ones. After 1 year, the model misfit with perfect MADTs is reduced to about 1 cm (Fig. 6, blue line). When the observations are considered as being perfect, the assimilation of new observations still brings the model free surface closer to true observations after a year. Indeed, the improvement observed between cycles 3 and 4 and the one between cycles 2 and 3 are alike (Fig. 6, blue line) conversely to the case of noisy observation assimilation (Fig. 6, red line).

It is worth noting that the analyses of the model free surface get closer to the true state not only in quiescent meso-scale regions, but also in active ones. In the experiment without assimilation, the misfit between the model free surface and the true reference trajectory at the end of the year 2001 exhibits large-scale structures of moderate intensity and meso-scale structures with maximum values of ± 68 cm in the Gulf Stream path (Fig. 7a). After a year of assimilation of noisy MADTs, the misfit is reduced everywhere (Fig. 7b). Maximum values are still found in the Gulf Stream region, but they are reduced to ± 22 cm. These residuals are close to the error added to the perfect MADTs (Fig. 1). The lack of large scale misfit after assimilation means that regional biases were successfully removed.

The MADT-carried information improves not only the model free surface, but also temperature, salinity and velocity fields away from the surface. The rms of the differences in temperature between the assimilation-free experiment and the reference is the largest below the mixed layer where it reaches 0.9°C (Fig. 8a). For the meridional velocity, it is the highest at the surface with values close to 0.11 m/s (Fig. 8b). These maximum errors are decreased to 0.55°C and 0.05 m/s by assimilation of noisy MADTs. One satisfying result comes from the vertical projection of the information contained in the MADTs. The temperature, salinity and velocities are mainly improved over the upper 1300 m where the initial error is the largest. For tracers, the typical relative decrease in rms between the states without and with assimilation reaches 30% near the surface, and 45% for the velocities. Below 3000 m, where the initial error is weaker, the decrease is in the range 10-20% for the tracers and 25-30% for the velocities. This is consistent with the observations made by Lea et al. (2006) about their regional and less energetic domain of the Irminger Sea. The vertical propagation is only due to the adjoint equations of the primitive equation model. It does not come from specific choices in the first guess error covariance matrix since, here, it is diagonal. In the case of sequential assimilation methods, specific developments need to be introduced to project the surface observations downwards. For instance, in an eddy-permitting model with

the same horizontal resolution, Killworth et al. (2001) projected vertically the surface geostrophic velocities derived from sea level anomalies through use of vertical regression coefficients between velocities at a depth, z , and surface geostrophic velocities (Oschlies and Willebrand, 1996). The projection is then statistical and may vary with time. When an adjoint model is used, the fluctuations of the sea surface are dynamically related to velocities and density (heat and salt content) at every time step like in the concurrent forward model. As suggested by the temporal trend in model-true MADT misfits (Fig. 6), the assimilation still extracts useful information from the observations, even after 1 year of assimilation. One should, however, note that improvements in the tracer and the velocity fields are usually decreasing from one cycle to the next. Although being unrealistic, the analysis made through assimilation of perfect MADTs (black curves of Fig. 8) gives significantly better results than those made with noisy MADTs; the improvement rate from one cycle to the next is always higher with perfect MADTs (not shown on Fig. 8 for clarity). This illustrates the potential offered by future improvements in both altimeter measurements and definition of a better geoid.

In order to gain more insight into the usefulness of extending the length of assimilation windows, the results obtained with a 180-day-long window (Dir180) were compared to those issued from the first two cycles of the Dir90 experiment (Fig. 8cd). With a 180-day window composed of 17 assimilated MADTs, the assimilation procedure is still able to modify the initial conditions to better fit the model trajectory with the observations. However, the rms of the differences in temperature and velocity with the reference state are 15% greater than those obtained with a 90-day-long window. The criterion to stop the minimisation iteration procedure was alike for Dir180 and Dir 90. Adding extra iterations caused no significant change in the results (even the use of a cold restart for which the history of cost function gradient is lost). The Dir180 was not carried on over a year. Such an experiment was performed in the real data assimilation section for a 120-day long window.

Examination of the latitudinal distribution of the misfits between the model and the reference trajectory shows a significant improvement of the model state by assimilation whenever large errors are found in the simulation without assimilation (Fig. 9). In particular, the large errors associated with the meso-scale activity in the Gulf of Mexico around 24°N and with the Gulf Stream and its North Atlantic Drift north of 30°N are significantly reduced by a factor up to three both in temperature and velocity. In a large equatorial region, there is no improvement, on average, for the temperature field (Fig. 9a). The SSH in the equatorial band is strongly modulated by winds and associated waves. Since there is no error in the wind stress, initial MADT misfits are below ± 3 cm along the Equator (Fig. 7a) and weaker than at higher latitudes. The assimilation then gets easily rid of these small initial misfits and ends up with MADT residuals below ± 1 cm (Fig. 7b) with only small changes in the temperature and salinity fields. Improvements in the velocity field are more homogeneous in latitude than those in the temperature field, even close to the Equator (Fig. 9b for the meridional velocity). The assimilation is unable to improve the oceanic state north of 60° of latitude. With a different method, Killworth et al. (2001) pointed out that, in regions without available observation, the assimilation failed to improve their eddy-permitting model. In our configuration, no data is assimilated north of 66° N. The error is reduced by the restoring conditions along the artificially closed boundary at 70°N, but this remains geographically limited to the restoring region (north of 69°N). Since the assimilated model state stays close to the simulation without assimilation north of 66°N, it tends to counteract the benefits of the assimilation close to that boundary where observations are available.

4 Assimilation of real MADTs for the period 2000-2002

Our twin experiments showed that enhancement in the quality of the analyses issued from MADT

assimilation can be expected over the first 12 months. This finding led us to carry out an assimilation exercise with only real MADTs for the period 2000-2002 to get an improved ocean state for further assimilation of Argo float data (subsequent paper). Although being useful to make adjustments and gain some insights into the method, twin experiments are always limited by their incomplete set-up. They may overestimate the performance of assimilation scheme in regard to the assimilation of real data. This led us to compare the results of twin experiments with those from real assimilation experiments so as to quantify the impact of error sources (model and surface forcing errors, error in the initial conditions) neglected in our previous use of synthetic data. The use of real MADTs follows the approach defined by the twin experiments: AVISO MADTs (v2.1.0) are assimilated over cycles of 90 days without using the QSVA method (Table 1). Each cycle is composed of 13 MADTs (one map per week). The restoring towards SST fields is active throughout the assimilation process. Like in the twin experiments, the diagonal of the background error matrix, B , corresponding to a given state variable depends on the model depth. For a given model depth, the corresponding element in B for temperature (resp. salinity) is calculated as the squared inverse rms of the differences in temperature (resp. salinity) between the assimilation-free model and hydrographic data (Argo floats, CTD cruises) over the period 2002-2006 (Fig. 2, dashed lines). The vertical profiles of these rms indicate that the errors in the first guess for real data assimilation are twice to 3-times those in the twin experiments (dashed lines against solid lines). For the velocity field lacking of available 3D observations for the period under consideration, the corresponding profiles from the twin experiments (Fig. 2, blue solid lines) are simply doubled.

Consistently with the twin experiments, the assimilation of real MADTs strongly decreases the error present in the model without assimilation (Fig. 10). Indeed, in the case of no assimilation, the model-data misfit, determined from the observation cost function, J_o , is 6 to 10 times the expected error on MADTs. It is worth recalling that the J_o -scaling error on real MADTs and the one used in the twin experiments are equal (Fig. 1). After assimilation, the model free surface is almost in the error bar since the normalised

cost function is slightly above unity over the 3-year assimilation. The fact that the normalised cost function is above unity, instead of oscillating around or slightly below that value as observed in the twin experiments, suggests a slight underestimation of MADT error. The assimilation is mostly constrained by the observation cost, J_o , which accounts for 85% of the total cost for the first assimilation window (the remaining 15% are associated with the background cost J_b), and 95% for the subsequent windows. The use of an assimilation window of 120 days for the year 2000 (Fig. 10, blue curve) led to the increase, by 17% on average, of the observation cost, J_o of the optimised model trajectory with respect to that with the 90-day-long one. Moreover, most of the domain proved to be concerned by this relative increase. The experiment with the 120-day window was, thus, ended.

The main difference between our twin experiments and the real data assimilation comes from the amplitude of the error in the model trajectory initialised from the first guess of the initial conditions (first iteration of the assimilation process). In the twin experiments, this model trajectory was initialised from the state on January 1, 2002 instead of the one on January 1, 2001. Hence the model misfit during the first iteration in the twin experiment (Fig. 7a) results from only moderate errors in the initial state, errors composed of inter-annual variability and model drift. In the context of real data assimilation, the error in the model trajectory during the first iteration (Fig. 11a) comes from both the presence of large errors in the initial state taken from a climatology and the accumulation of model and surface forcing errors during the 5-year-long spin-up. One should note that a longer spin-up period would have led not only to a better quasi-steady oceanic state, but also to a larger accumulation of model error, which is not desirable. As expected, real data assimilation starts from an oceanic state with larger errors than in the twin experiments, especially at large-scales. The normalised cost function on observations before application of any assimilation treatment shows that only 30% (vs 54% for the twin experiments) of the model free surface falls within the error bar on the observations (Fig. 11c). Despite this disadvantage, the assimilation strongly reduces the large scale biases (Fig. 11b). The largest model misfit associated with a

meso-scale features in the Gulf Stream path is also decreased from 1 to 0.6 m. This behaviour is consistent with the twin experiments. After assimilation, 69% (72% for the twin experiments) of the model free surface fall within the error bar on observations (Fig. 11d). The Gulf Stream region (75-40°W, 35-50°N) exhibits local patches where misfit residuals are twice the MADT error (i.e. normalised cost function equals 4 to 5 on Fig. 11d). This is not surprising since the model dynamics fails to resolve the energetic eddies of that region. In the Gulf Stream region, free surface misfits in the assimilation-free state are 30% greater than the rms of the SSH anomalies (19 cm) as diagnosed from AVISO maps. On the other hand, in the assimilated solution they are 20% lower. Though the estimated solution is not free of model error, it proved to be able to partly extract the SSH signal, even in the Gulf Stream region. Hence, the 4D-variational method tested and refined in the twin experiments does not fail when applied to real data assimilation despite larger errors in the first guess and the presence of model and surface forcing errors.

In agreement with the twin experiments, improvements in the model free surface also result in a better ocean stratification. This is evidenced by comparing the model against the Argo float network (Fig. 12). In order to get a significant number of Argo profiles covering a broad part of the model domain, the comparison is made for the year 2002 for which 4661 profiles passed our quality control (Fig. 12a). The unavailability of the exact oceanic sea state averaged on the model grid led us to estimate the representativity error associated with Argo profiles (Fig. 12bc, black lines). This representativity error is a vertical average of local standard deviations calculated with Argo temperature and salinity profiles that are within the same model grid cell and within the same period (± 10 days). It represents the mean sub-grid scale variability of Argo profiles that the model is unable to represent. Hence any significant departure from this representativity error can be interpreted as errors in the representation (properties, geographical distribution) of model water masses. The largest errors are found below the mixed layer and decrease with depth. Further to MADT assimilation, the improvement in temperature reaches 0.5°C over the first 500

m and starts to gently decrease with depth (Fig. 12b). The salinity is improved by 0.025 psu over the first 1200 m and by 0.06 psu just below the mixed layer (Fig. 12c). Below 1500 m improvements in both temperature and salinity are slight. It is worth noting that the different behaviours in temperature and salinity errors in the mixed layer are due to the restoring condition applied to the surface temperature while the surface salinity evolves freely. MADT assimilation causes a 20 to 40% (resp. 10 to 30%) reduction of the error on temperature (resp. salinity) over the first 1400 m with respect to the representativity error. Such improvements are consistent with those found in the twin experiments (Fig. 8a). The results of our twin experiments (Fig. 8b) let us expect a significant enhancement of the velocity field over the model domain by AVISO MADT assimilation. But providing evidence of this increase is impossible in the case of real data assimilation since it would require observations at every depth on the whole domain.

Evidence of improvements in the surface velocity field is provided by a comparison against the velocity field derived from sea surface height. Figure 13 shows the annual averaged geostrophic kinetic energy calculated from AVISO MADTs and from the model free surface for the year 2002. The assimilation-free model exhibits errors typically observed in eddy-permitting models: a detachment of the Gulf Stream from the coast to far north between 39° and 40°N, a well defined Azores current surface signature from 60°W up to Gibraltar Strait, a symmetric distribution of kinetic energy on both sides of the Rykjanes Ridge, too much kinetic energy in the southern part of the Gulf of Mexico associated with the regular production and trajectory of eddies from the Loop Current (Fig. 13b). These characteristics are in disagreement with the kinetic energy observed with altimeter data (Fig. 13a). After assimilation (Fig. 13c), the surface velocity field and its associated kinetic energy is more consistent with the observed MADTs. The Gulf Stream separates from the coast further south between 37° and 38°N. There is still a slight overshoot at the coast, but the Gulf Stream path East of 73°W closely follows its imprint on the MADTs. The meanders seen on the averaged MADTs between 65°W and 55°W are also reproduced in the assimilated solution. In

the assimilated experiment, the Azores current is much closer to observations than it was in the first guess solution. The asymmetry along the Reykjanes Ridge is also marked in the assimilated solution. On the shortcoming side of the assimilated solution, one can note a lack of kinetic energy in the Gulf Stream East of 65°W (but still better than without assimilation) and in the Northwest Corner. Though the consistency observed between the state with assimilation and the observations was expected since the model dynamically adjusts its currents to the signal extracted from the assimilated MADTs, it is satisfying to note that, at least, the final performance is close to the one from the twin experiments despite the model- and surface forcing-errors.

The comparison of the 2002 annual mean kinetic energy from “eddies” in the altimeter data and the model state without assimilation highlighted major discrepancies. “Eddies” are defined as fluctuations from the annual mean of 2002 . The main discrepancy is the underestimation by 5 to 10 of the eddy kinetic energy (EKE) along the Gulf Stream path, its North Atlantic Drift and in many other parts of the subtropical gyre (Fig. 14ab). The Irminger Basin and the Labrador Sea also lack EKE. These imperfections are typical of eddy-permitting models. The assimilation partly compensates for those imperfections: the level of EKE is higher and broader distributed in the Gulf Stream path, though being still 4 to 6 fold less than in the observations (Fig. 14ac). Moreover, the EKE of the Azores Current is now present and relatively consistent with observed levels, and there is globally more EKE in the subtropical gyre. However the Irminger Basin and Labrador Sea are still lacking EKE.

In order to quantify the changes brought by the assimilation in key quantities such as the overturning circulation, we compared the model state to the Ovide 2002 hydrographic cruise. During this cruise, more than 90 CTD casts took place between June 18 and July 10, 2002 from Greenland to Portugal (Fig. 12a, red line). Comparisons between CTD data and model- temperature and -salinity averaged over the cruise period highlight the same type of improvements on the first 1500 m after MADT assimilation (Fig. 15ab). The large error in the salinity profile between 800 and 1400 m comes from the influence of the

Mediterranean overflow, whose main core is correctly located at 1000-1200 m, but with too much salt. As transport of the overturning circulation across the section was computed from an inverse stationary model constrained by ADCP data (see Lherminier et al. 2007 for details), we compared this independent estimate calculated as a function of potential density to our model estimate before and after assimilation (Fig. 16). The estimate by Lherminier et al. (2007) showed a maximum of 16.2 ± 1 Sv at $s_1=32.2$ kg m⁻³. In the model, this maximum is increasing from 13.3 Sv at $s_1=32.1$ kg m⁻³ before assimilation to 15.2 Sv at $s_1=32.15$ kg m⁻³ after assimilation. It is worth noting that, in our model, there is no net transport across the section since the model domain is closed at 70°N conversely to the finding of 1-Sv northward transport in the study by Lherminier et al. (2007) (see their 3.4 section). Thus, the assimilation of MADTs provides an analysis of the overturning circulation consistent with the independent estimate; it results from the improvement of the 3D velocity field shown in the twin experiments.

Another important key quantity for the oceanic circulation is the heat transport. In the model without assimilation, the maximum of the 2001-2002 annual mean heat transport is weak and broadly extended in latitude from 5°N to 30°N (Fig. 17a). This is expected from an unconstrained model at this resolution (e.g. Willebrand et al. 2001). The maximum reaches 0.94 pW at 13°N, whereas independent estimates based on inversions of hydrographic data usually showed a maximum close to 1.2-1.3 pW (Fig. 17a, Ganachaud et al. 2003; Macdonald 1998). After MADT assimilation, the analyses show a remarkable consistency between the model heat transport and the independent estimates poleward of the Equator. A maximum heat transport of 1.38 ± 0.07 pW is reached at 12°N. The 0.07 pW corresponds to the inter-annual variability of the annual mean heat transport for the period 2001-2002 (Fig. 17a, shaded region). At 24°N, where the 2001-2002 inter-annual variability is weak, our heat transport of 1.21 ± 0.01 pW is closer to the estimate by Ganachaud et al. (2003) than to the one by Macdonald (1998). On the other hand, it is consistent with the error bars found by both teams. Other estimates at this latitude are in favour of a value greater than 1.2 pW at 24°N (see Fig. 3 in Ganachaud et al. 2003). South of the

Equator, the agreement is less obvious though the model heat transport remains consistent with the error bar produced by Ganachaud et al. (2003). If the model had to fail here, it might be due to the presence of restoring boundary conditions at 20°S. However, other independent estimates (see Fig. 3 in Ganachaud et al. 2003) show a heat transport close to 0.6 pW south of the Equator, which is in agreement with our analyses. These comparisons have to be considered with care: indeed, the years under study in the investigations at the origin of these independent estimates were based on the WOCE period (early 90's).

5 Conclusions

This study showed that a 3-month-long window is an efficient time-scale to identify the initial state and provides an optimised model trajectory consistent with maps of absolute dynamic topography in our eddy-permitting configuration. The use of a short window minimises the accumulation of model errors and limits the problem due to model non-linearities. The price to pay with short assimilation windows is an increase in the number of discontinuities in the optimised model trajectory. This drawback has, however, to be moderate given that the shorter the assimilation window length is, the smaller the amplitude of the discontinuities is because of a reduction in the model error accumulation to be corrected. A comparison of analyses issued from short- and long-assimilation windows would enable one to determine whether shortcomings of long windows (with increased viscosity and diffusivity coefficients) are weak enough to take advantage of the increased number of observations and time scales.

Discontinuities are induced by changes in the initial state, which are not in dynamical balance. The adjustment of the velocity field with the stratification does not create non physical jumps in key quantities such as the heat transport. For instance, at the beginning of an assimilation cycle, there is no sign of spurious anomalies in the time evolution of the heat transport at 24°N between 2001 and 2002 diagnosed from 15-day-long averages (Fig. 17b). On the other hand, the time variability of the heat transport in the estimated solution remains closely related to the solution without assimilation. The assimilation treatment

has, however, affected the intensity of the heat transport. The lack of jumps from the temporal averages under study means that the changes in the initial states induced by assimilation and observed between two consecutive cycles are sufficiently small to limit their signature in time. Moreover, from one cycle to the next one, the features of the assimilated observations, i.e. their dynamical content, number and distribution, remained alike. Since the performance of the assimilation is not dependent on the number of cycles as shown in Fig. 10, this means that one can expect to get a sufficiently good dynamical consistency between two adjacent cycles

The QSVA method did not prove to be more efficient than a simple direct assimilation method (Dir90). Both approaches lead to observational residuals that are consistent with the error bar, which is what is expected from an assimilation scheme. It means that the $1/3^\circ$ model is sufficiently linear over a 90-day long period. QSVA might be more advantageous for higher resolution assimilation configurations with increased non-linearity. Furthermore, one should note that this study was restricted to selected possibilities of a window length; starting from a window shorter than 30 days and/or increasing the window more progressively could be more efficient.

Our twin experiments showed that the MADT assimilation reduced not only large scale biases, but also meso-scale errors in the temperature, salinity and velocity fields. Further to major improvements, a depth of 1500 m was reached for tracers and velocities. Such a vertical scale is consistent with a constraint of the first baroclinic mode by altimeter data. This is true for most of the ocean where the variability of the main thermocline is dominant. Despite the easier framework of the twin experiment (no surface forcing and model errors, smaller errors in the initial conditions), real data assimilation exhibited the same improvements for the tracer fields. The integrated transports along the Ovide 2002 hydrographic section is also consistent with the independent estimate by Lherminier et al. (2007). This finding is an additional evidence of the assimilation-induced enhancement in the representation of the velocity field. A comparison

of q - S regional basin diagrams in the model without and after assimilation (not shown) evidenced only weak modifications. This is consistent with the results of our comparison to Argo CTD profiles (Fig. 12bc), i.e. the keeping of residual errors twice to 3-fold the representativity error despite improvements in fields compared to the case without assimilation. This means that most of the improvements observed in the temperature and salinity fields are associated with a better position of the structures (fronts, meso-scale) rather than with improved water mass characteristics. Therefore, the assimilation of in situ temperature and salinity profiles from the Argo network and from hydrographic cruises is necessary to improve the water mass properties.

Acknowledgments

I would like to thank Anthony Weaver who provided the core of the assimilation system and some assistance at the start of the project, Catherine Kermabon for help on the quality control of the Argo data, Herlé Mercier for discussions and support and anonymous reviewers for their comments and suggestions. This study received support from the LEFE/ASSIM "Les enveloppes fluides et l'environnement / Assimilation" programme, from the Mercator project and from IFREMER. The IDRIS (Institut du Développement et des Ressources en Informatique Scientifique) supercomputer facilities were used to make the calculations.

Table

| Twin experiments | | | | | |
|-------------------------|--------|-------------------|-----------------------------|------------------|-----------------------|
| Name | Type | MADT Observations | Window length (days) | Number of cycles | Total duration (days) |
| Dir30 | Direct | S+N | 30 | 3 | 90 |
| Dir90 | Direct | S+N | 90 | 4 | 360 |
| Dir90P | Direct | S P | 90 | 4 | 360 |
| Dir180 | Direct | S+N | 180 | 1 | 180 |
| QSVA | QSVA | S+N | Extended from 30 to 90 days | 1 | 90 |
| Real Data | | | | | |
| | Direct | AVISO 2000-2002 | 90 | 12 | 1080 |
| | Direct | AVISO year 2000 | 120 | 3 | 360 |

Table 1: List of twin and real data assimilation experiments. For twin experiments, maps of absolute dynamic topography (MADTs) are noisy (N) or perfect (P) maps of synthetic data (S).

Figure captions

Figure 1: a) Error on the mean sea level (Rio et al. 2005), b) error on maps of sea level anomaly (m). The latter is calculated as the rms of the differences between Jason-ENVISAT maps and tracks from Topex/Poseidon and Jason-1 over the year 2004. Note that, north of 66° latitude, data are unavailable.

Figure 2: rms of the differences between the first guess (January 1, 2002) and the true state (January 1, 2001) for each level model (solid lines); rms of the differences between the model without assimilation and hydrographic data from years 2002-2006 (dashed lines). Lower (resp. upper) abscissa gives the temperature scale (T: temperature) (S: resp. salinity; u: zonal velocity and v: meridional velocity).

Figure 3: Example of noise (m) added to a synthetic sea surface height map for the twin experiments. Maximum values reach ± 60 cm in the Gulf Stream region (colours saturated at ± 15 cm for clarity).

Figure 4: rms of the differences between true MADTs (no noise) and their model equivalent over the first 90 days (one map every 10 days) when the model equivalent is taken from the experiment: without assimilation (black line), with 3 cycles of 30-day-long assimilation period (blue), with 1 cycle of 90-day-long assimilation period starting from the optimised state of the first 30-day assimilation cycle (QSVA, red), with 1 cycle of 90-day-long assimilation period without prior assimilation (Dir90, green). Note that the blue line is discontinuous at the beginning of a cycle (first assimilated MADT) since the assimilation is restarted every 30 days. The dotted line is a forecast from the optimised state of the first 30-day cycle.

Figure 5: rms of the model-reference misfit in temperature calculated over the whole domain for the beginning of the assimilation period (fortnight 1), fortnight 3 and the end of the assimilation period (fortnight 6). The model state is the analysis either from the QSVA techniques or from the first cycle of the 90-day-long assimilation period without prior assimilation (Dir90).

Figure 6: rms of the differences between true MADTs and their model equivalent over 1 year when the model equivalent is taken from the run: without assimilation (black line), with 4 cycles of 90 days that

assimilate noisy (Dir90, red) and perfect (blue) MADTs. Note that red and blue lines are discontinuous at the beginning of a cycle.

Figure 7: Model - true MADT misfit for the last 10-day period of the year 2001 (i.e. after a year of assimilation): a) model without assimilation, b) with assimilation (Dir90).

Figure 8: rms of the model-reference misfit for a) temperature and b) meridional velocity calculated over the whole domain for the last fortnight of the 4 cycles. The model state is taken from the model without assimilation (no assim.) and with assimilation of noisy MADTs (assim.) using a 90-day-long window (Dir90). The rms for the last fortnight of the 4th cycle calculated with the assimilation of perfect MADTs is also shown (black curve).

(continuation): c) same as a), d) same as b), for March and June along with the model-reference misfit using model results from an assimilation based on a 180-day window without prior assimilation (Dir180, dotted lines).

Figure 9: zonally-averaged rms of the model misfits to the reference a) temperature and b) meridional velocity over the whole domain for the last fortnight of the fourth cycle (last fortnight of 2001). The model state is taken from the model without assimilation (magenta) and with assimilation of noisy MADTs (Dir90, black).

Figure 10: MADT cost function, J_o , normalised by the error on MADTs for the model without assimilation (black), with 90-day-long assimilation cycles (red) for the period 2000-2002, and with 120-day-long assimilation cycles (blue) for the year 2000. The beginning of each cycle of assimilation is indicated by a circle. A cost of 1 denotes that the model free surface agrees with the MADTs within the error bar.

Figure 11: Model-data misfit for the last MADT of the year 2000 (i.e. after 1-year assimilation): a) model

without assimilation, b) with assimilation. Equivalent normalised cost function c) without assimilation, d) with assimilation.

Figure 12: a) Location of 4661 Argo temperature and salinity profiles (blue dots) and Ovide 2002 CTD casts (red dots) during the year 2002; rms of the differences between Argo data and the model without (blue) and with (red) assimilation of MADTs for b) temperature and c) salinity; an estimate of the representativity error on the in situ profile is also provided (black).

Figure 13: Surface geostrophic kinetic energy averaged over the year 2002 from a) Aviso sea surface height, b) model free surface without assimilation, c) same as b) but with assimilation.

Figure 14: Surface geostrophic “eddy” kinetic energy averaged over the year 2002 from a) Aviso sea surface height, b) model free surface without assimilation, c) same as b) but with assimilation. “Eddy” velocities are defined as departures from the 2002 annual mean velocity field.

Figure 15: rms of the differences between Ovide2002 CTD and the model without (blue) and with (red) assimilation of MADTs for a) temperature and b) salinity.

Figure 16: Cumulated transports from the densest to the lightest water masses across the Ovide 2002 hydrographic line as a function of potential density referenced to 1000 dbar. Estimates come from the model without (blue) and with (red) MADT assimilation and from an independent inverse stationary model (Lherminier et al. 2007, black). Note that, following the model velocities, transports are positive when oriented to the north and east.

Figure 17: a) Heat transport from independent inversions of hydrographic data (Ganachaud and Wunsch, 2003; Macdonald 1998) and from an average over the period 2001-2002 in the model with assimilation and without assimilation ($1 \text{ pW} = 10^{15} \text{ Watts}$). The shaded region represents the 2001-2002 inter-annual variability of the model with assimilation. Vertical coloured lines are the error bar associated with the

independent inversions. b) Time history of the heat transport at 24°N in the model with and without assimilation. Circles represent the start of assimilation cycles; they are reported on the curve without assimilation for comparison.

References

- 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. Ocean.*, 23, 1363-1388.
- Blum J., B. Luong, and J. Verron, 1998: Variational assimilation of altimeter data into a non-linear ocean model: temporal strategies, in: ESAIM Proceedings, *Contrôle et Équations aux Dérivées Partielles.* 4, 21-57.
- Bouffard, J., Vignudelli, S., Cipollini, P. and Menard, Y., 2008: Exploiting the potential of an improved multimission altimetric data set over the coastal ocean. *Geophysical Research Letters*, 35, (10), L10601. (doi:10.1029/2008GL033488)
- Brasseur, P., E. Blayo, and J. Verron, 1996: Predictability experiments in the North Atlantic Ocean: Outcome of a quasi-geostrophic model with assimilation of TOPEX/POSEIDON altimeter data, *J. Geophys. Res.*, 101(C6), 14,161–14,173.
- Ferron, B., and J. Marotzke, 2003: Impact of 4D-variational assimilation of WOCE hydrography on the meridional overturning circulation of the Indian Ocean, *Deep-sea Res. II*, 50, 2005-2022.
- Forget, G., 2005: Profils Argo et assimilation 4D-var pour le suivi climatique de l'océan Nord Atlantique, Thèse de Doctorat, n°1059, Université de Bretagne Occidentale, 143pp.
- Forget, G., H. Mercier, and B. Ferron, 2008: Combining Argo profiles with a general circulation model in the North Atlantic. Part 2: Realistic transports and improved hydrography, between spring 2002 and spring 2003, *Ocean Mod.*, 20, 17-34, doi:10.1016/j.ocemod.2007.06.002.
- Ganachaud, A. and C. Wunsch, 2003: Large scale ocean heat and freshwater transports during the World Ocean Circulation Experiment, *J. Climate.*, 16, 696-705.
- Gebbie, G., P. Heimbach, and C. Wunsch, 2006: Strategies for Nested and Eddy-Permitting State Estimation, *J. Geophys. Res.*, 111, C10073, doi:10.1029/2005JC003094.
- Hoteit, I., B. Cornuelle, A. Köhl, and D. Stammer, 2006: Treating strong adjoint sensitivities in tropical eddy-permitting variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131 (613), 3659-3682.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S-K Yang, J.J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DEO AMIP-II Reanalysis (R-2): 1631-1643, *Bul. of the Atmos. Met. Soc.*, 83, 1631-1643. DOI: 10.1175/BAMS-83-11-1631.
- Killworth P., C. Dieterich, C. Le Provost, A. Oschlies and J. Willebrand, 2001: Assimilation of altimetric data into an eddy-permitting model of the North Atlantic, *Progress in Oceanog.* , Vol 48, 123-162.
- Köhl, A., and J. Willebrand, 2002: An adjoint method for the assimilation of statistical characteristics into eddy-resolving ocean models, *Tellus*, 54A, 406–425.
- Köhl, A., and J. Willebrand, 2003: Variational assimilation of SSH variability from TOPEX/POSEIDON and ERS1 into an eddy-permitting model of the North Atlantic, *J. Geophys. Res.*, 108,

doi:10.1029/2001JC000982.

S. Laroche and P. Gauthier, 1998: A validation of the incremental variational data assimilation in a two-dimensional turbulent flow, *Tellus*, 50A, 557-572.

Lea, D. J. H. W. N. Thomas, and R. F. Gasparovic, 2006: Observability of the irvinger sea circulation using variational data assimilation, *Quart. J. Royal Met. Soc.*, 32, 1545-1576.

Le Dimet F.-X., Talagrand O., 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects, *Tellus* 38A, 97-110.

Lherminier, P., Mercier, H., Gourcuff, C., Alvarez, M., Bacon, S., Kermabon, C., 2007. Transports across the 2002 Greenland–Portugal Ovide section and comparison with 1997, *J. Geophys. Res.* 112 (C07003), doi:10.1029/2006JC003716.

Macdonald, A. M., 1998: The global ocean circulation: A hydrographic estimate and regional analysis. *Progress in Oceanography*, Vol. 41, Pergamon, 281-382

Madec, G., P. Delecluse, M. Imbard, and C. Lévy, 1998: OPA 8.1 Ocean General Circulation Model reference manual. Note du Pôle de modélisation, Institut Pierre-Simon Laplace, N°11, 91pp.

Madsen, K. S., J. L. Høyer, and C. C. Tscherning, 2007: Near-coastal satellite altimetry: Sea surface height variability in the North Sea–Baltic Sea area, *Geophys. Res. Lett.*, 34, L14601, doi:10.1029/2007GL029965

Masuda, S., T. Awaji, N. Sugiura, Y. Ishikawa, K. Baba, K. Horiuchi, and N. Komori, 2003: Improved estimates of the dynamical state of the North Pacific Ocean from a 4 dimensional variational data assimilation, *Geophys. Res. Letters*, 30, NO. 16, 1868, doi:10.1029/2003GL017604, 2003

Mazloff, M., 2008: The Southern Ocean meridional overturning circulation as diagnosed from an eddy permitting state estimate. Ph.D. thesis, Massachusetts Institute of Technology and the Woods Hole Oceanographic Institution, Cambridge, MA.

Morrow, R., and P. De Mey, 1995: Adjoint assimilation of altimetric, surface drifter, and hydrographic data in a quasi-geostrophic model of the Azores Current, *J. Geophys. Res.*, 100(C12), 25,007–25,025.

Pires C., R. Vautard, and O. Talagrand, 1996: On extending the limits of variational assimilation in nonlinear chaotic systems, *Tellus*, 48A, 96-121.

Reynaud, T., Le Grand, P., Mercier, H., Barnier, B., 1998. A new analysis of hydrographic data in the Atlantic and its application to an inverse modeling study. *Intern. WOCENews*. 32, 29–31.

Reynolds, R. W. and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, 7, 929-948.

Rio, M.-H., Schaeffer, P., Lemoine, J.-M., Hernandez, F., 2005: Estimation of the ocean Mean Dynamic Topography through the combination of altimetric data, in-situ measurements and GRACE geoid: From global to regional studies, *Proceedings of the GOCINA international workshop*, Luxembourg.

Roullet, G., and G. Madec, 2000: Salt conservation, free surface, and varying levels: A new formulation

for ocean general circulation models. *J. Geophys. Res.*, 105, 23 927–23 942.

Schröter, J., U. Seiler, and M. Wenzel, 1993: Variational assimilation of Geosat data in an eddy-resolving model of the Gulf Stream extension area, *J. Phys. Ocean.*, 23, 925-953.

Smith, W. H. F. and D. T. Sandwell, 1997: Global Seafloor Topography from Satellite Altimetry and Ship Depth Soundings, *Science*, 277, 1956-1962.

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C.N. Hill and J. Marshall, 2002: The global ocean circulation during 1992 --1997, estimated from ocean observations and a general circulation model. *J. Geophys. Res.*, 107, C9 3118, doi:10.1029/2001JC000888

Stammer, D., C. Wunsch, R. Giering, C. Eckert, P. Heimbach, J. Marotzke, A. Adcroft, C.N. Hill and J. Marshall, 2003: Volume, heat and freshwater transports of the global ocean circulation 1993 --2000, estimated from a general circulation model constrained by World Ocean Circulation Experiment (WOCE) data . *J. Geophys. Res.*, 108, C1 3007, doi:10.1029/2001JC001115

Swanson, K.L., R. Vautard, and C. Pires, 1998: Four-dimensional variational assimilation and predictability in a quasi-geostrophic model *Tellus*, 50A, 369-390.

Tanguay, M., P. Bartello, and P. Gauthier, 1995: "Four-dimensional data assimilation with a wide range of scales." *Tellus*, 47A, 974-997.

Thacker, W. C. and R. B. Long, 1988: Fitting dynamics to data. *J. Geophys. Res.*, 93, 1127-1239.

Titau, O., A. Vidard, I. Souopgui, and F.-X. Le Dimet, 2010: Assimilation of Image Sequences in Numerical Models, *Tellus*, 62A , 30-47.

Vialard, J., A.T. Weaver, D.L.T. Anderson and P. Delecluse, 2003: Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean, Part 2 : physical validation. *Mon. Weather Rev.*, 131, 1379-1395.

Weaver, A.T., J. Vialard, and D.L.T. Anderson, 2003: Three- and four-dimensional variational assimilation with a general circulation model of the tropical Pacific Ocean, Part 1 : formulation, internal diagnostics and consistency checks. *Mon. Weather Rev.*, 131, 1360-1378.

Willebrand, J., B. Barnier, C. Boning, C. Dieterich, P. Hermann, P. D. Killworth, C. Le Provost, Y. Jia, J.M. Molines, and A.L. New, 2001: Circulation characteristics in three eddy-permitting models of the North Atlantic, *Prog. Oceanogr.*, 48, 123-161.

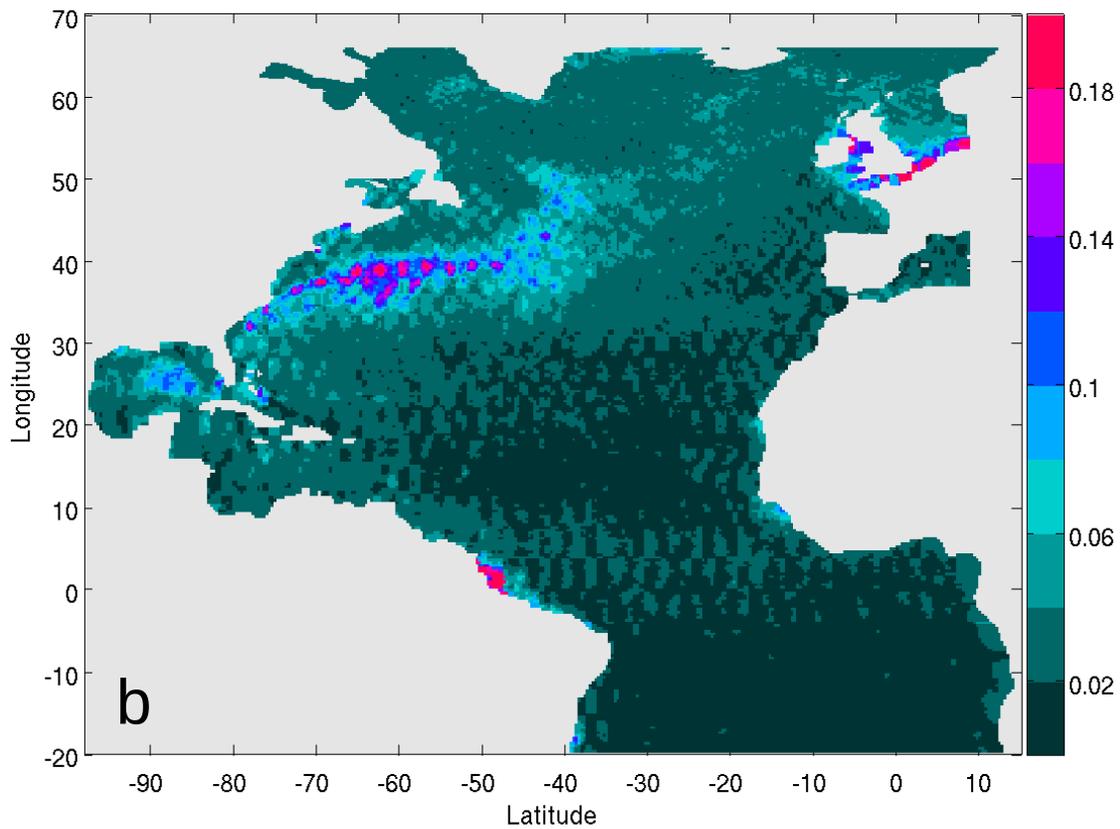
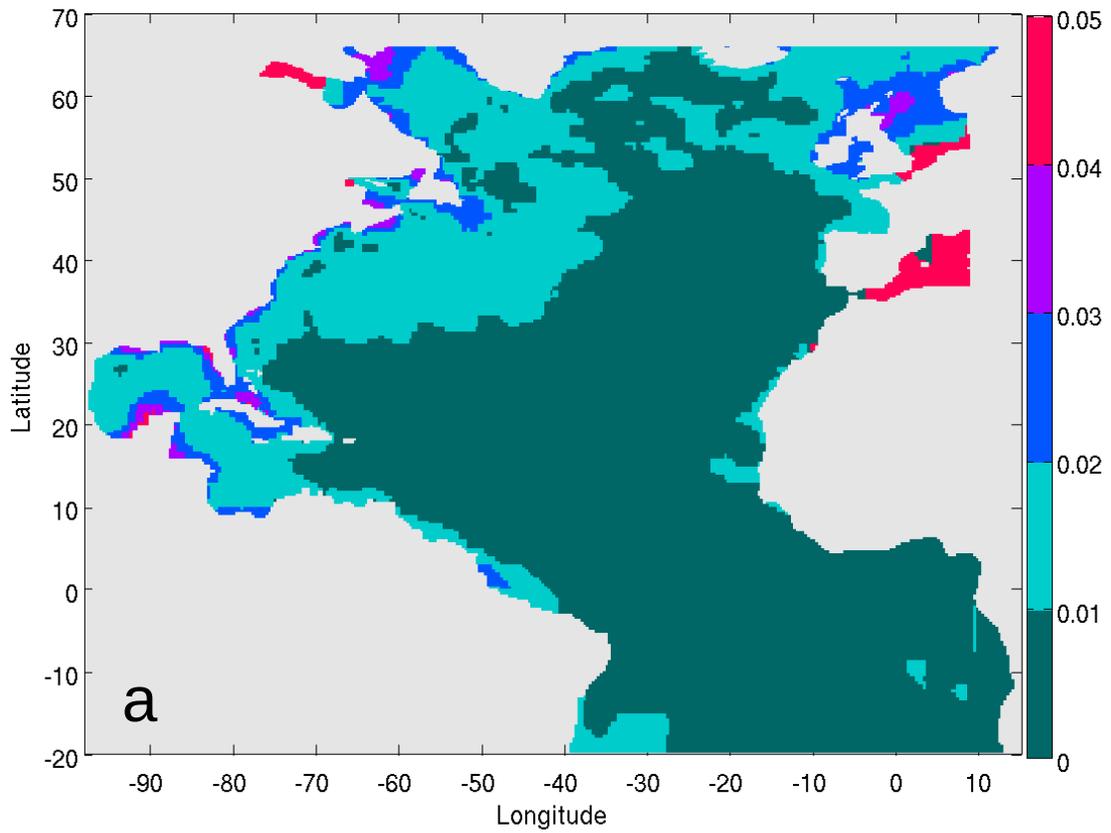


Figure 1: a) Error on the mean sea level (Rio et al. 2005), b) error on maps of sea level anomaly (m). The latter is calculated as the rms of the differences between Jason-ENVISAT maps and tracks from Topex/Poseidon and Jason-1 over the year 2004. Note that, north of 66° latitude, data are unavailable.

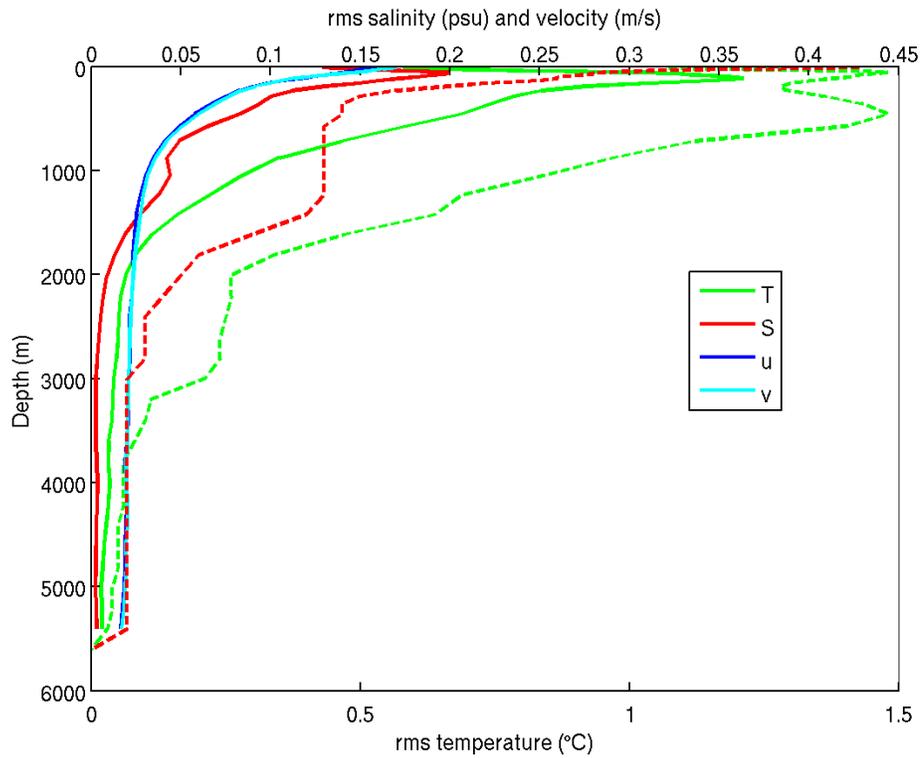


Figure 2: rms of the differences between the first guess (January 1, 2002) and the true state (January 1, 2001) for each level model (solid lines); rms of the differences between the model without assimilation and hydrographic data from years 2002-2006 (dashed lines). Lower (resp. upper) abscissa gives the temperature scale (T: temperature) (S: resp. salinity; u: zonal velocity and v: meridional velocity).

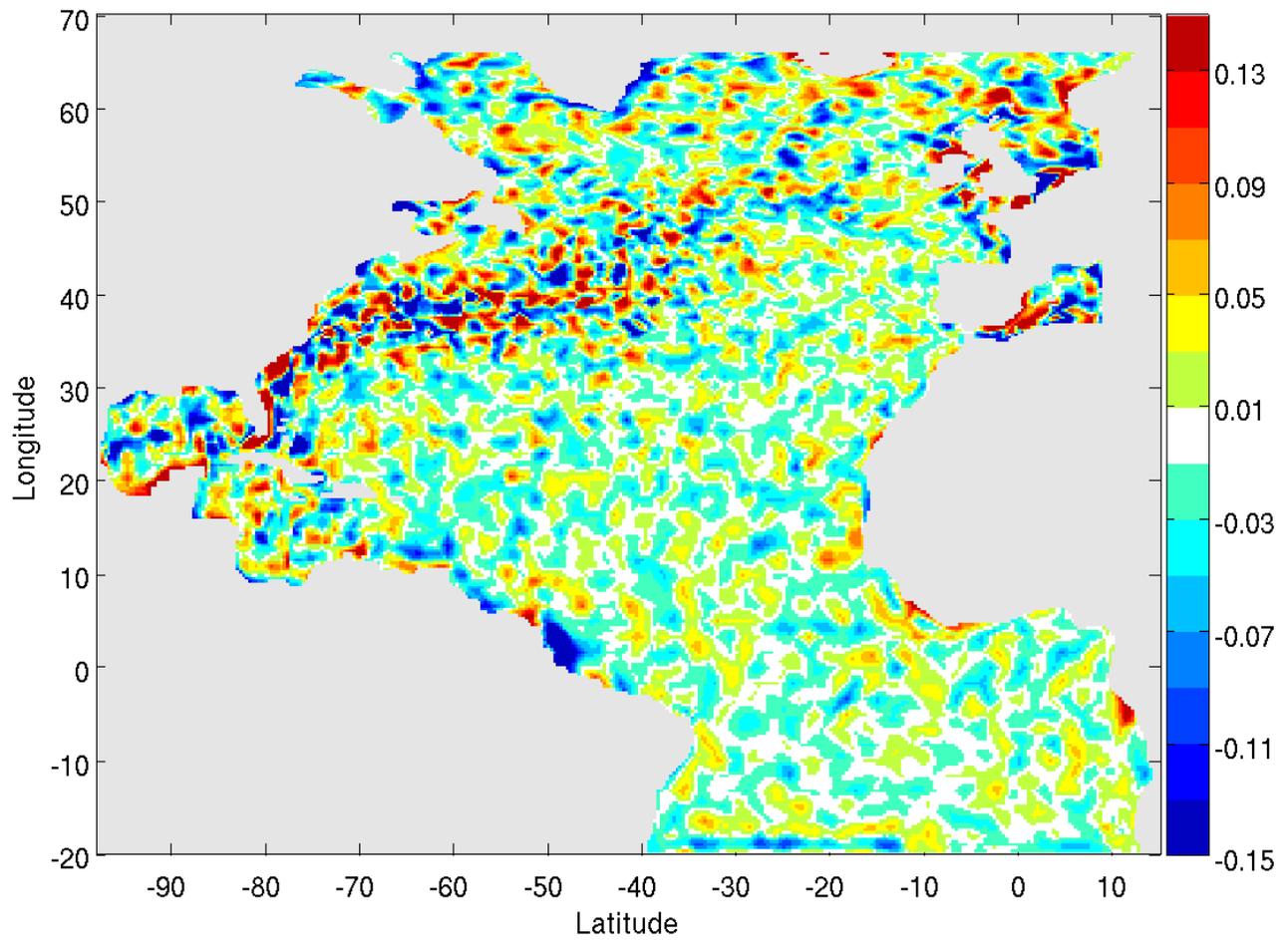


Figure 3: Example of noise (m) added to a synthetic sea surface height map for the twin experiments. Maximum values reach ± 60 cm in the Gulf Stream region (colours saturated at ± 15 cm for clarity).

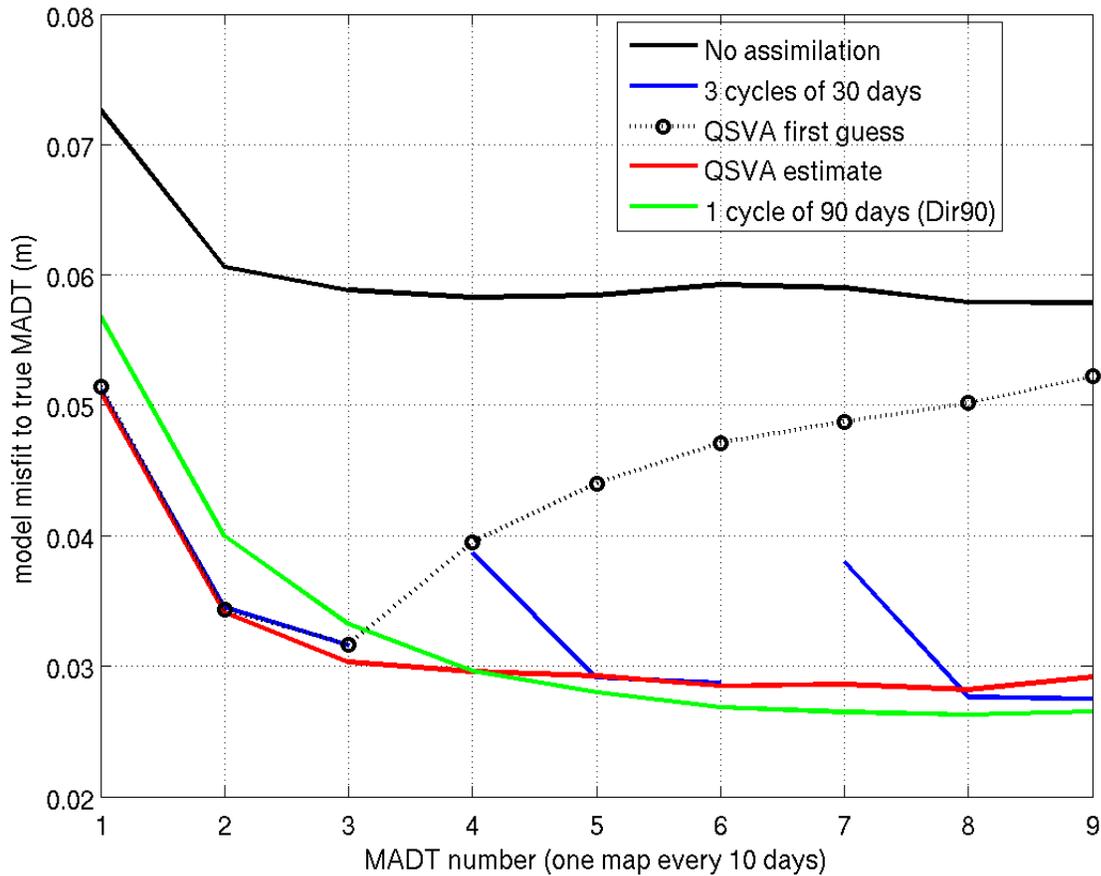


Figure 4: rms of the differences between true MADTs (no noise) and their model equivalent over the first 90 days (one map every 10 days) when the model equivalent is taken from the experiment: without assimilation (black line), with 3 cycles of 30-day-long assimilation period (blue), with 1 cycle of 90-day-long assimilation period starting from the optimised state of the first 30-day assimilation cycle (QSVa, red), with 1 cycle of 90-day-long assimilation period without prior assimilation (Dir90, green). Note that the blue line is discontinuous at the beginning of a cycle (first assimilated MADT) since the assimilation is restarted every 30 days. The dotted line is a forecast from the optimised state of the first 30-day cycle.

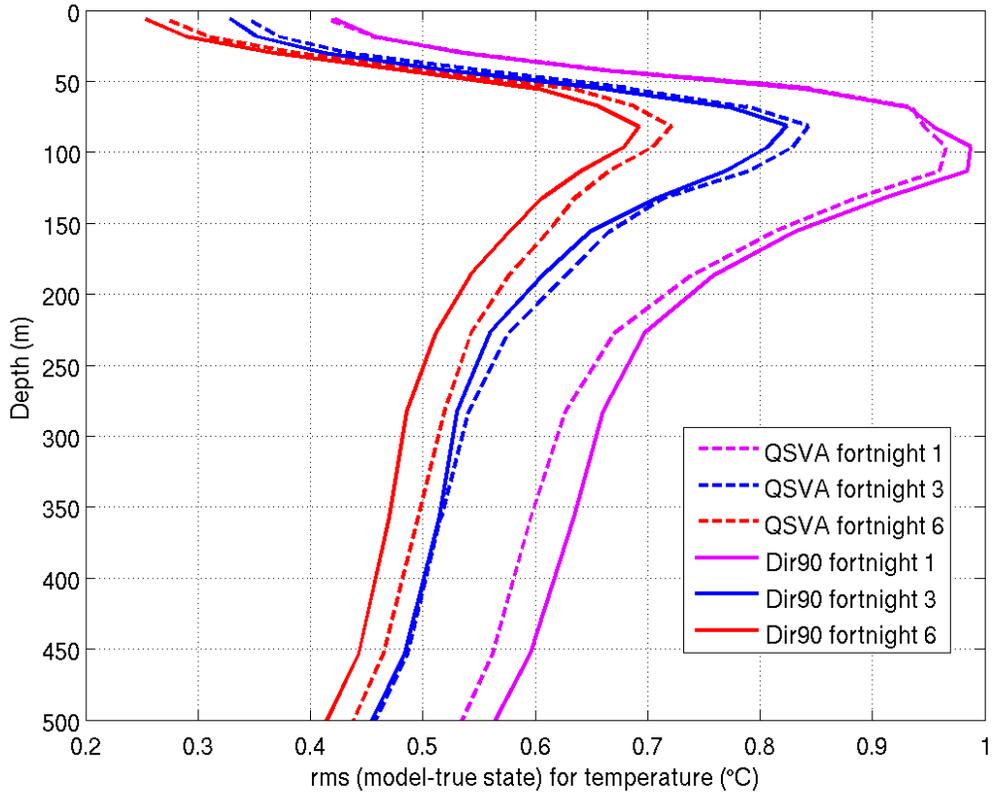


Figure 5: rms of the model-reference misfit in temperature calculated over the whole domain for the beginning of the assimilation period (fortnight 1), fortnight 3 and the end of the assimilation period (fortnight 6). The model state is the analysis either from the QSVA techniques or from the first cycle of the 90-day-long assimilation period without prior assimilation (Dir90).

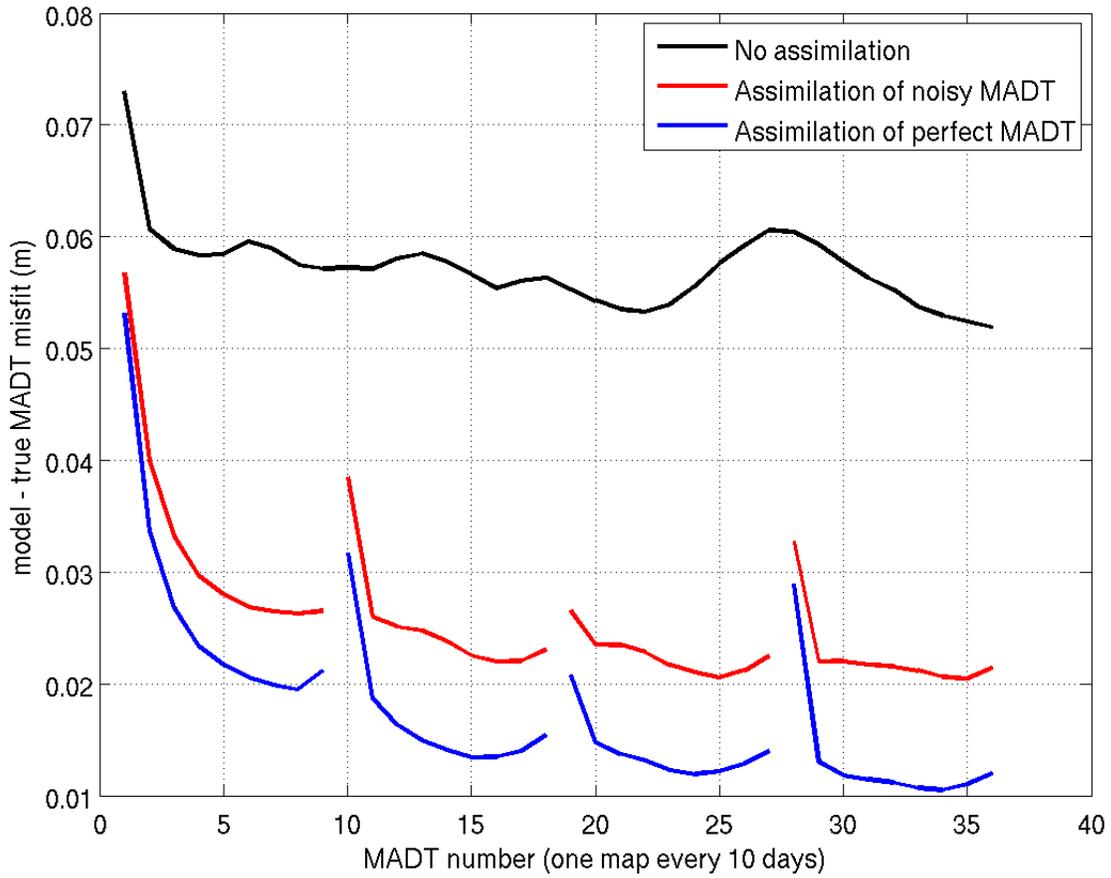


Figure 6: rms of the differences between true MADTs and their model equivalent over 1 year when the model equivalent is taken from the run: without assimilation (black line), with 4 cycles of 90 days that assimilate noisy (Dir90, red) and perfect (blue) MADTs. Note that red and blue lines are discontinuous at the beginning of a cycle.

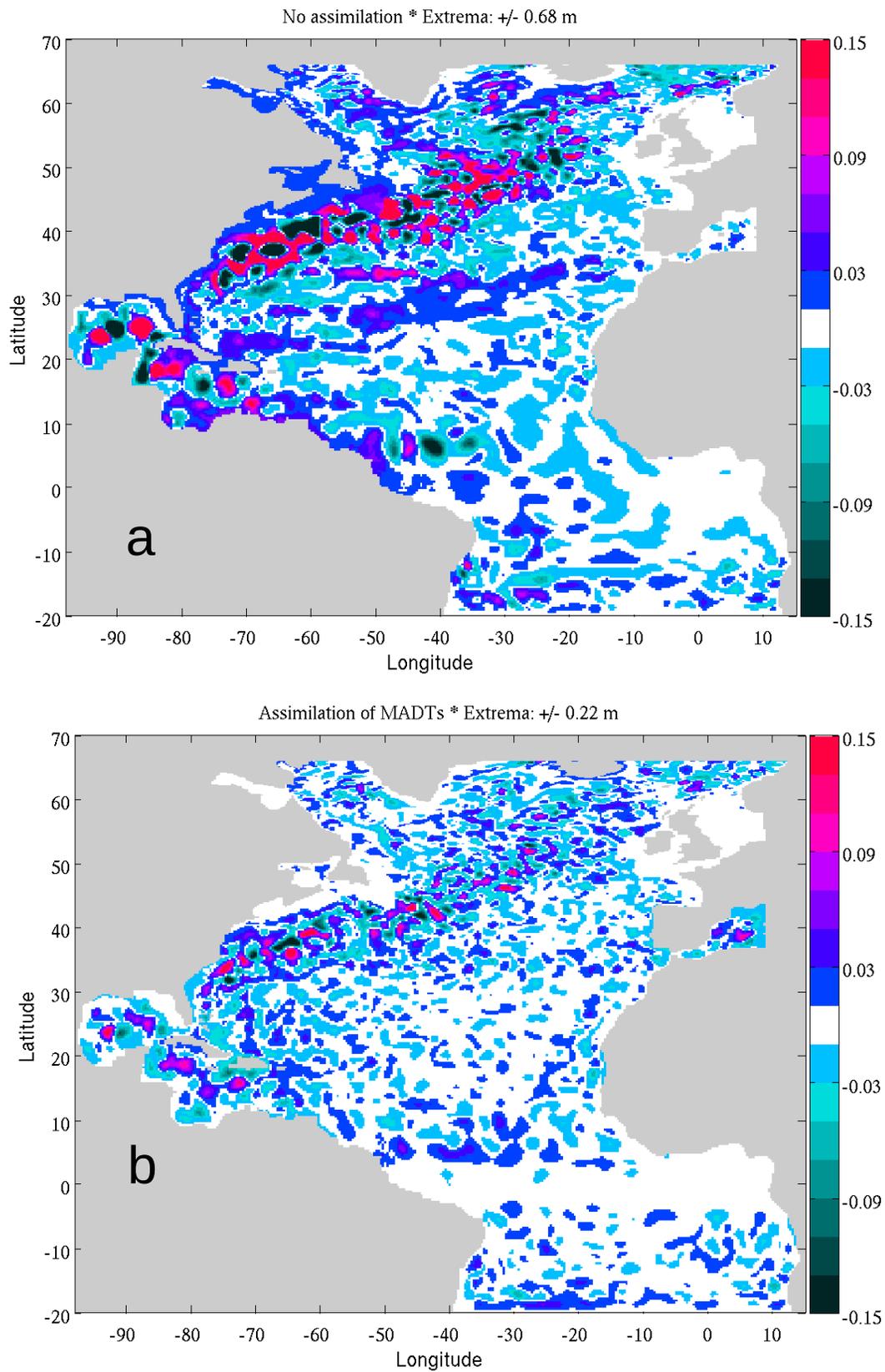


Figure 7: Model - true MADT misfit for the last 10-day period of the year 2001 (i.e. after a year of assimilation): a) model without assimilation, b) with assimilation (Dir90).

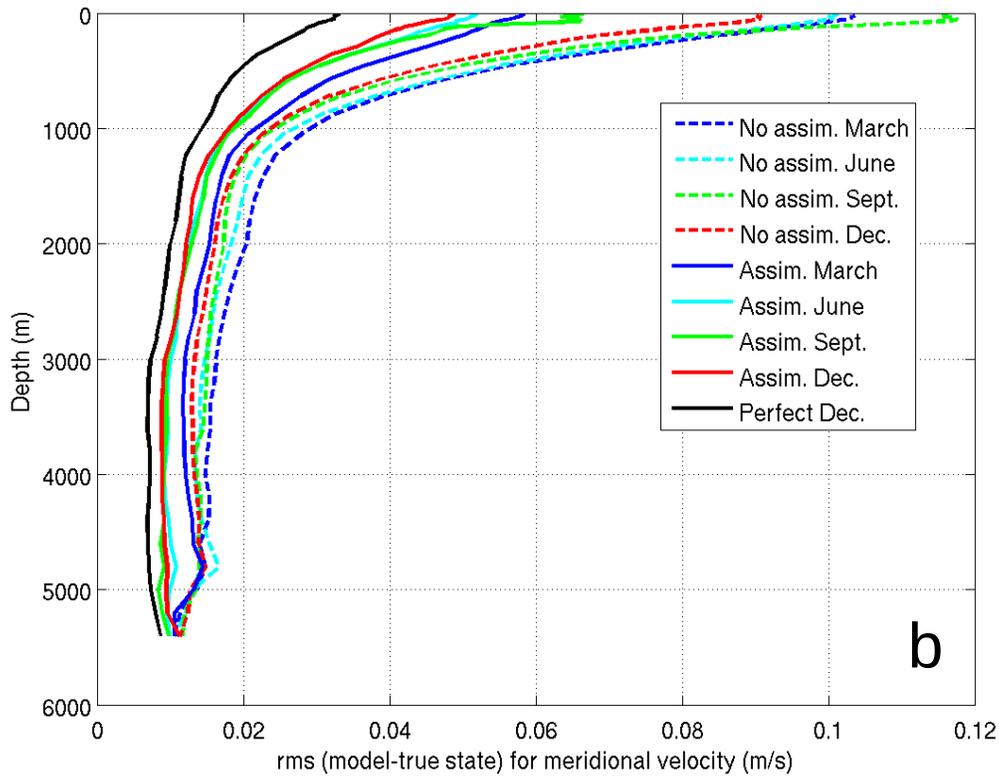
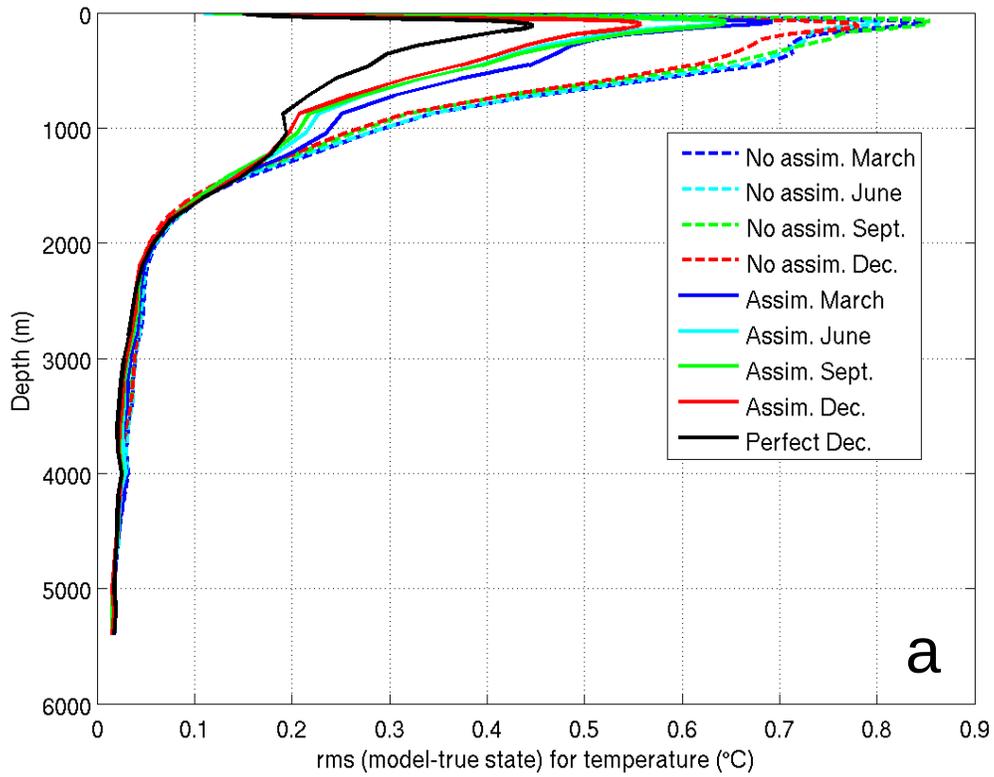


Figure 8: rms of the model-reference misfit for a) temperature and b) meridional velocity calculated over the whole domain for the last fortnight of the 4 cycles. The model state is taken from the model without assimilation (no assim.) and with assimilation of noisy MADTs (assim.) using a 90-day-long window (Dir90). The rms for the last fortnight of the 4th cycle calculated with the assimilation of perfect MADTs is also shown (black curve).

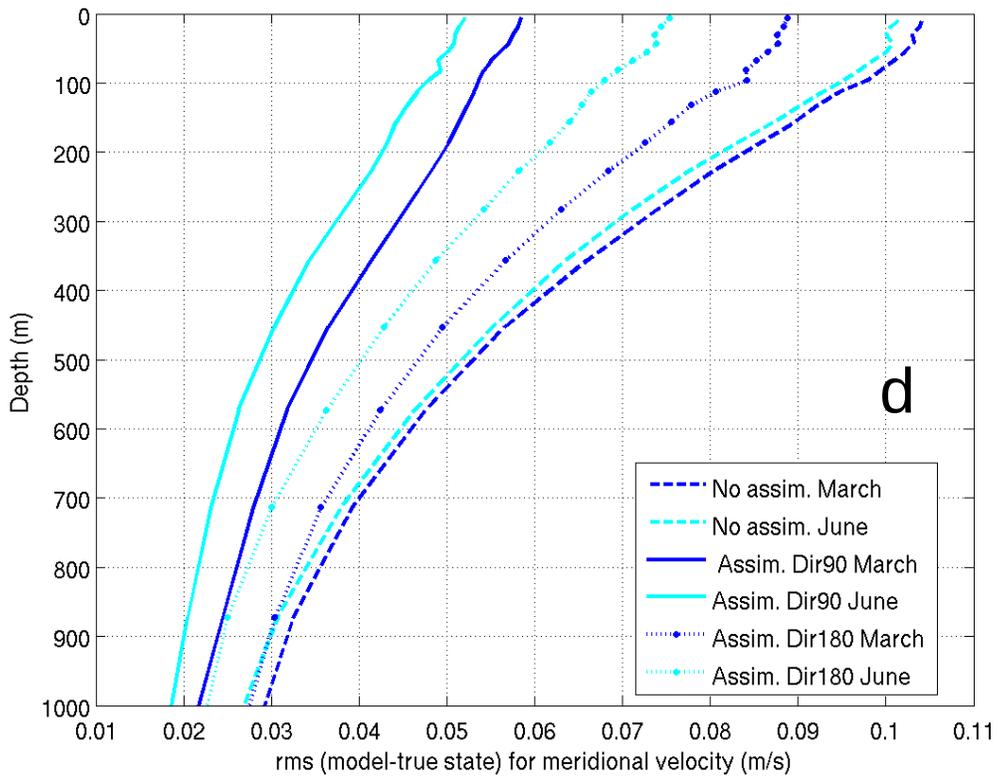
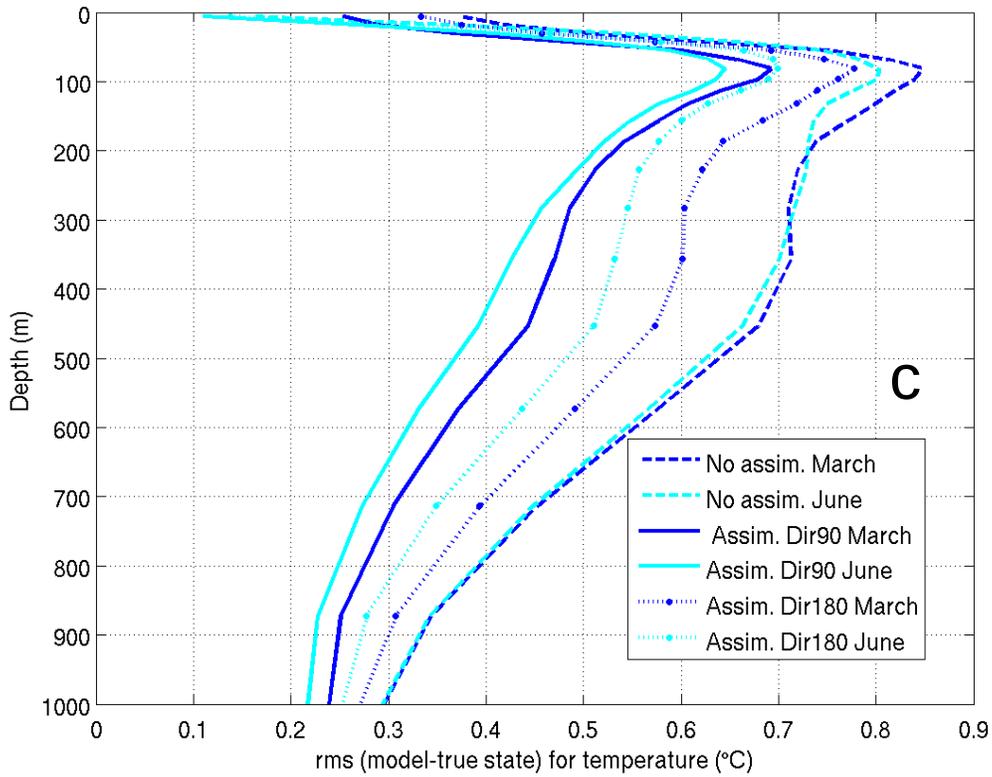


Figure 8:(continuation): c) same as a), d) same as b), for March and June along with the model-reference misfit using model results from an assimilation based on a 180-day window without prior assimilation (Dir180, dotted lines).

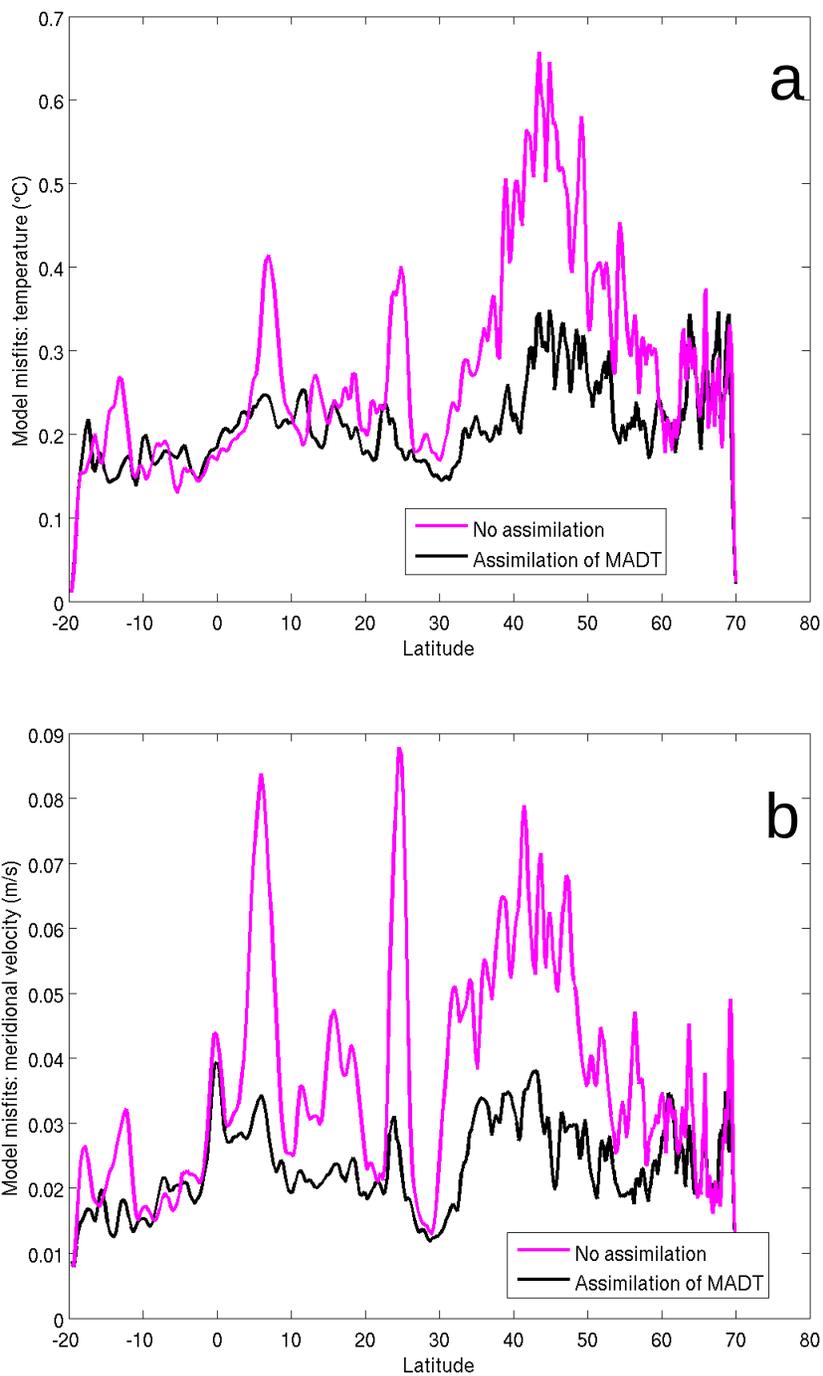


Figure 9: zonally-averaged rms of the model misfits to the reference a) temperature and b) meridional velocity over the whole domain for the last fortnight of the fourth cycle (last fortnight of 2001). The model state is taken from the model without assimilation (magenta) and with assimilation of noisy MADTs (Dir90, black).

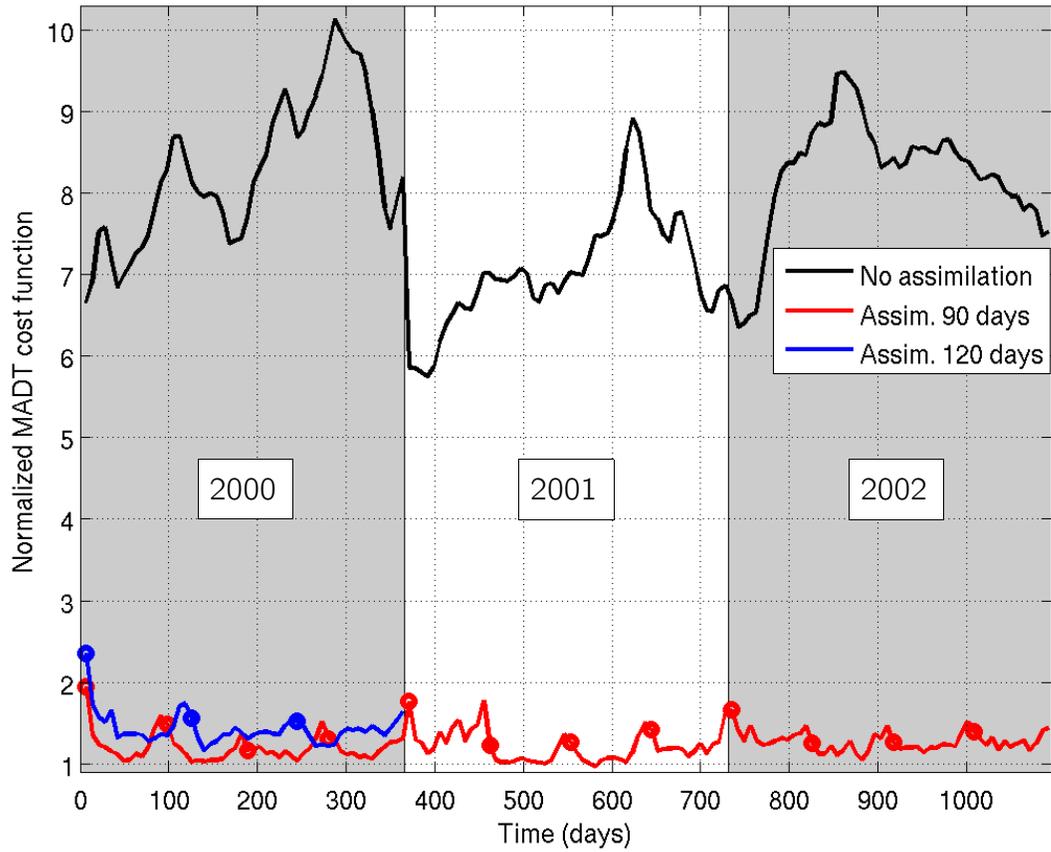


Figure 10: MADT cost function, J_o , normalised by the error on MADTs for the model without assimilation (black), with 90-day-long assimilation cycles (red) for the period 2000-2002, and with 120-day-long assimilation cycles (blue) for the year 2000. The beginning of each cycle of assimilation is indicated by a circle. A cost of 1 denotes that the model free surface agrees with the MADTs within the error bar.

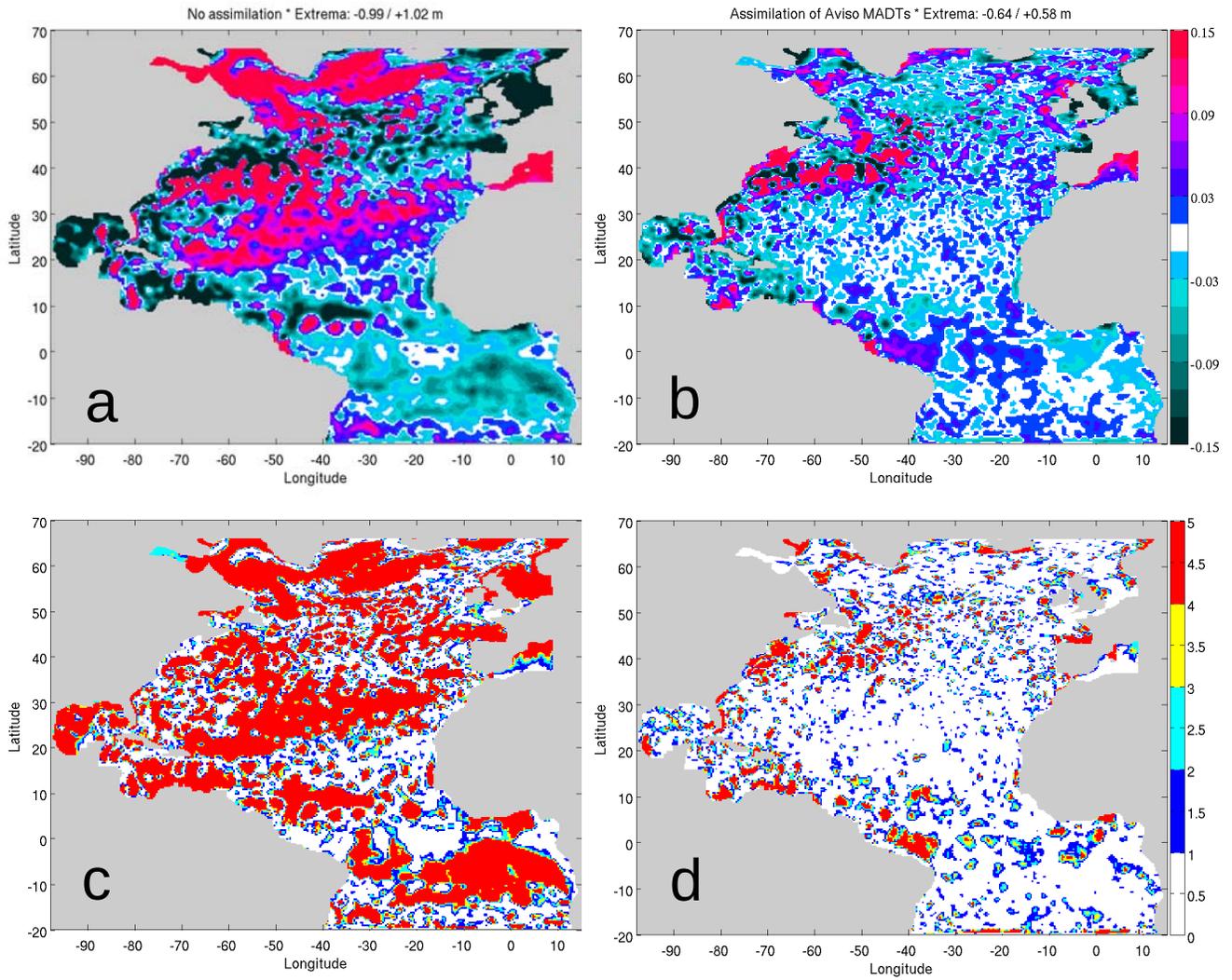


Figure 11: Model-data misfit for the last MADT of the year 2000 (i.e. after 1-year assimilation): a) model without assimilation, b) with assimilation. Equivalent normalised cost function c) without assimilation, d) with assimilation (e.g. a factor of 4 means that the difference between the model and MADTs is $\sqrt{4}$ larger than the MADT error provided on Fig. 1).

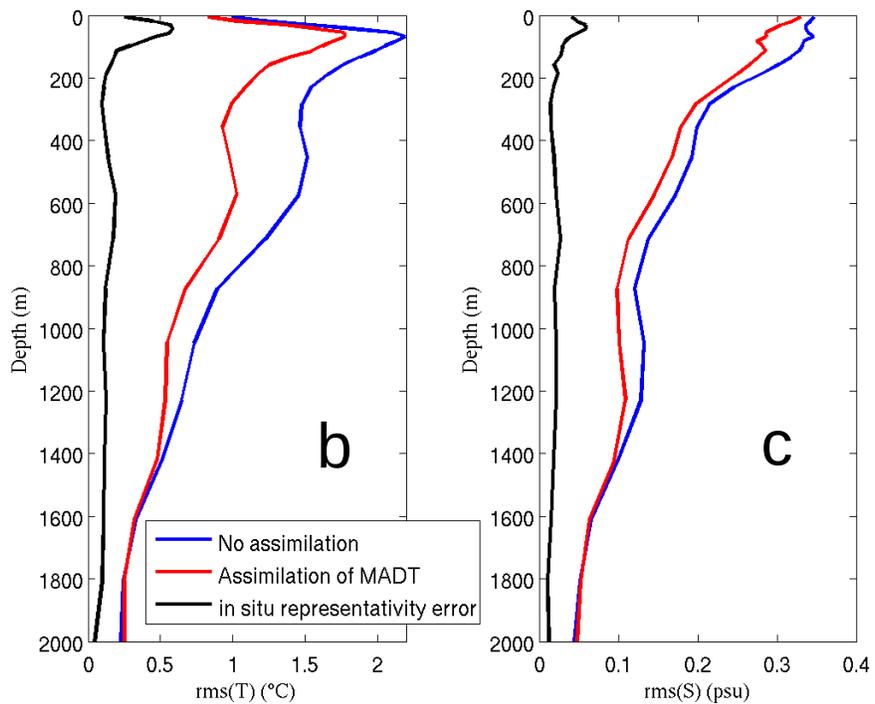
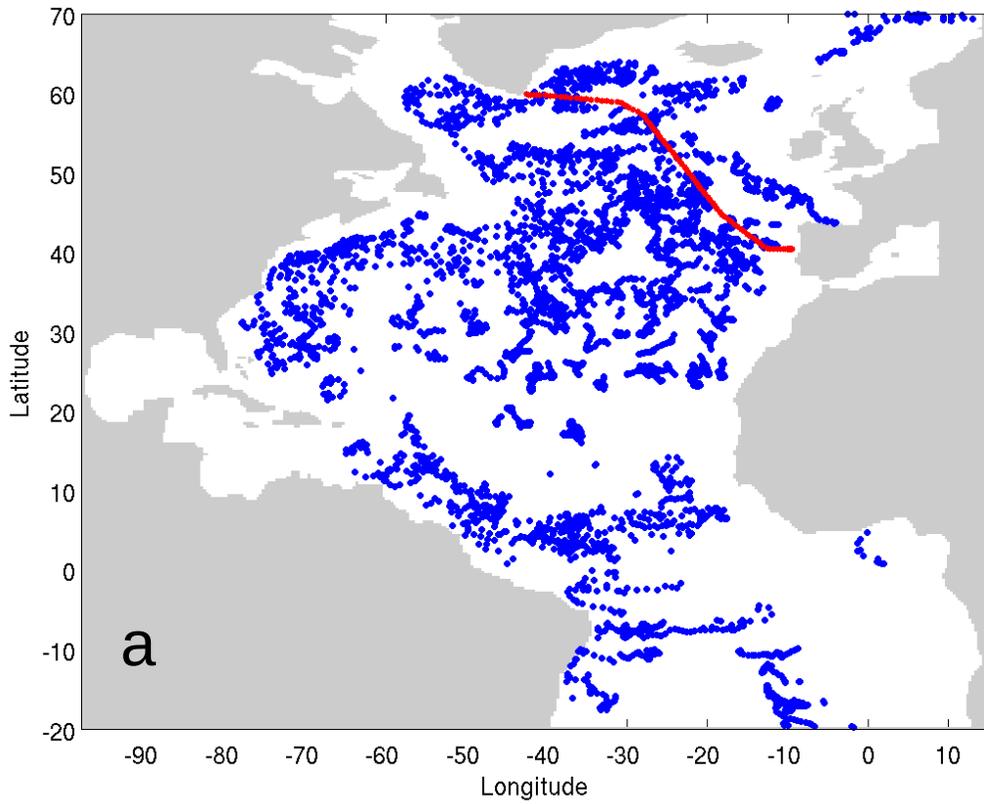


Figure 12: a) Location of 4661 Argo temperature and salinity profiles (blue dots) and Ovide 2002 CTD casts (red dots) during the year 2002; rms of the differences between Argo data and the model without (blue) and with (red) assimilation of MADTs for b) temperature and c) salinity; an estimate of the representativity error on the in situ profile is also provided (black).

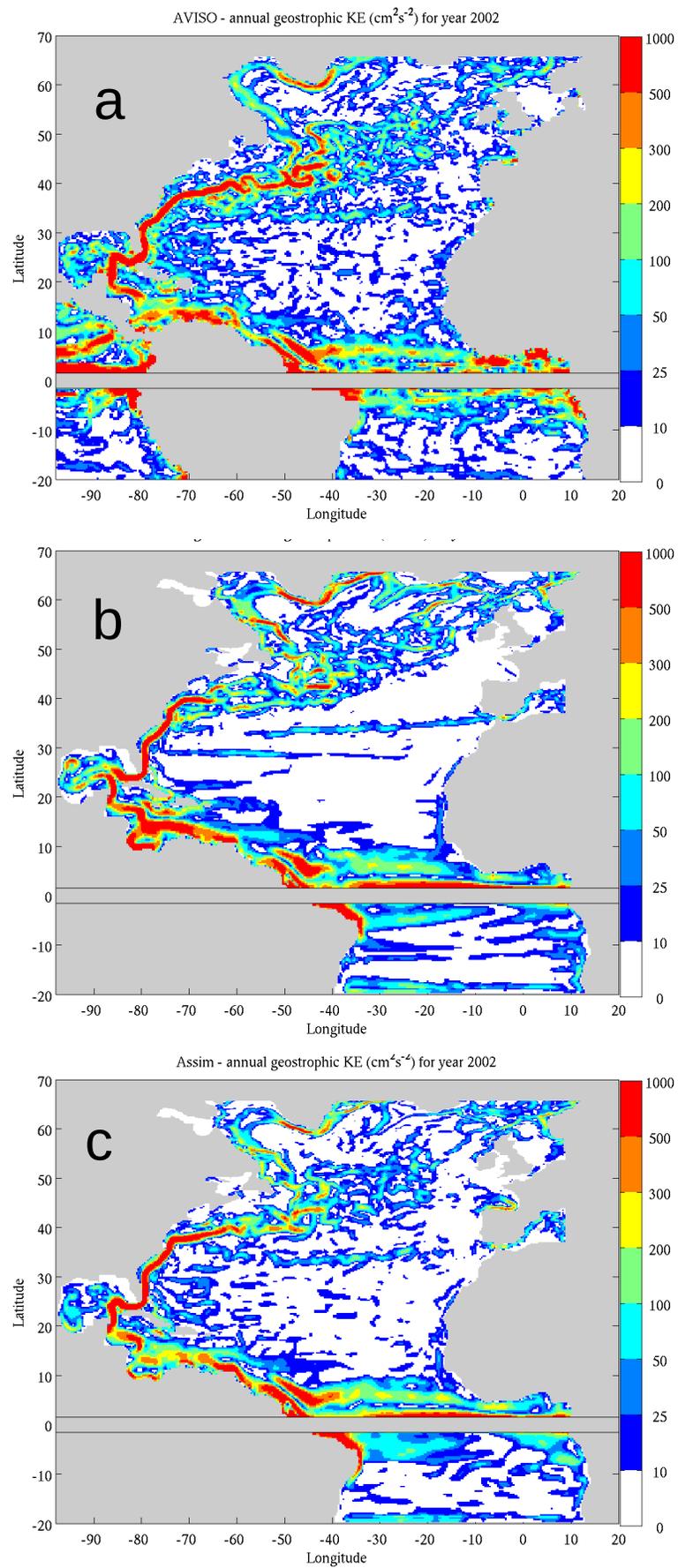


Figure 13: Surface geostrophic kinetic energy averaged over the year 2002 from a) Aviso sea surface height, b) model free surface without assimilation, c) same as b) but with assimilation.

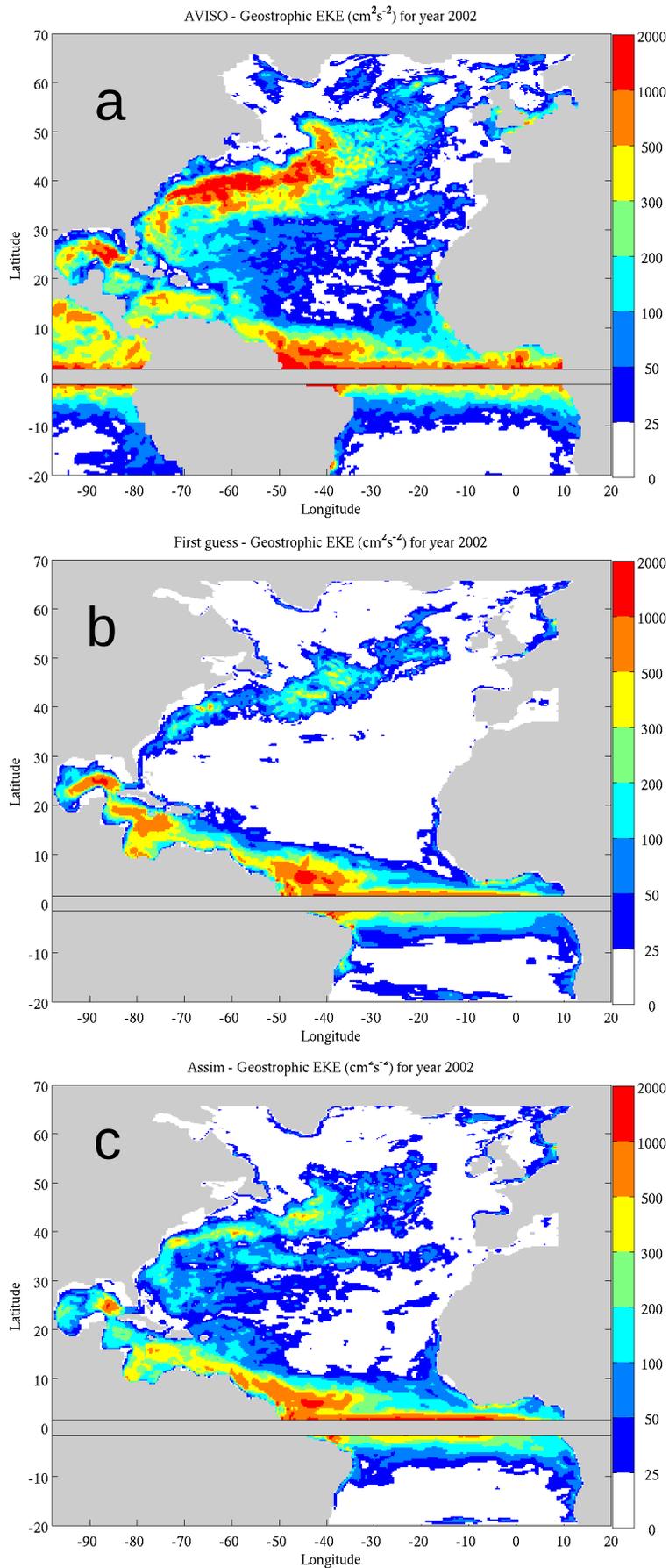


Figure 14: Surface geostrophic “eddy” kinetic energy averaged over the year 2002 from a) Aviso sea surface height, b) model free surface without assimilation, c) same as b) but with assimilation. “Eddy” velocities are defined as departures from the 2002 annual mean velocity field.

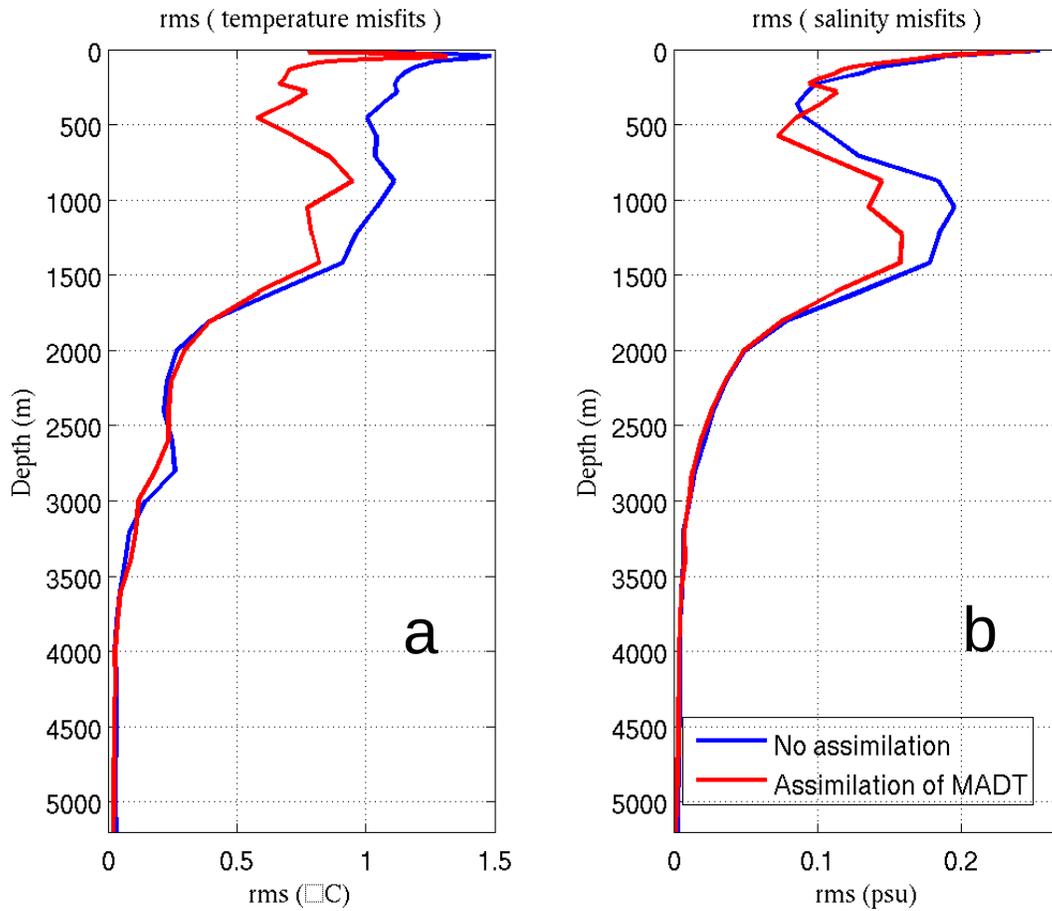


Figure 15: rms of the differences between Ovide2002 CTD and the model without (blue) and with (red) assimilation of MADTs for a) temperature and b) salinity.

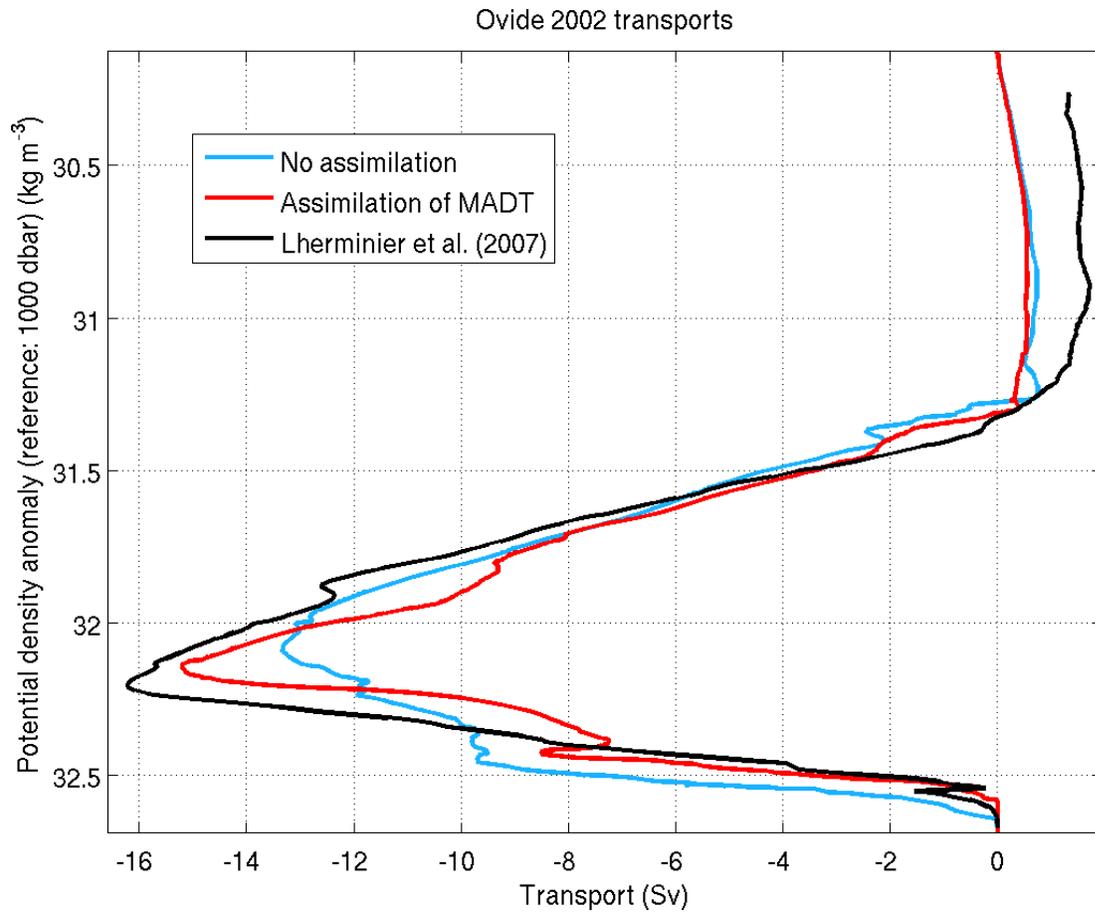


Figure 16: Cumulated transports from the densest to the lightest water masses across the Ovide 2002 hydrographic line as a function of potential density referenced to 1000 dbar. Estimates come from the model without (blue) and with (red) MADT assimilation and from an independent inverse stationary model (Lherminier et al. 2007, black). Note that, following the model velocities, transports are positive when oriented to the north and east.

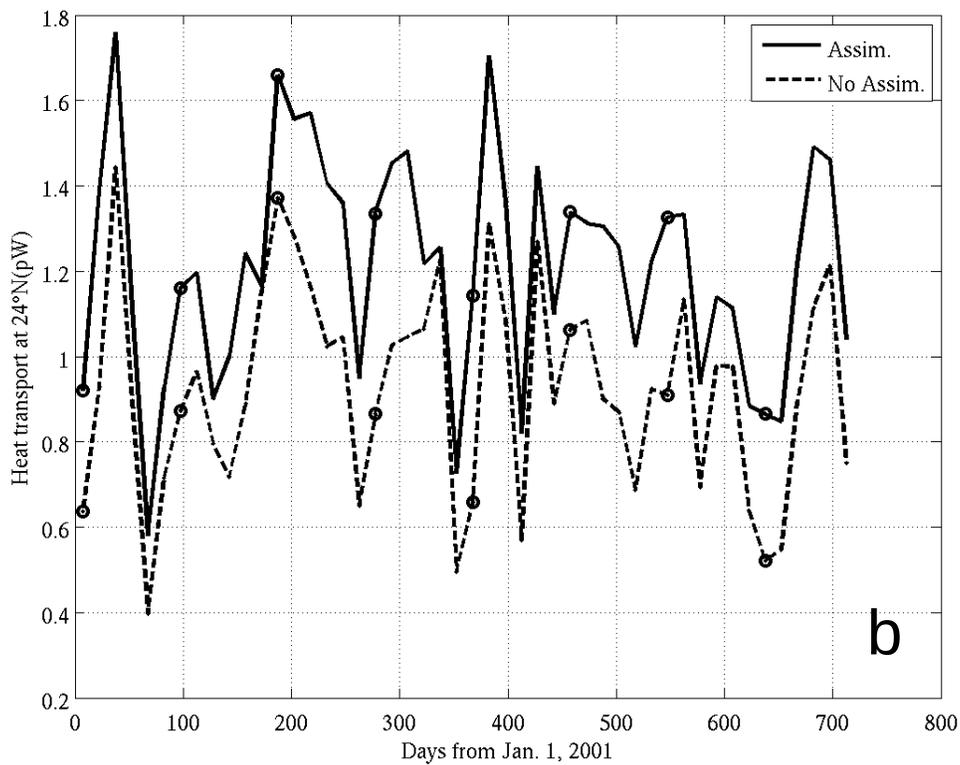
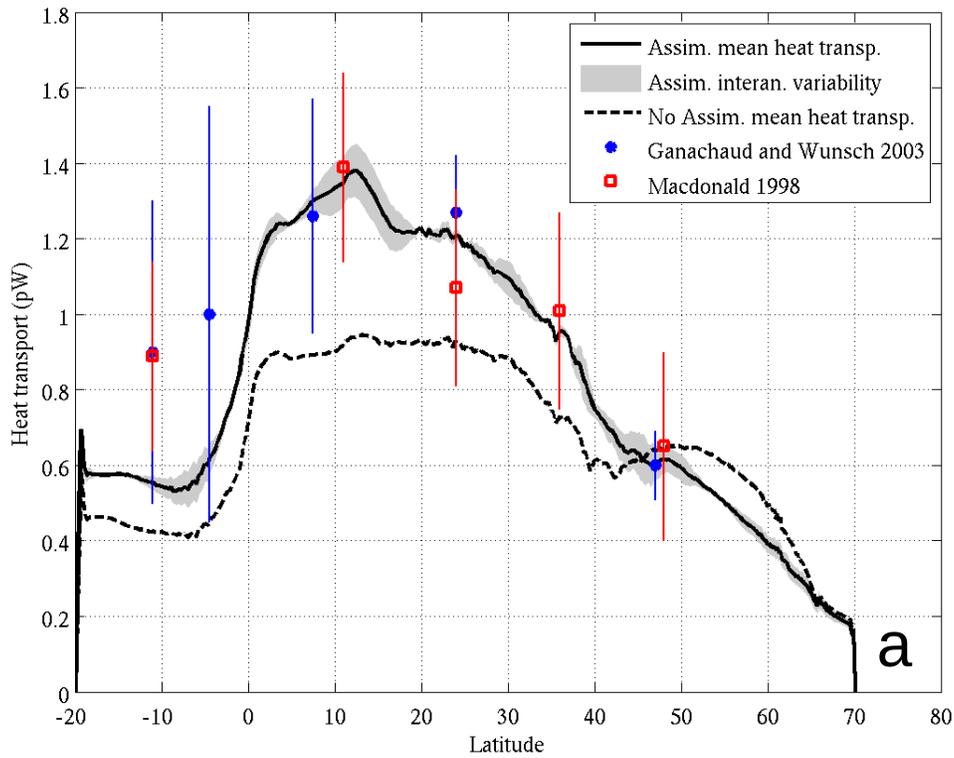


Figure 17: a) Heat transport from independent inversions of hydrographic data (Ganachaud and Wunsch, 2003; Macdonald 1998) and from an average over the period 2001-2002 in the model with assimilation and without assimilation (1 pW = 10^{15} Watts). The shaded region represents the 2001-2002 inter-annual variability of the model with assimilation. Vertical coloured lines are the error bar associated with the independent inversions. b) Time history of the heat transport at 24°N in the model with and without assimilation. Circles represent the start of assimilation cycles; they are reported on the curve without assimilation for comparison.