ELSEVIER

Journal of Marine Systems 35 (2002) 1–9

JOURNAL OF MARINE SYSTEMS

www.elsevier.com/locate/jmarsys

Design of an array of profiling floats in the North Atlantic from model simulations

S. Guinehut*, G. Larnicol, P.Y. Le Traon

CLS Space Oceanography Division, 8-10 rue Hermès, 31526 Ramonville St. Agne, France

Received 1 December 2000; accepted 22 November 2001

Abstract

This study aims to analyze the contribution of different profiling float arrays to the description of the 3-D large-scale thermohaline fields. It uses outputs and profiling float simulations derived from a primitive equation model of the North Atlantic Ocean. An optimal interpolation method is used to reconstruct the large-scale and low-frequency variability of the temperature and salinity fields at different depths from various sets of simulated temperature and salinity profiles. Both the covariance function and noise-to-signal ratio are derived from the analysis of the model fields. The differences between the reconstructed fields and the reference (model) fields are then analyzed. As expected, the results are sensitive to the a priori definition of the large-scale and low-frequency reference field. They show, however, that a 3° array with a 10-day cycle (*Argo* "nominal" resolution) can retrieve most of the variance of the large-scale and low-frequency temperature and salinity signals as observed by a 1/3° primitive equation model. The comparison between Eulerian and Lagrangian arrays shows only a slight deterioration in the results due to the spatial dispersion of the floats and due to the loss of 20 of them which left the model domain during the experiment.

Keywords: Ocean; Array design; Profiling float; Model; North Atlantic Ocean

1. Introduction

Argo is a global array of temperature/salinity profiling floats which is planned as a major component of the ocean observing system (Roemmich et al., 1999). It will complement the existing remote sensing and in situ observing systems, and will serve both climate and operational oceanography applications through the GODAE and CLIVAR programs (Smith and Lefebvre, 1997; Le Traon et al., 1999; CLIVAR, 1997). Based on the experience from the present observing system and a priori knowledge of the ocean variability, the *Argo* Science Team (1998) concluded that a global array of approximately 3000 floats, with a uniform 3° grid in latitude and longitude profiling every 10 days, is practical and should be the initial target for *Argo*. This initial design is expected to evolve as the scientific requirements are refined and more quantitative studies on the array's contribution are performed. A more precise design of the array is, however, difficult to achieve. First, there is no unique answer and the design will generally depend on the signals to be observed (e.g. meridional heat flux, heat content, large-scale temperature and salinity) and on their required space/time resolution and accuracy. Second, the optimization problem itself is a complex

* Corresponding author. Fax: +33-561-393-782.
E-mail address: stephanie.guinehut@cls.fr (S. Guinehut).

0924-7963/02/$ - see front matter
PII: S0924-7963(02)00042-8

one which ultimately should require the use of sophisticated data assimilation techniques (e.g. Hackert et al., 1998; Marotzke et al., 1999) and take into account the availability of other data sets (e.g. satellite altimetry). These techniques are, however, still in their infancy and are not yet tractable at high resolution. Results also are often sensitive to the model/assimilation system used to evaluate the observing system.

Given the complexity of the array design problem, it is useful and necessary to analyze the contribution of the data themselves using simple techniques such as optimal interpolation (e.g. Festa and Molinari, 1992; Smith and Meyers, 1996). This is the approach adopted in this study. Our objective is mainly to define a general methodology to provide a quantitative estimate of how different profiling float arrays can contribute to the observation of the large-scale and low-frequency temperature and salinity variations. We will use outputs and profiling float simulations from a primitive equation model of the North Atlantic. From various sets of temperature and salinity simulated profiles, the large-scale and low-frequency variability of the temperature and salinity fields will be reconstructed using an optimal interpolation method, and the contribution of the different arrays will be quantified.

The paper is organized as follows. Data and methods are presented in Section 2. Statistical results for different Eulerian arrays are presented in Section 3. Section 4 shows the sensitivity of the results to the a priori definition of the large-scale reference field. Section 5 compares the results obtained with Eulerian and Lagrangian arrays to estimate the deterioration in the results due to the dispersion of floats. Main conclusions and perspectives are given in Section 6.

2. Data and method

2.1. Data

We use model outputs and profiling float simulations from the MERCATOR/CLIPPER projects (Blanchet et al., 1999; Treguier et al., 1999, 2001; Lumpkin et al., submitted for publication). The model is a primitive equation model of the North Atlantic Ocean with a 1/3° horizontal resolution and 43 vertical levels. It uses realistic forcing from ECMWF reanalysis. A 1-year simulation period (year 1989) is used for the study.

The initial float distribution is on a 1° grid and the floats profile every 10 days from top to bottom from their parking depth (500 and 2000 m). The profiling float simulations are performed in line.

2.2. Method

The objective is to reconstruct the large-scale and low-frequency variability of the temperature (T) and salinity (S) fields at different depths from simulated T and S profiles corresponding to different types of Eulerian and Lagrangian arrays (e.g. from a 1° array to a 5° array). The main issue is to analyze how the a priori defined large-scale signal can be mapped from sparse measurements with a low signal-to-noise ratio (mainly due to mesoscale variability).

2.2.1. Reference fields

The first step consists of calculating the model a priori large-scale reference fields at different depths. Anomalies of the T and S fields relative to a 5-year mean (1989–1993) are first calculated at each depth. The anomalies are then separated into a large-scale part and a "mesoscale" part using a 2-D Loess smoother (Greenslade et al., 1997) with a cut-off wavelength of 10° in latitude and longitude. This approximately corresponds to averages over $6 \times 6°$ boxes (Greenslade et al., 1997) and is supposed to be representative of the signals that can be mapped from a global profiling float array. An example is given in Fig. 1b for an instantaneous field of T at 200 m. The filter appears to retain the main large-scale features of the T and S fields while mainly removing the mesoscale signals. The large-scale signal is the reference field that we want to reconstruct.

2.2.2. Objective analysis method

Simulated T and S profiles corresponding to different types of Eulerian and Lagrangian arrays (e.g. from a 1° array to a 5° array) are subsampled from the model fields every 10 days and at different depths (20, 200 and 1000 m). An objective analysis method (Bretherton et al., 1976) is then used to reconstruct the large-scale signal variations from these simulated data.

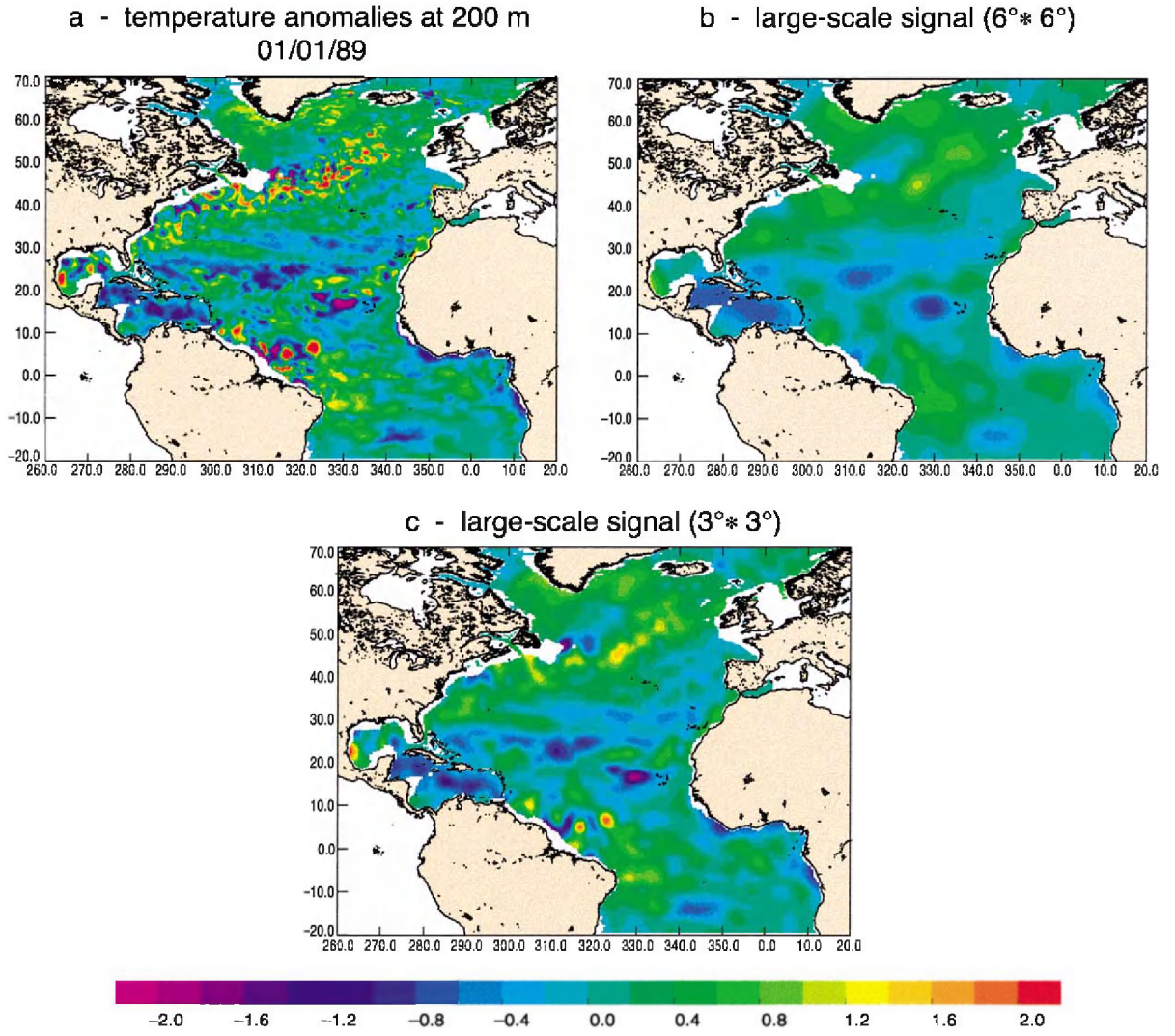

Fig. 1. (a) Instantaneous temperature anomalies at 200 m (1/1/1989), (b) the corresponding large-scale signal defined using a 2-D Loess smoother with a 10° cut-off wavelength and (c) the corresponding large-scale signal defined using a 2-D Loess smoother with a 5° cut-off wavelength.

The covariance functions of the large-scale T and S signals at different depths are derived on a 1° grid from the analysis of the model large-scale fields (Fig. 1b) over the year 1989. The resulting model mesoscale signal is used to determine the noise-to-signal ratio on the same grid. Thus, both covariance and noise-to-signal ratio depend on the geographical position and depth. Values are very similar for the T and S fields. Near the sea surface (20 m), the large-scale signal dominates with low noise-to-signal ratios. Below the mixed layer, the large-scale signal is, however, much smaller and the noise-to-signal ratio increases considerably. The noise-to-signal ratio varies from 0.1 at 20 m to 10 at 1000 m.

This large noise-to-signal ratio means that aliasing of the unresolved mesoscale signal is potentially a serious problem. The very subject of our paper is to analyze the impact of this mesoscale noise on the large-scale thermohaline retrieval.

Analyses are performed every 10 days over a 1-year period (year 1989), but since the objective is to retrieve the low-frequency variability of the large-scale signal, monthly means are calculated. They are then compared to the monthly means derived from the model large-scale reference fields.

3. Some statistical results for different Eulerian arrays

Three regular Eulerian arrays are first tested: $1 \times 1°$, $3 \times 3°$ (*Argo* "nominal" resolution) and $5 \times 5°$ for the T and S fields at 20, 200 and 1000 m.

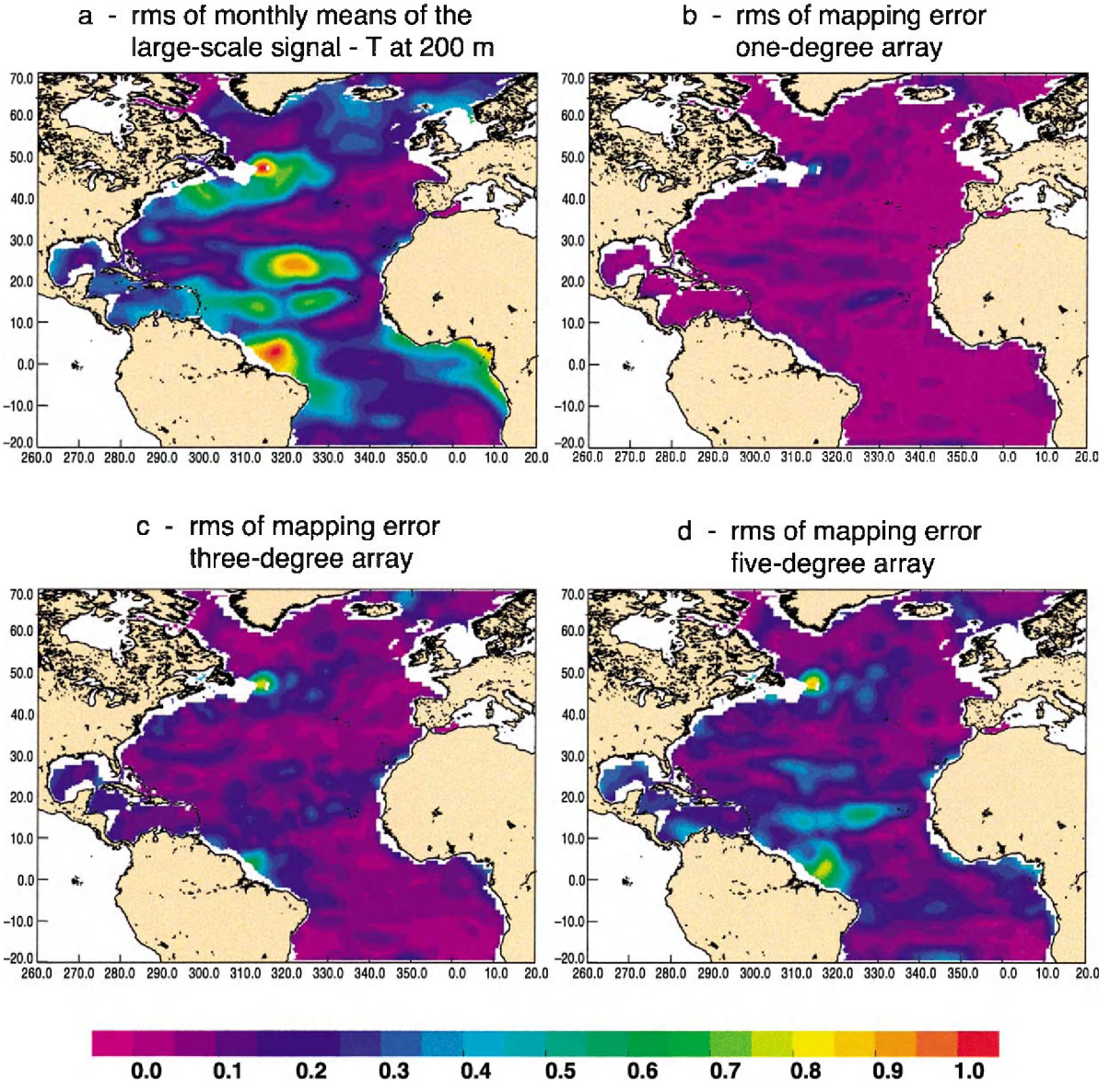

Fig. 2. (a) The rms of the monthly means of the large-scale signal of temperature anomalies at 200 m and the rms of the mapping error for the three arrays (b–d) (in °C). The large-scale signal is defined here using a 2-D Loess smoother with a 10° cut-off wavelength. Statistics are calculated over a 1-year period (year 1989).

Fig. 2 shows the rms of the monthly means of the large-scale signal of temperature anomalies at 200 m and the rms of the mapping error for the three arrays. The large-scale signal is almost perfectly reconstructed with the 1° array (Fig. 2b), whereas the 5° array misses several important structures (e.g. the large signals generated by the model near the mouth of the Amazon; Fig. 2d). The 3° array seems to be a good compromise and captures the main large-scale signals. The rms of the mapping error with the 3° array is almost everywhere much lower than the rms of the signal.

Table 1 compares the mean mapping error (in rms and as percentage of signal variance) of the monthly means of the large-scale T field at 200 m for the three Eulerian arrays and over a 1-year period. As expected, the 1° array gives very good results: 96% of the signal variance is retrieved but it requires 4880 profiling floats over the North Atlantic Ocean. A 5° array is far less costly with 190 floats but is not really satisfactory because it only captures 63% of the signal variance. On the other hand, a 3° array, as proposed by the Argo project, requires only 530 profiling floats and allows a very good retrieval of the signal (82% of the variance).

Near the sea surface (20 m), the T field is a very large-scale signal with a large variance (rms = 2.08 °C, see Table 1). It corresponds to the SST response to the (large-scale) atmospheric heat fluxes. This signal is almost perfectly retrieved with a 5° array (98% of the variance). The errors on the T signals decrease with depth for the different arrays (0.16 °C at 20 m for a 3° array to 0.14 °C at 200 m and 0.05 °C at 1000 m) but the T variance decreases much more rapidly (2.08, 0.34 and 0.08 °C, respectively, at the same depths). As a result, the error relative to the signal variance increases considerably with depth. A 3° array is then required at 200 m (82% of the signal retrieved). However, it only captures 66% of the signal variance at 1000 m (Table 1) due to the very low temperature variance at this depth and the very low signal-to-noise ratio.

Results obtained for the S field are very similar to those obtained with the T field (Table 1) at 200 and 1000 m. Near the sea surface, salinity spatial scales are smaller than temperature scales and the signals are less easily reconstructed.

4. Sensitivity of the results to the a priori definition of the large-scale field

The results obtained in Section 3 depend, of course, on the a priori definition of the large-scale

Table 1
Mapping errors of the monthly means of the large-scale temperature and salinity fields at different depths and for the three arrays

Array	rms of mapping error (°C)	Mapping error as percentage of signal variance	rms of mapping error (psu)	Mapping error as percentage of signal variance
	T at 20 m—rms = 2.08 °C		S at 20 m—rms = 0.19 psu	
1 × 1° (4880 floats)	0.09	0.2	0.05	6.3
3 × 3° (530 floats)	0.16	0.6	0.05	7.4
5 × 5° (190 floats)	0.24	1.4	0.08	18.9
	T at 200 m—rms = 0.34 °C		S at 200 m—rms = 0.06 psu	
1 × 1°	0.07	4.1	0.01	5.4
3 × 3°	0.14	17.5	0.02	15.8
5 × 5°	0.21	36.7	0.04	36.7
	T at 1000 m—rms = 0.08 °C		S at 1000 m—rms = 0.01 psu	
1 × 1°	0.02	4.0	0.003	4.3
3 × 3°	0.05	34.4	0.008	28.3
5 × 5°	0.07	61.3	0.01	59.6

The large-scale signal is defined using a 2-D Loess smoother with a 10° cut-off wavelength. The rms of the T and S signals are also indicated. Statistics are calculated over a 1-year period (year 1989).

and low-frequency signal (space/time means over approximately $6 \times 6°$ boxes and 1 month—hereafter referred to as the reference experiment). In this section, we study the sensitivity of the results to this a priori choice to determine the mapping accuracy according to the spatial scale of the signals to be mapped.

The reference fields are now constructed using a 2-D Loess smoother with a cut-off wavelength of 5° in latitude and longitude. This approximately corresponds to the averages over $3 \times 3°$ boxes, i.e. half the size as in Section 3. An example is given in Fig. 1c for an instantaneous field of T at 200 m. It can be compared to the example in Fig. 1b. The filter preserves smaller scale features like the signals generated by the model near the mouth of the Amazon and in the North Atlantic Current system. The corresponding mesoscale signal captures very small-scale features and is less energetic than the one presented in the reference experiment.

In a similar manner, both covariance functions and noise-to-signal ratio are derived from the analysis of the model large-scale and mesoscale fields for the year 1989. As expected, the spatial correlation scales for the T and S fields are everywhere smaller than those in Section 3. The noise-to-signal ratio is also everywhere smaller due to a lower "mesoscale" field variance and a larger variance of the large-scale field. The noise-to-signal ratio varies from 0.05 at 20 m to 3 at 1000 m.

The same analyses as in Section 3 are then performed for a 3° regular Eulerian array for the T and S fields at 20, 200 and 1000 m. Table 2 compares the mean mapping error (in rms and as percentage of signal variance) of the monthly means of the large-scale T and S fields at 20, 200 and 1000 m. These results can be compared to those in Table 1.

As expected, the results are not as good as those obtained by the reference experiment. For the T field at 200 m, the mapping error was 0.14 °C. It is now 0.23 °C. Even if the rms of the large-scale T signal increases from 0.34 to 0.42 °C, only 70% of the signal is retrieved compared with 80% in the reference experiment. At the surface (20 m), more than 90% of the variance of the T and S signals is still retrieved but only between 55% and 60% at 1000 m (it was between 65% and 70% in the reference experiment). Compared to the reference experiment (Fig. 2a and c), the rms of the monthly means of the large-scale signal of the T anomalies at 200 m (Fig. 3a) and the rms of the mapping error (Fig. 3b) are larger everywhere and show smaller scale features.

The larger errors can be explained by the difficulty of reconstructing smaller scale signals (here, scales below 6°) with the same number of observations (i.e. 530 floats for the 3° array). However, even with this rather conservative definition of the large-scale signal, a 3° array is able to retrieve a large portion of the large-scale T and S signals. The rms of the mapping error of T at 200 m (Fig. 3b) is thus, for example,

Table 2
Mapping errors of the monthly means of the large-scale temperature and salinity fields at different depths and for a 3° array

Array	rms of mapping error (°C)	Mapping error as percentage of signal variance	rms of mapping error (psu)	Mapping error as percentage of signal variance
	T at 20 m—rms = 2.10 °C		S at 20 m—rms = 0.23 psu	
$3 \times 3°$	0.25	1.4	0.08	13.0
	T at 200 m—rms = 0.42 °C		S at 200 m—rms = 0.08 psu	
$3 \times 3°$	0.23	29.8	0.04	27.0
	T at 1000 m—rms = 0.12 °C		S at 1000 m—rms = 0.02 psu	
$3 \times 3°$	0.08	44.7	0.01	36.2

The large-scale signal is defined using a 2-D Loess smoother with a 5° cut-off wavelength. The rms of the T and S signals are also indicated. Statistics are calculated over a 1-year period (year 1989).

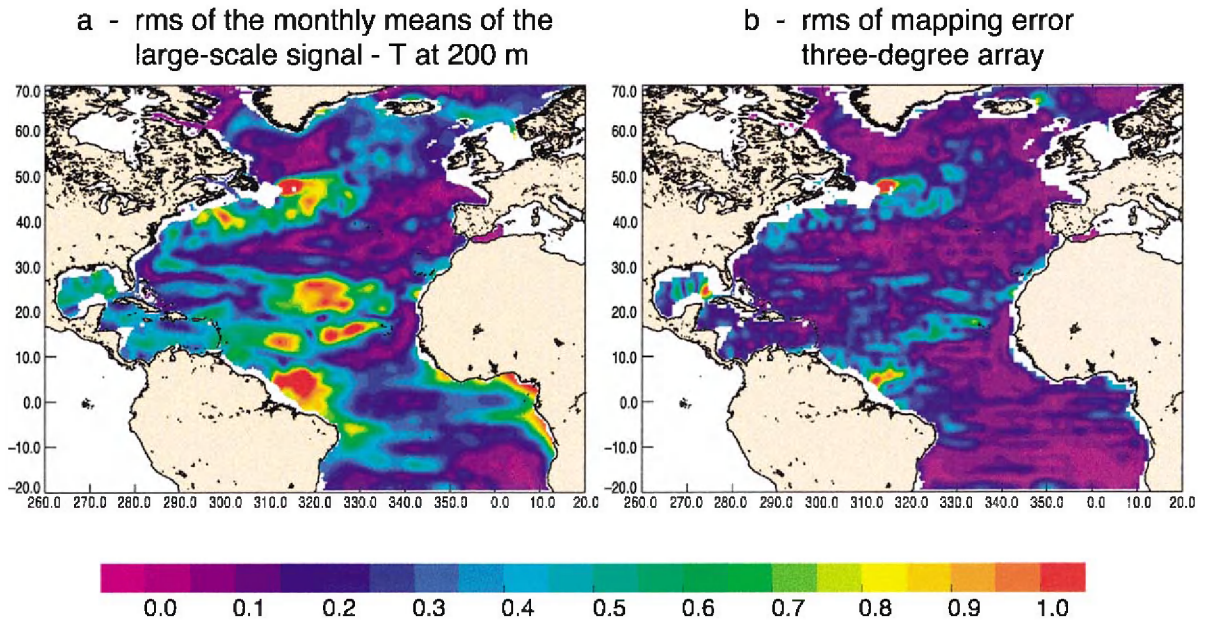

Fig. 3. (a) The rms of the monthly means of the large-scale signal of temperature anomalies at 200 m and the rms of the mapping error for a 3° array (b) (in °C). The large-scale signal is defined here using a 2-D Loess filter with a 5° cut-off wavelength. Statistics are calculated over a 1-year period (year 1989).

significantly smaller than the large-scale signal (Fig. 3a).

5. Comparison of Eulerian and Lagrangian arrays

Differences between Eulerian and Lagrangian 3° arrays are now analyzed. Our aim is to estimate the deterioration in results due to the loss or the dispersion of floats. Our analysis uses the Lagrangian float simulations at a depth of 500 m. We chose to analyze the 500 m simulations [(rather than 2000 m, which is the recommended parking depth for *Argo* floats (Roemmich et al., 1999)] to have a sufficient (and probably more realistic) estimation of the actual dispersion of the floats (Fig. 4) during the 1-year period.

The time evolution of the T mapping error at 200 m (Fig. 5) shows only a slight deterioration in the results for the Lagrangian array compared to the Eulerian array. Differences are of the order of 0.01 °C, which is small compared to the signal variance (about 2.5% of the signal variance; the reference field is constructed in the same manner as in the reference experiment). In only 1 year, the profiling floats did not disperse/converge too much and all the floats continue to provide useful information. Part of the differences can actually be explained by the loss of 20 floats which left the model domain during the experiment.

The rms of the mapping error (not shown) is very similar to that in Fig. 2c, with smaller errors where Lagrangian floats converged and larger errors where floats dispersed.

6. Conclusions and perspectives

Our study suggests that a 3° array of profiling floats cycling every 10 days (*Argo* "nominal" resolution) can retrieve most of the variance of the large-scale and low-frequency T and S signals as observed by a 1/3° primitive equation model. For space/time means over approximately $6 \times 6°$ and 1 month, more than 90% of the variance of the T and S signals is retrieved at 20 m, 80% at 200 m and between 65% and 70% at 1000 m. Results deteriorate with a more conservative definition of the large-scale signal (means over $3 \times 3°$ boxes) but they still show that a large portion of the large-scale signal can be retrieved. The comparison

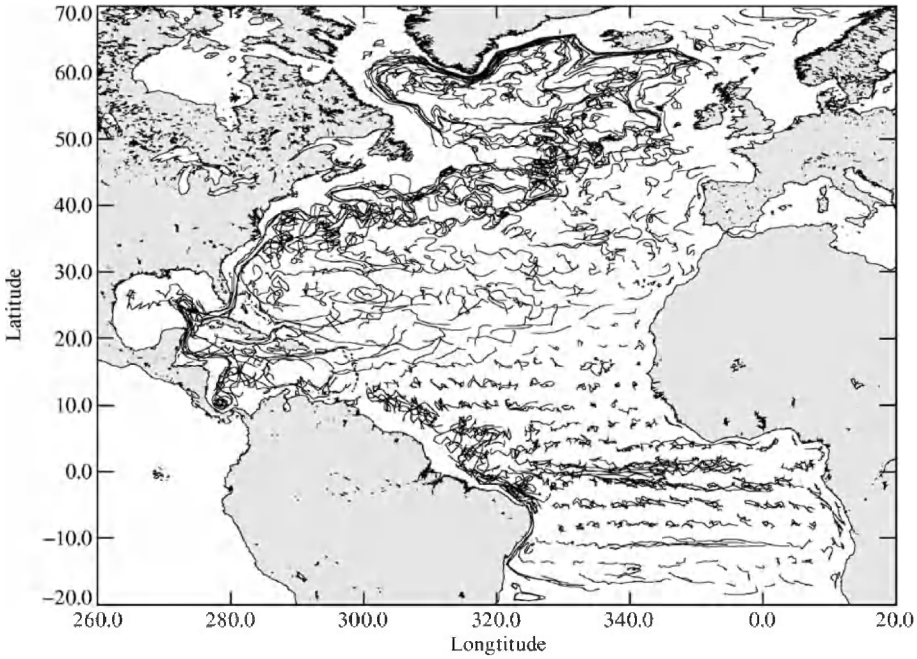

Fig. 4. Floats trajectories at 500 m over 1 year.

between Eulerian and Lagrangian arrays shows only a slight deterioration in the results due to the spatial dispersion of floats and due to the loss of 20 of them during the experiment.

These results are very different from those derived from the simple Nyquist theory. The latter only holds for regular grids and perfect measurements while the noise-to-signal ratios (mainly due to mesoscale variability) derived from the model vary from 0.1 to values as high as 10. Should we have applied the simple Nyquist theory, we would have found that a 3° array resolves wavelengths larger than 6° and periods longer than 20 days. We found, however, that we only resolve (i.e. here, map with a "good" accuracy) wavelengths larger than 12° and periods longer than 2 months.

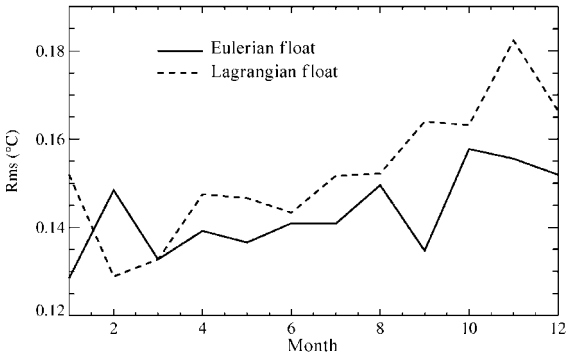

Fig. 5. The rms of the monthly mean T mapping error at 200 m as a function of month (in °C) for the 3° Eulerian and Lagrangian arrays. The large-scale signal is defined using a 2-D Loess smoother with a 10° cut-off wavelength.

The representation of the mesoscale variability was limited to that of a 1/3° primitive equation model. This is a limitation of the study, and thus, our conclusions may not hold in the "real" ocean. In the near future, we plan to extend this study using a higher resolution model over a longer period (4 years). This will allow us to better estimate the impact of (1) the mesoscale field on the large-scale signal retrieval (we believe, however, that our results would not change much in areas of low variability but change only in areas of high variability (for example, the Gulf Stream region)) and (2) the Lagrangian dispersion of the floats over their expected lifetime (loss of instruments, non-uniform temporal and spatial coverage). A more systematic analysis of sensitivity to

the a priori definition of the large-scale and low-frequency signal should also be done to determine the mapping accuracy according to the space (e.g. from 3° to 10°) and time (e.g. from 10 days to 1 year) scales of the signals to be mapped.

Note, finally, that only profiling float data were used here to estimate the large-scale T and S fields. It is clear, however, that in the future, the best use of profiling float data will be when they are combined with other data sets and models through effective data assimilation techniques. The development and application of such techniques is the main objective of GODAE (Smith and Lefebvre, 1997). The combination of profiling float data with satellite altimetry should, in particular, be instrumental in reducing aliasing due to the mesoscale variability.

Acknowledgements

We would like to thank the MERCATOR PAM team, particularly Laurence Fleury and Yann Drillet, for providing us with the model and float simulations. The study was partly funded by CNES and the EC as part of the Gyroscope project (EVK2_CT_2000_00087).

References

Argo Science Team, 1998. On the design and implementation of Argo: an initial plan for a global array of profiling floats. International CLIVAR Project Office Report 21, GODAE Report 5. GODAE International Project Office, Melbourne Australia, 32 pp.

Blanchet, I., De Prada, T., Drillet, Y., Fleury, L., Perez, H., Siefridt, L., Tranchant, B., 1999. The North Atlantic/Mediterranean Mercator Prototype. The Ocean Observing System for Climate, International Symposium, Saint Raphaël, France.

Bretherton, F., Davis, R.E., Fandry, C.B., 1976. A technique for objective analysis and design of oceanographic experiments applied to MODE-73. Deep-Sea Res. 23, 559–582.

CLIVAR, 1997. A Research Programme on Climate Variability and Predictability for the 21st Century, World Climate Research Programme. WCPR No. 101, WMO/TD No. 853, ICPO No. 10.

Festa, J.F., Molinari, R.L., 1992. An evaluation of the WOCE volunteer observing ship—XBT network in the Atlantic. J. Atmos. Oceanic Technol. 9, 305–317.

Greenslade, D.J.M., Chelton, D.B., Schlax, M.G., 1997. The midlatitude resolution capability of sea level fields constructed from single and multiple satellite altimeter data sets. J. Atmos. Oceanic Technol. 14, 849–870.

Hackert, E.C., Miller, R.N., Busalacchi, A.J., 1998. An optimized design for a moored instrument array in the tropical Atlantic ocean. J. Geophys. Res. 103, 7491–7509.

Le Traon, P.-Y., Rienecker, M., Smith, N., Bahurel, P., Bell, M., Hurlburt, H., Dandin, P., 1999. Operational oceanography and prediction—a GODAE perspective. The Ocean Observing System for Climate, International Symposium, Saint Raphaël, France.

Lumpkin, R., Treguier, A.-M., Speer, K., April 2000. Lagrangian eddy scales in the northern Atlantic Ocean. J. Phys. Oceanol., submitted for publication.

Marotzke, J., Giering, R., Zhang, Q.K., Stammer, D., Hill, C.N., Lee, T., 1999. Construction of the adjoint MIT ocean general circulation model and application to Atlantic heat transport sensitivity. J. Geophys. Res. 104, 29529–29547.

Roemmich, D., Boebel, O., Desaubies, Y., Freeland, H., King, B., LeTraon, P.-Y., Molinari, B., Owens, B., Riser, S., Send, U., Takeuchi, K., Wijffels, S., 1999. *Argo*: the global array of profiling floats. The Ocean Observing System for Climate, International Symposium, Saint Raphaël, France.

Smith, N.R., Lefebvre, M., 1997. The "Global Ocean Data Assimilation Experiment" (GODAE). Monitoring the Oceans in the 2000s: An Integrated Approach, International Symposium, Biarritz, October 15–17.

Smith, N.R., Meyers, G., 1996. An evaluation of expendable bathythermograph and tropical atmosphere–ocean array data for monitoring tropical ocean variability. J. Geophys. Res. 101, 28489–28501.

Treguier, A.-M., Reynaud, T., Pichevin, T., Barnier, B., Molines, J.-M., de Miranda, A.P., Messager, C., Beismann, J.O., Madec, G., Grima, N., Imbard, M., Le Provost, C., September 1999. The CLIPPER project: high resolution modeling of the Atlantic. Int. Woce Newsl. 36, 3–5.

Treguier, A.-M., Barnier, B., de Miranda, A.P., Molines, J.-M., Grima, N., Imbard, M., Madec, G., Messager, C., Reynaud, T., Michel, S., 2001. An eddy permitting model of the Atlantic circulation: evaluating open boundary conditions. J. Geophys. Res. 106, 22115–22130.