On the accuracy of North Atlantic temperature and heat storage fields from Argo

R. E. Hadfield,¹ N. C. Wells,¹ S. A. Josey,¹ and J. J-M. Hirschi¹

Received 20 July 2006; revised 15 August 2006; accepted 11 September 2006; published 25 January 2007.

[1] The accuracy with which the Argo profiling float dataset can estimate the upper ocean temperature and heat storage in the North Atlantic is investigated. A hydrographic section across 36°N is used to assess uncertainty in Argo-based estimates of the temperature field. The root-mean-square (RMS) difference in the Argo-based temperature field relative to the section measurements is about 0.6°C. The RMS difference is smaller, less than 0.4°C, in the eastern basin and larger, up to 2.0°C, toward the western boundary. In comparison, the difference of the section with respect to the World Ocean Atlas (WOA) is 0.8°C. For the upper 100 m, the improvement with Argo is more dramatic, the RMS difference being 0.56°C, compared to 1.13°C with WOA. The Ocean Circulation and Climate Advanced Model (OCCAM) is used to determine the Argo sampling error in mixed layer heat storage estimates. Using OCCAM subsampled to typical Argo sampling density, it is found that outside of the western boundary, the mixed layer monthly heat storage in the subtropical North Atlantic has a sampling error of 10–20 Wm⁻² when averaged over a 10° x 10° area. This error reduces to less than 10 Wm⁻² when seasonal heat storage is considered. Errors of this magnitude suggest that the Argo dataset is of use for investigating variability in mixed layer heat storage on interannual timescales. However, the expected sampling error increases to more than 50 Wm⁻² in the Gulf Stream region and north of 40°N, limiting the use of Argo in these areas.


1. Introduction

[2] The relative importance of the ocean and atmosphere for redistributing heat around the globe has been the subject of much research [Vonder Haar and Oort, 1973; Oort and Vonder Haar, 1976; Carrisimo et al., 1985; Bryden, 1993; Trenberth and Solomon, 1994; Trenberth and Caron, 2001]. Estimates of the ocean and atmosphere heat transports made by Trenberth and Caron [2001] suggest that poleward Ocean Heat Transports (OHT) are dominant only between 0 and 17°N. However, a decomposition of the total poleward heat transport into OHT, dry static atmospheric energy transport, and latent heat transport associated with the meridional fresh water transport indicates that the oceans and atmosphere may contribute about equally in maintaining the global heat balance [Bryden and Imawaki, 2001]. In order to gain a full understanding of the role the ocean plays within the global climate system, it is necessary to consider variations in the Ocean Heat Content (OHC) and the exchange of heat across the air-sea interface in addition to the OHT.

[3] In our study, we aim to make some progress towards determining the capabilities of the recently deployed Argo array of profiling floats for investigating the subsurface temperature field and rates of change in mixed layer heat content (henceforth the mixed layer heat storage). We focus on the North Atlantic Ocean, which deserves particular attention for its role in maintaining the present day global climate and because of its expected sensitivity to changes in global climate [Banks and Wood, 2001]. Although the Argo dataset provides subsurface temperature observations of the upper 2000 m, the results in the upper ocean are of particular interest due to its interactions with the atmosphere.

[4] Many previous studies have attempted to quantify the upper ocean heat storage in the North Atlantic. Gabites [1950] relied on a sea surface temperature dataset combined with a simplified model of the seasonal thermocline (due to the lack of subsurface temperature observations). Bryan and Schroeder [1960], Oort and Vonder Haar [1976], Lamb and Bunker [1982], Hsiung et al. [1989], Levitus [1982], and Levitus and Boyer [1994] used large numbers of expendable bathythermograph (XBT) profiles and ship data. These earlier studies are affected by a spatial sampling bias due to concentration of observations along shipping routes. Bryan and Schroeder [1960] observed that some locations only had data collected during 1 or 2 years and that this could lead to a temporal bias due to decadal changes in heat...
storage. In addition, the accuracy of the XBT dataset is dependent on application of corrections to fall rates [Willis et al., 2003].

Satellite data have been used alongside subsurface observations to try and overcome the spatial and temporal bias inherent in subsurface observational data alone [White and Tai, 1995; Chambers et al., 1999]. Such studies use the correlation between steric height, which can be extracted from altimetric height observations, and subsurface temperature observations to obtain estimates of the subsurface temperature with better spatial coverage than the directly measured values. However, this regression method is not very accurate in regions where steric height signals are weak. Guinehut et al. [2004] found that using the regression method on sea surface temperature and altimeter data could explain less than 40% of the signal variance at a depth of 200 m in the North Atlantic subtropical and subpolar gyres.

Models (OGCMs).

In the following, section 2 provides information on the Argo dataset. The methods of interpolation and heat storage calculation are outlined in sections 3. Our results regarding the accuracy of Argo-based temperature and heat storage estimates are presented in section 4. Finally, in section 5 we summarize our main findings and discuss their implications.

2. Data

This study utilizes temperature profiles obtained from free-drifting Argo floats. The Argo project aims to have one float in every 3° square of ocean, equivalent to roughly 3000 floats globally. Although this target has not yet been met for the World Ocean, the coverage provided by the current (2500) active floats is close to the target density for some regions. The target number of floats for the North Atlantic is about 350; with each float sampling one profile every 10 days these floats would provide 1050 profiles per month. The coverage of the Argo array at the end of 2005 sampled around 760 profiles per month in the North Atlantic (Figure 1).

It should be noted that Figure 1 shows values after both the Argo real time quality control and additional, more stringent quality control procedures that we have developed have been applied. In addition, about 24% of profiles have undergone Argo delayed mode quality control. The quality control checks developed in this study flagged (1) profiles that have no readings in the upper 10 m and (2) profiles that contain temperature readings that are outside a predetermined range. This range varies with location and is calculated using the World Ocean Atlas (WOA) [Stephens et al., 2001]. It is defined as five times the standard deviation of all temperature profiles in a 10-degree box centered on the Argo profile location, or 10°C, whichever was greater. Profiles failing these checks were subjected to visual inspection. Where possible, profiles were corrected. Correctable profiles typically show temperature “jumps” where the temperature reading appears bad at only one or two depths throughout the water column. It is believed that these jumps are caused by transmission errors and not sensor problems. Erroneous profiles that cannot be adjusted are removed from the dataset. During quality control all profiles were interpolated onto regular 10 dbar surfaces. Profiles with missing readings in the upper 100 m were removed from the dataset if extrapolation was not possible. Extrapolation was considered possible only if the gradient in the temperature field is less than a critical value. This value varies with the depth of the shallowest profile observation and is calculated as 0.2(B-1)-1°C m-1, where B is the depth bin of the shallowest reading; B = 1 at 10 m, B = 2 at 15 m, and so on. It was possible to extrapolate 69% of profiles with missing readings in the upper 100 m. The quality control procedures applied in this study led to a
rejection of bad profiles that were not quality controlled by the Global Data Centre but were not able to identify, or correct profiles in which temperature readings were drifting by a small amount. However, the expected drift is only 0.005°C year⁻¹ [Bohme, 2003] and so should not greatly affect results. On average 30 (40) profiles were removed each month by our (Argo) quality control procedures.

[11] The Argo sampling rate for 2002 and the 7-year period 1999–2005 is plotted in Figure 2. Note 2002 is of particular interest as we focus on this year later to determine the accuracy of Argo-based estimates of heat storage. The coverage of temperature profiles from Argo is highly variable throughout the North Atlantic. The region off the west coast of Africa is poorly sampled, while further north the data density is high. There is also variability in the spatial sampling densities between years. The sampling density during 2002 and thereafter is higher than the average for the years 1999–2005 due to the low number of floats deployed during the early years of the Argo project. Note also that although the profile distribution during 2002 is similar to the average for 1999–2005, it is a year of relatively low sampling in parts of the western boundary; thus our estimates of sampling related uncertainty in the heat storage estimates for this region may have been lower had a wider range of years been employed.

3. Method
3.1. Optimal Interpolation
[12] In order to undertake analysis of the temperature and heat storage fields, the Argo float data are optimally interpolated to cruise station positions and times and regular monthly 2° grids, respectively. The Optimal Interpolation (OI) method is based on the Gauss-Markov theorem which gives a linear estimate that is unbiased, is optimal in the least squares sense and provides an estimate of the error variance [Brewer et al., 1976; McIntosh, 1990; Wong et al., 2003; Bohme and Send, 2005]. In this study the covariance of the data is assumed to be Gaussian, with the decay scale determined by four correlation parameters: a longitudinal scale (Lx), a latitudinal scale (Ly), a cross-isobath scale (Φ), and a temporal scale (Δt).

[13] The objective estimate of the temperature at each grid point, for each 10 dbar surface is given by

$$T^{obj} = T_{WOA} + w \cdot (T - T_{WOA})$$

(1)

$T$ denotes the “N” profiles closest in space and time to the grid point being interpolated to. $N$ is the number of data points used in the interpolation. We investigated the dependence of the interpolation on $N$ and found that beyond $N = 40$ there was no significant difference to the interpolated value. We therefore set $N$ equal to 40 for this analysis. $T_{WOA}$ is the monthly mean field, taken to be the climatological value in the World Ocean Atlas [Stephens et al., 2001], and $w$ is the weighting matrix. Each element of $T$ is made up of a true signal ($s$) and some random noise ($η$).

[14] The weighting,

$$w = Cdg \cdot Cdd^{-1}$$

(2)

where $Cdg$ is the data-grid covariance,

$$Cdg = \langle s^2 \rangle \cdot \exp \left\{ - \frac{D_{ij}^s}{Lx^2} - \frac{D_{ij}^s}{Ly^2} - \frac{F_{ij}^s}{Φ^2} - \frac{D_{ij}^s}{Δt^2} \right\}$$

(3)

and $Cdd$ is the data-data covariance,

$$Cdd_{ij} = \langle s^2 \rangle \cdot \exp \left\{ - \frac{D_{ij}^s}{Lx^2} - \frac{D_{ij}^s}{Ly^2} - \frac{F_{ij}^s}{Φ^2} - \frac{D_{ij}^s}{Δt^2} \right\}$$

(4)

$$s^2 = \frac{1}{N} \sum_i (T_i - T_{WOA})^2,$$

(5)

$$η^2 = \frac{1}{2N} \sum_i (T_i - T_0)^2.$$
spatial distances between the floats and the grid point being mapped to in the zonal and meridional directions, respectively; $F$ is the fractional distance in potential vorticity representing the cross-isobath separation [Böhme and Send, 2005]; and $Dt$ is the temporal separation. The subscripts $i$ and $j$ refer to Argo profiles, while the subscript $g$ refers to the grid point. $F$, the cross-isobath separation, is calculated using the following formula,

$$
F = \frac{|PV(a) - PV(b)|}{\sqrt{PV^2(a) + PV^2(b)}}
$$

where $PV$ is the barotropic potential vorticity, $fH$, $f$ is the Coriolis parameter and $H$ is the full ocean depth. The inclusion of the cross-isobath separation considers the tendency of ocean currents to follow the bathymetry. In the calculation of $F$ for the data-grid covariance, $C_{dg}$, $a$ and $b$ represent the float profile positions, $i$, and the grid position, $g$, respectively, while for the calculation of $F$ for the data-data covariance, $C_{dd}$, $a$ and $b$ represent pairs of floats with the shortest spatial-temporal separations, $i, n$.

[15] Length scales ($L_x, L_y$) applied during interpolation in the literature vary from 150 km [Lavender et al., 2005], up to 1000 km [Böhme and Send, 2005]. To some extent, the choice of length scales is dependent on the sampling density of the dataset to be interpolated and the desired use of the gridded field. Longer length scales are typically used in applications involving investigation of a large-scale field (with smaller-scale features averaged out).

[16] We have estimated temperature correlation length scales using Argo temperature data and the method of Lavender et al. [2005]. Anomaly profiles are first calculated by removing the seasonal cycle, taken to be the WOA climatological temperature for the month and position closest to each Argo profile position. More than 43,000 profiles obtained between January 1999 and December 2005 are used in this analysis. For each pair of anomaly temperature profiles measured within a set timescale of one another, the spatial separation of the observations is calculated. The timescale needs to be short enough so that data pairs are correlated in time but long enough to ensure that there are enough data pairs to calculate the correlations; a timescale of 10 days is used here. Pairs are placed into bins defined by their spatial separation and the temperature anomaly correlation coefficient squared ($r^2$) of pairs in each bin is then calculated. Spatial separations of 0–500 km are considered with a bin size of 25 km. For example, in the first spatial bin, which considers pairs of floats separated by less than 10 days in time and 0–25 km in space, there are 10,524 data pairs. For plotting, the spatial correlations are normalized by the correlation of the first spatial bin.

[17] The results based on data collected throughout the whole North Atlantic are shown in Figure 3a. Two spatial scales can be identified; one reflecting the mesoscale,
portrayed by the rapid decrease in correlation at short spatial lags, and one reflecting the climatological scale. This climatological scale is depicted by the long “tail” in the correlation plot and represents the length scale of the background field (i.e., ignoring mesoscale features). The mesoscale, identified as the length scale at which the normalized correlation values level out, varies throughout the water column with longest scales of around 200 km in the surface decreasing to 150 km at 100 m. The climatological scale, identified by the maximum significant correlation values, exceeds 500 km at 20 m. However, below 50 m correlations are insignificant at lags of more than 180 km. The percentage of grid points with data within different distances and 1 month (discussed in detail below) is also plotted in Figure 3a. On the basis of this data availability, we have selected values for $L_x$ and $L_y$ of 500 km. This value is consistent with the upper limit of the values given above.

Length scales have been calculated for subregions of the North Atlantic; the “western basin” (0–40°N, west of 40°W), the “eastern basin” (0–40°N, east of 40°W), and the “subpolar” region (40–70°N) (Figures 3b–3d). The length scales vary between the different locations considered. The correlation remains significant at spatial lags of more than 500 km and 450 km in the subpolar and eastern basins respectively but is insignificant at length scales of 250 km in the western basin. As for the whole ocean basin, the correlation on depth surfaces below 20 m are typically insignificant at spatial lags exceeding 200 km. Despite this observed variation in scales, for simplicity, length scales are set to a constant 500 km throughout the North Atlantic.

White [1995] found somewhat larger values for the length scale than our Argo-based estimates. He computed covariance statistics in different ocean basins using a global dataset of 549,149 temperature observations collected over 13 years between 1979 and 1991. Depth surfaces of 0 m, 200 m, and 400 m were considered. Consistent with our study a depth-dependence was found, with values ranging from 800 km at the surface to 550 km at 400 m. While these length scales are longer than the length scales applied in our study, the availability of Argo data enables application of our chosen length scales.

Eulerian autocorrelation (temporal) scales have been calculated in much the same way as the correlation length scales. In this instance, pairs of floats that are separated by less than a set length scale (100 km) are binned according to their temporal separation. Values of $r^2$ of the anomaly temperature field were calculated for the data pairs in each time bin. The normalized correlations, plotted in Figure 4, show depth dependence, with e-folding scales of around 38 days in the surface 30 days at 100–200 m and 44 days at 1000 m. The dependence on depth varies between regions, but the mean depth timescale varies little, being close to 30 days in both the eastern and western subregions and close to 35 days in the subpolar North Atlantic. White [1995] also found a minimum in temporal length scales at 200 m.

In order for the OI scheme to perform an accurate interpolation of the initial field, the applied length scales need to be larger than or roughly equal to the minimum distance between data points [Piterbarg et al., 1991]. If the timescales and length scales applied are much less than the minimum distance between data points, the output field will simply be the mean field used in the OI. In order to test that the estimated correlation and autocorrelation length scales and timescales will effectively consider the sampled field,
where the terms on the left-hand side represent the local heat storage, the horizontal advection of heat, the velocity shear covariance term, and the vertical entrainment term. The notation employed is \( h \), the depth of the upper ocean; \( t \) time; \( T_a \) and \( v_a \), the temperature and horizontal velocity vertically averaged between the sea surface and depth \( h \); \( T \) and \( \bar{v} \), the deviations in temperature and velocity from the vertical average; \( T_{-h} \) and \( v_{-h} \), the temperature and velocity at depth \( h \); \( \nabla (\partial / \partial x, \partial / \partial y) \), the horizontal gradient; \( x \) and \( y \), the eastward and northward coordinates; \( w_e \), the entrainment velocity; \( \rho \), the potential density; \( C_p \), the specific heat capacity; \( Q \), the air-sea exchange of heat; and \( Q_{-h} \), the diffusive heat flux through the base of the layer. In this study we are concerned with determining the accuracy of mixed layer heat storage estimates from the Argo dataset. We focus on the left-hand side of equation (8). We will show below that the second and third terms in equation (8) are relatively small; hence we consider only the local heat storage and vertical entrainment terms in our analysis (equation (9)).

\[
\rho C_p \left[ h \frac{\partial T_a}{\partial t} + w_e (T_a - T_{-h}) \right]. \tag{9}
\]

The entrainment velocity \( w_e \), may be rewritten as

\[
w_e = \frac{\partial h}{\partial t} + v_{-h} \cdot \nabla h + w_{-h}, \tag{10}
\]

where \( w_{-h} \) is the vertical velocity at depth \( h \). Stevenson and Niiler [1983] set \( h \) equal to the depth of a fixed isotherm below the mixed layer depth. Such a surface approximates a material surface and the term in (10) is therefore minimized and can be neglected. There are two main limitations with applying such an approach in this study. The first lies in choosing which isotherm to consider, this is particularly difficult in the subpolar North Atlantic where the 0\(^\circ\)C isotherm outcrops and in the Gulf Stream region where surface temperature gradients are steep. The second and less arbitrary limitation of this method is that since the isotherm should lie beneath the mixed layer, errors are integrated over a larger area, thus reducing the accuracy of the estimated heat storage term. This is particularly true during the summer months when the mixed layer shoals considerably. For these reasons we chose to quantify the terms in (9) over the mixed layer, identified as the depth at which the temperature is more than 0.2\(^\circ\)C cooler than the 10 m temperature [De Boyer Montégut et al., 2004]. We should therefore include the entrainment terms in our analysis. Several earlier studies in the literature have also taken this approach [Wang and McPhaden, 1999; Foltz et al., 2003; Foltz and McPhaden, 2005].

[23] While it is possible to estimate a horizontal velocity field from the Argo dataset, our main concern is with the errors in heat storage arising from inaccuracies in the estimated temperature field based on a subsampled dataset. We therefore neglect the second two terms on the right-hand side of equation (10). In order to assess the importance of these terms, we have used the horizontal velocity field and mass conservation from the OCCAM model to estimate their magnitude. With mean velocities of 0.01 m s\(^{-1}\) and 20 m yr\(^{-1}\) in the horizontal and vertical,
respectively, and $\rho$ and $C_p$ set to constants of $1027 \text{ kg m}^{-3}$ and $3986 \text{ J kg}^{-1} \text{C}^{-1}$, the last two terms contribute 15% and 3% toward the entrainment term (equation (10)) and the heat storage (equation (9)), respectively.

[24] Heat is not conserved in our formulation since we do not consider all processes responsible for supplying heat to the region in equation (8). In particular we neglect horizontal heat advection. Although in some regions of the North Atlantic we expect this term to be large, Gill and Niler [1979] found that heat input is mainly stored locally in the North Atlantic and horizontal heat advection by the mean flow is not particularly important. We estimate the magnitude of the heat advection term using OCCAM for a box roughly $10^5 \times 10^2$, centered on $35^\circ \text{N}, 55^\circ \text{W}$, bounded to the north by the Gulf Stream. At this location this term is an order of magnitude less than the terms represented in equation (9). The contribution from the velocity shear covariance term, which is typically neglected in the literature due to difficulties in obtaining reliable estimates, is also thought to be relatively small [Swenson and Hansen, 1999].

4. Results

[25] In this section we present the main results of our analysis. We begin by investigating the typical uncertainty in Argo-based estimates of the ocean temperature field using measurements made on a hydrographic section at $36^\circ \text{N}$. We then consider whether the Argo array has a sufficiently high sampling density to obtain accurate estimates of mixed layer heat storage. Given the limited number of observational estimates of monthly heat storage, we have adopted a model-based approach, subsampling an eddy-permitting numerical model at the Argo resolution in order to determine the likely sampling errors in the Argo-based estimates. In addition, we also examine how errors in heat storage estimates vary with (1) the length of time over which storage is calculated, from 1 to 6 months and (2) the sampling density in the context of past and likely future numbers of available float profiles.

4.1. Accuracy of Argo-Derived Temperature Section

[26] We investigate the errors associated with using the Argo dataset to estimate upper ocean temperature with a focus on whether the dataset can usefully be employed to determine temperature across a typical North Atlantic hydrographic section. One of the aims here is to obtain some idea of likely errors in the Argo-based temperature field during periods when sections are unoccupied. For our analysis, Argo float temperature profiles are interpolated to positions of cruise conductivity-temperature-depth (CTD) stations in order to obtain a temperature section. Comparisons are made between the interpolated Argo section and the actual cruise CTD section. Although we have analyzed a single example cruise section, we expect that results will be generally applicable to other sections in regions with similar Argo sampling density. The cruise section considered is across $36^\circ \text{N}$, which was surveyed in May and June of 2005 [McDonagh et al., 2006]. The cruise track is shown in Figure 6. Positions of Argo float profiles interpolated to create the Argo-based estimate of the temperature section are also shown. In addition to showing the spatial distribution of floats used in the Argo-based section, Figure 6 provides an indication of the temporal spread of the data with dots and crosses representing profiles sampled within 30 days and more than 30 days before/after the cruise CTD, respectively. The spatial and temporal spread of Argo data covers an area of more than $2.5 \times 10^6 \text{ km}^2$ and a period of more than 5 months.

[27] The average station separation of cruise CTDs is 50 km; this is an order of magnitude smaller than the correlation length scales applied in the OI for the surface waters. Therefore the actual cruise section will resolve much smaller scale features than the Argo section. In order to remove some of this small-scale variability, the cruise section was smoothed using a 5-point moving average. It should be noted that despite this smoothing, the hydrographic section is likely to contain features due to internal waves and mesoscale eddies. The Argo dataset may also sample these small-scale features, thereby introducing an additional source of error. The error estimates of the Argo-based temperature field that we present are therefore an upper limit.

[28] We have plotted the signal (equation (5)) and noise (equation (6)) statistics across the cruise section in Figures 7a and 7b, respectively. It can be seen that the signal ($s$) generally exceeds the noise ($n$). This is also apparent from the signal-to-noise ratio (SNR) plotted in Figure 7c, with values of more than one throughout most of the North Atlantic. Schiller et al. [2004] tentatively suggest that a SNR of one or more provides statistically significant and useful information about Argo-sampled data in the ocean. The noise is less than 0.5°C in the eastern basin and below 1000 m in the western basin. However, there are two main regions centered at 500 m in the western basin where the noise exceeds 2°C. This estimate of the noise is probably pessimistic because it averages the differences over all station pairs where signal variance might contribute to the average [Fukumori and Wunsch, 1991].

[29] The isotherms across $36^\circ \text{N}$ based on the cruise CTD data and the interpolated Argo-based estimate of the temperature field are shown in Figures 8a and 8b. The Argo data is able to resolve much of the temperature structure evident in the cruise section, with isotherm depths similar between the two and the most notable differences near the western boundary. We have also examined the effects of varying the parameters in the OI scheme in an effort to better represent the physics in the western boundary and thus reduce the errors in the Argo-based estimate of the temperature field. To this end we have (1) reduced the correlation length scales from 500 km to 200 km and (2) employed a baroclinic potential vorticity function in which $H$ is taken to be the depth of the $15^\circ \text{C}$ isotherm instead of the full ocean depth. These sections are not plotted since neither of these alterations led to notable improvements in the level of agreement between the Argo-based and hydrographic temperature fields. Isotherms across $36^\circ \text{N}$ have been plotted using temperature data from the WOA for the month of May (Figure 8c). The climatological isotherms are much smoother than those in the cruise and Argo-derived sections.

[30] In order to quantify the level of agreement between the Argo-derived temperature field and the hydrographic measurements, difference plots of the Argo-based and WOA temperature fields relative to the cruise values are
Figure 6. Position of the cruise track (thick line) and the temperature profiles used to estimate the temperature field along the hydrographic section; circles indicate positions of profiles sampled within 30 days of the cruise CTD, pluses indicate positions of profiles sampled more than 30 days before or after the cruise CTD. The temporal spread of Argo data used in the OI spans 62 days before the cruise CTD to 72 days after the cruise CTD.

Figure 7. (a) Signal, (b) noise, and (c) the signal to noise ration (SNR) across 36°N (Cape Hatteras–Mediterranean). Figures 7a and 7b are in degrees Celsius and Figure 7c is unitless. Contours of 1 are shown in Figure 7c.
shown in Figure 9. Also shown is the difference between the hydrographic section and a temperature field obtained by applying the OI scheme to all the CTD station data in place of Argo profile data (Figure 9c). The latter calculation was carried out to give an indication of the errors associated with the OI scheme. Over much of the section, the Argo-based estimates of temperature agree with the cruise measurements to within 0.5°C. However, there are several regions in the 500–1000 m layer west of about 40°W where the differences exceed this value (Figure 9a). Furthermore at the western boundary, west of 74°W, the temperature is more than 2°C warmer in the Argo section than in the cruise section. As expected, the climatological values from the WOA typically show larger differences from the cruise section than in the cruise section. Although some of the observed differences between the cruise and Argo-based temperature fields in the western boundary may be due to aliasing caused by the long interpolation distances used in the OI, low sampling by Argo floats in the Gulf Stream, as shown in Figure 6, also contributes. The low sampling also indicates why the earlier use of shorter length scales (200 km instead of 500 km) in the OI did not reduce differences between the two temperature fields; in the absence of data collected within the applied length scales, the interpolated field relaxes to the WOA temperature field, which as already noted has large differences with respect to the cruise measurements in the western boundary.

[32] We have further quantified the level of agreement between the various temperature fields by calculating the RMS difference relative to the original cruise section data for each dataset as a function of station number averaged over the 0–1500 m depth interval (Figure 10a). The mean RMS difference across the whole section for the Argo based estimate (using the original OI parameters) is 0.58°C, while RMS differences at hydrographic stations in the Gulf Stream are more than 2°C. These RMS differences are lower than those found when the hydrographic data are compared with the WOA climatological values, for which the section average RMS difference is 0.84°C. Thus the Argo-based temperature estimates across the section are in better agreement with the hydrographic measurements than the WOA climatology by more than 0.25°C. For the upper 100 m, the improvement with Argo is more dramatic, the RMS difference being 0.56°C, compared to 1.13°C with WOA, i.e., there is an improvement by a factor of two. These results indicate that Argo can provide useful information on the temperature field at this latitude when no

Figure 8. Temperature sections across 36°N (Cape Hatteras–Mediterranean). Contours show temperature in 2.5°C intervals. (a) Determined from hydrographic CTD measurements made during the occupation of the section, (b) estimated from Argo float data using the Optimal Interpolation scheme, and (c) estimated from climatological temperature profiles in the WOA closest to the CTD stations. The dashed lines indicate discontinuities in the latitude at which the WOA temperature profile was taken, caused by changes in the latitude of the CTD stations.
Figure 9. Temperature section difference plots with respect to the hydrographic section data. Contours of 0.5, 1, 2, 3, 4, and 5 degrees Celsius are drawn. (a) Argo-hydrographic section, (b) WOA based estimates-hydrographic section, and (c) hydrographic section created by applying the OI scheme to the CTD data—original hydrographic section.

Figure 10. (a) Mean RMS temperature difference for the upper 1500 m at each station across the cruise section of the Argo (solid thick), WOA (solid thin), and interpolated cruise (dashed) data relative to the original cruise data; (b) depth-averaged temperature differences for the upper 1500 m at each station of Argo data relative to the original cruise data. Positive values indicate a warm bias in the Argo section.
The analysis of the 36°N cruise section provided an indication of likely errors in Argo-based estimates of the ocean temperature in relatively well-sampled regions. We now consider the accuracy of Argo-based estimates of heat storage and in particular their dependence on the number of available profiles. In the absence of detailed observational estimates of heat storage, we have adopted a model-based approach and have made use of the Ocean Circulation and Climate Advanced Model (OCCAM) [Webb et al., 1998; Coward and de Cuevas, 2005]. The OCCAM model is based on the Bryan-Cox-Semtner general ocean circulation model [Bryan, 1969; Semtner, 1976; Cox, 1984] and is a development of the Geophysical Fluid Dynamics (GFDL) Modular Ocean Model (MOM) code [Pacanowski et al., 1990]. The version of OCCAM considered is a 1/4° × 1/4° run for 1985–2003 with 66 levels in the vertical and full surface forcing based on a modified set of atmospheric variables from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis [Kalnay, 1996]. Several recent analyses have demonstrated that this model can be usefully employed to study the circulation of the North Atlantic [e.g., Marsh et al., 2005].

For our analysis, the model temperature field is subsampled to the Argo float locations and then interpolated onto a regular monthly 2° × 2° grid using the OI scheme. This is done for all months between January 2000 and March 2002. The solid line in Figure 11 shows the time series of mixed layer heat storage calculated using model temperature data subsampled at the Argo float locations. The dashed line uses fully sampled OCCAM temperature profiles. For each individual plot, the x-axis shows the month within 2002 and the y-axis the mixed layer heat storage change in W m⁻². The analysis of the 36°N cruise section provided an indication of likely errors in Argo-based estimates of the ocean temperature in relatively well-sampled regions. We now consider the accuracy of Argo-based estimates of heat storage and in particular their dependence on the number of available profiles. In the absence of detailed observational estimates of heat storage, we have adopted a model-based approach and have made use of the Ocean Circulation and Climate Advanced Model (OCCAM) [Webb et al., 1998; Coward and de Cuevas, 2005]. The OCCAM model is based on the Bryan-Cox-Semtner general ocean circulation model [Bryan, 1969; Semtner, 1976; Cox, 1984] and is a development of the Geophysical Fluid Dynamics (GFDL) Modular Ocean Model (MOM) code [Pacanowski et al., 1990]. The version of OCCAM considered is a 1/4° × 1/4° run for 1985–2003 with 66 levels in the vertical and full surface forcing based on a modified set of atmospheric variables from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis [Kalnay, 1996]. Several recent analyses have demonstrated that this model can be usefully employed to study the circulation of the North Atlantic [e.g., Marsh et al., 2005].

For our analysis, the model temperature field is subsampled to the Argo float locations and then interpolated onto a regular monthly 2° × 2° grid using the OI scheme. This is done for all months between January 2000 and March 2002. The solid line in Figure 11 shows the time series of mixed layer heat storage calculated using model temperature data subsampled at the Argo float locations. The dashed line uses fully sampled OCCAM temperature profiles. For each individual plot, the x-axis shows the month within 2002 and the y-axis the mixed layer heat storage change in W m⁻².
December 2003. The subsampled interpolated temperature field and the full OCCAM temperature field are used with equations (9) and (10) to obtain two sets of estimates of the heat storage. Although Argo data samples the upper 2000 m, the heat storage is only calculated over the mixed layer, since this is the part of the ocean whose properties are set directly by the atmosphere.

The difference between the heat storage calculated from the full and subsampled, interpolated temperature field is indicative of the combined error associated with sampling limitations of Argo and the OI scheme. Although the OCCAM model has a resolution of 0.25°C, the subsampled data has been gridded onto a 2°C grid to reduce computational cost. For consistency the full OCCAM temperature field is also subsampled onto a 2°C grid; henceforth this field is referred to as the fully sampled field to avoid confusion with the subsampled and then interpolated field (referred to simply as the subsampled field). It is recognized that use of a higher resolution could affect the results. However, owing to the high computational cost involved, this is not investigated in the present study. Both the fully sampled and the subsampled heat storage estimates are averaged into 10° × 10° boxes before comparisons are made as we anticipate that given the number of profiles currently available, Argo is unlikely to be able to quantify mixed layer heat storage at scales smaller than about 10 degrees.

Time series of mixed layer heat storage obtained from the fully sampled and subsampled temperature fields for each 10° × 10° box in the North Atlantic are shown in Figure 11. Only 1 year of data is shown. We plot the result for the year 2002. This year has been selected as it is the penultimate year of the OCCAM model run and so is the most recent year for which profiles from the following year can be considered in the OI. A clear seasonal cycle is evident in the majority of the boxes considered with small increases in heat storage during summer and larger reductions during winter (Figure 11). In many regions, differences between the heat storage over the mixed layer calculated from the subsampled and the fully sampled OCCAM temperature fields are small, of the order 10–20 Wm⁻². This indicates that the lower resolution associated with the Argo based subsampling does not introduce serious errors into the retrieved heat storage. This is particularly true of the area 20–30°N, 75–25°W and

Figure 12. Annual mean RMS difference between the subsampled and fully sampled OCCAM estimates of heat storage over the mixed layer based on (a) the subsampled temperature field and the MLD based on the subsampled temperature field, (b) the subsampled temperature field and the MLD based on the full temperature field, and (c) the full temperature field and the MLD based on the subsampled temperature field, (d) shows the difference between Figure 12a and Figures 12b and 12c (i.e., the nonlinearity). The main North Atlantic region and the subtropical subset are also shown in Figures 12c and 12d, respectively.
mean RMS difference north of 35°N increases to more than 100 Wm$^{-2}$ in the northwest North Atlantic.

[40] Decomposition of the RMS differences into errors associated with estimating the temperature field (Figure 12b) and errors associated with estimating the mixed layer depth (Figure 12c) indicates that the former is generally the higher source of error. This is particularly true of the Gulf Stream region where errors arising from inaccuracies in the estimated temperature field result in heat storage errors of more than 200 Wm$^{-2}$. The error arising from inaccuracies in the estimated mixed layer depth are typically less than 20 Wm$^{-2}$. However, within the Labrador basin where large fluctuations in mixed layer depth occur, the errors arising from estimating the mixed layer depth from a subsampled temperature field become large, exceeding 200 Wm$^{-2}$. It should be noted that the two sources of error are not linear. The discrepancy between the total RMS difference (Figure 12a) and the sum of the RMS differences arising from the two error sources (Figures 12b and 12c) is shown in Figure 12d. Nonlinear effects are small in the subtropics but exceed 100 Wm$^{-2}$ in the Labrador Sea.

[41] Adopting a similar approach to that taken for the hydrographic section analysis, we have applied the OI scheme to the fully sampled OCCAM temperature field to estimate the error associated with the interpolation method. The RMS difference between estimates of the monthly mixed layer heat storage determined from the fully sampled OCCAM output before and after the OI scheme is applied is shown in Figure 13. Values are less than 5 Wm$^{-2}$ over much of the study region south of 40°N but become large in the western boundary and subpolar region. The large RMS differences in the latter areas probably result from aliasing of temperature signals with length scales shorter than the applied 500 km correlation length scale. This result indicates the need for higher sampling densities (which will enable application of shorter length scales in the OI) if Argo is to be useful for identifying variability in these regions.

[42] We have also investigated the dependence of the OI scheme on the selected correlation length scales and timescales by running a number of experiments and briefly summarize the results here. For each experiment, one of the selected correlation parameters was altered, while the others remained constant. Comparison of the different output fields reveals the sensitivity of the scheme to the correlation parameters. The different experiments are detailed in Table 1. The standard deviation of the monthly mixed layer heat storage for 2002 obtained from the full range of experiments is shown in Figure 14. In the subtropical region, where the change in mixed layer heat storage can

**Figure 13.** RMS differences between the monthly mixed layer heat storage for 2002 determined from fully sampled OCCAM before and after the optimal interpolation scheme is applied.

30–40°N, 65–15°W (outlined in Figure 12d), henceforth referred to as the subtropical subset. However, further north and west the differences may exceed 50 Wm$^{-2}$, particularly in the winter when the mixed layer is deep. The boxes covering the region 40–50°N, 65–35°W and 50–60°N, 55–25°W show largest differences between heat storage estimated from the two different temperature fields. The heat storage in this region extends off the vertical scale in Figure 11 with values exceeding 900 Wm$^{-2}$ and 600 Wm$^{-2}$ based on the subsampled and fully sampled temperature fields, respectively. The large differences between the fully sampled and subsampled estimates in the northwest of the study area are probably caused by a combination of (1) the inability of the model field subsampled at Argo resolution to capture the large oscillations in monthly mixed layer depths that occur in this region and (2) the integration of errors over the deep winter mixed layers.

[38] To differentiate between these sources of error, the heat storage is estimated using two additional methods where the errors associated with estimation of the temperature field and estimation of the mixed layer depth are isolated. The first method uses the subsampled temperature field with mixed layer depth estimated from the full model temperature field (i.e., the errors arise from estimating the temperature field using the subsampled dataset). The second method uses the full model temperature field with the mixed layer depth estimated from the subsampled temperature field (i.e., the errors arise from estimating the mixed layer depth using the subsampled temperature).

[39] In order to show the spatial variation of the error associated with subsampling in more detail, the RMS differences for 2002 between the monthly mixed layer heat storage determined from the subsampled and fully sampled OCCAM temperature fields is calculated for $10^6 \times 10^6$ boxes centered on individual $2^6 \times 2^6$ boxes within the study region. The total RMS difference for the estimated heat storage is shown in Figure 12a. The RMS differences are less than 10 Wm$^{-2}$ at 15–25°N, 70–50°W and less than 20 Wm$^{-2}$ south of 35°N and west of 30°W. The total annual

**Table 1.** Experimental Runs Undertaken in Order to Determine the Sensitivity of the Interpolated Field to the Applied Correlation Length Scales

<table>
<thead>
<tr>
<th>Run</th>
<th>Lx, Ly, km</th>
<th>$\Phi$</th>
<th>$\Delta t$, days</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>500</td>
<td>0.25</td>
<td>30</td>
</tr>
<tr>
<td>2a</td>
<td>600</td>
<td>0.25</td>
<td>30</td>
</tr>
<tr>
<td>2b</td>
<td>200</td>
<td>0.25</td>
<td>30</td>
</tr>
<tr>
<td>3a</td>
<td>500</td>
<td>0.35</td>
<td>30</td>
</tr>
<tr>
<td>3b</td>
<td>500</td>
<td>0.15</td>
<td>30</td>
</tr>
<tr>
<td>4a</td>
<td>500</td>
<td>0.25</td>
<td>50</td>
</tr>
<tr>
<td>4b</td>
<td>500</td>
<td>0.25</td>
<td>20</td>
</tr>
</tbody>
</table>

*Values are parameters at the surface.*
indicating a strong dependence on OI parameters. The standard deviation of the monthly mixed layer heat storage is of order 10–20 W m⁻² or less for much of the subtropical North Atlantic, we now consider in more detail the relationship between this uncertainty and the number of Argo profiles available for the analysis. In order to do this, we have randomly subsampled the OCCAM model to different data densities. The sampling densities considered span the range 50–480 profiles per 10° × 10° box per year. This is equivalent to 42–420 floats in the North Atlantic. This range was chosen to cover the Argo float density during 1999 (110 floats) and the Argo target float density (344 floats). A total of 10 sampling densities were considered at regular intervals. For each of the 10 sampling densities, the OCCAM temperature field was randomly subsampled 10 times and then interpolated onto the full model grid using the OI scheme. RMS differences between the fully sampled OCCAM mixed layer heat storage and the subsampled interpolated OCCAM mixed layer heat storage were then calculated for each of the sampling densities. This was done for (1) the main North Atlantic region and (2) the subtropical subset of 10 degree boxes.

[45] The variation of the RMS difference between the monthly mixed layer heat storage for 2002 determined from the fully sampled OCCAM model and from interpolated subsampled model fields with number of profiles is shown in Figure 16. The RMS difference for the main North Atlantic region at the Argo target density is 25±1 W m⁻². The RMS difference values for the subtropical subset are smaller and fall by about a third from 18 W m⁻² at the 1999 Argo distribution to 12 W m⁻² at the target number of Argo floats. Error bars typically span 3–4 W m⁻² in the subtropics. This range in errors at a given sampling density is equivalent to the reduction in error in heat storage estimates associated with an increase in sampling from the 1999 resolution to that in 2004. This indicates that the distribution of profiles, as well as the number of profiles, affects the accuracy of Argo-based heat storage estimates.

Figure 14. The standard deviation of the monthly mixed layer heat storage for 2002 obtained from the full range of experiments with different OI parameters (See Table 1).

Figure 15. Variations in RMS difference between mixed layer heat storage determined from the fully sampled OCCAM model and from interpolated subsampled model fields with temporal scale over which the rate is estimated. Results are plotted for the main North Atlantic region (dashed) and the subtropical subset (solid). Error bars are two standard errors.

4.3. Sensitivity to the Sampling Density

[44] Having shown that the uncertainty due to sampling at the Argo resolution in estimated monthly mixed layer heat storage is of order 10–20 W m⁻² or less for much of the
The variation in sampling uncertainty obtained for the main North Atlantic region and the subtropical subset highlights the need for an array which is not simply random in nature but which addresses the requirements of a higher sampling density in regions of higher-temperature variability.

5. Discussion and Conclusion

In this study we have investigated the accuracy of North Atlantic temperature and heat storage fields from the Argo profiling float dataset. In particular we have considered the limitations placed by the number of available floats on the accuracy with which Argo can be used to estimate these fields. The work described here forms part of a broader study, which will investigate variability in heat storage and the mixed layer heat budget during the Argo period.

We have compared observations of the temperature field across 36°N obtained from a hydrographic section with the corresponding field derived for the same time as the section from Argo using an optimal interpolation method. Agreement between the hydrographic and Argo-based temperature fields to within 0.5 °C was typically found in the eastern basin with higher differences in the western basin, particularly within the boundary current where errors exceed 2 °C. Nevertheless, the difference between the hydrographic and Argo values was significantly smaller than the difference between hydrography and values from the WOA climatology. The improvement was particularly large in the upper 100 m, where the WOA RMS differences were more than twice as large as those from Argo. This result indicates that Argo data may be useful for capturing variability in the temperature field across zonal sections in the absence of a dedicated research cruise and thus demonstrates the potential for using Argo to monitor changes in ocean properties. We anticipate that this result will hold for other regions of the North Atlantic with similar sampling densities and thus that Argo offers exciting possibilities for investigating temperature variability that may be linked to changes in the meridional overturning circulation [Bryden et al., 2005].

Other published studies have employed a model-based approach to investigate the extent to which various sampling arrays can resolve the temperature field. Schiller et al. [2004] used a global version of MOM with an average resolution of 2° × 1°, while Guinehut et al. [2003] employed a 1/3° primitive equation model, which was subsequently extended to 1/6° by Guinehut et al. [2004]. All three studies found that the target Argo array could replicate a large amount of the seasonal variability in the temperature field; in particular, Guinehut et al. [2003] found that 82% of the signal variance could be accounted for by using a 3 × 3° sampling grid. However, unlike our analysis, these earlier studies all adopted a regular sampling array in their experiments. It is clear from Figure 2, which shows the distribution of Argo floats, that this is an unrealistic assumption and our results are the first that we are aware of to employ a sampling technique that reflects the actual float distribution.

Our analysis of subsampled temperature fields from the OCCAM model has shown that in the subtropical North Atlantic, the Argo project provides temperature data at a spatial and temporal resolution that results in a sampling uncertainty in mixed layer heat storage of order 10–20 W m⁻². The error gets smaller as the period considered increases and at seasonal timescales is reduced to 7 ± 1.5 W m⁻². Within the Gulf Stream and subpolar regions,
the sampling errors are much larger and thus the Argo dataset will be less useful in these regions for investigating variability in the mixed layer heat storage.

[51] We have also used randomly subsampled output from OCCAM to simulate varying numbers of Argo floats in order to investigate how the accuracy of Argo-derived mixed layer heat storage estimates is likely to vary with the number of floats. As anticipated, an increase in the number of profiles available for use in the optimal interpolation reduces the RMS difference between estimates of the heat storage from the fully sampled and subsampled versions of OCCAM. In the subtropics, this reduction is by a factor of 1/3 from 18 W m⁻² at the 1999 Argo resolution to 12 W m⁻² at the target resolution. In addition, it was found that the accuracy with which the Argo dataset could quantify the mixed layer heat storage was dependent on the distribution of the float profiles.

[52] In conclusion, our study has shown encouraging agreement, typically to within 0.5°C, between the Argo-based temperature field at 36°N and the corresponding field from a hydrographic section. Furthermore, the model analysis demonstrated that within the subtropical North Atlantic, sampling of the temperature field at the Argo resolution results in a level of uncertainty of around 10–20 W m⁻² at monthly timescales, falling to 7 W m⁻² at seasonal timescales. This is sufficiently small that it should allow investigations of variability in this region, on these timescales.

[53] Acknowledgments. The Argo data were collected and made freely available by the International Argo Project and the national program that contributes to it (http://wwwargo.ucsd.edu, http://argo.jcommops.org). Argo is a pilot program of the Global Ocean Observing System. We are grateful to Elaine McDonagh, the PSO of the 36°N cruise, for making the hydrographic section data available, to Andrew Coward and Beverly deCuesas for making output from OCCAM freely available, and to the reviewers for their invaluable feedback. This work is funded by a NERC rapid grant NER/S/S/2003/11906. The contribution by Josey was partially supported by a separate NERC rapid grant NER/T/S/2002/00427.

References

Oort, A. H., and T. H. Vonder Haar (1976), The observed annual cycle in the ocean-atmosphere heat balance over the northern hemisphere, J. Phys. Oceanogr., 6, 781–800.


R. Hadfield, J. J.-M. Hirschi, S. A. Josey, and N. C. Wells, School of Ocean and Earth Science, National Oceanography Centre, European Way, Southampton, SO14 3ZH, UK. (reh1@noc.soton.ac.uk)