

## XBT Temperature Errors during French Research Cruises (1999–2007)

G. REVERDIN

*L'OCEAN/IPSL, CNRS/UPMC, Paris, France*

F. MARIN

*LEGOS, IRD, Brest, Plouzané, France*

B. BOURLÈS

*LEGOS, IRD, Cotonou, Bénin, France*

P. LHERMINIER

*LPO, Plouzané, France*

(Manuscript received 16 July 2008, in final form 27 April 2009)

### ABSTRACT

Data from French cruises in 1999–2007, a period during which Deep Blue (DB) or T7 expendable bathythermographs (XBTs) were deployed, and for which ancillary temperature data are available in the northeast Atlantic and equatorial Atlantic regions, are examined. There was a total of 16 cruises with XBTs launched between conductivity–temperature–depth (CTD) stations; during most of these, as well as during three additional cruises that were also considered, intake temperature was measured. XBT data from two voluntary observing ships in the North Atlantic subpolar gyre for which intake temperature was measured were also investigated. There is an XBT cold bias due to stirring of a stratified upper layer by the ship, resulting in differences between XBT temperatures at 3–5 m and intake measurements. This is most pronounced for midlatitude spring or summer cruises, when it averages about 0.10°C. When these situations are removed, the comparisons clearly indicate positive biases in XBT temperature measurements in 1999–2006, with individual cruise averages generally between 0° and 0.1°C, and a tendency to have larger biases when surface temperature is high. In addition, a positive depth-estimate bias of the XBTs in the upper thermocline (on the order of 4 m) is identified, as well as a depth overestimation through the profile, averaging 1.7% (1.2%) for the equatorial (midlatitude) cruises (with respect to a previously published depth estimate).

### 1. Introduction

Data from expendable bathythermographs (XBTs) launched from ships form the core of the subsurface ocean temperature data available between the 1970s and the early 2000s (Levitus et al. 2005). There have been recent indications from studies on ocean heat content variability that the errors in temperature profiles derived from XBTs in the archived datasets are still large,

and that they might have evolved in time, inducing, if not corrected, spurious decadal variability (Willis et al. 2004; Willis et al. 2009; Wijffels et al. 2009; Gouretski and Koltermann 2007; V. V. Gouretski 2008, personal communication).

Most of the XBTs in the data archives were produced with Sippican probes (Lockheed Martin Sippican, Inc.), except in a few regions, as in the northwestern Pacific. In recent decades, a large part of the data was acquired numerically using an analog-to-digital (A-D) converter from Sippican. For a long time, XBT data have been known to be associated with both temperature biases and errors in the depth estimates. An early synthesis on the net effect of these errors on vertical profiles of temperature for early XBT datasets of the 1970s is

---

*Corresponding author address:* G. Reverdin, Laboratoire d'Océanographie et de Climatologie par Expérimentation et Analyse Numérique, Institut Pierre Simon Laplace, Université Pierre et Marie Curie, Case 100, 4 Pl. Jussieu, 75252 Paris CEDEX 05, France.

E-mail: gilles.reverdin@locean-ipsl.upmc.fr

DOI: 10.1175/2009JTECHO655.1

presented in Heinmiller et al. (1983). The specific effect of depth-estimation error was later analyzed by Hanawa and Yoritaka (1987), Hanawa and Yoshikawa (1991), and Hanawa et al. (1995), who recommended a new depth equation to correct the average bias in the depth estimates for each probe type. It was recently suggested from regional and global studies that this equation was not adequate for post-1995 XBT data (D. Snowden et al. 2008, unpublished manuscript; Wijffels et al. 2009), whereas Reseghetti et al. (2007) indicated for Mediterranean Sea profiles that the equation of Hanawa et al. (1995) is valid there.

Roemmich and Cornuelle (1987) investigated the error in XBT temperature and its digitization, and ways to reduce the error by calibrating each individual probe and letting the probe rest in a bucket of seawater before launch. They identified a net average temperature bias for T7 probes on the order of  $0.02^{\circ}\text{C}$  for uncalibrated data of the mid-1980s, albeit with a high individual scatter of at least  $0.07^{\circ}\text{C}$  and a slight increase of the bias from  $10^{\circ}$ – $15^{\circ}\text{C}$  to  $20^{\circ}$ – $25^{\circ}\text{C}$  (by roughly  $0.01^{\circ}\text{C}$ ). Gouretski and Koltermann (2007) suggested that temperature biases can be much larger and have recently been in the range of  $0.1^{\circ}$ – $0.3^{\circ}\text{C}$ . They also suggest that the temperature bias dominates the error budget of heat content changes over depth-estimate biases. This is larger than the uncertainty provided for the probes by Sippican ( $0.1^{\circ}\text{C}$ ); examples of probes with much larger errors are reported in different publications, but they are usually isolated cases (Hanawa et al. 1995; Roemmich and Cornuelle 1987). In recent (2004–05) data (Reseghetti et al. 2007), it was found that the temperature bias of Deep Blue (DB) XBTs increases with temperature by  $0.01^{\circ}\text{C}$  as  $T$  increases from  $12^{\circ}$  to  $22^{\circ}\text{C}$ . There are also indications that it might have changed in time (F. Reseghetti 2008, personal communication), with an increase from the early 1990s to 2004–05, but returning to a smaller value in a 2007 test. The larger temperature bias values in 2004–05 on the order of  $0.05^{\circ}$ – $0.06^{\circ}\text{C}$  were also found independently by D. Snowden et al. (2008, unpublished manuscript) in a recent intercomparison experiment. There might also be in the more recent data a slight dependency of the bias with pressure (a  $0.01^{\circ}\text{C}$  increase roughly between surface and 850 m), something that was already suggested in the earlier data (Roemmich and Cornuelle 1987). These bias estimates are all much less ( $0.02^{\circ}$  to  $0.06^{\circ}\text{C}$ ) than the range implied by Gouretski and Koltermann (2007).

In what follows, we will focus on DB and T7 XBTs that have been most used in the last two decades. We will investigate the temperature bias at 3–5-m depth in recent data collected in the Atlantic Ocean. Most of these data were collected during 18 hydrographic cruises

on French research vessels (from 1999 to 2008) during which T7 or DB XBTs were dropped between successive CTD stations or occasionally between expendable CTD (XCTD) drops (six cruises) (Table 1, Fig. 1). For most of these cruises (except two), we also have a reliable temperature measurement (to within  $0.01^{\circ}\text{C}$ ) from a sensor inserted in an intake pipe near the bow of the ship at a depth of 3–4 m. We also use sets of XBTs from three additional cruises for which CTDs and XBTs cannot be compared in a similar manner, but for which there is also a reliable intake measurement. We finally investigate sets of DB or T7 XBT data from two voluntary observing ships (VOSs) during which an intake temperature measurement was measured from water pumped at a depth of 4–5 m (VOS *Skogafoss* on line AX02 in 1994–96 with an intake on the side near the middle of the ship and VOS *Nuka Arctica* on line AX01 in 2005–May 2007 from an intake at the bow of the ship).

The analysis of the temperature bias will mostly rely on comparisons of temperature near the intake depth. Because the CTDs and XBTs were not dropped simultaneously (most of the time), we do not have the possibility of using details in the profiles to separate depth biases from temperature biases, as was done in Hanawa et al. (1995), for example. Then, after removing the estimated temperature biases, we will estimate average depth biases for each cruise for which comparison with CTDs is available, and compare these depth biases with other published results.

## 2. Data

### a. XBT measurements

The data are produced through a circuit including the XBT probe launched from the ship, a launcher, and an A–D interface board connected to it with the return loop through grounding to the ship's hull and the seawater. The probes considered here were all DB or T7 probes produced by Sippican (or, more recently, by Lockheed Martin Sippican). We will not differentiate these two types in this study, as they are identical for the sea-going part of the probe. Probes were commonly stored before launching inside the ship, in a space mostly air-conditioned for tropical cruises or warmed for winter midlatitude cruises or high-latitude VOS data (for a subset of VOS *Skogafoss* data, the probes were stored outside in the hour before launching). The launcher was usually a hand-held model, except on the VOS *Skogafoss*. The A–D interface board was usually either an MK12 or MK21 model. On the VOS *Nuka Arctica*, it was a French Protecno–Argo system. For these systems, we expect a small start-up electronic transient [on the order of 0.1-s  $e$ -folding time (Kizu and Hanawa 2002)]

TABLE 1. List of cruises, with ship, region, make of A–D converter board, dates of XBTs, and average temperature at 5-m depth in the DB-T7 XBTs of the cruise.

Cruise	Ship	Region	Converter type	Date XBTs	Avg $T$ (at 5 m)
Eq99	R/V <i>Atalante</i>	Eq Atlantic	MK12	16 Jul–18 Aug 1999	26.7°C
Eq2000	R/V <i>Atalante</i>	Eq Atlantic	MK12	27 Jul–19 Aug 2000	24.2°C
Pirata8	R/V <i>Atalante</i>	Eq Atlantic	MK12	25–27 Nov 2000	24.4°C
P0	R/V <i>Thalassa</i>	NE Atlantic	MK12	30 Sep–7 Oct 2000	19.2°C
P1	R/V <i>Thalassa</i>	NE Atlantic	MK12	5–23 Feb 2001	14.1°C
P1DTX	R/V <i>D'Entrecasteaux</i>	NE Atlantic	MK12	7–21 Feb 2001	14.2°C
P2	R/V <i>Atalante</i>	NE Atlantic	MK12	24 Mar–11 Apr 2001	14.1°C
P2DTX	R/V <i>D'Entrecasteaux</i>	NE Atlantic	MK12	27 Mar–10 Apr 2001	14.3°C
P3	R/V <i>Thalassa</i>	NE Atlantic	MK12	26 Aug–13 Sep 2001	21.4°C
Pirata10	R/V <i>Atalante</i>	Eq Atlantic	MK12	7–21 Dec 2001	26.7°C
Ovide2002	R/V <i>Thalassa</i>	N Atlantic	MK12	13–18 Jun 2002	9.4°C
Ovide2004	R/V <i>Thalassa</i>	N Atlantic	MK12	6–10 Jun 2004	10.3°C
Pirata12	R/V <i>Suroit</i>	Eq Atlantic	MK12	29 Jan–18 Feb 2005	28.0°C
Egee1	R/V <i>Suroit</i>	Eq Atlantic	MK12-MK21	10 Jun–4 Jul 2005	24.3°C
Egee2	R/V <i>Suroit</i>	Eq Atlantic	MK21	5–28 Sep 2005	25.3°C
Egee3	R/V <i>Atalante</i>	Eq Atlantic	MK21	30 May–27 Jun 2006	27.1°C
Egee4	R/V <i>Antea</i>	Eq Atlantic	MK21	20 Nov–1 Dec 2006	26.9°C
Egee5	R/V <i>Antea</i>	Eq Atlantic	MK21	7–22 Jun 2007	27.0°C
Egee6	R/V <i>Antea</i>	Eq Atlantic	MK21	7–24 Sep 2007	25.2°C
Ovide2008	R/V <i>Thalassa</i>	N Atlantic	MK21	20 Jun–8 Jul 2008	12.5°C
Pirata18	R/V <i>Antea</i>	Eq Atlantic	MK21	2–16 Sep 2008	24.7°C
AX02	VOS <i>Skogafoss</i>	N Atlantic	MK12	Apr 1994–Jul 1995	7.5°C
AX01	VOS <i>Nuka Arctica</i>	N Atlantic	Proteco	Apr 2004–Jul 2007	9.3°C

and thermal-mass effect of the probe, so that reported temperatures very close to the surface are influenced by the probe initial value. However, based on visual examination of late-night profiles and the storage conditions, we do not expect it to affect the temperature near intake depth at 3–4 m by more than 0.01°C.

All XBT profiles were collected with the ship moving, usually at 10 kt ( $5.14 \text{ m s}^{-1}$ ) for the research vessels and the VOS *Skogafoss*, and at 12–16 kt for VOS *Nuka Arctica*. This is different from other studies on XBT biases for which XBTs were dropped with the ship in station, and in a sense it is more representative of the common conditions of use from moving ships. Most of our conclusions will rely on the hydrographic cruises, but there are significant differences between the conditions of operation during the research cruises and on merchant vessels, which could result in systematic differences, and which is the reason why we also investigate two VOS datasets.

During the research cruises on the French research vessels, the launch from the lower ship deck is done close to the 2.5-m height above sea level recommended by Sippican. On the two merchant vessel datasets, XBTs were launched from the bridge (also near the aft of the ship) at heights above sea level of 12 m (VOS *Skogafoss*) and 18 m (VOS *Nuka Arctica*). This difference in launching height between research vessels and VOS might imply a change of initial speed in the water and therefore of

the extent of the initial-temperature transient and of depth estimates near the sea surface. [Reseghetti et al. (2007) found a small but not statistically significant sensitivity to launch height.] The XBT temperature data near the surface are also influenced by mixing in the wake of the ship.

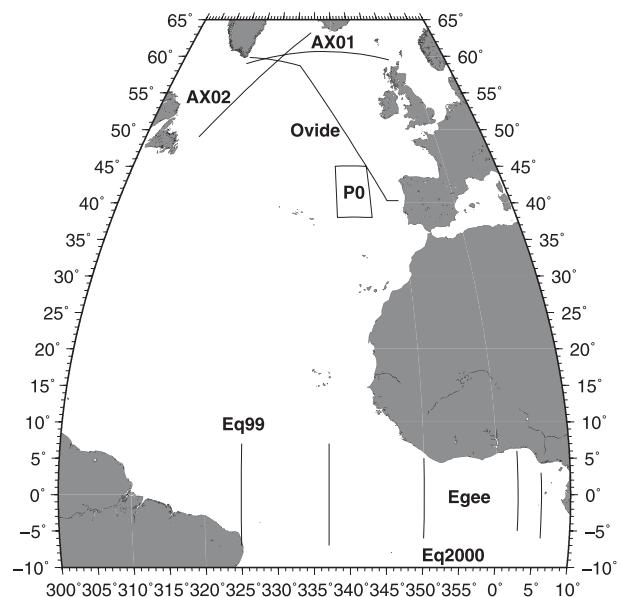


FIG. 1. Map of the Atlantic Ocean and of the regions sampled by the cruises and the VOS datasets (list in Table 1).

### b. CTD measurements

The CTD measurements are done on most of these cruises with a calibrated 911*plus* Seabird CTD mounted in a rosette, and temperature is expected to be accurate to within 0.001°C. The profiles are acquired as the ship is in station or slowly drifting because of wind or current constraints. We consider the downcasts done usually soon after the ship is in station, and we use reduced data transferred to the databases most commonly at a 1-dbar step. In some of the cruises in the tropical Atlantic that we use, data were interpolated upward from the first available depth, usually near 3–5-m depth, and we removed that nonmeasured part of the reported profiles. We also expect the near-surface part of the CTD profiles to be influenced by mixing induced by the ship and the rosette.

### c. XCTD measurements

In a few (six) cruises, we also use data from XCTD profiles, dropped as a substitute for CTDs (in the case of P1 and P2 on the R/V *D'Entrecasteaux*, when the weather was not considered good enough to do a CTD cast). The XCTDs are launched with the ship moving, as for an XBT (using an MK21 A–D converter board). The reported XCTD temperature is expected to be usually accurate to better than 0.05°C. The top 3–5 m of the profiles are usually not retained in the files we used, and depth is estimated on an XCTD in a similar way to the method used for an XBT assuming a drop rate equation. Systematic error in estimated depth was found to be less than 5 m above 800-m depth by Mizuno and Watanabe (1998).

### d. Intake temperature measurements

On research vessels, the intake temperature measurements are done usually in an intake water pipe located in the bow of the vessel, soon (up to 1 m) after the water enters the ship's bow. The measured temperature is usually not strongly influenced by a ship's warming with respect to outside seawater, except when flow rate in the pipe is very small (as on the R/V *Thalassa*). On the two merchant vessels we examined, the measurement was done somewhat farther (up to 3 m) from the intake within the pipe, and some small warming is expected. The intake is also located somewhat deeper below the sea surface on the two merchant vessels (4–5 m), and on the VOS *Skogafoss* it was not located at the bow, but on the side near the middle of the ship. For the VOS *Skogafoss*, we had the opportunity on a few occasions to verify its accuracy with measurements obtained from a towed Aquapack CTD system associated with a towed continuous plankton recorder [CPR; operated by the Sir

Alister Hardy Foundation for Ocean Science (SAHFOS), Plymouth, United Kingdom]. These indicated bad intake data in November 1994 and January 1995, and that in early October 1994, the intake temperature was 0.045°C too warm. (The same intake temperature sensor was used in 1994 and 1995, and we will assume that this provides an average estimate of the intake temperature warming.) We have no such comparison for the VOS *Nuka Arctica*, for which the temperature sensor was a calibrated Hart Scientific 1512 probe.

The temperature sensors on the research vessels were often SBE38 sensors calibrated beforehand, which should provide accurate readings to better than 0.01°C based on the later calibrations. We used the comparison with the CTD data as a way to verify how well it represents outside water temperature at the same depth and to correct possible biases. Based on that, we corrected on a nearly daily basis the intake temperature data from one cruise on the R/V *Thalassa* (P0), and corrected for an average bias in the intake temperature data for the four other R/V *Thalassa* cruises (P3, Ovide2002, Ovide2004, Ovide2008). These biases are probably caused by insufficient flow in the pipe. In addition, incorrect calibration coefficients were corrected on four R/V *Atalante* cruises (Pirata8, P1, P2, Egee3). For the remaining 12 cruises, the average difference is on the order of 0.01°C or less (usually, intake temperature was larger than CTD temperature at the same depth), and we assume that this intake measurement is unbiased. After corrections, 80% of individual CTD or XCTD temperatures at the intake depth are within 0.01°C of the corrected intake temperature.

## 3. Temperature and depth biases

### a. Surface temperatures

We report comparisons of XBT temperature with intake temperature ( $T_i$ ) at the same depth and of the 3–5-m averaged temperature of XBTs ( $T_{\text{XBT}}$ ) with the temperature interpolated linearly in time between successive CTD ( $T_{\text{CTD}}$ ) or XCTD ( $T_{\text{XCTD}}$ ) profiles.

Because we are dealing with a near-surface layer that can be stratified, we expect that profiles collected in conditions with near-surface stratification will present more scatter or even a different bias than when carried in a nearly homogeneous layer. The near-surface stratification estimated from the XBT or the CTD profiles depends somewhat on the area or the season. For example, for Eq99, the temperature decrease between 3 and 10 m averages less than 0.01°C, whereas for Eq2000, it averages 0.07°C. This later rather large near-surface stratification is typical of most cruises in the Gulf of

Guinea, whereas the midlatitude cruises present less temperature stratification in this layer, except in late spring or summer. Furthermore, the comparisons with the intake temperatures indicate, both for XBTs and for CTDs, instances where  $T_i$  is significantly larger than the profile temperature reported at that depth. This is particularly the case with the three late spring and summer midlatitude cruises for which between 20% and 30% of the profiles present large differences with  $T_i$  at intake depth. These profiles, collected mostly during daytime or early night and with weak wind, contribute to a negative cruise-averaged bias of  $T_{\text{XBT}}$  at the intake depth compared to  $T_i$  (respectively,  $-0.077^\circ$ ,  $-0.099^\circ$ , and  $-0.161^\circ\text{C}$  for P3, Ovide2002, and Ovide2004). During these cruises, we also find CTD profiles with large differences with  $T_i$ , and they contribute to a  $-0.030^\circ\text{C}$  average  $T_{\text{CTD}}$  bias. There is no doubt that these stratified near-surface layers are not correctly sensed by XBTs dropped behind the ship or even in CTD casts, because of mixing induced by the ship. There are also at least three other (tropical) cruises, where this effect of stratification is sensed in individual XBT profiles, with a net contribution on the order of  $-0.030^\circ\text{C}$  for the comparison with  $T_i$  (much less for CTDs).

In the following, we remove all instances of stratified layers (based on the XBT and CTD profiles, or anomalously large differences with  $T_i$ ) to estimate the XBT temperature bias for each cruise. Nevertheless, there is the possibility that residual stratification could contribute to average cold biases on the order of  $-0.01^\circ\text{C}$  at intake depth. To minimize the uncertainty on the estimated XBT temperature error, we will consider the median of the distribution of individual differences and not the average. (We also considered the average of the 10th–90th percentile values which were very close to the median estimate, and we associate the rms uncertainty estimated for this average to the median.)

The standard deviations of individual differences  $T_i - T_{\text{XBT}}$  are less than  $0.02^\circ\text{C}$  for all datasets, and therefore the uncertainties on the median difference are small for each set. An alternative to estimate  $T_{\text{XBT}}$  bias is to compare  $T_{\text{XBT}}$  with the interpolated CTD (or XCTD) profiles. However, the standard deviations of those differences often exceed  $0.1^\circ\text{C}$  at 3- or 5-m depth. This results in a less reliable estimate of XBT temperature bias. It also (not surprisingly) indicates that a large part of the differences between XBT and interpolated pair of CTDs (or XCTDs) originates from spatial or temporal variability. For each cruise, the different estimates of the XBT temperature bias are always compatible with the estimated uncertainties, with the largest differences happening for cruises P1 and P2 (Table 2). On cruises when there were comparisons with both CTDs and

XCTDs, the two results were also similar to within the uncertainties.

In the process of carrying this comparison, we found two sets of data that were not homogeneous (cruises Pirata12 and Egee1). For Egee1, the problem (not reported in the final dataset) originated from oxidized contacts in a switch box (resulting in added resistance and a negative temperature bias). This effect (averaging close to  $-1^\circ\text{C}$  at the surface, but highly variable) was corrected near the end of the cruise, and only the probes dropped after that are considered in Table 2. For Pirata12, there is also a subset of DB profiles on 5–7 February 2005 that present much smaller errors than before or after. [We refer to it as subset “b” (i.e., Pirata12b), whereas subset “a” (Pirata12a) is for the other probes that were either from the same DB batch or were T7s provided by NOAA or Coriolis.]

The estimated XBT temperature error is very variable from cruise to cruise (Table 2). Only two sets present temperature biases close to or larger than  $0.1^\circ\text{C}$  (Pirata8 and Pirata12a). The average bias is never negative, but cruises at lower SSTs tend to have smaller biases than those at higher SSTs in 1999–2006 (Fig. 2). Strangely enough, the 2007–08 cruises on board the R/V *Antea* in the warm equatorial Atlantic (Egee5, Egee6, Pirata18) also present a smaller bias. There is no relation with the A–D converter board (either MK12 or an MK-21, although the two cruises with serious problems, Pirata12 and Egee1, used an MK12 converter). Furthermore, comparisons by D. Snowden et al. (2008, unpublished manuscript) suggest that there should be little difference between different boards.

Interestingly, the comparison with the VOS *Skogafoss* data from 1994–95 indicates a small  $0.015^\circ\text{C}$  surface bias (Table 2). This assumes that the measured intake temperature was biased positive by  $0.045^\circ\text{C}$ , based on an October 1994 comparison with towed Aquadock CPR measurements. This tends to be less than on the research cruises we have examined but within the scatter of the different biases (notice also that the average temperature of  $7.5^\circ\text{C}$  is much less than for the other comparisons). The other VOS dataset (AX01) from 2004–07 indicates a  $0.02^\circ\text{C}$  positive  $T_{\text{XBT}}$  bias with respect to  $T_i$ . We do not know if the intake temperature measurement presents a positive bias, but such effect is likely, so the XBT temperature bias is at least  $0.02^\circ\text{C}$ , therefore probably larger than the estimated bias of the VOS *Skogafoss* data from 1994–95. This is consistent with the increase in temperature bias reported between the early 1990s and 2004–05 by Reseghetti et al. (2007), although here this is compounded by the use of different launchers and A–D converter boards. Notice also the outlier of Ovide2004 with a very low bias.

TABLE 2. XBT temperature bias, estimated either as the difference between XBT and ship intake temperatures ( $dT_i$ , line 1), or between XBT and CTD or XCTD temperature (respectively,  $dT_{ctd}$  and  $dT_{xctd}$ ). For each cruise, median estimate ( $m$ ), estimated error ( $e$ ), and number of data ( $n$ ) are reported on separate lines. Notice that for P1DXTX and P2DXTX, comparisons were done with CTD at 10-m depth, whereas for other cruises comparison is at 3–5 m.

Cruise		Eq1999 1999	Eq2000 2000	Pirata8 2000	P0 2000	P1 2001	P1DXTX 2001	P2 2001	P2DXTX 2001	P3 2001
$dT_i$	$m$	0.082	0.060	0.097	0.066	0.044		0.042		0.058
	$e$	0.005	0.004	0.010	0.008	0.004		0.004		0.007
	$n$	37	53	7	14	38		55		17
$dT_{ctd}$	$m$	0.091	0.063	0.080	0.079	0.006	0.025	0.020	0.054	0.060
	$e$	0.012	0.014	0.040	0.050	0.030	0.018	0.015	0.025	0.018
	$n$	45	48	9	9	25	57	24	33	18
$dT_{xctd}$	$m$	0.075	0.077						0.005	
	$e$	0.010	0.010						0.015	
	$n$	35	44						5	
Cruise		Pirat10 2001	Ovide 2002	Ovide 2004	Pirat12a 2004	Pirat12b 2004	Egee1 2005	Egee2 2005	Egee3 2006	Egee4 2006
$dT_i$	$m$	0.080	0.058	0.005	0.219	0.083	0.089	0.062	0.077	0.075
	$e$	0.005	0.005	0.005	0.007	0.004	0.006	0.003	0.004	0.002
	$n$	37	17	18	86	16	9	49	42	37
$dT_{ctd}$	$m$				0.216	0.105	0.120	0.055	0.060	0.080
	$e$				0.050	0.020	0.040	0.010	0.007	0.010
	$n$				6	6	7	51	44	38
$dT_{xctd}$	$m$						0.020		0.078	
	$e$						0.050		0.026	
	$n$						4		12	
Cruise		Egee5 2007	Egee6 2007	Ovide 2008	Pirat18 2008	AX02	AX01			
$dT_i$	$m$	0.053	0.030	0.055	0.045	0.015	>0.020			
	$e$	0.004	0.006	0.003	0.002	0.008	0.005			
	$n$	27	24	41	80	77	145			
$dT_{ctd}$	$m$	0.040	0.025							
	$e$	0.015	0.008							
	$n$	29	24							
$dT_{xctd}$	$m$		0.030							
	$e$		0.045							
	$n$		4							

*b. Subsurface temperatures*

We assume that the XBT temperature errors reported in Table 2 apply to all profiles of the same cruise and at all depths, and remove it from the data. We then investigate the profiles of differences between XBTs and pairs of CTDs (or XCTDs). For this, we first estimate XBT depth based on Hanawa et al.'s (1995) relationship (8 out of the 17 cruises that we considered were reported using the Sippican original relationship, and are thus first converted in the Hanawa et al. depth relationship). Then we interpolate linearly the two closest CTDs (XCTDs) to the position of the XBT. Usually, the XBTs were launched halfway between two CTD (XCTD) stations typically separated by 50 km and 3–4 h. Error on CTD depth is very small, whereas for XCTDs this is probably less than 5 m above 800 m (Mizuno and Watanabe 1998). Most of the resulting differences will originate from space–time variability (random contribution) and from errors in depth estimate or temperature, which could introduce systematic biases.

We summarize the comparisons for each cruise as an average profile of temperature differences (with an rms uncertainty) at 10-m step (see example in Fig. 3). Because of the large rms standard deviation in the temperature difference profiles (typically at least 0.1°C, even in the deepest part of the profiles, and more than 1°C in the upper thermocline), the uncertainties on the average profiles are typically at least 0.03°C in the deep part (see left side of Fig. 3 with separate scales in the upper 200 m and below that). For cruises with a small sample size, this can be much larger. Nonetheless, the resulting differences are usually much larger than the expected residual temperature errors (either from incorrect correction or a neglected pressure or temperature dependence of this correction; each of those was on the order of 0.01°C). We thus expect that the average profile of temperature differences results mostly from errors in depth estimates. In the right side of Fig. 3, we convert the temperature difference in a depth difference using the observed stratification (for P1, P2, and P1DXTX,

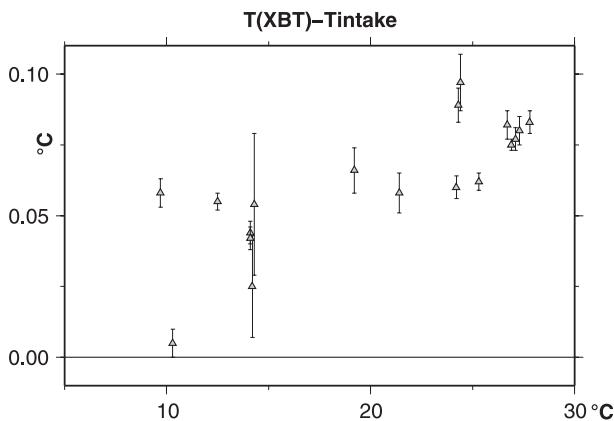


FIG. 2. Comparison of near-surface temperature from XBTs with corrected intake temperatures. For each cruise, the average difference and its associated error are indicated as a function of average intake temperature (see Tables 1 and 2). Notice that for the two “coldest” cruises Ovide2002 and Ovide2004, the intake temperature was corrected based on comparison with nonsimultaneous CTD measurements. The values for the two 2007 cruises Egee5 and Egee6, which exhibit smaller differences, have not been plotted.

stratification in the top 100 m is too small for such estimation).

In the upper thermocline, we commonly find positive biases in the region of maximum gradient, which indicates that depth is overestimated (a positive depth-estimate bias). This happens for cruises at different times (although not for the two 2007 cruises or P3 in 2001) and indicates an average depth bias of that layer of  $\sim 4.4$  m (Table 3), with individual cruise results scattered between 1.0 and 9.3 m. These estimated depth biases in the upper thermocline present little sensitivity to the actual temperature correction applied to the profile data. Their uncertainty is, however, often large (especially for cruises with a small sample size), as spatial variability will result in large differences at this depth between XBT profiles and the interpolated CTD profiles. Only a few cruises present significant differences from this average depth bias (P0 was larger, and Egee5 and Egee6 in 2007 were smaller).

The depth bias is much larger than what would result from the reported 0.1-s temperature sensor time constant (0.7 m) and is also larger than the 2-m depth bias reported by Reseghetti et al. (2007). This suggests that the actual descent rate of the probe just after having entered the water depends on the ship’s speed, possibly because probes do not enter the water as vertically when dropped from a moving ship. A 4-m difference was also reported from recent comparisons during a cruise (J. Gilson 2008, personal communication). During Eq99 and Eq2000, the depth differences found when comparing XBT and CTD (averaging 4.2 and 2.0 m, re-

spectively) are not found when comparing XBTs and XCTDs. This suggests that XCTDs might present the same depth bias as XBTs in the upper thermocline (they are deployed in the same conditions and use the same recorder board).

We now comment the profiles of temperature differences (or depth differences) below that. We separate the profiles of differences in XBT and CTD temperatures for a group of midlatitude northeast Atlantic cruises and a group of equatorial Atlantic cruises. In the first group (Fig. 4), P0 tends to deviate negatively (but with a large uncertainty due to the small sample size), and P3 tends to deviate positively. The other cruises are in between, but with a tendency to have an average positive deviation below 150 m. If interpreted as a depth bias, the average of the independent average estimates for each cruise yields a positive depth bias [even more when this is weighted by the number of individual comparison in each cruise (black line on Fig. 4)]. It therefore appears that Hanawa et al. (1995) tends to overestimate actual depth for these 2000–01 midlatitude cruises (this difference corresponds to an average 1.2% depth overestimation, but with clearly a large uncertainty).

In the second group (Fig. 5), all the cruises present positive biases (both in temperature and depth estimates). The two sets Pirata12a and -12b are close together (only Pirata12b is shown) and provide very high positive biases, which are quite separate from the others. The estimates for Pirata8 and Egee1 also show fairly large positive values. However, they are based on few profiles, therefore with large uncertainty estimates due to the small sample size. The scatter between the other cruises is not significant. When interpreted as depth-estimate biases, this shows usually an increased bias with depth. This can be summarized by an average residual profile of the 10 cruises (the average black curve with individual cruises weighted by the number of profiles) showing a positive bias that increases regularly with depth. Thus, the Hanawa et al. (1995) fall-rate equation seems to overestimate actual depth by 1.7% according to this average. This is less than the 3.36% of the Hanawa et al. correction from the “original” Sippican equation that is incorporated in the data. Of course, this is also sensitive to the accuracy of the correction in temperature that was applied to the whole profile based on the near-surface comparisons (a  $0.01^{\circ}\text{C}$  temperature converts into a  $\sim 2$  m depth error at 800 m for these cruises; indeed, if we were not correcting XBT temperature beforehand, the estimated depth overestimation would be larger by typically 10 m at 500–850-m depth).

We also find that the comparisons between XBTs and XCTDs usually support the sign of the average depth estimate bias below 300 m (well above the error in

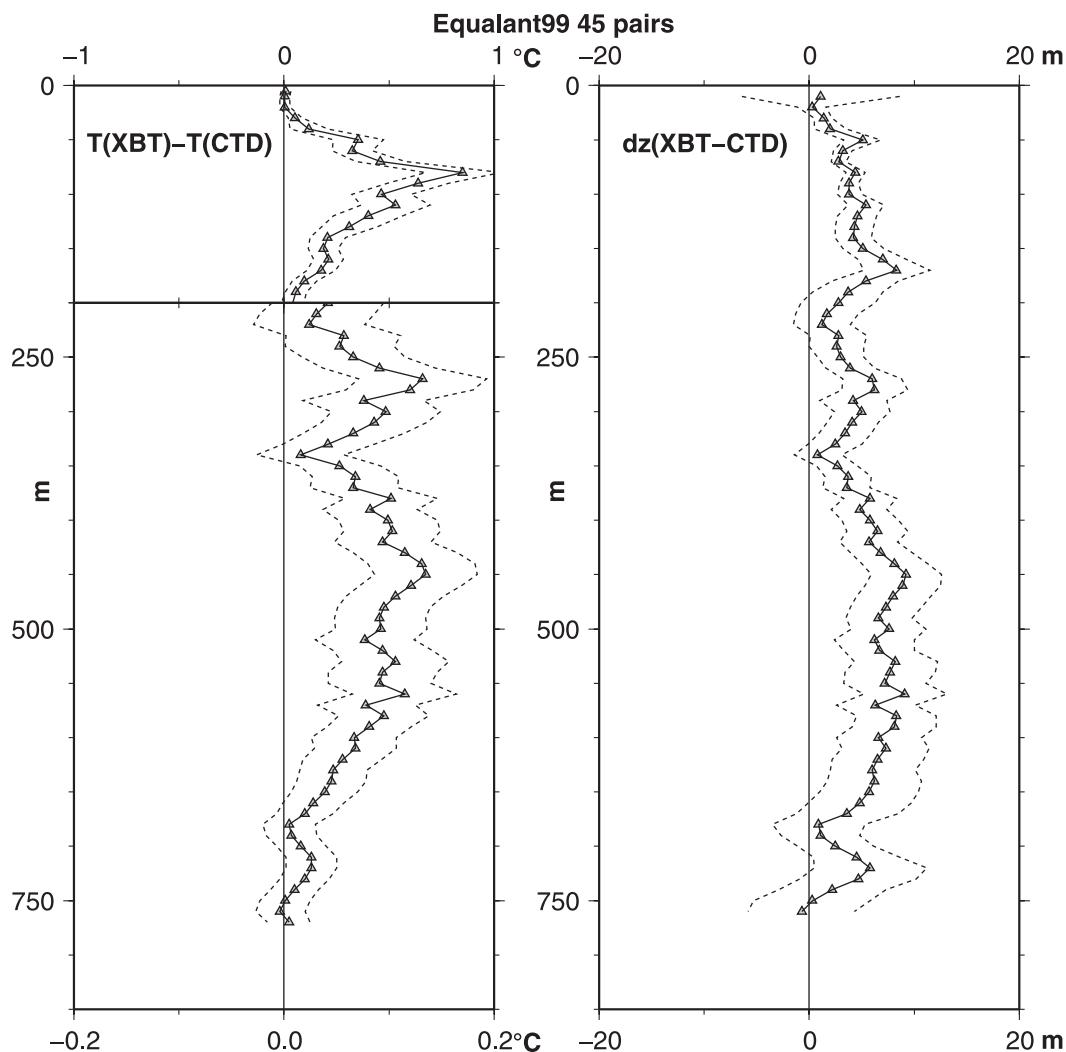


FIG. 3. Comparison for cruise Equalant99 of XBT temperatures with interpolated temperatures (after removing the estimated temperature bias): (left) average temperature difference and its associated error as a function of depth (with a change of scale at 200 m); (right) interpretation of those differences as a result of depth differences, assuming no temperature error (both average and rms error as a function of estimated depth).

the average temperature bias correction applied, at least for Eq99 and Eq2000). We commented on some cruise-to-cruise variability in the depth estimates. For Pirata12, we found (previous section) unusually large surface temperature biases, which suggests that the converter board might have been faulty. It is likely for that cruise that we did not correct fully the temperature biases by removing the average surface-based temperature bias. Based on the expectation that the depth-bias estimate also applies to this dataset, we could use the residuals to estimate an average temperature-dependent bias for that cruise. We have, however, no explanation for the changes between the two subsets Pirata12a and -12b.

#### 4. Conclusions

We find a significant near-surface XBT temperature bias, but with large cruise-to-cruise differences. The near-surface XBT temperatures (at 3 m or deeper) miss the near-surface warming sensed sometimes at the same depth by probes placed in intake pipe near the bow of the ship. The resulting negative bias could be of  $-0.10^{\circ}\text{C}$  in spring and summer at midlatitudes based on three cruises (the stratified layers are also partially missed by CTDs with a resulting average bias of  $-0.03^{\circ}\text{C}$ ). When removing these stratified situations, we estimate a cruise-average temperature bias that we apply to the whole profile. We usually find smaller temperature biases for

TABLE 3. Biases in upper-thermocline depth. Average by cruise of comparisons of the temperature of individual XBTs with interpolated CTDs at the depth of the upper thermocline. They are converted as a depth difference, assuming no temperature bias.

Cruise	Depth of upper thermocline (m)	No. of XBTs	$[T(\text{XBT}) - T(\text{CTD})]$ average ( $^{\circ}\text{C}$ , rms)	$Z(\text{XBT}) - Z(\text{CTD})$ average (m, rms error)
Eq99	50–80	45	0.267 (1.79)	4.2 (1.2)
Eq2000	20–40	55	0.220 (0.70)	2.0 (0.8)
Pirata8	60	9	1.02 (1.40)	8.0 (3.1)
P0	40	8	0.96 (0.80)	9.3 (1.3)
P2DTX	30	46	0.06 (0.25)	4.9 (3.3)
P3	40	24	0.2 (1.5)	1.0 (1.6)
Pirata12	50	6	1.20 (0.90)	8.0 (3.0)
Egee1	20	5	0.77 (0.83)	5.9 (2.8)
Egee2	40	51	0.31 (0.80)	2.4 (0.9)
Egee3	30–40	45	0.40 (0.68)	2.5 (0.5)
Egee4	40–50	38	0.48 (1.65)	3.5 (1.2)
Egee5	40	29	0.14 (0.97)	2.3 (0.8)
Egee6	40	28	0 (1.13)	1.3 (1.9)

cruises in temperate and high latitudes than for the equatorial Atlantic cruises. We cannot provide a reliable time history of temperature biases based on these comparisons because of two cruises that present a significantly different temperature bias (we have no particular explanations to offer for the large positive bias observed during part of the Pirata12 cruise). The other cruises do not suggest that temperature biases changed in the equatorial Atlantic between 1999–2002 and 2005–06, but the last three cruises suggest a decrease in 2007–08. There are also indications from earlier publications that temperature biases might have been less prior to the mid-1990s. We checked that on one cruise in 1983 in the equatorial Atlantic for which we had a comparable set of comparisons (FLOT on R/V *Marion Dufresne*), which indicates no significant temperature bias.

After correcting XBT temperatures from these estimated biases and using the Hanawa et al. (1995) depth equation, we compared the profiles with nearby CTD profiles. This indicates that XBT depth is often overestimated in the upper thermocline, located between 20 and 70 m in the set of cruises considered (all the XBTs were launched with the ship steaming). This 3–4-m average bias should also affect mixed layer depth, which would be overestimated by typically 5%–10% from XBTs. This effect is found for recent cruises, but we expect that this would also apply to earlier data. However, as the conditions of launching are different on VOS lines, it is not clear whether this estimated bias also applies to this larger set of data. What biases apply might be an issue for studies on interannual variability in mixed layer depth and other upper-ocean properties that often mix different types of data (de Boyer Montégut et al. 2004).

Except for one very anomalous cruise (Pirata12), the residual temperature differences found at depth are not very large and convert into positive depth-estimate

biases. This seems to vary a little from cruise to cruise, but the average of all these cruises is positive, both for the midlatitude cruise group and the equatorial Atlantic cruise group (for the equatorial cruises, a rather regular increase of depth bias with depth, suggestive an overestimation of 1.7%). This difference corresponds to what is suggested in recent dedicated intercomparison experiments (D. Snowden et al. 2008, unpublished manuscript). It is much less than in the statistical analysis of Wijffels et al. (2009) (~3.1%), but they did not correct for temperature bias, which has the right magnitude to explain the difference. Not having corrected for a temperature bias in their study, however, means that their corrected profiles are distorted at intermediate-near-surface levels compared to real ones. Because this cannot correct for  $T$  biases in the surface layer or in other weakly stratified layers in the upper ocean, this would also induce a small error in vertically integrated heat content. Assuming the temperature bias has changed by  $0.03^{\circ}\text{C}$  over 10 years, this could translate at most to an equivalent error in heat fluxes on the order of  $0.2 \text{ W m}^{-2}$ , which is small compared to expected net imbalances in air–sea heat flux, which exceed  $1 \text{ W m}^{-2}$ .

The depth error we estimate with respect to the Hanawa et al. (1995) fall-rate equation is, however, much larger than the one found in recent deployments in the Mediterranean Sea (Reseghetti et al. 2007). This could underline regional differences in the fall rate. The Hanawa et al. (1995) relationship was also deemed valid for the Indian Ocean by Thadathil et al. (1998), but for earlier data (cruises in 1994–97) close to when the change in depth bias was estimated by Wijffels et al. (2009). To check whether this transition also applies to the French cruise data, we will in future work attempt to apply the same methodology to a set of French cruises in the 1990s.

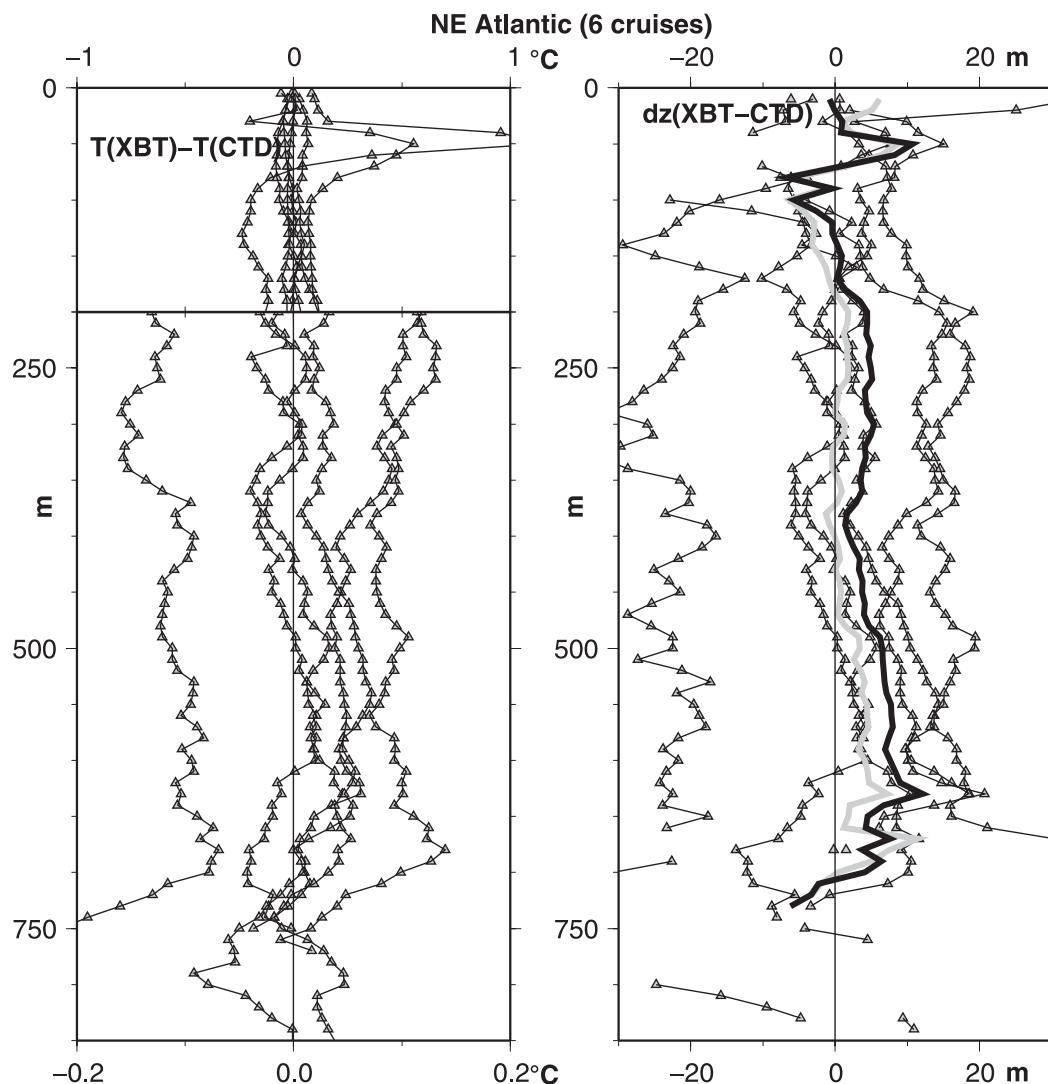


FIG. 4. (left) Average temperature difference profiles (with a change of scale at 200-m depth) and (right) average depth difference profiles for the six Programme d'Océanographie Multi-Disciplinaire Méso-Échelle (POMME) cruises in the northeast Atlantic. On the right, the gray line represents an average of these individual estimates, and the black line is an average weighted by the number of profiles in individual cruise estimates. Cruise P0 corresponds to the very negative profile and is based on a small number of profiles.

We find smaller temperature biases, both at surface and depth, than what is outlined in Gouretski and Koltermann (2007). However, the strong positive residuals for one cruise (Pirata12) clearly deviate from the others and are in the range proposed by Gouretski and Koltermann (2007). We also found one cruise with data transmitted (Egee1) that often presented large negative temperature biases due to poor contacts (but we did not include those data in the comparisons). The occurrence of such biased data (either positive or negative) clearly degrades the overall quality of the French cruise XBT dataset. It is not altogether clear what causes these errors that are not easily identified by usual quality control

of the data, or whether what we found here in the French research cruise datasets is representative of what is done in other countries during research cruises, on VOS lines, or by other users (navies). Such large errors or partially bad profiles left in the databases, or differences in the fall-rate equations used, could explain the differences with Gouretski and Koltermann (2007).

*Acknowledgments.* We are very grateful for the devoted effort of numerous colleagues to deploy XBTs during oceanographic cruises and to make these datasets available. The Service Hydrographique et Océanographique de la Marine (SHOM) provided the XBTs to

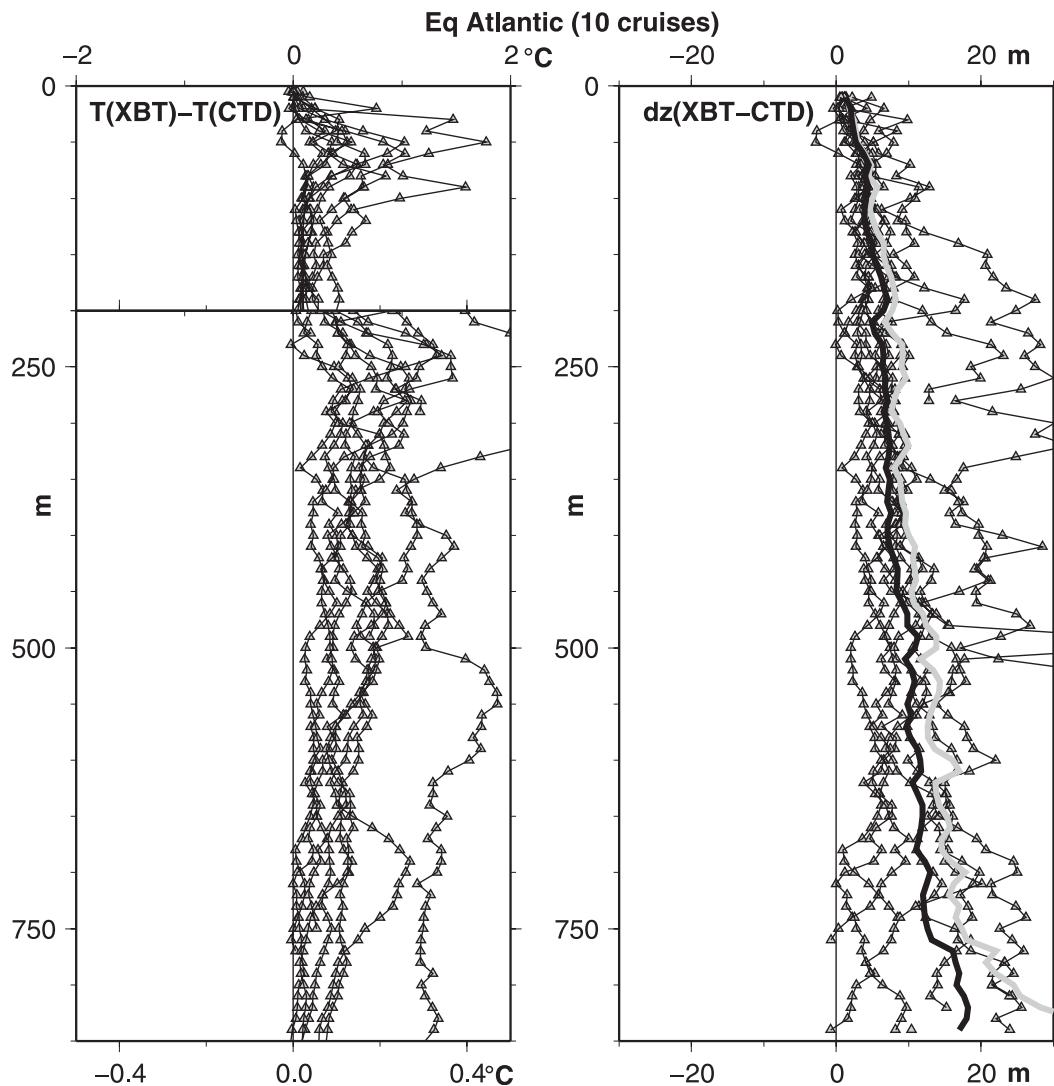


FIG. 5. Same as Fig. 4, but for 10 equatorial cruises (1999–2007). The large positive values correspond to cruises Pirata12b (Pirata12a is not plotted), Pirata8, and Egee1.

the six POMME cruises in the northeast Atlantic; and French national programs, Coriolis and the Institut National des Sciences de l'Univers (INSU), provided XBTs during the other cruises. UMS-025 of IRD provided supervision and personal support for the use of XBTs during the tropical Atlantic research cruises and on the VOS *Nuka Arctica*. The NOAA/Atlantic Oceanographic and Meteorological Laboratory (AOML) provided support for the XBTs and work on the VOS *Skogafoss*. This study was initiated and supported by Coriolis.

#### REFERENCES

- De Boyer Montégut, C., G. Madec, A. Fischer, A. Lazar, and D. Iudicone, 2004: Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res.*, **109**, C12003, doi:10.1029/2004JC002378.
- Gouretski, V. V., and K. P. Koltermann, 2007: How much is the ocean really warming. *Geophys. Res. Lett.*, **34**, L01610, doi:10.1029/2006GL027834.
- Hanawa, K., and Y. Yoritaka, 1987: Detection of systematic errors in XBT data and their correction. *J. Oceanogr.*, **43**, 68–76.
- , and Y. Yoshikawa, 1991: Re-examination of depth error in XBT data. *J. Atmos. Oceanic Technol.*, **8**, 422–429.
- , P. Rual, R. Bailey, A. Sy, and M. Szabados, 1995: A new depth-time equation for Sippican or TSK T-7, T-6 and T-4 expendable bathythermographs (XBT). *Deep-Sea Res. I*, **42**, 1423–1451.
- Heinmiller, R. H., C. C. Ebbesmeyer, B. A. Taft, D. B. Ilson, and O. P. Nikitin, 1983: Systematic errors in expendable bathythermograph (XBT) profiles. *Deep-Sea Res.*, **30**, 1185–1196.
- Kizu, S., and K. Hanawa, 2002: Start-up transients of XBT measurement. *Deep-Sea Res. I*, **49**, 935–940.
- Levitus, S., J. Antonov, and T. Boyer, 2005: Warming of the world ocean, 1955–2003. *Geophys. Res. Lett.*, **32**, L02604, doi:10.1029/2004GL021592.

- Mizuno, K., and T. Watanabe, 1998: Preliminary results of in-situ XCTD/CTD comparison test. *J. Oceanogr.*, **54**, 373–380.
- Reseghetti, F., M. Borghini, and G. M. R. Manzella, 2007: Factors affecting the quality of XBT data—results of analyses on profiles from the western Mediterranean Sea. *Ocean Sci.*, **3**, 59–75.
- Roemmich, D., and B. Cornuelle, 1987: Digitization and calibration of the expendable bathythermograph. *Deep-Sea Res.*, **34**, 299–307.
- Thadathil, P., A. K. Ghosh, and P. M. Muraleedharan, 1998: An evaluation of XBT depth equations for the Indian Ocean. *Deep-Sea Res. I*, **45**, 819–827.
- Wijffels, S. E., J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K. Ridgway, and J. A. Church, 2009: Changing expendable bathythermograph fall-rate and their impact on estimates of thermosteric sea level rise. *J. Climate*, **22**, 5657–5672.
- Willis, J. K., D. Roemmich, and B. Cornuelle, 2004: Interannual variability in upper ocean heat content, temperature, and thermosteric expansion on global scales. *J. Geophys. Res.*, **109**, C12306, doi:10.1029/2003JC002260.
- , M. Lyman, G. C. Johnson, and J. Gilson, 2009: In situ data biases and recent ocean heat content variability. *J. Atmos. Oceanic Technol.*, **26**, 846–852.