The Impact of Improved Thermistor Calibration on the Expendable Bathythermograph Profile Data

MARLOS GOES

Cooperative Institute for Marine and Atmospheric Studies, University of Miami, and National Oceanic and Atmospheric Administration/Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida

ELIZABETH BABCOCK

Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida

FRANCIS BRINGAS

National Oceanic and Atmospheric Administration/Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida

PETER ORTNER

Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, Florida

GUSTAVO GONI

National Oceanic and Atmospheric Administration/Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida

(Manuscript received 9 February 2017, in final form 9 June 2017)

ABSTRACT

Expendable bathythermograph (XBT) data provide one of the longest available records of upper-ocean temperature. However, temperature and depth biases in XBT data adversely affect estimates of long-term trends of ocean heat content and, to a lesser extent, estimates of volume and heat transport in the ocean. Several corrections have been proposed to overcome historical biases in XBT data, which rely on constantly monitoring these biases. This paper provides an analysis of data collected during three recent hydrographic cruises that utilized different types of probes, and examines methods to reduce temperature and depth biases by improving the thermistor calibration and reducing the mass variability of the XBT probes.

The results obtained show that the use of individual thermistor calibration in XBT probes is the most effective calibration to decrease the thermal bias, improving the mean thermal bias to less than 0.02°C and its tolerance from 0.1° to 0.03°C. The temperature variance of probes with screened thermistors is significantly reduced by approximately 60% in comparison to standard probes. On the other hand, probes with a tighter weight tolerance did not show statistically significant reductions in the spread of depth biases, possibly because of the small sample size or the sensitivity of the depth accuracy to other causes affecting the analysis.

1. Introduction

Expendable bathythermograph (XBT) data have provided an invaluable historical record of global upper-ocean temperature, and they still play a significant role in monitoring cross-transect currents and heat transport at mesoscale spatial resolution and on time scales up to decades. The importance of XBT data to the global inventory of temperature profiles results from their easy deployment and low cost. In an XBT profile, the depth $z(t)$ is estimated using a fall-rate equation (FRE):

$$z(t) = At - Bt^2, \tag{1}$$

where the coefficients A and B are both positive and dependent on the XBT type, and t is the time since the probe hits the water. Coefficient A is related to the

Corresponding author: Marlos Goes, marlos.goes@noaa.gov

DOI: 10.1175/JTECH-D-17-0024.1

terminal velocity of the probe, while coefficient B accounts for probe weight loss as the wire uncoils. Temperature is measured by a thermistor located at the probe's nose. As water passes through the nose, the resistance value in the thermistor is recorded and processed by the acquisition system and translated into a temperature record.

Systematic errors have been discovered in XBT data since the 1960s (Hazelworth 1966; Flierl and Robinson 1977; Seaver and Kuleshov 1982). A large effort by the scientific community has been dedicated to quantifying these errors by comparing XBT data with conductivity–temperature–depth (CTD) temperature profiles (Flierl and Robinson 1977; Anderson 1980; Hallock and Teague 1992), satellite altimetry observations (DiNezio and Goni 2010), and high-resolution bathymetry data (Good 2011; Gouretski 2012), among others.

A consensus had been achieved within the oceanographic community to update the coefficients of the FRE [Eq. (1)] provided by the manufacturer (NOAA 2002) with those derived from the comparisons of hundreds of pairs of XBT and CTD profiles (Hanawa et al. 1995, henceforth H95). More recent studies, however, have shown that these updated coefficients could be further improved, as discrepancies were found between ocean heat content estimates from numerical models and those calculated using historical XBT data corrected with the H95 coefficients (Bindoff et al. 2007). These discrepancies were partially explained by the detection of time-variable XBT biases (Gouretski and Koltermann 2007). Further studies revealed that XBT biases consist of systematic depth errors and an independent temperature bias (e.g., Gouretski and Reseghetti 2010; Cowley et al. 2013; Cheng et al. 2016).

Corrections in the FRE must take into account several factors: 1) new FRE coefficients that are time dependent (H95; Gouretski and Reseghetti 2010; Wijffels et al. 2008; DiNezio and Goni 2011; Cowley et al. 2013; Cheng et al. 2014), temperature dependent (Thadathil et al. 2002; Kizu et al. 2005; Cheng et al. 2014) and probe type dependent (Gouretski and Reseghetti 2010; Kizu et al. 2011; Cowley et al. 2013); 2) pure temperature biases independent from depth estimates (Cowley et al. 2013; Heinmiller et al. 1983; Reseghetti et al. 2007; Roemmich and Cornuelle 1987; Gouretski and Reseghetti 2010; Hamon et al. 2012; Cheng et al. 2014); and 3) depth offsets caused by the initial velocity of the XBTs in the water as a result of the deployment height or the conditions of the probe entry in the water (Gouretski and Reseghetti 2010; Cowley et al. 2013; Cheng et al. 2014; Bringas and Goni 2015; Abraham et al. 2014; Gorman et al. 2014; Shepard et al. 2014). Because of the multiplicity of these factors, the development of correction schemes has mostly relied on the constant assessment of errors using side-by-side XBT and CTD deployments, which can be very time consuming and also dependent on the quality (and actual comparability) of the data and the particular method used in the analysis (Hamon et al. 2012; Cheng et al. 2016).

Efforts to produce a ''climate quality'' XBT probe are underway, and some ideas proposed include adding one or more pressure switches (Goes et al. 2013) to reduce depth biases and to improve thermistor calibration (Reseghetti et al. 2007), and applying stricter controls upon probe weight and shape (Kizu et al. 2011). Such technical improvements could potentially reduce the need for the continuous development of bias corrections.

In collaboration with Lockheed Martin/Sippican (LMS), the largest manufacturer of XBT probes, NOAA/AOML performed several side-by-side XBT and CTD deployments. The XBT probes used were the Deep Blue model, which is currently the one most utilized for oceanographic purposes (Cheng et al. 2016). A subset of the probes featured tighter controls of their physical properties in addition to better calibrations, which are expected to improve the accuracy of their temperature and depth estimates.

The main objective of this paper is to examine the potential of such physical and calibration improvements to reduce systematic errors in XBT temperature and depth estimates. This manuscript is organized as follows. In section 2 we explain the cruise data collected, probe properties, and corrections. In section 3 we combine all the cruise data and examine the significance of the temperature and depth bias reductions. In section 4 we present our conclusions and recommendations.

2. Data and methods

a. Data

The data used in this study were collected during three hydrographic cruises in the North Atlantic (Fig. 1). In the first cruise, carried out in February 2012 for the Western Boundary Time Series project (WBTS2012), 21 standard Deep Blue (DB) probes [serial numbers (S/N) 1182082–1182105] and 22 DB probes in which the standard thermistors were replaced with specially screened thermistors (so-called experimental probes; S/N 1182106–1182129) were deployed along six CTD stations. The screening process guaranteed that the residual difference between the measured and bath temperatures (T_{bath}) was smaller than 0.05°C. This experiment aims to quantify the temperature bias reduction as a result of improvements in the thermistor physical properties.

The second experiment was performed in November/December 2013 during the Prediction and Research

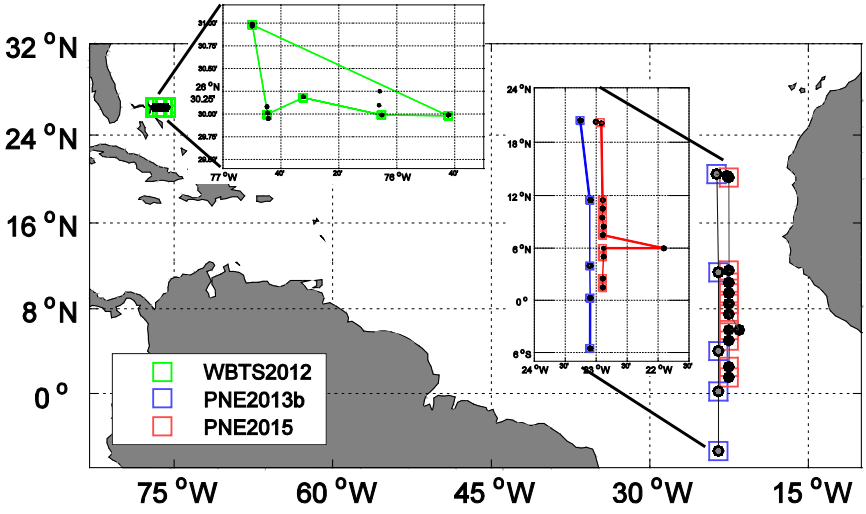

FIG. 1. Location of the side-by-side deployments. Three cruises analyzed include WBTS2012 (green squares), PNE2013b (blue squares), and PNE2015 (red squares). Squares represent the locations of the CTD casts and the dots the XBT deployments.

Moored Array in the Tropical Atlantic (PIRATA) Northeast Extension cruise (PNE2013b). In this experiment 96 DB XBT probes were deployed, collocated with CTD stations, comprising three types of probes: 1) standard (S/N 1212792–1212815), 2) experimental (S/N 1212720–1212743), and 3) ''tighter weight tolerance'' (TWT; S/N 1212744–1212791). The TWT probes featured screened thermistors identical to the experimental probes but with an improved weight tolerance (reduced variance) of the probe. The nominal weight of the Deep Blue probe nose cones is 575 g, with a standard weight tolerance of ±1.0 g according to the manufacturer's specifications (G. Johnson, LMS, 2017, personal communication). The nose weight tolerance for TWT probes deployed during the PNE2013b was reduced to ±0.1 g (Fig. 2). The nominal wire weight of the probe's spool in the DB probes is 105 g, with a tolerance of ±1.5 g. In the TWT probes, this wire weight tolerance was reduced to ±1.0 g. No such constraint was applied with respect to the plastic body of the probe, because it is not only much lighter (approximately 51 g) but also neutrally buoyant. Therefore, the standard probes have a total manufacturer tolerance of ±2.5 g, and the TWT constraint reduces this weight tolerance by approximately half. A previous study examined the physical properties of the XBT probe (Reseghetti et al. 2007) and measured a higher total weight variability of ±5.0 g in standard probes; thus, the manufacturer's tolerance may underestimate the weight variability of the probe. Besides examining the different probes, this experiment allowed us to test three different approaches to correcting the data [see section 2b(3) for details].

The third experiment was carried out in December 2015 (PNE2015). During this experiment, a total of 44 standard DB probes (S/N 41–84) were deployed along CTD stations. This experiment was performed as

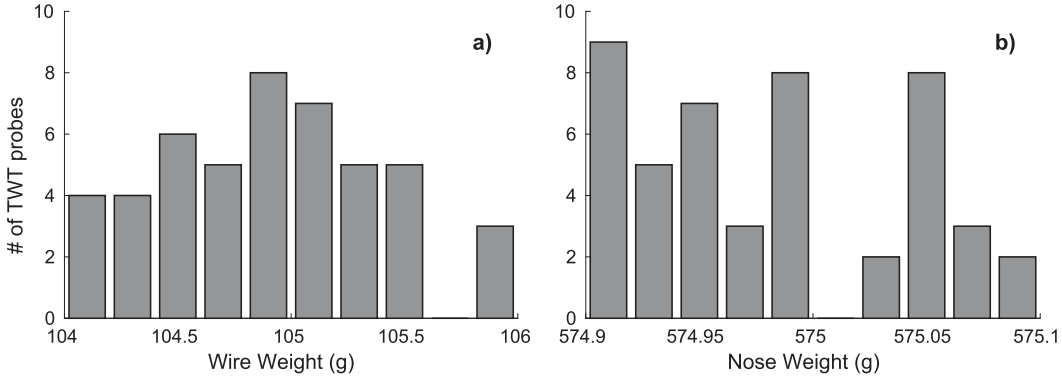

FIG. 2. Histogram showing the distribution of (a) wire and (b) nose weights (g) of the TWT probes deployed during the PNE2013b sea trial.

an additional test for the thermistor calibration (as described in section 2b). The details of the experiments are summarized in Table 1. All three cruises were conducted aboard R/V *Ron H. Brown*, using a manual XBT launcher, and an acquisition system that consisted of a Sippican MK21 readout card and a PC. The estimated mean deployment height above sea level was 4.4 m, which according to Bringas and Goni (2015) can produce a maximum depth offset of approximately 50 cm—much smaller than the standard deviation of the depth offset estimates presented here (section 3b).

b. Methods

1) CTD versus XBT comparison: The temperature gradient method

We quantify the errors in the XBT data using CTD data as the ground truth. The CTD model used in the sea trials is the Sea-Bird Scientific SBE 911, with a nominal temperature accuracy of 0.001°C and a nominal depth resolution of 0.015 m. The actual accuracy (and comparability) of the CTD profile varies depending on the sensor position in the rosette because of small-scale turbulence and the time difference between the XBT and CTD deployments, because of internal waves in the ocean. Strong differences were sometimes observed between the corresponding CTD upcasts and downcasts. In this study we use the downcast profile, because the sensor was located on the bottom of the rosette. The average depths of the CTD casts were around 1500 m. The XBTs were deployed either before or after a CTD cast, with time differences of less than 2 h since the CTD cast was initialized.

Systematic errors in XBT measurements can be approximated as

(i) a pure temperature bias, which is unrelated to depth biases and may be produced by the sensor or acquisition system, including a poorly calibrated thermistor, wire resistance imbalance, cables, or analog-to-digital (A/D) conversion (Roemmich and Cornuelle 1987);
(ii) a depth offset, which is linked to the initial orientation and fall speed of the XBT in the water, and also an offset in the time response of the thermistor or acquisition systems (e.g., Thresher 2014);
(iii) a linear depth bias, which is caused by inaccurate fall-rate coefficients (e.g., Flierl and Robinson 1977; Hanawa and Yoritaka 1987).

Following previous studies (e.g., Goes et al. 2013), we define the two depth errors—a depth offset Z_0 and a depth linear bias Z_d–such that

$$z_{\mathrm{XBT}} - z_{\mathrm{CTD}} = Z_0 + Z_d z_{\mathrm{CTD}} \pm \varepsilon_z, \tag{2}$$

and a pure temperature bias approximated by a temperature offset T_0 that is calculated after the depth errors are removed from the temperature profile:

$$T_{\mathrm{XBT}} - T_{\mathrm{CTD}} = T_0 \pm \varepsilon_T. \tag{3}$$

The residuals ε_z and ε_T are assumed to be randomly distributed with zero mean and uncorrelated, although the latter assumption is rarely met. No attempt was made to model the dependence of these biases [Eqs. (2) and (3)] on the local temperature (i.e., viscosity), so a potential correlation may still remain in the residuals within a profile.

The depth errors are calculated against the H95 FRE, using a temperature gradient method (e.g., Hanawa et al. 1994; H95). This method compares temperature gradients of the XBT and CTD profiles within a certain depth range (window), and locates the mean depth of the window that produces the best match. The criteria for the best XBT depth match is where 1) the RMSE is minimum, 2) the correlation is maximum, and 3) the mean temperature difference (DT) is less than a threshold of the 95th percentile of DT in the whole profile. Constraint 3 is done to restrict the best window locations to a nearby depth relative to the CTD cast. This optimization is performed twice, first with a moving depth window of 50 m and later with a window of 90 m and the DT threshold relaxed to DT + 1°C. Before applying this method, the data are interpolated to a depth step of 1 m, and filtered with a 7-point median and an 11-point Hanning window. XBT profiles that do not reach the depth of 600 m—or those that present too many spikes caused by wire insulation leaks (see Cook and Sy 2001)—are manually excluded from the analysis. The corrected depth is also filtered with an 11-point Hanning window. The depth error parameters are estimated using a least squares fit between 100 and 680 m, where the method performance is better and only the profiles that produce a good match (i.e., correlation) are included in the analysis.

2) The *T*–*R* equation

The temperature-resistance (*T*–*R*) equation for a thermistor is given by the modified form of the Steinhart and Hart (1968) equation:

$$T = \frac{1}{\left[A_0 + A_1 \log(R) + A_2 \log(R)^2 + A_3 \log(R)^3\right]} - 273.15, \tag{4}$$

where temperature T is given in degrees Celsius and the resistance (R) is given in ohms. The current values of the constant parameters used by Sippican are $A_0 = 1.290\,123 \times 10^{-3}$, $A_1 = 2.332\,252\,9 \times 10^{-4}$, $A_2 = 4.579\,129\,3 \times 10^{-7}$, and $A_3 = 7.162\,559\,3 \times 10^{-8}$.

TABLE 1. Summary of the analyzed experiments. This table is divided horizontally by probe type and vertically by cruise. For each cruise, the S/N, profile number (corresponding to Figs. 5, 6), and corrections applied are specified.

	Probe type		
	Standard	Experimental	TWT
Characteristic	DB probe	DB probe with screened thermistors (measured and bath temperatures differ less than 0.05°C)	DB probe with screened thermistors and TWT of 1.1g
Cruise		WBTS2012	
S/N	1182082–1182105	1182106–1182129	X
Profile cast number	1–15	81–99	X
Correction applied	Time constant	Thermistor calibration Time constant Wire imbalance	X
Cruise		PNE2013b	
S/N	1212792–1212815	1212720–1212743	1212744–1212791
Profile cast number	16–39	100–118	119–165
Correction applied	Thermistor calibration Time constant Wire imbalance	Thermistor calibration Time constant Wire imbalance	Thermistor calibration Time constant Wire imbalance
Cruise		PNE2015	
S/N	41–84	X	X
Profile cast number	40–80	X	X
Correction applied	Thermistor calibration Time constant	X	X

The temperature precision in typical XBT temperature recorders is truncated to one decimal digit because of the precision restrictions of the equipment. In our analysis, we use the full precision resulting from the T–R equation [Eq. (4)], since vertical gradients of temperature are better represented this way (Fig. 3b). This may potentially improve the comparison with the CTD using the gradient method.

3) CORRECTIONS APPLIED TO THE PROFILE DATA

In addition to the different probes used in the sea trials [see section 2b(1)], three postprocessing corrections to the XBT data are used when possible. These corrections are intended to counteract some of the biases that may be produced by the XBT system. The corrections are 1) wire imbalance correction, 2) static bath thermistor calibration, and 3) thermal time constant. They are detailed below.

(i) Correction 1: Wire imbalance

The thermistor in the XBT probe is physically connected by two wires (''A'' and ''B,'' respectively) to the probe's spool. The thermistor is located in the loop between wires A and B. A wire balance resistor is located inside the canister and is intended to cancel the differential resistance of the two leads (Fig. 4). However, there may be a residual unbalanced resistance that is dependent on the environmental temperature at launch, and it would cause an offset to the resistance/temperature profile. To correct for the residual wire imbalance, the wire resistances were measured for leads A and B at a given temperature. Before the probe deployment, the balance resistance was once again measured. The resistance of lead A was then subtracted from the resistance of lead B (B minus A) and from the measured resistance. This result was then added to the resistance profile measured by the probe. The new resistance profile is used as an input in the R–T equation [Eq. (4)] to calculate the new temperature values (Fig. 3).

(ii) Correction 2: Thermistor characterization (calibration)

The thermistor characterization is performed by measuring the thermistor resistance in a tightly controlled temperature bath (Georgi et al. 1980). The ratio between the measured and ideal resistance values at bath temperature $T_{\text{bath}} = 15$°C is used as a multiplying factor to correct the whole temperature profile. The thermistor R–T equation is then used to retrieve the ''calibrated'' temperature from the calibrated probe resistance data (Fig. 3).

(iii) Correction 3: Thermal time constant

The thermal time constant τ is the time required to detect 63% of a step thermal signal in a thermistor following an exponential decay. Its value ranges from 60 to 130 ms, and here we assume its maximum value

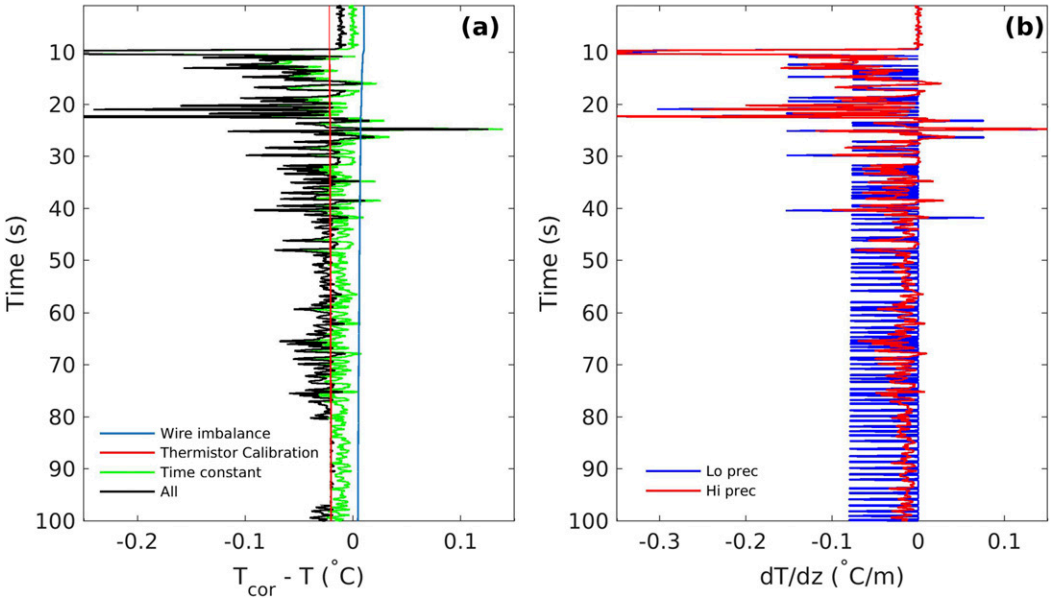

FIG. 3. (a) Difference between the corrected and original temperature profile against the time of descent for wire imbalance (blue), thermistor calibration (red), time constant (green), and all corrections (black). (b) Vertical temperature gradient against time for the same XBT profile. A T–R equation was applied in the corrections using the full resolution (red) instead of the truncated resolution common to XBT files (blue).

($\tau = 130$ ms). To correct for this effect, the temperature profile is shifted backward in time by τ seconds. The thermistor time constant correction can be mathematically represented by a bandpass filter $F = (0.008s + 1)/(0.13s + 1)$, with the low-pass filter defined as the Laplacian transform function $F_{\text{lp}} = 1/(\tau s + 1)$, where s is the operator variable, multiplied by a high-pass filter function $F_{\text{hp}} = (\tau_h s + 1)$. For the high-pass filter, the value $\tau_h = 0.008$ s was selected empirically to give stability to the filtering and minimal residual errors. To perform these calculations, we linearly interpolated the XBT data to produce between 25 and 50 times the initial number of data points.

4) STATISTICAL METHODS OF DATA INTERCOMPARISON

We calculate errors among different probe types using different correction methods. The errors in each XBT

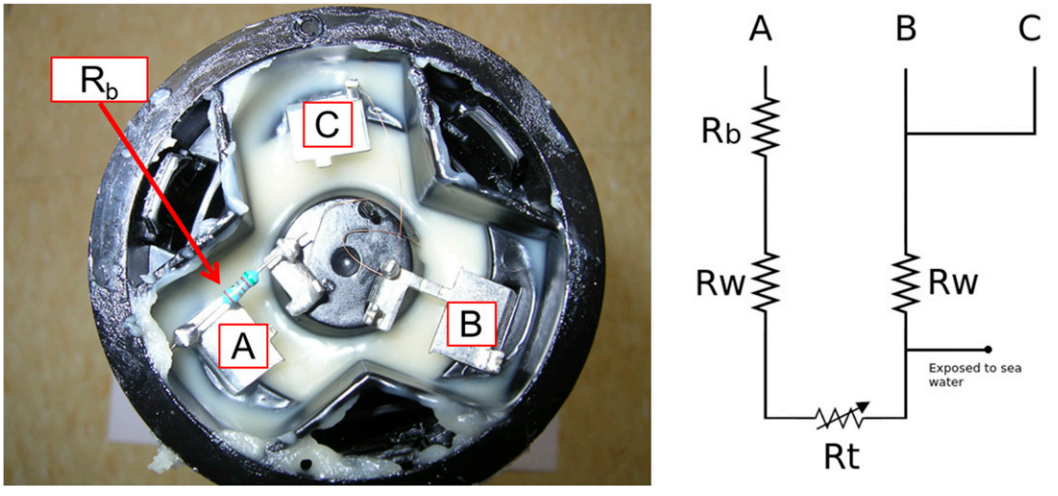

FIG. 4. (left) Location of the wire balance resistor (Rb) and the leads (A–C) on the top of the XBT canister. Balance resistor can be located in either one of the A and B leads. (right) Circuit diagram of an XBT, including the wire resistances (Rw), the thermistor (Rt), and Rb. (Drawing by Pedro Pena.)

profile relative to the CTD data are estimated using the temperature gradient method described in section 2b(1). The statistical error estimate and comparison between different probes and/or corrections were performed using an analysis of variance (ANOVA) method (Gelman 2005). The ANOVA method is a multilinear model fit used to compare the means of different numerical populations and to determine the relative importance of different sources of variation in a dataset. In our analysis we decomposed the mean of each XBT error parameter population ($Y[i]$, $i = 1{:}n$ samples) between two sets of predictors,

$$Y[i] = \mu + \alpha_j[i] + \beta_k[i] + \alpha\beta_{jk}[i] + \varepsilon_i, \qquad (5)$$

which are the probe type (α_j; $j = 0{:}1$ probes), where 0 is the standard probe and 1 is the modified probe; and correction (β_k; $k = 0{:}4$ corrections), where 0 is used for the uncorrected error values. The coefficient μ is the base case, which is used for the standard probe with no correction applied. Thus, the parameters α_j and β_k are set to zero in the base case ($\alpha_0 = \beta_0 = 0$), α_0 is the estimated difference between the modified probe and the standard probe, and $\beta_1{:}\beta_4$ are the estimated differences between the four types of corrected values and the uncorrected values. We also accounted for interaction terms $\alpha\beta_{jk}$, whereby a certain correction might produce different outcomes for different probes. The residuals ε_i are assumed to be normally distributed [$N(0,\sigma_j^2)$], with zero mean and variance σ_j^2 dependent on the probe type. The statistics of the coefficients are estimated in a Bayesian framework. Uninformative normal priors are used for the ANOVA coefficients, and uninformative Gamma priors are used for the variances of each probe type. The calculations are performed using a Monte Carlo (MCMC) method with the Windows Bayesian Inference Using Gibbs Sampling (WinBUGS) software (Lunn et al. 2000), using two Markov chains of 20 000 iterations (and a burn-in of 1000 samples).

The differences in the mean between the modified probes and corrections relative to the standard probe are significant if the magnitude of the coefficients α_j and β_k differ from zero given their respective standard errors. For the variances σ_j^2, the significance of the differences between the standard and modified probes is given in terms of probability. The probability that the error in the modified probe population is lower than the error in standard probe population [$P(\sigma_j^2 < \sigma_0^2)$] is modeled within the MCMC, by calculating the difference in variance between the errors of the two probe populations $\Delta v = \text{step}(\sigma_1^2 - \sigma_0^2)$ using a step function, which assumes the value $\Delta v = 1$ if $(\sigma_1^2 < \sigma_0^2) \geq 0$, and $\Delta v = 0$ if $(\sigma_1^2 < \sigma_0^2) < 0$. The percentage difference in the number of cases in which the step function assumes values 1 or 0 gives the relative improvement between the two populations. The same ANOVA approach was used to evaluate errors in depth ($Y[i] = z$[probe, correction] $- z_{\text{CTD}}$) and errors in temperature ($Y[i] = T$[probe, correction] $- T_{\text{CTD}}$).

The experimental and TWT probes were grouped together as "experimental" for the temperature analysis, and the standard and experimental probes were grouped together as "standard" for the depth analysis.

3. Results

We compare the side-by-side CTD and XBT data and examine how the temperature and depth biases in the XBT are sensitive to corrections and probe improvements. The data from the three cruises are analyzed together to improve statistical robustness and to assess, via an ANOVA, the relative significance of improvements resulting from probe type and correction method. In the ANOVA method, we use the probes and corrections as the factors [see section 2b(4) for details].

a. Temperature improvements

The sensitivity of the thermal accuracy was analyzed by comparing the biases between the standard and experimental (including TWT) probes. The thermal bias was calculated after subtracting from each profile the depth biases estimated using the gradient method. The temperature offset estimates are positive for practically all probes deployed without any corrections applied (gray bars in Fig. 5). This warm bias has been previously reported in the historical record (e.g., Gouretski and Koltermann 2007; Kizu and Hanawa 2002; Reverdin et al. 2009; Cowley et al. 2013) and is partially due to uncalibrated thermistors (Szabados and Wright 1989).

1) PROBE TYPE

A clear distinction can be seen between the standard probes (casts 1–80 in Fig. 5a) and the probes that feature screened thermistors (i.e., experimental and TWT). In general, the magnitude of the thermal biases is larger for the standard probes than for the experimental probes and is sometimes above the manufacturer's stated tolerance of 0.1°C. The results from the ANOVA for the temperature offset (Fig. 6a) indicate that the mean bias and standard error for the uncorrected standard probes is $T_0 = 0.073° \pm 0.004$°C, whereas for the experimental and TWT probes it is reduced to $T_0 \sim 0.035° \pm 0.004$°C. Averaged over all the corrections listed in Fig. 6a, the overall bias for the

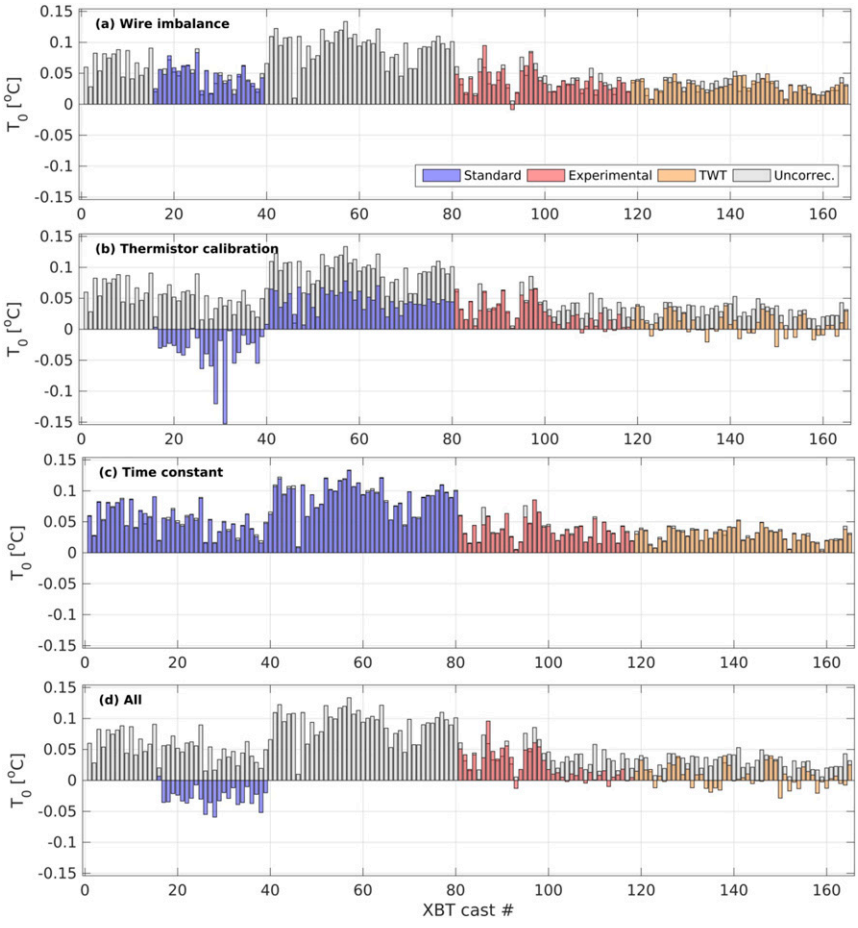

FIG. 5. Vertical bar plot of T_0 for all probes analyzed. Each panel compares T_0 before any correction (gray bars) against T_0 calculated after each correction (colored bars), for (a) wire imbalance, (b) thermistor calibration, (c) thermal time constant, and (d) all corrections together. Each colored bar is for a different probe type, and the x axis represents the individual deployments (in order of deployment), which are clustered by probe type.

experimental probes is $T_0 = 0.022° \pm 0.001°C$ (Fig. 6b), which is just slightly less than the average for the standard probes ($T_0 = 0.030° \pm 0.002°C$). This is due to the rebound of errors after the thermistor correction in some of the standard probes, as we shall see next.

2) CORRECTIONS

Each panel in Fig. 5 shows the sensitivity of the thermal bias to the different corrections. Although the overall effect of the wire imbalance correction is to reduce the thermal bias slightly, neither the wire imbalance nor the thermal time constant corrections are very efficient in reducing T_0 (Figs. 5a,c). Figure 6a reinforces this result for the experimental probes. For the standard probes, using these two corrections results in an apparent reduction of T_0 relative to the uncorrected estimates, but this is mostly driven by the larger population of uncorrected probes. Conversely, the thermistor calibration can change the thermal biases significantly, and it is the dominant factor when all corrections are applied, as shown by the close resemblance of the two results (Figs. 5b,d). For the experimental and TWT probes, the thermistor calibration was able to reduce the mean thermal bias very efficiently, from an initial $0.035 \pm 0.004°$ to $0.009 \pm 0.002°C$ (Fig. 6a). The thermistor calibration also reduced the mean temperature bias of standard probes considerably, from the initial $0.073° \pm 0.004°$ to $0.02° \pm 0.003°C$. Interestingly, the standard probes still exhibit large T_0 variability after the thermistor calibration and when all corrections are applied (Fig. 6a). Indeed, the calculated standard deviation of T_0 is $\sigma = 0.034°C$ for the standard probes and it is reduced to $\sigma = 0.014°C$ for the experimental and TWT probes (Fig. 7a; Table 2), which accounts for a 100% likelihood of variance reduction toward standard probes.

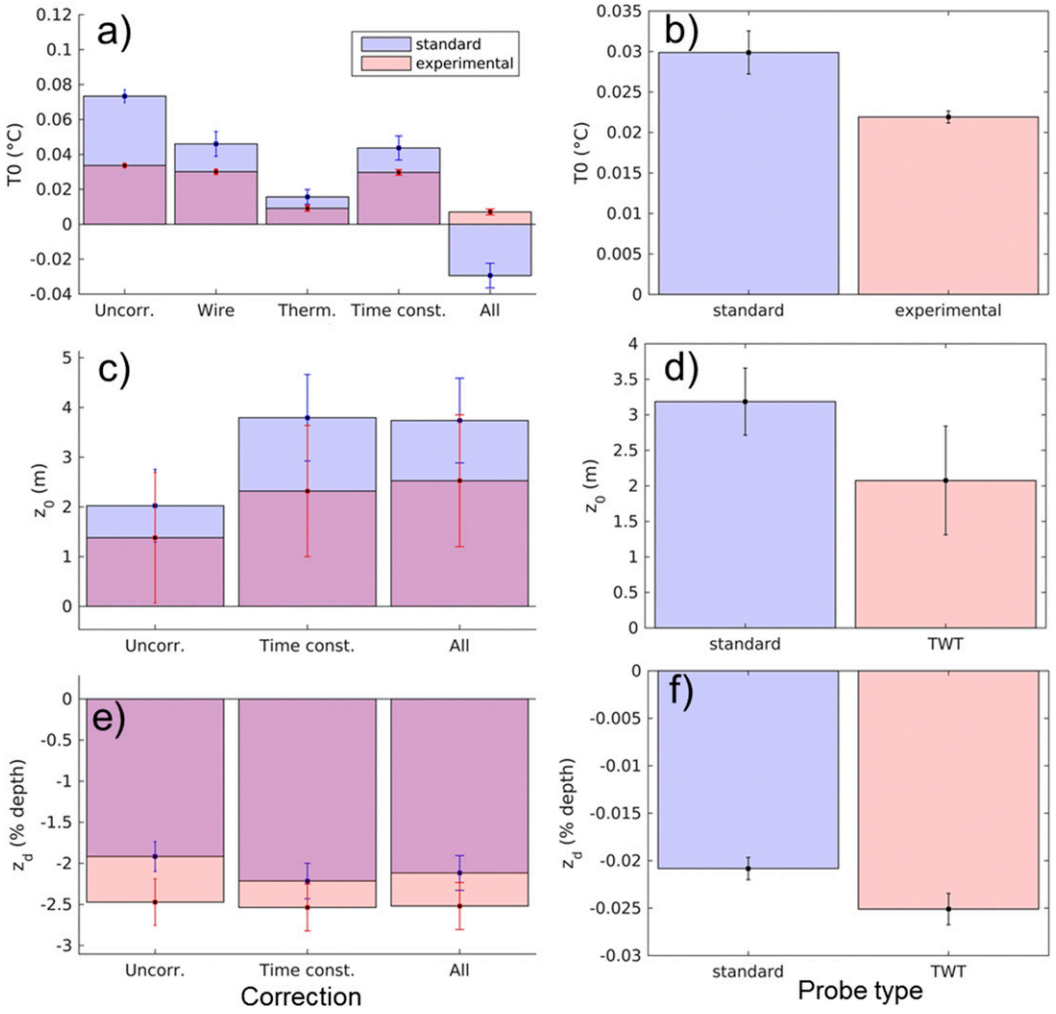

FIG. 6. Distributions of the mean XBT measurement biases (solid bars) and their standard errors (error bars) from the ANOVA analysis for the data of all three cruises together. Marginal distributions of the parameters relative to (left) the correction applied and (right) the probe types.

This strong variability is revealed in the standard probe casts 16 to 39 from the PNE2013b cruise (Fig. 5b). In those casts T_0 had large positive values for the uncorrected thermistors and reversed its sign after correction, becoming strongly negative. Since these are the only casts to which we were able to apply all the corrections together, this inversion is reflected in Fig. 6a. One possible explanation for this variability is that the thermistor calibration considered only the temperature value at $T_{\text{bath}} = 15°\text{C}$, yet the standard thermistors used in the probes during the PNE2013b cruise present different biases at different temperatures. We note that the temperature residuals (T minus T_{bath}) for both experimental and TWT probes from the PNE2013b cruise (Fig. 8), as well as for the probes that carry this information in the WBTS12 and PNE2015 cruises, have negligible differences at all depths. For the standard probes used in the PNE2013b cruise, however, the residual temperature differences are dependent on the bath temperature at which they were taken, within a range of 0.05°C measured at $T_{\text{bath}} = 0°\text{C}$, and within ~0.1°C at $T_{\text{bath}} = 15°$ and 35°C. Similar behavior was also observed for standard probes analyzed by Reseghetti et al. (2007).

b. *Depth improvements*

1) PROBE TYPE

The depth linear bias (Fig. 9a) is mostly negative for all probes analyzed (Fig. 9a), meaning that the probe's descent is generally slower than predicted by the H95 coefficients. The depth offset (Fig. 9b) is more randomly distributed, although these two parameters are highly negatively correlated ($R = -0.55$) (Figs. 9a,b,

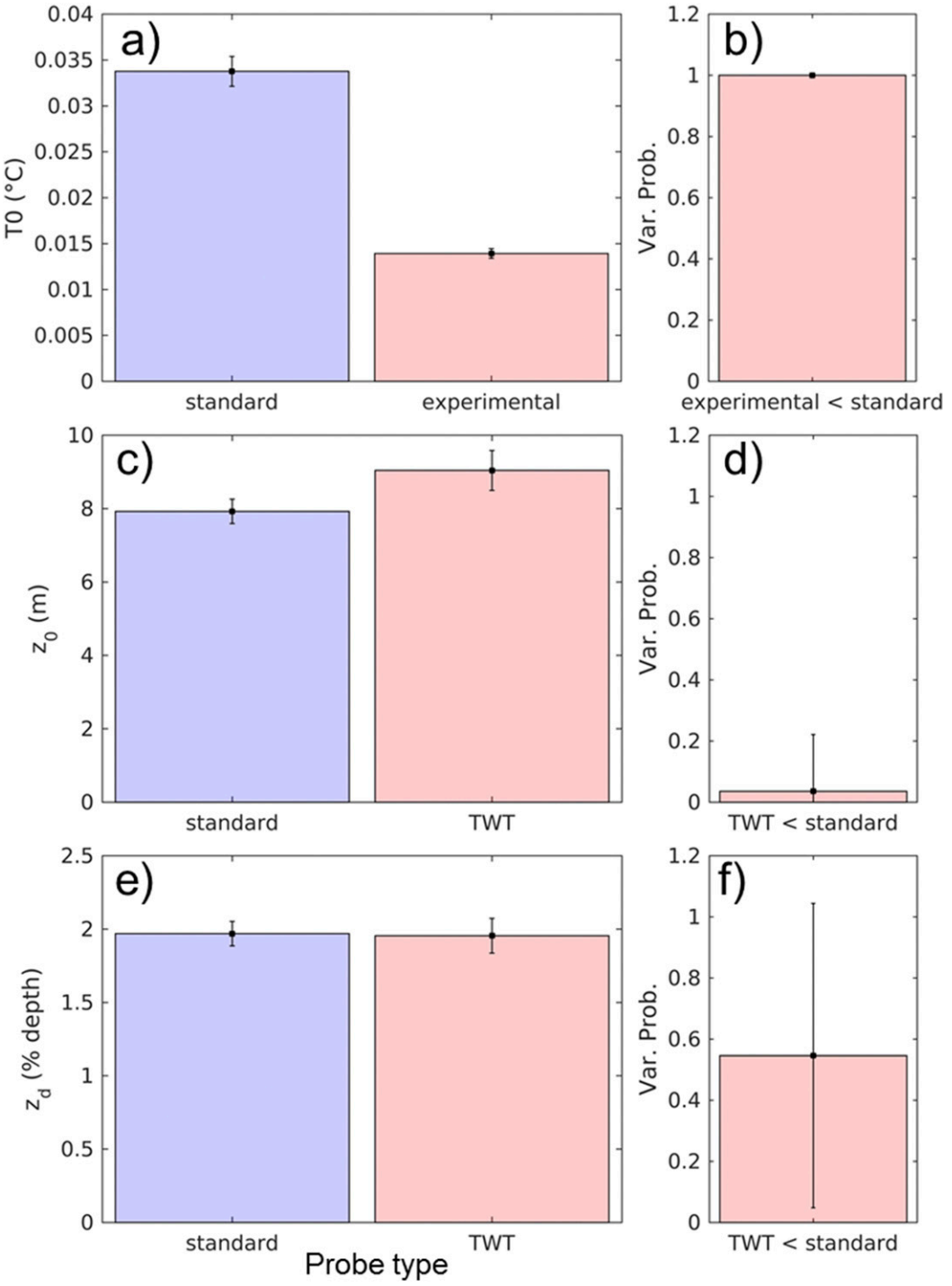

FIG. 7. (left) Standard deviation estimated for each probe type for (a) temperature offset, (c) depth offset, and (e) linear depth bias. (right) Probability for the hypothesis testing that the variance of the modified probe is smaller than the standard probe.

10a), which suggests there may be a common cause affecting the variability of both parameters. The overall depth offset (Fig. 6d) is reduced for the TWT probes (Z_0 = 2.0 ± 0.8 m) relative to the standard (and experimental) probes (Z_0 = 3.1 m ± 0.5 m). This difference could be explained by the differences in weight tolerance of the probes because the initial velocity of the probe as it touches the water is related to the probe's

TABLE 2. Statistical parameters estimated using the ANOVA methodology for the biases associated with the XBT measurements using the information of the three cruises analyzed. Values of T_0 are multiplied by 10. Parameters that are statistically significant are highlighted in bold. Modified probes are TWT for Z_0 and Z_d, and experimental for T_0.

			T_0		Z_0		Z_d	
Statistical parameter	Probe	Correction	Mean	STE	Mean	STE	Mean	STE
μ	Standard	—	***0.73***	0.04	***2.02***	0.73	***−1.92***	0.18
α [1]	Modified	—	***−0.40***	0.04	−0.64	1.51	***−0.56***	0.34
β [1]	—	Wire imbalance	***−0.27***	0.08	—	—	—	—
β [2]	—	Thermistor calibration	***−0.58***	0.06	—	—	—	—
β [3]	—	Time constant	***−0.30***	0.08	***1.77***	1.14	***−0.30***	0.28
β [4]	—	All	***−1.03***	0.08	***1.71***	1.13	−0.20	0.28
$\alpha\beta$ [1,1]	Modified	Wire imbalance	***0.24***	0.08	—	—	—	—
$\alpha\beta$ [1,2]	Modified	Thermistor calibration	***0.33***	0.06	—	—	—	—
$\alpha\beta$ [1,3]	Modified	Time constant	***0.26***	0.08	−0.83	2.19	0.23	0.49
$\alpha\beta$ [1,4]	Modified	All	***0.76***	0.08	−0.57	2.18	0.15	0.49
σ [0]	Standard	—	***0.34***	0.02	***7.93***	0.33	***1.97***	0.08
σ [1]	Modified	—	***0.14***	0.01	***9.04***	0.54	***1.96***	0.12

mass (Hallock and Teague 1992). However, there is a small likelihood that the TWT probes can reduce the variance of the depth offset (Fig. 7d) and therefore it may be more sensitive to the launch details, such as time difference between CTD and XBT casts (Fig. 10b) or environmental conditions.

The TWT probes did not produce significant changes in linear depth bias as would be expected. The linear depth bias mean (Fig. 6c) is similar for both standard ($Z_d = 2.1 \pm 0.1\%$ of depth) and TWT probes ($Z_d = 2.5 \pm 0.2\%$ of depth), which are within the range of previous estimates (e.g., Wijffels et al. 2008) and in agreement with the manufacturer tolerance of ±5 m and 2% of depth. The standard deviation estimated for the TWT and standard probes are similar, both approximately $\sigma = 2\%$ of depth (Fig. 7e; Table 2)—consistent with the 50% likelihood that their variances are different (Fig. 7f). Therefore, the reduced mass tolerance was not capable of constraining the spread of the linear depth bias.

2) CORRECTIONS

Of all the corrections applied to the XBT data, we consider herein only the results for the thermal time

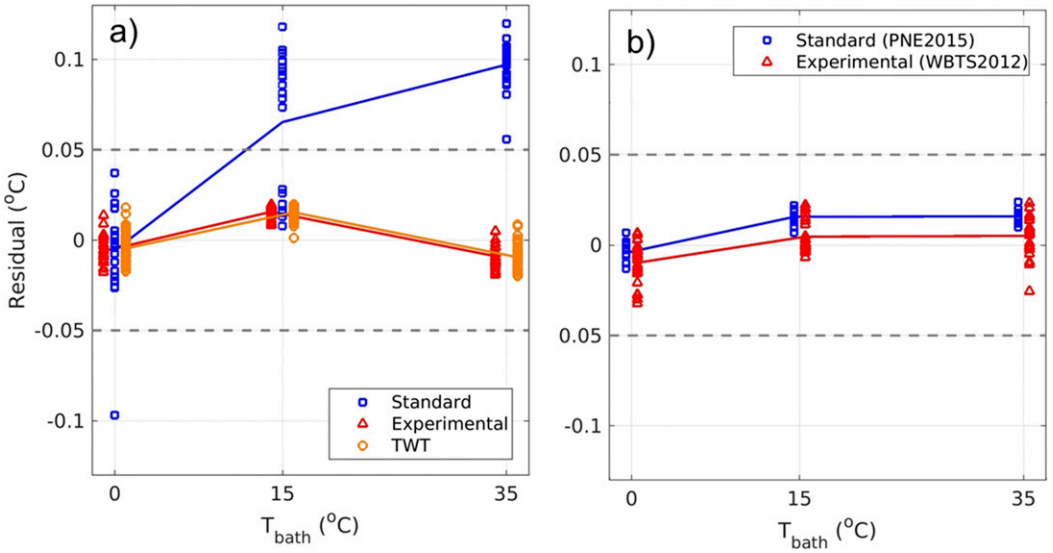

FIG. 8. Difference between measured temperature and bath temperature ($T - T_{bath}$) relative to T_{bath} (°C) in (a) PNE2013b cruise and (b) PNE2015 and WBTS2012 cruises. Bath temperatures are taken at 0°, 15°, and 35°C. Color/shapes refer to standard probes (blue squares), experimental probes (red triangles), and TWT probes (orange circles). Only the values at $T_{bath} = 15$°C are used in the thermistor calibration, and the threshold of 0.05°C used in the screened thermistors is highlighted (dashed gray lines).

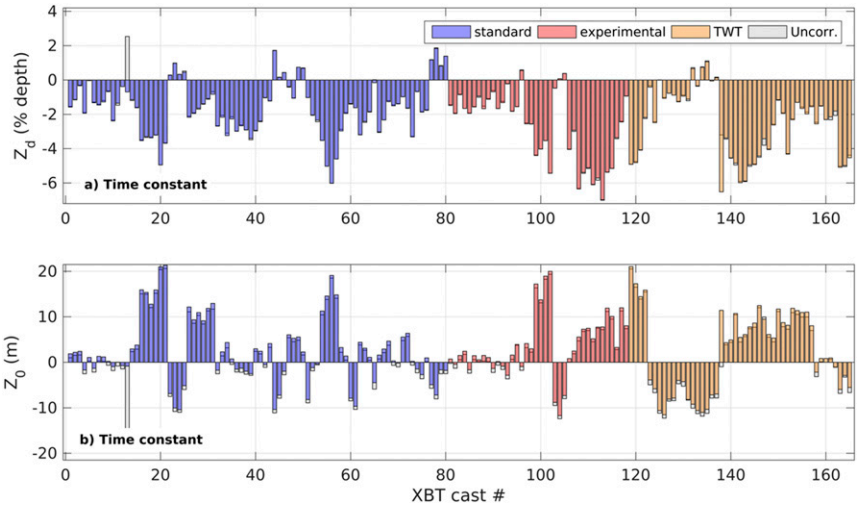

FIG. 9. (a) Depth linear bias (% of depth), and (b) depth offset (m) estimated for all probes analyzed. Gray bars are for the original values, and the different color bars are for the error estimates after the thermal time constant correction, with each color representing a different probe type. Individual deployments (in order of deployment), which are clustered by probe type, are represented on the *x* axis.

constant correction, which applies an upward shift in the temperature profile and would most likely influence the depth biases. The effect of this correction on the estimated depth offset (Fig. 9b) is a shift of approximately +1 m, as shown in Fig. 6c. The time constant correction tends to increase Z_0 in the analyzed population, since Z_0 estimated before the correction is positive on average (Fig. 6c). As Z_0 values range from positive to negative (Fig. 9b), this correction cannot solve the depth offset issue. Indeed, the depth offset is also a function of other factors, such as the initial velocity of the probe in the water, which depends on either the deployment height, the orientation of the probe as it touches the water, or the difference in time between the XBT and CTD deployment (e.g., Boyer et al. 2011; Bringas and Goni 2015). Although there is a potential relationship between the linear depth bias and the depth offset estimates, we conclude that the depth linear bias is not significantly affected by the thermal time constant correction (Fig. 6e).

4. Discussion and conclusions

In this study we investigated some potential probe enhancements to improve the accuracy of XBT data, one of several ongoing efforts to produce climate-quality

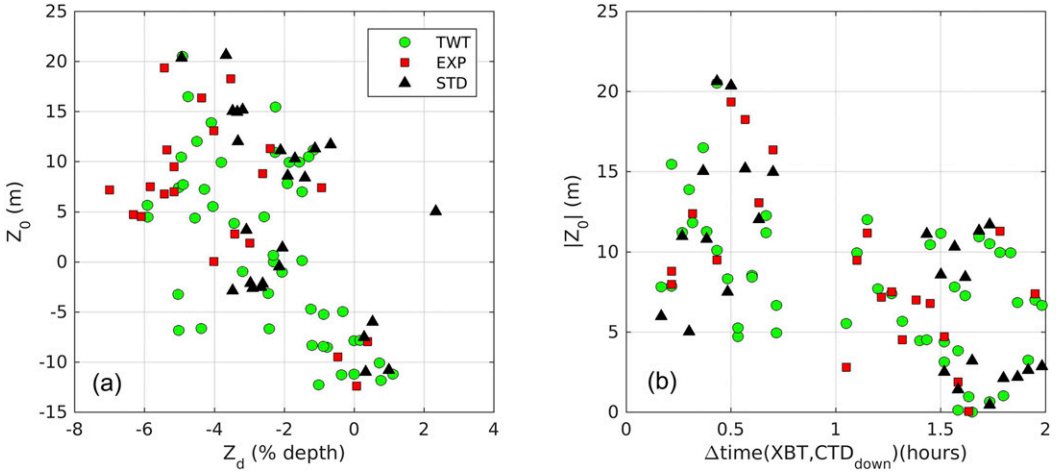

FIG. 10. (a) Relationship of Z_0 with Z_d, and (b) Z_0 magnitude with the time difference between the XBT and CTD casts for deployments of the PNE2013b cruise (see Fig. 1). Colors/shapes represent different probe types.

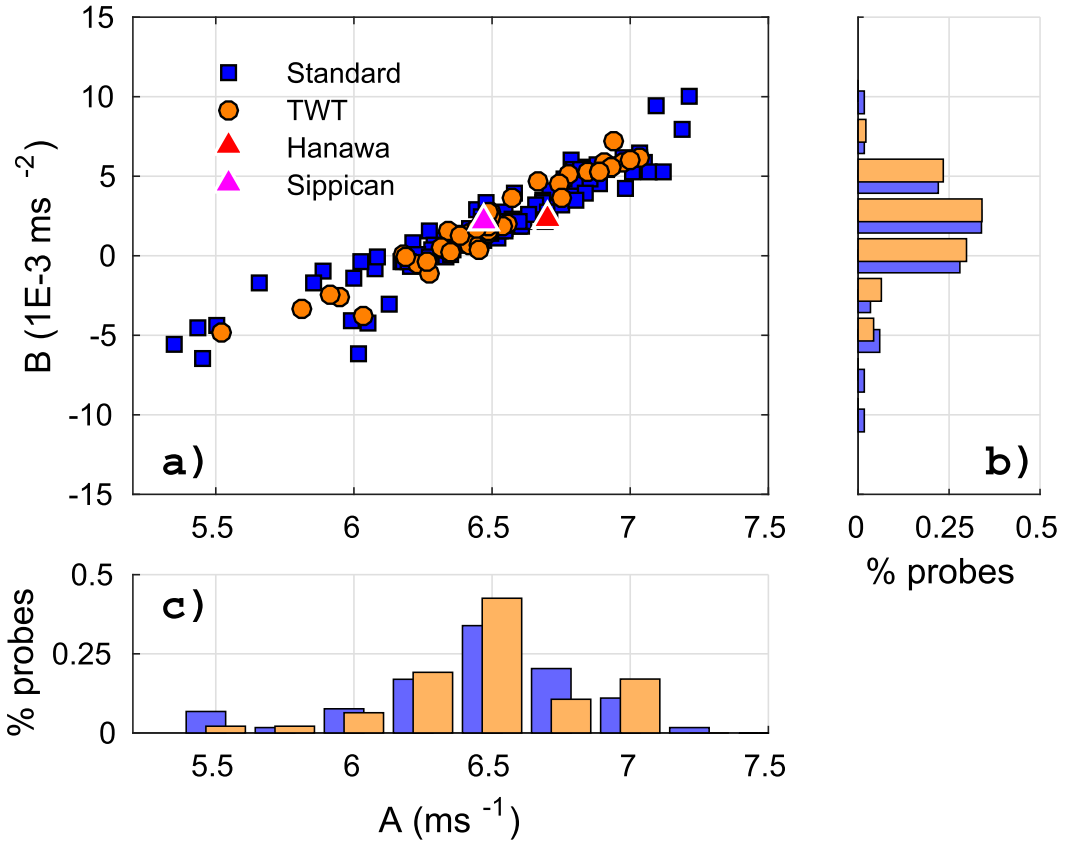

FIG. 11. (a) FRE coefficients A (m s^{-1}) and B (1E-3 m s^{-2}) estimated for the standard (and experimental, blue) and TWT (orange) probes deployed during the three cruises analyzed. H95 values used in the present study (red square), and Sippican's values as a comparison (magenta triangle). (b),(c) Respective parameter histograms normalized by the number of probes for the standard (orange) and TWT (blue) probes.

XBT probes. The objective of these experiments was to test how a stricter quality control during the production of the probes and postprocessing of the data may reduce the XBT thermal and depth biases. Three XBT probe types were tested: 1) standard, 2) experimental, and 3) tight weight tolerance (TWT) probes. In addition, three corrections were applied as part of the data postprocessing for the thermistor calibration: 1) wire imbalance, 2) manufacturer's thermistor calibration, and 3) thermal time constant corrections.

Our results show that the thermistor "calibration" has the strongest effect on correcting temperature biases. After its application, the mean warm bias is significantly reduced to 0.009°C for experimental probes and 0.02°C for standard probes, and the tolerance for the temperature offset is reduced to $|T_0| < 0.03$°C. The thermistor calibration overcorrected T_0 for the standard probe during one of the cruises. This was due to the strong temperature dependence of the thermistor accuracy in those probes, and there is a possibility that they were actually discarded from the thermistor screening process in that cruise, in which case a linear calibration with temperature may be necessary (Reseghetti et al. 2007). The thermal time constant correction did not produce significant changes in temperature or depth biases. Reseghetti et al. (2007) showed that the acquisition system may need ~0.6 s (4 m) before a probe detects a step signal (e.g., Kizu and Hanawa 2002). Increasing the sampling frequency in the recorder could reduce the detection time of a step signal. Our results show that the wire imbalance correction did not produce significant changes in temperature bias. Indeed, resistance residuals as a result of imbalanced wire resistance constitute <1% of the resistance reading in the profile.

With respect to the probe types, probes with screened thermistors (experimental and TWT) showed a smaller overall thermal bias (T_0 = 0.035°C) relative to the standard probes (T_0 = 0.073°C) and a more robust thermal bias reduction using a one-point thermistor calibration than the standard probes. TWT probes did not show considerable reduction in the mean or variance of depth biases, even though they showed smaller depth

offsets. Figure 11 shows the distributions of the coefficients A and B of the FRE [Eq. (1)] estimated for all profiles analyzed as a function of probe type. The mean (and standard deviations) of the FRE parameters A and B are 6.45 (0.44) ms^{-1} and 1.6 (3.5) $\times 10^{-3}$ for the standard probes, respectively; and 6.48 (0.32) ms^{-1} and 1.9 (2.8) $\times 10^{-3}$ ms^{-2} for the TWT probes. These values are not statistically different given the Student's t test. According to the values calculated theoretically by Seaver and Kuleshov (1982), the TWT probes used here (with a $\pm$1.1-g weight tolerance) would have a maximum depth error of 1.5 m, as opposed to an approximated 3-m error from the standard probes. This reduction was not detected in our experiments, which could be a caveat related to the precision of the temperature gradient method applied to estimate these errors. Other effects may be driving this variability, such as the nose roughness (which increases the drag in the water), air entrapped within the wire (which changed the buoyancy of the probe), among other factors such the speed and orientation of the probe as it hits the water, or the shape of the tail fin (e.g., Kizu et al. 2011; Abraham et al. 2014). Mostly likely, as our results suggest, it is the time difference between the CTD and XBT launches that is driving the variability of the depth bias (Fig. 10).

Our results suggest that further experiments should be performed focusing on the depth estimate improvement, in which both standard and TWT probe weights should be measured, and the deployment height and synchronization with the CTD casts should be tightly controlled. Additional measures may be necessary to further correct XBT depth biases. For instance, Goes et al. (2013) has shown that the inclusion of two pressure switches is an efficient way to correct depth estimates, although this results in higher probe costs. Additionally, FRE parameterization (Cheng et al. 2014; Bringas and Goni 2015) could be improved by including one extra term dependent on the deployment height to improve depth accuracy.

This study proposed and analyzed different corrections for the biases that affect XBT measurements. A potential application of our findings is that thermistor calibration can effectively reduce the pure temperature biases in future XBT records and therefore improve the accuracy of the ocean parameters measured by XBTs, especially for ocean heat content estimates. In addition, we presented a statistical platform that can be used in future studies of probe comparisons and uncertainty estimation.

Acknowledgments. The authors thank the engineers from Lockheed Martin/Sippican for their help with planning the experiments and for the interesting discussions. We also thank NOAA/AOML for the support during this work, and the PNE and WBTS scientists and crew for supporting the experiments. Goes, Bringas, and Goni were funded by the NOAA's Climate Program Office and NOAA/AOML. Goes, Ortner, and Babcock were partly funded by the University of Miami and the Cooperative Institute for Marine and Atmospheric Studies (CIMAS) under Cooperative Agreement NA17RJ1226.

REFERENCES

Abraham, J. P., J. M. Gorman, F. Reseghetti, E. M. Sparrow, J. R. Stark, and T. G. Shepard, 2014: Modeling and numerical simulation of the forces acting on a sphere during early-water entry. *Ocean Eng.*, **76**, 1–9, doi:10.1016/j.oceaneng.2013.11.015.

Anderson, E. R., 1980: Expendable bathythermograph (XBT) accuracy studies. Naval Ocean Systems Center Tech. Rep. TR 550, 202 pp.

Bindoff, N. L., and Coauthors, 2007: Observations: Oceanic climate change and sea level. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 385–432.

Boyer, T. P., and Coauthors, 2011: Investigation of XBT and XCTD biases in the Arabian Sea and the Bay of Bengal with implications for climate studies. *J. Atmos. Oceanic Technol.*, **28**, 266–286, doi:10.1175/2010JTECHO784.1.

Bringas, F., and G. Goni, 2015: Early dynamics of Deep Blue XBT probes. *J. Atmos. Oceanic Technol.*, **32**, 2253–2263, doi:10.1175/JTECH-D-15-0048.1.

Cheng, L., J. Zhu, R. Cowley, T. Boyer, and S. Wijffels, 2014: Time, probe type, and temperature variable bias corrections to historical expendable bathythermograph observations. *J. Atmos. Oceanic Technol.*, **31**, 1793–1825, doi:10.1175/JTECH-D-13-00197.1.

——, and Coauthors, 2016: XBT science: Assessment of instrumental biases and errors. *Bull. Amer. Meteor. Soc.*, **97**, 924–933, doi:10.1175/BAMS-D-15-00031.1.

Cook, S., and A. Sy, 2001: Best guide and principles manual for the Ships of Opportunity Program (SOOP) and expendable bathythermograph (XBT) operations. JCOMM Rep., 25 pp.

Cowley, R., S. Wijffels, L. Cheng, T. Boyer, and S. Kizu, 2013: Biases in expendable bathythermograph data: A new view based on historical side-by-side comparisons. *J. Atmos. Oceanic Technol.*, **30**, 1195–1225, doi:10.1175/JTECH-D-12-00127.1.

DiNezio, P., and G. Goni, 2010: Identifying and estimating biases between XBT and Argo observations using satellite altimetry. *J. Atmos. Oceanic Technol.*, **27**, 226–240, doi:10.1175/2009JTECHO711.1.

——, and G. J. Goni, 2011: Direct evidence of a changing fall-rate bias in XBTs manufactured during 1986–2008. *J. Atmos. Oceanic Technol.*, **28**, 1569–1578, doi:10.1175/JTECH-D-11-00017.1.

Flierl, G., and A. Robinson, 1977: XBT measurements of thermal gradients in the MODE eddy. *J. Phys. Oceanogr.*, **7**, 300–302, doi:10.1175/1520-0485(1977)007<0300:XMOTGI>2.0.CO;2.

Gelman, A., 2005: Analysis of variance—Why it is more important than ever. *Ann. Stat.*, **33**, 1–53, doi:10.1214/009053604000001048.

Georgi, D., J. Dean, and J. Chase, 1980: Temperature calibration of expendable bathythermographs. *Ocean Eng.*, **7**, 491–499, doi:10.1016/0029-8018(80)90048-7.

Goes, M., G. Goni, and K. Keller, 2013: Reducing biases in XBT measurements by including discrete information from pressure switches. *J. Atmos. Oceanic Technol.*, **30**, 810–824, doi:10.1175/JTECH-D-12-00126.1.

Good, S. A., 2011: Depth biases in XBT data diagnosed using bathymetry data. *J. Atmos. Oceanic Technol.*, **28**, 287–300, doi:10.1175/2010JTECHO773.1.

Gorman, J. M., J. P. Abraham, D. B. Schwalbach, T. S. Shepard, J. R. Stark, and F. Reseghetti, 2014: Experimental verification of drag forces on spherical objects entering water. *J. Mar. Biol. Oceanogr.*, **3** (2), doi:10.4172/2324-8661.1000126.

Gouretski, V., 2012: Using GEBCO digital bathymetry to infer depth biases in the XBT data. *Deep-Sea Res. I*, **62**, 40–52, doi:10.1016/j.dsr.2011.12.012.

——, and K. P. Koltermann, 2007: How much is the ocean really warming? *Geophys. Res. Lett.*, **34**, L01610, doi:10.1029/2006GL027834.

——, and F. Reseghetti, 2010: On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database. *Deep-Sea Res. I*, **57**, 812–834, doi:10.1016/j.dsr.2010.03.011.

Hallock, Z. R., and W. J. Teague, 1992: The fall-rate of the T7 XBT. *J. Atmos. Oceanic Technol.*, **9**, 470–483, doi:10.1175/1520-0426(1992)009<0470:TFROTT>2.0.CO;2.

Hamon, M., G. Reverdin, and P. Y. Le Traon, 2012: Empirical correction of XBT data. *J. Atmos. Oceanic Technol.*, **29**, 960–973, doi:10.1175/JTECH-D-11-00129.1.

Hanawa, K., and H. Yoritaka, 1987: Detection of systematic errors in XBT data and their correction. *J. Oceanogr. Soc. Japan*, **43**, 68–76, doi:10.1007/BF02110635.

——, P. Rual, R. Bailey, A. Sy, and M. Szabados, 1994: Calculation of new depth equations for expendable bathythermographs using a new temperature-error-free method (application to Sippican/TSK T-7, T-6 and T-4 XBTs). UNESCO Technical Papers in Marine Science 67, IOC Technical Series 42, 46 pp., http://unesdoc.unesco.org/images/0010/001035/103567e.pdf.

——, ——, ——, ——, and ——, 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, doi:10.1016/0967-0637(95)97154-Z.

Hazelworth, J. B., 1966: Quantitative analysis of some bathythermograph errors. ASWEPS Rep. 11, U.S. Naval Oceanographic Office Tech. Rep. TR-180, 27 pp.

Heinmiller, R. H., C. C. Ebbesmeyer, B. A. Taft, D. B. Olson, and O. P. Nikitin, 1983: Systematic errors in expendable bathythermograph (XBT) profiles. *Deep-Sea Res.*, **30A**, 1185–1197, doi:10.1016/0198-0149(83)90096-1.

Kizu, S., and K. Hanawa, 2002: Start-up transient of XBT measurement. *Deep-Sea Res. I*, **49**, 935–940, doi:10.1016/S0967-0637(02)00003-1.

——, S. I. Ito, and T. Watanabe, 2005: Inter-manufacturer difference and temperature dependency of the fall-rate of T-5 expendable bathythermograph. *J. Oceanogr.*, **61**, 905–912, doi:10.1007/s10872-006-0008-z.

——, C. Sukigara, and K. Hanawa, 2011: Comparison of the fall rate and structure of recent T-7 XBT manufactured by Sippican and TSK. *Ocean Sci.*, **7**, 231–244, doi:10.5194/os-7-231-2011.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter, 2000: WinBUGS—A Bayesian modelling framework: Concepts, structure, and extensibility. *Stat. Comput.*, **10**, 325–337, doi:10.1023/A:1008929526011.

NOAA, 2002: Code table 1770. [Available online at https://www.nodc.noaa.gov/woce/woce_v3/wocedata_1/woce-uot/document/wmocode.htm.]

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, doi:10.5194/os-3-59-2007.

Reverdin, G., F. Marin, B. Bourles, and P. L'Herminier, 2009: XBT temperature errors during French research cruises (1999–2007). *J. Atmos. Oceanic Technol.*, **26**, 2462–2473, doi:10.1175/2009JTECHO655.1.

Roemmich, D., and B. Cornuelle, 1987: Digitization and calibration of the expendable bathythermograph. *Deep-Sea Res.*, **34A**, 299–307, doi:10.1016/0198-0149(87)90088-4.

Seaver, G. A., and S. Kuleshov, 1982: Experimental and analytical error of expendable bathythermograph. *J. Phys. Oceanogr.*, **12**, 592–600, doi:10.1175/1520-0485(1982)012<0592:EAAEOT>2.0.CO;2.

Shepard, T., J. Abraham, D. Schwalbach, S. Kane, D. Siglin, and T. Harrington, 2014: Velocity and density effect on impact force during water entry of sphere. *J. Geophys. Remote Sens.*, **3**, 129, doi:10.4172/2169-0049.1000129.

Steinhart, J. C., and S. R. Hart, 1968: Calibration curves for thermistors. *Deep-Sea Res. Oceanogr. Abstr.*, **15**, 497–503, doi:10.1016/0011-7471(68)90057-0.

Szabados, M., and D. Wright, 1989: Field evaluation of the real-time XBT systems. *Proceedings of the Western Pacific International Meeting and Workshop on TOGA COARE*, J. Picaut, R. Lucas, and T. Delcroix, Eds., ORSTOM, 811–821.

Thadathil, P., A. K. Saran, V. V. Gopalakrishna, P. Vethamony, N. Araligidad, and R. Bailey, 2002: XBT fall rate in waters of extreme temperature: A case study in the Antarctic Ocean. *J. Atmos. Oceanic Technol.*, **19**, 391–396, doi:10.1175/1520-0426-19.3.391.

Thresher, A., 2014: Mk21 and Devil/Quoll timing test. *Fourth XBT Workshop: XBT Science and the Way Forward*, Beijing, China, Institute of Atmospheric Physics, Chinese Academy of Sciences, 9 pp., http://2014xbtworkshop.csp.escience.cn/dct/attach/Y2xiOmNsYjpwZGY6NDUxNTM=.

Wijffels, S. E., J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K. Ridgway, and J. A. Church, 2008: Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. *J. Climate*, **21**, 5657–5672, doi:10.1175/2008JCLI2290.1.