Journal of Oceanography, Vol. 65, pp. 287 to 299, 2009

Reevaluation of Historical Ocean Heat Content Variations with Time-Varying XBT and MBT Depth Bias Corrections

MASAYOSHI ISHII[1]* and MASAHIDE KIMOTO[2]

[1]*Frontier Research Center for Global Change, Japan Agency for Marine-Earth Science and Technology, Showamachi, Kanazawa-ku, Yokohama 236-0001, Japan*
[2]*Center for Climate System Research, The University of Tokyo, Kashiwa 277-8586, Japan*

(Received 16 February 2008; in revised form 26 November 2008; accepted 11 December 2008)

As reported in former studies, temperature observations obtained by expendable bathythermographs (XBTs) and mechanical bathythermographs (MBTs) appear to have positive biases as much as they affect major climate signals. These biases have not been fully taken into account in previous ocean temperature analyses, which have been widely used to detect global warming signals in the oceans. This report proposes a methodology for directly eliminating the biases from the XBT and MBT observations. In the case of XBT observation, assuming that the positive temperature biases mainly originate from greater depths given by conventional XBT fall-rate equations than the truth, a depth bias equation is constructed by fitting depth differences between XBT data and more accurate oceanographic observations to a linear equation of elapsed time. Such depth bias equations are introduced separately for each year and for each probe type. Uncertainty in the gradient of the linear equation is evaluated using a non-parametric test. The typical depth bias is +10 m at 700 m depth on average, which is probably caused by various indeterminable sources of error in the XBT observations as well as a lack of representativeness in the fall-rate equations adopted so far. Depth biases in MBT are fitted to quadratic equations of depth in a similar manner to the XBT method. Correcting the historical XBT and MBT depth biases by these equations allows a historical ocean temperature analysis to be conducted. In comparison with the previous temperature analysis, large differences are found in the present analysis as follows: the duration of large ocean heat content in the 1970s shortens dramatically, and recent ocean cooling becomes insignificant. The result is also in better agreement with tide gauge observations.

Keywords:
· Ocean temperature,
· XBT fall-rate equation,
· oceanographic observation,
· ocean heat content,
· objective analysis.

1. Introduction

Historical ocean temperature analyses have been conducted using in-situ oceanographic observations by purely statistical methods (Levitus *et al*., 2000; Ishii *et al*., 2003; Willis *et al*., 2004; Smith and Murphy, 2007). The analyses are defined on a spatially uniform grid at a regular interval, and they are available for decades. In addition, the observational noise in the analyses is filtered out efficiently through the processes of objective analysis that are adopted in each analysis. Therefore, the temperature analyses can easily be applied to studies of ocean climate changes. For example, global warming signals in the global oceans have been detected from these temperature analyses (Levitus *et al*., 2005; Ishii *et al*., 2006) and interannual variations of sea level (Lombard *et al*., 2005), and in verifying outputs of a model simulation for global warming signals (Sakamoto *et al*., 2005). The aforementioned temperature analyses depend largely on the historical oceanographic observations, although the methodologies of the analyses are different from one another. A recent study by Gouretski and Koltermann (2007) suggests significant positive temperature biases in historical expendable bathythermograph (XBT) and mechanical bathythermograph (MBT) observations, which affects the interdecadal variation of ocean heat content. These biases have not been fully taken into account in the above temperature analyses.

* Corresponding author. E-mail: ism@jamstec.go.jp
On leave from the Meteorological Research Institute of the Japan Meteorological Agency.

XBT observations have been conducted widely over the global oceans since the late 1960s since they permit an easier measurement of the temperature profile as compared with in-situ water sampling by hydrographic bottles. A cause of the biases pointed out by Gouretski and Koltermann (2007) is thought to be due to the absence of simultaneous pressure measurements with the temperature observation. Instead of the pressure measurement, the depths for the observed temperature values are given by fall-rate equations, attributed to each XBT probe type. The equation gives depths (d) as a function of elapsed time (t) from the instant when the probe touches the water surface:

$$d(t) = bt - at^2, \quad (1)$$

where a and b are constants, given individually for XBT probe types. Equation (1) denotes that the probe falls linearly with time and is simultaneously decelerated due to the decrease of probe weight as wire is lost. One possible cause of the temperature biases is positively biased depths given by the fall-rate equations.

In the mid 1990s, a fall-rate equation for popular XBT probe types, provided by several manufacturers, was replaced by a new one based on the results of field experiments (Hanawa *et al*., 1995). Observations by XBT were inspected for depth errors by comparison with accurate measurements by a conductivity, temperature, and depth (CTD) instrument made simultaneously with the XBT measurements. This revision had a significant impact on the long-term trend and detection of the global warming signal. When the correction is not applied to all XBT data, while already applied corrections to XBT data are also eliminated, annual global mean thermosteric sea level rise becomes smaller by 5–15 mm in most years after the mid-1960s (Ishii *et al*., 2006).

In the direct comparison between XBT and CTD measurements, the number of samples is generally of order 10 and the experiments were conducted in limited areas of the global oceans. Such fall-rate equations may thus not operate appropriately at any location in the global oceans at any time since the mid-1960s. Various sources of XBT error should also be taken into account. For instance, Kizu and Hanawa (2002) reported errors from data recorders employed in the XBT observation.

In this study, an XBT "depth" bias correction is introduced to eliminate the positive "temperature" biases from long-term XBT observations (Gouretski and Koltermann, 2007), which remain even after applying the fall-rate equation published by Hanawa *et al*. (1995). Similarly, a depth bias correction of MBT observation is also introduced. An ocean temperature analysis is conducted with these corrections after Ishii *et al*. (2006). The following sections describe the detail of the newly introduced XBT and MBT depth bias corrections, and the depth biases detected by this approach are presented. In our temperature analyses using the bias corrections, we demonstrate how the corrections affect ocean heat content on decadal and interannual time scales. Finally, the evidence supporting this approach is discussed.

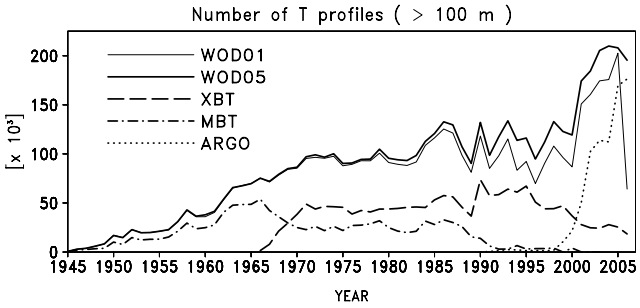

Fig. 1. Time series of the number of temperature profiles in which observations are available at depths greater than 100 m. The numbers are yearly counts for WOD01 (thin solid) and WOD05 (thick solid). Those for XBT (broken), MBT (dash-dotted), and Argo (dotted) in WOA05 are plotted together.

2. Observational Databases

The observed data and temperature and salinity climatology used in this study are the latest version of World Ocean Database (WOD05) and World Ocean Atlas (WOA05), respectively. The data sets are provided by the National Oceanographic Data Center of USA (NODC; Boyer *et al*., 2006). There are two types of data sets in WOD05: Observed level data and Standard level data, the former of which are used in this study. A near-real-time data archive made available through the Global Temperature-Salinity Profile Program (GTSPP) is also used, which compensates the data sparseness of WOD05 since 1990. The GTSPP data have been maintained by NODC under an international cooperation coordinated by the World Meteorological Organization and the Intergovernmental Oceanographic Commission. Furthermore, we added a set of XBT observations compiled by the Japan Oceanographic Data Center. These data are originally provided by the Japan Maritime Self-Defense Force (JMSDF) and are available from 1970 to 2003; they are not found in the above two data sets. Probe types T4 and T5 of Tsurumi Seiki have been used in the observations by JMSDF. The JMSDF data enable us to set up individual statistics for probe type T4 of Tsurumi Seiki.

Comparing the latest WOD05 with the 2001 edition (WOD01), the number of profiles available at depths greater than 100 m increases, particularly in the 1990s and recent years (Fig. 1). From the first appearance of

XBT observation in 1966 until the mid 1990s, XBT has been a major instrument in oceanographic observation. For a period before the XBT era, MBT observation is dominant at the level of 100 m depth, and MBT observations were relatively intensive in the 1970s and 1980s. In recent decades, Argo float data make up more than a half of the observational data since 2003, while XBT observations have been less frequent than before the mid-1990s.

3. XBT Depth Bias Correction

3.1 A bias model

The XBTs are the most error prone oceanographic observations (Gouretski and Koltermann, 2007), because the depths of temperature are not measured directly by the XBT itself, as described above. In this study we suppose that the XBT positive temperature biases have their origin in depth biases and that the depth bias equation takes the same form as Eq. (1):

$$\hat{d} = Bt + At^2. \qquad (2)$$

Letter $\hat{d}$ represents biases in depth. One reason for the introduction of Eq. (2) is because the elapsed time is the primary measurement in the XBT observation. The coefficients of the quadratic equation, B and A, are estimated in a least-square sense from all available samples of depth differences between XBT and accurate observations, such as CTD and hydrographic bottle, together with corresponding elapsed time, t. The elapsed time used above can be computed from the inverse of the fall-rate equation (1) for a given depth of observation.

In the following, coefficients B and A are calculated for each XBT probe type and for each year. When conducting an objective analysis for historical temperature changes, depths of XBT observation are corrected by subtracting depth biases provided by Eq. (2).

Before the evaluation of the depth bias in XBT observation, the XBT fall-rate equation proposed by Hanawa *et al*. (1995) is applied to XBT data of probe types, T4, T6, T7, and DEEP BLUE if necessary, and to a part of XBT data of unknown type, following Conkright *et al*. (2001). Hence, depth biases given by Eq. (2) are an additional correction to those of Hanawa *et al*. (1995).

Probe types are known for a half of XBT observations, but unknown for the other half, in the observational data sets used in this study (Table 1). Observations of unknown XBT probe types are also expected to be contaminated by positive temperature biases. In addition, our goal is not only to construct accurate XBT bias equations for each probe type, but also to eliminate temperature biases from XBT observations for an accurate historical temperature analysis. Hence, an additional type, *type unknown*, is introduced in the following. It is assumed that a single fall-rate equation is attributed to XBT observations of *type unknown*: that of Hanawa *et al*. which is Eq. (1) with $a = 2.25 \times 10^{-3}$ and $b = 6.691$.

Table 1. Fraction of the total number of profiles for each XBT probe type (Frac.; %) stored in the WOD05 and GTSPP data sets for the period from 1966 to 2006, and the number of samples used for constructing equations of the depth-bias correction for each probe type (rightmost column). The total number of profiles is 1,985,888.

No.	XBT type	Frac. [%]	Samples
1	T7 (Sippican)	10.6	310,883
2	T4 (Sippican)	22.5	195,679
3	T6 (Sippican)	0.4	18,921
4	T5 (Sippican)	0.4	69,619
5	T10 (Sippican)	1.9	21,824
6	DEEP BLUE (Sippican)	10.0	145,977
7	T4 (TSK)	1.1	8,657
8	T6 (TSK)	0.7	18,099
9	T7 (TSK)	0.8	12,599
10	XBT-7 (Sparton)	0.2	17,194
11	TYPE UNKNOWN	51.3	530,942

3.2 Box-averaged temperature profiles

Observed data are averaged monthly in box-shaped regions of the global oceans for individual XBT probe types and a mixture of CTD and hydrographic bottle (CTD+BOTTLE) for the period from 1966 to 2006. From the box averages, samples of depth difference between XBT and CTD+BOTTLE observations are taken for the estimation of the coefficients of Eq. (2). Before the box averaging, all observations are inspected and selected through quality control procedures adopted in our objective analysis scheme (Ishii *et al*., 2003, 2006).

The horizontal size of the box is set to 1° longitude by 1° latitude over the entire globe. Depth differences greater than 100 m were observed in actual XBT with CTD+BOTTLE observations at depths greater than 500 m. Therefore, the box is prepared up to 900 m depth for unbiased sampling of depth difference between XBT and CTD+BOTTLE box averages at 700 m depth. In addition, the thickness of the box is 10 m evenly from sea surface to 900 m so as to reduce the vertical interpolation error. In our experience, errors from the vertical interpolation of observed and box-averaged values become substantially smaller than the resultant depth biases when the thickness is 10 m.

In general, observations of XBT and CTD+BOTTLE used for the comparison are located far from each other, by a maximum of about 1.4-degree latitude distance, and there is a time lag within one month between them, too. One may also presume that oceanic temperature devia-

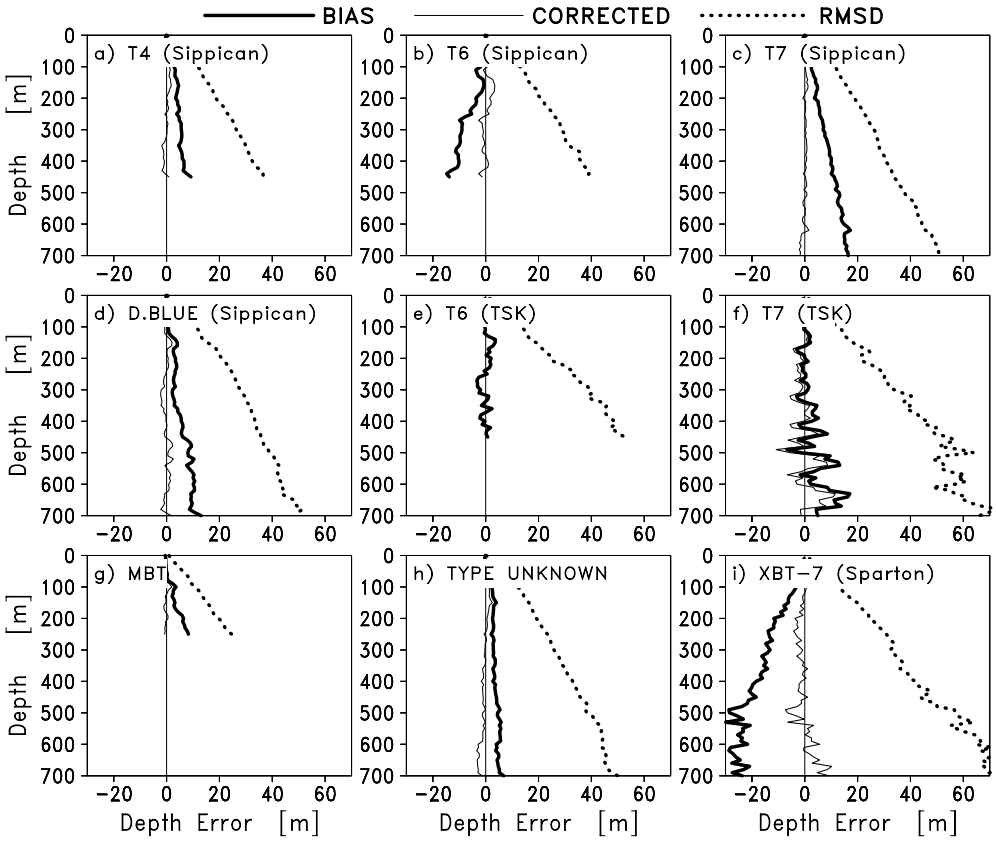

Fig. 2. Profiles of mean XBT depth biases (thick solid; BIAS), root-mean square differences between XBT and CTD+BOTTLE (dotted; RMSD), and residuals after the correction by using the depth bias equation (thin solid; CORRECTED) for each XBT probe type and MBT. Unit is meter.

tions from climatology have large spatio-temporal coherency exceeding 1°-latitude distance and one month (Reynolds and Smith, 1994; Ishii *et al*., 2003). Therefore, when computing the box averages we use observed deviations from the WOA05 climatology interpolated to the position and the date of the observation, rather than the full temperature values. This is expected to minimize errors in the box average, particularly in regions where the horizontal gradient of temperature is large and in months when the seasonal change is rapid.

Large observational errors still remains, due mainly to ocean meso-scale eddies, instrumental errors, and human errors. Some of these are thought to be random error, and hence they can be substantially reduced by using a number of samples when estimating the coefficients of the depth bias equation. The manufacturers state that the error in XBT temperature measurement is ±0.2°C, and that the depth accuracy is a minimum of 5 meters or 2% of depth (Kizu *et al*., 2005). By contrast, the accuracy of CTD and hydrographic bottle temperature observations is much higher than that of XBT, being within a range between ±0.003°C and ±0.02°C for temperature (Wyatt *et al*., 1967; Fahrbach, 1989), with a depth accuracy better than ±2 m for the CTD and within ±15 m for the hydrographic bottle (Wyatt *et al*., 1967; Fahrbach, 1989; Quadfasel *et al*., 1990). The hydrographic bottle error denotes the maximum for the depth measurement range: 0 to 6,000 m depth, and therefore smaller errors can be expected for depth measurements above 1,000 m depth. In addition, the accuracy in the actual observation seems to be much better than the nominal accuracy; that is, the error reduces by half, as Fahrbach (1989) reported. Moreover, these observations are made by trained members on ship and the sensors are generally regularly calibrated.

Within a calender month, observed deviations are averaged with weights which are the same as those used in a bilinear interpolation on the horizontal plane. Here, one observation is shared among the neighboring four boxes, and the weight is set to be largest when the location of observation is closest to the center of box. The sum of the weights for each observation is one. This procedure is expected to some extent to minimize errors from distances between the center of box and the position of observation. It also functions as a spatial filter by which

observational noise is somewhat reduced since observations in surrounding boxes are additionally used in the box averaging. Only box averages whose sum of weights exceeds 0.5 are used when taking a sample of depth difference in the procedure described below.

3.3 Sampling of depth difference

The sampling of the depth difference between XBT and CTD+BOTTLE is conducted with the above mentioned box averages as follows. The climatology subtracted before the box averaging is added back to the box averages. For an XBT box-mean temperature at a level of the box in the upper 700 m depth, a depth of the same temperature as that of the XBT is sought in a profile of CTD+BOTTLE box averages by interpolating linearly in the vertical. The difference between these two depths yields one sample for the estimation of the coefficients, together with elapsed time of the depth of the XBT observation. Such samples are taken from profiles of the XBT and CTD+BOTTLE box averages mutually in the same month and at the same box location. Similarly, an interpolated depth of an XBT profile is computed for a CTD observation at a level of the box in the upper 700 m depth. The above procedure is expected to minimize errors in linearly interpolated depths on average, particularly around the thermocline where the curvature of the temperature profile is large. The samples in the climatological mixed layer are not used, since the ocean temperatures fluctuate widely within a month there. It is also expected that XBT depth biases are small near the sea surface since the biases are assumed to be a function of the elapsed time, as in Eq. (2).

3.4 Mean depth biases

Figure 2 shows profiles of XBT depth bias (thick solid line) which are averages of the depth differences sampled by the above procedure over the whole period. The statistics are set up for individual XBT types. Although there are a number of XBT probe types produced by several manufacturers: Sippican, Sparton, Tsurumi Seiki (TSK) and others, the probe types appearing in the figure are the most dominant in WOD05 and GTSPP, as listed in Table 1. The figure shows that the biases are generally larger at greater depths. It seems acceptable to fit them to an appropriate linear or quadratic equation. Interestingly, the depths of XBT observation do not necessarily have positive deviations; rather, they depend on the XBT types. There also appears to be a dependence of depth bias on manufacturers; the depth biases of Sippican instruments are large in general, while those of Tsurumi Seiki are small. The mean depth biases averaged among all probe types are positive, however, and the positive depth biases appear to cause the positive temperature biases that Gouretski and Koltermann (2007) reported. According to our estimation, the depth bias reaches about 10 m at 700 m depth on average.

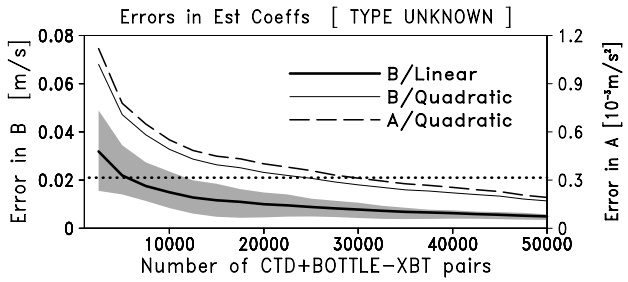

Fig. 3. Uncertainties expressed by 95% confidence interval in the coefficients of the quadratic equation for the XBT depth biases as a function of the number of pairs of XBT and CTD+BOTTLE. The figure is for the case of unknown XBT probe type, showing its uncertainty averaged over the entire period (1966–2006). Thick solid curve indicates errors in the coefficient of the linear equation. Shading indicates ranges between maximum and minimum 95% confidence levels during the whole period of the statistics. For the case of the quadratic equation, uncertainties in B and A are shown by thin solid and broken curves, respectively. The scale (m/s) for B is shown on the left-hand side and that (m/s^2) for A is on the right-hand side. Horizontal dotted line indicates 95% confidence intervals of Hanawa *et al.* (1995).

The depth biases are 3–4 times smaller than root-mean-square differences (RMSDs; dotted), and the latter are due mainly to observational and sampling noise. This implies that the XBT depth bias estimation suffers from noise. A number of samples are therefore required to construct the depth bias equation so as to overwhelm the noise. As discussed in the subsequent section, the lower limit of the number of samples is set to 8,000, and all probe types listed in Table 1 satisfy this threshold.

3.5 Bootstrap test

Either the linear or the quadratic equations for $\hat{d}$ is adopted when conducting the objective analysis of historical temperature. In the linear case, the quadratic term, At^2, is ignored. To determine which equation is finally used we introduced a non-parametric statistical test for errors in the estimated coefficients.

Because of the noisy box averages we prefer to use coefficients B and A, with small uncertainties rather than the coefficients that produce the best-fit regression to the noisy box averages. In addition, the error in depth difference used as a sample is generally larger than that of depth of both XBT and CTD+BOTTLE observations. A bootstrap approach is taken to test the statistical significance in the estimation of the coefficients. From this test we derive an appropriate number of pairs of XBT and CTD+BOTTLE used for the determination of XBT depth

bias equation. All available samples are collected within 10 years, centered at each year at first. Selecting a set of samples randomly at a given number of samples, the coefficients of the equation are determined by a least squares fit. A thousand trials are made for every 2,500 sample size from 2,500 to 50,000.

Figure 3 shows 95% confidence levels for estimated coefficients B and A as a function of the number of samples in the case of *type unknown*. The plotted values are the mean among all the tests over the whole period. Ranges between maximum and minimum 95% confidence levels during the whole period of the statistics are shown by shading only for the case of the linear equation (thick solid). The figure shows that the uncertainty in the coefficient of the linear equation is much smaller than in the case of the quadratic equation (thin solid and broken curves for coefficients, B and A, respectively). Values of 0.021 m/s and 0.00030 m/s^2 indicated by horizontal dotted line denote 95% confidence levels respectively for b and a of the Hanawa *et al*. fall-rate equation based on simultaneous XBT and CTD observations. These errors in b and a correspond to depth errors of 1.5–1.8 m at 500 m depth. The biases to be eliminated are of the order of 10 m, as seen in Fig. 2. To achieve uncertainty as small as that of Hanawa *et al*. (1995), the required sampling size is around 5,000 for the linear case on average, while it is about 30,000 in the quadratic case. The difficulty in the estimation of the quadratic equation is due to the noise in the box-averaged data, which are probably affected by various sources of error. In addition, the ranges between maximum and minimum levels for the quadratic case are double those of the linear case (figure not shown). This implies that many more samples than 30,000 are required to obtain coefficients with small uncertainty throughout the period. Data of *type unknown* could be a mixture of observations by various XBT probe types. Probably for this reason, statistics for data of a single type, such as T4, T7, and DEEP BLUE of Sippican, result in smaller 95% confidence levels and narrower ranges of maximum and minimum levels than those for *type unknown* in general.

The linear equation is therefore adopted for the removal of XBT depth bias. This enables us to set up depth-bias correction equations for XBT probe types listed in Table 1. The threshold of the number of pairs is set to at least 8,000 and 15,000 if possible. The 15,000 threshold ensures that the 95% confidence interval is always less than 0.02. In the cases of XBT probe types other than *type unknown*, the threshold is also good enough to ensure small estimation errors. A set of 8,000–15,000 samples is based on more than 85–165 observed profiles of XBT and CTD+BOTTLE individually. Thin solid curves in Fig. 2 show the residual of the differences between XBT and CTD+BOTTLE after applying the correction given by the linear equation. The depth biases are reduced substantially as a result.

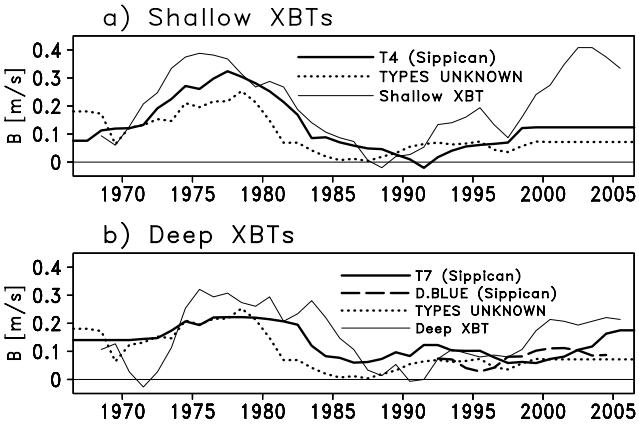

Fig. 4. Time series of coefficient B (m/s) of the linear equation for a) T4 (thick solid) and b) T7 and DEEP BLUE ("D. BLUE" in the figure) of Sippican (thick solid and broken, respectively), compared with those for shallow and deep XBTs given by Wijffels *et al*. (2008) (thin solid) and with those of *type unknown* (dotted). Values indicated by thin solid curves are obtained after multiplying fractional depth errors (depth error divided by depth) listed in their table by 6.691 which is b of the Hanawa *et al*. fall-rate equation. Line of DEEP BLUE before 1992 is not shown since they do not vary over this period.

3.6 Temporal changes in XBT depth biases

It is thought that there are various source of depth biases shown in Fig. 2. However, we do not know when and how such biases became mixed with XBT observation. Therefore, the coefficients of the linear depth bias equation are computed on a yearly basis for each XBT probe type. Samples in a given year are collected until the number of sample reaches 15,000, extending the range of the data search back and forth beyond the year under investigation. In order to avoid noisy temporal changes in coefficient B, samples for at least 5 years are used to determine the coefficient. If the sampling size is smaller than the threshold, 8,000, even after a search over the whole period, the coefficients are not computed.

Coefficients of B for T4, T7, and DEEP BLUE of Sippican and *type unknown* are shown in Fig. 4, compared with depth corrections for shallow and deep XBT probe types calculated by Wijffels *et al*. (2008). The shallow (deep) XBT is the one with maximum depth less (greater) than or equal to 550 m. In their approach, depth bias corrections are given by multiplying a factor by depth of observation. The factor is defined respectively for shallow and deep XBTs as a function of year. This approach correcting depth of observation is the same as that adopted

Table 2. Coefficient *B* of the time-varying XBT depth bias equation (2) defined for individual XBT probe types. Some probe types are shown by abbreviation; DB and UK denote DEEP BLUE and TYPE UNKNOWN, respectively.

Year	Sippican						TSK			Sparton	UK
	T7	T4	T6	T5	T10	DB	T4	T6	T7	XBT-7	
1966	0.061	0.175	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.181
1967	0.061	0.175	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.181
1968	0.061	0.074	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.184
1969	0.061	0.090	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.176
1970	0.062	0.102	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.035
1971	0.143	0.190	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.093
1972	0.167	0.206	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.118
1973	0.200	0.228	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.217
1974	0.177	0.266	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.184
1975	0.194	0.263	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.205
1976	0.211	0.305	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.176
1977	0.234	0.322	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.333
1978	0.222	0.319	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.245
1979	0.224	0.319	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.239
1980	0.217	0.259	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.197
1981	0.198	0.164	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.075
1982	0.138	0.124	−0.178	0.067	0.059	0.037	0.051	0.014	0.055	−0.317	0.006
1983	0.136	0.133	−0.178	0.068	0.059	0.037	0.051	0.014	0.055	−0.317	0.046
1984	0.083	0.072	−0.178	0.109	0.059	0.037	0.051	0.014	0.055	−0.317	0.044
1985	0.065	0.081	−0.178	0.109	0.059	0.037	0.051	0.014	0.055	−0.317	0.034
1986	0.036	0.053	−0.178	0.109	0.059	0.037	0.051	0.014	0.055	−0.317	−0.015
1987	0.011	0.063	−0.178	0.109	0.055	0.037	0.051	0.014	0.055	−0.317	−0.019
1988	0.064	0.046	−0.178	0.109	0.056	0.037	0.051	0.014	0.055	−0.317	0.003
1989	0.097	0.022	−0.178	0.109	0.081	0.037	0.051	0.014	0.055	−0.318	0.057
1990	0.105	−0.016	−0.178	0.109	0.082	0.037	0.051	0.014	0.055	−0.318	0.059
1991	0.087	−0.022	−0.178	0.106	0.082	0.037	0.051	0.014	0.055	−0.318	0.060
1992	0.130	−0.040	−0.178	0.127	0.082	0.037	0.051	0.014	0.055	−0.317	0.066
1993	0.128	0.078	−0.178	0.136	0.082	0.031	0.051	0.014	0.055	−0.286	0.074
1994	0.109	0.089	−0.179	0.167	0.089	0.074	0.051	0.014	0.055	−0.258	0.074
1995	0.086	0.101	−0.255	0.095	0.090	0.016	0.051	0.014	0.055	−0.247	0.056
1996	0.061	0.100	−0.244	0.121	0.090	0.032	0.051	0.015	0.055	−0.171	0.033
1997	0.073	0.106	−0.240	0.143	0.100	0.036	0.051	0.015	0.064	−0.171	0.036
1998	0.038	0.121	−0.238	0.170	0.086	0.093	0.051	0.009	0.064	−0.171	0.057
1999	0.057	0.180	−0.239	0.130	0.086	0.128	0.051	−0.004	0.064	−0.171	0.053
2000	0.065	0.183	−0.235	0.142	0.083	0.120	0.051	−0.009	0.065	−0.171	0.123
2001	0.085	0.183	−0.233	0.108	0.055	0.100	0.051	−0.011	0.049	−0.171	0.120
2002	0.098	0.183	−0.233	0.021	0.054	0.095	0.051	0.001	0.039	−0.171	0.120
2003	0.156	0.183	−0.233	0.014	0.054	0.094	0.051	−0.018	0.038	−0.171	0.120
2004	0.172	0.183	−0.233	0.032	0.054	0.086	0.051	−0.017	0.038	−0.171	0.120
2005	0.180	0.183	−0.233	0.032	0.054	0.087	0.051	−0.017	0.038	−0.171	0.120
2006	0.180	0.183	−0.233	0.032	0.054	0.083	0.051	−0.017	0.038	−0.171	0.120

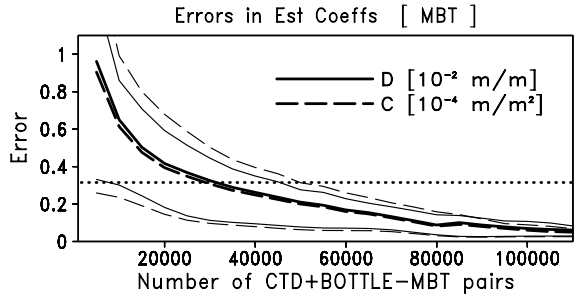

Fig. 5. Uncertainties expressed by 95% confidence interval in the coefficients of the quadratic equation for the MBT depth biases as a function of the number of pairs of MBT and CTD+BOTTLE. Solid and broken curves indicate errors in coefficients D (10^{-2} m/m) and C (10^{-4} m/m^2), respectively. Thick lines denote the averages over a period from 1945–2003, and the lower and upper lines respectively show minimum and maximum in the period of statistics. Horizontal dotted line is equivalent to that in Fig. 3, that is 95% confidence intervals of Hanawa *et al*. (1995).

here. All time series commonly show maximum values of B in the latter half of the 1970s, with coefficients decreasing in the latter half of the 1980s and increasing in recent years. It is necessary to derive a depth bias correction, even for *type unknown*. When B is 0.3 m/s, the observed depth is corrected by about –24 m at 500 m depth. The differences among probe types, including *type unknown*, exceed 95% confidence levels of less than 0.02 in most years. The coefficient of T4 and *type unknown* is mostly invariant in time after 2000 due to poor sampling, and similarly in the case of T7 before 1970. Around these periods there are great differences between the coefficients obtained by Wijffels *et al*. (2008) and this study.

Table 2 contains the coefficients of depth biases, B of the linear equation, for probe types for which the number of samples exceeds 8,000 (Table 1). In the case of probe type T6 of Tsurumi Seiki (TSK), the coefficient is always less than the estimation error, 0.02 m/s.

4. MBT Depth Bias Correction

In the case of MBT, the approach to the depth bias correction is the same as that of XBT, except for a bias equation containing a function of depth of observation rather than elapsed time, as in the case of XBT. The depth biases ($\hat{d}$) of MBT reported by Gouretski and Koltermann (2007) are represented by a quadratic function of depth (z) as follows:

$$\hat{d} = Dz + Cz^2. \quad (3)$$

Coefficients C and D are calculated in a least-squares sense with samples of depth differences between box-averaged MBT and CTD+BOTTLE at depths above 250 m depth (Fig. 2g) for the period from 1945 to 2006. No individual MBT instrument types are considered here, so a single pair of coefficients are computed as a function of the year. Although linear bias equations are used for XBT, quadratic bias equations are adopted for the elimination of MBT depth biases. This is because the fit with sampled depth biases is better than that of the linear equation and because many samples are available for estimating the coefficients. Figure 5 shows a bootstrap test for MBT depth bias. Thick lines indicate mean values of the 95% confidence levels for D (10^{-2} m/m; solid) and C (10^{-4} m/m^2; broken), and upper and lower thin lines denote the maximum and minimum of the 95% confidence levels, respectively. Values indicated by these lines are interpretable as errors in meters at 100 m depth. The horizontal dotted line is drawn as a confidence level equivalent to the 95% confidence interval at 100 m depth of the Hanawa *et al*. fall-rate equation.

The error in bias estimated by Eq. (3) is a sum of errors originating from both the first and second terms of

Table 3. Coefficients of the time-varying MBT depth bias equation: $\hat{d} = Dz + Cz^2$. Units of coefficients C and D are 10^{-4} m/m^2 and 10^{-2} m/m, respectively. The coefficients before 1950 and after 1994 are the same as those for 1948 and 1996, respectively.

Year	C	D	Year	C	D
1950	2.71	–0.57			
1951	2.75	–0.66			
1952	2.75	–0.66			
1953	2.76	–0.68			
1954	2.75	–0.66			
1955	2.02	–0.16			
1956	1.98	–0.11	1976	1.64	0.54
1957	1.64	0.04	1977	1.59	0.34
1958	1.61	–0.01	1978	2.96	–0.94
1959	1.87	–0.43	1979	2.30	–0.22
1960	1.64	0.00	1980	2.80	–0.85
1961	1.18	0.51	1981	2.86	–1.02
1962	1.00	0.98	1982	2.61	–1.00
1963	1.05	1.11	1983	2.15	–0.83
1964	1.35	0.90	1984	2.79	–1.44
1965	0.62	1.52	1985	2.22	–1.11
1966	0.59	1.58	1986	1.82	–0.98
1967	0.67	1.36	1987	1.84	–1.17
1968	0.73	1.26	1988	2.02	–1.37
1969	0.33	1.68	1989	2.00	–1.48
1970	1.03	0.80	1990	2.16	–2.07
1971	1.26	0.71	1991	1.72	–1.61
1972	1.14	0.95	1992	1.75	–1.66
1973	1.08	0.88	1993	1.74	–1.65
1974	1.02	0.96	1994	1.56	–1.49
1975	1.27	0.75			

Table 4. Versions of the temperature analysis compared in this study.

Ver.	Changes
6.2	Ishii *et al*. (2006): use of WOD01 and WOA01, and no XBT and MBT depth bias correction
6.3	Same as ver. *6.2* except use of WOD05 and WOA05
6.6	Same as ver. *6.3* except introduction of XBT and MBT depth bias correction
6.6X	Same as ver. *6.3* except introduction of only XBT depth bias correction
6.7	Same as ver. *6.6* except replacement of temperature climatology

Eq. (3). In the case of 80,000, the maximum error is expected to be nearly equal to the error of the Hanawa *et al*. fall-rate equation. The threshold number of samples is set to 80,000 in the case of the estimation of the MBT bias equation.

Table 3 displays the time series of coefficients C and D from 1950 to 1994. The values for C and D can be regarded as biases in meters at 100 m depth, evaluated respectively by the second and first terms of Eq. (3). The MBT depth bias at 100 m depth, as a sum of the linear and quadratic components, is about 2 m before 1980, gradually decreasing close to zero toward 1990. The MBT bias profile generally has a parabolic shape. The linearity of the bias profile is relatively large during the latter half of the 1960s.

5. Depth-Bias Corrected Ocean Heat Content

Several ocean temperature analyses have been conducted for the period from 1945 to 2006, following the version number *6.2* of Ishii *et al*. (2006). In the objective analysis, the temperatures are defined on a $1° \times 1°$ grid at 16 levels from the sea surface to 700 m depth. The analyzed temperatures are deviations from the WOA05 climatology, optimally estimated by a three dimensional variational minimization technique (Ishii *et al*., 2003).

In version *6.3*, a temperature analysis is conducted using the latest observational database (WOD05) and climatology (WOA05) in order to use more observations that are available after 1990 (Fig. 1). Version *6.2* used WOD01 and WOA01. The long-term averages of temperature analysis with the XBT depth bias correction (ver. *6.6*) differ significantly from the WOA05 climatological temperature, as shown below. Hence, we introduced another temperature climatology, computed from the average of temperature analysis of ver. *6.6* over the recent 40 years from 1961 to 2000. An analysis with the new climatology is called version *6.7* hereafter. In order to see the impact of the MBT depth bias correction, an additional analysis is introduced, in which only the XBT depth bias correction is adopted (ver. *6.6X*). The versions of the temperature analysis described above are summarized in Table 4.

Ocean heat content (OHC) discussed in this section

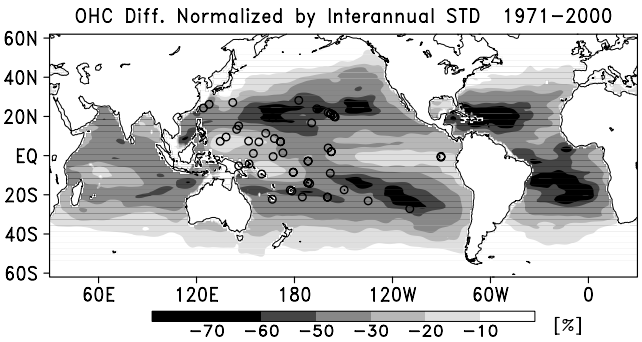

Fig. 6. Geographical distribution of difference in ocean heat content (OHC) computed from temperature analyses of ver. *6.6* minus ver. *6.3*. The negative differences are shaded. The difference at each grid is normalized by the standard deviation of the interannual variation of OHC. Circles indicate the locations of the tide gauge stations used for the verification in Fig. 8.

is computed from the monthly temperature analysis above 700 m depth and WOA05 monthly salinity climatology. A grid-wise OHC at latitude ϕ is defined as a vertical integration of heating profile:

$$\int_{-700}^{0} \rho(T, S_c) C_p \left(T - T_{ref}\right) R^2 \cos(\phi) dz, \qquad (4)$$

where $\rho(T, S_c)$ is the density of sea water for temperature T and climatological salinity S_c at depth z, C_p is a specific heat of sea water, T_{ref} is a reference temperature, and R is the radius of the earth. Although there are various choices of T_{ref} (e.g., Levitus and Antonov, 1997; Antonov *et al*., 2004), they are not so essential as to affect results presented in the following. In order to see explicit OHC changes caused by the depth bias equation, T_{ref} is set to zero in this study.

A temperature bias derived from the XBT depth bias is proportional to the vertical temperature gradient which varies from 0.02 to 0.2°C/m along the climatological thermocline. The gradient is mostly less than 0.05°C/m at depths below 300 m, while large gradients exceeding 0.1°C/m appear around 100 m depth in the tropics along

Table 5. Linear trends of global (60°S–60°N) annual mean OHC (10^{22} J/yr) and thermosteric sea level (mm/yr) estimated from each versions of temperature analysis respectively for a long-term period (1951–2005) and the latest decade (1993–2005).

Ver.	OHC		Thermosteric sea level	
	1951–2005	1993–2005	1951–2005	1993–2005
6.2	0.141 ± 0.035	0.296 ± 0.230	0.300 ± 0.067	0.805 ± 0.448
6.3	0.127 ± 0.035	0.369 ± 0.154	0.262 ± 0.063	0.899 ± 0.304
6.6	0.105 ± 0.034	0.590 ± 0.152	0.236 ± 0.066	1.24 ± 0.298
6.6X	0.075 ± 0.034	0.581 ± 0.152	0.180 ± 0.064	1.23 ± 0.297
6.7	0.147 ± 0.029	0.582 ± 0.151	0.294 ± 0.057	1.23 ± 0.295

the seasonal thermocline.

After introducing the XBT and MBT depth bias corrections in ver. *6.6*, cooling in the ocean temperatures are largely seen along thermocline depths in the subtropics (Fig. 6). This is reasonable because the depth biases become larger at greater depths and because the main thermocline is located around 200–400 m depth in the subtropics, while it shallows in the tropics. This result motivates us to replace WOA05 used in the objective analysis of ocean temperature by another temperature climatology in order to minimize gaps before and after the late 1960s, when the XBT observation started. The MBT bias correction contributes to some extent to the OHC differences, as discussed later. Deviations from the climatology are analyzed in the temperature analysis of Ishii *et al*. (2003, 2006). The new climatology, which is unbiased, is expected to yield an unbiased temperature analysis, especially in a data-sparse period.

Figure 7 shows the annual global mean time series of ocean heat content computed from all versions of the temperature analysis. The necessity of replacing the temperature climatology, mentioned in connection with Fig. 6 above, is reconfirmed by the figure; large differences are seen since the late 1960s between the analyses of vers. *6.3* (dotted) and *6.6* (dash-dotted). The differences mostly originate from the XBT depth biases, as the contribution of the MBT depth bias correction (solid line with circle) is relatively large before 1970 but small in later decades. By changing the climatology, an unbiased analysis is realized in the ver. *6.7* analysis (thick solid) before 1970.

Table 5 summarizes linear trends in OHCs and thermosteric sea levels derived from each version of the analysis. The thermosteric sea level is a term describing sea level affected by thermal expansion and contraction of sea water, and it is computed with the temperature analysis in the upper 700 m depth and climatological salinity (WOA05). The OHC trend of the ver. *6.7* analysis for 1955–2005 does not change significantly in comparison with ver. *6.2*. By contrast, the trend for the most recent 13 years (1993–2005) almost doubles, for the following reason: from the mid-1990s, the OHC differences between vers. *6.3* and *6.7* become small because the number of XBT observations gradually decreases (Fig. 1). This makes OHC larger in recent years than before the mid-1990s, and results in a trend that is double the previous value. The replacement of the climatology in ver. *6.7* contributes to some extent, but it is not a main factor over this period, as seen in Fig. 7. When the MBT bias correction is not introduced (ver. *6.6X*), the long-term trend becomes smaller significantly in comparison with ver. *6.6*.

A notable difference in OHC appears during the 1970s among the analyses. In the analyses prior to ver. *6.6*, OHC increases early in the 1970s and its anomaly stays largely positive throughout the decades. In version *6.7*, OHC similarly increases from 1970, but is slightly depressed in the latter half of the 1970s, increasing again toward the maximum in 1980. This result does not contradict figure 2 of Gouretski and Koltermann (2007). Sippican probes T4 and T7 were dominant during these three periods and they display large differences in depth from CTD+BOTTLE (Fig. 4).

In the previous analysis (ver. *6.2*), great ocean cooling appears since 2003 with a sizable amplitude (an OHC reduction indicated by the thin line in Fig. 7). This is partly due to a large cold bias in a small fraction of Argo floats as reported by Willis *et al*. (2008). These Argo floats measure lower temperatures than expected, particularly in the Atlantic Ocean. In ver. *6.3* and later these Argo observations were withheld from the objective analysis. The difference between versions *6.2* and *6.3* since 2003 is mainly caused by the problematic Argo data, while differences between WOD01 and WOD05 used in the two analyses yield far fewer changes throughout the period, as shown in Fig. 7. Moreover, in ver. *6.6*, the recent ocean cooling becomes insignificant thanks to the introduction of the XBT depth bias correction. The reason for this is the same as that for the doubled OHC trend mentioned above. According to a study in which global mean sea level changes due to thermal expansion and contraction (thermosteric) are estimated from satellite sea surface height observations excluding ocean mass changes by

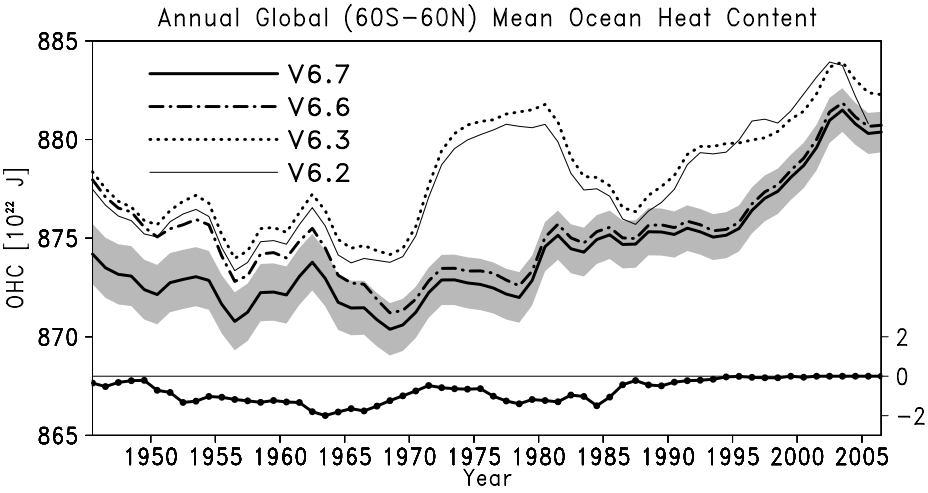

Fig. 7. Time series of annual global mean ocean heat contents (10^{22} J) computed from the previous (ver. *6.2*; V6.2; thin solid) and present temperature analyses: dotted for ver. *6.3* (V6.3), dash-dotted for ver. *6.6* (V6.6), and thick solid for ver. *6.7* (V6.7), integrated from sea surface to 700 m depth. Shading denotes one-sigma errors, following the curve of the ver. *6.7* OHC. Changes in OHC by the MBT depth bias correction are denoted by line with solid circles in the lower part of the figure. The values are differences (10^{22} J) between the versions *6.6* and *6.6X* analyses, and the scale is shown in the right hand side of the figure. A binomial filter with weights 0.25, 0.5, and 0.25, is applied to all time series in the figure for a better visibility of their differences. The ver. *6.2* (V6.2) OHC is padded by 11×10^{22} J that is originated from the difference of climatology, WOA01 and WOA05, used in the objective analysis.

satellite gravity measurements (Lombard *et al*., 2007), such cooling does not appear in time series of the satellite-derived thermosteric sea level.

6. Discussion

As seen in Figs. 6 and 7, the XBT depth bias correction affects the interannual to interdecadal variations of ocean heat content (OHC) as well as long-term averages of ocean temperature. However, it is hard to verify these changes directly by other oceanographic observations, the spatio-temporal coverage of which is not sufficient for such an investigation in general. We have made a trial of such a verification using sea level observations measured by tide gauges. The tide gauge observations used are research quality data compiled by the University of Hawaii Sea Level Center.

Tide gauge observations suffer from various local effects, e.g., ocean tide, coastal waves, crustal movement, especially at stations along continental coasts. In addition, dramatic changes due to the introduction of the depth bias correction were hardly seen in statistics between tide gauge data and thermosteric sea levels estimated from the temperature analysis at each tide gauge station, compared with table 1 and figure 5 of Ishii *et al*. (2006). Hence, sea levels averaged over all available tide gauge stations located on islands were compared with thermosteric sea levels at dates and positions corresponding to the tide gauge observations. The thermosteric data are computed

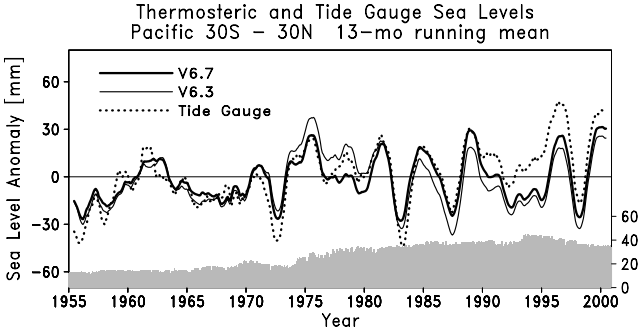

Fig. 8. Time series of monthly tide gauge sea level anomaly (dotted) and thermosteric sea level anomalies estimated from the ver. *6.3* analysis without the XBT and MBT depth bias correction (thin solid) and from that of the ver. *6.7* with the correction. The values are 13-month running means of sea level at tide gauge stations available in latitudes from 30°S to 30°N of the Pacific Ocean (marked by circles in Fig. 6). Bars denote the number of tide gauge stations available for each month, and its scale is shown on the right-hand side.

from the monthly temperature analysis and the WOA05 monthly salinity climatology. A target area for the verification is the tropical and subtropical Pacific between 30°S and 30°N, where the cooling due to the depth bias correction appears to be large (Fig. 6) and where the thermosteric effect is larger than the contribution of sa-

linity to sea level change (figure 9 of Ishii *et al*., 2006).

Figure 8 depicts the time series of monthly sea level anomalies for tide gauge observation (dotted) and the corresponding thermosteric sea level anomalies (thin and thick solid lines, respectively, for vers. *6.3* and *6.7*). Tide gauge stations are adopted for the comparison if the data are available in more than 120 months in the period from 1950 to 2005. The tide gauge stations used for the comparison are shown by circles in Fig. 6. The three time series agree well. In addition, the agreement between tide gauge and the ver. *6.7* analysis is better than that for ver. *6.3*, especially around 1975 and in years after 1985. The root-mean-square differences and correlation coefficients against the tide gauge observation are 16 mm and 0.77 for ver. *6.7*, respectively, while they are 19 mm and 0.68 for ver. *6.3*.

Following the discussion presented by AchutaRao *et al*. (2007), we produce two items of evidence in support of the disappearance of positively deviated OHC during the 1970s seen in the previous temperature analysis (ver. *6.2*). One point in evidence is that dynamical models arranged for reproduction of the 20th century climate with global warming gases and volcanic forcing could not reproduce anomalous OHCs over this period (figure 1 of their manuscript). The other is an abnormally rapid increase of OHC at the beginning of the 1970s, seen in the *6.2* and *6.3* analyses (Fig. 7). AchutaRao *et al*. (2007) concluded that recent ocean cooling since 2003 is artificial because state-of-art dynamical models are unable to replicate changes of OHC from 2003 to 2005 (their figure 3), based on statistics of 2-year OHC changes. A similar conclusion is drawn after their discussion, because the OHC change at the beginning of the 1970s ($+4.5 \times 10^{22}$ J per 2 years) is the same size as that for the recent ocean cooling (-4.5×10^{22} J per 2 years).

Following the result of the bootstrap test (Fig. 3), linear equations were chosen for the depth bias correction. We did an additional analysis with quadratic equations with the same threshold number of XBT-(CTD+BOTTLE) pairs as that for the linear equations. In this case, the estimation errors are twice as large as the linear case (Fig. 3). Although the selection of the order of the depth bias equation affects the grid-wise OHCs, the differences between the two analyses are mostly within 30%, and the global and basin-scale averages of OHC were almost the same. Therefore, at present, it is better to choose the linear equation for the removal of the depth bias since the biases defined by the quadratic equations involve large uncertainties.

There is no method for detecting depth biases in past XBT observations other than the approach taken in this study, or similar ones (e.g., Wijffels *et al*., 2008). Although the estimation of the bias equation by our approach suffers from noise in box averages of XBT and CTD+BOTTLE observations, the bootstrap test ensures that a reliable bias equation can be obtained by increasing the number of samples used for the estimation. Needless to say, there is a great advantage in Hanawa *et al*. (1995) in that they used simultaneous XBT and CTD observations.

7. Conclusion

Depth bias corrections for XBT and MBT are introduced to the historical temperature analysis of Ishii *et al*. (2006) in order to remove positive temperature biases in XBT and MBT observations identified by Gouretski and Koltermann (2007). The results of this study suggests the need to reevaluate interannual to decadal variations of ocean heat content as well as the global warming trend, particularly in recent decades.

Time-varying depth bias equations have been constructed for each XBT probe type and MBT to reduce the differences between bathythermographs and CTD+BOTTLE observations, although our knowledge is inadequate to specify the cause of the positive temperature biases. The correction depending on probe type is very necessary for a reasonable temperature analysis in regions of the ocean basins, since depth biases appear to be largely attributed to the probe types. Moreover, specific XBT probe types tend to be used in limited regions of the oceans. Furthermore, it is noteworthy that the depth bias equations obtained depend on the observational databases in this study. Accurate metadata on XBT probes are needed for a better understanding and estimation of XBT biases.

Observations by XBT as well as MBT are certainly important for long-term climate studies because of their spatio-temporal coverage. Continuous efforts should be made to use them better.

Acknowledgements

The authors wish to thank Prof. K. Hanawa (Tohoku Univ.), Dr. K. Mizuno (IORGC/JAMSTEC), Dr. N. Shikama (IORGC/JAMSTEC), and Prof. T. Suga (Tohoku Univ.) for fruitful comment and discussion on the oceanographic observations. They appreciate the great help of Mr. T. Umeda (JMA) and staff members of JMSDF, given when using domestic XBT observations. Thoughtful comments by three anonymous reviewers are thankfully acknowledged. Appreciation is extended to President H. Iwamiya and Mr. K. Amaike of the Tsurumi Seiki Co. Ltd., for valuable information on the history of XBT. Observational data used in this study we obtained from the following centers: WOA05, WOD05, and GTSPP from NODC/NOAA, tide gauge observation from University of Hawaii Seal Level Center, and JMSDF XBT data from JODC. This work was supported by the Ministry of Education, Culture, Sports, Science and Technology - Japan

through the Innovative Program of Climate Change Projection for the 21st Century.

References

AchutaRao, K. M., M. Ishii, B. D. Santer, P. J. Gleckler, K. E. Taylor, T. P. Barnett, D. W. Pierce, R. J. Stouffer and T. M. L. Wigley (2007): Simulated and observed variability in ocean temperature and heat content. *Proc. Natl. Acad. Sci.*, **104**, 10768–10773.

Antonov, J. I., S. Levitus and T. P. Boyer (2004): Climatological annual cycle of ocean heat content. *Geophys. Res. Lett.*, **31**, L04304, doi:10.1029/2003GL018851.

Boyer, T. P., J. I. Antonov, H. E. Garcia, D. R. Johnson, R. A. Locarnini, A. Mishonov, M. T. Pitcher, O. K. Baranova and I. V. Smolyar (2006): *World Ocean Database 2005*. NOAA Atlas NESDIS 60, ed. by S. Levitus (available at http://www.nodc.noaa.gov/OC5/WOD05/docwod05.html).

Conkright, M. E., R. A. Locarnini, H. E. Garcia, T. D. O. Brien, T. P. Boyer, C. Stephens and J. I. Antonov (2001): *World Ocean Atlas 2001: Objective Analyses, Data Statistics, and Figures*. NOAA Atlas NESDIS 42, 17 pp., CD-ROM, U.S. Government Printing Office, Washington, D.C.

Fahrbach, E. (1989): The use of electronic digital thermometers and pressure meters. *WOCE Newsletter No. 8*, 12–13.

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

Hanawa, K., P. Raul, 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 (XBTs). *Deep-Sea Res.*, **42**, 1423–1451.

Ishii, M., M. Kimoto and M. Kachi (2003): Historical ocean subsurface temperature analysis with error estimates. *Mon. Wea. Rev.*, **131**, 51–73.

Ishii, M., M. Kimoto, K. Sakamoto and S.-I. Iwasaki (2006): Steric sea level changes estimated from historical
subsurface temperature and salinity analyses. *J. Oceanogr.*, **61**, 155–170.

Kizu, S. and K. Hanawa (2002): Recorder-dependent temperature error of expendable bathythermograph. *J. Oceanogr.*, **58**, 469–476.

Kizu, S., H. Yoritaka and K. Hanawa (2005): A new fall-rate equation for T-5 expendable bathythermograph (XBT) by TSK. *J. Oceanogr.*, **61**, 115–121.

Levitus, S. and J. Antonov (1997): *Limatological and Interannual Variability of Temperature, Heat Storage, and Rate of Heat Storage in the Upper Ocean*. NOAA Atlas NESDIS 16, 186 pp.

Levitus, S., C. Stephens, J. I. Antonov and T. P. Boyer (2000): *Yearly and Year-Season Upper Ocean Temperature Anomaly Fields, 1948–1998*. NOAA Atlas NESDIS 40 (available from http://www.nodc.noaa.gov/OC5/PDF/ATLAS/nesdis40.pdf).

Levitus, S., J. I. Antonov, T. P. Boyer, H. E. Garcia and R. A. Locarnini (2005): Linear trends of zonally averaged thermosteric, halosteric, and total steric sea level for individual ocean basins and the world ocean, (1955–1959)–(1994–1998). *Geophys. Res. Lett.*, **32**, L16601, doi:10.1029/2005GL023761.

Lombard, A., A. Cazenave, P.-Y. Le Traon and M. Ishii (2005): Contribution of thermal expansion to present-day sea-level change revisited. *Global and Planetary Change*, **47**, 1–16.

Lombard, A., D. Garcia, G. Ramillien, A. Cazenave, R. Biancale, J. M. Lemoine, F. Flechtner, R. Schmidt and M. Ishii (2007): Estimation of steric sea level variations from combined GRACE and Jason-1 data. *Earth Planet. Sci. Lett.*, **254**, 194–202.

Quadfasel, D., N. Verch and J. Langhof (1990): Are mercury deep-sea reversing thermometers out of date? *Ocean Dynamics*, 145–152.

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

Sakamoto, T., H. Hasumi, M. Ishii, S. Emori, T. Suzuki, T. Nishimura and A. Sumi (2005): Responses of the Kuroshio and the Kuroshio Extension to global warming in a high-resolution climate model. *Geophys. Res. Lett.*, **32**, doi:10.1029/2005GL023384.

Smith, D. M. and J. M. Murphy (2007): An objective ocean temperature and salinity analysis using covariances from a global climate model. *J. Geophys. Res.*, **112**, C02022, doi:10.1029/2005JC003172.

Wijffels, S., J. Willis, C. M. Domingues, P. Baker, N. J. White, A. Cronell, K. Ridgway and J. A. Church (2008): Changing expendable bathythermograph fallrates and their impact on estimates of thermosteric sea level rise. *J. Climate*, **21**, 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**, C12036, doi:10.1029/2003JC002260.

Willis, J. K., J. M. Lyman, G. C. Johnson and J. Gilson (2008): In situ data biases and recent ocean heat content variability. *J. Atmos. Oceanic Tech.* (in press).

Wyatt, B., R. Still, D. Barstow and W. Gilbert (1967): *Hydrographic Data from Oregon Waters 1965*. Department Oceanography, School of Science, Oregon State Univ. Date Report No. 27.