Empirical correction of XBT data

M. Hamon[1], G. Reverdin[2], and P.-Y. Le Traon[1]

[1] Laboratoire d'Océanographie Spatiale, IFREMER, B.P.70 Plouzané, France

[2] LOCEAN, IPSL, CNRS/IRD/MNHN/UPMC, 4 place Jussieu 75252 Paris, France

Correspondent author : Gilles Reverdin, gilles.reverdin@locean-ipsl.upmc.fr

Abstract

We used a collocation method between XBT and CTD/OSD (Ocean Station Data including bottle cast and low resolution CTD) from WOD2005 to statistically correct the XBT fall rate. An analysis of the annual median bias on depth showed that it is necessary to apply a thermal correction, a second order correction on the depth as well as a depth offset representing measurement errors during XBT deployment. We separated data in several categories: shallow and deep XBTs and below or above 10°C of vertically averaged ocean temperatures (in the top 400m). We also processed separately XBT measurements in the western Pacific between 1968 and 1985 due to large regional biases. The estimates we find for these corrections deviate largely from other published estimates with some large variations in time of both linear and curvature terms in the depth corrections, and less time variation of the temperature correction for the deep XBTs. This analysis of heat content derived from corrected XBTs provides at first order a similar variability to other estimates from corrected XBTs and MBTs. It shows a fairly prominent trend in 0-700m ocean heat content of $0.39.10^{22}$J/year between 1970 and 2008.

1. Introduction

Identifying and quantifying the changes in ocean heat content is one of the most important research areas for the international oceanographic community. Due to its heat capacity, much larger than the other elements of the climate system, it is estimated that the oceans have absorbed more than 80% of the earth warming due to the anthropogenic increase of greenhouse gas concentration (Levitus et al, 2001, 2005). In the last few years, many studies have tried to accurately determine the evolution of the global ocean heat content (e.g. Gouretski and Koltermann, 2007; Wijffels et al., 2008; Levitus et al., 2009). These studies identified systematic differences between the different instruments used to collect ocean temperature profiles that need to be corrected. Since 1966, expendable Bathy-Thermographs (XBT) mostly launched from ships, have been used to measure the upper ocean's temperature, and constitute the most important source of upper ocean data between the late 1960s and 2000. The XBT system does not directly measure depth. The accuracy of the depth associated with each temperature depends on an equation which converts the time elapsed since the probe entered the water to depth.

We will not review in details the issues with the determination of this equation, which can be found in Hanawa et al (1995) for early researches and in Gouretski and Reseghetti (2010; later on, GR10) for a more recent review. The parabolic character of the fall rate equation was initially validated by the observation of the fall of a probe in a fresh water tank at a homogeneous temperature. The linear term is a terminal velocity of the instrument and the second-order term is there to take into account mass changes of the probes as the wire is spun during its fall. It has been known since the early uses of the probes that the fall rate should depend on the sea water physical characteristics, with for example a dependency on viscosity/temperature/density of sea water (Tadathil et al., 2002; Kizu et al., 2011). It has also been suggested early on that the assumption of a terminal velocity might not be always correct, in particular in the surface layer, and that this, compounded with time constant issues, can result in a depth offset (although the determination of this depth offset is not straigthforward, as discussed by di Nezio and Goni (2010), as it depends on many parameters hardly known and probably very variable). The weight and hydrodynamic characteristics of the probe/wire are known to strongly influence the fall rate equation. Seaver and Kuleshov (1982), for example, indicate that a weight uncertainty of 2% could induce 8.8m of depth error at 750 m. GR10 finds significant weight variations for probes manufactured after 1992 and there are strong suggestions based on dedicated comparisons done during cruises that the characteristics of the probes have changed in time (Hanawa et al. (1995) mostly late 1990-early 1990 data compared to di Nezio and Goni (2010) or Reverdin et al (2009) for early 2000s data) and between manufacturers (for example, Kizu et al., 2011). Different equations might also have been used to report the XBT profiles in the data bases, adding confusion on the accuracy of the profiles.

There are also subtle issues of temperature biases (associated with the probes, the electronics, circuitry, A/C converter…) (Roemmich and Cornuelle, 1987, for example), but that have not been so well documented (Reseghetti et al., 2007; Reverdin et al., 2009). Usually, these biases were shown to have little dependence on depth, although some systems have been known to result in a large bowing of the profiles at depth, thus a depth dependent bias. Furthermore, it is possible that other errors are left in the data base, even after quality control, for example erroneously warm portions of the profiles after the wire has touched the hull or when it is otherwise stretched (see summary in GR10), both happening more commonly near the end of the profiles and that could result in average depth-dependent temperature biases.

These different issues with the data of XBTs in the data bases explain why a large variety of approaches have been used to address the data biases, since it has been known that they contributed to anomalous low frequency variability (for example, the artificial ‘global’ heat content increase of the 1970s or the recent problems identified in Willis et al. (2009)).

Gouretski and Koltermann (2007) used an ocean climatology based on high quality data (Conductivity Temperature Depth (CTD) and Nansen casts) to identify biases in XBT observations. They found a positive bias by 0.2-0.4°C on average with some variations from year to year. Based on this study and further comparisons between data types, Wijffels et al (2008; later on W08) proposed a yearly multiplicative correction factor on the depth (hereafter W08). More recently, Levitus et al (2009) used a simpler temperature correction, subtracting to all XBTs, the annual median temperature bias obtained by comparing with the CTD climatology. Ishii and Kimoto (2009) estimated a new fall rate equation for each year, separating different kinds of XBTs. GR10 proposed a new correction for profiles in the World Ocean Database 2005 (WOD2005) using an average depth correction (but latitude dependent) added to a time-dependent temperature offset (hereafter GR10). Other interesting examples of statistical corrections of the temperature profiles data in recent years can be found in di Nezio and Goni (2010), based on comparison with Argo float data that combine temperature biases and changes in the fall equation involving both offsets and a change in the quadratic fall rate equation.

These different approaches result in fairly comparable reconstruction of vertically integrated heat content variability from WOD2005 data. On the other hand, because of very different profiles of corrections, they might differ in the vertical structure of the changes. Thus, there is still a need for further studies of the biases in the historical data sets, so that the different types of data can be combined more optimally. Here, we provide a comparison of the different corrections, and provide an alternative correction. The choice we made was of applying a time-dependent, but depth-independent temperature correction based on comparisons in the near surface layer, and then to correct the depths at which the data are reported with a time-dependent correction. This means two steps:

1. Correction of thermal bias:

$$T = T_{XBT} - T_{off} \qquad (1)$$

2. Correction of the depth bias:

$$Z = Z_{obs}(1 - A - BZ_{obs}) - Z_{off} \qquad (2)$$

A major difference with GR10 is that the two steps are done here independently, and thus, to provide a significant reduction of the biases, the coefficients in equations 1 and 2 have to be time-dependent. On the other hand, the spatial dependency of the depth correction is more crudely taken into account in our approach.

We will in part 2 present the data and the collocation method; remind in part 3 of what W08 depth corrections imply in terms of residual biases. Then, in part 4, we will discuss the thermal correction, and compare it with GR10, and discuss the remaining residuals. In part 5, we will discuss the depth corrections, before presenting the resulting heat content time series with the corrections adopted (section 6), and the conclusions (section 7).

2. Data and collocation method

In the current study we used temperature profiles of the World Ocean Database 2005 (hereafter WOD2005) where profiles have been interpolated to standard levels. The ocean was subdivided into 16 vertical levels from the surface to 700m depth. We used profiles that have been processed when identification was possible using the correction H95 (Hanawa et al, 1995). Instead of using two climatologies, one constructed with CTD and bottles profiles, and the other with XBT profiles, we used a collocation method to compare instruments. For each XBT, we selected all CTD and OSD

(Ocean Station Data) geographically distant by less than 1° of latitude and 2° of longitude and a time lag less than 15 days. Then, we compute a reference profile as the median of all CTD and OSD profiles selected in the region of collocation. The bias profile was then calculated by subtracting this reference profile from the XBT profile. We found that many comparisons corresponded to situations with an XBT deployed over the deep ocean and CTD stations over the shelf or the continental slope. Thus to avoid potential biases resulting from cross-shelf fronts, we also ensure that ocean depth where the XBTs have been deployed does not differ by more than 500 m from where CTDs have been deployed (we discuss this added condition in section 4.1). Finally we rejected collocated XBT and CTD profiles for which the resulting vertically-averaged bias was more than 1°C. This method allows us to retain about 10^4 profiles per year between the year 1968 and 2007 (Fig. 1). Following Levitus et al (2009), the median rather than arithmetic average was used, as it reduces the influence of outliers.

After correction of a temperature bias (eq. 1) in the median annual bias, we adjust the depth indicated by the original fall rate (Eq. 2). We compute the depth bias at each standard level with the first order approximation,

$$dZ = (T_{CTD} - T_{XBT}) * dZ/dT_{CTD} \qquad (3)$$

Then, we linearly interpolate to retrieve the temperature at standard levels. Our calculations of depth bias from collocated profiles leads to several preliminary comments.

First, we estimate the depth correction in an iterative process repeated three times because of the subtle non-linearities, when estimating the correction: we first calculate the raw depth bias and the fitted correction using the local gradient of temperature observed by the CTD (Eq. 2). However when we apply the correction to XBTs, we use the uncorrected gradient observed by XBT,

$$T_{corr}=T+d\,T_{XBT}/d\,Z \qquad (4)$$

Thus, there is a residual depth bias related to the difference of gradient between the uncorrected XBT and the CTD. It is necessary to start again the calculation of the depth bias until convergence when the corrected XBTs gradients will be statistically similar to the one of the CTDs.

Furthermore, it is important to realize that this average bias (see summary on Fig. 1, 2) contributes only a small portion of the total variance in the individual XBT bias profiles, which is dominated by the time-space variability (only reaching 1-2% near 300-400m depth). However, uncertainties in the median bias profiles that are considered in this study are small enough due to the inclusion of a large number of individual bias profiles (Fig. 1).

Once the depth bias vertical profile is obtained, we estimate the three coefficients in the parabolic depth correction equation 2, by a linear least square method.

3. Test of the W08 correction

The W08 correction is a linear correction where the ``true" estimated XBT depth Z_{true} is computed from the depth Z given with the original fall rate.

$$Z_{true}=Z(1-A) \qquad (5)$$

W08 separated the deep XBT profiles (hereafter called XBTD) reaching a depth larger than 500 m (in standard levels) which are predominantly T7 or Deep Blue, and the others, shallow XBT profiles (hereafter called XBTS) which are predominantly T4/T6 instruments. According to their study, W08 note a depth error near 400 m of 10 m for XBTS and half that for XBTD on average.

We first applied the W08 correction to our collocated profiles. Fig. 1 shows the yearly raw and W08 corrected median bias averaged vertically as a function of year, and the average bias profile. According to Gouretski and Koltermann (2007) and Wijffels et al (2008) there is a positive bias between vertically averaged XBT temperature and high quality data like CTD and OSD. This median bias varies with the year of deployment of the XBT. It varies between 0.2°C and 0.1°C during the end of the 60's until the beginning of the 80's. Then the bias stabilizes around 0.05°C. Moreover this evolution agrees with the results of Levitus et al (2009). This vertically averaged bias is partially corrected by W08 corrections (Fig. 1).

The 1-year median bias function of depth is not uniformly reduced while applying the W08 correction (Fig. 2). Obviously the linear depth equation correction can not correct the surface bias (Fig. 1). Sometimes, it can also be too large and induce a negative bias at some depths (Fig. 2). The comparison thus suggests that a linear depth correction is not sufficient to properly reduce the observed biases (as also commented by GR10).

4. Temperature correction

Comparing neighboring XBT and CTD/OSD profiles in the upper mixed layer, we usually observe a positive thermal bias between 10m and 30m. Following GR10, we selected close-by profiles with a weak temperature gradient in this upper layer (less than 0.0025°C/m). This criterium guarantees that the observed bias is more likely to be related to a temperature error than to incorrect estimation of depth (this was also checked after the corrections on the depth of the XBT profiles of section 5, with little difference in the results). We also restricted the difference in ocean depth of collocated profiles to 500 m.

As we did not find significant differences in this temperature bias related to sea surface temperature (between warm and cold seas), we decided to take into account only two categories (XBTS and XBTD) to estimate more robustly the temperature bias T_{off} (Tab. 1). The thermal offset associated with XBTS is largely positive between 1968 and 1985 (0.096°C on average) and is close to 0 afterward (Fig. 3). The thermal bias of XBTD varies less. A first maximum is reached between 1970 and 1980 (0.076°C on average) which decreases during the 80's and becomes again maximum at the end of the 90's (0.086°C on average between 1995 and 2000).

XBTs are more often deployed in high seas whereas CTDs can be launched close to the coast. In continental slope regions, comparisons between XBT and CTD profiles can yield unrealistic biases. As shown in Fig. 3, this criterion has a large impact on the calculation of the thermal offset, especially for XBTD for the period 1985-2000. It appears that including CTDs deployed on continental shelves induced an artificial negative bias. The time history of the thermal corrections we find for XBTS and XBTD are rather different, something we don’t have an explanation for.

We find that the thermal bias observed with our collocation is quite different from the one given in GR10 for XBTD after 1982. In GR10, the thermal offset becomes largely negative whereas our observed bias is much more constant and positive on the whole study period. For XBTS, on the other hand the low frequency evolution we find presents some similarities with the one in GR10, with a maximum during the 70's, a decrease until the end of the 90's and a slight increase afterward. However, the correction for a given year can be quite different, and the total range of corrections here is only half the one in GR10.

Alltogether, we also find that just correcting a thermal bias reduces the average biases more efficiently than applying a linear depth correction scheme as in W08. However, it leaves a time varying vertically-averaged temperature bias, and thus is not appropriate for heat content variability estimations. Furthermore, the rms deviation of vertical variations in annual bias profile is the same as in the raw data. This is usually better than with W08 correction (Fig. 1), but is nonetheless still very large. Further discussion of the different ways to compare the data is provided in section 6.

5. The residual XBT bias

5.1 Temperature dependence of the depth correction

We will try to refine the model of bias correction by examining in more detail the vertical and spatial structure of the XBT/CTD biases, after removal of the thermal bias identified in 4.1. We will first comment on the sensitivity to the profile temperature, separating XBTS and XBTD as done earlier. There seems to be a relation between depth bias and the temperature of the sea water where the probe had been deployed (Thadathil et al, 2002). Fig. 4 shows the depth bias at 100m as a function of average temperature between 0 and 200m for XBTS (in red) and XBTD (in blue) averaged over the study period. We notice an increase of the bias toward low temperatures, without finding a significantly different behavior between the two classes of XBTs (XBTS and XBTD). Thus, Fig. 4 illustrates the need to process XBTs in categories of temperature, but at this particular depth, XBTS from XBTD are not clearly distinguished.

Whereas (Fig. 4) suggests that fall rate depends continuously on temperature (viscosity), retaining only two 0-200m averaged temperature classes is a practical first choice, as the computation of the depth bias requires a sufficient number of profiles to be robust (Fig.5, but notice the relatively small number of comparisons at the lower temperatures). In practice, the two categories in vertically-averaged temperature overlap to avoid discontinuities between profiles deployed in water close to 10°C: when computing the correction for the high temperature class, we selected all XBTs deployed in water warmer than 8°C, whereas for the low temperature class, we selected XBTs deployed in water colder than 12°C.

5.2 Parabolic nature of the depth correction

As in W08, the behavior of XBTS and XBTD is found to be different in the deeper part of the profile. This suggests that the collocated profiles are better corrected by a parabolic function than by a linear correction.

This parabolic character varies with years, geographical area and the type of XBT. We thus computed a second order correction with a least square fitting process for each year of deployment and each class of XBT (Tables 2 and 3). The depth bias also has a different behavior in the first meters of the probe fall. Between the surface and 30m, the error deviates from its parabolic behavior, possibly due to the high variability of surface temperature added to low vertical gradient in the surface mixed layer producing high variability in the calculated dZ quantities.

As suggested by earlier studies (see GR10), the depth correction equation is far from linear. Fig. 6 represents the linear part Az as a function of the parabolic part at 400m depth Bz^2 (corrections based on equ. 2) for XBTS (circles) and XBTD (stars) at high temperature (upper panel) and at low temperature (low). Each sector represents a different behavior of the yearly median depth bias. At 0-order the two terms nearly compensate at that depth, and seem often rather correlated between low and high temperatures. However, between 1968 and 1980, the behavior of XBTS is clearly different with a positive parabolic part at low temperatures, and a negative contribution at high temperatures. At high temperatures, the behaviors of XBTS and XBTD fall rate are very different. In particular, until 1980, for XBTS, the parabolic part is positive and the linear part is negative, whereas it is the opposite for XBTD. Between 1985 and 1990, the behavior of the fitted bias deviates from other periods, in particular at the lower temperatures with both linear and parabolic parts negative for

XBTD. These differences of behavior justify the need to process with classes.

5.3 Depth offset

Without introducing a depth offset, the resulting temperature bias after temperature correction and parabolic temperature correction is still positive in the surface layer. In this layer, the depth bias calculation involving the local gradient of temperature is not very accurate, but this also seems to correspond to a positive median depth- bias which is not easily modeled by a parabolic function, and thus best considered as a depth bias. The sources of the offset are very varied and the information that would be necessary to accurately model the fall rate/temperature in the upper layer is not available. To overcome this lack of information, we opt for an empirical fitting proportionally adjusting the offset with the pre-corrected bias by the parabolic function. We used the yearly median depth error between 30m and 100m to statically correct the depth offset error. We chose to compute the offset in this thin layer because it corresponds to a compromise between the choice of a surface layer, where the calculation of a depth bias is not influenced by the fall rate error and the layer of maximum gradient.

$$Z_{off}=\langle dZ\rangle_{30-100} \qquad (6)$$

We note that the depth offset (Tab. 2 and 3) is usually positive and of a few meters with slightly larger values in low temperature waters and around 2000. Those results are also consistent with Reverdin et al (2009) for the period 1999-2007 for a subset of XBT data deployed during French research cruises, and with the results of GR10.

5.4 Specific western Pacific case

After the global bias analysis by collocation, it is possible that residual biases may result from regional variations, as could happen in the Kuroshio or Gulf Stream region, due to sampling issues, or because of processing or regional differences in the systems used. Measurements close to Japan and in the western Pacific basin (the northwest Pacific region bounded by 180°E and north of 20°S) show after the corrections a strong negative bias during the period 1968/1985, in particular near 300 m (Fig. 7top), but less afterwards (Fig 7 middle). This negative bias has a vertical profile and time history (Fig 7 bottom), and implies that these XBTs are poorly corrected by the globally-derived parabolic term. We were not able to distinguish further the reason for this time-dependent regional anomaly (also commented in W08). Because this is a large-enough region, it has an impact on the global correction estimates, which we consider detrimental. Thus, we separated these regional profiles in another category, which increased the robustness of the correction estimated for the other data (Tab. 4). The coefficients A and B calculated for these particular XBTs are quite different from those calculated for the other classes, in particular for XBTD. The parabolic coefficient A of the XBTD correction is largely positive in the first years and decreases until 1985 and the linear coefficient B is strongly negative and increases with time. This behavior is specific to those regional XBTD. Note also that the depth offset is much smaller for those XBTs than in other regions of the world. Furthermore, the temperature offset (Tab. 1, where it is not estimated separately for XBTS and XBTD, as the two estimates were not statistically significant) is also much smaller than in other regions.

6. Implications of the correction approach on heat content

Figures 8 and 9 illustrate the raw median bias, the bias corrected with only the thermal correction (Eq. 1) and corrected with the depth+thermal correction (Eq. 1+ Eq. 2). Not only is the time-averaged residual bias profile rather small (within +-0.02°C at all depths), but its time variability show that the job is rather good at all times and depths with biases rarely exceeding 0.05°C. This is not due to overfitting, although in the last five years the number of collocated pairs starts been too small. This figure illustrates that we have removed average biases in the collocated XBT profiles, but our method selected only a small part of the entire database (about 10%). As we do not know if this sample is representative of the whole dataset, there is no guarantee that this can extrapolate to the remaining 90% profiles.

Following Wijffels et al (2008) and Levitus et al (2009), we also estimated a median depth bias on Mechanical Bathy-Thermograph (MBT). Using the same methodology, we performed a second order correction added to an offset. We also separated MBT deployed at high and low temperatures. For those probes, the selected threshold was 12°C for the median temperatures calculated between the surface and 100m depth.

With the globally corrected database, we map the observations on a latitude and longitude grid (4°x8°). The annual mean anomalies of temperature are obtained by subtracting the WOD2005 climatology (Locarnini et al., 2006). In a given year, we attributed to empty boxes the value of the annual mean anomaly of all full boxes for that year. Except after 2002 if Argo float profiles are included, the majority of boxes are empty (Fig. 10, lower panel), and if only XBTs are considered, there is less than 40% of boxes with data except in the 1990s (and a large portion of the empty boxes are in the southern hemisphere, in particular in the southern ocean).

The 0-700m integrated heat content (OHC) calculated from the corrected XBT database (green curve on Fig. 10, upper panel) presents large differences with OHC from the raw XBT database (red). This confirms that the local warming in the 70's was an artefact of the positive biases in XBT temperatures (as commented in Domingues et al, 2008; Ishii and Kimoto, 2009; Levitus et al, 2009; W08; GR10). Furthermore, we note that OHC calculated from corrected XBTs is very close to OHC calculated with the WOD2005 (excluding Argo profiler and mooring data) (such agreement is also found when considering specific layers like 0-400m or 400-700m), whereas there were large differences between OHC from the uncorrected XBTs and from the entire uncorrected data set. This indicates that our correction estimated with the subset of XBT profiles that could be collocated with the CTD/bottle casts holds for the entire database, at least in order to estimate vertically/spatially integrated variability.

The time series of OHC (since 1970) presents large similarities with the one from Levitus et al. (2009), and a little less so to OHC from Ishii and Kimoto (2009) or Domingues et al. (2008) (Fig. 10, middle panel) or with GR10. This is attained despite large differences in how the data are corrected, but also how they are combined to provide an ocean heat content time series, as the results are quite sensitive to insufficient sampling. For example, the large difference after 2002 between the blue curve on the upper panel with the one of the middle panel is related to the incorporation for the later of the much better sampled Argo data. The new correction results in a linear trend of OHC of 0.39 10^{22}J/year between 1970 and 2008 (0.44 10^{22}J/year, without the Argo and mooring data for recent years). These are rather different from the trends in OHC for WOD2005 without correction (0.48 10^{22}J/year). This is larger than for Ishii and Kimoto (2009) OHC, but less than in Domingues et al (2008) OHC. Of course, the estimated OHC is strongly dependent on the assumption to fill missing boxes with the annual anomaly for that year, as much southern hemisphere boxes were not sampled in the early periods. There are larger differences in OHC variability before 1975 with the other estimates, but data coverage starts to be really insufficient for that period.

7. Conclusions

We considered 6 different XBT classes to compute a globally second order correction on depth. We chose to separate XBTS and XBTD mostly related to T4 and T7 during the study period. We also separated XBTs deployed in cold or warm water (colder or warmer than 10°C on average between the surface and 200m) due to the dependence on temperature of the behavior of the XBT fall rate. A parabolic correction was not sufficient, and it was necessary to apply offsets: One thermal offset only depending on the XBT type applied to the temperature profiles and a depth offset. Both are estimated in the upper layer. We also found that the results are sensitive to the choice of maximum difference of oceanic depth between collocated profiles or on criteria on minimal thermal gradients (here the thermal offset is estimated with a thermal gradient lower than 0.0025°C/m, whereas the upper limit is 0.005°C/m in Gouretski and Reseghetti, 2010). Although our goal was to produce global estimates, large residual biases induced us to treat separately XBTs launched in the western Pacific basin between 1968 and 1985. This specific situation has also been discussed in W08 based on the depth error at 400m.

This particular correction, but also some of the time variations of the different corrections applied illustrate that XBT dataset in WOD2005 present basically unexplained biases that need to be corrected empirically. The method used for these corrections, unfortunately, influences the vertical or spatial structure of the low-frequency variability portrayed by these data. What we propose is a set of corrections among other possible corrections. The separation in only two temperature categories for the depth correction equation is obviously an oversimplification, as illustrated in GR10. However, there are not enough collocated data to further investigate this temperature dependency for the depth correction. An alternative would be to estimate the average temperature dependence as done in GR10, but still accept a time variability of the depth correction. We also are aware of documented biases that affect specific types of probes (Kizu et al., 2011, for example) and that we did not take into account. Clearly, the corrected dataset will thus retain spatially/time varying biases.

Although the mode of correction of the data strongly differs between different published studies, they exhibit similarities in the portrayed vertically-integrated heat content variability. However, because of the differences in the methods applied, we expect larger differences when considering the vertical structure of the variability. For example, W08 do not use a thermal offset, and thus there is little correction near the surface. GR10 do not use time-varying depth correction equation, and thus their estimates of thermal offset are at times rather different from ours, in particular for deep-reaching XBTs.

We corrected the MBT database with the same methodology to obtain an entire corrected database. We were able to compute a revised 0-700m spatially-integrated heat content (OHC) and a corresponding new estimate of its linear trend over time. These calculations support the result of other recent papers that the anomalous increase of OHC during the 70's originated from uncorrected XBT biases. The spatially-integrated results are very sensitive to insufficient sampling (in particular in the southern hemisphere), as illustrated by the change in integrated ocean heat content (0-700m) after 2002 when adding the ARGO float data. Thus, the different mapping methods used to estimate OHC by the different authors certainly result in large differences in OHC time variability (see, for example, the discussion based on synthetic data in Lyman and Johnson, 2008).

Acknowledgements

We thank Karina Von Schuckmann, Clement de Boyer Montegut and Cécile Cabanes for useful comments and discussions during the preparation of this manuscript. Mathieu Hamon work is funded by IFREMER and Meteo France as part of a Coriolis Phd Grant. Coriolis is a joint infrastructure of IFREMER, INSU, IRD and SHOM. The research leading to these results has received funding from the European Community's Seventh Framework Program FP7/2007-2013 under grant agreement n° 218812 (MyOcean).

References

Di Nezio, P.N. and G.J. Goni, 2010: Identifying and estimating biases between XBT and Argo observations using satellite altimetry. *J. Atmos. Ocean. Technol*., **27**, 226-240.

Domingues C. M., J. A. Church, N. J. White, P. J. Gleckler, S. E. Wijffels, P. M Barker. and J. R. Dunn, 2008: Improved estimates of upper-ocean warming and multi-decadal sea level rise. *Nature*, **453**, 1090-1093.doi:10.1038/nature07080

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

Gouretski V. 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**, 812-833.

Hanawa K., P Rual, R. Bailey. A Sy and M. Szabados, 1995: A new depth-time equation for Sippican or TSK T-7, T-6 and T-4 expendable bathythermographs (XBT). *Deep Sea Res. I*, **42**, 1423-1451.

Ishii M., and M. Kimoto, 2009: Reevaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections. *J. Oceanogr*., **65**, 287-299.doi: 10.1007/s10872-009-0027-7.

Kizu, S., C. Sukigara, and K. Hanawa, 2011: Comparison of the fall rate and structure of recent T-7 XBT manufactured by Sippican and TSK. *Ocean Science*, **7**, 231-244.

Levitus, S;, H.I. Antonov, J. Wang, T.L. Delworth, K.W. Dixon, and A.J. Broccoli, 2001: Anthropogenci warming of Earth’s climate system. *Science*, **21**, 292, 265-270, doi:10.1126/science.1058154.

Levitus S., J Antonov, and T. Boyer, 2005: Warming of the world ocean, 1955-2003. *Geophys. Res. Lett.*, **32**, L02604. doi: 10.1029/2004GL021592

Levitus S., J Antonov, T. Boyer, R.A. Locamini, H.E. Garcia and A.V. Mishonov, 2009: Global ocean heat content 1955-2008 in light of recently revealed instrumentation problems. *Geophys. Res. Lett.*, **36**, L07608. doi: 10.1029/2008GL037155.

Locarnini, R.A., A.V. Mishonov, J.I. Antonov, T.P. Boyer, and H.E. Garcia, 2006: World Ocean Atlas 2005, Volume 1: Temperature. S. Levitus, Ed. NOAA Atlas NESDIS 61, U.S. Government Printing Office, Washington D.C., 182 pp.

Lyman, J.M., and G.C. Johnson, 2008: Estimating annual global upper-ocean heat content anomalies despite irregular in situ ocean sampling. *J. Clim*., **21**, 5629-5641.

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, 2007.

Reverdin G., F. Marin, B. Bourles, and P. L'Herminier, 2009: XBT temperature errors during French research cruises (1999-2009). *J. of Atmos. and Oceanic Technol*., **26**, 2462-2473.

Seaver G. A., and S. Kuleshov, 1982: Experimental and Analytical Error of the Expendable Bathythermograph. *J. Phys. Oceanogr*., **12**, 592-600.

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.

Wijffels, Susan 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

Willis, J.,K., J.M. Lyman , G.C. Johnson, and J. Gilson, 2009: In situ data biases and recent ocean heat content variability. *J. Atmos. And Ocean. Tech*., **26**, 846-852. Doi:10.1175/2008JTECHO608.1.

Figure captions

Figure 1: XBT-CTD 0/700m median raw bias (blue) and corrected W08 (green). Upper panel, vertical integral between 0 and 700m (average value (curve) with vertical standard deviation (bars:)). The number of yearly collocated pairs is indicated with the black dotted line. Lower panel: vertical profiles of median raw bias and of median bias after correction by W08 averaged over the study period (average curve and colored range within one standard deviation).

Figure 2: Evolution of XBT-CTD median raw bias (above) and corrected by W08 (under) as a function of depth and time.

Figure 3: Median XBT-CTD thermal bias for XBTS (left) and XBTD (right). Thermal bias observed from collocated data within bathymetry criterion is represented by the red line, with the bathymetry criterion, by the green line and the thermal offset of GR10, by the black line.

Figure 4: Median XBT-CTD depth bias at 100m depth as a function of the integrated temperature between 0 and 200m depth for XBTS (red) and XBTD (blue). The width provides the range of the +-one standard deviation of the uncertainty (standard deviation divided by the square root of the number of selected pairs).

Figure 5: Number of collocated pairs for the four classes of XBT function of time. XBTS(D)H/XBTS(D)L correspond to shallow (deep) XBTs deployed in high and low temperatures.

Figure 6: Linear part (coefficient B(t)) an a function of parabolic part (A(t)x 400) at 400 m depth for XBTS (large dots) and XBTD (stars). The upper panel includes XBTs deployed at high temperatures and the lower panel, XBTs deployed at low temperatures. Years are indicated with the color bar.

Figure 7: Residual average bias in XBT-CTD temperature at 300m between 1968 and 1985 (top panel), and between 1985 and 2007 (middle panel); (lower panel) Evolution of XBT-CTD median bias in the western Pacific (the region is bounded by 180°E and 20°S). This corresponds to XBT data corrected by a global parabolic correction, as a function of depth and time.

Figure 8: Upper panel with profiles of median raw bias (green), corrected by the thermal offset (blue) and corrected by a parabolic correction added an offset (red) as a function of depth averaged over the study period 1968/2007 (curve (average) with the color band corresponding to the one standard deviation range of the time series variability at each depth). Lower panel: the time series of vertically averaged bias (curve (average) with vertical bars (range of one vertical standard deviation)).

Figure 9: Evolution of XBT-CTD median raw bias (upper), corrected by the thermal offset (middle), and corrected by a parabolic correction added to an offset (lower) as a function of depth and time.

Figure 10: Ocean heat content (OHC) integrated between the surface and 700m depth (upper panel) calculated using the entire raw dataset (red), the entire corrected dataset excluding Argo floats (blue), and only using raw XBTs (black) and corrected XBTs (green) (the average seasonal cycle from Locaranini et al. (2006) was removed to the gridded data). Middle panel: OHC time series from different studies, including one for the corrected entire dataset including Argo floats (blue); the time series are reported relative to their respective time average. Lower panel: percentage of the oceanic volume covered by 4°x8° boxes including XBT data (dotted line), all WOD2005 data (full black line) or WOD2005 data with the exclusion of ARGO and mooring data (full red line).

Table 1: Thermal offset (°C) as a function of time for XBTS and XBTD, and western Pacific XBTs

(north of 20°S and west of 180°E).

Table 2: Coefficients of the parabolic correction and the depth offset for XBTD deployed in low (DL) and high (DH) temperature waters.

Table 3: Coefficients of the parabolic correction and the depth offset for XBTS deployed in low (SL) and high (SH) temperature waters.

Table 4: Coefficients of the parabolic correction and the depth offset for XBTD (DWP) and XBTS (SWP) deployed in the western Pacific Ocean.

Table 1: Thermal offset (°C) as a function of time for XBTS and XBTD, and western Pacific XBTs (north of 20°S and west of 180°E).

Year	ToffsetD	ToffsetS	ToffsetWP
1968	0.049	0.084	-0.006
1969	0.052	0.083	0.002
1970	0.051	0.080	0.020
1971	0.067	0.084	0.036
1972	0.081	0.087	0.043
1973	0.090	0.087	0.049
1974	0.098	0.091	0.055
1975	0.096	0.098	0.056
1976	0.089	0.105	0.051
1977	0.079	0.112	0.051
1978	0.070	0.118	0.055
1979	0.059	0.121	0.050
1980	0.059	0.119	0.047
1981	0.061	0.109	0.039
1982	0.062	0.099	0.031
1983	0.054	0.089	0.019
1984	0.056	0.082	0.014
1985	0.056	0.077	0.004
1986	0.053	0.077	
1987	0.049	0.073	
1988	0.059	0.062	
1989	0.056	0.049	
1990	0.049	0.034	
1991	0.047	0.021	
1992	0.051	0.011	
1993	0.051	0.009	
1994	0.059	0.011	
1995	0.074	0.015	
1996	0.087	0.014	
1997	0.090	0.011	
1998	0.093	0.005	
1999	0.093	0.000	
2000	0.082	-0.014	
2001	0.066	-0.036	
2002	0.058	-0.020	
2003	0.048	0.006	
2004	0.037	0.018	
2005	0.031	0.027	
2006	0.032	0.043	
2007	0.029	0.008	

Table 2: Coefficients of the parabolic correction and the depth offset for XBTD deployed in low (DL) and high (DH) temperature waters.

Year	**A**$_{DL}$	**B**$_{DL}$	**Offset**$_{DL}$	**A**$_{DH}$	**B**$_{DH}$	**Offset**$_{DH}$
1968	-0.000078	0.052	1.0	0.000006	-0.054	0.9
1969	-0.000016	0.024	1.5	0.000046	-0.039	0.9
1970	0.000040	-0.006	1.3	0.000066	-0.025	0.9
1971	0.000038	-0.014	1.7	0.000039	-0.010	0.8
1972	-0.000004	-0.008	2.1	-0.000006	0.015	0.6
1973	-0.000082	0.021	2.4	-0.000021	0.023	0.5
1974	-0.000151	0.062	2.7	0.000003	0.008	0.5
1975	-0.000137	0.075	2.9	0.000016	0.001	0.6
1976	-0.000049	0.052	2.9	-0.000015	0.021	0.4
1977	0.000037	0.018	2.5	-0.000057	0.048	0.1
1978	0.000070	-0.004	1.8	-0.000054	0.046	0.0
1979	0.000062	-0.004	1.3	-0.000021	0.029	0.0
1980	0.000025	0.026	1.3	0.000002	0.019	0.1
1981	0.000019	0.026	1.3	-0.000004	0.017	0.2
1982	0.000102	-0.070	0.9	-0.000026	0.019	0.4
1983	0.000167	-0.156	0.4	0.000013	-0.009	0.6
1984	0.000118	-0.145	0.6	0.000071	-0.048	0.7
1985	0.000006	-0.084	1.4	0.000054	-0.039	0.9
1986	-0.000053	-0.037	2.3	0.000015	-0.012	1.2
1987	-0.000033	-0.021	2.9	0.000003	-0.005	1.6
1988	-0.000003	-0.018	3.1	0.000014	-0.015	1.9
1989	0.000014	-0.025	2.8	0.000022	-0.022	1.8
1990	0.000016	-0.027	2.4	0.000011	-0.010	1.6
1991	0.000025	-0.031	2.2	0.000012	-0.008	1.4
1992	0.000025	-0.026	2.0	0.000016	-0.006	1.0
1993	0.000001	-0.003	1.8	0.000004	0.006	0.6
1994	0.000003	0.001	1.7	0.000000	0.006	0.4
1995	0.000038	-0.030	2.0	0.000007	-0.006	0.4
1996	0.000066	-0.068	2.5	0.000006	-0.009	0.6
1997	0.000040	-0.073	3.2	-0.000006	-0.001	0.9
1998	-0.000012	-0.063	3.9	-0.000000	-0.014	1.5
1999	-0.000004	-0.080	4.6	0.000026	-0.041	1.9
2000	0.000043	-0.107	4.8	0.000023	-0.033	1.8
2001	0.000091	-0.131	4.2	-0.000008	0.002	1.3
2002	0.000110	-0.132	3.2	-0.000014	0.013	0.8
2003	0.000060	-0.073	1.8	0.000002	0.004	0.6
2004	-0.000031	0.010	0.6	0.000001	0.001	0.4
2005	-0.000109	0.059	0.1	-0.000022	0.011	0.4
2006	-0.000138	0.071	0.6	-0.000041	0.029	0.5
2007	-0.000011	0.012	2.2	-0.000032	0.033	0.7

Table 3: Coefficients of the parabolic correction and the depth offset for XBTS deployed in low (SL) and high (SH) temperature waters.

Year	A_{SL}	B_{SL}	$Offset_{SL}$	A_{SH}	B_{SH}	$Offset_{SH}$
1968	-0.000109	-0.018	3.4	0.000100	-0.033	2.4
1969	0.000024	-0.079	3.0	0.000155	-0.072	2.0
1970	0.000014	-0.053	3.0	0.000146	-0.070	2.1
1971	-0.000080	0.012	2.3	0.000047	-0.020	1.5
1972	-0.000165	0.073	1.7	-0.000032	0.027	1.0
1973	-0.000147	0.089	1.5	-0.000012	0.029	0.9
1974	-0.000069	0.071	1.6	0.000062	-0.006	1.2
1975	-0.000069	0.074	1.4	0.000078	-0.021	1.4
1976	-0.000137	0.096	1.0	0.000027	-0.003	1.3
1977	-0.000198	0.120	0.6	0.000004	0.015	0.9
1978	-0.000216	0.123	0.4	0.000034	0.005	0.6
1979	-0.000216	0.110	0.5	0.000051	-0.008	0.7
1980	-0.000222	0.097	0.7	0.000063	-0.014	0.9
1981	-0.000235	0.083	1.0	0.000097	-0.030	1.3
1982	-0.000250	0.071	1.1	0.000114	-0.041	1.6
1983	-0.000264	0.059	1.2	0.000099	-0.041	1.7
1984	-0.000245	0.041	1.3	0.000072	-0.034	1.8
1985	-0.000158	0.019	1.7	0.000054	-0.025	1.7
1986	-0.000039	-0.014	2.2	0.000056	-0.032	1.6
1987	0.000047	-0.042	2.6	0.000038	-0.029	1.5
1988	0.000067	-0.046	2.7	0.000013	-0.019	1.3
1989	0.000036	-0.037	2.4	0.000006	-0.012	1.0
1990	0.000010	-0.034	2.2	0.000009	-0.006	0.8
1991	-0.000002	-0.033	2.2	-0.000003	0.003	0.6
1992	0.000003	-0.035	2.4	-0.000018	0.012	0.5
1993	0.000023	-0.041	2.5	0.000022	-0.002	0.5
1994	0.000009-	0.036	2.3	0.000069	-0.020	0.7
1995	-0.000049	-0.009	2.0	0.000035	-0.006	0.8
1996	-0.000083	0.015	1.8	-0.000056	0.023	0.8
1997	-0.000009	-0.003	2.1	-0.000111	0.028	0.8
1998	0.000141	-0.051	2.8	-0.000082	0.003	1.1
1999	0.000263	-0.089	3.5	-0.000023	-0.016	1.3
2000	0.000295	-0.093	3.7	-0.000004	-0.007	1.3
2001	0.000210	-0.063	3.4	-0.000005	0.003	1.3
2002	0.000044	-0.008	3.1	0.000061	-0.036	1.9
2003	0.000020	0.026	3.0	0.000154	-0.092	2.6
2004	0.000113	0.014	2.7	0.000184	-0.085	2.4
2005	0.000030	0.027	2.0	0.000083	-0.006	1.0
2006	-0.000138	0.070	0.9	-0.000102	0.082	-0.1
2007	-0.000304	0.128	-0.1	-0.000236	0.123	-0.3

Year	A_{DWP}	B_{DWP}	$Offset_{DWP}$	A_{SWP}	B_{SWP}	$Offset_{SWP}$
1968	0.000430	-0.113	0.0	0	0	0
1969	0.000320	-0.090	0.1	0	0	0
1970	0.000159	-0.055	0.1	0	0	0
1971	0.000078	-0.035	0.1	0	0	0
1972	0.000052	-0.019	-0.0	0.000021	-0.046	-0.1
1973	0.000047	-0.011	0.0	-0.000023	-0.013	0.3
1974	0.000080	-0.025	0.1	0.000034	-0.018	0.2
1975	0.000091	-0.024	0.1	0.000097	-0.015	0.2
1976	0.000077	-0.013	0.2	0.000101	-0.019	0.3
1977	0.000032	-0.001	0.5	0.000039	-0.019	0.6
1978	0.000015	-0.011	0.9	0.000036	-0.037	0.9
1979	0.000048	-0.029	1.2	0.000078	-0.054	1.0
1980	0.000052	-0.027	1.2	0.000026	-0.028	0.9
1981	0.000046	-0.025	1.2	0.000024	-0.015	0.8
1982	0.000055	-0.037	1.4	0.000066	-0.017	0.8
1983	0.000083	-0.059	1.5	0.000041	-0.008	0.7
1984	0.000111	-0.070	1.4	0.000022	-0.003	0.2
1985	0.000103	-0.056	1.1	-0.000004	0.009-	0.7

Table 4: Coefficients of the parabolic correction and the depth offset for XBTD (DWP) and XBTS (SWP) deployed in the western Pacific Ocean.

Figure 1: XBT-CTD 0/700m median raw bias (blue) and corrected W08 (green). Upper panel, vertical integral between 0 and 700m (average value (curve) with vertical standard deviation (bars:)). The number of yearly collocated pairs is indicated with the black dotted line. Lower panel: vertical profiles of median raw bias and of median bias after correction by W08 averaged over the study period (average curve and colored range within one standard deviation).

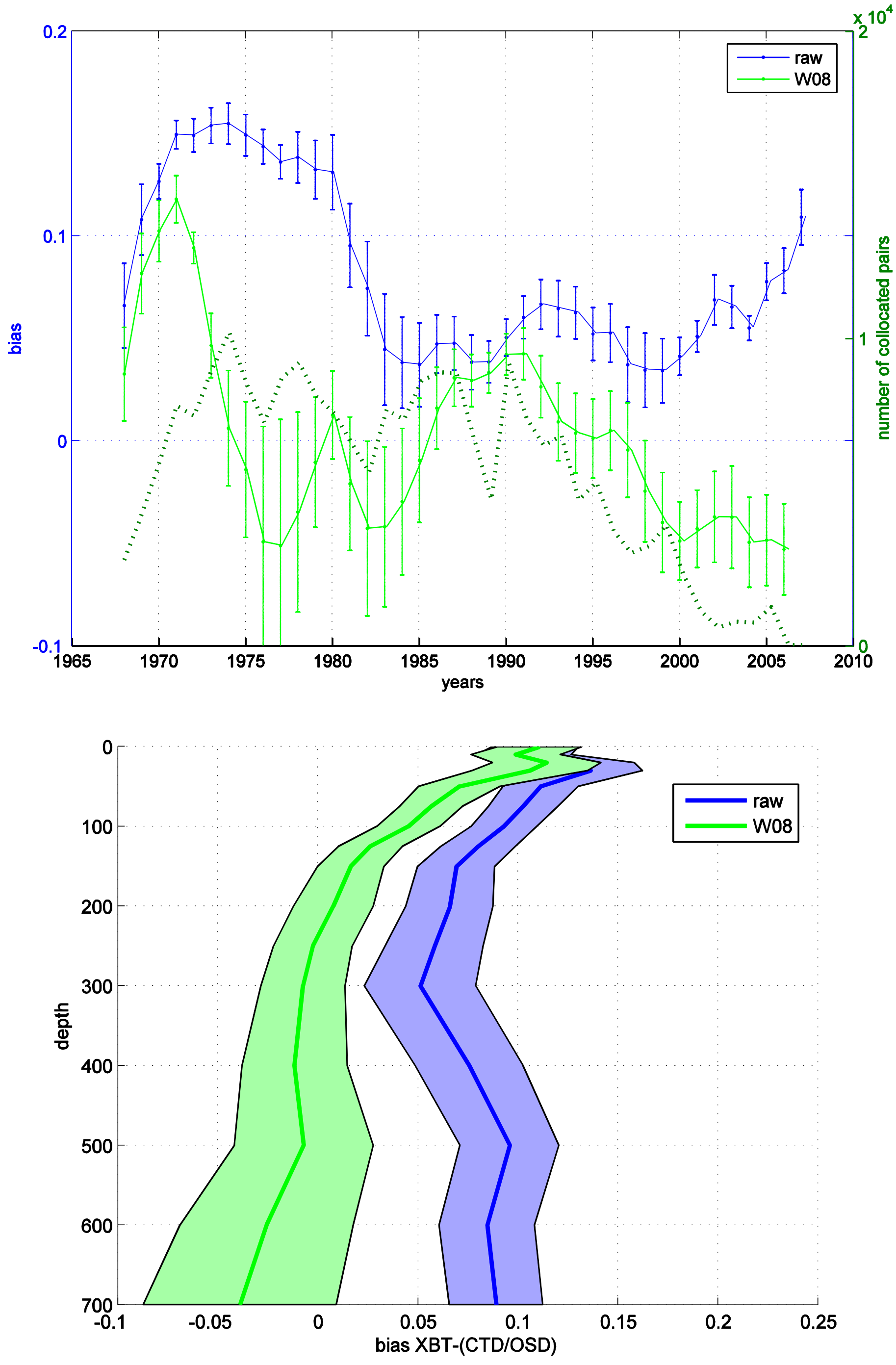


Figure 2: Evolution of XBT-CTD median raw bias (above) and corrected by W08 (under) as a function of depth and time.

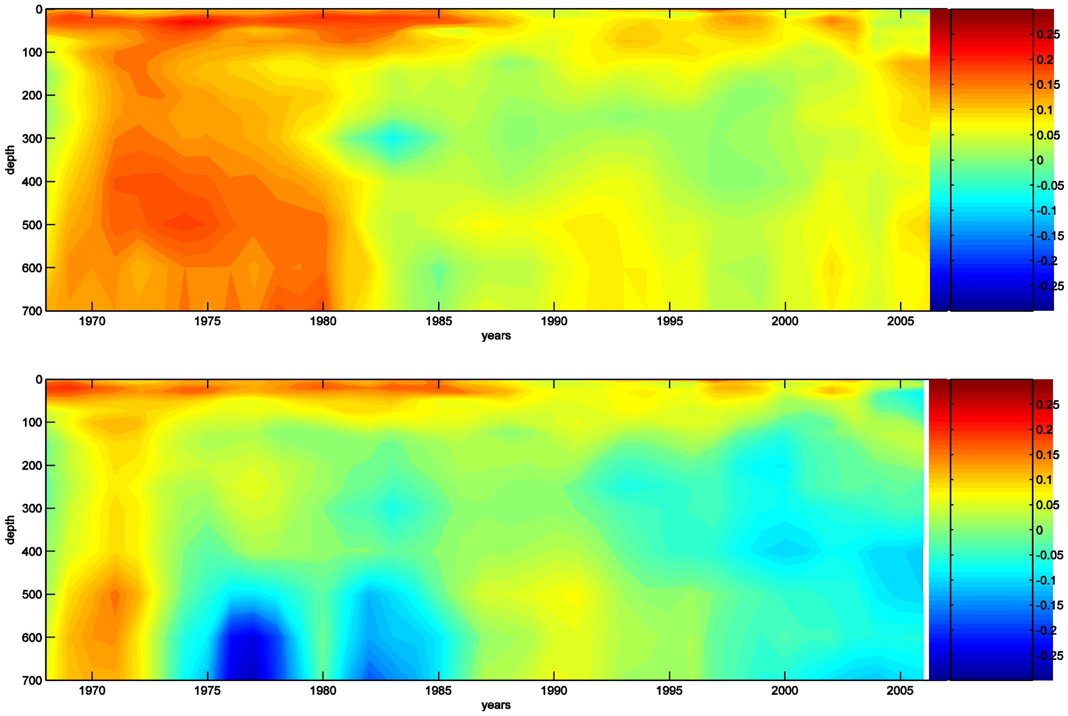
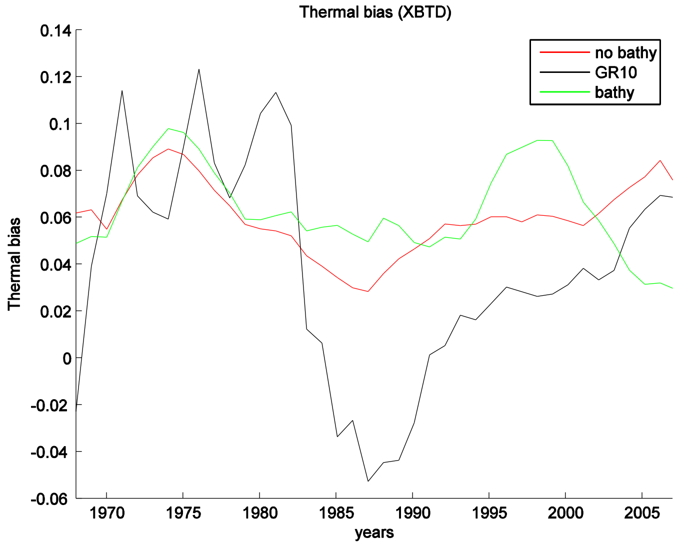


Figure 3: Median XBT-CTD thermal bias for XBTS (left) and XBTD (right). Thermal bias observed from collocated data within bathymetry criterion is represented by the red line, with the bathymetry criterion, by the green line and the thermal offset of GR10, by the black line.



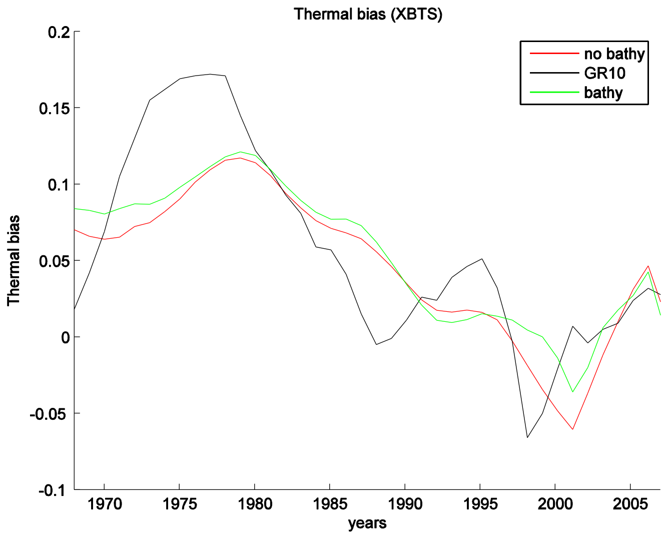


Figure 4: Median XBT-CTD depth bias at 100m depth as a function of the integrated temperature between 0 and 200m depth for XBTS (red) and XBTD (blue). The width provides the range of the +-one standard deviation of the uncertainty (standard deviation divided by the square root of the number of selected pairs).

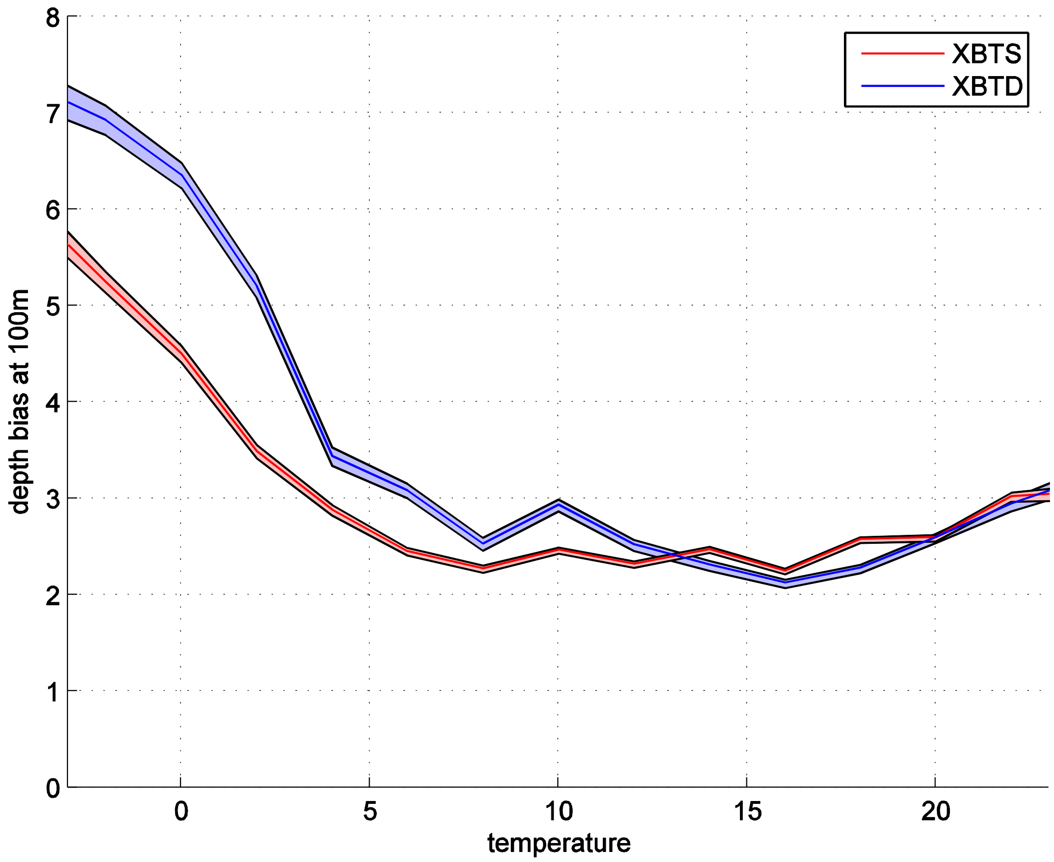


Figure 5: Number of collocated pairs for the four classes of XBT function of time. XBTS(D)H/XBTS(D)L correspond to shallow (deep) XBTs deployed in high and low temperatures.

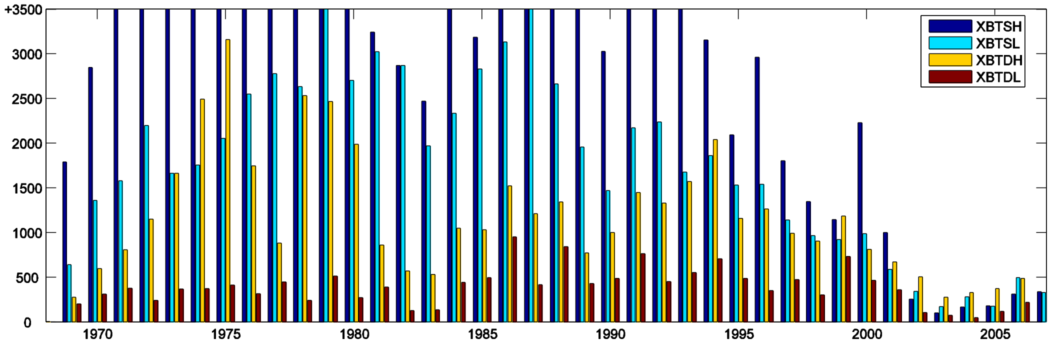


Figure 6: Linear part (coefficient B(t)) an a function of parabolic part (A(t)x 400) at 400 m depth for XBTS (large dots) and XBTD (stars). The upper panel includes XBTs deployed at high temperatures and the lower panel, XBTs deployed at low temperatures. Years are indicated with the color bar.

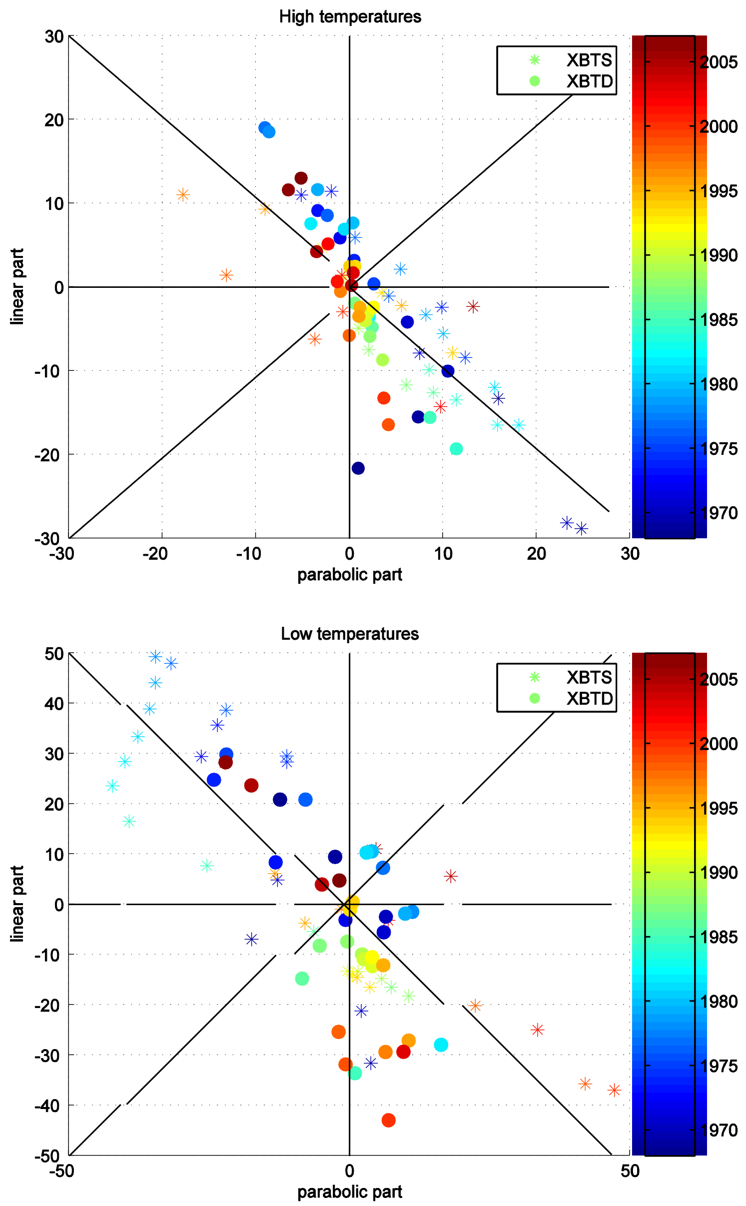


Figure 7: Residual average bias in XBT-CTD temperature at 300m between 1968 and 1985 (top panel), and between 1985 and 2007 (middle panel); (lower panel) Evolution of XBT-CTD median bias in the western Pacific (the region is bounded by 180°E and 20°S). This corresponds to XBT data corrected by a global parabolic correction, as a function of depth and time.

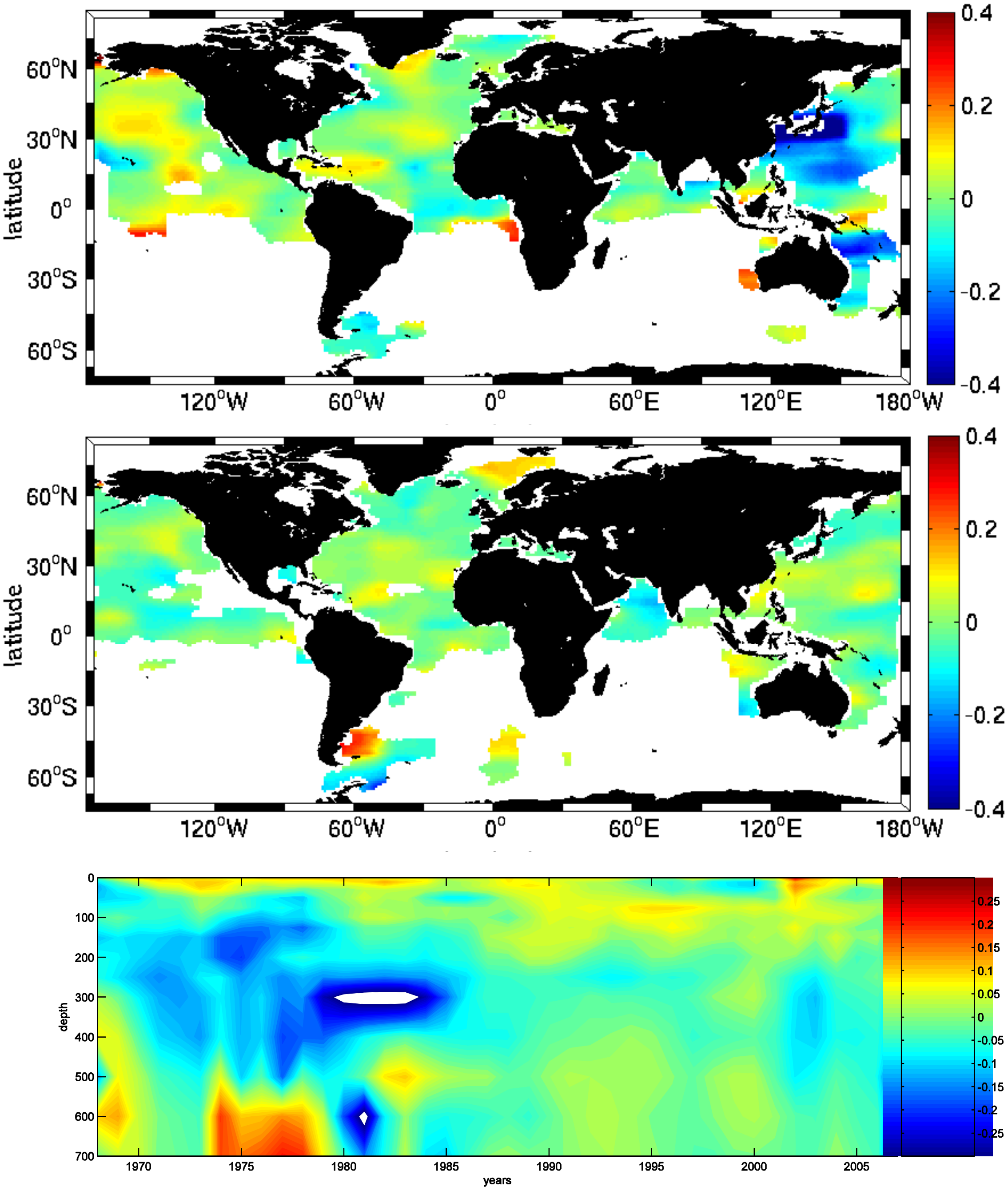


Figure 8: Upper panel with profiles of median raw bias (green), corrected by the thermal offset (blue) and corrected by a parabolic correction added an offset (red) as a function of depth averaged over the study period 1968/2007 (curve (average) with the color band corresponding to the one standard deviation range of the time series variability at each depth). Lower panel: the time series of vertically averaged bias (curve (average) with vertical bars (range of one vertical standard deviation)).

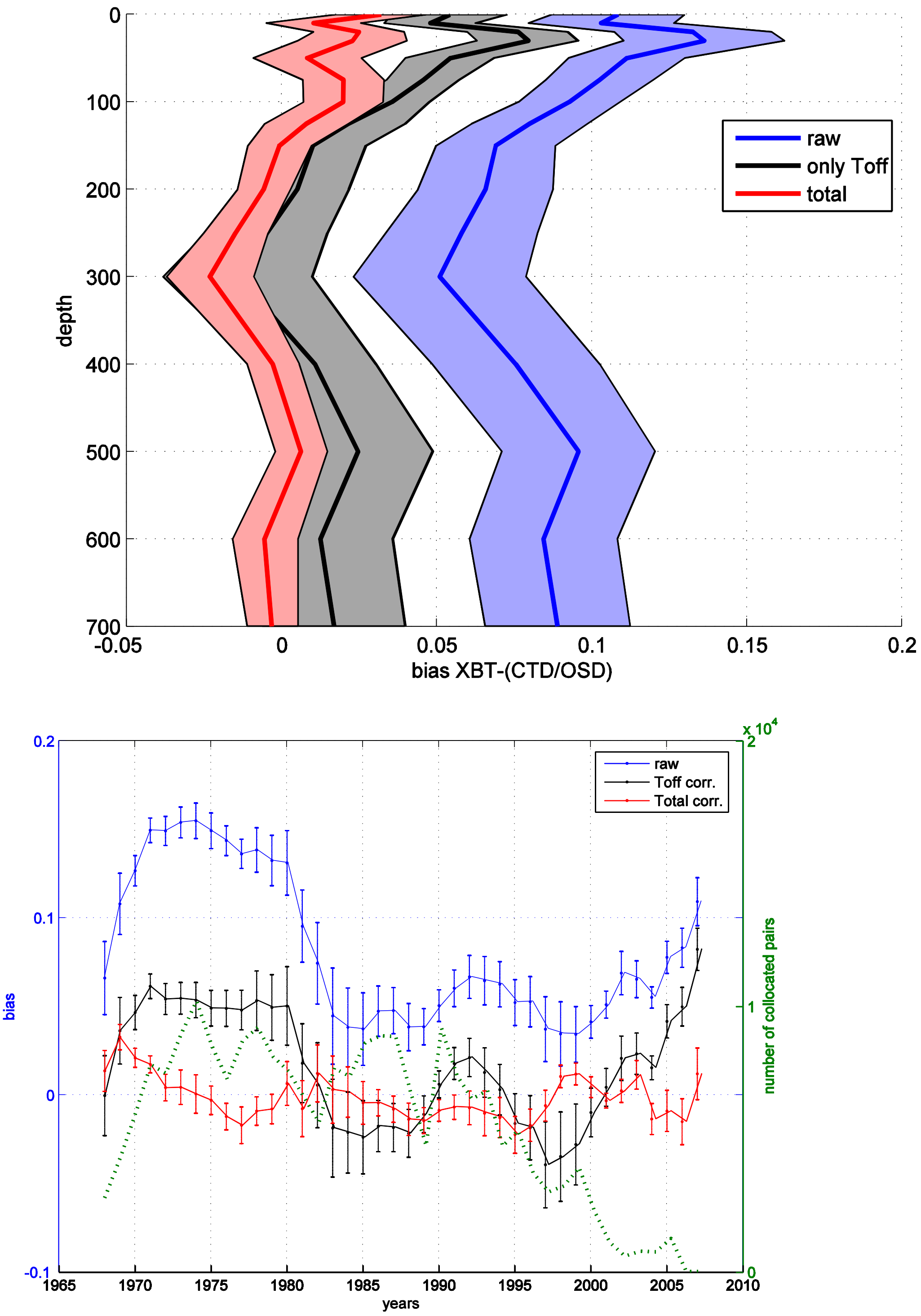


Figure 9: Evolution of XBT-CTD median raw bias (upper), corrected by the thermal offset (middle), and corrected by a parabolic correction added to an offset (lower) as a function of depth and time.

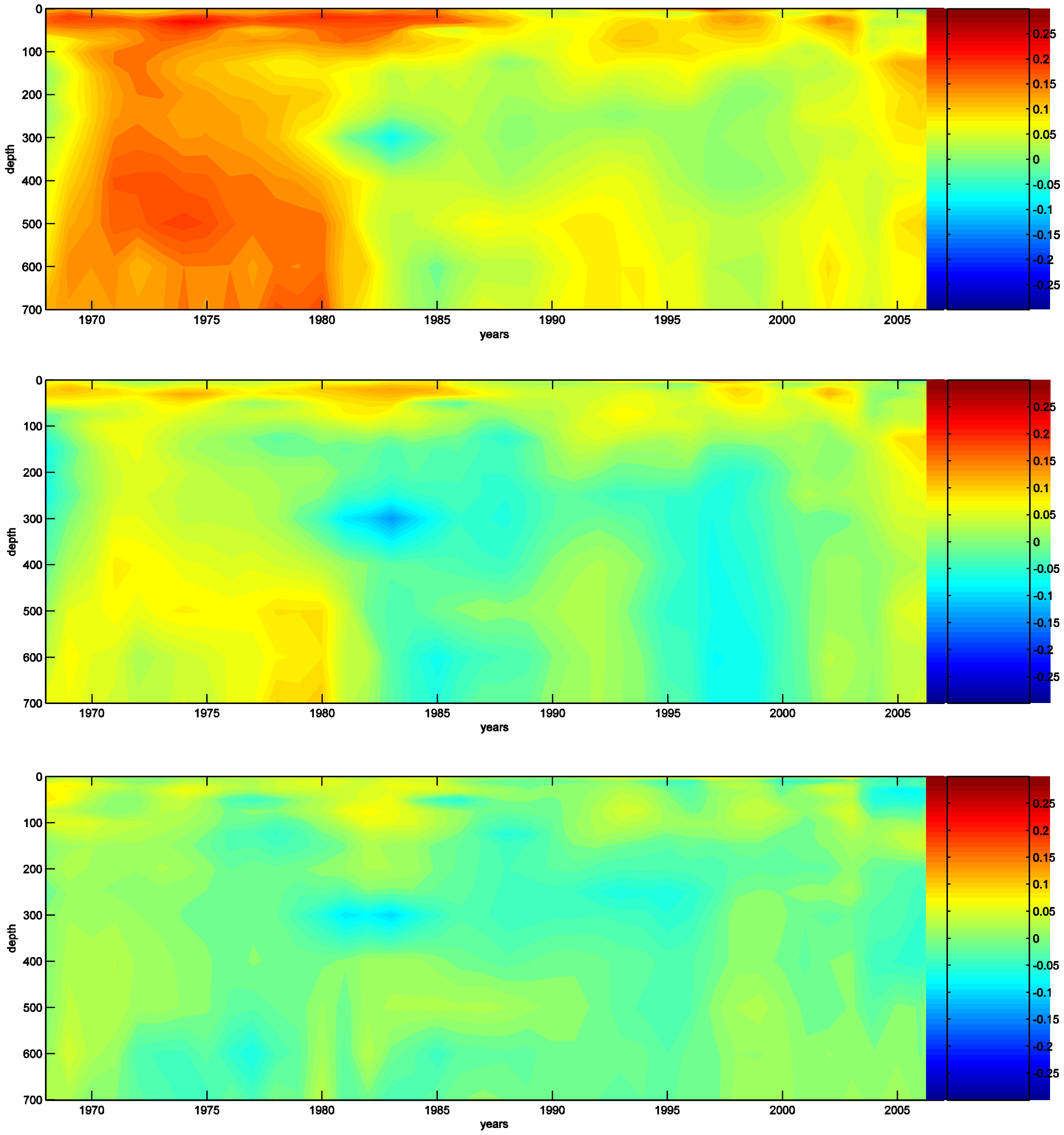


Figure 10: Ocean heat content (OHC) integrated between the surface and 700m depth (upper panel) calculated using the entire raw dataset (red), the entire corrected dataset excluding Argo floats (blue), and only using raw XBTs (black) and corrected XBTs (green) (the average seasonal cycle from Locarnini et al. (2006) was removed to the gridded data). Middle panel: OHC time series from different studies, including one for the corrected entire dataset including Argo floats (blue); the time series are reported relative to their respective time average. Lower panel: percentage of the oceanic volume covered by 4°x8° boxes including XBT data (dotted line), all WOD2005 data (full black line) or WOD2005 data with the exclusion of ARGO and mooring data (full red line).