

## Sensitivity of Global Upper-Ocean Heat Content Estimates to Mapping Methods, XBT Bias Corrections, and Baseline Climatologies\*

TIM BOYER,<sup>a</sup> CATIA M. DOMINGUES,<sup>b,c,d</sup> SIMON A. GOOD,<sup>e</sup> GREGORY C. JOHNSON,<sup>f</sup> JOHN M. LYMAN,<sup>f,g</sup> MASAYOSHI ISHII,<sup>h</sup> VIKTOR GOURETSKI,<sup>i</sup> JOSH K. WILLIS,<sup>j</sup> JOHN ANTONOV,<sup>k</sup> SUSAN WIJFFELS,<sup>l</sup> JOHN A. CHURCH,<sup>l</sup> REBECCA COWLEY,<sup>l</sup> AND NATHANIEL L. BINDOFF<sup>b,c,d,l</sup>

<sup>a</sup> NOAA/National Centers for Environmental Information, Silver Spring, Maryland

<sup>b</sup> Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia

<sup>c</sup> Antarctic Climate and Ecosystems Cooperative Research Institute, Hobart, Tasmania, Australia

<sup>d</sup> Centre of Excellence for Climate System Science, Australian Research Council, Hobart, Tasmania, Australia

<sup>e</sup> Met Office, Exeter, United Kingdom

<sup>f</sup> NOAA/Pacific Marine Environmental Laboratory, Seattle, Washington

<sup>g</sup> Joint Institute for Marine and Atmospheric Research, University of Hawai'i at Mānoa, Honolulu, Hawaii

<sup>h</sup> Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

<sup>i</sup> Center for Earth System Research and Sustainability, CliSAP, Integrated Climate Data Center, University of Hamburg, Hamburg, Germany

<sup>j</sup> Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

<sup>k</sup> University Corporation for Atmospheric Research, Boulder, Colorado

<sup>l</sup> Commonwealth Scientific and Industrial Research Organization, Hobart, Tasmania, Australia

(Manuscript received 9 November 2015, in final form 30 March 2016)

### ABSTRACT

Ocean warming accounts for the majority of the earth's recent energy imbalance. Historic ocean heat content (OHC) changes are important for understanding changing climate. Calculations of OHC anomalies (OHCA) from in situ measurements provide estimates of these changes. Uncertainties in OHCA estimates arise from calculating global fields from temporally and spatially irregular data (mapping method), instrument bias corrections, and the definitions of a baseline climatology from which anomalies are calculated. To investigate sensitivity of OHCA estimates for the upper 700 m to these different factors, the same quality-controlled dataset is used by seven groups and comparisons are made. Two time periods (1970–2008 and 1993–2008) are examined. Uncertainty due to the mapping method is 16.5 ZJ for 1970–2008 and 17.1 ZJ for 1993–2008 (1 ZJ =  $1 \times 10^{21}$  J). Uncertainty due to instrument bias correction varied from 8.0 to 17.9 ZJ for 1970–2008 and from 10.9 to 22.4 ZJ for 1993–2008, depending on mapping method. Uncertainty due to baseline mean varied from 3.5 to 14.5 ZJ for 1970–2008 and from 2.7 to 9.8 ZJ for 1993–2008, depending on mapping method and offsets. On average mapping method is the largest source of uncertainty. The linear trend varied from 1.3 to 5.0 ZJ yr<sup>-1</sup> (0.08–0.31 W m<sup>-2</sup>) for 1970–2008 and from 1.5 to 9.4 ZJ yr<sup>-1</sup> (0.09–0.58 W m<sup>-2</sup>) for 1993–2008, depending on method, instrument bias correction, and baseline mean. Despite these complications, a statistically robust upper-ocean warming was found in all cases for the full time period.

---

\*Pacific Marine Environmental Laboratory Contribution Number 4261 and Joint Institute for Marine and Atmospheric Research Contribution Number 16-393.

---

Corresponding author address: Tim Boyer, Ocean Climate Laboratory, National Centers for Environmental Information, 1315 East-West Highway, Silver Spring, MD 20910.  
E-mail: tim.boyer@noaa.gov

### 1. Introduction

Ocean heat content (OHC) is an essential climate variable for understanding interannual, decadal, and longer time-scale changes in the climate system and the earth's energy budget. The earth's energy budget is currently imbalanced (Wild et al. 2013 and references therein; Trenberth et al. 2014), with an estimated +0.5 to +1.0 W m<sup>-2</sup> net top-of-the-atmosphere (TOA) energy

flux over the first decade of the twenty-first century (Trenberth et al. 2014). Oceans cover more than 70% of the earth's surface. Their extensive interface with the atmosphere, combined with their high heat capacity, low albedo, and great depth, has resulted in ocean sequestration of more than 90% of the earth's excess heat over recent years (Roemmich et al. 2015) and recent decades (Bindoff et al. 2007; Rhein et al. 2013). This net buildup of excess heat in the ocean mitigates surface warming, but it increases ocean temperature, contributing to sea ice melt (Liu and Curry 2010; Polyakov et al. 2010), melting at the marine terminations of ice sheets (Straneo and Heimbach 2013; Rignot et al. 2013), changes in atmospheric and oceanic circulations (e.g., Hoerling et al. 2001; Toggweiler and Russell 2008), and sea level rise (Church et al. 2011). Since the 1970s, about one-third of the observed global mean sea level rise is explained by thermal expansion of the ocean's volume caused by increases in OHC (Church et al. 2013).

Here we focus on changes to OHC anomaly (OHCA) because of its importance to the global energy balance of the earth. While OHCA is expressed in joules, the rate of change of OHCA is often expressed in terms of watts per square meter applied over the earth's surface area ( $5.1 \times 10^{14} \text{ m}^2$ ) for comparisons with and constraints on net TOA energy fluxes.

OHCA can be calculated from in situ measurements by differencing ocean temperature profile data from a baseline climatology, multiplying this temperature anomaly by density and the specific heat of seawater, and volume integrating this product. Several research groups routinely produce OHCA estimates from in situ data that are widely used for ocean and climate science, model evaluations, outreach activities, and governmental reports. All of these estimates show that the global ocean has significantly warmed in the upper 700 m during the past decades; however, they can disagree in variability and warming rates (Rhein et al. 2013; Abraham et al. 2013), although some differences are not statistically significant. Differences arise because, in practice, there are many sources of uncertainty in OHCA estimation (Lyman et al. 2010; Palmer et al. 2010), and groups deal with them in different ways. Two relevant issues in OHCA estimation are associated with the evolving nature of the technology used to gather subsurface ocean temperature profiles (e.g., mix of instrumental accuracies and biases as well as quality control procedures) and the uneven distribution of the observational coverage in time, depth, and geography (Abraham et al. 2013; Gould et al. 2013).

Before the Argo program (Roemmich et al. 2009) array of autonomous profiling floats, which began

implementation around 2000, became near global, most ocean temperature profile data were obtained from instrumentation dropped over the side of a ship (Johnson and Wijffels 2011; Abraham et al. 2013). This practice limited ocean temperature profile measurements in hard-to-reach areas away from major shipping routes and during seasons of inclement weather, resulting in an uneven distribution of temperature profiles in time and space. To estimate global integrals of OHCA, algorithms must be developed to put temperature anomaly–OHCA information on a regular grid, cope with data-sparse regions, and smooth out temporal and spatial variability. These algorithms will hereafter be referred to as “mapping methods.” Prior to the middle of the twentieth century, ocean temperature profile data are too sparse to calculate OHCA with any mapping method without very large uncertainties (Gouretski et al. 2012). From the middle of the twentieth century to present, mapping method is still a source of uncertainty in OHCA estimates from in situ data (Gregory et al. 2004; AchutaRao et al. 2006).

Lyman et al. (2010) compared global OHCA estimates in the upper 700 m for 1993–2008 to quantify sources of uncertainty in OHCA calculations, including combined quality control and expendable bathythermograph (XBT) measurement bias corrections, climatological reference, observational sampling, and mapping method. Their calculations were all relative to comparisons using the same representative mean (PMEL\_R) mapping method from Lyman and Johnson (2008), except in the mapping method case. Uncertainty due to mapping method was quantified by matching pairs of OHCA estimates that had the same XBT corrections between those originally published by individual research groups and those mapped by using the PMEL\_R routine, although original estimates included differences other than just mapping method (e.g., baseline climatology and non-XBT data). Lyman et al. (2010) concluded that the largest source of uncertainty was found to be the combination of quality control and XBT bias corrections. Observational sampling uncertainties include horizontal and vertical variations in sampling (Lyman and Johnson 2008), changes in vertical sampling resolution (spacing between adjacent measurements; Cheng and Zhu 2014a), and changes in the ocean-observing system (Cheng and Zhu 2014b).

Here we investigate the sensitivity of global OHCA estimates in the upper 700 m for 1970–2008 and 1993–2008 to mapping methods, XBT bias corrections, and baseline climatologies. Our sensitivity experiments extend the earlier work of Lyman et al. (2010) by directly quantifying uncertainty attributable to mapping method and separately quantifying uncertainty attributable to

different baseline climatologies. Our analyses also cover a longer period, 1970–2008, including years prior to 1993, when data were sparser and hence global estimates more susceptible to larger mapping uncertainties. However, we do not address the influence of quality control of in situ temperature in detail. Monthly baseline climatological means are used in all instances to remove the seasonal cycle and minimize the impact on OHCA calculations. Not all extant mapping methods are examined here (e.g., [Cheng and Zhu 2014b](#); [Cheng et al. 2015](#)). Observational sampling changes and their effects on uncertainty are not addressed directly here, although they create differences arising from mapping methods and baseline climatologies.

Below we briefly describe the mapping methods routinely used by the seven research groups involved in this study, as well as their baseline climatologies, and provide an overview of the available set of XBT bias adjustments. [Section 2](#) provides details on the sensitivity experiments performed in this study. Results are found in [section 3](#) and conclusions in [section 4](#).

#### *a. Details of mapping methods*

The mapping methods used in each of the estimates of OHCA analyzed here are described below.

##### 1) LEVITUS (LEV) MAPPING

[Levitus et al. \(2000\)](#) compute OHCA for the upper 300 m and for the upper 3000 m for the years 1948–98 using 5-yr (pentadal) running means as the compositing time period and monthly climatologies calculated from all available data from the *World Ocean Atlas 2001* ([Stephens et al. 2002](#), hereafter *WOA01*; years 1772–1998) as the long-term mean. This work shows an increase in OHCA over the years under study and provides an estimate of OHC changes for the second half of the twentieth century. The procedure outlined in [Levitus et al. \(2000\)](#) has been modified since then, deepening the shallower depth layer from 300 to 700 m and shortening the compositing time period to 1 year ([Levitus et al. 2005](#)). Two subsequent updates on this work were published ([Levitus et al. 2009, 2012](#)), which use a baseline monthly climatology calculated using data only for the years 1955–2006. Pentadal, yearly, and seasonal OHCA are calculated based on the method outlined in these works and posted online every three months ([http://www.nodc.noaa.gov/OC5/3M\\_HEAT\\_CONTENT/](http://www.nodc.noaa.gov/OC5/3M_HEAT_CONTENT/)). Currently, this time series covers the years 1955–2015. The LEV method used by [Levitus et al. \(2000, 2005, 2009, 2012\)](#) is an objective analysis technique developed by [Cressman \(1959\)](#) for meteorological data, modified by [Barnes \(1964\)](#). The technique calculates a correction factor at each grid point based on

the weighted difference between a first-guess value and a mean value for each grid point within a radius of influence. The first-guess field is zero (degrees Celsius for temperature anomalies, joules for OHCA, and so forth). The mean value is the difference between an in situ value and the baseline monthly climatology averaged over all such differences within a 1° latitude by longitude grid box. The grid point is the center of the 1° grid box. Weights are calculated based on distance from the grid point within the radius of influence ([Barnes 1964](#)). The correction factor is applied to the first-guess field at each grid point, a 5-point smoother ([Shuman 1957](#)) is then applied, and the procedure is repeated twice more, with progressively smaller radii of influence preserving smaller-scale signals when possible. The three radii of influence are 880, 660, and 440 km, approximately 8°, 6°, and 4° of latitude and longitude at the equator. Full details, including a complete description of the weight function, can be found in [Locarnini et al. \(2010\)](#).

##### 2) ISHII (ISH) MAPPING

[Ishii et al. \(2006\)](#) calculates global OHCA for the upper 700 m using subsurface temperatures computed with an analysis based on a variational minimization scheme ([Derber and Rosati 1989](#)) with spatiotemporal covariance of background error. Mixed layer temperatures are constrained by sea surface temperature (SST; [Ishii et al. 2005](#)). The analysis scheme is detailed in [Ishii et al. \(2003\)](#). The Tikhonov term in [Ishii et al. \(2003\)](#) is replaced by background error decorrelation as a function of both horizontal and vertical distances for later OHCA calculations ([Ishii et al. 2006](#); [Ishii and Kimoto 2009](#)). Spatial decorrelation scales vary with depth. The surface horizontal decorrelation scale is given as a cosine function of latitude, with 900 km at the equator and a minimum of 300 km. At the surface the vertical decorrelation scale is 10 m and temporal decorrelation scale is 15 days. Monthly differences from a baseline climatological monthly mean were calculated in the analysis scheme, binning all data within 120 days of the 15th day of each month.

##### 3) WILLIS (WIL) MAPPING

[Willis et al. \(2004\)](#) calculates OHCA for the years 1993–2003 for the upper 750 m of the global ocean using altimeter data to compensate for sparse in situ temperature profile data. As outlined in [Willis et al. \(2003\)](#) a two-term covariance with large (920 km) and small (100 km) scales is used to map the in situ data to all grid points. The covariance function is chosen based on global wavenumber–frequency spectrum for altimetric sea surface height ([Zang and Wunsch 2001](#)). The seasonal cycle

is removed by subtracting out baseline climatological means. Altimeter fields of sea surface height (AVISO; Ducet et al. 2000) are then used to modify the objectively mapped in situ values using the following:

$$\langle \text{INSIT} \rangle \text{ difference estimate} = \langle \text{INSIT} - (\alpha \text{AH}) \rangle + \langle \alpha \text{AH} \rangle, \quad (1)$$

where INSIT is the in situ field, AH is the altimeter sea surface height field, and  $\alpha$  is a time-averaged regression coefficient for AH onto INSIT. The angled brackets denote objective mapping as described above. Thus, Eq. (1) results in an objective mapping of a quantity (INSIT) using altimetric sea surface height, represented in terms of INSIT as a first guess. In the case of OHCA, INSIT is derived from a mean difference between observed and baseline mean monthly temperature.

#### 4) PMEL MAPPING

Pacific Marine Environmental Laboratory (PMEL) mapping methods are derived from the objective mapping methods of Willis et al. (2003, 2004). As with all other methods, seasonal cycle is removed by subtracting out baseline climatological monthly means. Instead of using sea surface height anomalies from satellite altimeter data to compensate for sparse in situ ocean temperature profile data as in WIL, the PMEL method calculates a representative mean, which effectively assumes that unsampled regions have the same anomalies as the mean anomaly of the sampled regions (Lyman and Johnson 2008). This method is used to compensate for areas where the distance between measurements is large compared with mapping correlation length scales or the amount of data are insufficient to overcome the noise-to-signal ratios (Lyman and Johnson 2008). The initial global estimate in the PMEL method is the mean of the maps (PMEL\_M; called SI in Lyman and Johnson 2008). In the mean of the maps, the objective map relaxes back toward the initial guess of zero anomalies in data-sparse regions. The second step in the mapping is the calculation of the global representative mean (PMEL\_R), wherein data-sparse areas are effectively given a value equal to the mean value for all well-represented areas (Lyman and Johnson 2008). While the PMEL\_M product should be viewed as an intermediate step to the final global integral of PMEL\_R, both are retained here so that PMEL\_M can be compared against other objectively mapped products.

#### 5) MET OFFICE HADLEY CENTRE DATASET MAPPING

The Met Office Hadley Centre produces monthly objectively analyzed temperature fields using a modified

version of the method outlined in quality control procedures for the Enhanced Ocean Data Assimilation and Climate Prediction (ENACT) and ENSEMBLES projects (Ingleby and Huddleston 2007). This mapping method (hereafter referred to as EN) is a modified successive correction scheme that approximates optimal interpolation (Lorenc et al. 1991; Bell et al. 2000). The background for the analysis for each month is the combination of climatology and a damped persistence of anomalies from the previous month's analysis:

$$B_{i+1} = C_{i+1} + \alpha(A_i - C_i), \quad (2)$$

where  $C_i$  is the climatological mean for month  $i$ ,  $A_i$  is the analyzed value for the given month-year,  $\alpha$  is a factor, here set to 0.9 (based on experimental best fit), and  $B_{i+1}$  is the background for the following month. This creates a first-guess field that has damped persistence from the previous month-year, relaxing to climatology where there are no data. The combination of the background and the observations is governed by estimates of observation error variance and background error covariance. The covariance is parameterized using two second-order autoregressive (SOAR) functions with horizontal scales of 300 and 400 km. The former is extended near the equator to 1500 km. The analysis scheme used in this study corresponds to that used to produce the EN3 version 2a dataset (Ingleby and Huddleston 2007) (<http://www.metoffice.gov.uk/hadobs/en3>).

#### 6) DOMINGUES (DOM) MAPPING

The CSIRO–Antarctic Climate and Ecosystems Cooperative Research Centre (ACE CRC)–Institute for Marine and Antarctic Studies (IMAS) research group produces global OHCA estimates for the upper 700 m of the ocean by summing two depth-integrated layers, 0–300 and 300–700 m (Domingues et al. 2008; Church et al. 2011). In situ temperature observations from Argo floats, bottles, CTDs, and XBT-corrected data (Wijffels et al. 2008) are first converted into thermosteric sea level. Anomalies are obtained using a historical climatology that includes annual, semiannual, and linear trend terms (Alory et al. 2007; Wijffels et al. 2008). Thermosteric sea level anomaly profiles are depth integrated for the layers mentioned above. If more than one profile exists in the same  $1^\circ \times 1^\circ$  grid box, within  $65^\circ\text{N}$ – $65^\circ\text{S}$ , their median is used. These  $1^\circ$  estimates are then projected onto the first 30 EOFs, derived from the satellite record altimeter (since 1993), to reconstruct spatially complete fields, at monthly time scales, from the 1950s onward. The reconstruction method is based on a reduced-space optimal interpolation (Kaplan et al. 2000), which formalism provides estimates of errors on

the basis of the in situ data distribution and uncertainties (instrumental and geophysical errors) as well as ocean eddy variability based on satellite altimeter data. Finally, thermosteric sea level anomaly fields are converted into OHCA using coefficients obtained from a spatially variable linear regression between OHCA and thermosteric sea level anomalies on a  $10^\circ \times 10^\circ$  grid, for each of the two depth layers (similarly to Willis et al. 2004). To reduce unwanted noise, monthly estimates are yearly averaged and then a 3-yr running mean is applied. The version history of these reconstructed global OHCA and thermosteric sea level anomalies, including the latest available estimates, are found online ([http://www.cmar.csiro.au/sealevel/thermal\\_expansion\\_ocean\\_heat\\_timeseries.html](http://www.cmar.csiro.au/sealevel/thermal_expansion_ocean_heat_timeseries.html)).

### 7) GOURETSKI (GOU) MAPPING

Gouretski et al. (2012) calculate OHCA from median temperature anomalies for a depth layer over  $5^\circ$  grid boxes. The depth layers are 0–20 and 0–400 m. Temperature profiles are first interpolated to 1-m increments. The 0–400-m vertical temperature average is obtained using correlations between the mean temperature of the entire layer to mean temperatures in shallower levels. The correlations are calculated based on high-resolution CTD temperature profiles. The mean layer temperature anomaly is then differenced from a baseline monthly mean. Baseline monthly mean layer temperatures are calculated on a  $0.25^\circ$  grid from temperature profiles measured during 2001–10. The baseline mean is calculated using the following weighting function:

$$W = (R^2 - r^2)/(R^2 + r^2), \quad (3)$$

where  $R$  is the radius of influence, set to 111 km, and  $r$  is the distance from observation to each grid point. The global average layer temperature anomaly is then the sum of all  $5^\circ$  median layer temperature anomalies divided by the number of  $5^\circ$  grid boxes containing data. This is converted to OHCA over the entire ocean surface area, thus assuming a mean global layer anomaly for all  $5^\circ$  ocean grid boxes without data. This assumption is similar to that used in PMEL\_R and by Palmer et al. (2007).

#### *b. XBT bias adjustments*

Temperature profiles collected by XBTs make up a large part of the available subsurface temperature data from 1970 to 2003. Without these data, calculations of OHCA for in situ profiles would be very difficult for these decades. Even now, when the Argo program is the dominant observing system, XBTs make up on the order

of 10% of available subsurface temperature profiles each year. The depths of temperature measurements from XBTs are estimated from elapsed time since launch using a fall-rate equation. It was known early on that there were systematic errors in the XBT data (Flierl and Robinson 1977). Hanawa et al. (1995) introduced a correction for the most commonly used XBTs [Sippican and Tsurumi Seiki (TSK) T-4, T-6, and T-7 models] based on a time-invariant fall-rate equation to mitigate the biases. However, Gouretski and Koltermann (2007) showed that XBT biases vary with year, and thus corrections need to be time dependent.

Since then, a number of XBT corrections have been proposed for depth (Wijffels et al. 2008, hereafter W08 for their Table 1 corrections; Ishii and Kimoto 2009, hereafter I09 when referring to XBT corrections; Good 2011, hereafter G11), for temperature (Levitus et al. 2009, hereafter L09 when referring to XBT corrections), and for both depth and temperature (Gouretski and Reseghetti 2010, hereafter GR10; Hamon et al. 2012; Gouretski 2012, hereafter GO12; Cowley et al. 2013, hereafter C13 when referring to the thermal gradient XBT correction). Of the different calculations of OHCA, LEV uses L09; ISH, PMEL\_M, and PMEL\_R use I09; WIL uses Wijffels et al. (2008, their Table 2); DOM uses W08; and GOU and EN use GR10. The XBT bias, when modeled as a fall-rate correction only, tends to report water parcels as being deeper than their true depths. Since water temperature generally decreases with depth, the deeper displacement increases the reported temperature at the deeper depth, thereby increasing the OHCA relative to an unbiased measurement. Splitting the XBT types into shallow ( $<550$  m) and deep ( $>550$  m) as per W08, the depth bias is approximately 10 m at 400-m depth for the shallow probes and 10 m at 700-m depth for the deeper probes averaged over all years, but with significant year-to-year variations. Viewing the XBT bias as pure (but depth dependent) temperature offset (L09) results in a temperature correction that is on the order of  $+0.1^\circ\text{C}$  averaged over the 0–700-m depth range. C13 estimate a mean temperature bias of  $+0.05^\circ\text{C}$  with a widely varying depth bias by probe and year, but generally smaller than the depth-only corrections of W08. All XBT bias correction methods are time (year to year) varying in depth and/or temperature correction. Differences among methods have implications for OHCA calculations.

Cheng et al. (2016) report on recommendations by the XBT community that XBT corrections account for five factors (depth bias, pure temperature bias, temperature-induced depth bias, temperature-induced temperature bias, and depth offset). Only one correction scheme



(Cheng et al. 2014) directly accounts for all five of these factors explicitly. Work is still in progress to quantify which XBT bias correction method works best. For experimental simplicity and clarity of results, rather than using all extant XBT correction methods, XBT correction methods were selected from the categories of depth correction, temperature correction, and depth-plus-temperature correction. The importance of initial velocity variations of the XBT probe due to drop height has recently been documented using a numerical method (Abraham et al. 2012) and empirically from water tank experiments (Bringas and Goni, 2015). C13 (and Cheng et al. 2014) account for this effect in the depth offset corrections.

As previously indicated, Lyman et al. (2010) show that the differences due to XBT corrections and quality control combined are the largest source of uncertainty in OHCA calculations. In the present work, we quantify uncertainties owing to XBT corrections but without the complicating influence of quality control. We make a comparison of OHCA using the different mapping methods and the same XBT corrections. We also compare OHCA using each mapping method with different XBT corrections to see how each mapping method is affected by the XBT correction used. Since there is no consensus yet on the best XBT correction scheme to implement, the W08 XBT corrections, as the longest in-use XBT correction for OHCA (Domingues et al. 2008), will be used for comparison of mapping methods with different baseline climatologies.

### c. Baseline climatologies

Also not part of the actual mapping method, but of vital importance to the OHCA calculation, is the baseline mean subtracted from observations to compute anomalies. Ishii and Kimoto (2009, their Table 5) find a 10% difference in the linear trend of OHCA ( $0.14 \text{ ZJ yr}^{-1}$ ;  $1 \text{ ZJ} = 1 \times 10^{21} \text{ J}$ ) from 1951 to 2005 when using the *World Ocean Atlas 2005* (Locarnini et al. 2006, hereafter *WOA05*) in place of *WOA01*. The difference between *WOA05* and *WOA01* is in the input data, the *World Ocean Database 2005* (Boyer et al. 2006, hereafter *WOD05*) and the *World Ocean Database 2001* (Conkright et al. 2002, hereafter *WOD01*), respectively. *WOD05* contains data from 7 additional years (1999–2005), additional historical data, and additional quality control. However, *WOD05* and *WOD01* are used as input data for temperature anomalies as well in the respective cases, so some of the difference in OHCA found by Ishii and Kimoto (2009) may be due to these data differences. Lyman et al. (2010) compare OHCA using baseline climatologies calculated using data from year

periods 1993–2003 and 2005–08 for their comparison and find negligible differences in OHCA curves due to baseline climatology after adjustment for the warmer ocean in the 2005–08 period. Lyman and Johnson (2014) show a striking difference in OHCA for the PMEL\_M mapping method when using a baseline climatology for 2005–10 from Argo data compared to the same climatology adjusted to reflect estimates of mean ocean temperature differences in 1955 relative to 2005–10. The difference in baseline does not affect their final product, the PMEL\_R method (Lyman and Johnson 2014).

The mapping methods described above all use baseline monthly mean temperature climatologies over different time periods. LEV currently uses monthly *WOA09* as a baseline (Levitus et al. 2009). *WOA09* monthly climatologies are the means of five decadal (or near decadal) climatologies (1955–64, 1965–74, 1975–84, 1985–94, and 1995–2006), thus producing a mean climatology of 1955–2006 that is not skewed toward any one decade. The most recent PMEL maps (Johnson et al. 2015) use a monthly baseline climatology of 2004–13 derived from Argo data only. OHCA are then adjusted to reflect differences from the period 1993–2013. The most current ISH (Ishii and Kimoto 2009) uses a baseline climatology encompassing years 1961–2000. DOM uses a climatology that has the long-term linear trend of temperature change removed (Alory et al. 2007). The EN uses a time-varying baseline mean that incorporates persistence of temperature change from previous years, as described above. All of the above use climatological monthly mean temperatures at discrete depths on a  $1^\circ$  grid. GOU uses climatological monthly mean vertically averaged temperature on a  $1/4^\circ$  grid. Similar to Ishii and Kimoto (2009) and Lyman et al. (2010), OHCA in this study is calculated relative to different baseline climatologies. These calculations are used to assess the impact of baseline climatology on each of the mapping methods.

## 2. Methods

To perform the intercomparison experiments for this study, a uniform set of input data is generated. To prepare the data for mapping, the method outlined in Domingues et al. (2008, their supplementary material) is followed. For 1970–2004 the EN 3 version 2a database (EN3v2a; Ingleby and Huddleston 2007) of in situ temperature profiles is used for all bottle, CTD, and XBT data. For 2000–08, the real-time and delayed mode temperature profiles from Argo profiling floats as in Barker et al. (2011) are used. Only temperature profile data from the above-mentioned instrument classes are used. Data from moored buoys, drifting buoys, towed

CTDs, and mechanical bathythermographs (MBTs) are not used, following the data choice of Domingues et al. (2008). Inclusion/exclusion of specific ocean profile data types will have an effect on OHCA uncertainty. It should be noted that the Argo array still had geographic gaps in its target coverage until 2006/07 and was not designed initially to cover continental shelves, marginal seas, or areas where ice is present (Roemmich et al. 2015). Of the years in this study with only Argo data (2005–08), year 2005 should not be considered to have full Argo coverage. After removal of gross outliers using a simple check (Domingues et al. 2008), quality flags set in the EN3v2a quality control procedures (Ingleby and Huddleston 2007) are used almost exclusively for bottle, CTD, and XBT data. If a profile had a measurement with a quality flag other than good, the entire profile was excluded from use. Additionally, profiles are not used if they have coarse vertical resolution (as in Willis et al. 2004) or if they terminate at depths shallower than 100 m. For each temperature profile OHCA is calculated relative to a baseline climatology for 0–300 m and for 300–700 m. The 0–700-m OHCA is the sum of the OHCA for each of the two layers. Splitting of the ocean into two layers follows Domingues et al. (2008) and allows for the inclusion of profile data that do not reach to the full 700-m study depth. Vertical sampling depth contributes to OHCA uncertainties. Each of the mapping methods is applied to this same input OHCA dataset to calculate a  $1^\circ$  latitude–longitude bin gridded field of OHCA between  $65^\circ\text{S}$  and  $65^\circ\text{N}$  for each year 1970–2008 ( $5^\circ$  gridded bins in GOU method, years 1993–2008 only for WIL method). The years are chosen to start with the advent of the XBT, which provided increased global data coverage for 0–300 m for 1970 onward compared to previous years (Church et al. 2011; Gleckler et al. 2012; Abraham et al. 2013; Lyman and Johnson 2014). Since the only difference between the gridded fields produced is the mapping method, the effect of different mapping methods on OHCA calculations can be readily examined. Furthermore, by using the same input dataset as described above, with different sets of XBT corrections and base climatologies, an examination of the relative effect of these factors on the different mapping methods can be evaluated.

The different test cases using the same input data (EN3v2a + Barker Argo) are as follows, using as naming convention CY\_T\_XXX for each case. The CY\_T portion of the name refers to the baseline climatology used, whereas the XXX portion refers to the XBT bias correction (see Table 1 for a complete listing).

Uncertainty is represented by the yearly standard deviation among grouped cases. There are seven mapping methods (eight for 1993–2008), six XBT correction

schemes, and three baseline climatologies representing the different cases studies here, so the standard deviation is based on relatively few separate cases, especially for the baseline means. Furthermore, as discussed in Lyman et al. (2010), the different cases are not completely independent since they use the same input dataset. To better quantify confidence in the standard deviation as uncertainty, effective degrees of freedom (EDOF) are given for each case. The value of the EDOF is always smaller than the simple degrees of freedom for standard deviation since it accounts for the interdependence of the datasets (Lyman et al. 2010, based on Bretherton et al. 1999).

#### a. Baseline climatologies

Three monthly mean temperature climatologies—two historical (C1\_H and C2\_H) and one modern (C3\_M)—were used to calculate OHCA (without seasonal variability) for each temperature profile from the input dataset versions for the sensitivity tests (Table 1). The historical climatologies (C1\_H and C2\_H) were originally mapped by Alory et al. (2007) and are based on historical observations from the EN3v1d database (Ingleby and Huddleston 2007)—the previous version to EN3v2a—merged with recent Argo data (Barker et al. 2011). Cheng and Zhu (2015) show that historical climatologies can have inconsistencies in the mean year for different regions of the ocean based on sampling patterns (e.g., good coverage most years in the Northern Hemisphere and good coverage in the Southern Hemisphere only during the Argo era). If mean year is different for different regions, OHCA calculations will be inconsistent among the regions. As part of the Alory et al. (2007) mapping technique, a long-term linear trend was removed at each grid point to reduce the impact of these sampling biases from the observing system, ameliorating the inconsistent mean year problem. At the same time, inclusion of Argo data helped to reduce aliasing of seasonal signals into historical changes, particularly in high-latitude areas of the Southern Hemisphere (not shown), which are historically poorly sampled in winter. C1\_H contains temperature data from bottles, CTDs, and Argo floats, whereas C2\_H also includes bias-corrected XBT from W08. By including XBT data, C2\_H has a greater number of profiles and thus a better sampling coverage than C1\_H. On the other hand, C2\_H is burdened with uncertainties due to XBT bias corrections. The modern monthly mean temperature climatology (C3\_M), originally mapped by Locarnini et al. (2013), is dominated by Argo data for 2005–12 but also includes data from bias-corrected XBTs (Levitus et al. 2009), moored buoys, drifting buoys, gliders, and towed CTDs from the WOA13 (Locarnini et al. 2013). Although its 8-yr period is

TABLE 1. Summary of the sensitivity experiments performed in this study in terms of variations in XBT data (see section 1b), climatological baselines, and mapping methods (see section 1a). [Climatological baseline cases include H: historical period, only bottles, CTDs, and Argo data (C1\_H); 2: historical period, as with H but including XBT data corrected by W08 (C2\_H); and M: modern period (2005–12), Argo and concurrent non-Argo data from the WOA13.]

XBT cases		Mapping Methods							
		DOM	EN	LEV	GOU	ISH	PMEL_M	PMEL_R	WIL
W08	No XBT correction	H	—	H	H	H	H	H	H
	Depth	H, 2, M	H, 2, M	H, 2, M	H, 2, M	H, 2, M	H, 2, M	H, 2, M	H, 2, M
I09		H	H	H	H	H	H	H	H
L09	Temperature	H	H	H	H	H	H	H	H
GO12	Depth + temperature	H	H	H	H	H	H	H	H
C13		H	H	H	H	H	H	H	H
G11	Depth (bathymetry approach)	H	H	H	H	H	H	H	H

relatively short, C3\_M presents a more temporally uniform temperature field, which is not biased to different epochs depending on location, as for many historical climatologies; C3\_M also represents a warmer ocean than both C1\_H and C2\_H. (C3\_M global OHC is 157 ZJ higher than C1\_H and 141 ZJ higher than C2\_H). The latter two are representative of longer baseline periods with cooler means.

### b. XBT bias corrections

The XBT corrections examined are as follows: XXX = none, W08, L09, I09, G11, GO12, and C13. Each case is described above in section 1b. There are more available XBT corrections, but those used here represent a cross section of approaches—temperature corrections only (L09), depth corrections only (W08, I09, and G11), and depth and temperature corrections (GO12 and C13). “None” is the case where no XBT correction [other than Hanawa et al. (1995) where necessary] is applied. Since it is clear that there are XBT biases that necessitate correction (Gouretski and Koltermann 2007), results from the none case are presented for perspective but are not included in any calculations of uncertainties.

## 3. Results

The global OHCA time series created as part of the sensitivity experiments (Table 1) are organized into three groups to assess the spread introduced:

- by mapping methods for each XBT correction (section 3a),
- by XBT corrections for each mapping method (section 3b), and
- by baseline climatologies for each mapping method (section 3c).

In the figures, the same historical climatology (C1\_H) underpins the results shown for (i) and (ii) and the same XBT correction (W08) for (iii).

### a. Effect of mapping methods

The globally integrated OHCA estimates using eight different mapping methods using six different XBT corrections and the no-XBT-correction case, relative to the C1\_H climatology (Figs. 1a–f), exhibit apparent differences for all years over the 1970–2008 period. Since we use the same climatology for all estimates, no shifting to a common mean is required for purposes of comparison here. The uncertainty of OHCA due to XBT corrections is discussed below. Those features that are common to Figs. 1a–e (but not necessarily the no-XBT-correction case) are evaluated as differences attributable to mapping method.

In the Argo-data-only years (2005–08) ISH mapping produces the lowest globally integrated OHCA; WIL, DOM, PMEL\_R, and GOU produce the highest OHCA values; while LEV, PMEL\_M, and EN are similarly clustered in the middle (Fig. 1). Some mapping method–XBT correction combinations have peak OHCA in 2004 with a minimum in 2007 before increasing in 2008. The last year for this experiment in which XBT data were used was 2004, pointing to possible OHCA calculation uncertainty due to choice of input ocean profile data. For earlier years, with more XBT data, ISH, PMEL\_M, and EN are usually very similar, within about 20 ZJ of each other. LEV, DOM, PMEL\_R, and GOU generally track along with the other three OHCA mappings but with some larger year-to-year spikes (e.g., the increase of 79 ZJ from 1976 to 1977 followed by a decrease of 32 ZJ in 1978 in the LEV W08 case), most often of higher OHCA than ISH, PMEL\_M, and EN. DOM yearly estimates contain large spikes, which are usually reduced with a 3-yr running mean average (e.g., Domingues et al. 2008; Church et al. 2011). This smoothing, however, is not applied here, the better to allow intercomparisons. These spikes are not entirely physical but are due to the specific mapping method.



The range of OHCA estimates from the eight mapping methods is given in terms of standard deviations (Fig. 2). Calculated EDOF associated with the standard deviations of mapping methods are approximately 4 for years 1970–2008 and 3 for years 1993–2008 (Table 2). The differences among standard deviations for each XBT correction case are discussed below, focusing on mapping method differences. For the six XBT corrections the standard deviation for the different mapping methods is 16.5 ZJ over the 39-yr period, with a local minimum of 10.0 ZJ in 1976 and 1985 and a local maximum above 30.0 ZJ in 1997 and 2001. Mean standard deviations among estimates using different mapping methods over 1970–2008 for different XBT corrections range from 14.0 ZJ for W08 to 17.9 ZJ for C13. Over 1993–2008, the range of OHCA mapping method standard deviations for different XBT corrections is larger (12.5 ZJ for W08 to 20.0 ZJ for L09), and the average is 17.1. The average standard deviations—16.5 for 1970–2008 and 17.1 for 1993–2008—are used as the overall estimates of the uncertainty due to mapping method. This uncertainty is a measure of the spread of OHCA values among the mapping methods within a year. It is not exactly comparable to the 9.0-ZJ standard error of the mean estimated by Lyman et al. (2010) for the 1993–2008 time period. In Lyman et al. (2010) uncertainties were given as standard errors, but we present them here as standard deviations to facilitate comparisons among the different mapping methods, which produce time series with different decorrelation time scales. The discrepancies among OHCA estimates from different mapping methods diminish and converge toward the end of the time series (Fig. 1), when global coverage improves during the Argo period, as discussed below. The standard deviations of the OHCA values from the mid- to late 2000s are below the average over the entire period analyzed (Fig. 2) but similar to values found in the early 1990s. Even with the increased global coverage of the Argo array, there are still differences due to mapping method, both for areas still without data coverage and for areas with sufficient data coverage.

Insight into the differences in OHCA estimates from the various mapping methods can be gained by exploring the yearly variations of fractional coverage for ocean temperature data for each mapping method (Fig. 3). Fractional coverage, the fraction of total ocean area with sufficient information to estimate OHCA (Lyman and Johnson 2008), is specific to each mapping method. Alternatively, this quantity can be thought of as what fraction of a globally uniform signal is recovered by each mapping routine given the spatial and temporal data distribution. In general, fractional coverage is relatively

high during the 1990s, when the World Ocean Circulation Experiment (WOCE) collected substantial numbers of ocean temperature profiles around the world, and even higher during the last few years of the record, when the global Argo array of profiling CTD floats builds to its initial target strength. The GOU mapping generally has the smallest fractional coverage because it requires data within a 5° grid box to assume that coverage is sufficient, but after the final step, its fractional coverage becomes equal to 1 in every year. PMEL\_M and ISH methods are similar, probably because they both use objective mapping methods that incorporate similar correlation length scales and additionally give less weight to anomalies in regions where there are few data points. The EN method, while using spatial decorrelation length scales similar to the other methods, adds temporal persistence beyond a year to their monthly estimates as described above so that even one year after a data point is collected, it still has a 28% effect on the anomaly estimate at that location in the absence of new data. Hence, the EN fractional coverage is somewhat higher than those of PMEL\_M and ISH. The LEV maps use the longest spatial correlation length scales of all and also do not take signal-to-noise ratios into account in their smoothing; they thus result in large yearly fractional coverages, lower only than WIL and PMEL\_R. Since the PMEL\_R estimates assume that the mean anomalies of sampled regions apply to unsampled regions, similar to the final GOU estimates, they also effectively have fractional coverage of 1 for every year after that step has been taken.

The DOM curve does not show fractional coverage as defined above but rather the percent of 1° ocean boxes with data, as their method does not use a spatial mapping radius. DOM compensates for this relative data paucity with a well-developed related field of sea level height on which to regress the temperature data for OHCA estimates. The relatively low percent coverage may also lead to some of the spikes seen in the DOM estimates. The spikes are not apparent in 3-yr means of OHCA, which is how results of the DOM mapping method are generally presented.

The different mapping methods yield different results regionally as well as in the global integral. The geographic distribution of OHCA for the year with highest percent coverage in this study (2008) varies by method (Fig. 4). Coverage is nearly global, owing to the Argo data, and extends at least to the 700-m depth, the limit for this study, in almost all areas of the ocean (Fig. 4a). La Niña conditions present in 2008 are represented by cooler tropical and subtropical eastern Pacific waters in all mappings except EN (Fig. 4d). Warming along the Indian Ocean Antarctic zone, warming over most of the Atlantic Ocean, and cooling

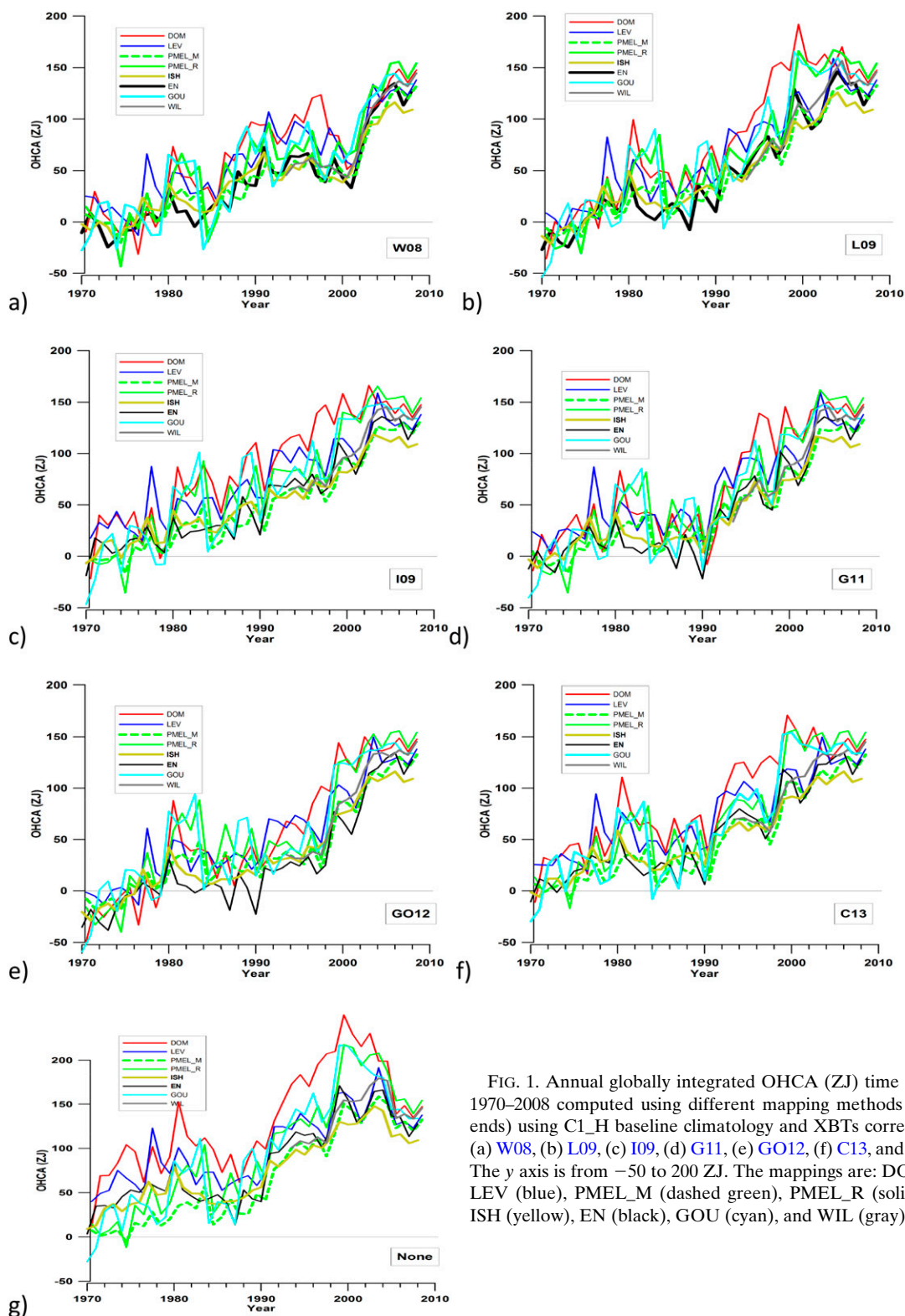


FIG. 1. Annual globally integrated OHCA (ZJ) time series for 1970–2008 computed using different mapping methods (see legends) using C1\_H baseline climatology and XBTs corrected with (a) W08, (b) L09, (c) I09, (d) G11, (e) GO12, (f) C13, and (g) none. The y axis is from  $-50$  to  $200$  ZJ. The mappings are: DOM (red), LEV (blue), PMEL\_M (dashed green), PMEL\_R (solid green), ISH (yellow), EN (black), GOU (cyan), and WIL (gray).

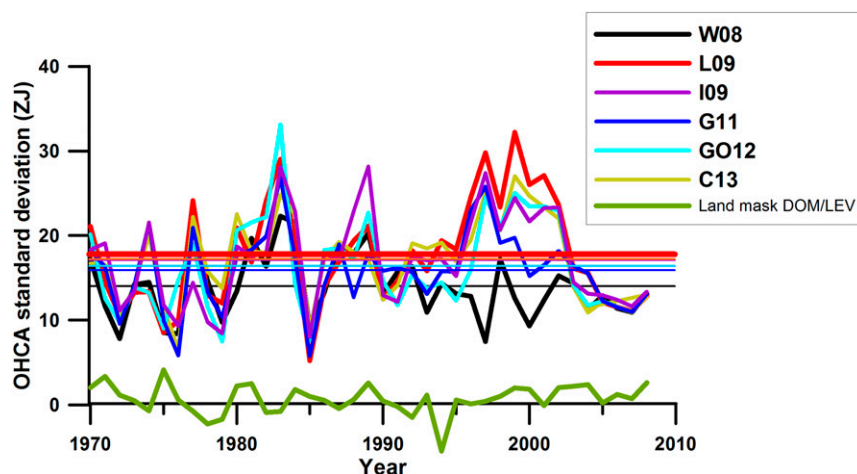


FIG. 2. Standard deviations of global OHCA (ZJ) for 1970–2008 from ensemble means of the eight different mapping methods for each of the six XBT bias corrections (see legend) referenced to the C1\_H baseline climatology. Horizontal lines represent the standard deviations means over the full time period. The green line denotes the OHCA difference between DOM and LEV results where land–ocean grids are represented in one or the other method, but not both (ocean definition differences).

in the central Indian Ocean are all represented to some degree in each mapping method. The DOM (Fig. 4b) and ISH (Fig. 4f) OHCA fields for 2008 are smoother than the LEV, PMEL\_M (Fig. 4e), and WIL (Fig. 4h) OHCA fields, which exhibit more small-scale features. The differences can be explained for each case. The EOF mapping using sea surface height patterns in DOM can only resolve the larger-scale features (Kaplan et al. 1998, 2000). The LEV mapping does not give less weight to sparse data (a lone data point within a radius of influence will fully describe the OHCA at a point in the absence of additional data points, whereas other mapping methods will weigh the singular data point at a value less than one, diminishing its influence), the PMEL\_M and WIL mappings include a small spatial correlation length scale, and the WIL mapping makes use of detailed fields of altimetric sea surface height,

all of which tend to preserve smaller-scale features. Whether some of these smaller-scale features are noise introduced by the point measurements of the in situ temperature profiles, are factors influencing sea surface height other than OHCA, or may legitimately represent the influences of eddies and other small-scale ocean features will not be investigated here. The different mapping methods are filtering noise and smoothing at different length scales and this accounts for some of the differences in OHCA from different mapping methods.

The fractional global coverage for 1980 ranges from 45% (GOU) to 81% (LEV) and in general is much lower than the coverage for 2008 (Figs. 3 and 5a). Areas with data, such as the North Pacific, do not have coverage below 300 m (Fig. 5a). The DOM-mapped OHCA for 1980 (Fig. 5b) is positive or close to 0.0 ZJ over most

TABLE 2. EDOF calculated following Lyman et al. (2010) for mapping method cases comparison, XBT correction cases, and baseline mean cases.

Mapping method cases			XBT correction cases			Baseline mean cases		
	1970–2008	1993–2008		1970–2008	1993–2008		1970–2008	1993–2008
W08	4.3	3.4	DOM	3.6	3.0	DOM	1.3	1.6
L09	3.6	3.6	LEV	3.6	3.1	LEV	1.4	1.3
I09	3.7	3.7	PMEL_M	4.2	2.9	PMEL_M	1.1	1.4
C13	3.9	3.7	PMEL_R	4.2	2.9	PMEL_R	1.8	1.8
GO12	3.9	3.5	ISH	3.6	2.9	ISH	1.2	1.6
G11	4.3	3.8	EN	3.9	3.1	EN	1.2	1.4
			GOU	3.9	2.9	GOU	1.9	1.3
			WIL		2.9	WIL		1.4

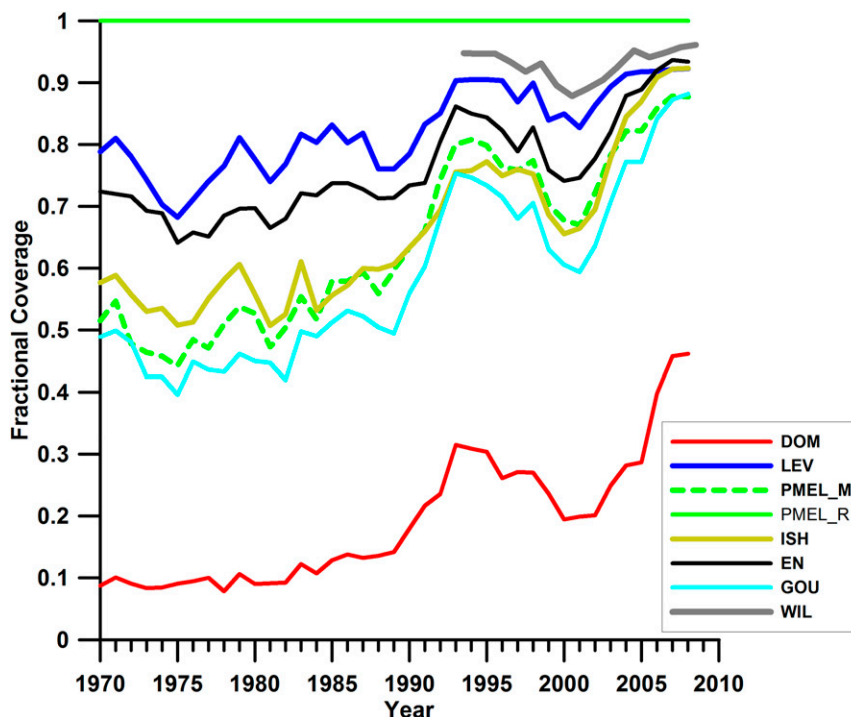


FIG. 3. Fraction of the global ocean considered sampled for temperature data in each year for each mapping method (see legend). The PMEL\_R and GOU method after its final infilling step are equal to 1 for every year (green solid line). (GOU curve before infill is shown here.) DOM curve (red) is not fractional coverage but indicates where data are available for  $1^\circ$  grid boxes without including additional spatial radius.

of the global ocean, with the exception of the far-western Pacific, east of the Philippines, along the track of the South Pacific convergence zone (roughly) and west of Australia. The DOM mapping also shows a coherent warm region in the central-east and northeast Pacific. Other mappings exhibit localized cooling in the northwest Atlantic and in the eastern Pacific. All mappings show warming greater than  $10 \times 10^{18}$  J in the central Indian Ocean, but this warming is particularly strong in LEV (Fig. 5c) and GOU (Fig. 5g), possibly owing to the influence on these mapping methods of the isolated data in the area (Fig. 5a). Despite the higher coverage in the Northern Hemisphere for 1980 compared to the Southern Hemisphere (Fig. 5a) the 18-ZJ difference in OHCA for the Northern Hemisphere between DOM and LEV is not that much smaller than the 24.5-ZJ difference for the Southern Hemisphere. In all maps, the Indian Ocean is mainly positive, with higher OHCA for LEV and GOU (Figs. 5c,g). Strong positive OHCA in all three basins of the Southern Ocean are only seen in DOM (Fig. 5b).

Each mapping method has slightly different land-ocean definitions, but only the DOM method excludes marginal seas and ice-covered areas. In addition, compared to the LEV mask, there are also some differences

along coastlines. (Figs. 4b and 5b, dark shading). Despite that, the effect of the different land-ocean definitions for DOM and LEV is not large with respect to integrated global OHCA in comparison to mapping method differences (Fig. 2, green curve).

Quality control of the in situ profile data is also important and can cause differences in OHCA using different mapping methods or keeping the mapping method constant. The 2008 OHCA field using LEV mapping with the World Ocean Database (WOD) quality control flags and *World Ocean Atlas 2009* baseline climatology (Fig. 4i) is different from the field produced by LEV with the Barker Argo data quality control (Fig. 4c) because of the sensitivity of the LEV mapping method to data quality control. While these fields are very similar on the large (basin) scale, there are some differences. For instance, the cool OHCA in the eastern Pacific is larger in the LEV method with WOD quality control (Fig. 4i) than in the LEV method with Barker Argo quality control (Fig. 4c). The 1980 OHCA field mapped with the LEV method using WOD quality control flags (Fig. 5h) has fewer small-scale features than that using EN3v2a quality control (Fig. 5c), although there are some anomalous features in Fig. 5h not in Fig. 5c (viz., the northern North Atlantic). None of the anomalous features are found



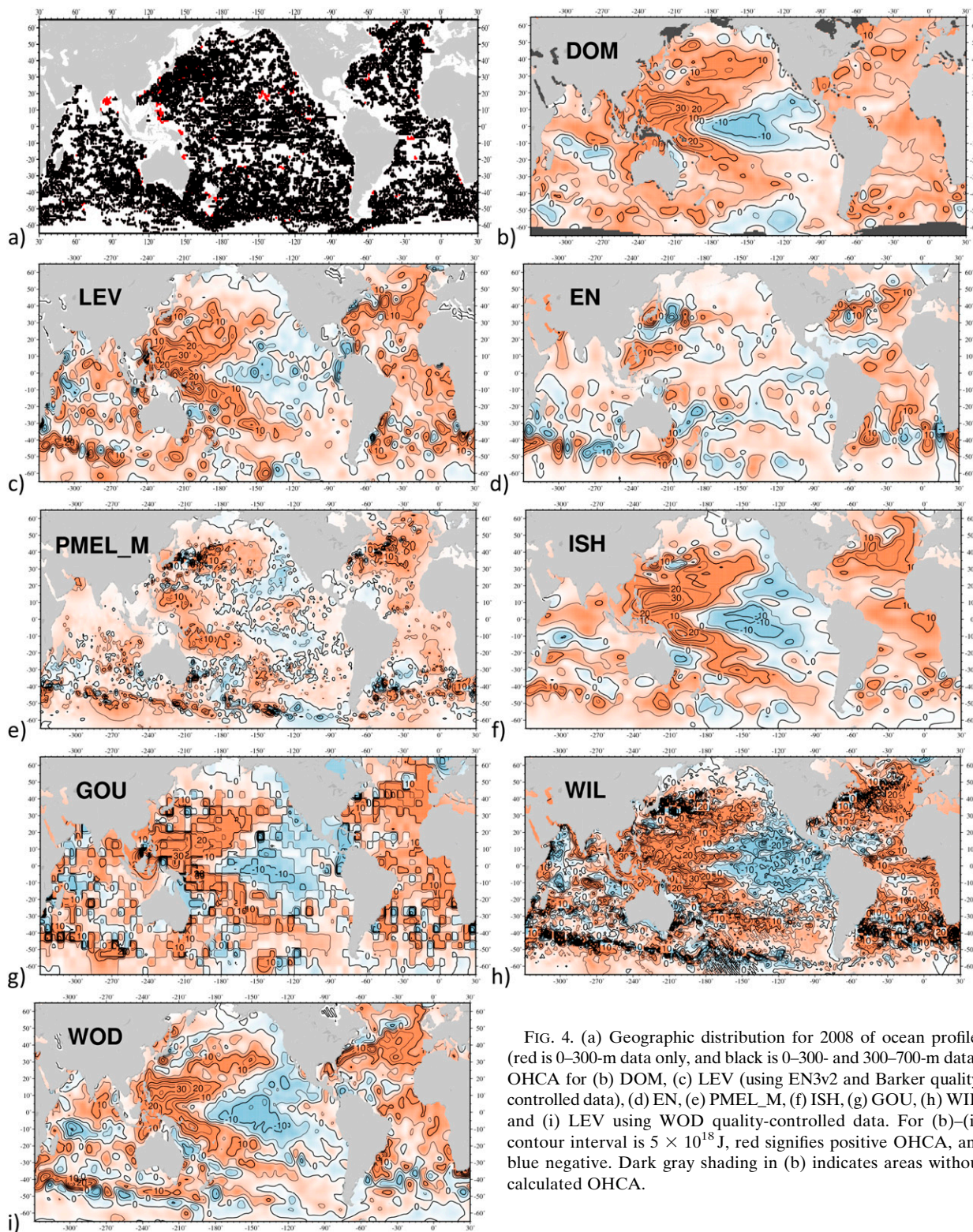


FIG. 4. (a) Geographic distribution for 2008 of ocean profiles (red is 0–300-m data only, and black is 0–300- and 300–700-m data). OHCA for (b) DOM, (c) LEV (using EN3v2 and Barker quality-controlled data), (d) EN, (e) PMEL\_M, (f) ISH, (g) GOU, (h) WIL, and (i) LEV using WOD quality-controlled data. For (b)–(i), contour interval is  $5 \times 10^{18}$  J, red signifies positive OHCA, and blue negative. Dark gray shading in (b) indicates areas without calculated OHCA.



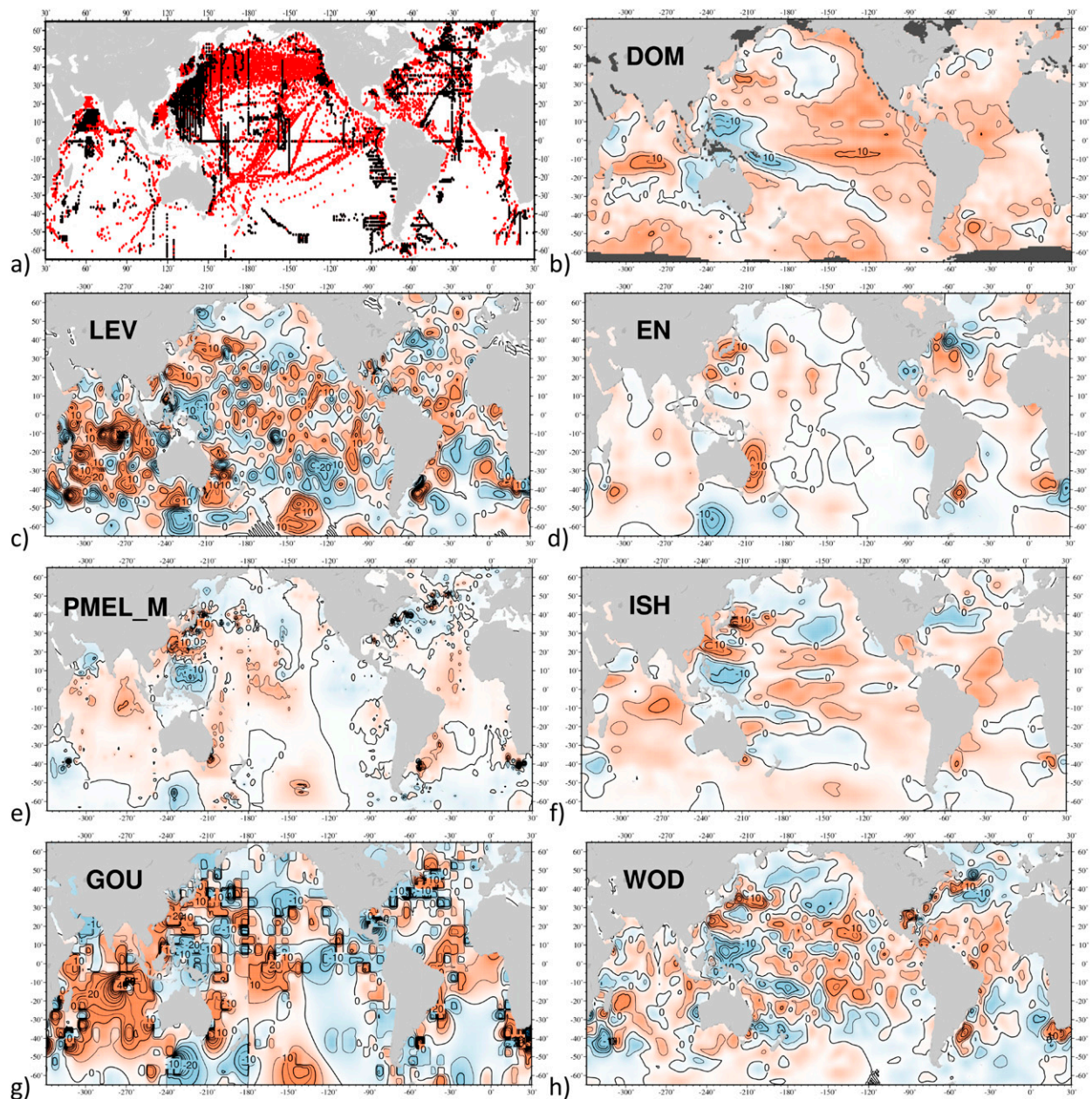


FIG. 5. As in Fig. 4, but for 1980 and with (a)–(h) only (without WIL mapping).

using the DOM mapping (Fig. 5b) or are as pronounced in the other mappings, which use large-spatial-scale bins or give less weight to spatially isolated profiles. Quality control of in situ temperature profile data is not explored further here as a source of difference in OHCA.

#### b. Effects of XBT corrections

Here we explore the effect of each XBT correction on the global integral of OHCA by mapping method (Fig. 6) relative to the C1\_H baseline climatology. Yearly values for different XBT corrections can vary

widely using the same mapping method. The W08 XBT correction is a clear outlier from 1999 to 2001 in most panels (e.g., Figs. 6a,c,d,e,g,h). Part of this deviation is explained by the use of different dataset versions. At least for DOM, the OHCA increase over 2000–04 when using the EN3v1d data (Domingues et al. 2008) is about 50% smaller than observed in Fig. 6a (based on EN3v2a data). For the LEV estimate, however, these differences are smaller (not shown). An investigation of the differences between EN3v1d and EN3v2a would be needed to understand these changes.

EDOFs for XBT correction for each mapping method are approximately 4 for years 1970–2008 and 3 for 1993–2008, less than the number of mapping methods investigated (seven for 1970–2008 and eight for 1993–2008; Table 1). Some mapping methods exhibit larger standard deviations for estimates using different XBT corrections in a given year, showing more sensitivity to the XBT correction used (Fig. 7). The DOM mapping method is the most sensitive to XBT correction method used, while ISH and PMEL\_M are the least sensitive, based on the mean standard deviation (Fig. 7). Again, we use the standard deviation of different XBT corrections for a given mapping method as a measure of the variations inherent in estimates due to XBT correction used for each mapping method (Fig. 7). The mean standard deviations for different XBT corrections applied to each mapping method are calculated excluding the years 2005–08, for which only Argo data were used. The mean standard deviation for XBT corrections for the ensemble of mapping methods is 12.1 ZJ for the period 1970–2004 and 14.9 ZJ for the period 1993–2004. Excluding years 1999–2001 because of the outliers in W08, these values are 11.6 ZJ for 1970–2004 and 12.2 ZJ for 1993–2004. These values excluding years 1999–2001 will be used in comparing XBT correction uncertainty. XBT correction is a larger factor than mapping method for DOM for both time periods, although nearly equal for 1993–2008. This result could arise because the DOM mapping method focuses on large-scale patterns, possibly making it more susceptible to changes in sampling in data-sparse areas and also because of variations in XBT corrections in these areas. For all other mapping methods, XBT correction engenders a smaller mean uncertainty than do uncertainties due to mapping method. XBT correction uncertainty was calculated using the C1\_H baseline climatology. This uncertainty was also calculated for C2\_H and C3\_M baselines for the DOM and LEV mapping methods (not shown). There was very little difference in XBT correction uncertainty between the C1\_H and C2\_H cases for DOM and LEV. However, the C3\_M case produced smaller yearly standard deviations for some years (not shown)—namely, those years with outliers, such as 1999–2001 (discussed above for the W08 correction), and also years such as 1990–91 with one or more XBT correction outliers. Use of the C3\_M baseline climatology ameliorated the effects of XBT correction outliers while giving nearly identical standard deviation for years with smaller differences between corrections. The mean standard deviation values for different XBT corrections using the C3\_M baseline climatology for DOM mapping method are 15.9 ZJ for 1970–2004 (compared

with 16.7 ZJ, excluding years 1999–2001) and 17.1 ZJ for 1993–2004 (compared with 18.9 ZJ, again excluding years 1999–2001). For LEV, mean standard deviations are 10.2 ZJ for 1970–2004 (11.1 ZJ using C1\_H) and 8.8 ZJ for 1993–2004 (10.4 ZJ using C1\_H). Why the C3\_M baseline lessens the effects of outlier XBT corrections is a question for further study.

### c. Effects of baseline climatologies

The final variable in the calculation of OHCA that we examine with respect to the different mapping methods is baseline climatology. For all mapping methods, a baseline monthly mean of temperature or sea level is subtracted from in situ temperature observations to calculate an anomaly value that is then used to map OHCA. All OHCA calculations to this point use the C1\_H climatological baseline mean. C1\_H uses a baseline temperature climatology calculated from all bottle, CTD, and Argo profiling floats for all available years. The climatological monthly means are constructed with a model that includes a linear trend with time to reduce temporal biases in the climatology owing to long-term temperature changes (Alory et al. 2007). For comparison, a similarly calculated baseline is used in C2\_H. C2\_H uses a baseline temperature climatology calculated in the same way as Alory et al. (2007), but with the addition of XBT data corrected using W08 estimates of fall-rate corrections. C2\_H has the benefit of greatly increased spatial and temporal data coverage with the disadvantage of the uncertainty introduced by the XBT bias and choice of correction. Cheng and Zhu (2015) recommend using an Argo-era climatology rather than a longer, historical time period. This recommendation arises because of the full geographical and temporal sampling during the Argo era compared with the nongeographically uniform time period (e.g., Southern Hemisphere with a different mean year of all observations relative to the Northern Hemisphere) found in historical data climatologies. The C1\_H and C2\_H climatologies calculate a trend factor for each ocean grid to make the mean year geographically uniform, mitigating this issue (Alory et al. 2007). To address differences due to the use of temporally limited data coverage representing the Argo period, baseline mean C3\_M uses the *World Ocean Atlas 2013* (WOA13) climatology from the 2005–12 time period (Locarnini et al. 2013). All instrument types are used, including XBTs corrected using the L09 method, but the time period is dominated by Argo measurements. OHCA in the 2005–12 time period is warmer than over the rest of the 1970–2008 time period under consideration and can have an effect on the results from some mapping methods (Lyman



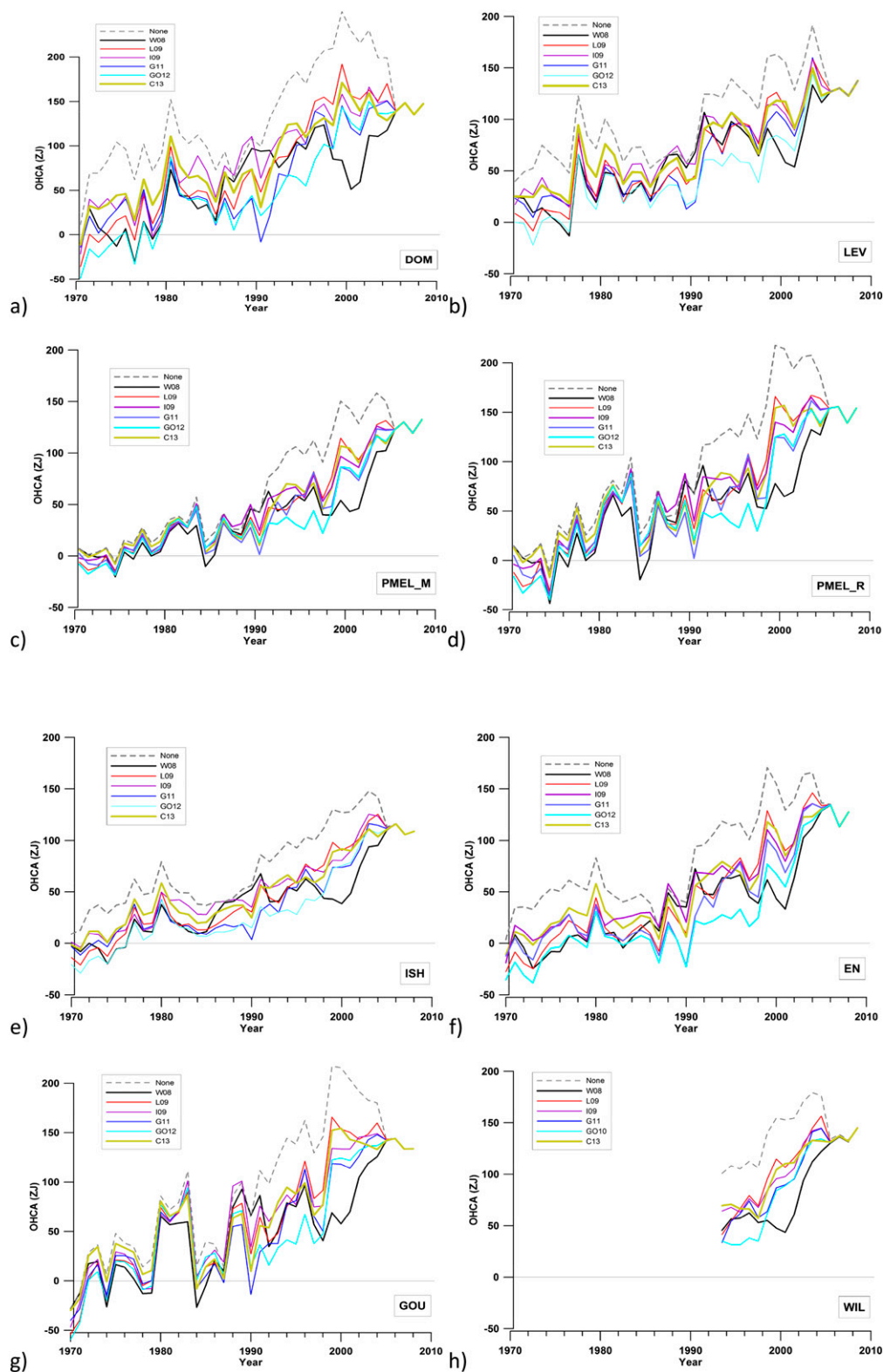


FIG. 6. Annual globally integrated OHCA (ZJ) time series for 1970–2008 for mappings using: (a) DOM, (b) LEV, (c) PMEL\_M, (d) PMEL\_R, (e) ISH, (f) EN, (g) GOU, and (h) WIL for 1993–2008, relative to the C1\_H baseline climatology, using different XBT corrections: none (dashed gray), **W08** (black), **L09** (red), **I09** (purple), **G11** (blue), **G012** (cyan), and **C13** (yellow). The y axis is from  $-50$  to  $200$  ZJ.

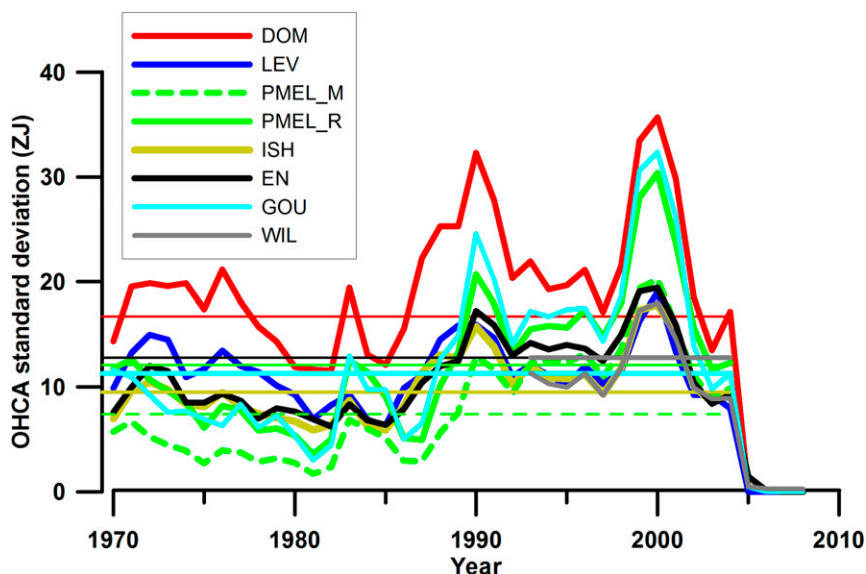


FIG. 7. Standard deviations of the yearly ensemble means of global OHCA (ZJ) for 1970–2008 using different XBT corrections for each mapping method (see legend) from 1970–2008 referenced to the C1\_H baseline climatology. Horizontal lines represent the mean standard deviations over the full time series.

and Johnson 2014). The global integral of OHC for CM\_3 is 157 ZJ warmer than CH\_1, while that for CH\_2 is only 16 ZJ warmer than CH\_1.

For simplicity, only the W08 XBT correction case is examined. Results using other XBT correction cases were nearly identical (in yearly standard deviation, means within 0.1 ZJ) for the DOM and LEV methods for each of the XBT correction cases but could possibly be different for other mapping methods (not calculated). We choose to adjust results from the three climatological cases for comparison. Since 2008 has the widest and most homogenous spatial and temporal coverage (and does not include XBT data) we adjust OHCA estimates for C2\_H and C3\_M by subtracting out the difference in OHCA between these cases and the C1\_H case for year 2008. The 2008 differences from C1\_H for C2\_H are less than 18 ZJ, but those for C3\_M are larger, being 167 ZJ for DOM, 103 ZJ for LEV, 118 ZJ for ISH, 154 ZJ for PMEL\_M, 167 ZJ for PMEL\_R, 154 ZJ for EN, 152 ZJ for GOU, and 149 ZJ for WIL.

After these 2008 offsets are applied, the mapping methods are still affected to differing degrees by the change in baseline climatology (Figs. 8a–g). For all mapping methods except GOU, the difference between OHCA using C1\_H and C2\_H is small, less than 5 ZJ for most years, never exceeding 16 ZJ. For GOU in the 1990s and early 2000s, the difference in OHCA between C1\_H and C2\_H cases is greater than 25 ZJ for most years. It is not clear why this occurs only with the

GOU mapping method. Differences between OHCA using the C1\_H and C3\_M baseline climatologies are more pronounced for all mapping methods than between the C1\_H and C2\_H cases. For the DOM, LEV, and PMEL\_R mapping methods, the C3\_M OHCA yearly values are generally lower than the C1\_H values, consistent with Cheng and Zhu (2015), more so for the DOM mapping method than the other two. For the PMEL\_M, ISH, and EN mapping methods, the C3\_M OHCA yearly values are higher than for the C1\_H OHCA values and the difference is larger for years earlier in the time series. The effect of the baseline mean on the PMEL\_M and ISH, methods, and variants on zero infill (populating data void areas with values which represent the absence of change), with the warmer baseline (2005–12 in C3\_M) result in smaller linear trends than the cooler historical climatologies (C1\_H, C2\_H), similar to the results of the experiment documented in Lyman and Johnson (2014).

The EDOF for baseline climatology comparison for each mapping method (Table 2) is very low, approximately 1. With that result in mind we turn to standard deviations of the estimates using different climatologies for the same mapping method (Fig. 9). WIL and LEV show a low sensitivity to baseline climatology, with LEV having mean standard deviations of 3.5 and 2.7 ZJ for 1970–2008 and 1993–2008, respectively (Table 3), and WIL having a 3.1-ZJ mean standard deviation for 1993–2008. PMEL\_R and ISH, have mean standard deviations

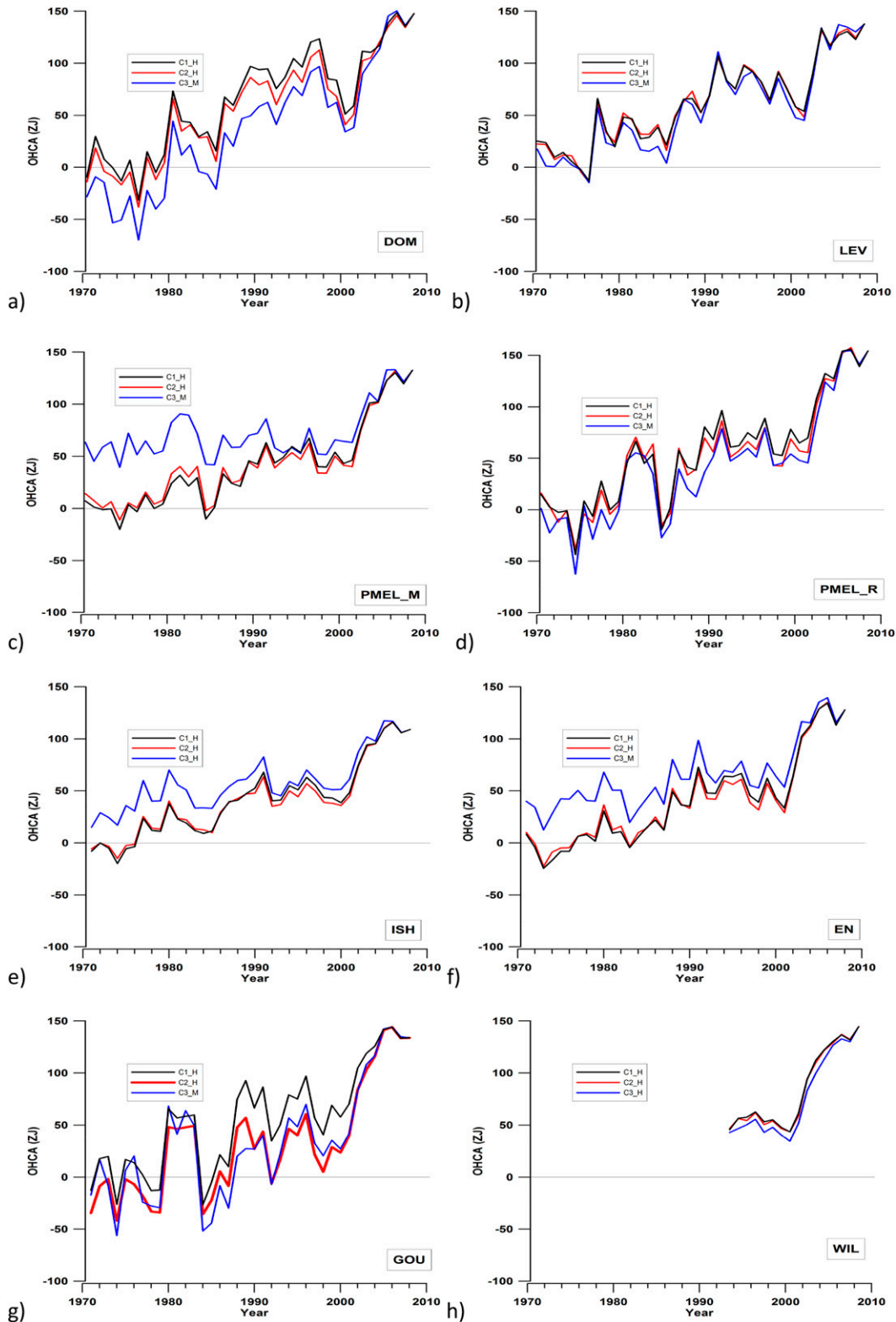


FIG. 8. Annual globally integrated OHCA (ZJ) time series for 1970–2008, using the W08 XBT correction relative to three climatological mean baselines—C1\_H (black), C2\_H (red), and C3\_H (blue)—for mappings using: (a) DOM, (b) LEV, (c) PMEL\_M, (d) PMEL\_R, (e) ISH, (f) EN, (g) GOU, and (h) WIL for 1993–2008. Individual curves have been shifted to overlie each other in 2008, the year with the best coverage (Fig. 3). The y axis is from  $-100$  to  $150$  ZJ.



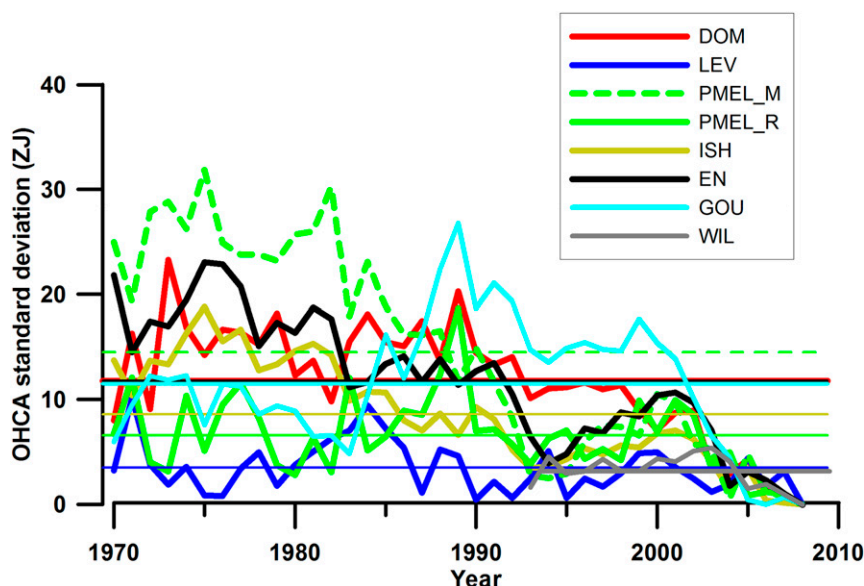


FIG. 9. Standard deviations of the yearly ensemble means of global OHCA (ZJ) for 1970–2008 based on the three different baseline climatologies (C1\_H, C2\_H, and C3\_H) for each mapping method (see legend). Horizontal lines represent mean standard deviations over the full time period.

due to changes in baseline climatology of 6.6 and 8.6 ZJ, respectively, for 1970–2008, higher than LEV but lower than for the other mapping methods. The results for ISH are similar to the results reported in Ishii and Kimoto (2009). DOM, EN, GOU, and PMEL\_M all have mean standard deviations due to baseline climatology greater than 10.0 ZJ for 1970–2008. The standard deviation of the GOU method has the largest differences between C1\_H and C2\_H cases in the 1993–2008 period.

Standard deviations are constructed here only from the three baseline cases for each mapping method for each year, a very small sample size (approximately 1 EDOF; Table 2), and are subject to our choice to force agreement at 2008, which impacts both the magnitudes

and temporal structures of the standard deviations. Subject to these caveats, the effects of different baseline mean in all mapping methods are smaller than the difference between the mapping methods themselves. However, relative importance of difference due to XBT correction and to baseline climatology is more year dependent, with XBT correction uncertainty larger in later years (1990s and 2000s) and baseline climatology uncertainty larger in earlier years (1970s and 1980s) for some mapping methods with zero infill variants (PMEL\_M and ISH) as well as EN. Therefore, the choice of the baseline climatology is of importance in estimating the uncertainty in OHCA with a widely varying dependence on the mapping method used. The choice of a latter-year

TABLE 3. Time-mean standard deviations of global OHCA (ZJ) associated with variations in XBT bias corrections and baseline climatologies for different mapping methods for 1970–2008 and 1993–2008 (in parentheses). The asterisk indicates that the WIL method can only be calculated from 1993–2008. The XBT bias correction standard deviation is only through 2004, since XBTs were not used in this experiment beyond that year; and also excludes years 1999–2001 because of obvious outliers in the W08 method for those years.

	Mapping method	XBT bias correction	Baseline climatology
DOM	16.5 (17.1)	16.7 (18.9)	11.8 (7.1)
LEV	16.5 (17.1)	11.1 (10.4)	3.5 (2.7)
PMEL_M	16.5 (17.1)	7.4 (11.5)	14.5 (4.7)
PMEL_R	16.5 (17.1)	12.1 (15.3)	6.6 (4.9)
ISH	16.5 (17.1)	9.5 (10.5)	8.6 (3.9)
EN	16.5 (17.1)	12.8 (15.1)	11.8 (5.6)
GOU	16.5 (17.1)	11.3 (15.2)	11.5 (9.8)
WIL*	— (17.1)	— (12.8)	— (3.1)
Average	16.5 (17.1)	11.6 (12.2)	9.8 (5.2)

TABLE 4. Linear trends of OHCA ( $\text{ZJ yr}^{-1}$ ) for 1970–2008 and 1993–2008 by mapping and test case with standard errors and degrees of freedom (in parentheses). Degrees of freedom are estimated as described in the text. The W08 case is equivalent to the C1\_H case. The none case is not included in the mean value in the final row. The uncertainty of the mean in the last row and column is the standard deviation of the included cases.

1970–2008									
	DOM	LEV	PMEL_M	PMEL_R	ISH	EN	GOU	Mean	
W08	3.9 ± 0.65 (9.7)	3.1 ± 0.48 (12.7)	3.2 ± 0.61 (7.1)	3.9 ± 0.65 (10.8)	2.9 ± 0.47 (7.5)	3.4 ± 0.60 (7.7)	3.7 ± 0.70 (11.8)	3.4 ± 0.4	
L09	5.0 ± 0.81 (6.8)	3.7 ± 0.44 (14.2)	3.9 ± 0.67 (4.4)	4.9 ± 0.75 (7.9)	3.5 ± 0.55 (4.8)	4.2 ± 0.78 (5.0)	4.8 ± 0.78 (10.7)	4.3 ± 0.6	
C13	3.7 ± 0.66 (8.7)	3.0 ± 0.44 (13.6)	3.5 ± 0.67 (4.9)	4.2 ± 0.77 (8.0)	2.9 ± 0.50 (5.1)	3.4 ± 0.70 (5.7)	4.0 ± 0.79 (10.1)	3.5 ± 0.5	
I09	4.1 ± 0.42 (17.0)	3.1 ± 0.35 (17.0)	3.7 ± 0.44 (7.6)	4.6 ± 0.48 (14.2)	3.1 ± 0.31 (8.97)	3.6 ± 0.48 (8.0)	4.4 ± 0.59 (14.7)	3.8 ± 0.6	
G11	4.1 ± 1.0 (5.3)	3.3 ± 0.69 (7.4)	3.5 ± 0.73 (5.1)	4.4 ± 0.85 (7.1)	3.1 ± 0.83 (3.3)	3.6 ± 1.1 (3.8)	4.1 ± 1.0 (7.9)	3.7 ± 0.4	
GO12	4.8 ± 0.75 (7.0)	3.7 ± 0.44 (13.4)	3.4 ± 0.91 (3.7)	4.3 ± 1.11 (5.2)	3.4 ± 0.79 (3.3)	3.7 ± 1.1 (3.5)	4.1 ± 0.87 (9.1)	3.9 ± 0.5	
None	3.9 ± 1.4 (5.5)	2.9 ± 0.78 (7.3)	4.4 ± 0.67 (6.2)	5.3 ± 0.95 (7.4)	3.1 ± 0.66 (5.2)	3.5 ± 1.0 (4.9)	5.1 ± 1.1 (7.9)	4.0 ± 0.9	
C2_H	3.9 ± 0.65 (9.7)	3.1 ± 0.49 (12.7)	2.9 ± 0.69 (6.6)	3.7 ± 0.78 (9.0)	2.7 ± 0.53 (6.7)	3.2 ± 0.66 (7.2)	3.7 ± 0.70 (11.8)	3.3 ± 0.4	
C3_M	4.7 ± 0.70 (9.8)	3.4 ± 0.61 (10.2)	1.3 ± 0.74 (5.8)	4.0 ± 0.82 (9.1)	2.0 ± 0.51 (7.2)	2.3 ± 0.58 (7.9)	3.6 ± 1.1 (9.0)	3.0 ± 1.1	
Mean	4.2 ± 0.4	3.3 ± 0.3	3.3 ± 0.8	4.4 ± 0.5	3.0 ± 0.4	3.4 ± 0.5	4.2 ± 0.5		
1993–2008									
	DOM	LEV	PMEL_M	PMEL_R	ISH	EN	GOU	WIL	Mean
W08	3.3 ± 3.4 (2.3)	4.0 ± 2.5 (3.0)	6.2 ± 3.0 (1.8)	7.0 ± 2.7 (2.4)	5.1 ± 2.4 (1.9)	5.8 ± 3.2 (2.2)	6.0 ± 2.7 (2.5)	7.1 ± 2.7 (1.7)	5.6 ± 1.3
L09	2.3 ± 2.9 (2.5)	3.9 ± 1.5 (5.2)	5.9 ± 1.1 (5.3)	6.7 ± 2.8 (3.0)	4.8 ± 1.2 (3.5)	5.6 ± 1.7 (4.6)	5.3 ± 2.9 (3.3)	7.0 ± 1.5 (3.2)	5.2 ± 1.4
C13	1.5 ± 1.6 (3.6)	3.2 ± 1.3 (5.5)	5.1 ± 1.2 (4.9)	5.7 ± 2.3 (3.8)	4.2 ± 0.88 (3.8)	4.7 ± 1.6 (4.5)	4.5 ± 0.23 (3.6)	6.1 ± 1.3 (3.0)	4.4 ± 1.4
I09	2.0 ± 1.3 (4.5)	3.3 ± 1.4 (5.0)	5.5 ± 1.0 (5.4)	6.1 ± 2.3 (3.0)	4.4 ± 1.0 (3.59)	5.0 ± 1.4 (4.6)	5.0 ± 1.8 (4.4)	6.3 ± 1.3 (2.9)	4.7 ± 1.4
G11	3.7 ± 0.95 (14)	4.2 ± 1.5 (5.0)	6.3 ± 1.3 (5.1)	7.1 ± 1.7 (5.3)	5.4 ± 1.1 (4.2)	6.2 ± 1.7 (4.8)	6.2 ± 1.8 (0.78)	7.5 ± 1.3 (3.6)	5.8 ± 1.2
GO12	6.0 ± 2.0 (2.8)	5.7 ± 1.5 (4.8)	8.0 ± 1.3 (4.7)	9.4 ± 2.4 (3.7)	6.7 ± 1.2 (3.1)	8.7 ± 1.7 (4.2)	8.4 ± 2.3 (3.7)	9.0 ± 1.6 (2.7)	7.7 ± 1.3
None	−2.8 ± 5.4 (1.6)	0.7 ± 2.2 (4.0)	2.5 ± 2.0 (3.2)	2.2 ± 5.0 (1.8)	1.8 ± 2.4 (2.0)	1.8 ± 2.2 (3.6)	0.53 ± 4.7 (2.0)	3.4 ± 3.0 (2.1)	1.3 ± 1.8
C2_H	4.2 ± 3.7 (2.1)	4.0 ± 2.6 (3.0)	6.6 ± 3.1 (1.8)	7.7 ± 3.1 (14)	5.6 ± 2.4 (1.9)	6.4 ± 3.2 (2.2)	6.0 ± 2.7 (2.5)	7.4 ± 3.1 (1.6)	6.0 ± 1.3
C3_M	5.6 ± 4.0 (2.0)	4.8 ± 3.1 (2.6)	6.0 ± 1.9 (2.5)	7.9 ± 3.6 (2.0)	4.9 ± 1.8 (2.3)	5.5 ± 2.2 (2.9)	8.2 ± 3.9 (1.9)	7.2 ± 2.8 (1.7)	6.3 ± 1.2
Mean	3.6 ± 1.5	4.1 ± 0.8	6.2 ± 0.8	7.2 ± 1.1	5.1 ± 0.7	6.0 ± 1.2	6.2 ± 1.3	7.2 ± 0.8	

(Argo period) climatology (recommended in [Cheng and Zhu 2015](#)) may be appropriate for some mapping methods but will cause larger uncertainties for methods with variants on zero infill if applied to data-sparse years. For other mapping methods, such as LEV and WIL, the uncertainties due to baseline climatology are small in comparison to other uncertainties. A further study, resolving the caveats listed above, would help to clarify the results given here.

#### d. Linear trend

Finally, we consider linear trend estimates for OHCA in the global ocean ([Table 4](#) and [Fig. 10](#)) for each mapping method and each case (XBT correction and baseline climatology) examined here. Here standard errors are given for linear trend uncertainties. Degrees of freedom are estimated for each time series by dividing the time series length by its decorrelation time scale and then subtracting two. The decorrelation time scale for each time series is estimated as twice the maximum value of the lagged autocorrelation of the residual computed by subtracting the linear trend from the time series ([von Storch and Zwiers 1999](#)). For 1970–2008 [[Table 4](#) (top) and [Fig. 10a](#)], the DOM linear trend is highest in five of the eight test cases including each baseline climatology, while PMEL\_R linear trend is highest in four (both have the same value for the W08 C1\_H case). The GOU, DOM, and PMEL\_R mapping methods form a distinct

group with an average linear trend over all cases greater than  $+4.0 \text{ ZJ yr}^{-1}$ . These results make sense for PMEL\_R and GOU since they both explicitly infill data-sparse regions with the average of well-sampled regions when constructing a global mean. PMEL\_M, ISH, and EN all reach their lowest value using the C3\_M baseline, the warmest climatology. This result makes sense as well since these mapping methods will tend to underestimate trends using warm climatologies when earlier, colder years are data sparse ([Lyman and Johnson 2014](#)). ISH consistently exhibits the lowest trend value for all XBT correction cases, excepting C3\_M. ISH, LEV, PMEL\_M, and EN form a group of mapping methods whose average linear trend over all eight cases is less than  $+3.5 \text{ ZJ yr}^{-1}$ . Finally, all cases for all mapping methods exhibit a warming trend within the standard error over the period 1970–2008, even the PMEL\_M results using the C3\_M baseline case ([Fig. 10](#)), which has the lowest linear trend of all cases at  $1.3 \pm 0.74 \text{ ZJ yr}^{-1}$ . The reasons for that low result have been carefully explained, along with the reason PMEL\_R is preferable for global integrals of OHCA ([Lyman and Johnson 2014](#)).

For the period 1993–2008 [[Table 4](#) (bottom) and [Fig. 10b](#)] WIL has the highest linear trend for all cases using the C1\_H baseline, except the GO12 XBT correction case, where PMEL\_R is higher. WIL also has the highest linear trend for the C2\_H case, while GOU has the highest linear trend for the C3\_M case. The highest

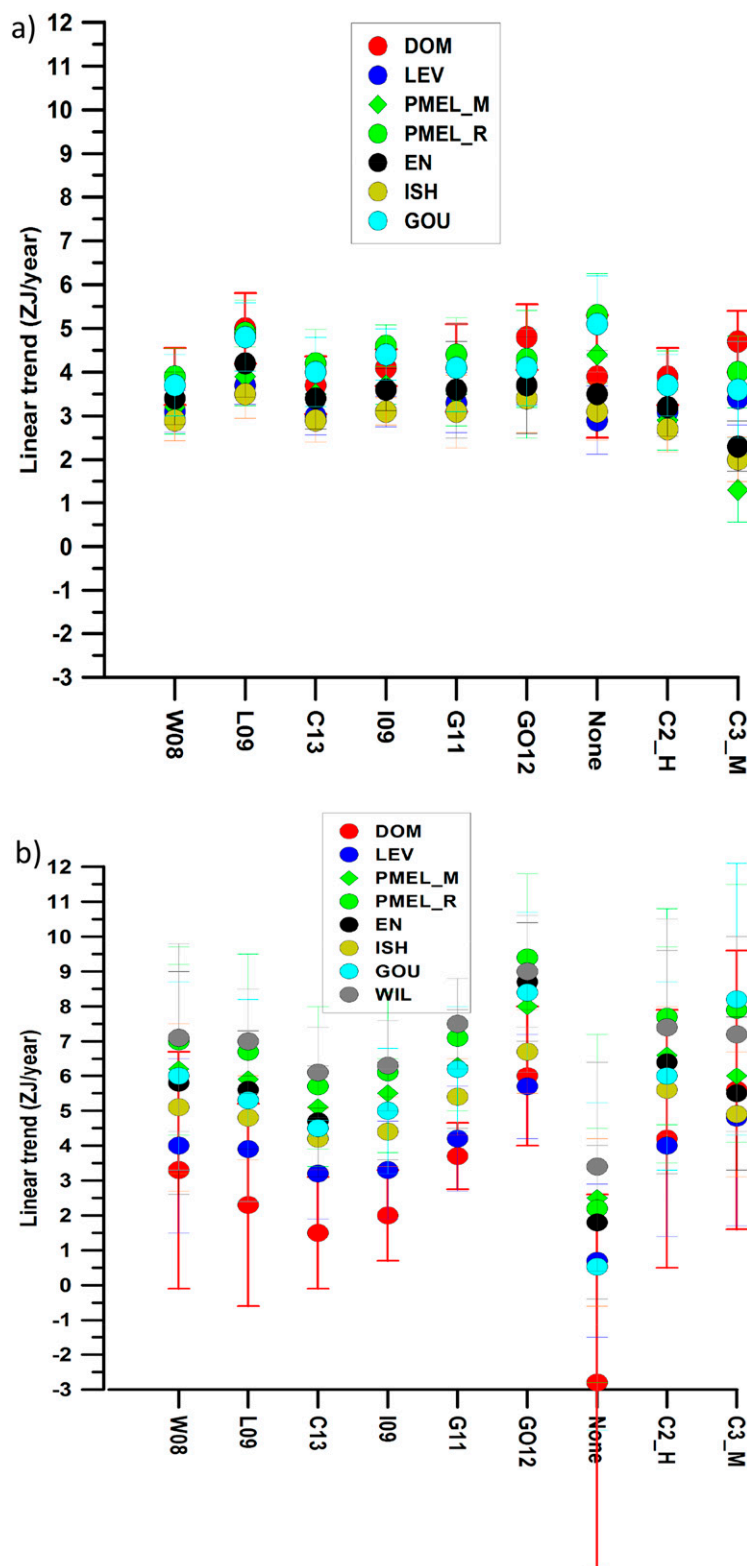


FIG. 10. Linear trends (filled circles) of annual globally integrated OHCA ( $\text{ZJ yr}^{-1}$ ) with standard errors of the mean (error bars) for (a) 1970–2008 and (b) 1993–2008 by mapping method (see legends) and for all test cases (see horizontal axis labels). The C2\_H and C3\_M cases used W08 XBT bias corrections. The W08, L09, C13, I09, G11, GO12, and none cases used C1\_H baseline climatology.

linear trend in the study is the PMEL\_R C1\_M case with the GO12 XBT correction ( $9.4 \pm 2.4 \text{ ZJ yr}^{-1}$ ). For 1993–2008, PMEL\_M, EN, and GOU all have average linear trends between  $+6.0$  and  $+6.2 \text{ ZJ yr}^{-1}$ , while DOM, LEV, and ISH all have average linear trends less than or equal to  $+5.1 \text{ ZJ}$ . PMEL\_R and WIL have the highest trend, both at  $+7.2 \text{ ZJ yr}^{-1}$  for 1993–2008. The DOM average linear trend for 1993–2008 over all cases is the lowest of all mapping methods, at  $+3.6 \pm 1.5 \text{ ZJ yr}^{-1}$ , while the DOM method produces the second highest average linear trend over 1970–2008 (with PMEL\_R highest), with ISH lowest for 1970–2008. While the standard error of the mean for cases over 1993–2008 are higher than the same cases over 1970–2008, warming trends are all positive within the standard error, except the W08, L09, and C13 (C1\_H) cases for the DOM method ( $3.3 \pm 3.4 \text{ ZJ yr}^{-1}$ ,  $2.3 \pm 2.9 \text{ ZJ yr}^{-1}$ , and  $1.5 \pm 1.6 \text{ ZJ yr}^{-1}$ , respectively).

For some perspective, the recent estimate of OHCA linear trend over 1970–2005 of  $5.6 \pm 0.15 \text{ ZJ yr}^{-1}$  from Cheng et al. (2015) is higher than any of the present results, although within uncertainty of all DOM, PMEL\_R, and GOU cases and all but one historical baseline climatology PMEL\_M case. This is due to a combination of all factors tested here (mapping method, XBT correction, and baseline mean) as well as input dataset/quality control. There are also three fewer years in Cheng et al. (2015), but their 1970–2014 linear trend ( $5.5 \pm 0.15 \text{ ZJ}$ ) is nearly identical to their 1970–2005 trend.

Of the five estimates of 1993–2014 OHCA linear trend from Johnson et al. (2015), all are within uncertainty of the most similar case from the present work for 1993–2008, with four higher linear trends [their PMEL–JPL–Joint Institute for Marine and Atmospheric Research (JIMAR) compared with PMEL\_R, their NODC compared with LEV, their CSIRO–ACE CRC–IMAS–University of Tasmania (UTAS) compared with DOM, and their Met Office Hadley Centre compared with EN] and one lower linear trend [their Meteorological Research Institute (MRI)–JMA compared with ISH]. Differences are due in part to the different time period covered but also to factors such as baseline climatology and dataset/quality control used.

#### 4. Conclusions

By holding input data constant, we isolate the sensitivity of global OHCA estimates in the upper 700 m of the ocean over a 39-yr period, 1970–2008, to mapping method, XBT bias correction, and baseline climatology.

Annual differences of roughly  $17.1 \text{ ZJ}$  in OHCA estimates can be attributed to differences in mapping methods used by seven groups (eight methods) actively performing such calculations (Table 3). These differences

in calculations are not simply due to infilling of data gaps but also to smoothing or filtering of fields geographically where data are present. The smoothing or filtering inherent in each mapping method is performed over different length scales. Quality flags for the input data will also affect calculated OHCA differently based on the length scale and amount of smoothing of each mapping method.

The choice of XBT bias correction has an effect on differences in calculated OHCA, as previously reported by Lyman et al. (2010). However, in the present study, we find that mapping method is generally the largest source of uncertainty (in six of seven calculations for 1970–2008 and five of eight calculations for 1993–2008, as shown in Table 3). Only for the DOM mapping method does the choice of XBT correction have a larger impact on OHCA estimates over both periods than does the overall choice of mapping method. For other mapping methods, the choice of XBT correction has a smaller impact for 1970–2008 (between  $8.0$  and  $12.6 \text{ ZJ}$ ; Table 3) than the choice of mapping method ( $16.5 \text{ ZJ}$ ). The relative uncertainty of mapping method and XBT correction for 1993–2008 is somewhat more dependent on mapping method, with DOM, PMEL\_R, and GOU having higher uncertainties for XBT correction than mapping method uncertainty. The XBT community is approaching, but has not yet reached, a consensus on the annually varying corrections of both depth (probe fall rate) and temperature to minimize the uncertainty of temperature measurements associated with XBT corrections (Cheng et al. 2016). Once a consensus is reached, this particular uncertainty should be further reduced (but not eliminated) as a source of uncertainty in historical and future OHCA calculations.

The choice of baseline climatological monthly mean also has an effect on OHCA calculations that is different for the different mapping methods. As Lyman and Johnson (2014) describe, for some mapping methods, this may be due to the relation between the infilling method and the warm time period over which the 2005–12 climatology was calculated. The warmer climatology will engender larger cool anomalies in earlier years (compared to a historical mean climatology), the effects of which are damped in the integral by infilling zero anomaly in data-sparse regions. Taking into account the small 2008 offset compared to total heat content offset between C1\_H and C3\_M climatologies ( $103$  vs  $157 \text{ ZJ}$ ), the LEV mapping method is not very sensitive to the baseline climatology used, despite using a variation on relaxation to climatology as infill method. For LEV, this insensitivity may be due to the more liberal definition used for the area over which the mapping method can incorporate data for each grid

box (Fig. 3). Choosing a proper baseline climatology can be method dependent, as the PMEL\_M method should not be (and is not) employed with a climatology consisting of recent data only, as per the reasons outlined in Lyman and Johnson (2014). For the DOM method, further investigation would reveal which baseline climatology produces the most realistic results, possibly on a regional level.

Mean linear trends for each mapping method from all cases listed in Table 1 (except none) range from +3.0 (ISH) to +4.4 ZJ yr<sup>-1</sup> (PMEL\_R) for the 1970–2008 period and from +3.6 (DOM) to +7.2 ZJ yr<sup>-1</sup> (WIL and PMEL\_R) for 1993–2008. Expressing these values in terms of watts per square meter, the unit in which the earth's heat budget is usually discussed, the trends are from 0.08 to 0.31 W m<sup>-2</sup> for 1970–2008 and from 0.09 to 0.58 W m<sup>-2</sup> for 1993–2008, using the global surface area of  $5.144 \times 10^{14}$  m<sup>2</sup>. These linear trends are small compared to the yearly uncertainties calculated here. However, all (XBT corrected) upper-OHCA estimates regardless of baseline climatology or mapping method show a distinctive increase over the 1970–2008 period and, except for the DOM method, a higher warming rate over the 1993–2008 period. Regardless of mapping method, OHCA calculations do reveal an increase in ocean heat content on both the multidecadal and decadal time scales.

We find that the uncertainty in calculation of OHCA from irregular spatial and temporal data distribution is larger than interannual variability in OHCA, but long-term trends are still statistically robust, in agreement with Lyman et al. (2010), even when extending the time series back to 1970 and including years with relatively sparse data distribution. The mapping method is shown here to be a source of uncertainty on the order of 17 ZJ. XBT corrections and baseline climatological mean add uncertainties dependent on the mapping method used, but at least the uncertainty due to the XBT correction should be reduced if a consensus can be reached on the optimum corrections. Furthermore, care is indicated in choosing the best baseline climatology, perhaps on a method-dependent basis.

The present work carries forward understanding of the uncertainties affecting in situ estimates of OHCA, but there is still much work to do. Future work will include the following:

- Refinement in understanding the role of the baseline climatology in OHCA uncertainty, with a fuller exploration of the relative effects of the more uniform spatial and seasonal coverage versus shorter time span inherent in Argo-float-era climatologies—for instance, a 10-yr Argo-period climatology constructed from subsampled model data is unbiased when compared to the

spatially complete model data for the same period, but a shorter-term Argo-era climatology results in larger errors than a longer-term climatology in some regions (Good 2016).

- Exploring the benefits and deficiencies in filtering OHCA signals at different spatial scales as well as differences due to nonseasonal signals (e.g., El Niño).
- Testing whether different grid sizes for averaging in situ data affect OHCA calculations (same mapping method, varying grid size).
- Quantifying the uncertainty introduced by dataset choices and quality control choices.
- Quantifying XBT correction schemes effectiveness compared with CTD reference data.
- Better understanding the effects of vertical sampling, both spacing between adjacent measurements and the absence or presence of data in different depth ranges.
- Augmenting the current historic temperature profile database especially for data-sparse regions.

In summary, although the present work helps to quantify the uncertainties in OHCA estimation from in situ oceanographic profile data, there is still much work to be done to refine the process and interpret the results.

*Acknowledgments.* TB was funded by the Climate Observations and Monitoring Program, National Oceanic and Atmospheric Administration, U.S. Department of Commerce. SG was supported by the Joint UK DECC/Defra Met Office Hadley Centre Climate Programme (GA01101). CMD was initially supported by the Australian Government's Business Cooperative Research Centres Programme through the Antarctic Climate and Ecosystems Cooperative Research Centre (ACE CRC) and subsequently by an Australian Research Council Future Fellowship (FT130101532). The research described in this paper was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. GCJ and JML were supported by NOAA Research and the NOAA Ocean Climate Observations Program. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA. Data used herein are available upon request.

## REFERENCES

- Abraham, J. P., J. Gorman, F. Reseghetti, E. M. Sparrow, and W. J. Mincowycz, 2012: Drag coefficients for rotating expendable bathythermographs and the impact of launch parameters on depth predictions. *Numer. Heat Transfer*, **62A**, 25–43, doi:10.1080/10407782.2012.672898.



- , and Coauthors, 2013: A review of global ocean temperature observations: Implications for ocean heat content estimates and climate change. *Rev. Geophys.*, **51**, 450–483, doi:[10.1002/rog.20022](https://doi.org/10.1002/rog.20022).
- AchutaRao, K. M., B. D. Santer, P. J. Gleckler, K. E. Taylor, D. W. Pierce, T. P. Barnett, and T. M. L. Wigley, 2006: Variability of ocean heat uptake: Reconciling observations and models. *J. Geophys. Res.*, **111**, C05019, doi:[10.1029/2005JC003136](https://doi.org/10.1029/2005JC003136).
- Alory, G., S. Wijffels, and G. M. Meyers, 2007: Observed temperature trends in the Indian Ocean over 1960–1999 and associated mechanisms. *Geophys. Res. Lett.*, **34**, L02606, doi:[10.1029/2006GL028044](https://doi.org/10.1029/2006GL028044).
- Barker, P. M., J. R. Dunn, C. M. Domingues, and S. E. Wijffels, 2011: Pressure sensor drifts in Argo and their impacts. *J. Atmos. Oceanic Technol.*, **28**, 1036–1049, doi:[10.1175/2011JTECH0831.1](https://doi.org/10.1175/2011JTECH0831.1).
- Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *J. Appl. Meteor.*, **3**, 396–409, doi:[10.1175/1520-0450\(1964\)003<0396:ATFMDI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1964)003<0396:ATFMDI>2.0.CO;2).
- Bell, M. J., R. M. Forbes, and A. Hines, 2000: Assessment of the FOAM global data assimilation system for real-time operational ocean forecasting. *J. Mar. Syst.*, **25**, 1–22, doi:[10.1016/S0924-7963\(00\)00005-1](https://doi.org/10.1016/S0924-7963(00)00005-1).
- Bindoff, N. L., and Coauthors, 2007: Observations: Oceanic climate change and sea level. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 385–432. [Available online at <https://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-chapter5.pdf>.]
- Boyer, T. P., and Coauthors, 2006: World ocean database 2005. NOAA Atlas NESDIS 60, 182 pp.
- Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Blade, 1999: The effective number of spatial degrees of freedom of a time-varying field. *J. Climate*, **12**, 1990–2009, doi:[10.1175/1520-0442\(1999\)012<1990:TENOSD>2.0.CO;2](https://doi.org/10.1175/1520-0442(1999)012<1990:TENOSD>2.0.CO;2).
- Bringas, F., and G. Goni, 2015: Early dynamics of Deep Blue XBT probes. *J. Atmos. Oceanic Technol.*, **32**, 2253–2263, doi:[10.1175/JTECH-D-15-0048.1](https://doi.org/10.1175/JTECH-D-15-0048.1).
- Cheng, L., and J. Zhu, 2014a: Uncertainties of the ocean heat content estimation induced by insufficient vertical resolution of historical ocean subsurface observations. *J. Atmos. Oceanic Technol.*, **31**, 1383–1396, doi:[10.1175/JTECH-D-13-00220.1](https://doi.org/10.1175/JTECH-D-13-00220.1).
- , and —, 2014b: Artifacts in variations of ocean heat content induced by the observation system changes. *Geophys. Res. Lett.*, **41**, 7276–7283, doi:[10.1002/2014GL061881](https://doi.org/10.1002/2014GL061881).
- , and —, 2015: Influences of the choice of climatology on ocean heat content estimation. *J. Atmos. Oceanic Technol.*, **32**, 388–394, doi:[10.1175/JTECH-D-14-00169.1](https://doi.org/10.1175/JTECH-D-14-00169.1).
- , —, R. Cowley, T. Boyer, and S. Wijffels, 2014: Time, probe type, and temperature variable bias corrections to historical expendable bathythermograph observations. *J. Atmos. Oceanic Technol.*, **31**, 1793–1825, doi:[10.1175/JTECH-D-13-00197.1](https://doi.org/10.1175/JTECH-D-13-00197.1).
- , —, and J. Abraham, 2015: Global upper ocean heat content estimation: Recent progress and the remaining challenges. *Atmos. Oceanic Sci. Lett.*, **8**, 333–338, doi:[10.3878/AOSL20150031](https://doi.org/10.3878/AOSL20150031).
- , and Coauthors, 2016: XBT science: Assessment of instrumental biases and errors. *Bull. Amer. Meteor. Soc.*, doi:[10.1175/BAMS-D-15-00031.1](https://doi.org/10.1175/BAMS-D-15-00031.1), in press.
- Church, J. A., and Coauthors, 2011: Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. *Geophys. Res. Lett.*, **38**, L18601, doi:[10.1029/2011GL048794](https://doi.org/10.1029/2011GL048794).
- , and Coauthors, 2013: Sea level change. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Cambridge University Press, 1137–1216. [Available online at [https://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5\\_Chapter13\\_FINAL.pdf](https://www.ipcc.ch/pdf/assessment-report/ar5/wg1/WG1AR5_Chapter13_FINAL.pdf).]
- Conkright, M. E., and Coauthors, 2002: World ocean database 2001. Volume 1: Introduction. NOAA Atlas NESDIS 42, 160 pp. [Available online at [ftp://ftp.nodc.noaa.gov/pub/data.nodc/woa/PUBLICATIONS/wod01\\_1.pdf](ftp://ftp.nodc.noaa.gov/pub/data.nodc/woa/PUBLICATIONS/wod01_1.pdf).]
- Cowley, R., S. Wijffels, L. Cheng, T. Boyer, and S. Kizu, 2013: Biases in expendable bathythermograph data: A new view based on historical side-by-side comparisons. *J. Atmos. Oceanic Technol.*, **30**, 1195–1225, doi:[10.1175/JTECH-D-12-00127.1](https://doi.org/10.1175/JTECH-D-12-00127.1).
- Cressman, G. P., 1959: An operational objective analysis scheme. *Mon. Wea. Rev.*, **87**, 367–374, doi:[10.1175/1520-0493\(1959\)087<0367:AOOAS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2).
- Derber, J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**, 1333–1347, doi:[10.1175/1520-0485\(1989\)019<1333:AGODAS>2.0.CO;2](https://doi.org/10.1175/1520-0485(1989)019<1333:AGODAS>2.0.CO;2).
- 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](https://doi.org/10.1038/nature07080).
- Ducet, N., P. Y. Le Traon, and G. Reverdin, 2000: Global high-resolution mapping of ocean circulation from TOPEX/Poseidon and ERS-1 and -2. *J. Geophys. Res.*, **105**, 19 477–19 498, doi:[10.1029/2000JC900063](https://doi.org/10.1029/2000JC900063).
- Flierl, G. R., and A. R. Robinson, 1977: XBT measurement of thermal gradients in the MODE eddy. *J. Phys. Oceanogr.*, **7**, 300–302, doi:[10.1175/1520-0485\(1977\)007<0300:XMOTGI>2.0.CO;2](https://doi.org/10.1175/1520-0485(1977)007<0300:XMOTGI>2.0.CO;2).
- Gleckler, P. J., and Coauthors, 2012: Human-induced global ocean warming on multidecadal timescales. *Nat. Climate Change*, **2**, 524–529, doi:[10.1038/nclimate1553](https://doi.org/10.1038/nclimate1553).
- Good, S. A., 2011: Depth biases in XBT data diagnosed using bathymetry data. *J. Atmos. Oceanic Technol.*, **28**, 287–300, doi:[10.1175/2010JTECH0773.1](https://doi.org/10.1175/2010JTECH0773.1).
- , 2016: The impact of observational sampling on time series of global 0–700 m ocean average temperature: A case study. *Int. J. Climatol.*, doi:[10.1002/joc.4654](https://doi.org/10.1002/joc.4654), in press.
- Gould, J., B. Sloyan, and M. Visbeck, 2013: *In situ* ocean observations: A brief history, present status, and future directions. *Ocean Circulation and Climate*, Vol. 103, Academic Press, 59–81.
- Gouretski, V., 2012: Using GEBCO digital bathymetry to infer depth biases in the XBT data. *Deep-Sea Res. I*, **62**, 40–52, doi:[10.1016/j.dsr.2011.12.012](https://doi.org/10.1016/j.dsr.2011.12.012).
- , and K. P. Koltermann, 2007: How much is the ocean really warming? *Geophys. Res. Lett.*, **34**, L01610, doi:[10.1029/2006GL027834](https://doi.org/10.1029/2006GL027834).
- , and F. Reseghetti, 2010: On depth and temperature biases in bathythermograph data: Development of a new correction scheme based on analysis of a global ocean database. *Deep-Sea Res. I*, **57**, 812–834, doi:[10.1016/j.dsr.2010.03.011](https://doi.org/10.1016/j.dsr.2010.03.011).
- , J. Kennedy, T. Boyer, and A. Köhl, 2012: Consistent near-surface ocean warming since 1900 in two largely independent observing networks. *Geophys. Res. Lett.*, **39**, L19606, doi:[10.1029/2012GL052975](https://doi.org/10.1029/2012GL052975).
- Gregory, J. M., H. T. Banks, P. A. Stott, J. A. Lowe, and M. D. Palmer, 2004: Simulated and observed decadal variability in ocean heat content. *Geophys. Res. Lett.*, **31**, L15312, doi:[10.1029/2004GL020258](https://doi.org/10.1029/2004GL020258).
- Hamon, M., G. Reverdin, and P.-Y. Le Traon, 2012: Empirical correction of XBT data. *J. Atmos. Oceanic Technol.*, **34**, 960–963, doi:[10.1175/JTECH-D-11-00129.1](https://doi.org/10.1175/JTECH-D-11-00129.1).

- 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, doi:[10.1016/0967-0637\(95\)97154-Z](https://doi.org/10.1016/0967-0637(95)97154-Z).
- Hoerling, M. P., J. Hurrell, and T. Xu, 2001: Tropical origins for North Atlantic climate change. *Science*, **292**, 90–92, doi:[10.1126/science.1058582](https://doi.org/10.1126/science.1058582).
- Ingleby, B., and M. Huddleston, 2007: Quality control of ocean temperature and salinity profiles—Historical and real-time data. *J. Mar. Syst.*, **65**, 158–175, doi:[10.1016/j.jmarsys.2005.11.019](https://doi.org/10.1016/j.jmarsys.2005.11.019).
- Ishii, M., and M. Kimoto, 2009: Reevaluation of historical ocean heat content variations with an XBT depth bias correction. *J. Oceanogr.*, **65**, 287–299, doi:[10.1007/s10872-009-0027-7](https://doi.org/10.1007/s10872-009-0027-7).
- , —, and M. Kachi, 2003: Historical ocean subsurface temperature analysis with error estimates. *Mon. Wea. Rev.*, **131**, 51–73, doi:[10.1175/1520-0493\(2003\)131<0051:HOSTAW>2.0.CO;2](https://doi.org/10.1175/1520-0493(2003)131<0051:HOSTAW>2.0.CO;2).
- , A. Shouji, S. Sugimoto, and T. Matsumoto, 2005: Objective analyses of sea-surface temperature and marine meteorological variables for the 20th century using ICOADS and the Kobe collection. *Int. J. Climatol.*, **25**, 865–879, doi:[10.1002/joc.1169](https://doi.org/10.1002/joc.1169).
- , M. Kimoto, K. Sakamoto, and S. I. Iwasaki, 2006: Steric sea level changes estimated from historical ocean subsurface temperature and salinity analysis. *J. Oceanogr.*, **62**, 155–170, doi:[10.1007/s10872-006-0041-y](https://doi.org/10.1007/s10872-006-0041-y).
- Johnson, G. C., and S. E. Wijffels, 2011: Ocean density change contributions to sea level rise. *Oceanography*, **24** (2), 112–121, doi:[10.5670/oceanog.2011.31](https://doi.org/10.5670/oceanog.2011.31).
- , and Coauthors, 2015: Ocean heat content [in “State of the Climate in 2014”]. *Bull. Amer. Meteor. Soc.*, **96** (7), S64–S68.
- Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.*, **103**, 18 567–18 589, doi:[10.1029/97JC01736](https://doi.org/10.1029/97JC01736).
- , Y. Kushnir, and M. A. Cane, 2000: Reduced space optimal interpolation of historical marine sea level pressure: 1854–1992. *J. Climate*, **13**, 2987–3002, doi:[10.1175/1520-0442\(2000\)013<2987:RSOIOH>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2987:RSOIOH>2.0.CO;2).
- Levitus, S., J. I. Antonov, T. P. Boyer, and C. Stephens, 2000: Warming of the World Ocean. *Science*, **287**, 2225–2229, doi:[10.1126/science.287.5461.2225](https://doi.org/10.1126/science.287.5461.2225).
- , —, and —, 2005: Warming of the World Ocean, 1955–2003. *Geophys. Res. Lett.*, **32**, L02604, doi:[10.1029/2004GL021592](https://doi.org/10.1029/2004GL021592).
- , —, —, R. A. Locarnini, 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](https://doi.org/10.1029/2008GL037155).
- , and Coauthors, 2012: World Ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys. Res. Lett.*, **39**, L10603, doi:[10.1029/2012GL051106](https://doi.org/10.1029/2012GL051106).
- Liu, J., and J. A. Curry, 2010: Accelerated warming of the Southern Ocean and its impacts on the hydrological cycle and sea ice. *Proc. Natl. Acad. Sci. USA*, **107**, 14 987–14 992, doi:[10.1073/pnas.1003336107](https://doi.org/10.1073/pnas.1003336107).
- Locarnini, R. A., A. V. Mishonov, J. I. Antonov, T. P. Boyer, and H. E. Garcia, 2006: *Temperature*. Vol. 1, *World Ocean Atlas 2005*, NOAA Atlas NESDIS 61, 182 pp.
- , —, —, —, O. K. Baranova, M. M. Zweng, and D. R. Johnson, 2010: *Temperature*. Vol. 1, *World Ocean Atlas 2009*, NOAA Atlas NESDIS 68, 184 pp.
- , and Coauthors, 2013: *Temperature*. Vol. 1, *World Ocean Atlas 2013*, NOAA Atlas NESDIS 73, 40 pp.
- Lorenc, A. C., R. S. Bell, and B. MacPherson, 1991: The Meteorological Office analysis correction data assimilation scheme. *Quart. J. Roy. Meteor. Soc.*, **117**, 59–89, doi:[10.1002/qj.49711749704](https://doi.org/10.1002/qj.49711749704).
- Lyman, J. M., and G. C. Johnson, 2008: Estimating annual global upper-ocean heat content anomalies despite irregular in situ sampling. *J. Climate*, **21**, 5629–5641, doi:[10.1175/2008JCLI2259.1](https://doi.org/10.1175/2008JCLI2259.1).
- , and —, 2014: Estimating global ocean heat content changes in the upper 1800 m since 1950 and the influence of climatology choice. *J. Climate*, **27**, 1945–1958, doi:[10.1175/JCLI-D-12-00752.1](https://doi.org/10.1175/JCLI-D-12-00752.1).
- , S. A. Good, V. V. Gouretski, M. Ishii, G. C. Johnson, M. D. Palmer, D. A. Smith, and J. K. Willis, 2010: Robust warming of the global upper ocean. *Nature*, **465**, 334–337, doi:[10.1038/nature09043](https://doi.org/10.1038/nature09043).
- Palmer, M. D., K. Haines, S. F. B. Tett, and T. J. Ansell, 2007: Isolating the signal of global ocean warming. *Geophys. Res. Lett.*, **34**, L23610, doi:[10.1029/2007GL031712](https://doi.org/10.1029/2007GL031712).
- , and Coauthors, 2010: Future observations for monitoring global ocean heat content. *Proc. OceanObs’09: Sustained Ocean Observations and Information for Society*, Venice, Italy, ESA WPP-306, European Space Agency, doi:[10.5270/OceanObs09.cwp.68](https://doi.org/10.5270/OceanObs09.cwp.68).
- Polyakov, I. V., and Coauthors, 2010: Arctic Ocean warming contributes to reduced polar ice cap. *J. Phys. Oceanogr.*, **40**, 2743–2756, doi:[10.1175/2010JPO4339.1](https://doi.org/10.1175/2010JPO4339.1).
- Rhein, M., and Coauthors, 2013: Observations: Ocean. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 255–316, doi:[10.1017/CBO9781107415324.010](https://doi.org/10.1017/CBO9781107415324.010).
- Rignot, E., S. Jacobs, J. Mouginot, and B. Scheuchl, 2013: Ice-shelf melting around Antarctica. *Science*, **341**, 266–270, doi:[10.1126/science.1235798](https://doi.org/10.1126/science.1235798).
- Roemmich, D., and Coauthors, 2009: The Argo program: Observing the global oceans with profiling floats. *Oceanography*, **22**, 34–43, doi:[10.5670/oceanog.2009.36](https://doi.org/10.5670/oceanog.2009.36).
- , J. Church, J. Gilson, D. Monselesan, P. Sutton, and S. Wijffels, 2015: Unabated planetary warming and its ocean structure since 2006. *Nat. Climate Change*, **5**, 240–245, doi:[10.1038/nclimate2513](https://doi.org/10.1038/nclimate2513).
- Shuman, F. G., 1957: Numerical methods in weather prediction: II. Smoothing and filtering. *Mon. Wea. Rev.*, **85**, 357–361, doi:[10.1175/1520-0493\(1957\)085<0357:NMIWPI>2.0.CO;2](https://doi.org/10.1175/1520-0493(1957)085<0357:NMIWPI>2.0.CO;2).
- Stephens, C., J. I. Antonov, T. P. Boyer, M. E. Conkright, R. A. Locarnini, T. D. O’Brien, and H. E. Garcia, 2002: *Temperature*. Vol. 1, *World Ocean Atlas 2001*, NOAA Atlas NESDIS 49, 167 pp.
- Straneo, F., and P. Heimbach, 2013: North Atlantic warming and the retreat of Greenland’s outlet glaciers. *Nature*, **504**, 36–43, doi:[10.1038/nature12854](https://doi.org/10.1038/nature12854).
- Toggweiler, J. R., and J. Russell, 2008: Ocean circulation in a warming climate. *Nature*, **451**, 286–288, doi:[10.1038/nature06590](https://doi.org/10.1038/nature06590).
- Trenberth, K. E., J. T. Fasullo, and M. A. Balmaseda, 2014: Earth’s energy imbalance. *J. Climate*, **27**, 3129–3144, doi:[10.1175/JCLI-D-13-00294.1](https://doi.org/10.1175/JCLI-D-13-00294.1).
- von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp.
- Wijffels, S. E., J. Willis, C. M. Domingues, P. Barker, N. J. White, A. Gronell, K. Ridgway, and J. A. Church, 2008: Changing expendable bathythermograph fall rates and their impact on estimates of thermosteric sea level rise. *J. Climate*, **21**, 5657–5672, doi:[10.1175/2008JCLI2290.1](https://doi.org/10.1175/2008JCLI2290.1).

- Wild, M., D. Folini, C. Schär, N. Loeb, E. G. Dutton, and G. König-Langlo, 2013: The global energy balance from a surface perspective. *Climate Dyn.*, **40**, 3107–3134, doi:[10.1007/s00382-012-1569-8](https://doi.org/10.1007/s00382-012-1569-8).
- Willis, J. K., D. Roemmich, and B. Cornuelle, 2003: Combining altimetric height with broadscale profile data to estimate steric height, heat storage, subsurface temperature, and sea-surface temperature variability. *J. Geophys. Res.*, **108**, 3292, doi:[10.1029/2002JC001755](https://doi.org/10.1029/2002JC001755).
- , —, and —, 2004: Interannual variability in upper ocean heat content, temperature, and thermosteric expansion on global scales. *J. Geophys. Res.*, **109**, C12036, doi:[10.1029/2003JC002260](https://doi.org/10.1029/2003JC002260).
- Zang, X., and C. Wunsch, 2001: Spectral description of low-frequency oceanic variability. *J. Phys. Oceanogr.*, **31**, 3073–3095, doi:[10.1175/1520-0485\(2001\)031<3073:SDOLFO>2.0.CO;2](https://doi.org/10.1175/1520-0485(2001)031<3073:SDOLFO>2.0.CO;2).