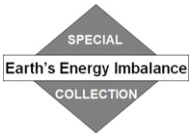

Improved Quantification of the Rate of Ocean Warming

LIJING CHENG,[a,b] GRANT FOSTER,[c] ZEKE HAUSFATHER,[d] KEVIN E. TRENBERTH,[e] AND JOHN ABRAHAM[f]

[a] *International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*
[b] *Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China*
[c] *Tempo Analytics, Old Town, Maine*
[d] *Energy and Resources Group, University of California, Berkeley, Berkeley, California*
[e] *National Center for Atmospheric Research, Boulder, Colorado*
[f] *School of Engineering, University of St. Thomas, Saint Paul, Minnesota*

(Manuscript received 15 November 2021, in final form 22 March 2022)

ABSTRACT: The increased concentrations of greenhouse gases in the atmosphere create an increase in Earth's thermal energy, which is mainly stored in the ocean. Quantification of the rate of increase in ocean heat content (OHC) is vital for understanding the current and future climate of Earth. Linear trend lines have been frequently used to quantify long-term rates of change, but are inappropriate because they cannot capture nonlinearity in trends, have large start- and end-point sensitivity, and the assumption of linearity is nonphysical. Here observed and model-based linear regressions with higher-order polynomial (quadratic), piecewise linear, and locally weighted scatterplot smoothing (LOWESS) are compared. Piecewise linear and LOWESS perform best in depicting multidecadal trends. It is shown that linear rates are valid for up to about 15-yr segments (i.e., it is valid to compute linear rates within a 15-yr time window). Using the recommended methods, ocean warming for the upper 2000 m increases from about 0 to 0.06 ± 0.08 W m^{-2} for 1958–73 to 0.58 ± 0.08 W m^{-2} for 2003–18, indicating an acceleration of ocean warming that happens in all four ocean basins and from near the sea surface to 2000 m. There is consistency between multimodel-mean historically forced climate models and observations, which implies that the contribution of internal variability is small for global 0–2000 m OHC. Notable increases of OHC in the upper ocean (i.e., 0–300 m) after about 1980 and the deeper ocean (300–2000 m) after the late 1980s are also evident. This study suggests alternative methods to those currently used to estimate ocean warming rates to provide a more accurate quantification of long-term Earth's energy changes.

SIGNIFICANCE STATEMENT: Quantifying long-term rates of change is needed to understand the time evolution of ocean warming and to assess the changing ocean and Earth's energy budgets. Linear trend lines have been frequently used but cannot capture nonlinearity in trends, and have large start- and end-point sensitivity. Based on an analysis of the statistical features of ocean heat content time series, this study proposes two alternative methods to quantify the rates of change, including piecewise linear fit and LOWESS. Robust increases in warming for the upper 2000 m detected through observational records and climate models from 1958 to 2020, indicate a robust acceleration of ocean warming. Slow penetration of heat from the upper ocean into the deeper ocean is also evident.

KEYWORDS: Climate change; Climate records; In situ oceanic observations; Time series; Uncertainty; Energy budget/balance

1. Introduction

Industrial-era human emissions of heat-trapping greenhouse gases have resulted in a clear, unequivocal, long-term warming of the planet (IPCC 2013). Much research to date has focused on quantifying the overall rate of warming and it is well established that the vast majority (>90%) of the extra heat from increasing greenhouse gases is absorbed by the oceans (Levitus et al. 2012; Rhein et al. 2013; Trenberth et al. 2014; Cheng et al. 2017; von Schuckmann et al. 2020; Gulev et al. 2022). Consequently, "global warming" is, in fact, mostly "ocean warming," which makes ocean heat content (OHC) and sea level rise the most vital indicators of climate change.

Estimation of the rates of change in OHC has been subject to large uncertainties because of the challenges of making multidecadal measurements with sufficient accuracy and spatial coverage. Detection of global and local low-frequency changes in the OHC are confounded by the impacts of both natural variability and observational shortcomings (Roberts et al. 2017). For instance, there are significant interannual fluctuations due to El Niño–Southern Oscillation (ENSO) (Roemmich and Gilson 2011; Cheng et al. 2019a) and other natural climate modes, such as the Indian Ocean dipole (Zhang et al. 2019), with typical periods of less than 10 years.

Denotes content that is immediately available upon publication as open access.

Corresponding author: Lijing Cheng, chenglij@mail.iap.ac.cn

DOI: 10.1175/JCLI-D-21-0895.1

Major volcanic eruptions induce additional interannual OHC fluctuations (Schmidt et al. 2018). Moreover, ocean data quality and data coverage are far from ideal, particularly before ~2005 when the Argo network was widely deployed (Abraham et al. 2013). These factors contribute to uncertainty in OHC records (Durack et al. 2014; Boyer et al. 2016; Cheng et al. 2019b), and led to a large spread (~100%) in OHC estimates for the upper 700 m (0.15–0.27 W m^{-2} within 1971–2010) in the Fifth Assessment Report of the IPCC (IPCC-AR5) in 2013 (700–2000 m warming rate is 0.09 W m^{-2}) (Rhein et al. 2013). Refinement of the uncertainty has been a key research focus since then, with recent progress assessed in the recent IPCC Special Report on the Ocean and Cryosphere (SROCC) (IPCC 2019) and the Sixth Assessment Report of the IPCC (IPCC-AR6) (Gulev et al. 2022). SROCC reported a 0–700 m ocean warming of 0.27 ± 0.05 W m^{-2} (0–2000 m: 0.41 ± 0.07 W m^{-2}) during 1971–2017 using the ensemble average of several data products (Bindoff et al. 2019). IPCC-AR6's assessment is 0.32 ± 0.10 W m^{-2} for 0–700 m and 0.16 ± 0.04 W m^{-2} for 700–2000 m over the 1971–2018 period. Thus, the 0–2000 m ocean warming rate is 0.48 W m^{-2} for IPCC-AR6, which is 17% higher than SROCC result (Table 5.1 in Bindoff et al. 2019), 33% higher than IPCC-AR5 (box 3.1 in Rhein et al. 2013) during the same period. This assessment generally confirms the argument in Cheng et al. (2019b) that ocean warming rates have been underestimated.

However, there are other issues. Traditionally, rates of OHC change have been quantified by a linear trend line applied over long time periods (Rhein et al. 2013; Bindoff et al. 2019). Linear regressions are not optimal when evident nonlinearities are present in the underlying data (Trenberth et al. 2007); linear trend lines can also be subject to large start- and end-point sensitivity. The IPCC-AR5 calculated the linear rate of OHC increase for 1971–2010 (Rhein et al. 2013; IPCC 2019), but there is no physical reason to expect linear changes over this period. Rather, there is an expectation of small rates of heating before about 1950 that increase in more recent years, driven by accumulated atmospheric greenhouse gases (Trenberth et al. 2007; Tokarska et al. 2019). SROCC improved the estimate by calculating linear rates in the two successive 25-yr periods: 1969–93 and 1993–2017; however, a linear regression is used in each period and the two trend lines are discontinuous, which is nonphysical. The IPCC-AR6 estimated the total ocean warming rate based on the difference between first and last annual mean value in a specific period (Gulev et al. 2022), which is strongly impacted by the year-to-year variations. Here we explore the advantages of alternative statistical models compared to the commonly used linear trends.

The present study revisits several key issues using improved updated datasets to better assess the rate of changes for OHC within 1971–2010 (the period highlighted in the IPCC-AR5), and also within the 1958–2018 and 1958–2020 periods over which reliable OHC data are available (Abraham et al. 2013; Cheng et al. 2017, 2020; Ishii et al. 2017; Levitus et al. 2012) using state-of-the-art observational OHC datasets, and climate model simulations from the Coupled Model Intercomparison Project phase 5 (CMIP5-MMM) (Taylor et al. 2012) to provide new insights.

2. Data and methods

a. Data

This study focuses on OHC for the upper 2000 m and selects several OHC estimates to illustrate the impact of the rate calculation method. The IPCC-AR5 chose 1971–2010 for OHC changes, and combined Domingues et al. (2008) (DOM) for warming in the upper 700 m (the largest among the five time series featured) with Levitus et al. (2012) (LEV) for warming in the 700–2000 m layer (DOM+LEV). This study uses an updated (2018) version of DOM (denoted as DOMu). Two other recent OHC analyses also incorporate improvements to instrument bias corrections and mapping methods: Ishii et al. (2017) (ISH) and Cheng et al. (2017) from the Institute of Atmospheric Physics (IAP). These data are selected in this study for illustration purpose because they better account for the "conservative bias" (in poorly observed areas, the reconstruction is biased to the climatology field) in many other estimates [see Cheng et al. (2019b) for a detailed discussion].

We also use the multimodel mean of the climate models in CMIP5-MMM that well represents the historical OHC changes (Gleckler et al. 2012; Cheng et al. 2016, 2019b). The list of models used is provided in Cheng et al. (2019b). Representative concentration pathway scenario with radiative forcing of 4.5 W m^{-2} at 2100 (RCP4.5) projections were used to extend CMIP5 historical simulations beyond 2005. However, a multimodel mean averages out the internal natural variability, although external variability, such as from volcanic eruptions and greenhouse gases, becomes emphasized. Although CMIP6 models are recently available they underestimate the ocean trend for the upper 2000 m according to the assessment in Eyring et al. (2022) and the significantly larger equilibrium climate sensitivity in many models needs to be accounted for when using CMIP6 data. Thus, we decided not to include CMIP6-MMM in this study. A comprehensive analysis of OHC change for all available observational estimates and CMIP5/CMIP6 models will be completed in a separate study.

b. Rate quantification method

This study aims to quantify the trends of OHC from decadal to multidecadal scales, which are more relevant to a climate change signal which grows in time, but allows for decadal natural variability too. We wish to exclude ENSO and other high-frequency variability. To separate signal (low-frequency variability) from noise (high-frequency variability) during the trend calculation, we apply four different regression methods. These can broadly be divided into two categories: full, which uses a model defining the signal from throughout the entire time span, and segmental, which can model local temporal variations without distorting global patterns. Our global models are linear and quadratic fits, for segment models we use piecewise linear fits (PLF) and locally weighted scatterplot smoothing (LOWESS).

PLF (Tomé and Miranda 2004) models the data as a continuous piecewise linear function, with slopes changing at breakpoints. Compared with a traditional linear regression fit, PFL

is a linear regression in each subperiod but with an additional constraint to ensure the continuity of the trend lines. The key of PLF is the determination of the timing and length of each segment, or equivalently the times of the breakpoints (the choice of span width will be discussed in the next section). Here segments of equal length are used over the entire time series to ensure that the method can be easily implemented. In some cases, one can include a constraint, such as adding a disposable portion before the first piece to enforce a specified small or zero OHC change in this piece. As the human-induced changes are small prior to the 1970s (Trenberth et al. 2007; Tokarska et al. 2019), making a reasonable assumption such as this can potentially better constrain the fit for the entire time series. And the contribution from natural variability may also be small because of the plateau in global mean surface temperature between the 1950s and 1970s.

LOWESS applies a weighted regression at each data point within a prescribed span width that depends on the time scales of interest (Cleveland 1979, 1981) (as tested in the next section). For the application of LOWESS, one also prescribes a span, the number of nearby data used in each locally weighted fit. Since the time sampling is even, this corresponds to a specific choice of time scale. A merit of LOWESS is continuity of the time derivative of the smoothed series, while for PLF the derivative is discontinuous at the breakpoints. The two together may be even more powerful as the PLF helps determine the span.

One caveat of the PLF and LOWESS methods is if a time series has a regime shift (abrupt change), they may not properly capture this shift. For global OHC, it is highly unlikely, because of the large thermal inertia of global ocean. Regional OHC changes may be impacted by regional and fast ocean mass formation driven by anomalous surface fluxes or winds.

These methods have been widely utilized and proven in previous studies that quantified the rate of sea level rise (e.g., Visser et al. 2015; Zhu et al. 2021), global surface temperature increase (e.g., Cahill et al. 2015), ice sheet loss, glaciers melting, etc. A comprehensive review is provided in Visser et al. (2015). Previously, some other statistical approaches have been used to quantify the rate from the observational time series (Zhu et al. 2021). For instance, Cazenave et al. (2014) and Nerem et al. (2018) removed the ENSO and PDV variability by a multivariable empirical orthogonal function (EOF) analysis. However, for OHC, there are substantial observational uncertainties from interannual to decadal scales, and the impacts of key modes on OHC are a function of depth. Thus the PDV and ENSO related variability cannot be as reliably determined and then removed from the raw time series.

c. *Choice of span width in PLF and LOWESS*

Choices of the time-span impact the results of both PLF and LOWESS. Wider spans can only mimic slower, simpler patterns because the model has few degrees of freedom (DOF; i.e., a PLF with n segments uses a model with $n + 1$ DOF). Shorter spans increase the DOF of the model, thus there are more short-term fluctuations and a closer fit to the original time series (i.e., 5-yr compared with 15- or 20-yr windows). While many DOFs can give an apparently excellent fit, the DOF left to the residuals is reduced and the classic statistical trap called "overfitting" occurs. Therefore, the fit should be as parsimonious as possible, but not so sparse as to lose significant information. Furthermore, LOWESS is a host of weighted regressions which narrow the effective "window" width for a given choice of time scale. Hence, the effective time scale for LOWESS is less than the specified window.

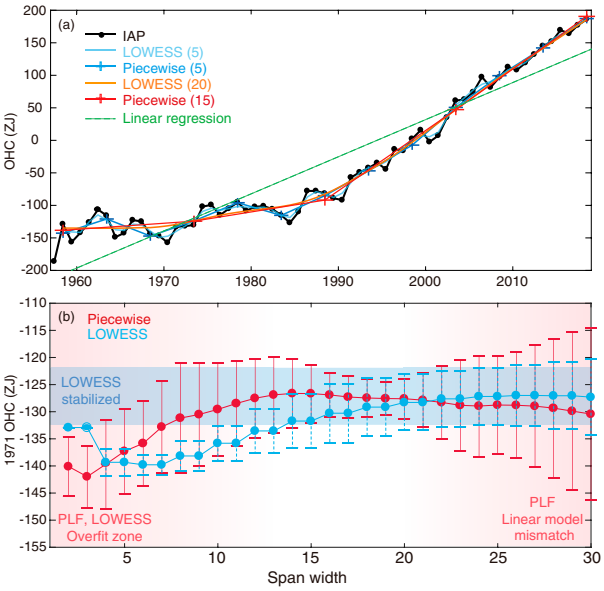

FIG. 1. Impacts of the span width on LOWESS and PLF methods. (a) LOWESS and PLF applied to IAP data within 1958–2018 using different choices of span width, as indicated in the legend. (b) Values of OHC at 1971 from LOWESS (blue) and PLF (red) using various span width (2–30 years) and various start points within 1958–71 (14 choices). The error bars (2-sigma range) are derived through the 14 runs with different start points. in (b) the pink shading on the left reflects models that overfit the time series whereas the right-hand shading indicates large error bars. The horizontal illustrative blue shaded band indicates the stabilized zone of LOWESS results where they are statistically indistinguishable and the error bars have similar magnitude (the shading corresponds to the error bar when span width is 25 years).

To illustrate the impact of different choices of span width in PLF and LOWESS analyses, we apply both fits to the OHC data for 1958–2018 using all possible span widths from 2 to 30 years, and all possible start times within 1958–71 (Fig. 1b). The estimated values in 1971 have 2-sigma error bars calculated according to the sensitivity test on the starting points (method see section 3d). A short span width (less than ~8 years for both PLF and ~13 years for LOWESS) gives lower 1971 estimates, because the 1970/71 OHC in IAP is lower than the adjacent data (Fig. 1a), indicating the impact of interannual fluctuations. At time scales less than 10 years, overfitting is dominant although it can better reproduce interannual variations from volcanic eruptions or ENSO events. Here, 10–20 years is the "sweet spot" where the interannual fluctuations are muted and filtered values stabilize. For time

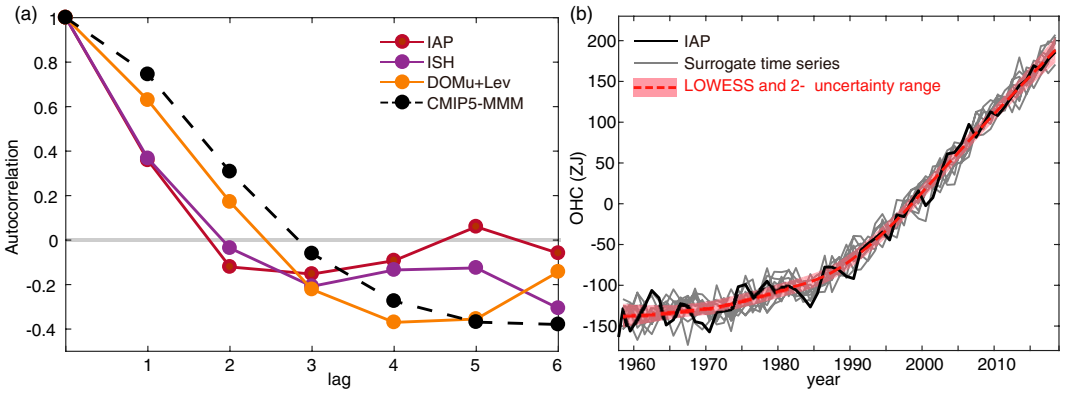

FIG. 2. (a) Autocorrelation of residuals from a linear fit to OHC time series within 1958–2018. (b) Example of 10 surrogate series (gray) by MC simulation based on IAP time series and LOWESS method. The original time series is shown in black. The LOWESS trend line is presented in dashed red, with the 95% confidence interval (2 times the standard deviation) in pink based on 1000 surrogates.

scales > 13 years, the results are statistically indistinguishable for LOWESS and the error bars are very similar, suggesting that the choice of time scale is not critical within this window for OHC records. However, the error bars get bigger for PLF with increasing span width (Fig. 1b) because a wider span provides smoothing that is too strong. Beyond the 20-yr time scale, values again decrease, now because of too few DOF in the model. Results (Fig. 1) are robust using other data and other years.

Information criteria can select between various time scales, but we apply physical considerations (i.e., accounting for interannual variability associated with ENSO and volcanos) as well. Because we emphasize longer-term trends relevant to the slow-varying signals, fluctuations less than 8 years should be removed. Based upon all of these considerations, a reasonable choice of the span is ~10–20 years for PLF and slightly longer for LOWESS because of its lower effectiveness of smoothing. A test provided in the appendix confirms different "effectiveness" of smoothing by PLF compared with LOWESS, the difference is 5–15 years. Second, the test indicates that the best choices of span width is 8–16 years for PLF and 13–25 years for LOWESS. Also, breakpoints at times within 3 years of a major volcanic eruption (Agung 1963/64, El Chichón 1982, Pinatubo 1991), or within 2 years of super El Niños (1982/83, 1997/98, 2015/16) should be avoided if possible.

The time period should also be chosen according to the data availability and the reliability of the record, for instance, from 1955 to 1960 mechanical bathythermograph (MBT) data provided stable coverage for the upper 200 m (Abraham et al. 2013), in the late 1960s (1966–70) expendable bathythermograph (XBT) data extended to below 200 m (~450 m within 1966–85 and ~750 m within 1985–2000) (Abraham et al. 2013), during 2005–07 Argo float observations in the open ocean extended down to 2000 m (Riser et al. 2016). Spatial coverage has also evolved. Therefore, many individual studies choose one of these dates as the start point of their estimates (Domingues et al. 2008; Levitus et al. 2012; Ishii et al. 2017; Cheng et al. 2017).

With these statistical and physical consideration, this study adopted 15 (25) years for PLF (LOWESS) when examining the full 1958–2018 OHC record, which is to say the record is quasi linear for about 15 years. Hence, four pieces were used: 1958–73–88–2003–18; 1958 is used because reliable OHC estimates start from the late 1950s (Cheng et al. 2017), and 1958 is a point when the MBT system helped achieved reasonable coverage. The last piece (2003–18) is reasonably homogeneous with regard to the Argo observing system (Riser et al. 2016). Around 1988, there was another observation system transition from shallower XBTs (terminal depth of 400–550 m) to deeper XBTs (terminal depth of 700–800 m) and its presence as another series marker is fortuitous. Note that using 15-yr periods gets the time series to 2018, but as more data are added, it is reasonable to extend the last interval slightly to 2020.

d. *Uncertainty estimate*

To obtain the uncertainty range of the fitted trend lines, uncertainty around the trend lines of LOWESS are estimated by Monte Carlo simulations under the assumption of a given mean (LOWESS trendline), variance and autocorrelation structure estimated directly from the underlying datasets (Visser et al. 2015, 2018; Chandler and Scott 2011). The variance is calculated from the OHC time series after removing a LOWESS-smoothed trendline (as in Foster and Rahmstorf 2011). The autocorrelation of residuals are analyzed within 1958–2018 for all four products (DOMu+LEV for the 1971–2018 period) (Fig. 2). The lag one autocorrelation (i.e., 1-yr lag) is within 0.35–0.76 for four datasets. Differences appear among different datasets, partly because the time variation of the interannual variability (i.e., ENSO diversity) (Capotondi et al. 2015; Timmermann et al. 2018) and partly because of the error in interannual OHC records (Boyer et al. 2016). It may also be due to error in estimating autocorrelation. Another reason for the uncertainty is the use of annual time series rather than monthly as ENSO is phase locked in northern winter and always persists to the next January and February, and hence, the annual mean is not a proper

representation of the ENSO signal. These inconsistencies are responsible for the differences in error bars provided in literature. The lag two correlations are reduced sharply for OHC residuals (Fig. 2) and are negative for two of the datasets, probably arising from ENSO episodes. This characteristic of the autocorrelation implies that the application of autoregressive model of order 1—AR(1) is more appropriate here than Autoregressive Moving Average Model ARMA(1, 1). The AR(1) linear model predicts the present value of a time series using the immediately prior value in time, assuming that the errors are independently distributed with a normal distribution. Hence, this is simple persistence. ARMA(1, 1) provides a description of a stationary stochastic process in terms of two polynomials, one for the AR(1) and the second for the moving average—MA(1). Thus, the new values of a time series are predicted by the values in present and past times.

Based on this autocorrelation analysis, a surrogate OHC series is formed by simulating a new residual series based on the AR(1) process, and adding it to the estimated LOWESS line. Then a LOWESS trendline is estimated for each surrogate. This process is repeated 1000 times and 1000 trendlines are available. The 95% confidence interval for the trendline is calculated based on ±2 times the standard deviation ($\pm 2\sigma$) of all 1000 trendlines of the surrogates. An illustration of 10 of such surrogates can be found in Fig. 2 for IAP data. Results for other data (not shown) are similar. As using more surrogates (i.e., 2000 or 5000) has negligible impact on the estimate of confidence interval, 1000 cases are used here. Furthermore, the uncertainty in the rate of the OHC is estimated by the 1000 LOWESS trendlines: 1) calculating the rate based on the difference between first and last annual mean value of the LOWESS trendline in a specific period; 2) calculating ±2 times the standard deviation ($\pm 2\sigma$) of the 1000 rate values.

For linear regression in this study, the ARMA(1, 1) model is used when both lag one and lag two autocorrelations are positive, and the AR(1) model is used when lag one autocorrelation is positive and the lag two autocorrelation is negative. Simple standard error assuming white noise is used when the lag one autocorrelation is negative. Only positive autocorrelations are taken into account, because theoretically the lag one or two correlations in AR(1) and ARMA(1, 1) models should be positive. Our method does not require an explicit estimation of the error sources (i.e., systematic error, sampling error, climatology error, etc.), which are not independent, and the error covariance is unknown.

In addition to computing uncertainty levels based on regression, we also estimated the sensitivity of the 1971–2010 warming rate to start and end dates as the time window is extended 0–years back from 1971 and 0–6 years forward from 2010 for a total of 49 individual calculations. The two standard deviation (2σ) is calculated based on these 49 results to quantify this sensitivity.

3. Results of tests on the methods: Sensitivity to start and end points

a. Linear and quadratic fits and their caveats

First, a traditional linear fit to OHC 0–2000 m time series is applied within the 1971–2010 window (Fig. 3a) because it is the simplest and most widely used method (as in the IPCC-AR5) (Rhein et al. 2013). Figure 3a uses the IAP OHC estimate as an example (Cheng et al. 2017), though all recent estimates assessed show similar results. It is apparent that a linear fit is not a satisfactory model for OHC, as it results in very large residuals: 159 ZJ (i.e., 10^{21} J) (norm of residuals) with a "quadratic-like" feature. For comparison, the total OHC increase in the upper 2000 m within 1971–2010 is about 260 ZJ. A quadratic fit to the residuals from the linear fit is statistically significant ($R^2 = 0.79$, $p = 0.0021$ for IAP) (Fig. 3a) thereby rejecting the null hypothesis that a linear fit is a proper model. This is to be expected, as all OHC times series exhibit slower warming in earlier years and faster after the 1990s.

The linear trend is also sensitive to the start and end points, with ±6-yr shifts resulting in up to 23% differences in the mean warming rate for the CMIP5-MMM for 1971–2010 (0.34–0.43 W m^{-2}, with the central estimate of 0.39 W m^{-2}) (Fig. 3b). All data also show a large sensitivity to a linear fit (IAP ~ 20%) (Ishii et al. 2017) (ISH, ~ 17%) and updated (Domingues et al. 2008) (DOMu, ~ 16%). This start and end point sensitivity also implies that the IPCC-AR6's method of estimated the total ocean warming rate within 1971–2018

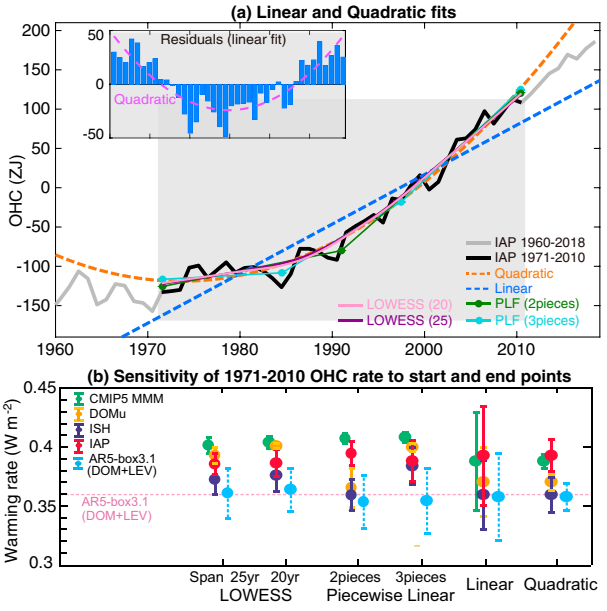

FIG. 3. (a) Examination of the traditional linear and quadratic fits applied to IAP OHC 0–2000 m time series (black line) for the 1971–2010 period (gray shading). The residuals (blue bars) from the linear fit are shown in the inner box, with a quadratic fit (purple) applied to the residuals. (b) Sensitivity of the mean warming rate within 1971–2010 to the start and end points. The mean rate (dots) obtained by various statistical fits to OHC data within 1971–2010 from the IPCC-AR5, CMIP5-MMM, and three more recent OHC datasets. The mean rate is calculated by differencing 2010 and 1971 OHC values of the trendline. The error ranges are derived through the sensitivity test for the start and end points (calculated by 2-sigma range of the OHC rates with different start and end points).

based on the difference between 1971 and 2018 values is uncertain (Gulev et al. 2022).

These results show that a single linear trend is an inappropriate model of observed OHC change while models allowing nonlinear behavior provide more reasonable estimates. A common choice is a higher-order polynomial function, for instance a quadratic fit to sea level time series to examine its acceleration (Nerem et al. 2018; Cazenave et al. 2018) and an estimate of acceleration/deceleration emerges directly. While the quadratic fit of Fig. 3a is much better than a linear fit, showing less end-point sensitivity (Fig. 3b), it develops unphysical values outside of the data period (Fig. 3a).

b. *New trend calculation methods: Sensitivity to start and end points*

Only if the signal is truly linear (or parabolic) will a straight-line (or quadratic) fit give best results. In addition to the uncertainty due to noise (i.e., short-term variability), uncertainty due to model mismatch can be considerably larger. Hence, we tested "segmental" methods, which treat the high-frequency behavior as noise, while allowing enough DOF for the models to accommodate more complex changes than a straight line or parabola.

The sensitivity of PLF and LOWESS to the start and end points (Fig. 3b) are both much less than linear or quadratic fits, with 2-sigma error ranges always within 0.02 W m^{-2} (the total range is less than 8% of the total warming rate for ISH, less than 5% for IAP, DOMu, and CMIP5-MMM). For 1971–2010, the four OHC data (DOMu, IAP, ISH, CMIP5-MMM) have higher mean warming rates (0.36–0.41 W m^{-2} with a median of 0.39 W m^{-2}) using PLF and LOWESS than AR5 (0.24–0.36 W m^{-2}, Fig. 3b), supporting the conclusions of recent reports (Cheng et al. 2019b; Durack et al. 2014; IPCC 2019). The underestimation in AR5 estimates (except DOM) is mainly due to data sparse regions where zero anomalies were assumed along with the choice of methods and start/end points (Fig. 3b).

Although PLF show smaller impacts of start and end points than linear trends, the choice of span width is important. Figure 3b also presents results for two arbitrary choices of span width for PLF: two pieces (~20 yr span width) and three pieces (~13 yr span width). Some data show notable sensitivity to the choice of span width; for example, ISH data show ~0.03 W m^{-2} differences for PLF and ~0.04 W m^{-2} differences for DOMu data. This confirms the importance of the span width choice.

4. Results on applications to changes since 1958

Here we used PLF and LOWESS method with optimized span width to derive the rates of ocean warming and exploit the results of the revised filters.

Five-part PLFs were applied for the period 1958–2018 to three OHC records (ISH, IAP, CMIP5-MMM) (Figs. 4a–c, Table 1). DOMu is not used here because the updated version is available only after 1971. The three OHC records show significant ocean warming since 1958 (Fig. 4) using both PLF and LOWESS: the spread of 0–2000 m ocean warming rate given by different datasets changes from 0.01–0.17 W m^{-2} (1958–73), to 0.13–0.27 W m^{-2} (1973–88), to 0.43–0.58 W m^{-2} (1988–2003), and finally to 0.58–0.71 W m^{-2} (2003–18). Especially, IAP data show a long-term rate increase from 0.06 ± 0.08 W m^{-2} (1958–73) to 0.58 ± 0.08 W m^{-2} (2003–18) using LOWESS.

Different datasets give slightly different rates of warming, particularly for the earlier periods, indicating the impact of data scarcity and quality. Before 1958, much fewer data make OHC estimates unreliable (Cheng et al. 2017), but both IAP and CMIP5-MMM suggest a faster warming within 1945–57 than 1958–73 (Figs. 4a,c), indicating the impact of Mount Agung eruption in 1963/64. To test the sensitivity of the results to early period uncertainties, an assumption is applied before 1958 that the rate of OHC change was zero. Indeed, given the uncertainties it is not possible to demonstrate otherwise (Fig. 4b and Table 1). Adding a zero-trend piece before 1958 slightly improves the consistency among datasets in the first 1958–73 piece but has little impact on the later periods. Here we set the 1943–57 OHCs to the 1958 value, but making other choices (e.g., assigning a mean value of 1957–59) would only slightly impact the results.

As LOWESS ensures continuity of the time derivative of the smoothed series, it can also be used to estimate rates (Fig. 5a) by applying centered differences to the LOWESS smoothed OHC series. The increase of OHC rate over time is pronounced after the late 1950s for all three datasets (Fig. 5a).

However, decadal variations are also apparent (Fig. 5) [ISH data (not shown) show similar decadal variations]. For instance, the OHC rate is nearly constant from the early 2000s to the present, consistent with Argo-data-based analyses (Johnson et al. 2017) and other indirect methods (Meyssignac et al. 2019). CMIP5-MMM shows a higher rate of warming since 2000, probably because volcanic eruptions after 2000 have not been taken into account (Gleckler et al. 2016; Outten et al. 2015). The stabilized warming rate during the 1960s, from the late 1970s–80s and shape increase during 1990s shown in all three data, is at least partly due to volcano eruptions (Gleckler et al. 2016), as revealed by sea level records (Fasullo et al. 2016) and CMIP5-MMM. The broad consistency between CMIP5-MMM and ISH and IAP data also implies that multidecadal internal variability (such as Pacific decadal variability or Atlantic multidecadal variability) does not have detectable imprints on the global OHC time series. Nevertheless, the impacts of PDV and AMV on global OHC are still open questions.

The global ocean warming better reflects the effects of external forcing because individual basins (Fig. 5b) are linked by transports or exchanges of heat; the latter come about through changes in winds and surface fluxes of heat. While all four major basins contributed to the increase of global warming rate in the past six decades (Fig. 5b), exchanges between Indian and Pacific basins through the Indonesian Throughflow contributed to the lack of a significant warming trend in the Indian Ocean in the 1980s to the early 1990s (Li et al. 2018).

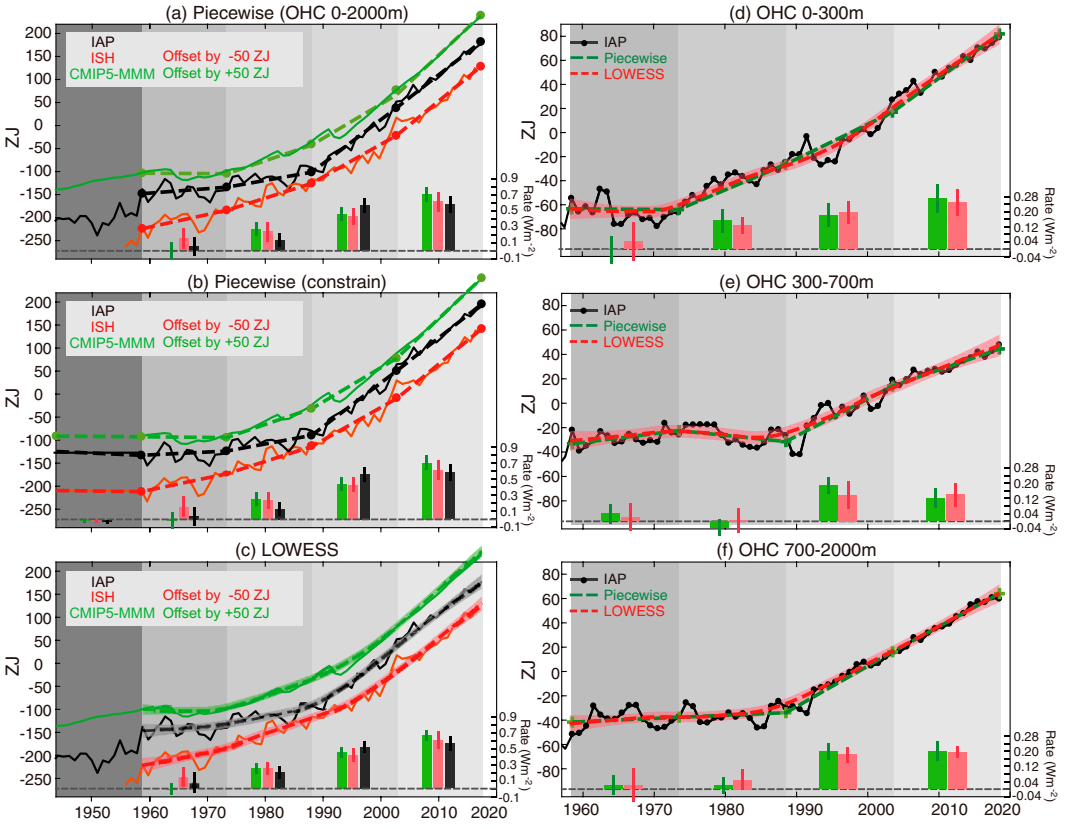

FIG. 4. Application of PLF and LOWESS to the OHC records. (a) Four-piece PLF, (b) five-piece PLF with a disposable piece (1943–58) as a constraint, and (c) LOWESS applied to the two recent OHC datasets (ISH and IAP) and to the CMIP5-MMM. Four-piece PLF and LOWESS applied to the IAP data (black dotted line) for three ocean layers: (d) 0–300, (e) 300–700, and (f) 700–2000 m. The fitted time series are shown in dashed curves with 2-sigma ranges of LOWESS results shown in shading (using Monto Carlo method). For PLF, the breakpoints are shown by dots in (a) and (b) and by crosses in (d)–(f). The mean warming rates (W m^{-2}, averaged over Earth's surface) is provided with bars, calculated by differencing starting and ending OHC values of the trendline, and 2-sigma error bars are shown in sticks. The background gray shading indicates the pieces.

The southern oceans (>35°S) and Pacific Ocean show the earliest onset of the acceleration since the 1970s and the early 1980s, respectively, and are also linked. Observations and models indicate that the Southern Ocean has warmed since 1970 due to surface heat flux and heat transport (Meredith et al. 2019; Swart et al. 2018; Armour et al. 2016). However, uncertainty remains for the changepoint of the Southern Ocean OHC. The Atlantic Ocean warming rate increased after 1990, in line with an independent reconstruction based on sea surface temperatures (Zanna et al. 2019), suggesting that both the vertical penetration of surface anomalies and ocean circulation change [i.e., meridional overturning circulation (MOC)] contributed (Zanna et al. 2019).

TABLE 1. Ocean warming rates (0–2000 m) for 1958–2018 using PLF with and without a constraint, as well as LOWESS. The mean rate is calculated by differencing the breakpoints in each piece (units: W m^{-2}, averaged over Earth's surface). Two-sigma errors are included.

	CMIP5-MMM			ISH			IAP		
Time period	PLF	PLF constraint	LOWESS	PLF	PLF constraint	LOWESS	PLF	PLF constraint	LOWESS
1943–58		−0.01 (0.02)			−0.01 (0.02)			−0.03 (0.02)	
1958–73	0.00 (0.10)	−0.01 (0.10)	0.02 (0.06)	0.17 (0.12)	0.16 (0.12)	0.16 (0.08)	0.06 (0.11)	0.04 (0.11)	0.06 (0.08)
1973–88	0.26 (0.08)	0.26 (0.08)	0.25 (0.04)	0.24 (0.10)	0.24 (0.10)	0.27 (0.06)	0.13 (0.09)	0.14 (0.09)	0.18 (0.05)
1988–2003	0.45 (0.08)	0.45 (0.08)	0.46 (0.04)	0.43 (0.10)	0.43 (0.10)	0.43 (0.06)	0.58 (0.09)	0.57 (0.09)	0.53 (0.05)
2003–18	0.71 (0.09)	0.71 (0.09)	0.66 (0.06)	0.62 (0.12)	0.62 (0.12)	0.61 (0.08)	0.59 (0.10)	0.59 (0.10)	0.58 (0.08)

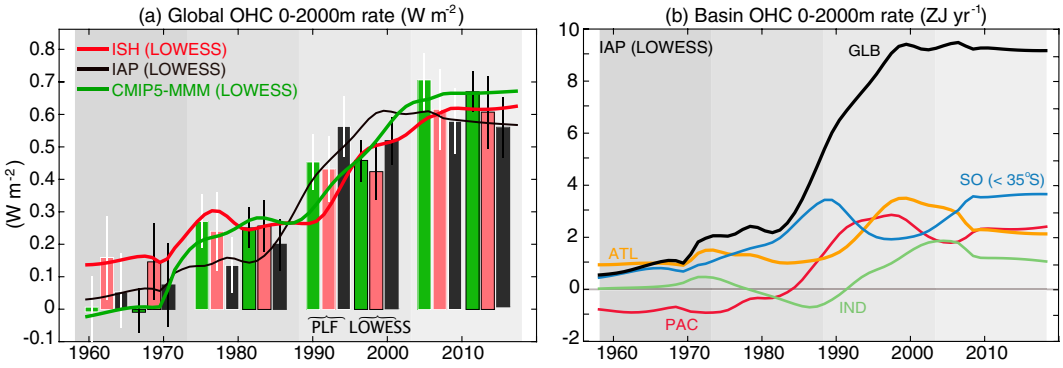

FIG. 5. Rate of ocean warming. (a) Upper-2000-m OHC since 1958 based on LOWESS-smoothed series using centered differences (W m^{-2} averaged over Earth's surface). The continuous rate estimates are compared with those in each time window from Fig. 4 using both PLF (white bars) and LOWESS (black bars). The background gray shading, as in Fig. 4, shows the 15-yr spans. (b) OHC rate (ZJ yr^{-1}) in the four major ocean basins.

Both four-piece PLF and LOWESS have been applied to IAP data since 1958 for the three vertical layers (0–300, 300–700, and 700–2000 m separately) to illustrate the changes of ocean warming over depth (Figs. 4d–f). For the 0–300 m layer, IAP data show very weak warming during the 1960s, followed by a significant warming after the early 1970s (Fig. 4d). The OHC change within 0–300 m is consistent with the evolution of sea surface temperatures, which also show weak or no warming during the 1950s to the 1960s and a significant warming after the mid-1970s (Trenberth et al. 2007). However, below 300 m, the OHC warming before the late 1980s is either insignificant or very weak (Figs. 4e,f), but becomes large and significant after the late 1980s. The mean warming rates in the last two pieces (1988–2002, 2003–18) range within ~0.12–0.19 W m^{-2} for 300–700 m and ~0.18–0.21 W m^{-2} for 700–2000 m. There is a 1–2-decade delay of the warming of the deep layers (below 300 m) behind the upper ocean (0–300 m). It may not be a coincidence that XBTs began to be deployed to depths greater than 400 m in the late 1980s, and hence, there is a question about whether this signal is real or not. Otherwise, the interpretation is that the ocean warming did not penetrate below about 300 m until the late 1980s. These aspects can be explored further by considering regional signals.

This ocean warming trend and the warming penetration into the deep layers can be further seen through global mean temperature anomaly evolution from the sea surface to 2000 m (Fig. 6a). For example, relative to a 1958–62 baseline, the global mean 0.02°C contour (Fig. 6a) temperature anomalies first occurred near the sea surface in the late 1970s, and penetrated gradually into deep layers before reaching ~2000 m around 1997. Similar penetration can be seen for 0.1°C contour temperature anomalies, which is now about ~900 m deep (Fig. 6a). CMIP5 models show similar penetration, and suggest that greenhouse gas forcing is the major driver, with some offsetting effects by aerosols and volcanic forcing (Bilbao et al. 2019).

The rate of global temperature change for each layer based on LOWESS-smoothed temperature time series (Fig. 6b) shows a mostly positive warming rate over all depths after the late 1950s. Consistent with previous findings, the increase of warming rate is more pronounced from late 1980s to the 1990s mainly in the upper 1500 m, while below there, both warming and its rate change are less evident. The subsurface (300–1500 m) cooling between the 1970s and the late 1980s could be noise (insufficient observations, large data uncertainty) or natural variability before the warming kicks in the late 1980s at these levels. Volcano-induced cooling (~1983) could also contribute. CMIP5-MMM also show such cooling,

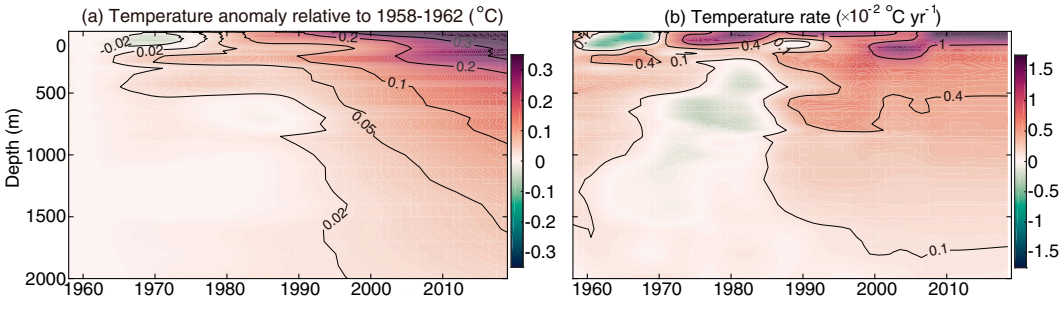

FIG. 6. (a) Global mean ocean temperature evolution at each depth from the sea surface to 2000 m, relative to 1958–62, with temperature anomalies (black contours). (b) Rate of temperature change at each depth for LOWESS-smoothed temperature anomaly time series using centered differences, with temperature rates (black contours).

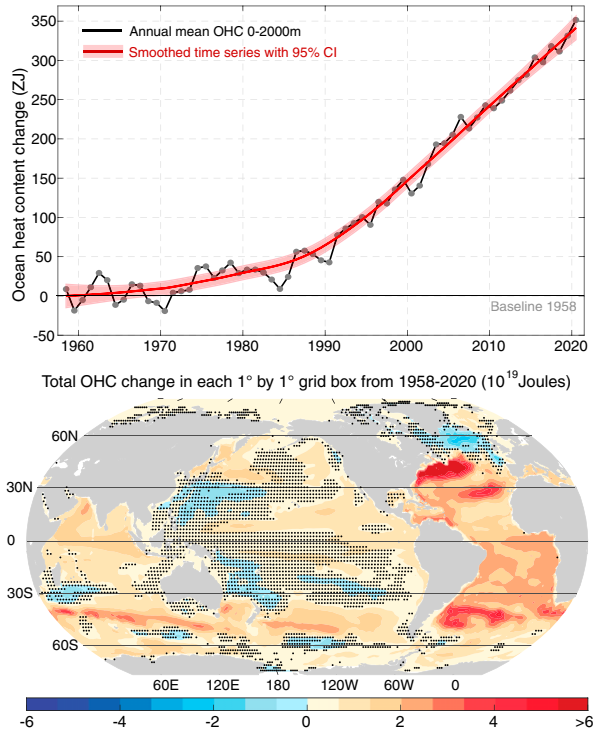

FIG. 7. (top). Global upper-2000-m OHC from 1958 through 2020 as the dotted line and the smoothed time series based on LOWESS method (with the 95% CI in pink shading). (bottom) The total OHC change in each 1° × 1° grid box calculated based on the difference between the start and end points of the LOWESS-smoothed time series. The areas with insignificant change (90% CI) are stippled.

implying that it is likely due to external forcing (GHGs and aerosols) rather than internal variability.

5. Update of ocean and Earth's energy budget since 1958

The above developments were based upon data through 2018, and with two more years of data available, we revisit the ocean and Earth's energy budget since 1958. Accordingly, Fig. 7 shows the updated OHC time series along with the LOWESS fit for the entire record of 1958–2020 based on the IAP dataset. Only the IAP time series is used here and the results will be updated regularly in the future. Note that 16.1 ZJ yr^{-1} is 1 W m^{-2} for the global domain. With the new approach, the total ocean warming for the upper 2000 m is 341.3 ± 21.0 ZJ from 1958 to 2020 (with the 95% confidence interval), equal to a mean ocean warming rate of 0.34 ± 0.02 W m^{-2}. For 1971–2018 (used in IPCC-AR6), the total OHC increase is 310.8 ± 15.4 ZJ for the upper 2000 m.

The LOWESS approach also provides an opportunity to quantify the local ocean heat budget by minimizing the impact of year-to-year variability. The physical arguments applied for global OHC generally also apply on regional scales. For example, the consideration of the timing of ENSO and volcanic eruption impacts, and other multidecadal-scale changes.

The regional OHC in each 1° × 1° grid box has been estimated by applying LOWESS to local OHC time series and then calculating the difference between 1958 and 2020 (Fig. 7). The Atlantic Ocean (within 50°S–50°N) and the Southern Ocean store more heat than in the other basins, probably associated with the deep convection and subduction processes that is more effective transporting heat into the deep layers. Another factor is ENSO which acts to move heat out of the tropical western Pacific thereby moderating global warming (Cheng et al. 2019a). Most regions on the west side of the Pacific Ocean within 40°S–40°N show an insignificant warming. There are several cold spots in the background of broadscale warming, including the northwest Pacific, southwest Pacific, and subpolar North Atlantic Oceans. Some (subpolar North Atlantic) are linked to the fingerprint of the AMOC slowdown (Caesar et al. 2018).

Estimating deep ocean change below 2000 m is more challenging than for the upper ocean because of the limited in situ observations. Thus, an estimate of a linear trend is the state-of-the-art (Purkey and Johnson 2010) and the uncertainty is substantial. Combining the upper-2000-m change with the updated estimate from Purkey and Johnson (2010) for the below-2000-m change (1.15 ± 0.57 ZJ yr^{-1} after 1991 and assuming zero change before 1990) yields a total ocean warming of 374.7 ± 41.1 ZJ from 1958 to 2020 and 341.9 ± 34.4 ZJ from 1971 to 2018.

Here we reassess the heat uptake due to atmosphere warming, land warming, and ice melt based on the updated time series in von Schuckmann et al. (2020) (used in AR6 as well). The atmosphere heat content time series shows significant interannual variations, which is not present in the land and ice time series mainly because of the data limitations. Based on the LOWESS method, all these time series are smoothed and long-term variations are revealed. A test adjusting the time span by ±5 years shows negligible impact on our estimates. The estimated total atmosphere warming, land warming, and ice melt from 1958 to 2020 are 5.2 ± 1.7, 25.6 ± 0.3, and 17.0 ± 1.1 ZJ, respectively. Consequently, the total Earth's energy budget (Fig. 8) is estimated to be 422.5 ± 41.2 ZJ from 1958 to 2020 (382.7 ± 34.4 ZJ for 1971–2018), where the uncertainty is quantified assuming the four contributions are independent. However, because of the lack of observations before the 1990s for sea ice volume and ice sheet change (Trenberth 2022), the abovementioned estimate on ice melt contribution is likely a notable underestimation for the uncertainty. Similarly, the land warming estimate depends on the uneven distribution of borehole data and lacks detailed time series, so the above may also underestimate the uncertainty. It should be updated in future after properly accounting for the heterogeneity of the land (Trenberth 2022). Using more reliable land–atmosphere–ice datasets during the 2005–19 period, Trenberth (2022) estimated a nonocean contribution of 30 TW (14.2 ZJ). Combined with the results for OHC here this yields an EEI of 153.9 ZJ with the ocean heat uptake of 139.7 ZJ for 2005–19. Thus, the ocean contributes to ~91% of the EEI from 2005 to 2019.

The central estimate and uncertainty of EEI from this study are smaller than IPCC-AR6 (435 ± 110 ZJ for 1971–2018).

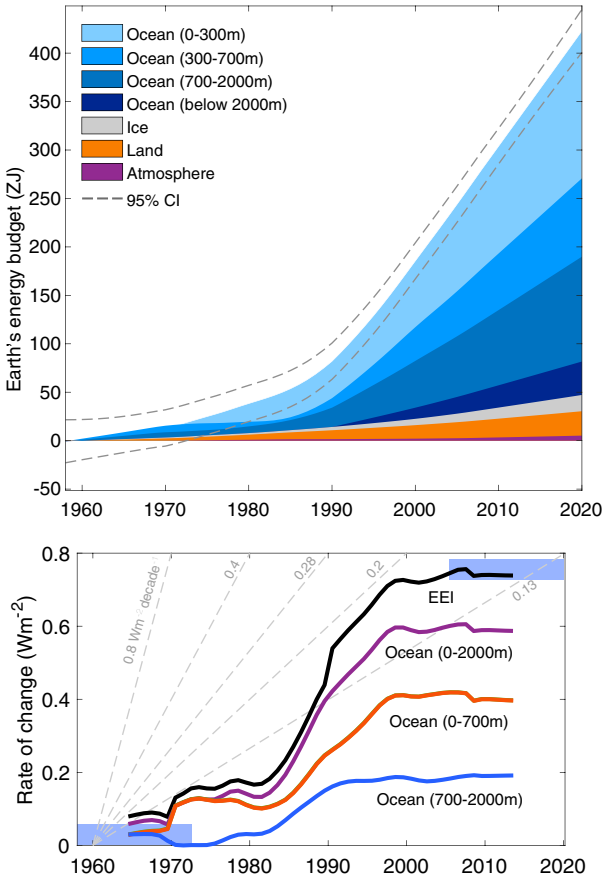

FIG. 8. (top) Estimates of Earth's heat imbalance (ZJ) from the observed changes of different components from 1958 to 2020 based on LOWESS method. (bottom) Rates of change (EEI, OHC 0–2000 m, OHC 0–700 m, OHC 700–2000 m) since 1958 based on LOWESS-smoothed series using centered differences (W m^{-2} averaged over Earth's surface). The span width is 25 years with an effective smoothing of 15 years (denoted by blue shading).

The possible reasons are 1) AR6 combined DOM for 0–700 m and ISH for 700–2000 m to give an estimate of OHC for the upper 2000 m and 2) AR6 calculated the difference between 2018 and 1971 to estimate the rate of change, which is sensitive to start and end points (Fig. 3). The benefit of the LOWESS-based method is the reduction of the impact of the year-to-year variations; 3) AR6 used the ensemble spread among different data products to quantify the uncertainty, which overestimates the true uncertainty because some products are subject to large bias (Durack et al. 2014; Cheng et al. 2019b).

The new estimate suggests a dramatic increase of EEI (Fig. 8, lower panel) from 1980s to early 2000s, manifested in the increased rate of ocean warming within both the 0–700 and the 700–2000 m layers. The maximum rate of increase can be ~0.8 W m^{-2} $decade^{-1}$ around 1990. Loeb et al. (2021) identified an increase of EEI from 2005 to 2019, our result does not necessarily contradict this because the impact of year-to-year variability (such as ENSO and data uncertainty) has not been filtered out. As our LOWESS method effectively smooths out variability less than 15 years, changes within 15 years are not resolved.

6. Discussion and conclusions

A linear trend over long periods (greater than about 15 years) is generally inappropriate for OHC change given the increasing radiative forcing of the climate system due to human activities. It is demonstrated that a single linear or quadratic trend line is very susceptible to the starting/ending conditions and is unable to capture potential acceleration in OHC, and consequently we recommend against their use. However, a linear trend may be useful for intervals from 10 perhaps up to 25 years.

These results are consistent with hypotheses linking the changing rates of increase to increasing human influences. Accordingly, it is essential to have statistical models that emulate these expectations. Moreover, they should be robust and fairly easily applied with transparent metrics of how well they fit the data. This philosophy applies generally, and here it is applied to OHC data. Here we recommend 15 years (25 years for LOWESS) as a useful compromise between resolving the change (increase) in rates and not overfitting the OHC data. Moreover, this time scale conveniently avoids all of the issues with ENSO and volcanic events. All of these results are improvements over a single linear trend because they greatly improve the fit to the data and reduce the uncertainty in the OHC rate estimates. Moreover, they can be tuned to largely eliminate the effects of internal variability.

We then applied PLF and LOWESS to three recent OHC datasets and to CMIP5 OHC estimates and the observations are shown to closely match those of CMIP5-MMM, where a clear acceleration in OHC is also found. This consistency also confirms that the estimated OHC trends are signatures of external forcing (anthropogenic and volcanic) rather than internal variability. A robust increase of ocean warming for the upper 2000 m has occurred since 1958 from about 0 to 0.06 ± 0.08 W m^{-2} for 1958–73 to 0.58 ± 0.08 W m^{-2} in 2003–18.

With the new methods, the rates of OHC change and EEI since 1958 have been recalculated and updated. The total ocean warming for the upper 2000 m is 341.3 ± 21.0 ZJ from 1958 to 2020 (with the 95% confidence interval). The new estimate suggests a dramatic increase of ocean heat uptake and EEI from 1980s to early 2000s. For the most recent period with better data quality (2005–19) and another estimate of land–ice–atmosphere heat content (Trenberth 2022), the EEI is estimated to 153.9 ZJ (10.99 ZJ yr^{-1}) with the ocean heat uptake of 139.7 ZJ (9.98 ZJ yr^{-1}) for 2005–19. This estimate is slightly lower than that using von Schuckmann et al. (2020) in Fig. 8, indicating uncertainty in land–ice–atmosphere heat content.

Note that adding an error bar to a rate estimate is highly desirable to quantify the uncertainty. However, it is not straightforward as there can be systematic biases and structural errors as well as sampling and random errors. This study used a statistical approach. Structural errors may be based on

the physical understanding of the error sources: for example, taking account of the errors due to instrumental bias, gap-filling methods, and choice of climatology to define the final OHC estimate uncertainty (Lyman et al. 2010). Moreover, one can use the ensemble range of individual estimates from different studies to provide an estimate of the uncertainty (Boyer et al. 2016). But this approach only provides an idea of the uncertainty due to the methodology used to compute OHC but it does not give insight on structural errors, as structural errors are by definition common to most methodologies. Physical constraints, such as conservation of mass and energy may also provide insights (Trenberth et al. 2016). Different methodology choices result in a divergence of the uncertainty range in OHC literature (Domingues et al. 2008; Lyman et al. 2010; Levitus et al. 2012; Rhein et al. 2013; Johnson et al. 2017; Cheng et al. 2019b; Meyssignac et al. 2019). Future refinement of the error estimate is highly recommended with the following improvements: (i) use of monthly anomaly time series rather than annual means; (ii) revisiting the error structure when data accumulate: i.e., when there are decades of Argo data; (iii) development of an understanding of the sources of error in OHC estimates; (iv) completion of an intercomparison of the uncertainty estimates based on different methods.

Acknowledgments. This study is supported by the National Natural Science Foundation of China (Grants 42076202, 42122046), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB42040402), and National Key R&D Program of China (2017YFA0603202). The National Center for Atmospheric Research is sponsored by the National Science Foundation. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Cheng et al. 2019b) for producing and making available their model output. The Argo Program is part of the Global Ocean Observing System.

Data availability statement. The IAP data and PLF/LOWESS code used in this study can be found in http://www.ocean.iap.ac.cn/, ISH data in https://climate.mri-jma.go.jp/pub/ocean/ts/, DOM data in https://www.cmar.csiro.au/sealevel/thermal_expansion_ocean_heat_timeseries.html, and LEV data in https://www.nodc.noaa.gov/OC5/3M_HEAT_CONTENT/. The Argo data were collected and made freely available by the International Argo Program and the national programs that contribute to it (http://www.argo.ucsd.edu, http://argo.jcommops.org).

APPENDIX

Test on the Response Function of PLF and LOWESS

Here we provide an idealized test comparing the difference between PLF and LOWESS in their effectiveness of smoothing (i.e., response function). An idealized OHC time series (OHC_{ideal}), with a long-term linear upward trend and different scales of variabilities representing by several cosine functions, are constructed as follows:

$$OHC_{ideal} = OHC_{trend} + OHC_{multi\text{-}decadal} + OHC_{inter\text{-}annual} + Noise,$$

where OHC_{trend} represents the long-term increase in OHC, which is set to a constant linear trend of 5.6×10^{21} J yr^{-1} (0.35 W m^{-2} averaged over Earth's surface) according to the linear trend of the observational OHC change at upper 2000 m after 1958. Suppose that $OHC_{interannual} = A_1 \cos(2\pi/B_1)$, representing the interannual variability in OHC with a period of B_1, and the amplitude of the interannual variability A_1 is simply set to 17.9×10^{21} J with standard deviation of $OHC_{interannual}$ of ~12.6×10^{21} J. The latter is defined as the standard deviation of the residual time series after removing LOWESS-smoothed time series (span width = 25 years) from the original IAP OHC series. B_1 is set to 8 years to represent the interannual variability such as ENSO (Fig. A1a).

Suppose that $OHC_{multidecadal} = A_2\cos(2\pi/B_2)$, representing the low-frequency (multidecadal) variability in OHC that regulates the long-term linear trend in OHC_{trend}. We set A_2 to 47×10^{21} J with standard deviation of $OHC_{interannual}$ of ~33×10^{21} J, which is roughly calculated by the standard deviation of the residual series between LOWESS smoothed time series (bandwidth = 25 years) and the linear trendline of the observational time series (IAP). We set B_2 to 60 years, to mimic the multidecadal variability in OHC change from 1958 to 2016 (Fig. A1a).

"Noise" represents the noise and other unrepresented variability in the time series, which is set to the white noise with a mean of 0 J and standard deviation of 12.6×10^{21} J. In this case, the interannual variability and the noise have similar magnitude as indicated by previous studies such as Boyer et al. (2016).

The question we explore here is, How well can PLF and LOWESS represent the low-frequency variability in this noisy time series of OHC_{ideal}, given the truth of $OHC_{truth} = OHC_{trend} + OHC_{multidecadal}$. Here the truth OHC series include both a linear trend and multidecadal variability, mimic the nonlinear upward trend in OHC record. Therefore, we applied PLF and LOWESS to OHC_{ideal} and compared the resultant fitted series (denoted as OHC_{fit}) with OHC_{truth} (Fig. A1a).

The length of the idealized OHC time series is 2000 years to get a robust estimate of the error (using 3000 and 5000 years give near identical results). Figure A1 presents the first 60 years of simulation, compared with the IAP observational time series from 1958 to 2018. The similarity of the idealized and observed time series indicates that our simulated time series is a reasonable representation of the real world that can be used for the test. However, we note that one caveat is the real OHC time series has an irregular period of the fluctuations on different time scales (i.e., ENSO period ranges from 3 to 8 years and the two episodes are asymmetric; there are also interventions by volcanos), which is not fully represented by the idealized time series. Also, this test neglects the reduction of errors on interannual scales in time series after 2005, when Argo was

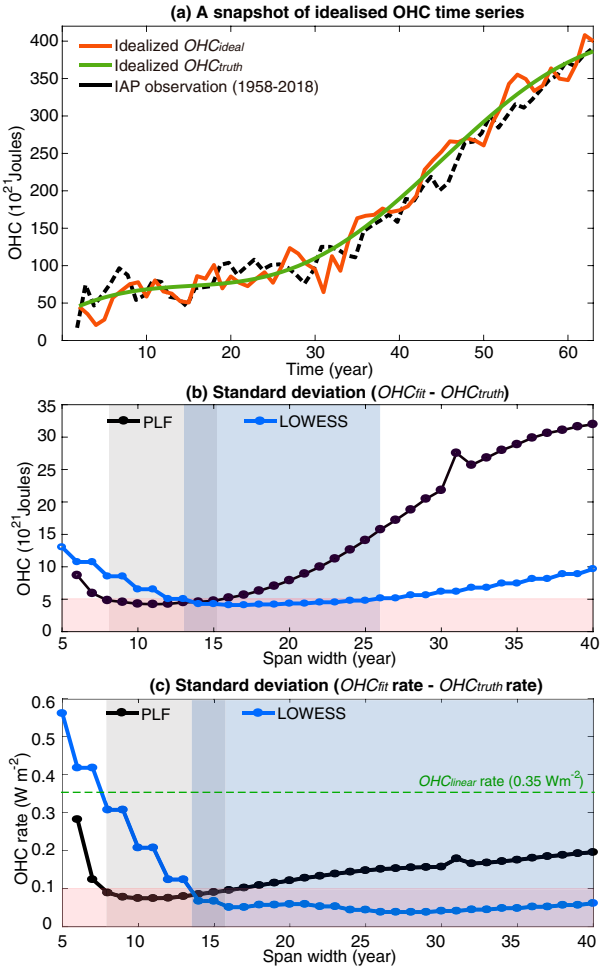

FIG. A1. Test on the PLF and LOWESS based on idealized OHC time series. (a) The first 60 years of one idealized simulation of OHC (OHC_{ideal}; in orange). The truth (OHC_{truth}; in green) contains low-frequency variation of OHC in which we are interested. For comparison, the IAP observational time series within 1958–2018 is included (in dashed black line). (b) Standard deviations of the residual series between OHC_{truth} and OHC_{fit} after application of PLF (black) and LOWESS (blue) using different choice of span width from 5 to 40 years. The shading denotes the choices of span width with smaller error than 5×10^{21} J. (c) Standard deviations of the residual series between OHC_{truth} rate and OHC_{fit} rate (here rate is calculated using centered difference) against span width after the application of PLF (black) LOWESS (blue). The rate is shown as W m^{-2} averaged over Earth's surface. For comparison, the rate of OHC_{trend} is constant and included, which is 0.35 W m^{-2}.

established, in this case, the error using our test is an upper limit for the Argo period.

Both PLF and LOWESS used in this study aim to remove the short-term fluctuations in OHC time series with periods less than 10 years. Therefore, in this idealized test, we examine how wide a span in PLF and LOWESS can effectively remove the function by calculating the standard deviation between the fitted time series (OHC_{fit}) and the truth (OHC_{truth}) (Fig. A1b). At time scales much below 7 years, interannual variation ($OHC_{interannual}$) was not efficiently removed so the error is large (Fig. A1b). From 8 to 16 years of span width for PLF (13–25 years for LOWESS), it appears a "sweet spot" where the interannual fluctuations are muted and values stabilize (Fig. A1b). Beyond the 17 years for PLF (26 years for LOWESS) time scale, error again increases, because the multidecadal variation ($OHC_{multidecadal}$, with period of ~60 years) could not be well represented (Fig. A1b).

This test first indicates different "effectiveness" of smoothing by PLF compared with LOWESS, the difference is 5–15 years. Second, the test indicates that the best choices of span width is 8–16 years for PLF and 13–25 years for LOWESS. Slightly tuning the periods of variability B_1 and B_2 to 5–10 and 50–70 years, respectively, offsets the error structure by several years, but does not change the key results. This supports our choice of 15 years for PLF and 25 years for LOWESS in the main context. We have chosen the higher end of the optimal window taking account of the irregular periods of fluctuation in real OHC time series.

As we are interested in the rate of OHC, so it is also valuable to further provide the accuracy of the OHC rate estimate. Figure A1c presents the standard deviation of the difference between the rate of fitted time series (OHC_{fit} rate) and the truth (OHC_{truth} rate). Within the optimal choice of span width, the accuracy of OHC_{fit} rate using PLF is within 0.1 W m^{-2}, and within 0.75 W m^{-2} for LOWESS. For comparison, in this idealized OHC representation, we have assumed a constant 0.35 W m^{-2} linear trend (OHC_{trend}), and the maximum multidecadal regulation simulated by $OHC_{multidecadal}$ is 0.3 W m^{-2}.

REFERENCES

Abraham, J. P., and Coauthors, 2013: A review of global ocean temperature observations: Implications for ocean heat content estimates and climate change. *Rev. Geophys.*, **51**, 450–483, https://doi.org/10.1002/rog.20022.

Armour, K. C., J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom, 2016: Southern Ocean warming delayed by circumpolar upwelling and equatorward transport. *Nat. Geosci.*, **9**, 549–554, https://doi.org/10.1038/ngeo2731.

Bilbao, R. A. F., J. M. Gregory, N. Bouttes, M. D. Palmer, and P. Stott, 2019: Attribution of ocean temperature change to anthropogenic and natural forcings using the temporal, vertical and geographical structure. *Climate Dyn.*, **53**, 5389–5413, https://doi.org/10.1007/s00382-019-04910-1.

Bindoff, N. L., and Coauthors, 2019: Changing ocean, marine ecosystems, and dependent communities. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, H.-O. Pörtner et al., Eds., Cambridge University Press, 477–587.

Boyer, T. C., and Coauthors, 2016: Sensitivity of global ocean heat content estimates to mapping methods, XBT bias corrections, and baseline climatology. *J. Climate*, **29**, 4817–4842, https://doi.org/10.1175/JCLI-D-15-0801.1.

Caesar, L., S. Rahmstorf, A. Robinson, G. Feulner, and V. Saba, 2018: Observed fingerprint of a weakening Atlantic Ocean

overturning circulation. *Nature*, **556**, 191–196, https://doi.org/10.1038/s41586-018-0006-5.

Cahill, N., S. Rahmstorf, and A. C. Parnell, 2015: Change points of global temperature. *Environ. Res. Lett.*, **10**, 084002, https://doi.org/10.1088/1748-9326/10/8/084002.

Capotondi, A., and Coauthors, 2015: Understanding ENSO diversity. *Bull. Amer. Meteor. Soc.*, **96**, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1.

Cazenave, A., H.-B. Dieng, B. Meyssignac, K. von Schuckmann, B. Decharme, and E. Berthier, 2014: The rate of sea-level rise. *Nat. Climate Change*, **4**, 358–361, https://doi.org/10.1038/nclimate2159.

——, and Coauthors, 2018: Global sea-level budget 1993–present. *Earth Syst. Sci. Data*, **10**, 1551–1590, https://doi.org/10.5194/essd-10-1551-2018.

Chandler, R. E., and M. Scott, 2011: *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences*. Wiley, 392 pp., https://doi.org/10.1002/9781119991571.

Cheng, L., K. E. Trenberth, M. D. Palmer, J. Zhu, and J. P. Abraham, 2016: Observed and simulated full-depth ocean heat-content changes for 1970–2005. *Ocean Sci.*, **12**, 925–935, https://doi.org/10.5194/os-12-925-2016.

——, ——, J. Fasullo, T. Boyer, J. Abraham, and J. Zhu, 2017: Improved estimates of ocean heat content from 1960–2015. *Sci. Adv.*, **3**, e1601545, https://doi.org/10.1126/sciadv.1601545.

——, ——, ——, M. Mayer, M. Balmaseda, and J. Zhu, 2019a: Evolution of ocean heat content related to ENSO. *J. Climate*, **32**, 3529–3556, https://doi.org/10.1175/JCLI-D-18-0607.1.

——, J. Abraham, Z. Hausfather, and K. E. Trenberth, 2019b: How fast are the oceans warming? *Science*, **363**, 128–129, https://doi.org/10.1126/science.aav7619.

——, and Coauthors, 2020: Record-setting ocean warmth continued in 2019. *Adv. Atmos. Sci.*, **37**, 137–142, https://doi.org/10.1007/s00376-020-9283-7.

Cleveland, W. S., 1979: Robust locally weighted regression and smoothing scatterplots. *J. Amer. Stat. Assoc.*, **74**, 829–836, https://doi.org/10.1080/01621459.1979.10481038.

——, 1981: LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *Amer. Stat.*, **35**, 54, https://doi.org/10.2307/2683591.

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, https://doi.org/10.1038/nature07080.

Durack, P., P. J. Gleckler, F. Landerer, and K. E. Taylor, 2014: Quantifying underestimates of long-term upper-ocean warming. *Nat. Climate Change*, **4**, 999–1005, https://doi.org/10.1038/nclimate2389.

Eyring, V., and Coauthors, 2022: Human influence on the climate system. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, in press.

Fasullo, J. T., R. S. Nerem, and B. Hamlington, 2016: Is the detection of accelerated sea level rise imminent? *Sci. Rep.*, **6**, 31245, https://doi.org/10.1038/srep31245.

Foster, G., and S. Rahmstorf, 2011: Global temperature evolution 1979–2010. *Environ. Res. Lett.*, **6**, 044022, https://doi.org/10.1088/1748-9326/6/4/044022.

Gleckler, P. J., and Coauthors, 2012: Human-induced global ocean warming on multidecadal timescales. *Nat. Climate Change*, **2**, 524–529, https://doi.org/10.1038/nclimate1553.

——, P. J. Durack, R. J. Stouffer, G. C. Johnson, and C. E. Forest, 2016: Industrial-era global ocean heat uptake doubles in recent decades. *Nat. Climate Change*, **6**, 394–398, https://doi.org/10.1038/nclimate2915.

Gulev, S., and Coauthors, 2022: Changing state of the climate system. *Climate Change 2021: The Physical Science Basis*, V. Masson-Delmotte et al., Eds., Cambridge University Press, in press.

IPCC, 2013: *Climate Change 2013: The Physical Science Basis.* Cambridge University Press, 1535 pp., https://doi.org/10.1017/CBO9781107415324.

——, 2019: Summary for policymakers. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, H.-O. Pörtner et al., Eds., Cambridge University Press, 3–35.

Ishii, M., Y. Fukuda, S. Hirahara, S. Yasui, T. Suzuki, and K. Sato, 2017: Accuracy of global upper ocean heat content estimation expected from present observational data sets. *SOLA*, **13**, 163–167, https://doi.org/10.2151/sola.2017-030.

Johnson, G. C., and Coauthors, 2017: Ocean heat content [in "State of the Climate in 2016"]. *Bull. Amer. Meteor. Soc.*, **98**, S66–S69, https://doi.org/10.1175/2017BAMSStateoftheClimate.1.

Levitus, S., and Coauthors, 2012: World Ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys. Res. Lett.*, **39**, L10603, https://doi.org/10.1029/2012GL051106.

Li, Y., W. Han, A. Hu, G. A. Meehl, and F. Wang, 2018: Multidecadal changes of the upper Indian Ocean heat content during 1965–2016. *J. Climate*, **31**, 7863–7884, https://doi.org/10.1175/JCLI-D-18-0116.1.

Loeb, N. G., G. C. Johnson, T. J. Thorsen, J. M. Lyman, F. G. Rose, S. Kato, 2021: Satellite and ocean data reveal marked increase in Earth's heating rate. *Geophys. Res. Lett.*, **48**, e2021GL093047, https://doi.org/10.1029/2021GL093047.

Lyman, J. M., and Coauthors, 2010: Robust warming of the global upper ocean. *Nature*, **465**, 334–337, https://doi.org/10.1038/nature09043.

Meredith, M., and Coauthors, 2019: Polar regions. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, H.-O. Pörtner et al., Eds., Cambridge University Press, 203–320.

Meyssignac, B., and Coauthors, 2019: Measuring global ocean heat content to estimate the Earth energy imbalance. *Front. Mar. Sci.*, **6**, 432, https://doi.org/10.3389/fmars.2019.00432.

Nerem, R. S., B. D. Beckley, J. T. Fasullo, B. D. Hamlington, D. Masters, and G. T. Mitchum, 2018: Climate-change–driven accelerated sea-level rise detected in the altimeter era. *Proc. Natl. Acad. Sci. USA*, **115**, 2022–2025, https://doi.org/10.1073/pnas.1717312115.

Outten, S., P. Thorne, I. Bethke, and Ø. Seland, 2015: Investigating the recent apparent hiatus in surface temperature increases: 1. Construction of two 30-member Earth system model ensembles. *J. Geophys. Res. Atmos.*, **120**, 8575–8596, https://doi.org/10.1002/2015JD023859.

Purkey, S. G., and G. C. Johnson, 2010: Warming of global abyssal and deep Southern Ocean waters between the 1990s and 2000s: Contributions to global heat and sea level rise budgets. *J. Climate*, **23**, 6336–6351, https://doi.org/10.1175/2010JCLI3682.1.

Rhein, M., and Coauthors, 2013: Observations: Ocean. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 255–315.

Riser, S. C., and Coauthors, 2016: Fifteen years of ocean observations with the global Argo array. *Nat. Climate Change*, **6**, 145–153, https://doi.org/10.1038/nclimate2872.

Roberts, C. D., M. D. Palmer, R. P. Allan, D. G. Desbruyeres, P. Hyder, C. Liu, and D. Smith, 2017: Surface flux and ocean

heat transport convergence contributions to seasonal and interannual variations of ocean heat content. *J. Geophys. Res. Oceans*, **122**, 726–744, https://doi.org/10.1002/2016JC012278.

Roemmich, D., and J. Gilson, 2011: The global ocean imprint of ENSO. *Geophys. Res. Lett.*, **38**, L13606, https://doi.org/10.1029/2011GL047992.

Schmidt, A., and Coauthors, 2018: Volcanic radiative forcing from 1979 to 2015. *J. Geophys. Res. Atmos.*, **123**, 12 491–12 508, https://doi.org/10.1029/2018JD028776.

Swart, N. C., S. T. Gille, J. C. Fyfe, and N. P. Gillett, 2018: Recent Southern Ocean warming and freshening driven by greenhouse gas emissions and ozone depletion. *Nat. Geosci.*, **11**, 836–841, https://doi.org/10.1038/s41561-018-0226-1.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, https://doi.org/10.1175/BAMS-D-11-00094.1.

Timmermann, A., and Coauthors, 2018: El Niño–Southern Oscillation complexity. *Nature*, **559**, 535–545, https://doi.org/10.1038/s41586-018-0252-6.

Tokarska, K. B., G. C. Hegerl, A. P. Schurer, A. Ribes, and J. T. Fasullo, 2019: Quantifying human contributions to past and future ocean warming and thermosteric sea level rise. *Environ. Res. Lett.*, **14**, 074020, https://doi.org/10.1088/1748-9326/ab23c1.

Tomé, A. R., and P. M. A. Miranda, 2004: Piecewise linear fitting and trend changing points of climate parameters. *Geophys. Res. Lett.*, **31**, L02207, https://doi.org/10.1029/2003GL019100.

Trenberth, K. E., 2022: *The Changing Flow of Energy Through the Climate System.* Cambridge University Press, 319 pp.

——, and Coauthors, 2007: Observations: Surface and atmospheric climate change. *Climate Change 2007: The Physical Science Basis*, S. Solomon et al., Eds., Cambridge University Press, 235–336.

——, J. T. Fasullo, and M. Balmaseda, 2014: Earth's energy imbalance. *J. Climate*, **27**, 3129–3144, https://doi.org/10.1175/JCLI-D-13-00294.1.

——, ——, K. von Schuckmann, and L. Cheng, 2016: Insights into Earth's energy imbalance from multiple sources. *J. Climate*, **29**, 7495–7505, https://doi.org/10.1175/JCLI-D-16-0339.1.

Visser, H., S. Dangendorf, and A. C. Petersen, 2015: A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate." *J. Geophys. Res. Oceans*, **120**, 3873–3895, https://doi.org/10.1002/2015JC010716.

——, ——, D. P. V. Vuuren, B. Bregman, and A. C. Petersen, 2018: Signal detection in global mean temperatures after "Paris": An uncertainty and sensitivity analysis. *Climate Past*, **14**, 139–155, https://doi.org/10.5194/cp-14-139-2018.

von Schuckmann, K., and Coauthors, 2020: Heat stored in the Earth system: Where does the energy go? *Earth Syst. Sci. Data*, **12**, 2013–2041, https://doi.org/10.5194/essd-12-2013-2020.

Zanna, L., S. Khatiwala, J. M. Gregory, J. Ison, and P. Heimbach, 2019: Global reconstruction of historical ocean heat storage and transport. *Proc. Natl. Acad. Sci. USA*, **116**, 1126–1131, https://doi.org/10.1073/pnas.1808838115.

Zhang, L., W. Han, Y. Li, and N. S. Lovenduski, 2019: Variability of sea level and upper-ocean heat content in the Indian Ocean: Effects of subtropical Indian Ocean dipole and ENSO. *J. Climate*, **32**, 7227–7245, https://doi.org/10.1175/JCLI-D-19-0167.1.

Zhu, Y., G. T. Mitchum, K. S. Doran, D. P. Chambers, and X. Liang, 2021: Distinguishing between regression model fits to global mean sea level reconstructions. *J. Geophys. Res. Oceans*, **126**, e2021JC017347, https://doi.org/10.1029/2021JC017347.