nature climate change

LETTERS

PUBLISHED ONLINE: 18 JANUARY 2016 | DOI: 10.1038/NCLIMATE2915

Industrial-era global ocean heat uptake doubles in recent decades

Peter J. Gleckler[1]*, Paul J. Durack[1], Ronald J. Stouffer[2], Gregory C. Johnson[3] and Chris E. Forest[4]

Formal detection and attribution studies have used observations and climate models to identify an anthropogenic warming signature in the upper (0–700 m) ocean[1–4]. Recently, as a result of the so-called surface warming hiatus, there has been considerable interest in global ocean heat content (OHC) changes in the deeper ocean, including natural and anthropogenically forced changes identified in observational[5–7], modelling[8,9] and data re-analysis[10,11] studies. Here, we examine OHC changes in the context of the Earth's global energy budget since early in the industrial era (circa 1865–2015) for a range of depths. We rely on OHC change estimates from a diverse collection of measurement systems including data from the nineteenth-century *Challenger* expedition[12], a multi-decadal record of ship-based *in situ* mostly upper-ocean measurements, the more recent near-global Argo floats profiling to intermediate (2,000 m) depths[13], and full-depth repeated transoceanic sections[5]. We show that the multi-model mean constructed from the current generation of historically forced climate models is consistent with the OHC changes from this diverse collection of observational systems. Our model-based analysis suggests that nearly half of the industrial-era increases in global OHC have occurred in recent decades, with over a third of the accumulated heat occurring below 700 m and steadily rising.

The ocean has stored more than 90% of the Earth's uptake of heat associated with greenhouse-gas-attributed warming since 1970 (refs 14,15). Multi-model comparisons with observed changes in global OHC (refs 3,16) have primarily focused on the upper (0–700 m) ocean, where historical measurements of temperature are relatively abundant and evidence suggests that most of the ocean heating has occurred[15]. Below 700 m, little is known about the consistency between temperature measurements and model simulations. There are substantial barriers to such comparisons however, progress is required and is critical for estimating changes in the Earth's energy budget[6,9], as well as in equilibrium climate sensitivity and the transient climate response[17,18].

We examine changes in global OHC occurring since the industrial revolution, comparing observational estimates of global surface–bottom OHC changes with results from a recent suite of state-of-the-art climate model simulations contributed to the fifth phase of the Coupled Model Intercomparison Project (CMIP5; Methods). To enable comparisons with previous estimates (Methods), we study the upper (0–700 m), intermediate (700–2,000 m) and deep (>2,000 m) layers separately, comparing observational estimates in these layers with CMIP5 'Historical' simulations that are forced with time-varying natural (solar and volcanic) and anthropogenic forcings (for example, CO_2, sulphate aerosols, and land use). We also make use of pre-industrial control (piControl) runs to adjust for simulation drifts, and projections made with the Representative Concentration Pathway (RCP) 8.5 that enable us to make comparisons with more recent measurements (Methods). Our analysis emphasizes results from the CMIP5 Historical+RCP8.5 multi-model mean (MMM; Methods), which reduces some of the effects of model errors, forcing uncertainties, and internal variability.

Simulated OHC increases over the industrial era (circa 1865–2015; Fig. 1a) are principally caused by increases in greenhouse gas forcing, with periods of temporary cooling resulting from the large 1883 eruption of Krakatoa and smaller eruptions of Mount Tarawera (1886), Mount Bandai (1888) and Santa Maria (1902). Responses to large eruptions later in the twentieth century (Mount Agung, 1963; El Chichon, 1982; and Mount Pinatubo, 1991) are clearly evident in the upper layer in both models and observations, but their modelled impact on the deeper layers is much less than for Krakatoa. This is in part because the more recent eruptions occur against a backdrop of rapid anthropogenic warming, but also because deficiencies in the experimental framework lead to prolonged cooling at depth following large eruptions early in the simulation, notably Krakatoa[19].

Annual time series of upper-ocean OHC changes[16,20,21] extend back to circa 1960 (Fig. 1b). Although near-global, these estimates rely on various methods for 'infilling' in data-sparse regions, which are substantial, especially in earlier years[22], and for correcting well-recognized instrument biases[23]. Recent work suggests that the overall heat uptake in these estimates may be biased low as a result of the extremely sparse Southern Hemisphere data earlier in the record[24]. As with the earlier (CMIP3) generation of models[3], throughout much of the 1970–2005 observational record the CMIP5 MMM upper-layer OHC is bounded by these observational estimates, but since the year 2005 the simulated heat uptake starts to become greater than observed. One possible explanation is that post-2000 the volcanic forcing applied to all CMIP5 models is essentially zero, and thus these simulations have not experienced the cooling effects of the moderate eruptions that have occurred during this century[25]. Quantifying the impact of the missing twenty-first-century volcanic forcings on OHC changes will require a new generation of simulations with updated forcing estimates; however, a simple correction estimate can be made based on recent observational work (Methods). Applying this adjustment (thin black line, Fig. 1a,b) reduces the total industrial-era heat uptake by 7% and yields a closer correspondence with the higher end observational estimate[16].

[1]Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, 7000 East Avenue Livermore, California 94550, USA. [2]Geophysical Fluid Dynamics Laboratory, Princeton University Forrestal Campus, 201 Forrestal Road Princeton, New Jersey 08540-6649, USA. [3]NOAA/Pacific Marine Environmental Laboratory, 7600 Sand Point Way NE, Seattle, Washington 98115-6349, USA. [4]Departments of Meteorology and Geosciences & Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, State College, Pennsylvania 16802, USA. *e-mail: gleckler1@llnl.gov

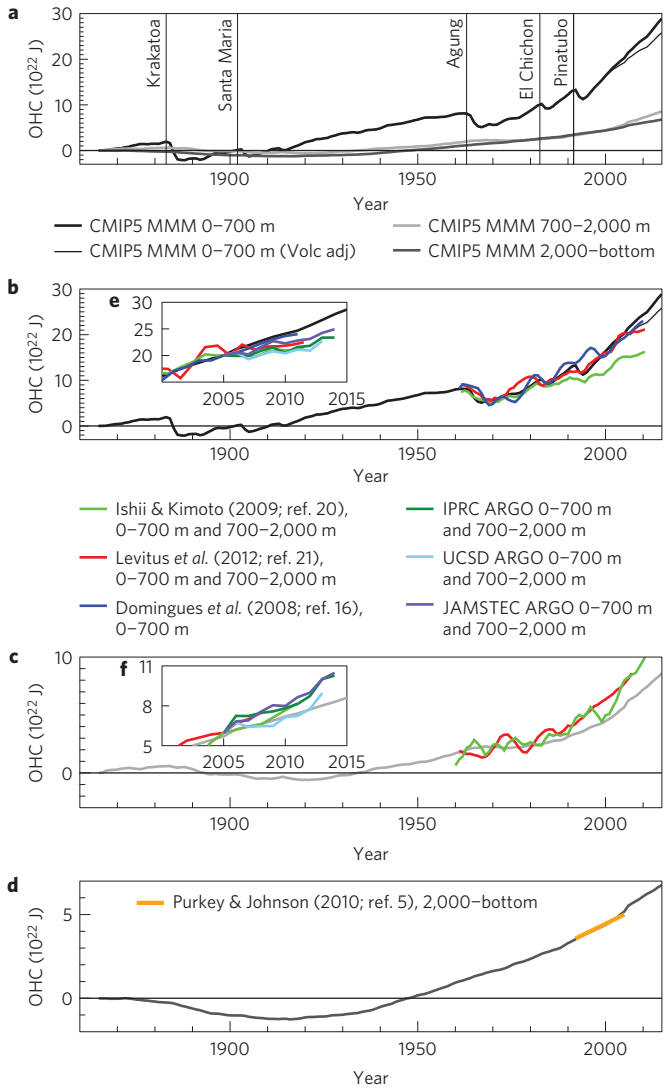

Figure 1 | Simulated and observed industrial-era changes in global OHC for three depth ranges (0–700 m, 700–2,000 m and >2,000 m). **a**, Model-only results (CMIP5 MMM) are for all three depth layers. **b**–**d**, CMIP5 MMM and observations are shown for the individual layers; 0–700 m (**b**), 700–2,000 m (**c**) and >2,000 m (**d**). The CMIP5 MMM 0–700 m and 700–2,000 m results are further compared in insets **e** and **f**, and include three Argo-only estimates (refs 13,34,35), with all observational results set to the MMM value in year 2006 for comparison). An adjusted version (Volc adj; see Methods) of the 0–700 m MMM result is shown in **a** and **b** to provide a simple measure of the impact of missing twenty-first-century volcanic forcing in the CMIP5 simulations. The CMIP5 MMM is obtained from Historical simulations and extended to 2015 using RCP8.5 simulations starting in 2005 (Methods). CMIP5 historical MMM time series are set to zero at 1865. For observations, the upper (0–700 m) and intermediate (700–2,000 m) estimates are set to equal the MMM in 1971, and the deep estimate originates from the 1992 CMIP5 MMM deep value.

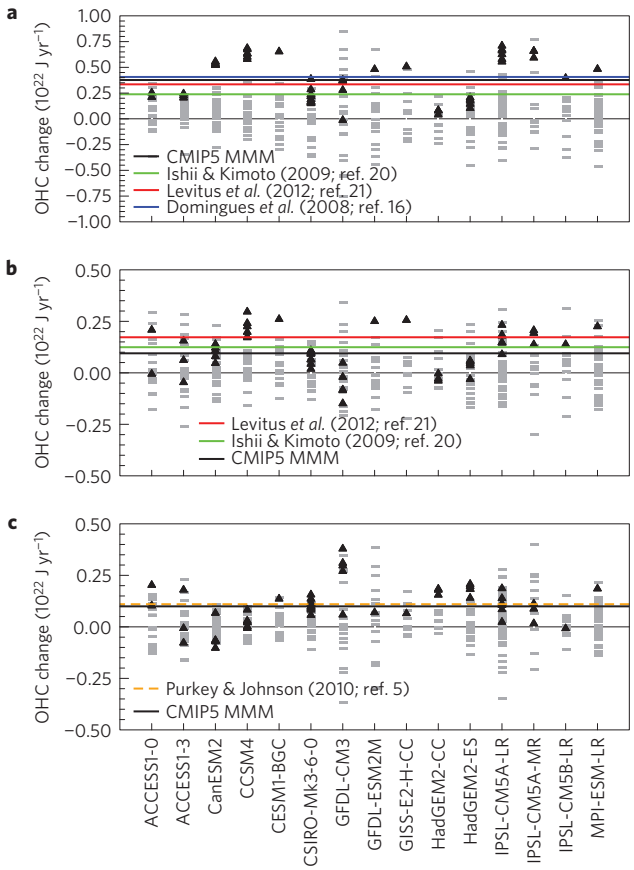

Figure 2 | Simulated and observed global OHC linear trends (35 year; 1971–2005). **a**, 0–700 m range. **b**, 700–2,000 m range. **c**, >2,00 m depth. CMIP5 historical trends for 1971–2005 are shown for individual models (black triangles) including results from multiple realizations where available. Model-based natural variability estimates are computed from non-overlapping 35-year trends of each model's piControl simulation, as shown by thin grey horizontal dashes. Trends estimated from available observations in each layer are shown as coloured horizontal lines: upper and intermediate. Observed trend estimates in the upper and intermediate layer are based on 35 years (1971–2005) whereas the deep layer is based on 14 years (1992–2005) and scaled to units consistent with others. Trend estimates end in 2005 to make use of a larger sample of model simulations (multiple realizations).

Observational estimates of OHC changes in the intermediate layer[20,21] are based on far fewer measurement profiles than are available for the upper ocean[23], with substantial differences evident by contrasting the two available estimates (Fig. 1c). The CMIP5 MMM intermediate layer warming is nonetheless similar in magnitude to these estimates.

Post-2005, substantially improved estimates of OHC are available from the near-global deployment of Argo profilers that provide surface–2,000-m temperature measurements with a much greater sampling density[13], particularly in the Southern Hemisphere. Combining pre- and post-Argo time series is challenging, evidenced by rapid yet coincident changes in both observational coverage and upper-layer OHC from 2000 to 2005 (refs 22,26). As there are ongoing efforts by data experts to improve and merge the pre- and post-Argo measurement systems, it is difficult to gauge the role of global OHC changes in the recent hiatus of surface warming. To mitigate these data inconsistencies for our comparisons with models, we set all observational data to the CMIP5 MMM value in 2006 (Fig. 1e,f), including three data sets that are comprised of modern data only from Argo profilers (Methods). These results suggest that during the well-sampled 2005–2015 period the CMIP5 MMM upper-layer warming is on the high end of the range of observational estimates whereas in the intermediate layer it is on the lower end of the observational range. These comparisons with the limited record of the more highly sampled Argo data are useful; however, there may be substantial decadal-scale variability present in the observations whereas variability is largely averaged out in the MMM (ref. 8).

The deep (2,000 m–bottom) layer OHC change estimate[5] is based on a 1992–2005 linear trend, constructed from differences between full-depth transoceanic section measurements taken during the 1990's World Ocean Circulation Experiment and more recent repeats of these sections by the Global Ocean Ship-based Hydrographic Investigations Program (GO-SHIP). The spatial coverage of these repeat sections, although global, is sparse, and as a result uncertainty estimates for the abyssal trend are nearly as large as the trend itself[5]. The remarkable agreement between the deep OHC change estimate and the CMIP5 MMM (Fig. 1d) may thus be fortuitous. Nevertheless, our view is that these observations are invaluable; for the first time we can quantitatively assess global OHC changes in the deep ocean.

To provide perspective on the forced and unforced contributions to OHC changes, we compare 1971–2005 observed time series (Supplementary Fig. 2) and trends (Fig. 2a–c) to both the CMIP5 historical and pre-industrial controls of each individual model. We include multiple realizations of each model where available, and in most cases find that the impact of natural variability (gauged by the spread expressed in the multiple realizations) is not as large as the inter-model spread (Fig. 2a–c). In the upper layer, the warming trend of some models lies above the range of observational trend estimates, some below, and five models have at least one realization that lies within the spread of observational estimates (Fig. 2a). The case is similar in the intermediate layer where several models have late-twentieth-century cooling that may result from excessive aerosol indirect effects[27].

Previous work based on the comparison of observed and simulated time- and global-mean vertical profiles of temperature suggests that these models are weakly stratified and imply a systematic overestimate of ocean heat uptake[28]. With the OHC trends of historically forced individual models straddling the observational estimates in each layer (Fig. 2), we find no evidence of this systematic positive heat uptake bias across models. This discrepancy may be partially explained by previous work that highlights additional observational uncertainties in long-term OHC change estimates due to sparse measurement coverage[24]. However, it is difficult to reconcile the results from these different analyses as the ocean gains heat through multiple poorly observed mechanisms, making their relative contributions to global heat uptake difficult to quantify.

The late-twentieth-century observed and historically modelled trends in upper OHC are well separated from the distribution of model-based 35-year trend estimates sampled from piControl simulations (Methods; Fig. 2a). There are some exceptions to this pattern, where the distribution of trends is large (GFDL-CM3) or the historical trends are small (both HadGEM2 simulations). For many models, the distinction between historically forced OHC trends and internal variability is less clear in the intermediate and deep layers (Fig. 2b,c). However, returning to the MMM context, we extend the signal-to-noise analysis of previous work[3] to include the intermediate and deep ocean, and find that for the CMIP5 MMM global OHC multi-decadal S/N ratios are highly significant in all three depth layers (Supplementary Fig. 4). More rigorous comparisons of observations with these model-based results will be required for deep ocean detection and attribution work, including expanding subsampling techniques[3,29–31] of upper-ocean warming to address the severe data paucity in the deep ocean.

Comparisons of measurements from the *Challenger* expedition (1872–1876) and Argo (2004–2010) suggest that during the industrial era the global ocean has absorbed 33 (±14) 10^{22} J of additional heat in the upper 700 m (ref. 12). The total CMIP5 MMM heat uptake by the upper ocean since the beginning of the simulated industrial era (~1865–2015) is 29 (±12) 10^{22} J, with a large spread of results in the industrial-era heat uptake of individual models (Fig. 3) primarily resulting from differences in model formulation and external forcings. The range of total heat uptake in

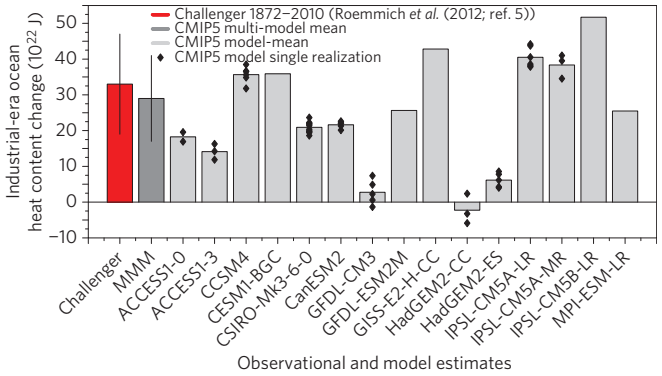

Figure 3 | Observed and simulated industrial-era ocean heat uptake. Observed value and uncertainty estimate[12] is based on measurement differences from the *Challenger* expedition (1872–1876) and Argo (2004–2010). Results from CMIP5 historical simulations for individual models, many with multiple realizations, are for 1865–2005. The CMIP5 MMM is based on the individual model ensemble means, with ±1 s.d. spread shown (thin black vertical lines).

multiple realizations from individual models suggests that internal variability is less of a factor on this longer timescale, which includes a progressively stronger external forcing.

As an approximation to changes in the Earth's global energy budget, we are interested in the evolution of heat uptake by the oceans, which, given changes in observation methods and data coverage over the entire industrial era, are difficult to assess with observations alone. However, having found the CMIP5 MMM historically forced changes in OHC to be consistent with a range of observationally based estimates over a range of selected time periods, we consider the model results for examining the entire period from 1865 to 2015. To do so, we normalize the evolution of simulated OHC in the individual models to their total (1865–2015) ocean heat uptake in an attempt to extract as much useful information from the model suite as possible. Using the CMIP5 MMM heat accumulation (Fig. 4, integrated from the deep through intermediate to upper ocean), we estimate that half of the total global ocean heat uptake since 1865 has accumulated since 1997—nearly coincident with the beginning of the observed surface warming hiatus—with a contribution of 35% from below 700 m. The substantial role of the deeper ocean is consistent with estimates made on shorter timescales using alternative methods[32]. Our simple estimate of the missing twenty-first-century volcanic forcing would suggest that the date of '50% heat uptake' occurs about three years earlier; other deficiencies in how volcanic forcings are included in the experimental design could also play a role[19]. The CMIP5 MMM does not capture the recent hiatus in surface warming[33], which may in part be due to the missing twenty-first-century volcanic forcings[25]. However, as emphasized, the uncertainties in global OHC during the transition to Argo make it difficult to access observational OHC changes during the hiatus period.

We use a carefully chosen suite of observational analyses to compare with global OHC changes simulated by climate models in the upper (0–700 m), intermediate (700–2,000 m) and deep (>2,000 m) ocean. These include multi-decadal estimates derived from ship-based *in situ* mostly upper-ocean measurements, the more recent near-global Argo profiles to intermediate depth, full-depth transoceanic sections, and differences between Argo and data collected during the nineteenth-century *Challenger* expedition. Global OHC changes simulated by the MMM of current-generation climate models are largely consistent with this set of observational analyses. Our model-based analysis suggests that much of the industrial-era global ocean heat uptake has occurred in recent decades, with a fractional contribution from

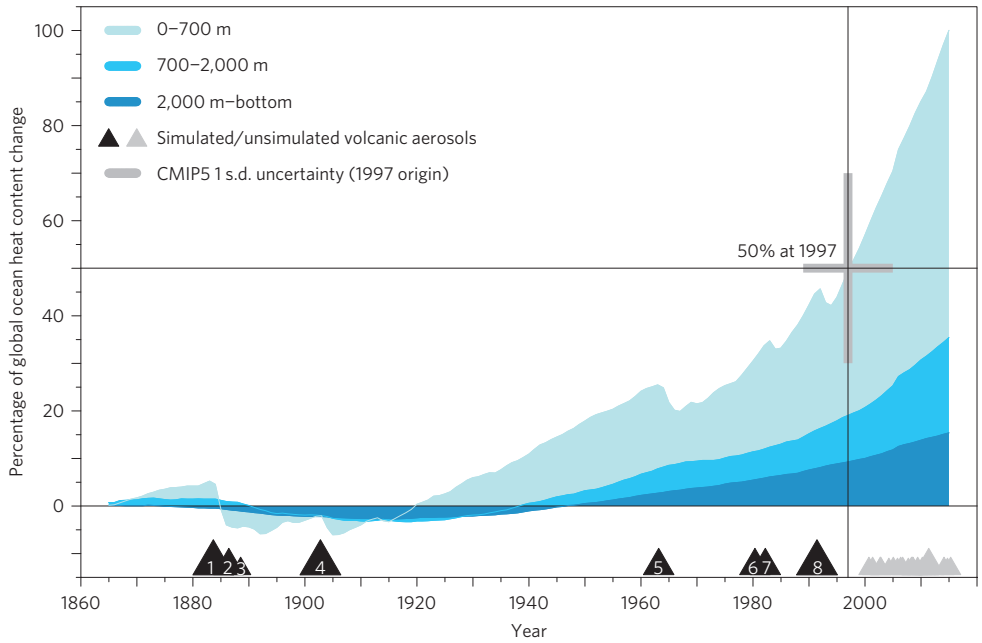

Figure 4 | Ocean heat uptake (percentage of total 1865–2015 change) for the CMIP5 MMM layers. The three shaded wedges are combined similarly to the AR5 change in global energy inventory (ref. 15; Box 3.1 Fig. 1). The thick vertical grey bar represents a ±1 s.d. spread from the CMIP5 simulations about the year (1997) at which the MMM heat uptake reaches 50% of the net (1865–2015) industrial-era increase, and the thick horizontal grey bar indicates the CMIP5 ±1 s.d. spread in the year at which 50% the total accumulated heat is reached. Black (forcing included) and grey (forcing not included) triangles represent major twentieth- and twenty-first-century volcanic eruptions with magnitude represented by symbol size (Supplementary Information).

the deeper ocean (below 700 m) that at present accounts for about 35%, and is rapidly increasing (Fig. 4). Our collective examination of available observations and models highlights the crucially important monitoring provided by the Argo programme, the value of both full-depth profiles from 'Deep Argo' (as more ocean warming penetrates below the sampling depth of Argo), and the repeat sampling by the GO-SHIP programme.

Methods

Methods and any associated references are available in the online version of the paper.

Received 14 October 2015; accepted 9 December 2015; published online 18 January 2016

References

1. Barnett, T. P. *et al.* Penetration of human-induced warming into the world's oceans. *Science* **309,** 284–287 (2005).
2. Palmer, M. D., Good, S. A., Haines, K., Rayner, N. A. & Stott, P. A. A new perspective on warming of the global oceans. *Geophys. Res. Lett.* **36,** L20709 (2009).
3. Gleckler, P. J. *et al.* Human-induced global ocean warming on multidecadal timescales. *Nature Clim. Change* **2,** 524–529 (2012).
4. Pierce, D. W., Gleckler, P. J., Barnett, T. P., Santer, B. D. & Durack, P. J. The fingerprint of human-induced changes in the ocean's salinity and temperature fields. *Geophys. Res. Lett.* **39,** L21704 (2012).
5. Purkey, S. G. & Johnson, G. C. Warming of global abyssal and deep Southern Ocean waters between the 1990s and 2000s: contributions to global heat and sea level rise budgets. *J. Clim.* **23,** 6336–6351 (2010).
6. Loeb, N. G. *et al.* Observed changes in top-of-the-atmosphere radiation and upper-ocean heating consistent within uncertainty. *Nature Geosci.* **5,** 110–113 (2012).
7. Llovel, W., Willis, J. K., Landerer, F. W. & Fukumori, I. *Nature Clim. Change* **4,** 1031–1035 (2014).
8. Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A. X. & Trenberth, K. E. Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nature Clim. Change* **1,** 360–364 (2011).
9. Palmer, M. D. & McNeall, D. J. Internal variability of Earth's energy budget simulated by CMIP5 climate models. *Environ. Res. Lett.* **9,** 034016 (2014).
10. Balmaseda, M. A., Trenberth, K. E. & Källén, E. Distinctive climate signals in reanalysis of global ocean heat content. *Geophys. Res. Lett.* **40,** 1754–1759 (2013).
11. Wunsch, C. & Heimbach, P. Bidecadal thermal changes in the Abyssal Ocean. *J. Phys. Oceanogr.* **44,** 2013–2030 (2014).
12. Roemmich, D., Gould, W. J. & Gilson, J. 135 years of global ocean warming between the Challenger expedition and the Argo Programme. *Nature Clim. Change* **2,** 425–428 (2012).
13. Roemmich, D. *et al.* Unabated planetary warming and its ocean structure since 2006. *Nature Clim. Change* **5,** 240–245 (2015).
14. Levitus, S., Antonov, J. & Boyer, T. Warming of the world ocean, 1955–2003. *Geophys. Res. Lett.* **32,** L02604 (2005).
15. Rhein, M. *et al.* in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) Ch. 3, 255–315 (IPCC, Cambridge Univ. Press, 2013).
16. Domingues, C. M. *et al.* Improved estimates of upper-ocean warming and multi-decadal sea-level rise. *Nature* **453,** 1090–1093 (2008).
17. Sokolov, A. P., Forest, C. E. & Stone, P. H. Sensitivity of climate change projections to uncertainties in the estimates of observed changes in deep-ocean heat content. *Clim. Dynam.* **34,** 735–745 (2010).
18. Church, J. A. *et al.* in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) Ch. 13, 1137–1206 (IPCC, Cambridge Univ. Press, 2013).
19. Gregory, J. M. *et al.* Climate models without preindustrial volcanic forcing underestimate historical ocean thermal expansion. *Geophys. Res. Lett.* **40,** 1600–1604 (2013).
20. Ishii, M. & Kimoto, M. Re-evaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections. *J. Oceanogr.* **65,** 287–299 (2009).
21. Levitus, S. *et al.* World ocean heat content and thermosteric sea level change (0–2000 m), 1955–2010. *Geophys. Res. Lett.* **39,** L10603 (2012).
22. Lyman, J. M. & Johnson, G. C. Estimating global ocean heat content changes in the upper 1800 m since 1950 and the influence of climatology choice. *J. Clim.* **27,** 1945–1957 (2014).
23. Abraham, J. P. *et al.* A review of global ocean temperature observations: implications for ocean heat content estimates and climate change. *Rev. Geophys.* **51,** 450–483 (2013).
24. Durack, P. J., Gleckler, P. J., Landerer, F. W. & Taylor, K. E. Quantifying underestimates of long-term upper-ocean warming. *Nature Clim. Change* **4,** 999–1005 (2014).
25. Santer, B. D. *et al.* Volcanic contribution to decadal changes in tropospheric temperature. *Nature Geosci.* **7,** 185–189 (2014).

26. Cheng, L. & Zhu, J. Artifacts in variations of ocean heat content induced by the observation system changes. *Geophys. Res. Lett.* **41,** 7276–7283 (2014).
27. Zhang, R. *et al.* Have aerosols caused the observed Atlantic multidecadal variability? *J. Atmos. Sci.* **70,** 1135–1144 (2013).
28. Kuhlbrodt, T. & Gregory, J. M. Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. *Geophys. Res. Lett.* **39,** L18608 (2012).
29. Gregory, J., Banks, H., Stott, P., Lowe, J. & Palmer, M. Simulated and observed decadal variability in ocean heat content. *Geophys. Res. Lett.* **31,** L15312 (2004).
30. Pierce, D. W. *et al.* Anthropogenic warming of the oceans: observations and model results. *J. Clim.* **19,** 1873–1900 (2006).
31. AchutaRao, K. *et al.* Variability of ocean heat uptake: reconciling observations and models. *J. Geophys. Res.* **111,** C05019 (2006).
32. Church, J. A. *et al.* Revisiting the Earth's sea-level and energy budgets from 1961 to 2008. *Geophys. Res. Lett.* **38,** L18601 (2011).
33. Flato, G. *et al.* in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. *et al.*) Ch. 9, 741–866 (IPCC, Cambridge Univ. Press, 2013).
34. *2015: IPRC Products Based on Argo Data* (International Pacific Research Center, 2015); http://apdrc.soest.hawaii.edu/projects/argo
35. Hosoda, S., Ohira, T. & Nakamura, T. A monthly mean dataset of global oceanic temperature and salinity derived from Argo float observations. Vol. 8, 47–59 (JAMSTEC (Japan Agency for Marine-Earth Science and Technology), 2008); www.jamstec.go.jp/ARGO

Acknowledgements

The work of P.J.G. and P.J.D., from Lawrence Livermore National Laboratory, is a contribution to the US Department of Energy, Office of Science, Climate and Environmental Sciences Division, Regional and Global Climate Modeling Program under contract DE-AC52-07NA27344. C.E.F. was partially supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, grants DE-SC0004956 (as a member of the International Detection and Attribution Working Group (IDAG)) and DEFG02-94ER61937 and by the National Science Foundation through the Network for Sustainable Climate Risk Management (SCRiM) under NSF cooperative agreement GEO-1240507. G.C.J. is supported by NOAA Research and the NOAA Ocean Climate Observations Program. We thank K. Taylor, B. Santer and J. Gregory for their helpful suggestions concerning our analysis. We acknowledge the sources of observed data used in this study: C. M. Domingues, M. Ishii and M. Kimoto, S. Levitus and T. Boyer, S. Purkey and G. Johnson, D. Roemmich and J. Gilson, S. Hosoda, T. Ohira and T. Nakamura and the International Pacific Research Center. We thank the climate modelling groups (listed in Supplementary Table 1) for producing and making available their model output.

Author contributions

P.J.G., R.J.S. and C.E.F. conceived the experimental design, P.J.G. and P.J.D. performed the analysis, G.C.J. provided data and expertise on deep ocean measurements, and P.J.G. wrote the paper with substantial input from all authors.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.J.G.

Competing financial interests

The authors declare no competing financial interests.

Methods

Observational data. We rely on a diverse collection of observational analyses employing data from the nineteenth-century *Challenger* expedition, a multi-decadal record of ship-based *in situ* mostly upper-ocean measurements, the more recent near-global Argo floats profiling to intermediate (2,000 m) depths, and full-depth repeated transoceanic sections. We use estimated OHC changes between the *Challenger* expedition and Argo period[12]; multi-instrument OHC estimates for the upper[16,20,21] and intermediate[20,21] ocean (1960–near present), and a deep-layer linear change estimate[5] (1992–2005). Three additional 'Argo-only' data sets[13,34,35] (2005–near present) are also used for the upper and intermediate layers and do not include any measurements made with eXpendable BathyThermographs (XBTs), which require bias corrections.

Model simulations. We make use of results from state-of-the-art climate models made available as part of the fifth phase of the Coupled Model Intercomparison Project, CMIP5 (ref. 36). We focus on comparing observations with historically forced (or 'Historical') simulations that nominally begin in 1850 and end in 2005. Where available, we make use of multiple realizations (varying only in initial conditions) of Historical simulations with the same model to help assess the impact of internal variability on our results. Our analysis also requires use of pre-industrial control simulation (piControl) from each model, and, following previous investigators[25], we use the RCP8.5 simulations, so that we can compare the models and observations beyond 2005 (the end of the Historical simulations). The choice of the RCP for the years 2006 to 2015 does not impact the results owing to the long ocean response timescales (ref. 18, Fig 13.8), and the largest number of projection simulations available are with RCP8.5. General information about CMIP5 forcing is included in the Supplementary Information.

Historical simulations from nearly 50 models have been contributed to CMIP5; however, there are only 15 for which all of the necessary data and metadata required for our analysis are available or verifiable. For example, removing simulation drift from the Historical simulations (below) requires a piControl from which each Historical run was initiated. For many models critical information about when the Historical run was initialized from the control run ('branch-time') is either missing or incorrect. Another data constraint relates to combining or 'stitching' the CMIP5 Historical simulations with a CMIP5 future scenario because we have found that it is not always clear which Historical simulation was used to initiate which future projection, or determined that it was inaccurately reported. In summary, missing or inaccurate metadata relating to drift removal and stitching (Historical runs with projections) contribute to reducing our pool of CMIP5 models. Further information about the simulations used in this study is included in the Supplementary Information.

The CMIP5 Historical+RCP8.5 multi-model mean (MMM) is produced after removing the drift from each simulation (below), averaging across all available realizations for each model, and then using these to average across each model. Thus, each model receives equal weight in the MMM irrespective of how many realizations exist.

Removing drift from model simulations. Residual drift associated with the incomplete spin-up of model control runs is removed from globally integrated time series. Following previous studies[3], quadratic fits are computed for the entire control, yielding a drift estimate (Supplementary Fig. 1). This drift is then removed from the original control, yielding an estimate of the modelled natural variability. For each Historical simulation, the contemporaneous section of the control-drift estimate is removed from the Historical simulation to isolate the response to external forcing.

Twenty-first-century forcing adjustment to the simulations. A known deficiency in the CMIP5 Historical simulations is that the volcanic forcing applied to most models did not include any twenty-first-century eruptions. Recent work has demonstrated[37] that available satellite databases neglect substantial amounts of volcanic aerosol between the tropopause and 15 km at middle to high latitudes and therefore underestimate total radiative forcing resulting from the recent eruptions. Incorporating these estimates into a simple climate model, they have estimated the global volcanic aerosol forcing since 2000 to be $-0.19 \pm 0.09\ \mathrm{Wm}^{-2}$. We convert this estimate to joules for the 2000–2015 period, and then apply this simple linear adjustment to the CMIP5 MMM to give us some measure of the impact of missing volcanic forcing. An accurate quantification of this deficiency would require a new suite of model experiments including twenty-first-century volcanic forcing.

Model-generated estimates of natural variability and trends. Distributions of 35-year trends from each model's piControl run (after drift removal) are produced from all non-overlapping segments. The distribution provides a backdrop of model-generated natural variability for each model in Fig. 2 to be contrasted with the Historical simulations and observations. The least-squares trends for the observations and Historical simulations in Fig. 2 are for the period 1971–2005; ending in 2005 permits us to make use of the multiple realizations of Historical simulations. An equivalent figure compares the models and observations for a longer period (1971–2013) with only one Historical+RCP8.5 realization per model (Supplementary Fig. 4), and is consistent with our conclusion that there is no evidence of a systematic bias in OHC trends of the CMIP5 models.

Volcanic explosive index. The estimated magnitude of observed volcanic eruptions (Fig. 4) is shown in terms of a volcanic explosivity index (VEI). Large triangles (for example, 1, 4 and 8) represent eruptions with a VEI = 6, with the smallest triangles representing eruptions with a VEI = 3. Triangle numbering relates to volcanic eruption identifiers contained in Supplementary Table 2.

References

36. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93,** 485–498 (2012).
37. Ridley, D. A. *et al.* Total volcanic stratospheric aerosol optical depths and implications for global climate change. *Geophys. Res. Lett.* **41,** 7763–7769 (2014).