Steric Sea Level Rise and Relationships with Model Drift and Water Mass Representation in GFDL CM4 and ESM4

JOHN P. KRASTING,[a] STEPHEN M. GRIFFIES,[a,b] JAN-ERIK TESDAL,[c] GRAEME MACGILCHRIST,[c,d]
REBECCA L. BEADLING,[e] AND CHRISTOPHER M. LITTLE[f]

[a] *NOAA/OAR/Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey*
[b] *Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, New Jersey*
[c] *Cooperative Institute for Modeling the Earth System, Princeton University, Princeton, New Jersey*
[d] *School of Earth and Environmental Science, University of St Andrews, St Andrews, United Kingdom*
[e] *Department of Earth and Environmental Science, Temple University, Philadelphia, Pennsylvania*
[f] *Verisk/Atmospheric and Environmental Research, Lexington, Massachusetts*

(Manuscript received 29 September 2023, in final form 16 July 2024, accepted 4 September 2024)

ABSTRACT: Density-driven steric seawater changes are a leading-order contributor to global mean sea level rise. However, intermodel differences in the magnitude and spatial patterns of steric sea level rise exist at regional scales and often emerge during the spinup and preindustrial control integrations of climate models. Steric sea level results from an eddy-permitting climate model, GFDL CM4, are compared with a lower-resolution counterpart, GFDL-ESM4. The results from both models are examined through basin-scale heat budgets and watermass analysis, and we compare the patterns of ocean heat uptake, redistribution, and sea level differ in ocean-only [i.e., Ocean Model Intercomparison Project (OMIP)] and coupled climate configurations. After correcting for model drift, both GFDL CM4 and GFDL-ESM4 simulate nearly equivalent ocean heat content change and global sea level rise during the historical period. However, the GFDL CM4 model exhibits as much as a 40% increase in surface ocean heat uptake in the Southern Ocean and subsequent increases in horizontal export to other ocean basins after bias correction. The results suggest regional differences in the processes governing Southern Ocean heat export, such as the formation of Antarctic Intermediate Water (AAIW), Subpolar Mode Water (SPMW), and gyre transport between the two models, and that sea level changes in these models cannot be fully bias-corrected. Since the process-level differences between the two models are evident in the preindustrial control simulations of both models, these results suggest that the control simulations are important for identifying and correcting sea level–related model biases.

KEYWORDS: Ocean; Ocean circulation; Ocean dynamics; Sea level; Ocean models

1. Introduction

Sea level rise presents one of the most existential challenges to coastal communities resulting from anthropogenic climate change. Steric expansion and freshwater mass increases dominate decadal- and centennial-scale sea level trends further modulated by spatially varying patterns of vertical land motion (Hamlington et al. 2020). The density and mass changes also alter ocean circulation, which drives additional dynamic sea level changes. Long-term sea level trends over the past century—both globally and locally—are evident in tide gauge records and will likely accelerate through the twenty-first century depending on the ultimate trajectory of future climate change (Fox-Kemper et al. 2021). The observed global mean sea level has risen 0.12 m over the presatellite era from 1901 to 1990, with a likely range of 0.07–0.17 m. This change is equivalent to an average rate of ~1.35 mm yr^{-1}, which has more than doubled to 3.16 mm yr^{-1} during the satellite altimetry period from 1993 to 2015 (Fox-Kemper et al. 2021). High-frequency events such as storm surges and tidal anomalies superimposed on this background trend of sea level rise provide more imminent threats to coastal communities on weather to seasonal time scales (e.g., Yin et al. 2020).

Thermosteric sea level rise accounts for approximately half of the observed global mean sea level rise, while mass contributions of melting glaciers and land ice account for nearly the other half (Fox-Kemper et al. 2021). Numerical ocean models can represent changes in ocean heat content and local steric effects (Griffies and Greatbatch 2012). However, other physical processes that ocean models do not yet represent influence sea level rise. For example, most models consider a fixed geoid and do not include the localized effects of vertical land motion changes from glacial isostatic adjustment and groundwater extraction (Griffies et al. 2014), which can also influence ocean circulation through gravity changes and ocean mass redistribution. The most significant uncertainty in future projections of sea level rise stems from poorly constrained estimates of glacial meltwater contributions from Greenland and Antarctica (Fox-Kemper et al. 2021; Sweet et al. 2022). The development of dynamic ice sheet models coupled with physical ocean and climate models will progress over the coming decade (Fox-Kemper et al. 2019) and provide critical boundary conditions for future sea level projections. In the meantime, idealized perturbation studies and model intercomparison efforts aim to assess the impact of meltwater on ocean circulation (Swart et al. 2023; Gregory et al. 2016). Offline and probabilistic

---

Corresponding author: John P. Krasting, john.krasting@noaa.gov

DOI: 10.1175/JCLI-D-23-0591.1

frameworks (e.g., Kopp et al. 2023, 2014) that account for vertical land motion and ice sheet melt are used to bridge the gap between numerical model output and future climate projections. These frameworks, however, still rely on coupled climate model projections for their ocean and climate response to forcing.

The kinematic sea level equation represents many key factors contributing to sea level and has proven to be a useful framework for diagnosing ocean model simulations:

$$\frac{\partial \eta}{\partial t} = \frac{Q_m}{\rho(\eta)} - \nabla \cdot \mathbf{U} - \int_{-H}^{\eta} \frac{1}{\rho} \frac{D\rho}{Dt} dz. \qquad (1)$$

According to this relation, the local time tendency of sea level (η) is determined by the addition or removal of mass Q_m, at the ocean surface, the convergence of the depth-integrated volume flux at a given location ($-\nabla \cdot \mathbf{U}$), and the column-integrated material time changes in seawater in situ density (ρ) that lead to steric expansion or contraction. Derivation of this relationship and examples of its application to ocean model simulations are detailed in Griffies and Greatbatch (2012) and Griffies et al. (2014).

Ocean model “spinup” aims to achieve a steady state, or “quasi-equilibrium” with the climate boundary conditions, with the ventilation time scales of the deep ocean dictating the length of this process (Stouffer et al. 2004). Equilibration time scales to surface radiative forcing vary from days for the surface layer to millennia for the oldest waters in the deep Pacific Ocean (Sarmiento and Gruber 2006; Stouffer and Manabe 1999). Throughout this adjustment process (i.e., spinup), internal process-level biases, which vary in nature among ocean models, imprint themselves on the diagnosed terms of this kinematic sea level framework. There are various sources of bias, including models’ dynamical formulations and their representation of ocean physics (Fox-Kemper et al. 2019). The convolution of these biases with the multicentennial equilibration process leads to model-dependent spatiotemporal patterns in preindustrial control simulations that are typically performed follow the models’ spinup. Further, differences in this background ocean state can influence climate models’ transient climate response to anthropogenic forcing, including ocean heat uptake (e.g., Newsom et al. 2023) and subsequent sea level rise.

Preindustrial control simulations are typically conducted under fixed preindustrial radiative forcing and provide a background climate state. This reference state serves as the initialization point for historical and future simulations of sea level rise. A common assumption is that these model biases are relatively stationary in time and that trends in sea level rise obtained from the long control runs can be removed from the historical and future projections in a process called “dedrifting” to yield more appropriate depictions of the models’ response to climate forcing. While dedrifting is a standard and often necessary step in analyzing climate model output, examining the drift present in the preindustrial control simulations provides essential information about the underlying processes relevant to sea level as ocean models advance in complexity and resolution.

It is common for ocean models to have some degree of temperature drift. This drift can occur in localized regions or depths and when considering the ocean’s global volume mean temperature (Irving et al. 2021). Although sea level is measured as a surface quantity, it is an integral of many different processes throughout the water column and is influenced by these local drift processes. The causes of these temperature drifts vary but are often associated with parameterized and resolved mixing and ventilation processes. These temperature drifts also are important for the baseline ocean state for future climate projections as nonlinearities in the equation of state for seawater are nonnegligible sources of differences among future sea level projections (Hallberg et al. 2013). Current generation models are less susceptible to spurious temperature drift than previous generations [e.g., Ocean Model Intercomparison Project (OMIP2); Griffies et al. 2009, Tsujino et al. 2020], although this drift must still be taken into account when diagnosing thermosteric sea level.

The refinement of horizontal grid spacing, sufficient to admit transient mesoscale eddies, poses new challenges in modeling sea level rise. Mesoscale eddying ocean models can directly resolve more features at shorter horizontal length scales, reducing the dependence on subgrid-scale parameterizations. However, current resolutions in global climate models are still insufficient to fully resolve these features at high latitudes (e.g., Hallberg 2013). Eddying ocean models must accurately represent water mass features and the ocean circulation correctly, especially in cases where ventilation is critical to changing the density of seawater and steric sea level properties. Choices regarding the vertical coordinate (z^*, isopycnal, hybrid, etc.) and resolution are also important for global circulation and representing interior water mass properties.

This study’s objective is to analyze sea level drift, and relevant underlying processes, in reanalysis-forced ocean-only and fully coupled preindustrial control simulations in two models with differing horizontal resolution developed by NOAA’s Geophysical Fluid Dynamics Laboratory (NOAA-GFDL). Section 2 describes the baseline configuration of the two models and the differences in their formulation. Section 3 describes the results of ocean-only and coupled model simulations performed with both models. Using basinwide heat budgets and water mass analyses, we examine the drivers of the steric sea level changes and investigate how preindustrial control run drifts lead to biases that potentially explain differences in the historical (1850–2014) heat budgets and sea level trends. Section 4 provides a summary and discussion of these results.

2. Materials and methods

a. Models and experimental design

The Ocean Model version 4 (OM4, Adcroft et al. 2019) developed at NOAA-GFDL is used in this study to evaluate the impact of model resolution on the preindustrial control sea level simulation. There are two configurations of OM4 with different nominal horizontal grid spacings: 0.25°, which was used in the formulation of the GFDL CM4 coupled model

(Held et al. 2019), and 0.5°, which formed the basis of the GFDL-ESM4 (Dunne et al. 2020a). Although ocean-only and coupled simulations exist for each ocean model resolution, the configurations are referred to by their coupled model names throughout this study.

The Modular Ocean Model version 6 (MOM6) code is the basis for the ocean component of CM4 and ESM4 (Adcroft et al. 2019). GFDL's Sea Ice Simulator version 2 (SIS2) and Lagrangian iceberg model (Martin and Adcroft 2010; Stern et al. 2016) complete the representation of the ocean–cryosphere system. The higher-resolution CM4 model is considered "eddy-permitting" and can explicitly represent the effects of mesoscale eddies in the tropics and some of the midlatitudes. However, insufficient resolution of the first baroclinic Rossby deformation radius (Hallberg 2013) limits high-latitude eddy activity in CM4. The mesoscale eddy kinetic energy parameterization (Jansen et al. 2015) represents the effects of eddy mixing throughout the global ocean in the lower-resolution ESM4 model. CM4 and ESM4 include a parameterization for the restratification effects of submesoscale eddy activity (Fox-Kemper et al. 2011). This parameterization was tuned to be slightly stronger in the ESM4 configuration to improve ventilation in the Southern Ocean (Dunne et al. 2020a).

CM4 and ESM4 share the same 75-level hybrid z^*-isopycnal vertical coordinate (Griffies et al. 2020). This coordinate has finer resolution at depth levels in the upper portion of the water column and transitions to isopycnal coordinates in the ocean interior. The z^* regime improves the representation of weakly stratified regions, while the isopycnal coordinate excels in preserving high-latitude and deep-water mass structures. The models are run with the Boussinesq approximation (see section 2b), and the global steric effects of sea level change must be inferred through offline analysis. However, the model includes a nonlinear free surface layer that responds to freshwater fluxes between the ocean and the rest of the climate system. The total sea level change over a given time frame is thus a combination of the dynamic sea surface height that the model directly simulates and a global steric adjustment diagnosed offline. Unlike CM4, the ESM4 model also includes a neutral diffusion parameterization. CM4 and ESM4 employ a parameterization for the mixing effects of internal tides (Melet et al. 2013). An energetics- based planetary boundary layer (ePBL) scheme (Reichl and Hallberg 2018) is used as well as a bottom boundary layer scheme based on Legg et al. (2006). Both models also include a geothermal heat flux at the bottom of the ocean that induces a buoyancy effect at the bottom boundary.

The CM4 model features a 100-km grid spacing, 33-level atmosphere model (AM4, Zhao et al. 2018) with a simplified atmospheric chemistry scheme. The land model (LM4.0) represents land surface hydrology, river routing, and soil and vegetative processes. CM4 includes the Biogeochemistry with Light, Iron, Nutrients, and Gas (BLING) reduced complexity biogeochemical model (Dunne et al. 2020b). BLING's effects are purely diagnostic and do not feed back onto the physical climate. The ESM4 model incorporates updates to the 100-km atmosphere (Horowitz et al. 2020), including an enhanced vertical resolution to 49 levels and a fully prognostic atmospheric chemistry scheme. Hydrological updates and a new interactive vegetation scheme with the climate system are part of an updated land model component (LM4.1) in ESM4. In terms of ocean biogeochemistry, ESM4 employs the Carbon, Ocean Biogeochemistry, and Lower Trophic (COBALT) model (Stock et al. 2020), which incorporates 33 tracers and represents the effects of changing ocean biogeochemistry feedback on the model's climate. The trade-offs between reduced ocean resolution and increased atmospheric chemistry and ocean biogeochemical tracers yield similar computational performance between the two models.

The ocean–sea ice components of CM4 and ESM4 participated in the Ocean Model Intercomparison Project (Griffies et al. 2016) as part of GFDL's contributions to the phase 6 of the Coupled Model Intercomparison Project (CMIP6; Eyring et al. 2016). Present-day temperature and salinity fields from World Ocean Atlas version 2013 (WOA13, Levitus et al. 2014) provided the initial hydrographic conditions and the Coordinated Ocean-ice Reference Experiments, phase II (CORE-II) reanalysis (Large and Yeager 2009; Dai and Trenberth 2002; Dai et al. 2009) under the OMIP1 protocol derived from observational datasets provides the surface boundary conditions spanning the period 1948–2007. The model spinup consisted of cycling the forcing 5 times, with the last cycle being the focus of analysis in this study.

Most of this study focuses on the fully coupled preindustrial control runs of both models. Similar to their ocean-only configurations, CM4 and ESM4 use WOA13 temperature and salinity fields as their initial conditions. Constant 1850 climate forcing provides the surface boundary conditions used to spin up both coupled models, consistent with the CMIP6 protocol (Eyring et al. 2016). The forcing includes fixed concentrations of well-mixed greenhouse gases, aerosols, background volcanic forcing, and a fixed solar constant. Given its relatively high computational cost, the CM4 coupled model was spun up for only 150 years until reaching a stable global sea surface temperature. To accommodate the longer time scales of the coupled carbon cycle, the ESM4 model's ocean component was spun up separately for 400 years before the first year of the coupled preindustrial control run began. The equilibration targets for ESM4 prior to the start of the preindustrial control were that the net radiation at the top of the atmosphere (TOA) stabilized to within 0.1 W m^{-2} of zero and that the net 100-yr smoothed air–sea CO_2 flux was less than 0.01 GtC (Orr et al. 2017).

b. The Boussinesq approximation and steric sea level

The OM4 ocean model employs a Boussinesq approximation in its formulation of the tracer and momentum equations and the physical parameterizations. The net effect is that the ocean velocity is assumed to be nondivergent; i.e., the flow is incompressible, so the model conserves volume rather than mass.

When considering global mean sea level, however, density variations cannot be ignored, as the thermal expansion of seawater is a dominant driver of contemporary and future change. The Boussinesq ocean model can simulate changes in ocean heat content, and therefore, local density changes that can impact ocean circulation are simulated. However, the global steric effect in a Boussinesq model must be diagnosed offline, and we follow the approximation detailed by Greatbatch (1994), Griffies and Greatbatch (2012), and Griffies et al. (2014):

$$\eta^B_{\text{diag}}(\tau) \approx \eta^B(\tau) + \frac{V^0}{A}\ln\left[\frac{\langle\rho\rangle^0}{\langle\rho(\tau)\rangle}\right]. \tag{2}$$

Here, $\eta^B_{\text{diag}}(\tau)$ is the adjusted, diagnosed sea level in a Boussinesq model at time τ, $\eta^B(\tau)$ is the unadjusted sea surface height diagnostic from the Boussinesq model, V^0 is the initial reference volume of the global ocean, A is the global surface area of the ocean, $\langle\rho\rangle^0$ is the initial global mean density, and $\langle\rho(\tau)\rangle$ is the global mean density at time τ. When considering global changes, this approximation is used throughout the manuscript to diagnose the steric, thermostatic, and halosteric contributions to global sea level. This calculation represents the global mean steric expansion. When this global adjustment is added to the models' dynamic sea level "zos," which by construction has a global mean of zero (Griffies et al. 2016), the result is a representation of thermodynamic sea level change that is more comparable to observations, especially on climate (decadal to centennial) time scales.

While the model does not explicitly resolve global steric changes, it is important to note that the local steric effects are represented. As the seawater density changes in a particular location or region, the influence of those changes on the circulation and dynamic sea level is experienced. While the Boussinesq approximation assumes the flow is incompressible, the ocean still experiences a spatially varying density field resulting from changes in temperature and salinity that influence dynamics at the local scale. In this manuscript, where we examine the spatial patterns of steric sea level change, we calculate the change in in situ density relative to a known reference state. Using a common reference density, we then convert the density difference to an effective steric height.

c. *Dedrifting processes*

Throughout this study, model drift is calculated and often removed from the analysis. In all cases, the drift is defined as the linear least squares fit to the 500-yr preindustrial control simulation following the initial model spinup (see the materials and methods section). For global mean steric sea level, the relevant products are first generated based on the methods described in section 3b from the raw temperature and salinity fields. Then, the linear trend is calculated from the control simulation (and removed from the historical simulations, if applicable). For all other fields, including the local steric effects, temperature time series, and heat budget terms, dedrifting is performed at the gridpoint level. The only exceptions to this protocol are the analysis of potential vorticity in section 3b and the analysis of northward heat transport in section 3c where we retain the full unadjusted fields to better highlight the model differences.

It is important to note that all the results based on ocean-only simulations presented in section 3a are not de-dedrifted. The OMIP experimental design (Griffies et al. 2016) does not specify a control simulation protocol for the reanalysis-forced ocean-only simulations. This makes dedrifting an OMIP simulation somewhat ill-posed.

d. *Analysis code*

An open-source Python-based package named "momlevel" contains the sea level analysis routines used in this study. This package aims to curate and standardize sea level analysis tools to support a broader community using MOM6 output. The momlevel package contains routines for diagnosing steric and dynamic sea level and supports data extraction at tide gauge locations. The simulations described in this study use the Wright (1997) equation of state, for which momlevel provides a Python implementation that includes vertical density gradient derivatives as well as thermal expansion and haline contraction coefficients. The code's modular design allows for future integration of other equation of state formulations and can be adapted to accommodate other ocean model output in the future (see the data availability statement for references).

3. Results and discussion

a. *Reanalysis-forced response*

The atmospheric reanalysis provides a prescribed atmospheric state to drive the ocean simulation in ocean–sea ice configurations. This framework allows for more direct comparison among ocean models in their ability to simulate anthropogenic heat uptake and the associated sea level rise by eliminating biases that may arise from coupling to an atmospheric model.

The patterns of steric sea level trends from the reanalysis-forced simulations performed with CM4 and ESM4 (Figs. 1a,c) are comparable with the previous generation CMIP5 CORE-II ensemble mean (Griffies et al. 2014). The global mean steric trend is removed from all models to allow for a better comparison of the patterns. The dominant feature is an east–west dipole in steric sea level across the tropical Pacific Ocean characterized by increases in the western part of the basin and decreases in the eastern part. Although a forced response to climate change cannot be ruled out, decadal variability likely plays a significant role in this pattern (e.g., Meehl et al. 2011; England et al. 2014; Delworth et al. 2015) as this pattern has reversed over the second half of the altimeter record (2006–19; Hamlington et al. 2021). Both models also show increases in steric sea levels in the deep convection regions of the North Atlantic and the Labrador Sea. Increases in these regions may indicate changes in rates of North Atlantic Deep Water formation and ventilation associated with the Atlantic meridional overturning circulation (AMOC) (Yin et al. 2009; Krasting et al. 2016). The richness of the eddy field resolved by CM4 is also evident by the patchy nature in the western boundary currents of the Gulf Stream and Kuroshio and along the northern front of the Antarctic Circumpolar Current (ACC).

The patterns of the total steric level trends are dominated by the upper-ocean thermosteric component (Figs. 1b,d), but CM4 and ESM4 differ from the observationally based estimates for the 0–700-m thermosteric sea level trend pattern

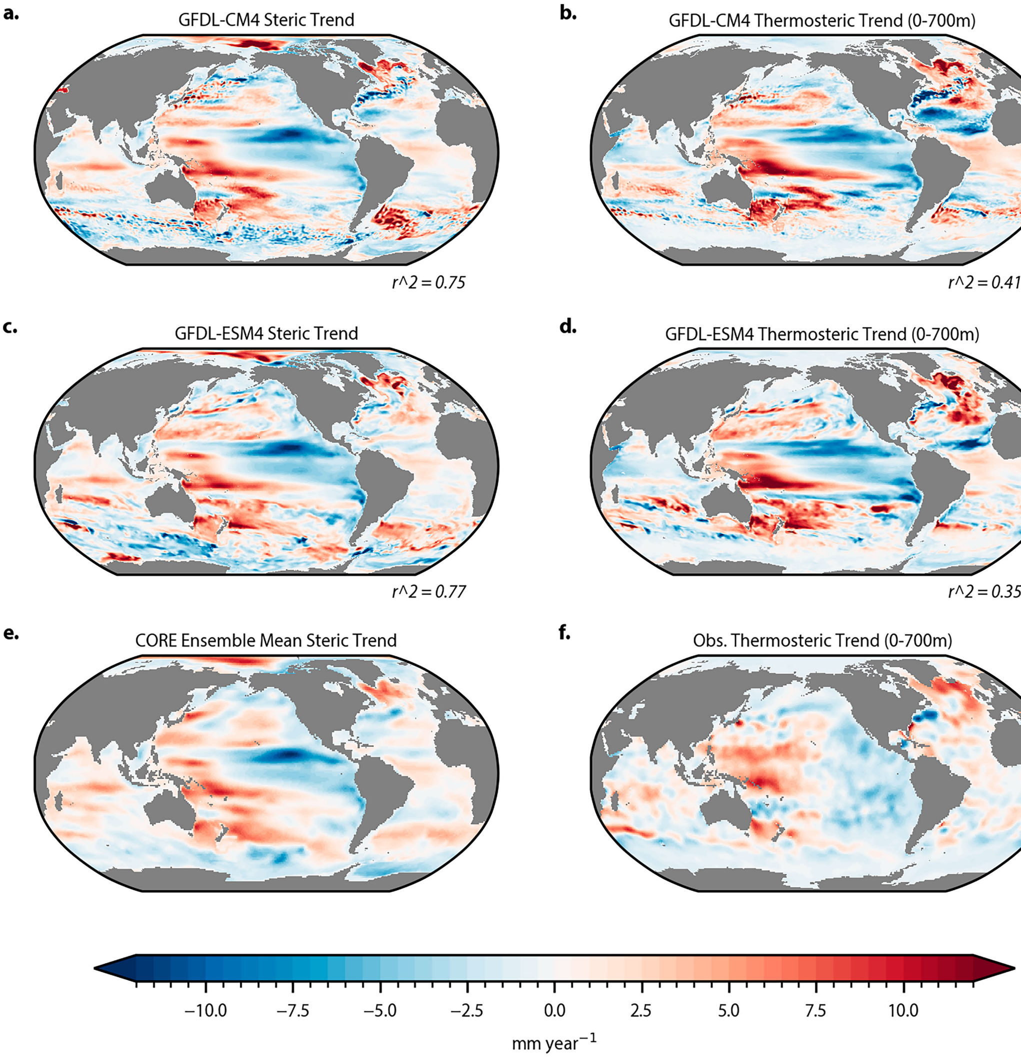

FIG. 1. Patterns of steric sea level trends between years 1993 and 2007 under CORE-II forcing. Steric sea level includes both thermosteric and halosteric components. Linear trends are calculated at each grid point (mm yr^{-1}). Pattern correlations (r^2) are noted for (a) CM4 and (c) ESM4 compared with the (e) CORE-forced ensemble mean of CMIP5 models from Griffies et al. (2014). Thermosteric sea level trends calculated between the surface and 700 m are shown (b),(d) for the same models and compared with observational estimates from (f) World Ocean Atlas (Levitus et al. 2012).

from World Ocean Atlas (Levitus et al. 2012; Fig. 1f) in several regions. The equatorial cold bias in the eastern tropical Pacific is prominent in both models. CM4 and ESM4 exhibit weaker sea level trends in the subtropical regions of the North Atlantic. Both models also exhibit larger than observed thermosteric sea level rise in the western South Pacific off the coast of Australia. Both models significantly underrepresent the observed rise in the globally averaged steric level when forced with reanalysis. The upper ocean (0–700 m) trends in thermosteric sea level over the period from 1993 to 2007 in CM4 (0.52) and ESM4 (0.57 mm yr^{-1}) represent approximately 74%–80% of the global steric change in these models (Fig. 2). However, the observed trend (1.04 mm yr^{-1}) from Levitus et al. (2012) is nearly double the trend of the models. Revised estimates from Domingues et al. (2008) indicate that this estimate from Levitus et al. (2012) may be biased low based on how infilling is performed in data-sparse regions (i.e., the Southern Ocean) and the global 0–700-m thermosteric sea level trend over this same period may be as high as 1.6 mm yr^{-1}. Regardless of which observational benchmark is used, CM4 and ESM4 underestimate the amount sea level rise due to thermal expansion over this period when forced with atmospheric reanalysis.

Underestimation of heat uptake by ocean models forced with reanalysis is a known issue with the OMIP experimental design. Griffies et al. (2014) discuss these issues in more detail, broadly falling into two categories. First, the reanalysis-derived boundary forcing helps eliminate biases arising from

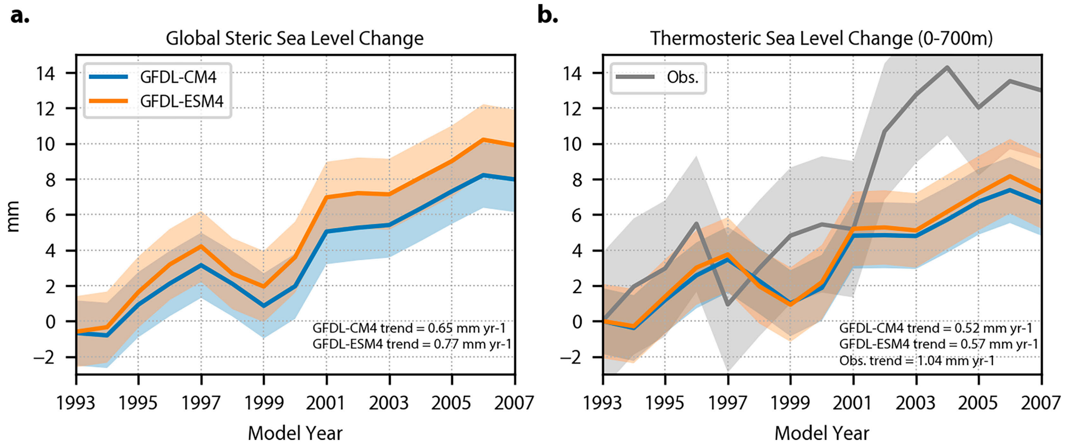

FIG. 2. Global steric sea level rise between years 1993 and 2007 under CORE-II forcing (a) for CM4 (blue) and ESM4 (orange). Steric sea level includes both thermosteric and halosteric components (mm). Global upper ocean (0–700 m) thermosteric sea level rise (b) for both models are compared with observations from World Ocean Atlas (Levitus et al. 2012). Shading denotes 1σ range for both models and observations.

coupling with imperfect atmospheric models. While this facilitates a cleaner comparison among ocean models for many features, it eliminates feedback processes with the atmosphere. These coupled Earth system feedbacks play an important role in ocean heat uptake and are partly responsible for the underestimation seen here. Second, the repeat-cycle experiment design further dampens the rate of ocean heat uptake over the late twentieth century. This is demonstrated by considering the long-term trend of volume-mean potential temperature across the cycles, commonly used to diagnose "model drift." As described in Adcroft et al. (2019), CM4 and ESM4 exhibit minimal temperature drift under CORE-II forcing, with the volume-mean potential temperature drift being 0.0006° for CM4 and 0.008°C century^{-1} for ESM4, both of which are notably lower compared to previous generation global ocean models (Griffies et al. 2009). Compared to the magnitude of the anthropogenic global ocean heat content increase between 1971 and 2018 [329–463 ZJ or $\sim$0.13°–0.18°C century^{-1}; Fox-Kemper et al. (2021), where 1 ZJ = 10^{21} J], the magnitude of this drift is small.

While the time series of volume-mean potential temperature from the models appear stationary in Adcroft et al. (2019), there is a notable decadal structure when plotted over a smaller range of temperatures (Fig. 3). Both models show an increase in temperature over the first three cycles and a decline over the remaining two cycles. As described in the materials and methods section, the models are initialized with observations of temperature and salinity that represent a mean over the period 1955–2014. There is an initial cooling as the warmer ocean in the model adjusts to the CORE boundary forcing that begins in 1948. There is also a cooling adjustment at the start of each new cycle of the CORE forcing. The cooling signal dominates the first several decades of each cycle before the anthropogenic warming signal emerges in these models beginning in the 1990s. To state more simply, throughout each cycle, the anthropogenic warming is offset by the cooling response at the start of each cycle. However, the differences in radiative forcing are not symmetric as the ocean stratification structure makes the ocean more effective at cooling than warming (Pudig et al. 2023). This leads to the overall cooling trend that emerges over the repeat cycles of CORE-II forcing.

The temporal evolution of the horizontally averaged annual temperature anomalies relative to year 1993 (Figs. 4a–c) at each depth level is in reasonable agreement with observations in the upper ocean (above 200 m) where the model is more strongly influenced by the surface boundary conditions that are imposed from reanalysis. The effects of the 1997/98 El Niño event, and subsequent La Niña event, are evident in both models along with a period of sustained warming after year 2000. When the anomalies are plotted relative to 1961, however, strong biases emerge in both models below 200 m. The aforementioned processes leading to long-term cooling resulting from the cyclic OMIP experimental design are

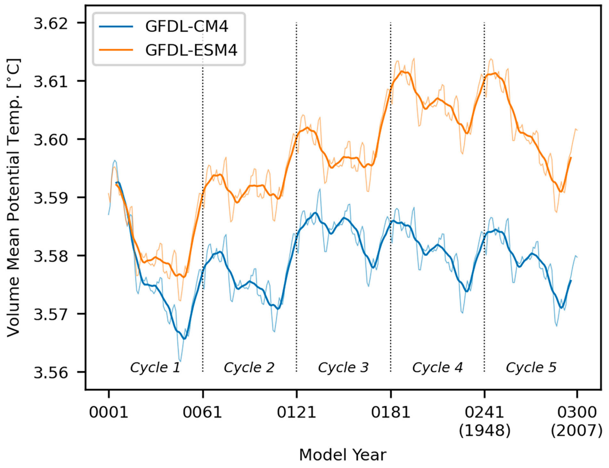

FIG. 3. Global volume mean potential temperature (°C) (10-yr smoothed) in CM4 (blue) and ESM4 (orange) over five repeat cycles of CORE-II interannual forcing.

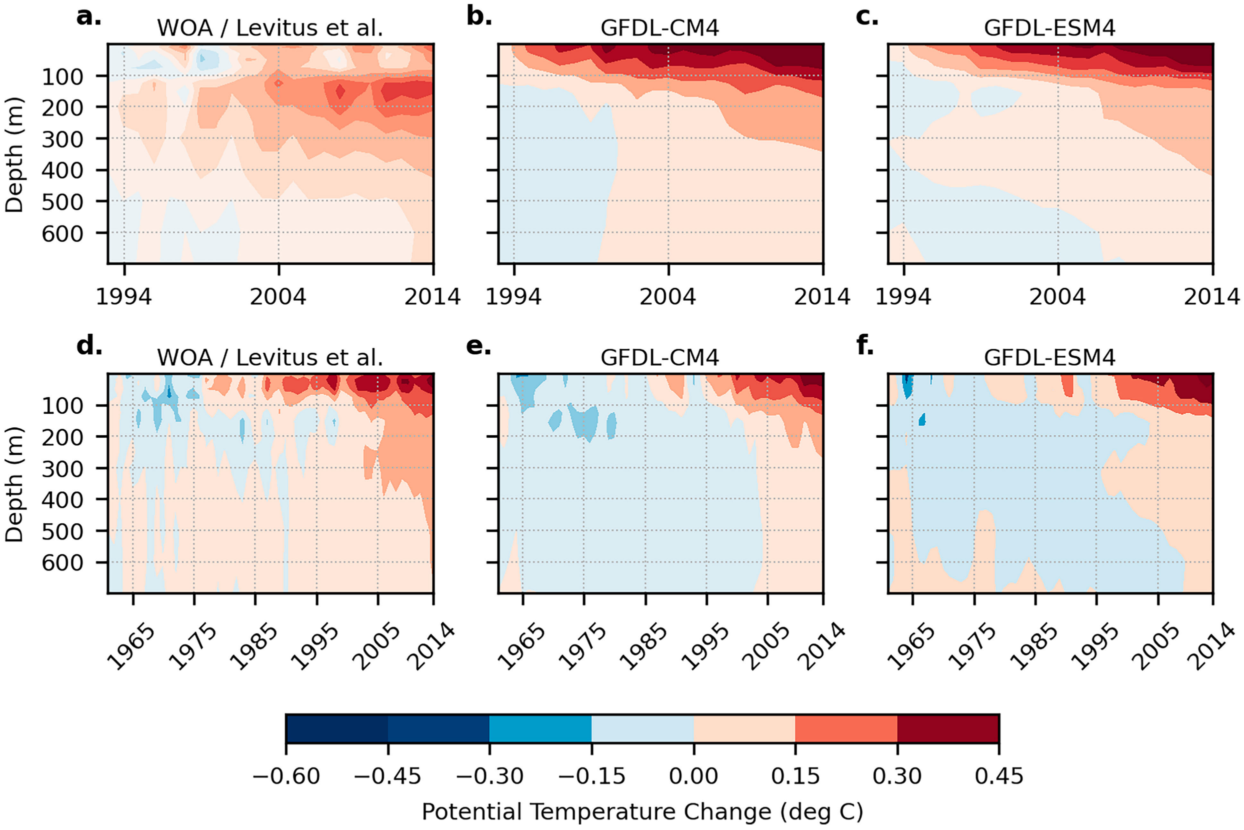

FIG. 4. Horizontally averaged annual mean potential temperature (°C) anomaly relative to (a)–(c) 1993 and (d)–(f) 1961 for World Ocean Atlas (Levitus et al. 2012), and the CORE-II forced CM4 and ESM4. Note the differing time axis for the top and bottom rows.

operating just below the boundary layer in both CM4 and ESM4 and are offsetting a significant amount of the 0–700-m warming taking place over the late twentieth century.

b. Coupled model drift

The volume-mean potential temperature drift expressed as a linear trend over the preindustrial control simulation is larger in the coupled version of CM4 (0.059°C century^{-1}) than in the coupled version of ESM4 (0.017°C century^{-1}). This difference is likely related to differences in the length of ocean spin-up time from the initial present-day temperature and salinity state (400 years for ESM4, 150 years for CM4). After 400 years into the preindustrial control run of CM4, the volume-mean potential temperature is approximately equal to year 1 of the ESM4 preindustrial control simulation (3.7°C), thus highlighting a decelerating drift as the model spinup time lengthens. On millennial time scales, the coupled climate system will achieve a quasi–steady state (Krasting et al. 2018). However, this state will differ between the two models based on their different atmospheric formulations and ocean model configurations. As demonstrated in both observations and the ocean-only simulations, sea level changes associated with ocean heat content changes (i.e., thermosteric) in the coupled models dominate the total steric sea level rise throughout the control simulations. The volume-mean potential temperature (Fig. 5d) is consistent with more steric sea level drift in CM4 than ESM4 as most of the sea level increase is due to thermosteric effects (Figs. 5a,b).

The global halosteric effect is minimal relative to the thermosteric and may be ignored. Using a conceptual framework of homogenizing two buckets of seawater—one with freshwater and one with salty water—(Gregory et al. 2019 appendix 2) demonstrates that the resulting thickness change is almost entirely barystatic (i.e., related to the change in mass). To reinforce this concept and for completeness, the global halosteric sea level drift is shown for both coupled models (Fig. 5c). Plotted on the same scale as the thermosteric drift, the halosteric drift is essentially zero. When plotted on an adjusted vertical axis, the magnitude of the global halosteric drift is slightly larger in ESM4 than in CM4. This is likely due to a coupled equilibration as the amount of snow on land increases during the control run in response to a minor perturbation to the near-infrared albedo of snow on ice that was introduced to address a warm bias over the Southern Ocean and improve deep water formation in this region (Dunne et al. 2020a; Zhang et al. 2020). While the global halosteric effect can be ignored, local trends in halosteric sea level can be considerable—especially in cold regions and high latitudes where the haline contraction coefficient (β) has a relatively greater effect than the low latitudes (e.g., Stewart and Haine 2016; Griffies et al. 2014). Large halosteric trends are especially evident in the Mediterranean Sea, where future model projections of salinity vary

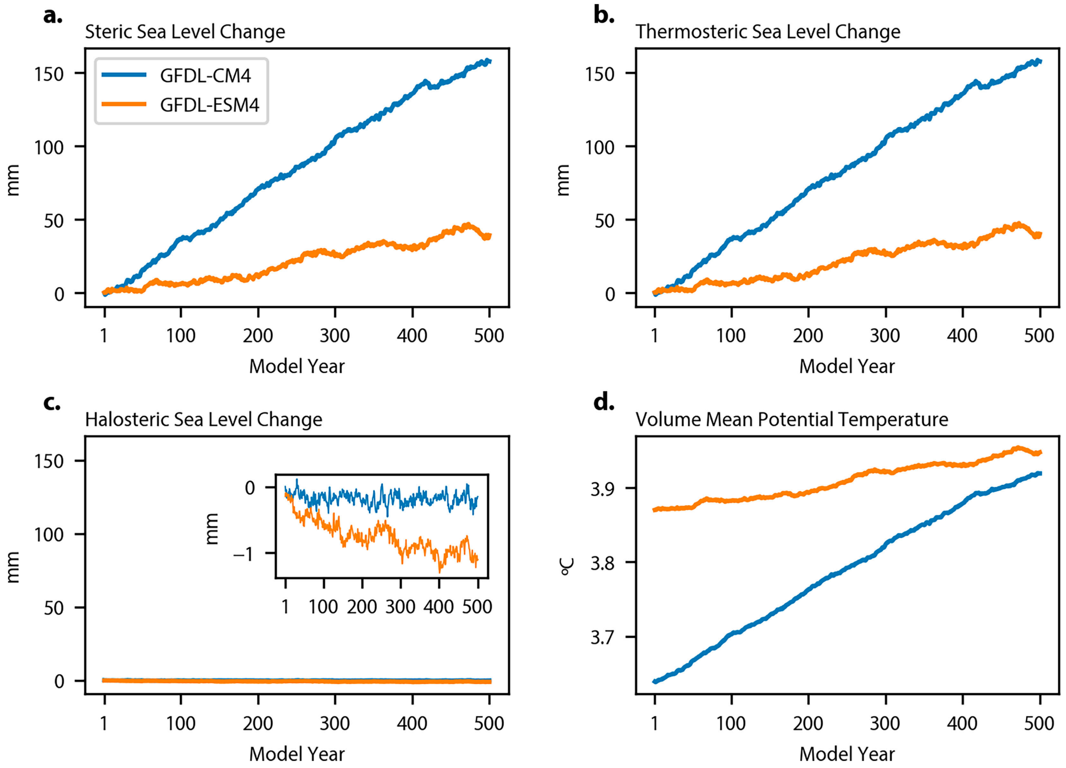

FIG. 5. Time series of steric sea level drift in preindustrial control simulations. (a) Global steric sea level drift is decomposed into the (b) thermosteric and (c) halosteric components (mm). CM4 is shown in blue, ESM4 is shown in orange. (d) Total ocean volume mean potential temperature drift is shown for comparison.

considerably as the changes in freshwater balance and net salt transport through the Strait of Gibraltar are uncertain (Soto-Navarro et al. 2020). The Sea of Japan/East Sea is another region where most CMIP6 models simulate negative salinity trends inconsistent with observations (Jin et al. 2023).

Consistent with the model differences in global mean steric sea level drift, the magnitudes of the regional patterns are larger in CM4 than in ESM4 (Fig. 6). The thermosteric trends dominate the regional drift patterns in both models and are relatively uniform globally, with a few notable exceptions. The regional patterns of thermosteric sea level show considerable drift in deep water regions. The signals of steric drift are less pronounced along the coastal shelf regions where total sea level trends are dominated by mass redistribution (Landerer et al. 2007; Yin et al. 2010; Griffies et al. 2014). In CM4, there is also a region of enhanced thermosteric sea level drift (Fig. 6c) accompanied by halosteric contraction (Fig. 6e) centered between approximately 40° and 30°S in the South Pacific Ocean.

Both models exhibit multicentennial-scale variability in the Southern Ocean deep convection (Held et al. 2019; Dunne et al. 2020a). The periodic release of built-up subsurface heat through full-column convection drives this mode of variability. The signature of the growth of this heat reservoir over time is most apparent in the steric and thermosteric drift patterns of ESM4, particularly in the vicinity of the Ross and Amundsen Seas, where deep convection occurs (Fig. 6d). There is less signal in the halosteric component of sea level (Fig. 6f) as this is primarily a thermally driven process. Both models also show substantial drift in a few of the marginal seas, most notably the Sea of Japan/East Sea, the Mediterranean Sea, and portions of the Caribbean Sea and Gulf of Mexico. While the underlying processes driving these trends are unclear, they have little impact on the global mean values of sea level drift, given their small surface area relative to the global ocean. However, these regions are of significant interest in rising sea levels due to other factors, such as vertical land motion in the case of the Gulf of Mexico. These coastal regions also include major population centers and warrant further investigation.

The in situ density drift throughout the water column highlights the depths at which changes drive the overall drift in steric height (Fig. 7). Most density decreases occur in the abyssal ocean in CM4 below 3000 m, consistent with declines in dense shelf water and Antarctic Bottom Water demonstrated in Tesdal et al. (2023). A few notable regions in CM4 with enhanced drift include the deep Southern Hemisphere tropics below 4000 m and in the Arctic basin. A pronounced tongue in CM4 in the southern Pacific Ocean between 500- and 1000-m depth is present at the boundary between the subtropical and subpolar waters. Several key water masses interact in this region. At higher latitudes, warm Circumpolar Deep Water (CDW) is cooled and freshened as it migrates equatorward to form Subpolar Mode Water (SPMW) and more dense Antarctic Intermediate Water (AAIW). Between

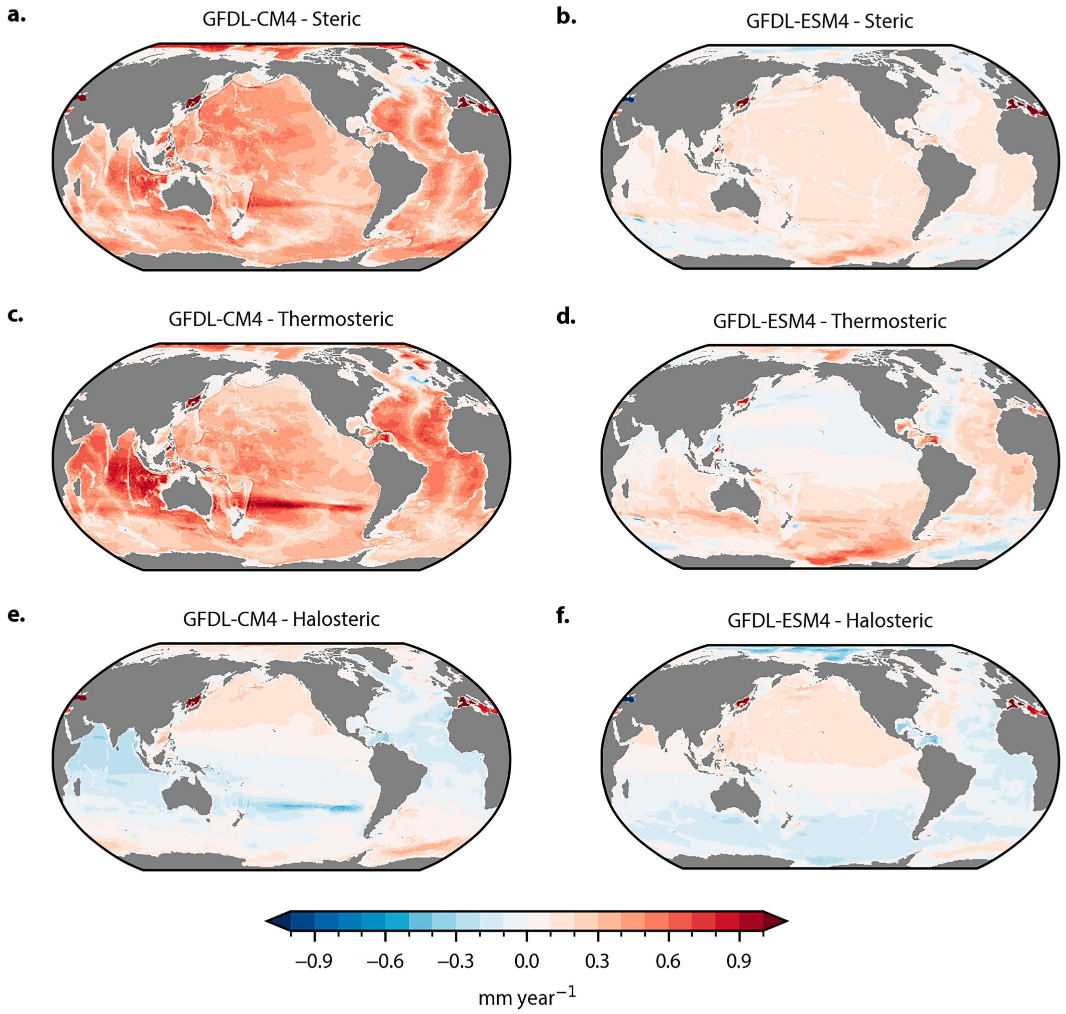

FIG. 6. Patterns of steric sea level drift in preindustrial control simulations. (a),(b) Steric sea level drift is decomposed into the (c),(d) thermosteric and (e),(f) halosteric components (mm yr^{-1}). (left) CM4. (right) ESM4.

40° and 30°S, these waters are further influenced by the broad basin-scale subtropical gyre circulation. The reasons for this signal in CM4 are complex, and we leave the exact details for future studies; however, there are two key processes that may be responsible. First, this result may be indicative of the higher-resolution model's inability to faithfully represent SPMW and AAIW formation. Both of these water masses deliver relatively cool, freshwater below to the subsurface. Without this ventilating water, the existing water column grows warmer and saltier. This lack of ventilation leads to relative warming of the water column, which manifests as an increase in the thermosteric sea level. Similarly, the lack of freshwater input increases the salinity over time, increasing the seawater density in this region and leading to halosteric contraction. Second, the higher-resolution model has more resolved transport from mesoscale eddies, leading to an intensification of the subtropical gyre circulation in the South Pacific. A more intensified gyre circulation is consistent with a downward transfer of heat from the warm ocean surface to deeper layers.

Zonally averaged buoyancy contributions to potential vorticity (PV*) can be used to identify these water mass structures in the South Pacific (McCarthy and Talley 1999) and highlight a distinct difference in the representation of SPMW and AAIW between CM4 and ESM4 (Fig. 8):

$$\mathrm{PV}^* = \frac{fN^2}{g}, \tag{3}$$

where f is the local Coriolis parameter, N^2 is the vertical stratification represented by the square of the Brunt–Väisälä frequency, and g is the gravitational constant. Note that potential vorticity also includes contributions from relative vorticity, but they are omitted from this analysis. For AAIW to form, upwelled (CDW) is modified by sea ice melt, precipitation, and heat fluxes north of the main ACC region and subducts underneath the relatively lower PV SPMW closer to the surface. In ESM4, AAIW is represented by a tongue of relatively higher PV water greater than 60 cm^{-1} s^{-1} penetrating to approximately 1000-m depth in the South Pacific Ocean. In

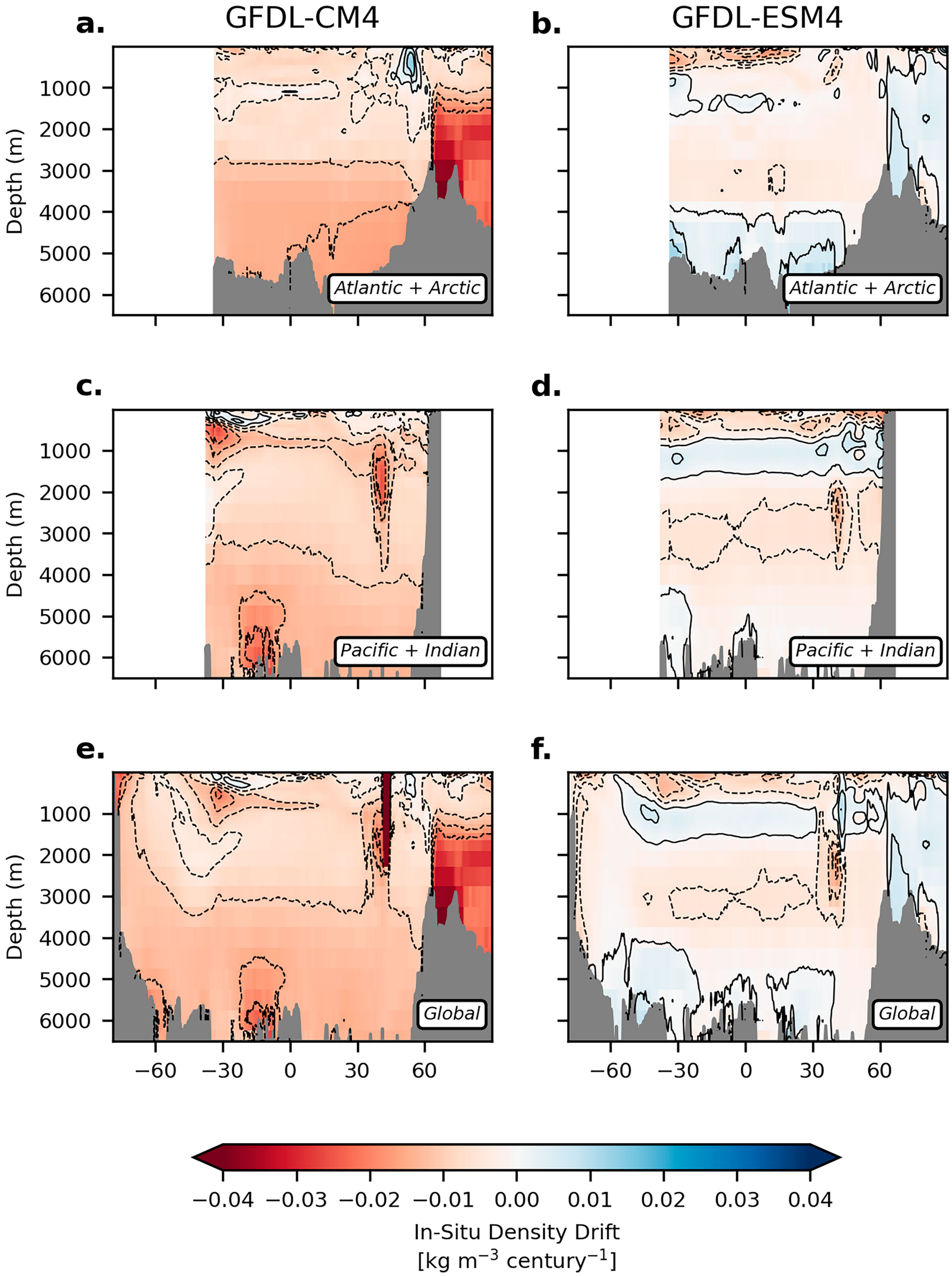

FIG. 7. Basin-averaged depth vs latitude sections of in situ density drift over the preindustrial control simulations for (left) CM4 and (right) ESM4 (kg m^{-3} century^{-1}.). Patterns are shown for the (a),(b) Atlantic + Arctic basins, (c),(d) Pacific + Indian basins, and (e),(f) global domain.

contrast, this PV signature of AAIW is absent and overly thick layers of SPMW are evident in the CM4 model. Enhanced eddy restratification is necessary to weaken near-surface mode water and allow intermediate water to subduct underneath it. This lack of eddy-induced restratification coincides with latitudes poleward of which CM4 experiences limitations in its ability to resolve mesoscale eddy activity explicitly (Hallberg 2013).

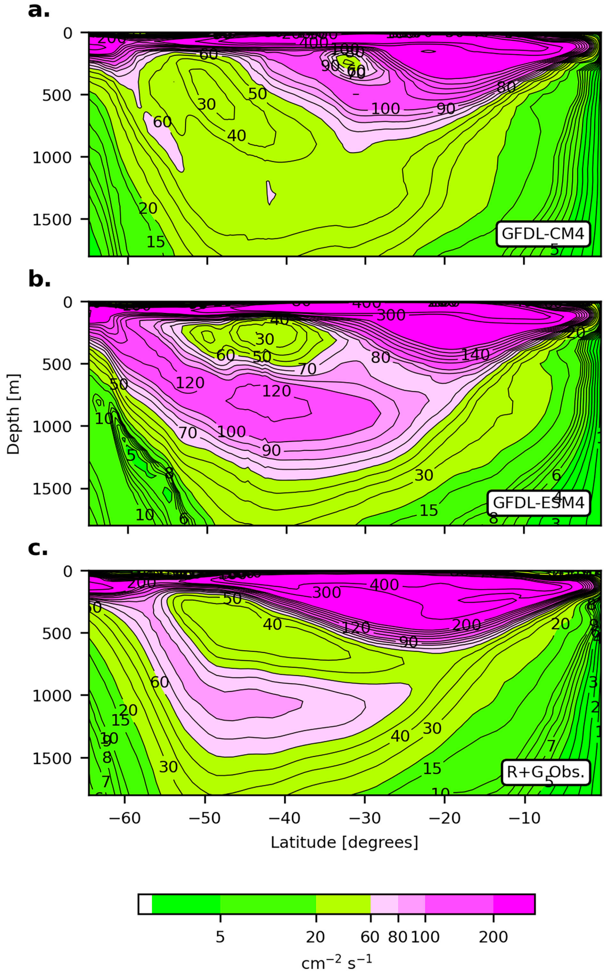

FIG. 8. Buoyancy contribution to potential vorticity (PV^*) averaged over South Pacific sector (100°W°180°) for years 1–500 of the preindustrial control simulation of (a) CM4, (b) ESM4, and (c) gridded ARGO climatology (Roemmich and Gilson 2009). AAIW is represented as the tongue of locally higher subsurface potential vorticity (greater than 60 cm^{-2} s^{-1}) between 60° and 30°S, while SPMW is the bolus of slightly lower PV (30–60 cm^{-2} s^{-1}) located just below the ocean surface layer. Note the nonmonotonic color scale used to highlight the potential vorticity structure of AAIW vs SPMW. The PV^* is calculated using Eq. (3) using the raw (i.e., not dedrifted) model output.

To highlight the dependence of the drift on potentially underrepresented eddy processes, the steric sea level drift at each grid point in both models was binned according to the first surface baroclinic mode Rossby radius of deformation (Fig. 9):

$$R_d = c/f, \tag{4}$$

$$c = \frac{1}{\pi}\int_{-H}^{0} N dz, \tag{5}$$

where c is the wave speed of the first baroclinic mode based on the vertical integral of the buoyancy frequency, N

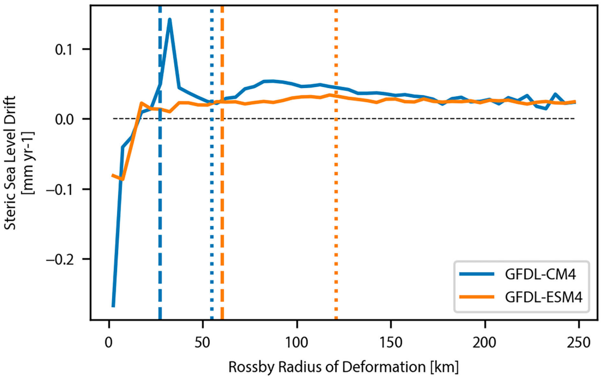

FIG. 9. Steric sea level drifts binned by the first baroclinic Rossby radius of deformation (mm yr^{-1}). CM4 is shown in blue, and ESM4 is shown in orange. The dashed vertical lines denote each model's nominal horizontal grid spacing (Δx, CM4 = 0.25 ≈ 28; ESM4 = 0.5 ≈ 55 km), while the dotted lines denote $2\Delta x$ values.

(Chelton et al. 1998). Maps of the deformation radius for first surface baroclinic mode highlight a meridional gradient of large values in the tropics (Fig. 10). The R_d decreases with latitude, where smaller-scale processes have a larger influence on ocean dynamics. High resolution is required in ocean models to resolve the combined influence of the Coriolis effect and vertical stratification on ocean dynamic processes in these regions.

Examining drift in the context of the Rossby deformation radius facilitates a comparison of the sea level drift rates to the models' ability to represent stirring and mixing due to explicitly resolved eddy activity in CM4 and to the parameterized transport in ESM4. Remarkably, both models consistently exhibit drift rates of about 0.03 mm yr^{-1} for length scales greater than their nominal grid spacing (Δx), which are denoted by vertical lines in Fig. 9. A value of twice the nominal grid spacing ($2\Delta x$) represents a theoretical limit of the models' ability to resolve smaller length-scale features (Hallberg 2013). Both models show an increase in the magnitude of drift at shorter length scales, indicating amplified sea level drift in regimes where models rely on parameterized mixing schemes (ESM4) or cannot explicitly resolve eddy activity (CM4). The importance of accurately representing the effects of mesoscale eddies is relevant to a model's ability to simulate sea level rise, particularly when that eddy activity plays an essential role in establishing ventilating water masses critical for delivering cooler, fresher water to the ocean interior. Further investigation is needed to completely understand the interplay between mesoscale eddy effects and high-latitude ventilating water masses.

c. *Historical sea level*

1) GLOBAL MEAN STERIC CHANGE

The process level differences between CM4 and ESM4 identified in the previous section lead to the question of how well the models agree on their response to anthropogenic

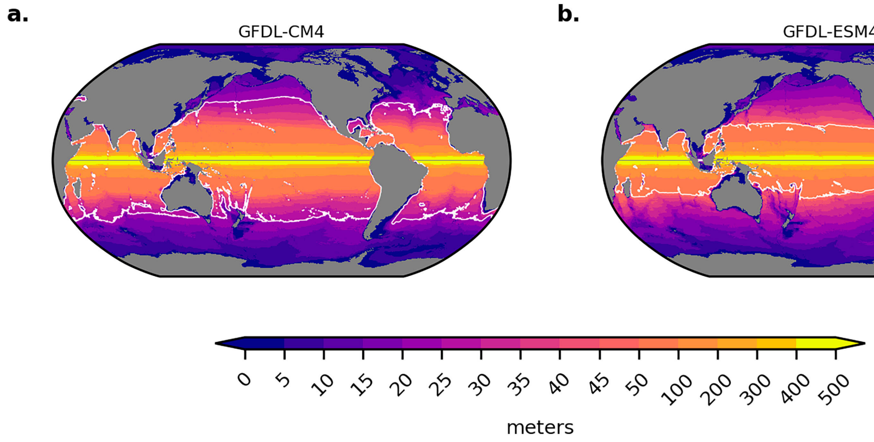

FIG. 10. Rossby deformation radius for first baroclinic mode for (a) CM4 and (b) ESM4. Note the nonmonotonic color scale used to emphasize midlatitude and high-latitude values. White contours denote where the Rossby deformation radius is equal to the horizontal grid spacing for CM4 (25) and ESM4 (50 km).

forcing over the historical period (1850–2014). In this section, both models are dedrifted using the linear trends in global mean steric, thermosteric, and halosteric sea levels calculated from the 500-yr preindustrial control simulations. The global mean historical sea level time series are shown for the CM4 and ESM4 model ensemble means and ranges (Fig. 11). The base period of 1995–2014 is used for the anomalies to align with Jevrejeva et al. (2020), which did not include either GFDL coupled models.

The global mean thermosteric sea level is driven by ocean heat content changes since the preindustrial era. The raw ocean heat content change (i.e., not accounting for drift) is nearly twice as large for the CM4 ensemble mean compared to ESM4 (Fig. 12, dashed lines). The difference between the models' raw heat content reflects the imprint of the stronger drift in CM4 diagnosed from the preindustrial control runs (section 3b) compared with ESM4. Once adjusted for drift, there is better agreement between CM4 and ESM4. Compared to CM4, ESM4 has more climate warming (and by extension ocean heat uptake) in the early twentieth century which could be related to its more sophisticated land model representation of land-use and land-cover changes. Differences between CM4 and ESM4 also emerge in the post-Pinatubo period that are possibly related to differences in the treatment of aerosols and atmospheric chemistry (Dunne et al. 2020a).

Like the CORE-II simulations (section 3a), warming-induced thermosteric sea level rise dominates the total steric response over the historical period (Fig. 11). Both models also agree with representing the late twentieth-century global thermosteric sea level increase consistent with the rest of the CMIP6 ensemble mean shown in Jevrejeva et al. (2020). The record includes a marked increase in global mean sea level rise beginning in the early 1990s that has been previously linked to a reduction in atmospheric aerosol burdens (Slangen et al. 2016) that were partially offset by the cooling effects of the Mt. Pinatubo volcanic eruption in 1991 (Fasullo et al. 2016; Church et al. 2005). The most considerable difference between the global mean response of both models occurs during the first half of the twentieth century, where ESM4 produces more thermosteric sea level rise than CM4, and the ensemble spread also increases. These sea level differences are consistent with the models' different responses in global

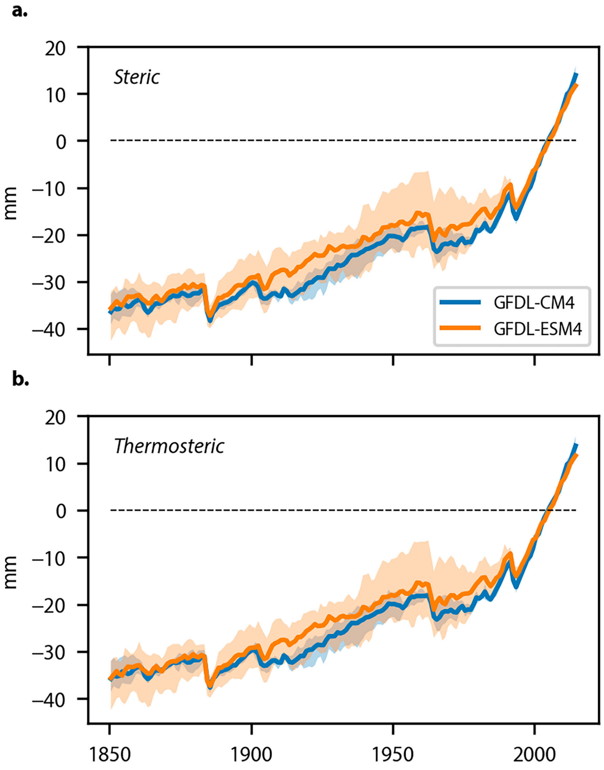

FIG. 11. (a) Global steric and (b) thermosteric sea level anomalies relative to 1995–2014 baseline for CM4 (blue) and ESM4 (orange) (mm). Shading represents the ensemble range, and solid line represents the ensemble mean ($n = 3$ for both models). As detailed in appendix 2 in Gregory et al. (2019), global steric sea level is dominated by global thermosteric sea level, with the global halosteric sea level tiny by comparison. As a result, the two panels are nearly identical.

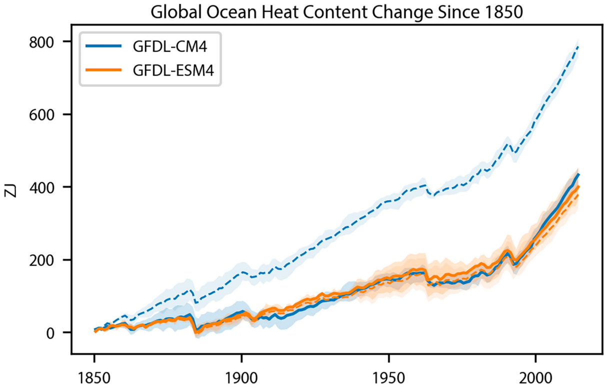

FIG. 12. Global ocean heat content change relative to 1850 for CM4 (blue) and ESM4 (orange) (ZJ). Solid lines represent the dedrifted values based on removing linear trends from the corresponding preindustrial control simulations, while the dotted lines represent the raw model output.

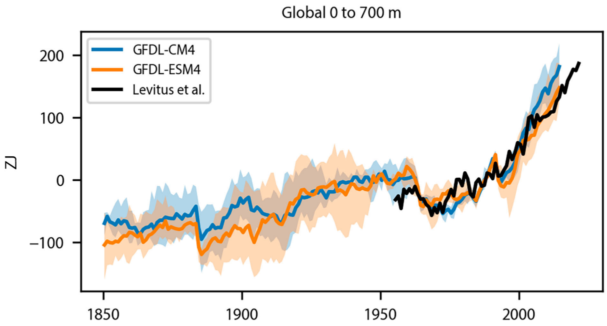

FIG. 13. Upper ocean (0–700 m) heat content anomaly relative to a 1951–2000 baseline for CM4 (blue), ESM4 (orange), and observations (black, Levitus et al.) (ZJ). Shading represents the ensemble range, and the solid line represents the ensemble mean ($n = 3$ for both models). Historical simulations from both models were dedrifted based on linear trends from their corresponding preindustrial control simulations.

mean surface air temperature (Dunne et al. 2020a) and likely reflect differences in the treatment of other coupled processes, such as aerosol forcing or land use and vegetation properties.

Much of the historical warming response over this period occurs in the upper ocean. The drift-adjusted 0–700-m heat content anomalies (1955–2006 base period, Fig. 13) are very similar between both models and are comparable to observationally based estimates (Levitus et al. 2012). For most of the historical period, the ensemble ranges overlap between CM4 and ESM4 except when they diverge in the early 2000s. The observed ocean heat anomalies in the early 2000s are consistent with the ensemble spread of ESM4 but are below the ensemble spread of CM4. Also, the models tend to overestimate the ocean heat content increase in the early 2000s.

The overestimation of heat content increase occurs in both models between the surface and 100 m in the early 2000s (Fig. 14). Both models are overly stratified with respect to temperature and do not represent the relatively stronger trends evident in the observations below 100 m. It is possible that some of the processes leading to subsurface cooling (e.g., parameterized restratification by submesoscale eddies) in the OMIP simulations (Fig. 4) may also be present in the historical simulations. Overall, the drift-adjusted historical simulations show better agreement with observations from 1961 to 2014 than their OMIP counterparts.

2) REGIONAL PATTERNS OF OCEAN HEAT CONTENT CHANGE

The previous section 3c(1) demonstrates the need for dedrifting when interpreting the global and regional historical sea level and ocean heat content changes from the GFDL CM4 and ESM4 models. After accounting for drift, the global mean ocean heat content change (Fig. 13) and thermosteric sea level agree very closely between both models. However, section 3b establishes that the ventilation pathways are different between both models, especially in the southern and South Pacific Oceans (Fig. 8). This section explores how the basin-scale patterns of drift-corrected historical heat content change differ between the two models while still resulting in similar global mean responses.

To answer this question, we consider the vertically integrated heat budget in the models, which is given by

$$\frac{dQ}{dt} = F - \rho_0 C_p \int_{-H}^{\eta} \nabla_z \cdot (\mathbf{u}\theta) dz + \text{residual}. \qquad (6)$$

The vertically integrated total heat content tendency (dQ/dt) equals the net surface (plus a much smaller geothermal) heat flux (F) plus the vertically integrated horizontal convergence of temperature advection at each grid point. These terms are diagnosed online in MOM6 as “opottemptend_2d,” “hfds,” and “T_advection_xy_2d,” respectively. The ESM4 model includes a parameterization for the effects of neutral diffusion, and this term (“opottemppmdiff”) is added into Eq. (6) for this model. This analysis has a residual term accounting for subgridscale processes, but this term is minimal on the basin-integrated scale and is not shown. In summary, this budget breaks apart the total heat tendency into the local component (i.e., surface heating) and the remote effects (i.e., import and export through circulation). Consistent with the rest of this study, these terms are drift-corrected by subtracting the linear trend of the time integral of each term based on the 500-yr control simulation.

It is important to note that this approach does not distinguish between the added anthropogenic heat versus internally redistributed heat considered by Todd et al. (2020), Gregory et al. (2016), and Zika et al. (2021), with the former term referring to the time integral of a surface heat flux anomaly relative to some unperturbed reference state—such as a preindustrial control simulation—when considering models. Instead, the analysis considers the changes at each model grid point, which can either result from heat directly added from an anthropogenic source or through passive advection. Thus, changes in the local heat content result from heat exchange with the atmosphere, moving existing heat into or out of the grid box, or a combination of both.

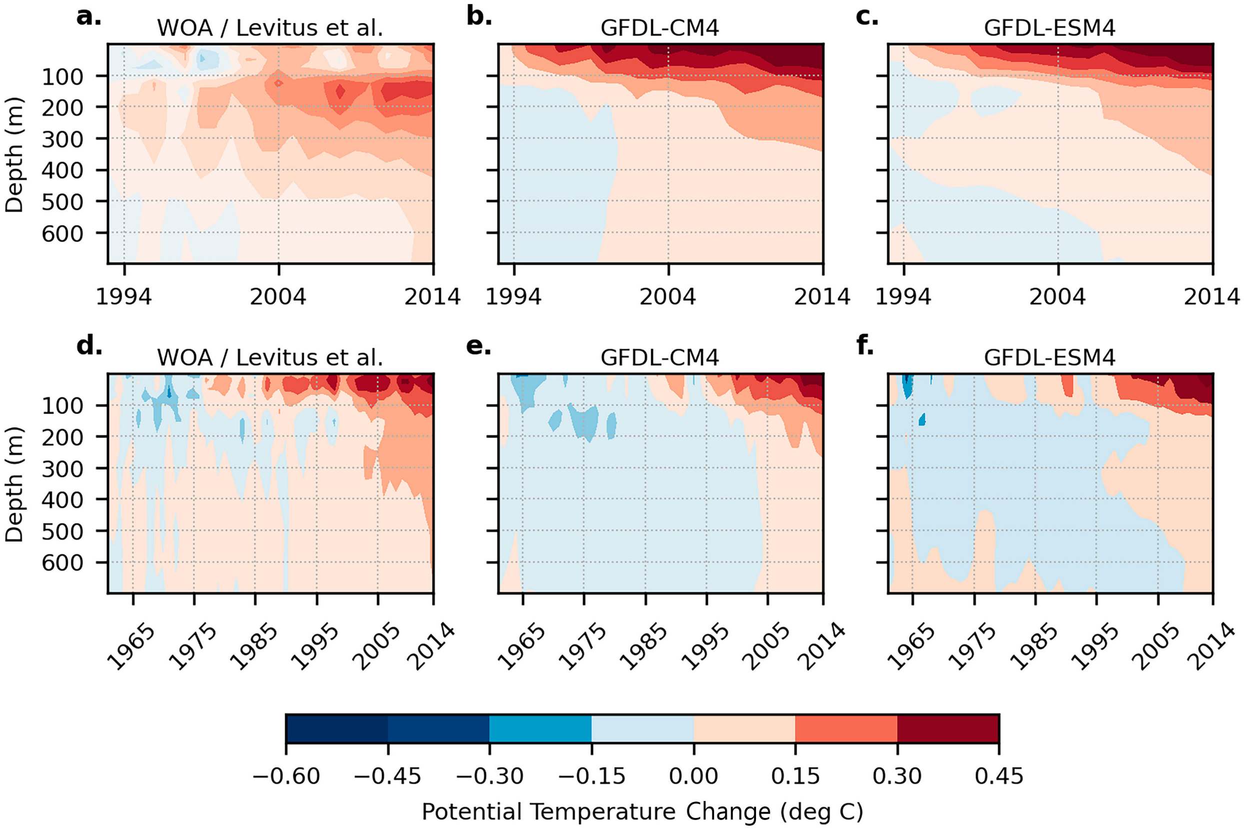

FIG. 14. Ensemble mean (n = 3) upper ocean (0–700 m) heat content anomaly relative to (a)–(c) 1993 and (d)–(f) 1961 for World Ocean Atlas (Levitus et al. 2012), CM4, and ESM4 historical simulations from both models were dedrifted at each grid point based on linear trends from their corresponding preindustrial control simulations (°C).

Figure 15 contains the ensemble-mean (n = 3) time-integrated terms of the vertically integrated heat budget integrated globally and by basin. Dedrifting relative to the preindustrial control was performed at the basin-integral scale. The error bars denote the ensemble range for each term. The global integrals confirm heat conservation in the models and indicate that the global net change in ocean heat content results from surface heating (i.e., the convergence term contribution vanishes). Comparing this term with the results of Fig. 13 indicates that nearly half of the anthropogenic heat uptake occurs in the upper 700 m of the ocean in both models, approximately 200 ZJ versus 400 ZJ based on Figs. 12 and 13 and consistent with previous studies (e.g., von Schuckmann et al. 2020).

The Southern Ocean exhibits the most significant increase in heat content over the entire historical period (1850–2014) through surface heating among the major ocean basins (CM4 = 350–420, ESM4 = 250–350 ZJ). The enhanced surface heating in CM4 is consistent with its colder initial state from being spun up for a shorter time than ESM4. Despite CM4's more extensive surface heating, both models retain approximately 100 ZJ in the net gain in the Southern Ocean. The horizontal divergence works to export the excess surface heating out of the basin in both models. The similarity of the models in the net heat content change likely reflects the heat accumulation in the upper ocean and their similar time scales of heat exchange with the deep ocean. When examining the patterns of net heat accumulation across the Southern Ocean, more heat is accumulating near the continental shelves near the Antarctic coastline in CM4 compared to ESM4 (Fig. 16c), particularly close to the Ross and Weddell Seas. The enhanced heat accumulation on the continental shelves in CM4 may be related to a more vigorous and better-resolved Antarctic Slope Current that isolates the shelf regions from the open oceans (Beadling et al. 2022) and reduced production and export of dense shelf water and Antarctic bottom water (Tesdal et al. 2023).

The Southern Ocean is the region of the most substantial intermodel differences in the heat budget. Although both models accumulate roughly 100 ZJ of heat in the basin, CM4 exhibits considerably more compensation between surface heating (400 ZJ) and convergence (−300 ZJ) compared to ESM4 (300 ZJ heating balanced by −200 ZJ convergence). It was noted in section 3b that the Southern Ocean is likewise a region of considerable differences between the two models in the preindustrial control drift, specifically in the representation of watermass formation, with CM4 exhibiting overly extensive SPMW formation and diminished AAIW. Differences in the models' representation of SPMW and AAIW are evident in the northward heat transport (Fig. 17) in the Pacific

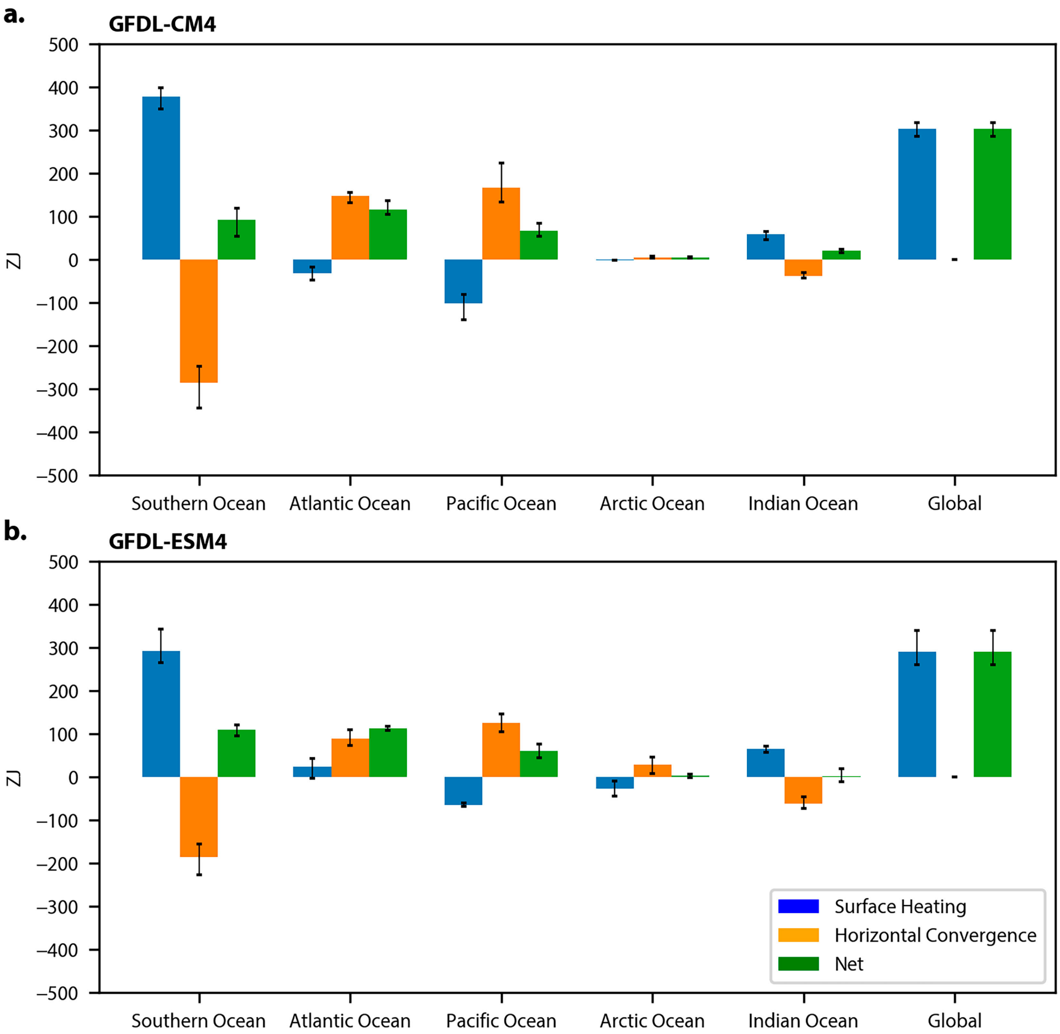

FIG. 15. Basin-integrated, time-integrated terms of the dedrifted vertically integrated heat budget. Terms are integrated over the period 1850–2014 (ZJ) for (a) CM4 and (b) ESM4. Black bars denote the range across the three-member ensemble performed with each model.

sector at the end of the historical simulations. In the subpolar region, poleward of 40°S, the vertical gradient of heat transport is weaker in CM4 than in ESM4. This result is consistent with the broader and deeper influence of SPMW in the CM4, as demonstrated in the previous section with the potential vorticity analysis (Fig. 8). ESM4 has more near-surface northward heat transport away from the main ACC region, which is critical for forming SPMW and AAIW. North of 40°S, CM4 has stronger northward heat transport at the surface and a sharper vertical gradient with depth. The enhanced heat transport is consistent with a stronger subtropical gyre circulation in the higher-resolution CM4 model than in ESM4, which exports more heat out of the Southern Ocean, as demonstrated by the basin-scale heat budget decomposition.

The differences in the processes by which heat is exported from the Southern Ocean identified between CM4 and ESM4 demonstrate the complexities between ventilating water masses and gyre circulations. First, model resolution and its influence on lateral mixing play an important role in stirring the near-surface SPMW and preventing overly thick mode water layers at the expense of deeper ventilating intermediate waters. This is a case where the improved performance in ESM4 might be related to its parameterization of these processes. Model vertical resolution may also be important to consider in this region. Second, similar or even stronger northward meridional heat transport can be maintained by an enhanced subtropical gyre circulation that potentially offsets mode and intermediate water biases in coupled models.

The net heat changes in the Atlantic and Pacific basins are comparable to the Southern Ocean. However, the primary source for Atlantic and Pacific heat changes is the horizontal convergence of heat that comes about through transport from

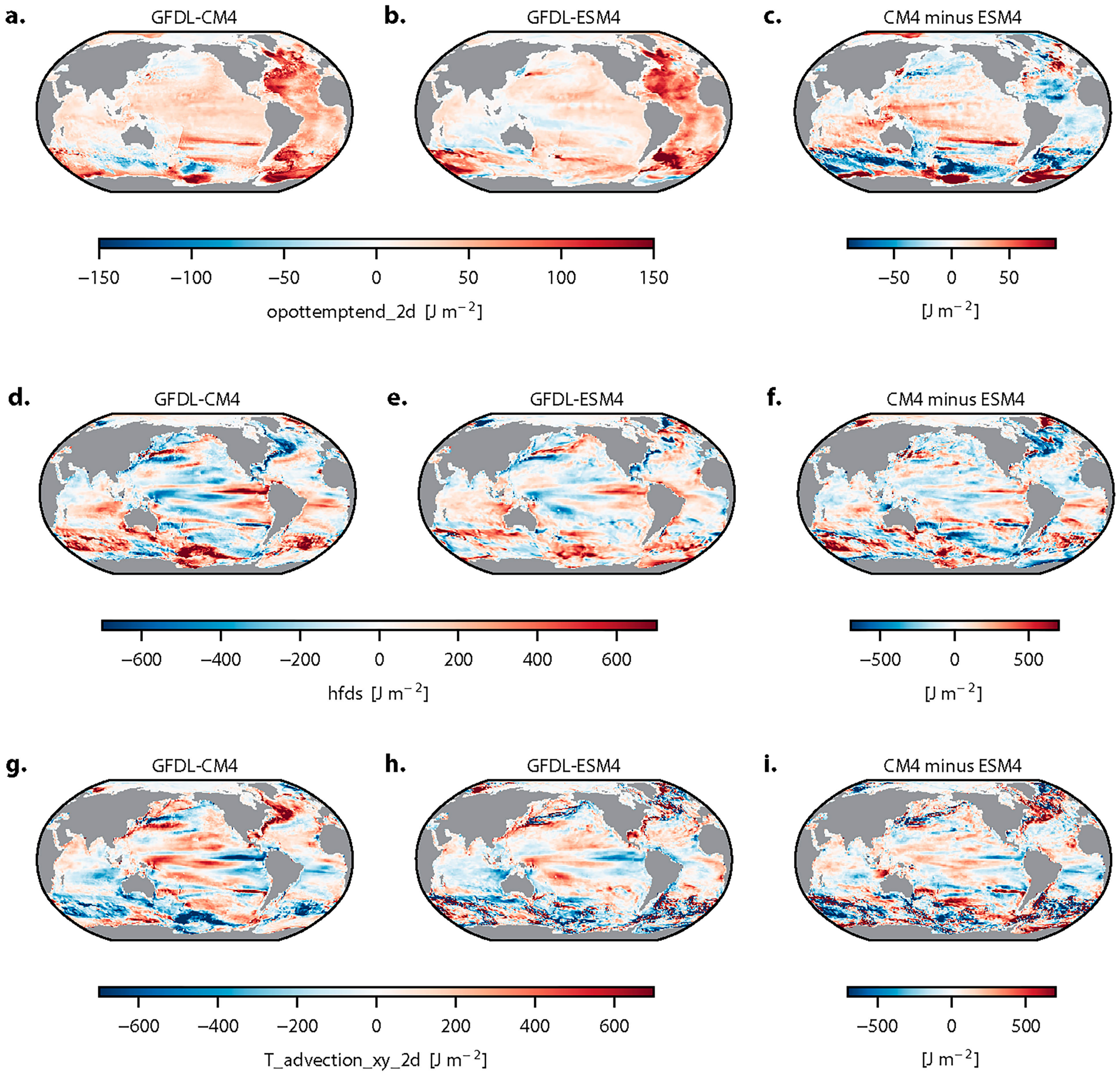

FIG. 16. Spatial patterns of the time-integrated terms of the vertically integrated heat budget. Terms are integrated over the period 1850–2014 (J m^{-2}). (a)–(c) The vertically integrated heat tendency at each grid cell is balanced by the (d)–(f) net heat flux across the ocean surface and (h),(i) net horizontal heat convergence. Results are shown for (left) CM4, (middle) ESM4, and (right) model difference.

the rest of the global ocean, primarily from the Southern Ocean (Figs. 15 and 16). In both models, there is a net heat loss through the surface from the ocean to the atmosphere in the Pacific basin (CM4 = 8–15, ESM4 8–10 ZJ). The western boundary currents, western tropical Pacific, and portions of the Maritime Continent are regions of greatest heat loss (Figs. 16d,e). Conversely, heat gains occur in regions of typically colder sea surface temperatures such as the eastern tropical Pacific, upwelling zones off the coast of South America, and locations north of the Kuroshio. The amplitudes of the heat gain are stronger in the eastern tropical Pacific in CM4 since the model is colder from being spun up for less time and has a more substantial cold tongue bias (Held et al. 2019).

In both models, the Atlantic exhibits ocean heat loss to the atmosphere near the Gulf Stream (Figs. 16d,e). Due to its finer resolution and improved bathymetry, CM4 has better Gulf Stream separation from the coastline of North America near Cape Hatteras compared to ESM4 (Adcroft et al. 2019). In ESM4, the surface heat loss signal extends farther north along the continental shelf regions adjacent to New England and Nova Scotia. CM4 also has a strong ocean heat loss signal in the Labrador Sea region associated with the diabatic cooling

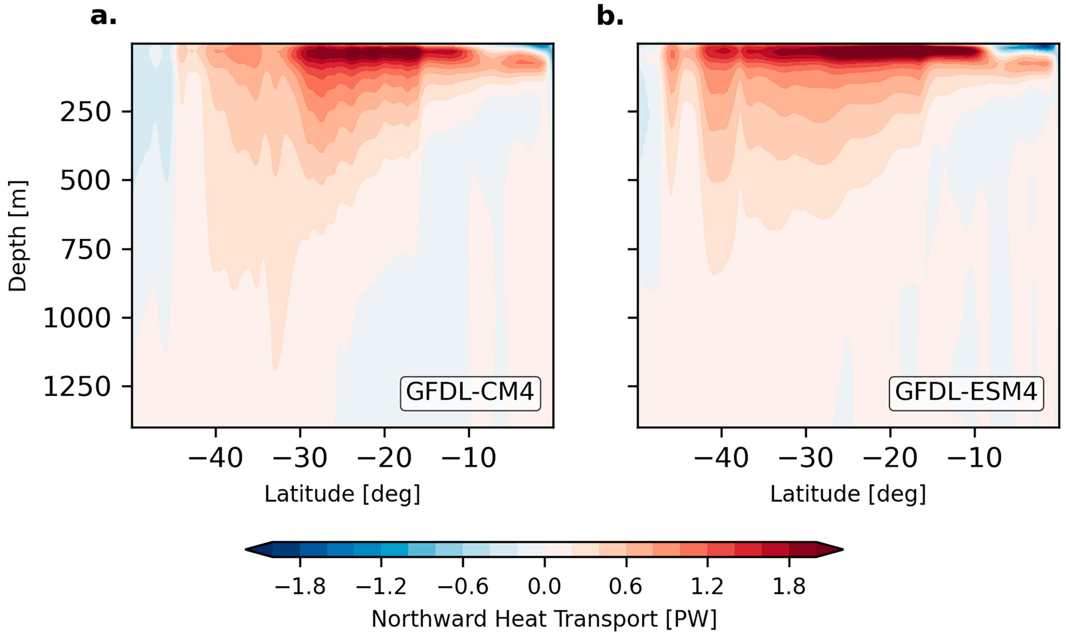

FIG. 17. Northward heat transport (PW) for the Pacific sector (180°–100°W) in (a) GFDL CM4 and (b) GFDL-ESM4. The heat transport in each figure is integrated from the ocean bottom vertically to each model level and then summed horizontally across the region. Results are shown for the ensemble mean ($n = 3$) for each model averaged from 2000 to 2014. Model results were not dedrifted to preserve their native structure.

of surface ocean waters and the formation of North Atlantic Deep Water. These processes are better represented in CM4 compared to ESM4 (Adcroft et al. 2019). Across the rest of the Atlantic basin, broad areas of heat gain in the tropics and subtropics dominate the basin-wide net accumulation of heat (Figs. 16a,b).

The remaining ocean basins have minimal net heat changes over the historical period. In the case of the Arctic Ocean, surface heat loss is balanced by the convergence of heat, although the magnitude of each of these terms is slightly larger in ESM4 versus CM4. Compared with the magnitude of the terms in the other basins, the balance of terms is similar in the Indian Ocean for both models. CM4 exhibits slightly reduced horizontal divergence, leading to more net heat accumulation in the Indian basin. The remaining seas worldwide contribute relatively little to the total ocean heat uptake.

4. Summary and conclusions

This study examines steric sea level in two climate models that differ primarily in their ocean horizontal grid spacing. The GFDL CM4 and ESM4 coupled climate models are based on the OM4 ocean model and have nominal grid spacings of 0.25° and 0.5°, respectively. Using both reanalysis-forced and fully coupled simulations, this study focuses on assessing the impacts of model drift and ventilation pathways on sea level. Both models in their reanalysis-forced ocean–sea ice configurations (Adcroft et al. 2019) agree well with previous generation CMIP5 ocean models in their patterns of late-twentieth-century steric sea level trends but underestimate the observed thermosteric sea level rise, particularly below 200 m.

Both models have minimal temperature drift in their reanalysis-forced configuration. However, this drift is not trivial when comparing ocean heat content and sea level change with the observational record. We show that the temperature drift results from a complex interaction of warming from surface boundary forcing and a cooling trend that becomes apparent over successive OMIP cycles. While the OMIP protocol generally provides a framework that is suitable for comparing output from ocean models with each other and to observations, there are specific challenges when using this experimental design to assess ocean heat content changes, and thus steric sea level, over the late twentieth century.

In both CM4 and ESM4, the ocean potential temperature drift in coupled configurations is larger compared to the reanalysis-forced configurations. The magnitude of the drift is also larger in CM4 compared to ESM4 and may be related to the CM4 model having a shorter spin-up time and more accelerated drift than the ESM4 counterpart. Over 500 years of preindustrial control forcing, differences in the steric sea level drift patterns emerge between the models. We examine the drift as a function of the local Rossby radius of deformation, which is used as a proxy for the length scales at which one would expect the model to fully resolve the effects of explicit ocean dynamics. In both models, the magnitude of the steric sea level drift is greatest at small Rossby length scales, which are typically found at higher latitudes. In the ESM4 model, this could indicate an imperfect parameterization of the effects of mesoscale eddies. In the CM4 model, which can explicitly resolve at least some of the mesoscale eddy effects at these resolutions, this result suggests insufficient horizontal resolution in these regions. In the CM4 model, a lack of eddy restratification at high latitudes is linked to a pronounced bias in the model's ability to represent the subsurface ventilating SPMW and AAIW water masses in the South Pacific. Without the presence of these colder water masses, there is a tendency for the middepths of the ocean to warm in the basin. Exacerbated by a potentially invigorated subtropical gyre circulation, this warming in CM4 leads to a positive steric sea level drift in this region.

It is typical—and necessary—to remove the effects of temperature drift when analyzing historical and future projections of sea level using coupled models. Once corrected for this drift, ESM4 and CM4 produce remarkably consistent global integrated heat content changes over the historical period that align well with the late-twentieth-century observational estimates. When considering this global agreement in the context of demonstrated biases in some high-latitude ventilating water masses, such as SPMW and AAIW in CM4, important pattern differences of ocean heat content change are evident between the two models. A vertically integrated, basin-scale heat budget analysis reveals that the CM4 and ESM4 models accumulate approximately the same amount of heat over the historical period. Still, there is approximately a 30% increase relative to ESM4 in the surface flux of heat into the Southern Ocean in the CM4 model. A comparable increase in the horizontal heat export from the Southern Ocean to the remainder of the ocean basins balances this enhanced uptake.

The offsetting mechanisms of heat transport mean that both CM4 and ESM4 have a similar amount of heat uptake in the Southern Ocean (Fig. 15), despite their markedly distinct stratification established by the unique patterns of drift during the model's preindustrial spin-up (Fig. 8). This result slightly contrasts the findings of Newsom et al. (2023). They note that across the CMIP5 and CMIP6 ensembles, intermodel differences in preindustrial stratification, particularly in the thermocline and mode waters of the Southern Ocean, are an overall good predictor of ocean heat uptake efficiency. The relative insensitivity of ocean heat uptake in CM4 and ESM4—once corrected for drift—to the Southern Ocean stratification highlights an important subtlety in the framework presented by Newsom et al. (2023)—the relationship between stratification and heat uptake efficiency becomes less clear when examining the details of individual, drift-corrected models

Overall, the results of the present study highlight the importance of using coupled preindustrial control runs to understand the nature of biases present in the pathways of oceanic heat, particularly in the case of high-latitude ventilating water masses that play an essential role in setting the thermal structure of the interior ocean. Since the temperature drift also depends on the length of the model spin-up, it argues for the need for a sufficient length of spin-up that is on the order of several centuries for the ocean to achieve some form of quasi-equilibrium to the climate forcing at the surface. This simulation length becomes an increasingly challenging task as the resolution and complexity of models increase. The computational demands of running such long simulations motivate reducing the time required for model spin-up.

The 0.25° grid is problematic for ocean modeling since it aims to bridge the gap between fully mesoscale eddy-resolving models and those requiring mesoscale eddy parameterizations (Hewitt et al. 2020; Griffies et al. 2015). For most of the global ocean, the 0.25° grid spacing reasonably captures the effects of eddy restratification and other benefits, including improved boundary currents and topographic representation, which continue to be advantageous at this resolution. At high latitudes, however, ocean models at this intermediate resolution could benefit from some form of eddy parameterization to account for the unresolved processes. Improvements in scale-aware mesoscale eddy parameterizations should continue to be a priority for the community even as higher horizontal resolutions are pursued (Hewitt et al. 2020).

While this manuscript focuses on the steric sea level drift and bias at the higher latitudes, additional areas in the mid and lower latitudes should be explored in more detail. These areas also include the Mediterranean Sea and the Sea of Japan/East Sea, which are complex regions for global models to simulate. There are also local steric effects simulated by the models that arise from the changes in density that have not been explored in much detail in this manuscript aside from the basin-scale heat budget presented here. Additional work is needed to understand how these biases that emerge through the long-term integration of model drift impact local and regional circulation patterns and manifest in dynamic sea level changes. Additionally, the length of integration for model spinup influences the mean baseline state of the ocean from which historical and future scenario simulations are performed. Nonlinearities in the seawater equation of state have been shown to substantially impact future sea level projections (Hallberg et al. 2013) and should be revisited in newer classes of ocean models.

The demand for products and services to inform coastal planning, mitigation, and adaptation efforts is increasing, and future projections of coastal sea level rise are germane to these efforts. In addition to statistical and dynamical downscaling techniques already employed to translate global model output to specific locations, output from eddying global ocean models will increasingly be used directly for decision-making purposes. While ocean models generally capture the global mean ocean heat content changes and steric sea level changes, biases remain at the regional and local levels. Leveraging the information of drift throughout a long control simulation provides a valuable source of information for interpreting the historical and future sea level responses from climate models.

Acknowledgments. The authors thank Dr. Jacob Steinberg and Dr. Lori Sentman for their helpful comments on earlier versions of the manuscript. They also thank the anonymous reviewers whose input greatly improved the presentation. RLB was additionally supported by the NOAA Climate and Global Change Postdoctoral Fellowship Award. JET and GAM were xsupported by U.S. DOE Subcontract B640108 under Prime Contract DE-AC52-07NA27344. GAM was additionally supported by UKRI Grant MR/W013835/1. CML was supported by NOAA Climate Program Office Grant NA22OAR4310112. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.

Data availability statement. The model output used in this study is available for download from NOAA-GFDL's contributions to CMIP6, hosted by the Earth System Grid Federation (https://esgf-node.llnl.gov/projects/cmip6/). The source code for MOM6 and both GFDL coupled models is available at https://github.com/NOAA-GFDL. The "momelevel" analysis code contains the routines used to generate the figures and analyses for the project. It is available for download at https://github.com/jkrasting/momlevel, PyPi, and the conda-forge distribution channels. Additional datasets for generating the figures in this manuscript are available from Zenodo (https://doi.org/10.5281/zenodo.14014193).

REFERENCES

Adcroft, A., and Coauthors, 2019: The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *J. Adv. Model. Earth Syst.*, **11**, 3167–3211, https://doi.org/10.1029/2019MS001726.

Beadling, R. L., and Coauthors, 2022: Importance of the Antarctic Slope Current in the Southern Ocean response to ice sheet melt and wind stress change. *J. Geophys. Res. Oceans*, **127**, e2021JC017608, https://doi.org/10.1029/2021JC017608.

Chelton, D. B., R. A. deSzoeke, M. G. Schlax, K. E. Naggar, and N. Siwertz, 1998: Geographical variability of the first baroclinic Rossby radius of deformation. *J. Phys. Oceanogr.*, **28**, 433–460, https://doi.org/10.1175/1520-0485(1998)028<0433:GVOTFB>2.0.CO;2.

Church, J. A., N. J. White, and J. M. Arblaster, 2005: Significant decadal-scale impact of volcanic eruptions on sea level and ocean heat content. *Nature*, **438**, 74–77, https://doi.org/10.1038/nature04237.

Dai, A., and K. E. Trenberth, 2002: Estimates of freshwater discharge from continents: Latitudinal and seasonal variations. *J. Hydrometeor.*, **3**, 660–687, https://doi.org/10.1175/1525-7541(2002)003<0660:EOFDFC>2.0.CO;2.

——, T. Qian, K. E. Trenberth, and J. D. Milliman, 2009: Changes in continental freshwater discharge from 1948 to 2004. *J. Climate*, **22**, 2773–2792, https://doi.org/10.1175/2008JCLI2592.1.

Delworth, T. L., F. Zeng, A. Rosati, G. A. Vecchi, and A. T. Wittenberg, 2015: A link between the hiatus in global warming and North American drought. *J. Climate*, **28**, 3834–3845, https://doi.org/10.1175/JCLI-D-14-00616.1.

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.

Dunne, J. P., and Coauthors, 2020a: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics. *J. Adv. Model. Earth Syst.*, **12**, e2019MS002015, https://doi.org/10.1029/2019MS002015.

——, and Coauthors, 2020b: Simple global ocean Biogeochemistry with Light, Iron, Nutrients and Gas version 2 (BLINGv2): Model description and simulation characteristics in GFDL's CM4.0. *J. Adv. Model. Earth Syst.*, **12**, e2019MS002008, https://doi.org/10.1029/2019MS002008.

England, M. H., and Coauthors, 2014: Recent intensification of wind-driven circulation in the Pacific and the ongoing warming hiatus. *Nat. Climate Change*, **4**, 222–227, https://doi.org/10.1038/nclimate2106.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

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.

Fox-Kemper, B., and Coauthors, 2011: Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations. *Ocean Modell.*, **39**, 61–78, https://doi.org/10.1016/j.ocemod.2010.09.002.

——, and Coauthors, 2019: Challenges and prospects in ocean circulation models. *Front. Mar. Sci.*, **6**, 65, https://doi.org/10.3389/fmars.2019.00065.

——, and Coauthors, 2021: Ocean, cryosphere and sea level change. *Climate Change 2021: The Physical Science Basis.*, V. Masson-Delmotte et al., Eds., Cambridge University Press, 1211–1362, https://doi.org/10.1017/9781009157896.011.

Greatbatch, R. J., 1994: A note on the representation of steric sea level in models that conserve volume rather than mass. *J. Geophys. Res.*, **99**, 12 767–12 771, https://doi.org/10.1029/94JC00847.

Gregory, J. M., and Coauthors, 2016: The Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP) contribution to CMIP6: Investigation of sea-level and ocean climate change in response to CO_2.forcing. *Geosci. Model Dev.*, **9**, 3993–4017, https://doi.org/10.5194/gmd-9-3993-2016.

——, and Coauthors, 2019: Concepts and terminology for sea level: Mean, variability and change, both local and global. *Surv. Geophys.*, **40**, 1251–1289, https://doi.org/10.1007/s10712-019-09525-z.

Griffies, S. M., and R. J. Greatbatch, 2012: Physical processes that impact the evolution of global mean sea level in ocean climate models. *Ocean Modell.*, **51**, 37–72, https://doi.org/10.1016/j.ocemod.2012.04.003.

——, and Coauthors, 2009: Coordinated Ocean-ice Reference Experiments (COREs). *Ocean Modell.*, **26** (1–2), 1–46, https://doi.org/10.1016/j.ocemod.2008.08.007.

——, and Coauthors, 2014: An assessment of global and regional sea level for years 1993–2007 in a suite of interannual CORE-II simulations. *Ocean Modell.*, **78**, 35–89, https://doi.org/10.1016/j.ocemod.2014.03.004.

——, and Coauthors, 2015: Impacts on ocean heat from transient mesoscale eddies in a hierarchy of climate models. *J. Climate*, **28**, 952–977, https://doi.org/10.1175/JCLI-D-14-00353.1.

——, and Coauthors, 2016: OMIP contribution to CMIP6: Experimental and diagnostic protocol for the physical component of the Ocean Model Intercomparison Project. *Geosci. Model Dev.*, **9**, 3231–3296, https://doi.org/10.5194/gmd-9-3231-2016.

——, A. Adcroft, and R. W. Hallberg, 2020: A primer on the vertical Lagrangian-remap method in ocean models based on finite volume generalized vertical coordinates. *J. Adv. Model.*

Earth Syst., **12**, e2019MS001954, https://doi.org/10.1029/2019MS001954.

Hallberg, R., 2013: Using a resolution function to regulate parameterizations of oceanic mesoscale eddy effects. *Ocean Modell.*, **72**, 92–103, https://doi.org/10.1016/j.ocemod.2013.08.007.

——, A. Adcroft, J. P. Dunne, J. P. Krasting, and R. J. Stouffer, 2013: Sensitivity of twenty-first-century global-mean steric sea level rise to ocean model formulation. *J. Climate*, **26**, 2947–2956, https://doi.org/10.1175/JCLI-D-12-00506.1.

Hamlington, B. D., and Coauthors, 2020: Understanding of contemporary regional sea-level change and the implications for the future. *Rev. Geophys.*, **58**, e2019RG000672, https://doi.org/10.1029/2019RG000672.

——, T. Frederikse, P. R. Thompson, J. K. Willis, R. S. Nerem, and J. T. Fasullo, 2021: Past, present, and future Pacific sea-level change. *Earth's Future*, **9**, e2020EF001839, https://doi.org/10.1029/2020EF001839.

Held, I. M., and Coauthors, 2019: Structure and performance of GFDL's CM4.0 climate model. *J. Adv. Model. Earth Syst.*, **11**, 3691–3727, https://doi.org/10.1029/2019MS001829.

Hewitt, H. T., and Coauthors, 2020: Resolving and parameterising the ocean mesoscale in Earth System models. *Curr. Climate Change Rep.*, **6**, 137–152, https://doi.org/10.1007/s40641-020-00164-w.

Horowitz, L. W., and Coauthors, 2020: The GFDL global atmospheric chemistry-climate model AM4.1: Model description and simulation characteristics. *J. Adv. Model. Earth Syst.*, **12**, e2019MS002032, https://doi.org/10.1029/2019MS002032.

Irving, D., W. Hobbs, J. Church, and J. Zika, 2021: A mass and energy conservation analysis of drift in the CMIP6 ensemble. *J. Climate*, **34**, 3157–3170, https://doi.org/10.1175/JCLI-D-20-0281.1.

Jansen, M. F., A. J. Adcroft, R. Hallberg, and I. M. Held, 2015: Parameterization of eddy fluxes based on a mesoscale energy budget. *Ocean Modell.*, **92**, 28–41, https://doi.org/10.1016/j.ocemod.2015.05.007.

Jevrejeva, S., H. Palanisamy, and L. P. Jackson, 2020: Global mean thermosteric sea level projections by 2100 in CMIP6 climate models. *Environ. Res. Lett.*, **16**, 014028, https://doi.org/10.1088/1748-9326/abceea.

Jin, S., H. Pan, and T. Xu, 2023: Assessment of the sea surface salinity simulation and projection surrounding the Asian waters in the CMIP6 models. *Atmosphere*, **14**, 726, https://doi.org/10.3390/atmos14040726.

Kopp, R. E., R. M. Horton, C. M. Little, J. X. Mitrovica, M. Oppenheimer, D. J. Rasmussen, B. H. Strauss, and C. Tebaldi, 2014: Probabilistic 21st and 22nd century sea-level projections at a global network of tide-gauge sites. *Earth's Future*, **2**, 383–406, https://doi.org/10.1002/2014EF000239.

——, and Coauthors, 2023: The Framework for Assessing Changes To Sea-level (FACTS) v1.0: A platform for characterizing parametric and structural uncertainty in future global, relative, and extreme sea-level change. *Geosci. Model Dev.*, **16**, 7461–7489, https://doi.org/10.5194/gmd-16-7461-2023.

Krasting, J. P., J. P. Dunne, R. J. Stouffer, and R. W. Hallberg, 2016: Enhanced Atlantic sea-level rise relative to the Pacific under high carbon emission rates. *Nat. Geosci.*, **9**, 210–214, https://doi.org/10.1038/ngeo2641.

——, R. J. Stouffer, S. M. Griffies, R. W. Hallberg, S. L. Malyshev, B. L. Samuels, and L. T. Sentman, 2018: Role of ocean model formulation in climate response uncertainty. *J. Climate*, **31**, 9313–9333, https://doi.org/10.1175/JCLI-D-18-0035.1.

Landerer, F. W., J. H. Jungclaus, and J. Marotzke, 2007: Regional dynamic and steric sea level change in response to the IPCC-A1B scenario. *J. Phys. Oceanogr.*, **37**, 296–312, https://doi.org/10.1175/JPO3013.1.

Large, W. G., and S. G. Yeager, 2009: The global climatology of an interannually varying air–sea flux data set. *Climate Dyn.*, **33**, 341–364, https://doi.org/10.1007/s00382-008-0441-3.

Legg, S., R. W. Hallberg, and J. B. Girton, 2006: Comparison of entrainment in overflows simulated by *z*-coordinate, isopycnal and non-hydrostatic models. *Ocean Modell.*, **11**, 69–97, https://doi.org/10.1016/j.ocemod.2004.11.006.

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.

——, and Coauthors, 2014: World Ocean Atlas 2013 (NCEI Accession 0114815). NOAA National Centers for Environmental Information, https://doi.org/10.7289/V5F769GT.

Martin, T., and A. Adcroft, 2010: Parameterizing the fresh-water flux from land ice to ocean with interactive icebergs in a coupled climate model. *Ocean Modell.*, **34**, 111–124, https://doi.org/10.1016/j.ocemod.2010.05.001.

McCarthy, M. C., and L. D. Talley, 1999: Three-dimensional isoneutral potential vorticity structure in the Indian Ocean. *J. Geophys. Res.*, **104**, 13 251–13 267, https://doi.org/10.1029/1999JC900028.

Meehl, G. A., J. M. Arblaster, J. T. Fasullo, A. Hu, and K. E. Trenberth, 2011: Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods. *Nat. Climate Change*, **1**, 360–364, https://doi.org/10.1038/nclimate1229.

Melet, A., R. Hallberg, S. Legg, and K. Polzin, 2013: Sensitivity of the ocean state to the vertical distribution of internal-tide-driven mixing. *J. Phys. Oceanogr.*, **43**, 602–615, https://doi.org/10.1175/JPO-D-12-055.1.

Newsom, E., L. Zanna, and J. Gregory, 2023: Background pycnocline depth constrains future ocean heat uptake efficiency. *Geophys. Res. Lett.*, **50**, e2023GL105673, https://doi.org/10.1029/2023GL105673.

Orr, J. C., and Coauthors, 2017: Biogeochemical protocols and diagnostics for the CMIP6 Ocean Model Intercomparison Project (OMIP). *Geosci. Model Dev.*, **10**, 2169–2199, https://doi.org/10.5194/gmd-10-2169-2017.

Pudig, M. P., C. K. Yung, J. D. Zika, and R. M. Holmes, 2023: Rectified ocean heat uptake from oscillatory surface forcing. *J. Climate*, **36**, 2663–2680, https://doi.org/10.1175/JCLI-D-22-0267.1.

Reichl, B. G., and R. Hallberg, 2018: A simplified energetics based Planetary Boundary Layer (ePBL) approach for ocean climate simulations. *Ocean Modell.*, **132**, 112–129, https://doi.org/10.1016/j.ocemod.2018.10.004.

Roemmich, D., and J. Gilson, 2009: The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program. *Prog. Oceanogr.*, **82**, 81–100, https://doi.org/10.1016/j.pocean.2009.03.004.

Sarmiento, J. L., and N. Gruber, 2006: *Ocean Biogeochemical Dynamics*. Princeton University Press, 503 pp., https://doi.org/10.1515/9781400849079.

Slangen, A. B. A., J. A. Church, C. Agosta, X. Fettweis, B. Marzeion, and K. Richter, 2016: Anthropogenic forcing dominates global mean sea-level rise since 1970. *Nat. Climate Change*, **6**, 701–705, https://doi.org/10.1038/nclimate2991.

Soto-Navarro, J., and Coauthors, 2020: Evolution of Mediterranean Sea water properties under climate change scenarios in the Med-CORDEX ensemble. *Climate Dyn.*, **54**, 2135–2165, https://doi.org/10.1007/s00382-019-05105-4.

Stern, A. A., A. Adcroft, and O. Sergienko, 2016: The effects of Antarctic iceberg calving-size distribution in a global climate model. *J. Geophys. Res. Oceans*, **121**, 5773–5788, https://doi.org/10.1002/2016JC011835.

Stewart, K. D., and T. W. N. Haine, 2016: Thermobaricity in the transition zones between alpha and beta oceans. *J. Phys. Oceanogr.*, **46**, 1805–1821, https://doi.org/10.1175/JPO-D-16-0017.1.

Stock, C. A., and Coauthors, 2020: Ocean biogeochemistry in GFDL's Earth System Model 4.1 and its response to increasing atmospheric CO_2. *J. Adv. Model. Earth Syst.*, **12**, e2019MS002043, https://doi.org/10.1029/2019MS002043.

Stouffer, R. J., and S. Manabe, 1999: Response of a coupled ocean–atmosphere model to increasing atmospheric carbon dioxide: Sensitivity to the rate of increase. *J. Climate*, **12**, 2224–2237, https://doi.org/10.1175/1520-0442(1999)012<2224:ROACOA>2.0.CO;2.

——, A. J. Weaver, and M. Eby, 2004: A method for obtaining pre-twentieth century initial conditions for use in climate change studies. *Climate Dyn.*, **23**, 327–339, https://doi.org/10.1007/s00382-004-0446-5.

Swart, N. C., and Coauthors, 2023: The Southern Ocean Freshwater Input from Antarctica (SOFIA) Initiative: Scientific objectives and experimental design. *Geosci. Model Dev.*, **16**, 7289–7309, https://doi.org/10.5194/gmd-16-7289-2023.

Sweet, W., and Coauthors, 2022: Global and regional sea level rise scenarios for the United States: Updated mean projections and extreme water level probabilities along U.S. coastlines. NOAA Tech. Rep. NOS 01, 111 pp., https://oceanservice.noaa.gov/hazards/sealevelrise/noaa-nos-techrpt01-global-regional-SLR-scenarios-US.pdf.

Tesdal, J.-E., G. A. MacGilchrist, R. L. Beadling, S. M. Griffies, J. P. Krasting, and P. J. Durack, 2023: Revisiting interior water mass responses to surface forcing changes and the subsequent effects on overturning in the Southern Ocean. *J. Geophys. Res. Oceans*, **128**, e2022JC019105, https://doi.org/10.1029/2022JC019105.

Todd, A., and Coauthors, 2020: Ocean-only FAFMIP: Understanding regional patterns of ocean heat content and dynamic sea level change. *J. Adv. Model. Earth Syst.*, **12**, e2019MS002027, https://doi.org/10.1029/2019MS002027.

Tsujino, H., and Coauthors, 2020: Evaluation of global ocean–sea-ice model simulations based on the experimental protocols of the Ocean Model Intercomparison Project phase 2 (OMIP-2). *Geosci. Model Dev.*, **13**, 3643–3708, https://doi.org/10.5194/gmd-13-3643-2020.

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.

Wright, D. G., 1997: An equation of state for use in ocean models: Eckart's formula revisited. *J. Atmos. Oceanic Technol.*, **14**, 735–740, https://doi.org/10.1175/1520-0426(1997)014<0735:AEOSFU>2.0.CO;2.

Yin, J., M. E. Schlesinger, and R. J. Stouffer, 2009: Model projections of rapid sea-level rise on the northeast coast of the United States. *Nat. Geosci.*, **2**, 262–266, https://doi.org/10.1038/ngeo462.

——, S. M. Griffies, and R. J. Stouffer, 2010: Spatial variability of sea level rise in twenty-first century projections. *J. Climate*, **23**, 4585–4607, https://doi.org/10.1175/2010JCLI3533.1.

——, ——, M. Winton, M. Zhao, and L. Zanna, 2020: Response of storm-related extreme sea level along the U.S. Atlantic coast to combined weather and climate forcing. *J. Climate*, **33**, 3745–3769, https://doi.org/10.1175/JCLI-D-19-0551.1.

Zhang, Y., M. Yamamoto-Kawai, and W. J. Williams, 2020: Two decades of ocean acidification in the surface waters of the Beaufort Gyre, Arctic Ocean: Effects of sea ice melt and retreat from 1997–2016. *Geophys. Res. Lett.*, **47**, e60119, https://doi.org/10.1029/2019GL086421.

Zhao, M., and Coauthors, 2018: The GFDL Global Atmosphere and Land Model AM4.0/LM4.0: 1. Simulation characteristics with prescribed SSTs. *J. Adv. Model. Earth Syst.*, **10**, 691–734, https://doi.org/10.1002/2017MS001208.

Zika, J. D., J. M. Gregory, E. L. McDonagh, A. Marzocchi, and L. Clément, 2021: Recent water mass changes reveal mechanisms of ocean warming. *J. Climate*, **34**, 3461–3479, https://doi.org/10.1175/JCLI-D-20-0355.1.