Uncertainty in Preindustrial Global Ocean Initialization Can Yield Irreducible Uncertainty in Southern Ocean Surface Climate

HANSI K. A. SINGH,[a] NAOMI GOLDENSON,[b] JOHN C. FYFE,[c] AND LORENZO M. POLVANI[d]

[a] School of Earth and Ocean Sciences, University of Victoria, Victoria, British Columbia, Canada
[b] Department of Atmospheric and Oceanic Sciences, University of California at Los Angeles, Los Angeles, California
[c] Canadian Center for Climate Modeling and Analysis, Victoria, British Columbia, Canada
[d] Department of Applied Physics and Applied Mathematics, and Department of Earth and Environmental Sciences, Columbia University, New York City, New York

ABSTRACT: How do ocean initial states impact historical and future climate projections in Earth system models? To answer this question, we use the 50-member Canadian Earth System Model (CanESM2) large ensemble, in which individual ensemble members are initialized using a combination of different oceanic initial states and atmospheric microperturbations. We show that global ocean heat content anomalies associated with the different ocean initial states, particularly differences in deep ocean heat content due to ocean drift, persist from initialization at year 1950 through the end of the simulations at year 2100. We also find that these anomalies most readily impact surface climate over the Southern Ocean. Differences in ocean initial states affect Southern Ocean surface climate because persistent deep ocean temperature anomalies upwell along sloping isopycnal surfaces that delineate neighboring branches of the upper and lower cells of the global meridional overturning circulation. As a result, up to a quarter of the ensemble variance in Southern Ocean turbulent heat fluxes, heat uptake, and surface temperature trends can be traced to variance in the ocean initial state, notably deep ocean temperature differences of order 0.1 K due to model drift. Such a discernible impact of varying ocean initial conditions on ensemble variance over the Southern Ocean is evident throughout the full 150 simulation years of the ensemble, even though upper ocean temperature anomalies due to varying ocean initial conditions rapidly dissipate over the first two decades of model integration over much of the rest of the globe.

KEYWORDS: Atmosphere-ocean interaction; Climate variability; Climate models; Ensembles; General circulation models; Model initialization

1. Introduction

Earth's climate system is variable over a range of time scales, from seconds to decades to millennia (Peixoto and Oort 1992). This abundant internal variability presents challenges for understanding the climate system's response to anthropogenic greenhouse gas emissions and other forcing agents: What part of the observed (or modeled) change in climate is due to the forcing, greenhouse gas or otherwise, and what part is due to the internal variability of the Earth system?

"Large ensembles" are an important tool for separating the forced response from internal variability. These ensembles are a sizeable collection of experiments using a single Earth system model (ESM) subjected to identical forcings but with different initial conditions. Because two ESM integrations forced identically will diverge even if they start from a nearly identical initial state, such a large ensemble may be used to create an array of possible climate trajectories. Differences between ensemble members are then attributable solely to internal variability in the model, while the mean evolution of all ensemble members is attributable to the forcing. In this framework, the actual trajectory of Earth's climate is just one of many possible trajectories that might arise from the applied forcing in a perfect model.

Large ensembles show that internal variability lends substantial uncertainty to future climate projections (Deser et al. 2012, 2014). In the 40-member Community Earth System Model Large Ensemble (CESM-LE; see Kay et al. 2015), for example, individual ensemble members exhibit significantly different global mean surface temperature trends even a century after initialization, and regional surface temperature trends show even greater variance between members. In the Arctic, where internal variability is particularly large, analysis of large ensembles suggests that much of the observed total sea ice area decline, warming, and change in precipitation is attributable to greenhouse gas forcing (Screen et al. 2014). However, variability in the atmospheric circulation may still account for up to half the observed downward trend in summer sea ice (Ding et al. 2017), since circulation changes that accompany Arctic warming are difficult to distinguish from internal variability (Screen et al. 2014; Wettstein and Deser 2014). Moreover, local trends in sea ice area are only attributable to greenhouse gas forcing in certain regions and over

---

certain seasons (England et al. 2019). Indeed, the precise timing of a sea ice–free Arctic in summer depends largely on the sequence of internal variability in a given ensemble member (Swart et al. 2015), and may depend very little on the emissions scenario (Jahn et al. 2016). Other studies show that internal variability is significant for such varied climate change indicators as Hadley cell expansion (Kang et al. 2013), atmospheric river landfall frequency (Hagos et al. 2016), and Southern Ocean carbon uptake (Lovenduski et al. 2016).

Because large ensembles have become an indispensable tool for understanding how the climate system evolves in the presence of internal variability, it is reasonable to consider just how these ensembles are constructed. Thus far, there are two commonly used methods for creating initial conditions to spawn large ensembles [as described by Stainforth et al. (2007)]: microinitialization, using tiny perturbations (i.e., of a magnitude similar to machine round-off error) in the atmospheric initial state, or macroinitialization, using different ocean starting states. Because large ensembles generally use either atmospheric microperturbations (see, e.g., the CESM-LE; Kay et al. 2015) or varying ocean initial conditions (see, e.g., the MPI Grand Ensemble, which uses ocean initial states sampled from a long control experiment; Maher et al. 2019) for their ensemble initialization, it is unclear whether the two methods yield a similar range of ensemble internal variability and, therefore, a similar spread in climate projections. Because each ESM has its own representation of internal climate variability, macroinitialization and microinitialization would need to be applied in the same ESM in order to assess whether constructing an ensemble through macroinitialization increases ensemble variance relative to one constructed solely through microinitialization.

The importance of the ocean state for driving Earth system evolution is already well recognized in other applications. In the field of decadal climate predictability, accurate ocean state initialization is of prime importance in determining the climate's trajectory (see, e.g., Latif and Keenlyside 2011; Bellucci et al. 2013; Meehl et al. 2014; Yeager and Robson 2017, and many others). Initialization of coupled climate models with a given phase of the Atlantic multidecadal oscillation (AMO), Pacific decadal oscillation (PDO), or both, partly determines the evolution of ocean temperature, salinity, and sea surface height over one or more decades (see, e.g., Griffies and Bryan 1997; Rodwell et al. 1999; Mochizuki et al. 2012; Chikamoto et al. 2013), and may enhance predictability of the extratropical circulation, the hydrologic cycle, and tropical Atlantic variability over seasonal, interannual, and decadal time scales (see, e.g., Dunstone et al. 2011; Simpson et al. 2019; Athanasiadis et al. 2020). Furthermore, climate model experiments also suggest that the ocean state may help drive multidecadal trends in Antarctic sea ice, including the expansion of Antarctic sea ice area over the satellite era (1979–2015; see Cavalieri et al. 1996, updated yearly) some have suggested that tropical–extratropical teleconnections mediated by the interdecadal Pacific oscillation may have facilitated Antarctic sea ice expansion over that period (Meehl et al. 2016), while others have pointed to the state of the Southern

Ocean as the implicating factor (Zhang et al. 2019; Singh et al. 2019).

Given this wealth of evidence that the ocean state impacts climate evolution, it is reasonable to hypothesize that large ensembles initialized from many different ocean states may exhibit variability not found in those initialized from a single ocean state. Indeed, one prior study exploring the matter suggests that initializing a large ensemble with a range of ocean initial conditions increases ensemble variance beyond that possible with only atmospheric microperturbations. Hawkins et al. (2016) used an Earth system model of intermediate complexity (EMIC) to show that a historically forced large ensemble starting from several distinct ocean initial states displayed significantly greater variance in global and regional temperature trends, even a century after initialization, compared to one starting from only a single ocean initial state. More specifically, the phase of the Atlantic meridional overturning circulation from which an ensemble member was initialized influenced Northern Hemispheric temperature trends, particularly in those regions proximal to the North Atlantic. Because these experiments utilized an EMIC rather than an ESM, however, there remains a question of whether such increased variability is a product of the greater sensitivity of simpler models to parameter and initial condition perturbations (as is the case for sea ice instability; see Wagner and Eisenman 2015), or whether such increased variability is also found in large ensembles of more comprehensive Earth system models.

The analysis performed by Hawkins et al. (2016) also highlights another key question: while the phasing of decadal-scale internal variability may give rise to differences in ocean initial conditions whose effects persist with time, might ocean initial conditions also differ in other substantive ways? Indeed, not all ESMs exhibit long time scale modes of internal variability from which to sample ocean initial states [see, e.g., the intermodel comparison of Southern Ocean overturning variability in Behrens et al. (2016)]. Specifically, the deep ocean (below 2000 m) is one area where there is significant uncertainty in present-day ocean state estimates (see Abraham et al. 2013; Cheng et al. 2017; Gasparin et al. 2020), which are only compounded when estimating the deep ocean state in the era prior to ocean observation networks, pre-2001 (see, e.g., Crowley et al. 2003). Moreover, anomalies in the deep ocean state may persist for centuries or millennia, given that long time scales of adjustment characterize the deep ocean. Gebbie and Huybers (2019), for example, show that cooling from the Little Ice Age (circa 1700 CE) is evident in the deep Pacific Ocean today, which significantly impacts the rate at which global ocean heat content is rising at present. Additionally, deep ocean temperatures in many ESMs frequently trend either up or down over long time scales, due to insufficient spinup or energy leaks in model components (Sen Gupta et al. 2012, 2013; Hobbs et al. 2016). These considerations suggest that uncertainty in deep ocean temperatures may be an important factor that impacts surface climate in a large ensemble, particularly over regions where the deep ocean more readily interacts with the surface (such as the North Atlantic

or the Southern Oceans; see Ganachaud and Wunsch 2000; Marshall and Speer 2012).

In this study, we examine the above by answering the following question: is ESM ensemble variance amplified by initializing members from different ocean states, particularly those with different deep ocean temperatures [rather than different phasing of modes of internal variability, as in Hawkins et al. (2016)], compared to initializing members with atmospheric microperturbations alone? More specifically, does the persistence of deep ocean temperature variance in a large ensemble increase uncertainty in projected surface climate, either globally or regionally? To answer these questions, we analyze the Canadian Earth System Model version 2 (CanESM; Arora et al. 2011) large ensemble, run with historical and RCP8.5 future scenario forcings (Taylor et al. 2012; Deser et al. 2020) from 1950 to 2100. This large ensemble is composed of five microensembles (consisting of 10 ensemble members each), where individual members of a given microensemble are initialized from an identical ocean state, but each microensemble is initialized from a distinct ocean state. Most importantly, as we will show, these ocean initial states differ by $\mathcal{O}(0.1)$ K in their deep ocean temperatures, a range made possible by deep ocean drift in the preindustrial control experiment from which microensembles were branched over 50-yr intervals. Both the unique structure of this 50-member large ensemble and the range of deep ocean temperatures generated by the ocean drift permit us to decompose the variance in the ensemble into a component due to the ocean initial state, notably variance in deep ocean temperatures and ocean heat content, and a component due to atmospheric microperturbations alone.

We begin our analysis of the CanESM2 large ensemble by evaluating how ocean initial conditions, including potential temperature and ocean heat content, differ between microensembles (section 3a). We then show how the ocean state evolves from 1850 to 2100 in each microensemble, and compute the extent to which ocean potential temperature variance in the full ensemble can be attributed to different ocean initial conditions (section 3b). Finally, we demonstrate that it is over the Southern Ocean where such initial conditions continue to impact ensemble variance in surface climate, up to 150 years following model initialization in 1950 (section 3c). In section 4, we conclude by discussing the implications of our findings for the design of large ensembles, and how climate system predictability may be limited by our imperfect knowledge of prior ocean states.

2. Methods

The Canadian Earth System Model, version 2 (CanESM2) is state-of-the-art, fully coupled, and has atmosphere, ocean, sea ice, and land components [described in detail in Arora et al. (2011)]. The atmosphere model, CanAM4 (von Salzen et al. 2013), utilizes a spectral dynamical core at T63 truncation, with a resolution of 1.875 at the equator; there are 35 vertical levels, which extend to 1 hPa. New parameterizations include a correlated-k radiative transfer scheme (Li and Barker 2005), a prognostic bulk aerosol treatment (Ma et al.

2010), and single-moment cloud microphysics (Khairoutdinov and Kogan 2000). The ocean model has 40 vertical levels with a nominal horizontal resolution of 1. It utilizes the K-profile parameterization for vertical mixing at the boundary layer (Large et al. 1994) and the GM90 parameterization for mixing by subgrid scale eddies along isopycnal surfaces (Gent and McWilliams 1990). The sea ice model is fully dynamic and thermodynamic, and both the land and ocean models include a prognostic carbon cycle (Christian et al. 2010).

CanESM2 compares favorably with other models participating in phase 5 of the Climate Model Intercomparison Project (CMIP5; see Taylor et al. 2012), in terms of its representation of both mean state climate and internal variability over seasonal to centennial time scales (Flato et al. 2014). Further studies indicate reasonable simulation of coupled modes of climate variability, including ENSO (see, e.g., Bellenger et al. 2014), the PDO (see, e.g., Yim et al. 2015), and Southern Hemispheric extratropical circulation features (including the SAM, jet position, and location of the maximum westerly wind stress; see Thomas et al. 2015). CanESM2 also simulates both the mean state and variability of meridional ocean heat transport well, including its gyre and overturning components (see Yang and Saenko 2012).

As illustrated in Fig. 1, ocean initial conditions for the 50-member CanESM2 large ensemble are constructed by branching five runs from different points in an 1850s preindustrial control experiment (Kirchmeier-Young et al. 2017). The first of the five branches (which will give rise to microensemble 5) starts after 2271 model-years of the preindustrial control simulation, and subsequent branches each begin 50 years after the previous branch (years 2321, 2371, 2421, and 2471, corresponding to microensembles 4, 3, 2, and 1, respectively). The preindustrial control has a top-of-atmosphere anomaly of 0.17 W m^{-2}, and the deep ocean is drifting by approximately -0.1 K (100 years)$^{-1}$ [see Fig. 1b; as documented for CMIP5-participating models in Hobbs et al. (2016)]. Because of this deep ocean drift, there is approximately a 0.2-K range in deep ocean temperatures (below 1500 m) between these branches.

Each of these five branches is subjected to identical historical forcings from years 1850 to 1950. Differences in deep ocean temperatures between branches, which are primarily attributable to ocean drift, persist as each branch is integrated forward in time, and give rise to a set of distinct ocean initial states. At year 1950, each of the five branched runs is subjected to 10 distinct sets of random microperturbations in the atmosphere (by using 10 different preset seeds for a random number generator employed in the model's cloud microphysics parameterization) to produce 10 ensemble members each. Hereafter, we use the term *microensemble* to refer to each set of 10 ensemble members that shares an identical ocean initial state at year 1950. As per the protocol of phase 5 of the Climate Model Intercomparison Project (CMIP5; see Taylor et al. 2012), all of these ensemble members are subjected to identical historical forcings (from 1950 to 2005) and the RCP8.5 scenario forcing (from 2006 to 2100, to yield a total nominal greenhouse gas forcing of 8.5 W m^{-2} by the end of the twenty-first century, relative to the preindustrial).
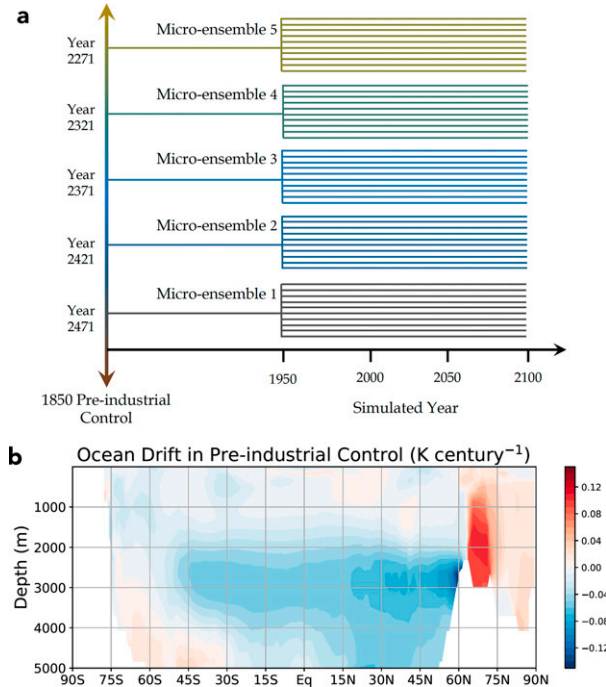
FIG. 1. (a) Initialization structure of the 50-member CanESM2 large ensemble. Five runs were branched at 50-yr intervals from the 1850 preindustrial control, and each was subjected to identical historical forcings from the period 1850 to 1950. At year 1950, each of the five runs was perturbed with 10 distinct random atmospheric microperturbations, which created 10 ensemble members per branched run. Each of these microensembles of 10 members were subjected to identical historical forcings (for the period 1950–2005), and then subject to the RCP8.5 future preindustrial control scenario forcing (Taylor et al. 2012) to year 2100. (b) Ocean potential temperature drift in the run (in K century^{-1}) computed as the linear trend in the zonal mean potential temperature over years 2271–2470.

a. Decomposition of ensemble variance

We now describe the process by which we estimate how much variance in the whole ensemble is attributable to ocean initial conditions, and how much is attributable to atmospheric microperturbations.

The variance σ_X^2 in a climatically relevant quantity X (such as temperature, surface fluxes, ocean heat content, or others) between all ensemble members over a given year is computed as

$$\sigma_X^2(t) = \frac{\sum_{i=1}^{n}[X(t) - \overline{X}(t)]^2}{n - 1}, \qquad (1)$$

where $\overline{X}(t)$ is the average of X across all ensemble members at year t, and n is the number of ensemble members (equal to 50 in the CanESM2 large ensemble). While this can be a function of time, we drop this time-dependent notation in the following description for the sake of clarity.

The total variance between ensemble members over a given year can be approximated as the sum of two variances: 1) the variance between microensembles, due to the different ocean states used to initialize each microensemble, is denoted by $\sigma_{X,\text{ocean}}^2$; and 2) the variance within microensembles, due to application of different atmospheric microperturbations in each ensemble member, is denoted by $\sigma_{X,\text{atmos}}^2$. In other words,

$$\sigma_X^2 = \sigma_{X,\text{ocean}}^2 + \sigma_{X,\text{atmos}}^2 + \varepsilon. \qquad (2)$$

In Eq. (2) above, the error ε includes the nonlinear interaction term; ε generally constitutes less than 5% of the total variance, which we drop for convenience. This approximation, inspired by the decomposition of variance performed by Hawkins and Sutton (2009), makes sources of ensemble variance simple to compute and easy to attribute, to first order.

The variance within microensembles, $\sigma_{X,\text{atmos}}^2$ is computed as the average of the variance within each microensemble:

$$\sigma_{X,\text{atmos}}^2 = \frac{1}{p}\sum_{k=1}^{p}\frac{\sum_{j=1}^{m}(X_{k,j} - \overline{X}_k)^2}{m - 1}, \qquad (3)$$

where $X_{k,j}$ is the value of X in the jth member of the kth microensemble, and $\overline{X}_k$ is the mean of X in microensemble k. In the above equation, m is the number of ensemble members in each microensemble (equal to 10 for the CanESM2 large ensemble), and p is the number of microensembles (five for the CanESM2 large ensemble). The variance between microensembles, $\sigma_{X,\text{ocean}}^2$, is computed as the variance of the individual microensemble means:

$$\sigma_{X,\text{ocean}}^2 = \frac{\sum_{k=1}^{p}(\overline{X}_k - \overline{X})^2}{p - 1}, \qquad (4)$$

where $\overline{X}$ is the mean of X in the entire ensemble (i.e., over all 50 members of the CanESM2 large ensemble).

Because individual ensemble members within each microensemble all start with identical ocean initial conditions at year 1950, the variance within microensembles, $\sigma_{X,\text{atmos}}^2$, is attributable solely to initial microperturbations (on the order of machine error) in the surface atmospheric temperature. Similarly, the variance between microensembles, $\sigma_{X,\text{ocean}}^2$, arises from the different ocean initial conditions in each microensemble; by considering the variance of the microensemble means, the impact of varying atmospheric microperturbations is averaged out. The fraction of the ensemble variance in X due to ocean initial conditions at time t can then be written as follows:

$$\chi_{\text{OcnICs}}(t) = \frac{\sigma_{X,\text{ocean}}^2(t)}{\sigma_X^2(t)}. \qquad (5)$$

We label $\chi_{\text{OcnICs}}(t)$ as statistically distinct from zero using a bootstrapped 90%-confidence approach as follows. For 100

realizations, we randomly assign each of the 50 ensemble members into 5 microensembles of 10 members each, and recompute the variance between microensembles ($\tilde{\sigma}^2_{X,\text{between}}$) and within microensembles ($\tilde{\sigma}^2_{X,\text{within}}$). These randomly resampled microensembles are synthetic, in that their members do not share the same ocean initial conditions as do members of the original microensembles. Therefore, nonzero values of $\tilde{\sigma}^2_{X,\text{between}}$ are attributable solely to chance, not to ocean initial conditions. We repeat the above randomization a total of 100 times to get 100 synthetic realizations of $\tilde{\sigma}^2_{X,\text{between}}$, to compare to the variance between the real microensembles, $\sigma^2_{X,\text{ocean}}$. We treat $\sigma^2_{X,\text{ocean}}$ as statistically different from $\tilde{\sigma}^2_{X,\text{between}}$ if $\sigma^2_{X,\text{ocean}} > \tilde{\sigma}^2_{X,\text{between}}$ at least 90% of the time, accepting a 10% possibility that the difference could be due to chance. We use a 90% confidence level, rather than the more customary 95% level, in order to avoid type II errors that are more likely to arise when comparing the variance of two quantities (see Von Storch and Zwiers 2002).

b. Deep ocean drift and de-drifting

The present study focuses on how differences in the deep ocean state, primarily generated by model drift, impact variance in surface climate over time in the CanESM2 large ensemble. However, not all the differences in ocean initial states between microensembles are due to deep ocean drift (i.e., are created from the deep ocean potential temperature trends shown in Fig. 1b); some may be due to internal variability in CanESM2. We employ de-drifting to distinguish between drift-generated differences between microensembles, and internal variability–generated differences. De-drifted results are shown in the online supplemental material.

For de-drifting each microensemble i, we assume that for some variable $X_i(t)$, the drift generates a linear trend dX/dt that can be computed from the preindustrial control. The impact of this drift on microensemble i will be an additional $(\Delta X_i)_{\text{drift}} = \Delta t_i(dX/dt)$, where Δt_i is the time span over which the ith microensemble has drifted, relative to the first microensemble. Therefore, we remove the drift by performing the following operation:

$$X_{i,\text{de-drift}}(t) = X_i(t) - \left.\frac{dX}{dt}\right|_{\text{piC}}\Delta t_i. \tag{6}$$

Note that this differs from the de-drifting procedures outlined in Sen Gupta et al. (2012) and Sen Gupta et al. (2013), where the authors de-drift to remove trends due to ocean drift from the forced response. Here, de-drifting serves to remove differences in the ocean state between microensembles that are attributable to ocean drift in the preindustrial control.

3. Results

a. Ocean initial conditions in the CanESM2 large ensemble

We begin by examining how ocean initial conditions at year 1950 vary between microensembles. Figure 2 shows the anomaly in the mean (zonally averaged) ocean potential temperature

in each microensemble, relative to the mean over all ensemble members (i.e., $[\overline{\theta}_k] - [\overline{\theta}]$, where $[\overline{\theta}_k]$ is the average zonal-mean potential temperature in microensemble k, and $[\overline{\theta}]$ is the average zonal-mean potential temperature in the full 50-member ensemble). At year 1950, there are several key areas where ocean initial temperatures differ significantly between microensembles: within the Arctic basin (poleward of 75°N), in the Northern Hemisphere subpolar oceans (between 60° and 75°N), and in the global deep ocean (below 1.5-km depth at latitudes south of 60°N). Further differences are also apparent in the upper ocean (above 500 m), particularly in the tropics and over the Southern Ocean (poleward of 45°S). While upper ocean temperature differences between microensembles arise from internal variability, deep ocean temperature differences are generated by drift in the preindustrial control experiment (see Fig. S1 in the online supplemental material; also recall section 2 and Fig. 1b).

We further note that there is little coherence between anomalies over different areas: individual microensembles are neither uniformly cooler than average globally nor uniformly warmer. For example, cool temperatures in the subpolar Northern Hemisphere may be associated with either cool anomalies in the global deep ocean (as in microensemble 1; Fig. 2a) or warm anomalies (as in microensemble 5; Fig. 2e).

In Fig. 3, we show the average initial ocean heat content anomaly per unit area (in 10^9 J m^{-2}) in each microensemble, relative to the average over the full ensemble (i.e., $\overline{\text{OHC}}_k - \overline{\text{OHC}}$). As expected, anomalies in ocean potential temperature result in significant differences in ocean heat content between microensembles. Over most latitudes, the average heat content anomaly in each microensemble is consistent with the potential temperature anomaly in the deep ocean (below 1.5 km): anomalously cool deep ocean temperatures in microensemble 1 (Fig. 2a) are accompanied by lower-than-average ocean heat content over much of the globe (Fig. 3a), while anomalously warm deep ocean temperatures in microensemble 5 (Fig. 2e) are accompanied by higher-than-average ocean heat content. Although anomalies in potential temperature in the deep ocean are small (below 2-km depth, there is less than a 0.2-K difference between microensembles 1 and 5, as shown in Fig. 2), ocean heat content anomalies are substantial (on the order of 10^9 J m^{-2}) because of the enormous volume of the deep ocean.

b. Ocean evolution in the CanESM2 large ensemble

In Fig. 4, we show the evolution of global ocean heat content from 1950 to 2100 in each microensemble, $\overline{\text{OHC}}_k$ (relative to the ensemble mean global ocean heat content from 1950 to 1970). At year 1950, the average global ocean heat content in each microensemble, relative to that in other microensembles, is consistent with the temperature and ocean heat content anomalies shown previously (recall Figs. 2 and 3). For example, microensemble 1 has, on average, the most anomalously cold deep ocean temperatures (Fig. 2a) and the lowest ocean heat content per unit area (Fig. 3a), relative to other microensembles; therefore, unsurprisingly, its average global ocean heat content is the lowest of the five
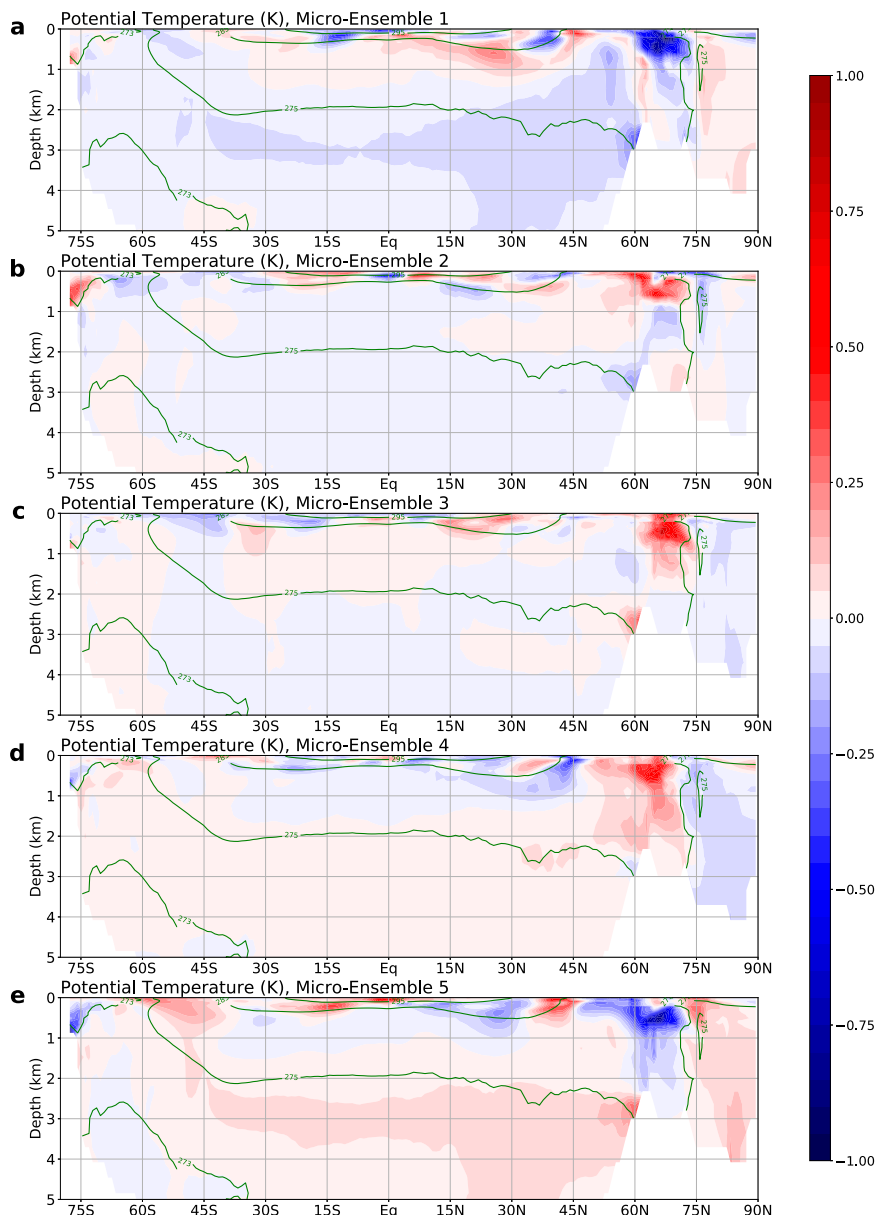
FIG. 2. Zonal-mean ocean potential anomaly (in K; shading) in (a)–(e) microensembles 1–5, respectively, at year 1950 relative to the mean potential temperature (contours at 273, 275, 285, and 295 K) in the full ensemble at year 1950.

microensembles (Fig. 4a, thick dark blue line). Similarly, microensemble 5 has, on average, the most anomalously warm deep ocean temperatures and highest ocean heat content per unit area, giving it the greatest average global ocean heat content of all microensembles (Fig. 4a, thick dark red line). The total range in global ocean heat content between microensemble means is approximately 350 ZJ at year 1950 (Fig. 4b; difference between thick dark red and dark blue lines). Much of the variance in ocean heat content between microensembles is attributable to deep ocean drift: when de-drifted, the range in global ocean heat content

between microensembles is only 60 ZJ at year 1950 (see Fig. S2).

The global ocean heat content remains relatively constant from years 1950 to 1980 in all ensemble members, but begins to increase after year 1980 (Fig. 4a). The rate at which global ocean heat content increases is not constant in time, but accelerates in all microensembles (Fig. 4a; the ocean heat content time series have positive curvature) as the forcing and rate of ocean heat uptake increase (Shi et al. 2018). As such, by year 2100, the global ocean heat content has increased by approximately 3500 ZJ due to (historical and RCP8.5)
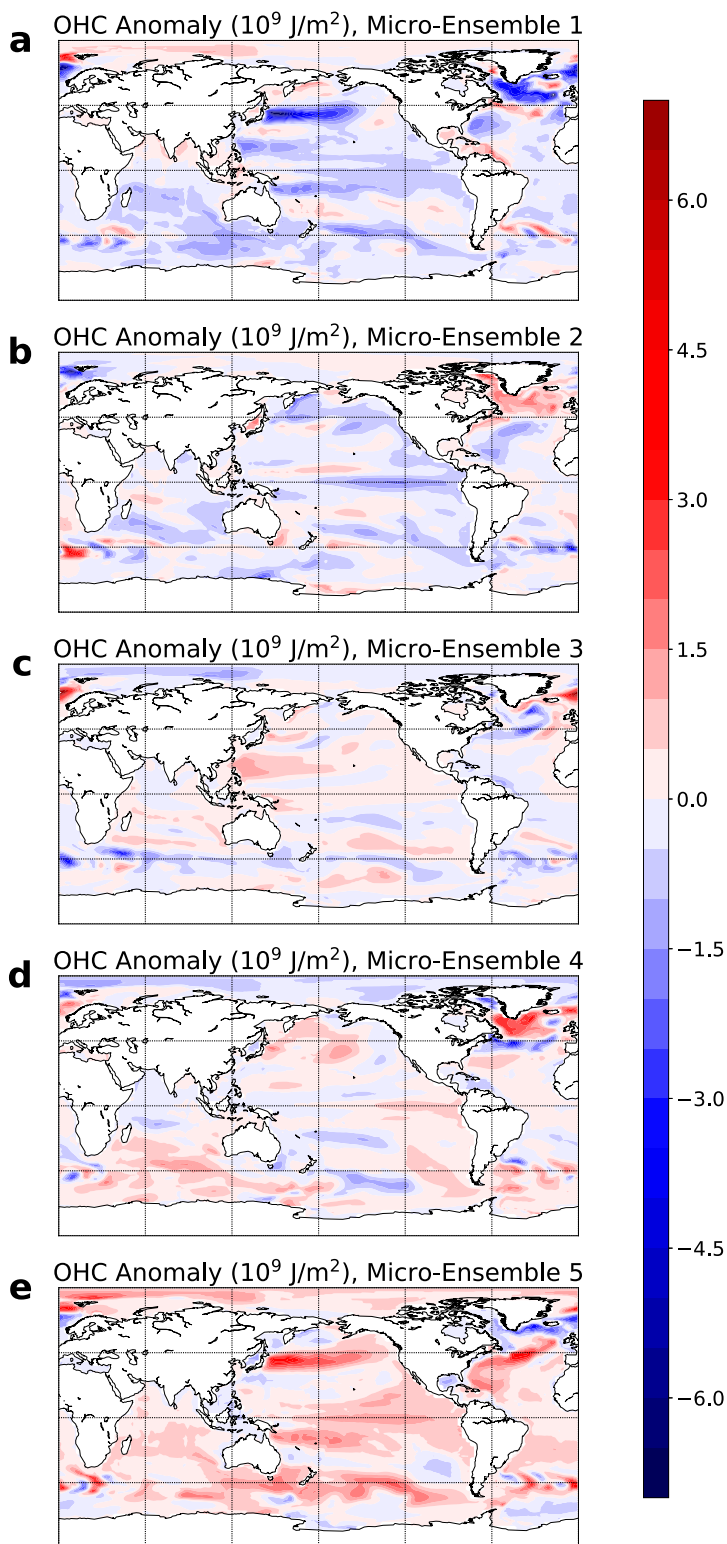
FIG. 3. Anomaly in ocean heat content per unit area (in 10^9 J m^{-2}) at year 1950 in (a)–(e) microensembles 1–5, respectively, relative to the mean ocean heat content in the full ensemble at year 1950; in other words, $\overline{\mathrm{OHC}}_k(t = 1950) - \overline{\mathrm{OHC}}(t = 1950)$.
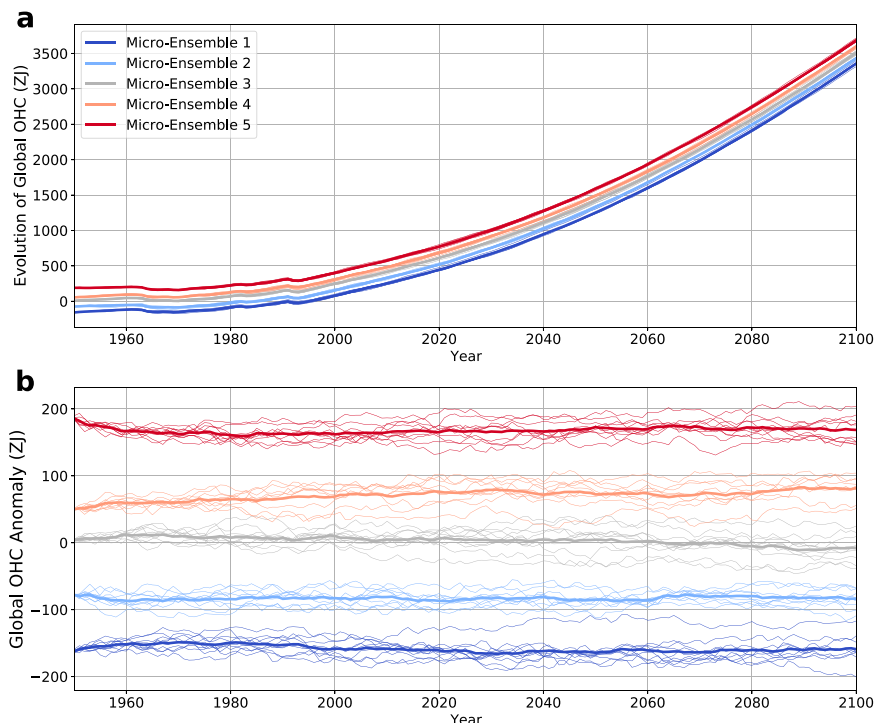
FIG. 4. Evolution of global ocean heat content in the CanESM2 large ensemble, color-coded by microensemble, with thin lines denoting individual ensemble members and thick lines denoting microensemble means [$\overline{\mathrm{OHC}_k}(t)$]. Shown are the (a) drift-corrected global ocean heat content in each ensemble member (in ZJ), relative to the ensemble-mean global ocean heat content over years 1950–70 and (b) the global ocean heat content anomaly (in ZJ) relative to the yearly ensemble-mean ocean heat content [i.e., $\overline{\mathrm{OHC}_k}(t) - \overline{\mathrm{OHC}}(t)$ for the kth microensemble, and $\mathrm{OHC}_i(t) - \overline{\mathrm{OHC}}(t)$ for the ith ensemble member]. For (a), we drift-correct by removing the linear trend in ocean heat content in the preindustrial control from each ensemble member [as described in Sen Gupta et al. (2012)].

forcings that have warmed the planet and increased global ocean temperatures.

Of particular note in Fig. 4b is that the ordering of the average global ocean heat content anomaly in each microensemble, $\overline{\mathrm{OHC}_k} - \overline{\mathrm{OHC}}$, remains constant with respect to other microensembles throughout the 150 years of the experiment: for example, the average global ocean heat content in microensemble 2 is always greater than that in microensemble 1 [i.e., $\overline{\mathrm{OHC}_1}(t) < \overline{\mathrm{OHC}_2}(t)$ for all t] and less than that in microensembles 3 through 5 [i.e., $\overline{\mathrm{OHC}_2}(t) < \overline{\mathrm{OHC}_{3,4,5}}(t)$ for all t]. This is also evident in individual ensemble members within each microensemble: for example, the global ocean heat content anomalies in individual ensemble members from microensemble 1 (Fig. 4b, thin dark blue lines) are always less than those in individual ensemble members in microensemble 2 (Fig. 4b, thin light blue lines). Indeed, only microensembles 3 and 4 show significant overlap between ocean heat content in individual ensemble members (Fig. 4b, compare thin gray and pink lines), although their microensemble means never overlap during the 150-yr experiment. Furthermore, the range of the microensemble means remains relatively constant at 350 ZJ up to year 2100, although the range of individual ensemble members

adds approximately 50 ZJ in additional variance over the course of the experiment (Fig. 4b; compare range of thick lines to range of thin lines). Such persistent differences in global ocean heat content between microensembles are largely attributable to differences in ocean initial conditions produced by ocean drift; when de-drifted, the range of the microensemble means is only 20 ZJ at year 2100 (see Fig. S2).

Figure 5 shows that the average global ocean heat content remains distinct in each microensemble because the mean potential temperature anomaly in the deep ocean in each microensemble [$\overline{\theta}_k(t) - \overline{\theta}(t)$; below 1.5 km] persists through the full 150 years of the experiment. Microensembles 1 and 2 always have cooler than average deep ocean potential temperature anomalies from 1950 to 2100 (Figs. 5a,b), although the magnitude of these cool anomalies appears to dissipate somewhat with time (particularly in microensemble 1; see Fig. 5a). Similarly, microensembles 4 and 5 have warmer than average deep ocean potential temperature anomalies, with larger anomalies near year 1950 than year 2100 (Figs. 5d,e). Unlike the deep ocean, upper ocean potential temperatures (above 1 km) do not persist for nearly so long: in all microensembles, most coherent upper ocean potential temperature anomalies
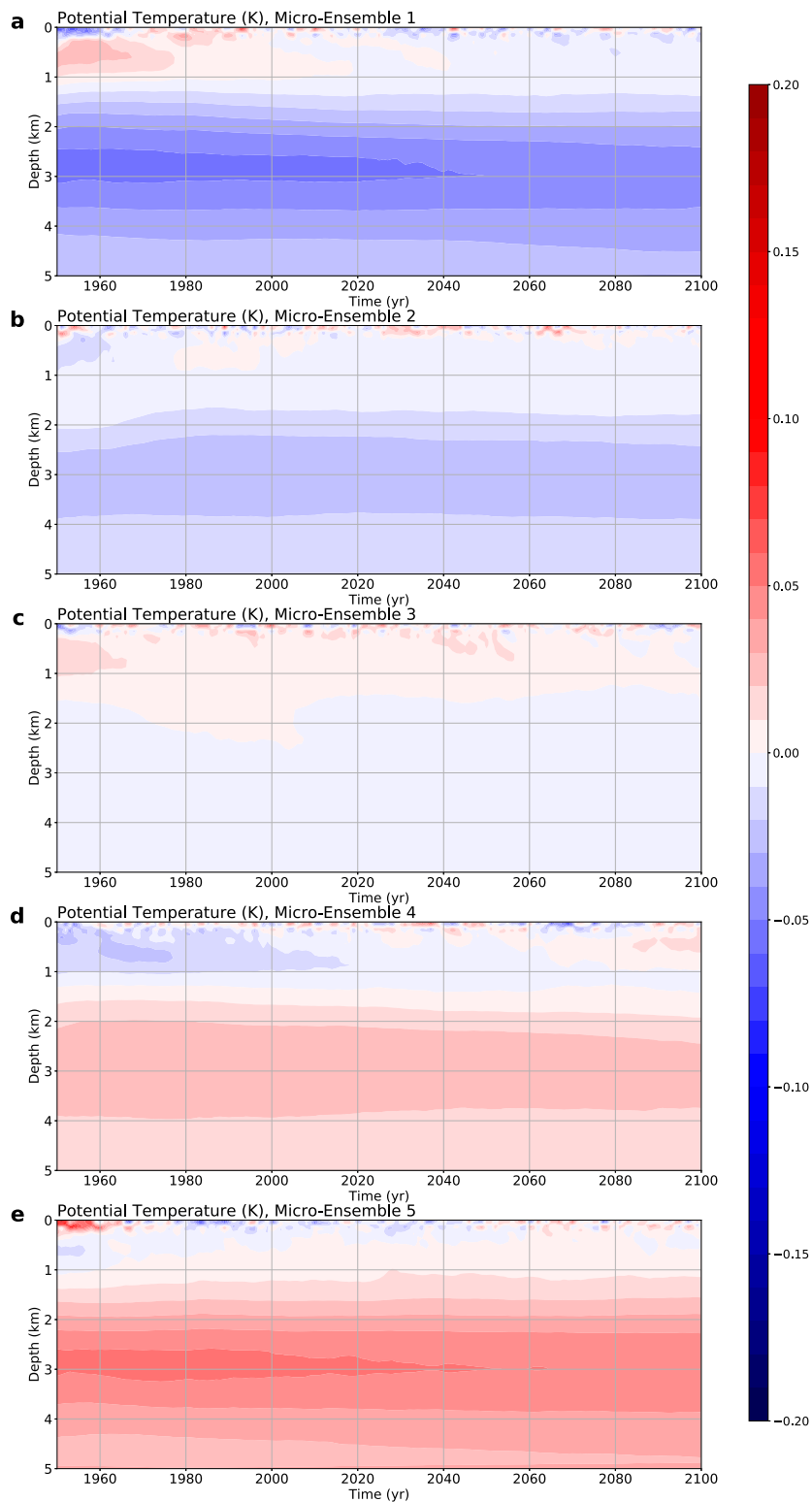
FIG. 5. Evolution of the area-weighted, globally averaged, ocean potential temperature anomaly (in K) in each microensemble. The anomaly is computed relative to the global mean potential temperature in the full ensemble each year.
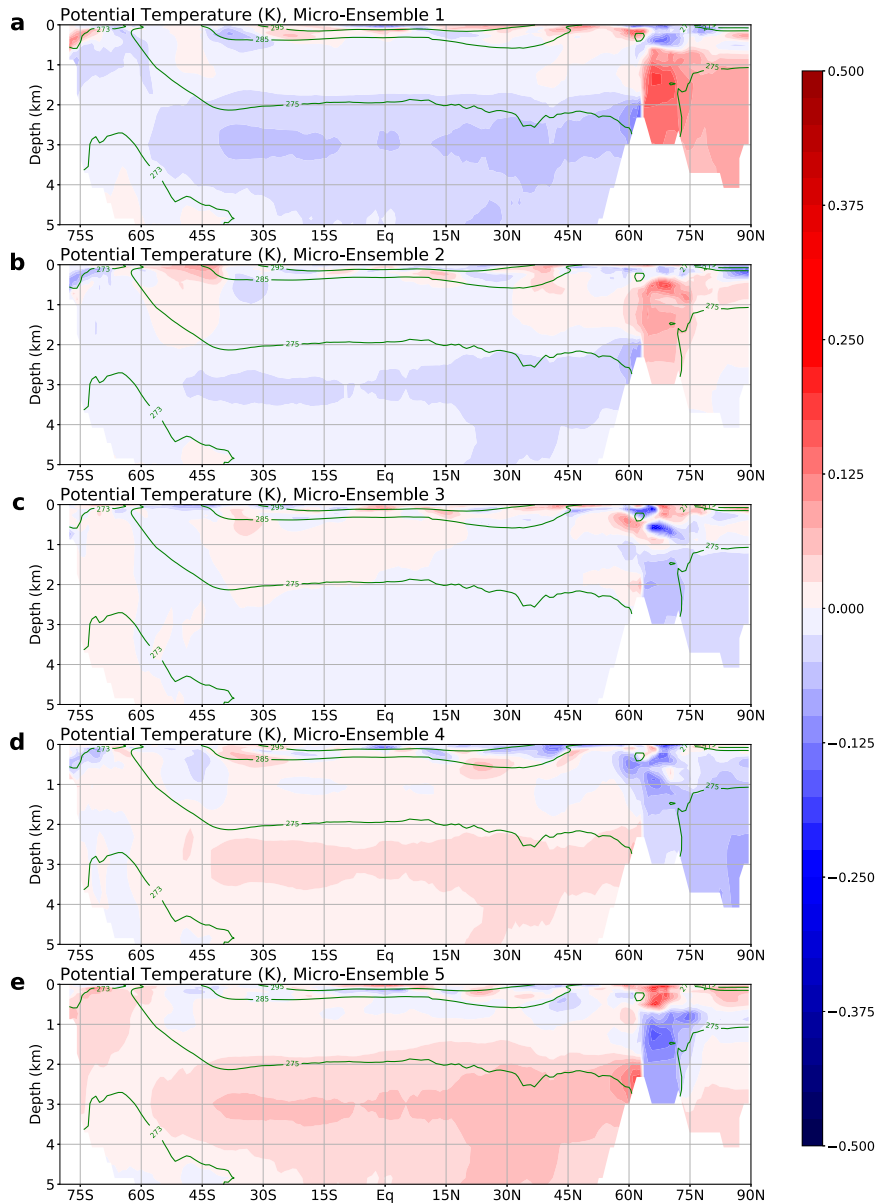
FIG. 6. Zonal-mean ocean potential anomaly (in K; shading) in each microensemble at year 2080 relative to the mean potential temperature (contours at 273, 275, 285, and 295 K) in the full ensemble at year 2080; in other words, $\overline{\mathrm{OHC}_k}(t = 2080) - \overline{\mathrm{OHC}}(t = 2080)$.

have dissipated by year 2000. Even though upper ocean temperatures dissipate over the course of several decades, the average global ocean heat content anomalies in each microensemble (and their constituent individual ensemble members) remain constant with time relative to each other because small (of magnitude 0.1 K) potential temperature anomalies in the deep ocean persist over century-long time scales.

Figure 6 shows the mean potential temperature anomaly at 2080 in each microensemble relative to that in the full ensemble [i.e., $\overline{\theta}_k(t = 2080) - \overline{\theta}(t = 2080)$], which illustrates how the deep ocean temperature differences identified at year 1950

(recall Fig. 2) persist over centennial time scales. In all microensembles, the deep ocean temperature anomalies (below 2000 m and south of 60°N) at year 2080 are of the same sign as those at year 1950, albeit of somewhat weaker magnitude (cf. microensembles in Fig. 6 with same microensembles in Fig. 2; note that the color bar range is twice as large in Fig. 2 as in Fig. 6). De-drifting confirms that differences in deep ocean temperatures between microensembles are predominantly due to ocean drift in the preindustrial control, not internal variability (see Fig. S3). On the other hand, upper ocean temperature anomalies in individual microensembles are substantially weaker at year 2080 than at year 1950, and
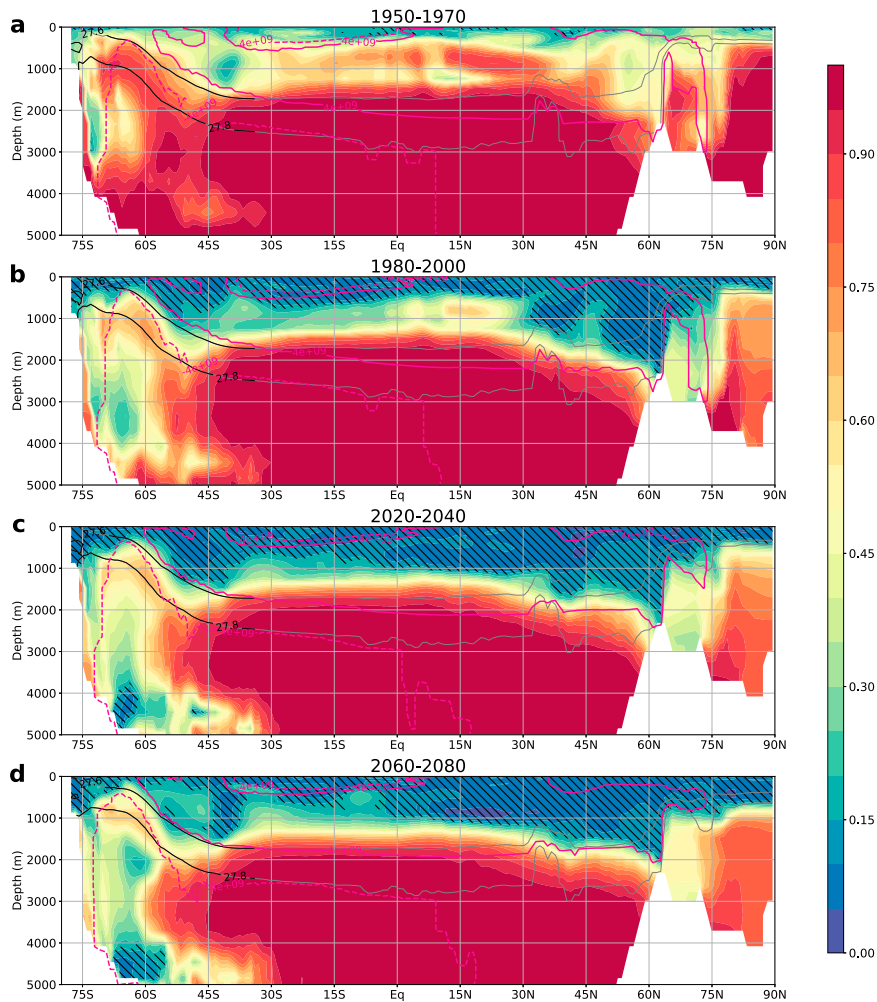
FIG. 7. Fraction of total variance in zonal-mean ocean potential temperature attributable to variance between microensembles, $\chi_{\text{OcnICs}} = \sigma^2_{\theta,\text{ocean}}/\sigma^2_\theta$, over four time periods spanning the full 150 years of the experiment: years (a) 1950–70, (b) 1980–2000, (c) 2020–40, and (d) 2060–80. Also shown are isopycnal contours (solid lines; at sigma levels 27.6 and 27.8 kg m^{-3}) and the ocean meridional mass overturning streamfunction (pink contours at $[-4, 4] \times 10^9$ kg s^{-1}). Hatched areas indicate that the fraction of ensemble variance attributable to ocean initial conditions is not statistically distinct from zero at $p < 0.1$.

are generally not of the same sign or spatially coherent with those at the start of the experiment. In the Arctic basin (poleward of 70°N), we do find some evidence of coherence in temperature anomalies from 1950 and 2080, although not in all microensembles: potential temperature anomalies are of the same sign through the course of the experiment in microensembles 1, 3, and 4, but are of different (or mixed) sign in ensembles 2 and 5.

ATTRIBUTION OF OCEAN STATE EVOLUTION TO ATMOSPHERE AND OCEAN INITIAL STATES

We now compute the fraction of the total variance in ocean potential temperature in the CanESM2 large ensemble that is attributable to ocean initial conditions, $\chi_{\text{OcnICs}} = \sigma^2_{\theta,\text{ocean}}/\sigma^2_\theta$ (i.e., the fraction of the total ensemble variance that is

between microensembles, as detailed in section 2a). Figure 7 shows this quantity from four 20-yr periods over the course of the experiment, and Fig. 8 shows a closer view of the top 2000 m of the water column. Figures S4 and S5 are analogous to Figs. 7 and 8, respectively, but show the variance attributable to ocean initial conditions following de-drifting of each microensemble (see section 2b).

Early in the experiment (from years 1950 to 1970; Figs. 7a and 8a), most ensemble variance in ocean potential temperature below 1500 m is between microensembles (i.e., $\sigma^2_{\theta,\text{ocean}} \gg \sigma^2_{\theta,\text{atmos}}$; note red and orange colors), indicating that it is attributable to the different ocean initial conditions in each microensemble. Even in the upper ocean (above 1000 m), at least half of the ensemble variance is attributable to these differences in ocean initial conditions, though this
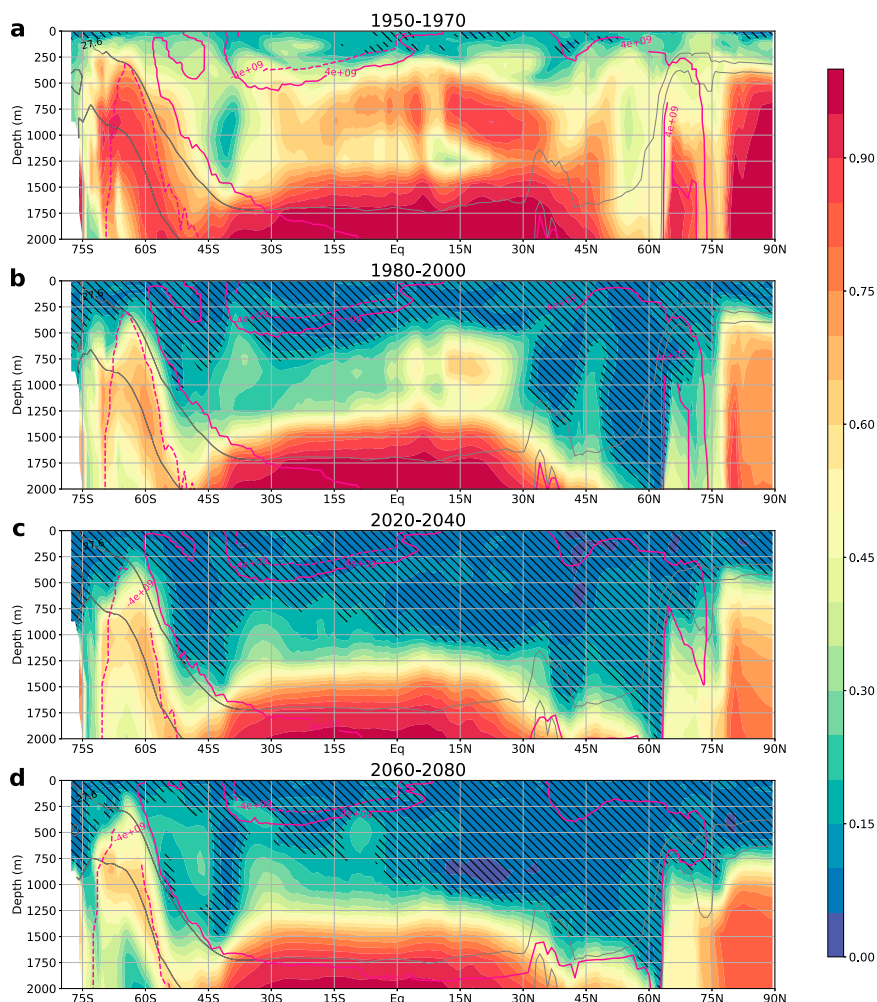
FIG. 8. As for Fig. 7, but only including the top 2000 m of the ocean.

varies by latitude and depth. Above 3500 m, approximately half of this variance in potential temperature arises from differences in ocean initial conditions due to model drift (cf. Figs. 7a and 8a with Figs. S4a and S5a).

By years 1980–2000 (Figs. 7b and 8b) and beyond (Figs. 7c,d and 8c,d), much of the ensemble variance in upper ocean potential temperatures (above 1000 m at most latitudes) is no longer attributable to differences between ocean initial states, but rather to atmospheric variability (note hatched blue and green areas, where the fraction of the variance attributable to ocean initial conditions is not statistically distinct from zero). At some latitudes, atmospheric variability penetrates even deeper into the ocean: in the subpolar Northern Hemisphere, circa 60°N, and also in the deep Southern Ocean, poleward of 60°S below 2000 m. This occurs because the subpolar North Atlantic and the Antarctic continental shelves are locales of weak vertical stratification and deep convection, which allows atmospheric anomalies to penetrate to depth at these latitudes. Indeed, we observe that the variance attributable to ocean initial conditions steadily decreases with time in the

deep Southern Ocean (compare, in succession, Figs. 7b–d), as anomalies attributable to atmospheric variability penetrate farther into the deep ocean along the descending branch of the deep overturning cell (Fig. 7, dotted purple lines).

On the other hand, nearly all ensemble variance in deep ocean temperatures, north of 50°S, is attributable to ocean initial conditions over the full 150 years of the experiment (Fig. 7, dark red regions below 2000 m). Approximately half of this variance attributable to ocean initial conditions is due to potential temperature differences between microensembles arising from deep ocean drift (cf. Figs. 7b–d with Figs. S4b–d). These persistent deep ocean temperature anomalies appear to be isolated from the surface at most latitudes, as only a small fraction of upper ocean temperature variance is attributable to ocean initial conditions. Therefore, persistent deep ocean temperature anomalies (recall Figs. 5 and 6) do not impact surface climate directly. Indeed, the upper ocean is highly stably stratified at most latitudes (Peixoto and Oort 1992), which effectively isolates deep ocean waters from those nearer the surface.

However, in the upper ocean between 60° and 70°S, we find that approximately 50% of ensemble variance is between microensembles over all time periods (Figs. 8a–d) and is therefore attributable to differences in ocean initial conditions. Indeed, we note a plume-like feature that emerges from the deep ocean around 2500 m, near 50°S, where most ensemble variance is due to ocean initial conditions, and follows sloping isopycnal surfaces to the upper ocean near 65°S (see orange and yellow shaded regions between black contours in Figs. 7 and 8). This feature is apparent over all time periods shown (although it does appear to weaken with time; cf. Figs. 8b,d), and is coincident with climatological upwelling of deep waters in the ascending branch of the lower cell of the oceanic meridional overturning circulation (Figs. 7 and 8, dashed pink contour at -4×10^9 kg s^{-1}; also see Marshall and Speer 2012). In other words, the lower cell of the meridional overturning circulation transports deep ocean temperature anomalies, attributable to ocean initial conditions, into the upper ocean circa 65°S. As a result, the Southern Ocean, between 55° and 70°S, is the primary locale where surface conditions are impacted directly by persisting deep ocean temperature anomalies, which are due to differences in ocean initial conditions between microensembles. We note that this upwelling feature is not present when microensembles are de-drifted prior to computing the variance attributable to ocean initial conditions (cf. Figs. S5b–d to Figs. 8b–d), indicating that it arises from differences in deep ocean temperatures between microensembles due to model drift.

We also note that only about half of the temperature variance in the Southern Ocean upwelling branch of the overturning circulation is attributable to ocean initial conditions (particularly over longer time scales; see Figs. 8b–d). This suggests that while persistent deep ocean temperature anomalies upwell along sloping isopycnal surfaces, adiabatic eddies also transport temperature anomalies from the surface to depth along these same isopycnal surfaces (see Gent and McWilliams 1990; Marshall and Speer 2012). Mixing with equatorward-flowing Antarctic Intermediate and Subantarctic Mode Waters (Rintoul 1991) likely also contributes further atmosphere-sourced temperature variance to these upwelling waters. Therefore, temperature anomalies that upwell from the deep ocean are responsible for about half the ensemble variance, while the rest is attributable to variability generated by atmospheric temperature anomalies mixed down from the surface.

c. Impact on surface climate

We now consider the impact of ocean initial conditions on ensemble variance in surface climate, focusing on quantities central to the forced evolution of the ensemble. These include upper ocean heat content, surface temperature trends, and air–sea fluxes, which govern the rate at which the ocean takes up excess heat. As described above, persistent deep ocean temperature anomalies (attributable to differences in ocean initial conditions, as shown in Figs. 7 and 8) primarily affect upper ocean temperature variance between 55° and 75°S. As expected, we find the greatest fraction of variance in upper

ocean heat content (reckoned from the surface to 300 m depth) attributable to ocean initial conditions circa these same Southern Ocean latitudes (Fig. 9a, which shows $\chi_{\text{OcnICs}} = \sigma^2_{\text{OHC,ocean}}/\sigma^2_{\text{OHC}}$; note area between pink horizontal lines, which delineate the Southern Ocean). This is evident over the entire course of the experiment, although it is greatest near the beginning of the experiment (circa year 1960), decreases thereafter, but increases again between years 2055 and 2095.

The primary mechanism by which converging ocean heat impacts the surface climate is through changes in surface turbulent (sensible and latent heat) fluxes (Sutton and Mathieu 2002). This relationship is apparent from the physics that governs evolution of the ocean mixed layer temperature T_o:

$$\rho c_w h_{\text{ML}} \frac{dT_o}{dt} = -\rho c_w h_{\text{ML}} \mathbf{v} \cdot \boldsymbol{\nabla} T_o + Q_{\text{sfc}}(T_o), \qquad (7)$$

where ρ is the density of seawater, c_w is its heat capacity, h_{ML} is the mixed layer depth, $\mathbf{v}$ is the advective velocity, and $Q_{\text{sfc}}(T_o)$ is the sum of the surface fluxes (positive into the ocean). In brief, the temperature evolution of the upper ocean depends on convergent temperature advection by fluid flow ($\rho c_w h_{\text{ML}} \mathbf{v} \cdot \boldsymbol{\nabla} T_o$) and energy loss or gain through surface fluxes [$Q_{\text{sfc}}(T_o)$]. Therefore, temperature anomalies that upwell from the deep drive the evolution of upper Southern Ocean temperatures, which then further impact surface fluxes. Turbulent surface fluxes, in particular, depend on the temperature difference between the ocean surface and overlying atmosphere, indicating that these respond to changes in upper ocean temperature.

Indeed, we find that the Southern Ocean, between 45° and 70°S, is the locale where the greatest fraction of ensemble variance in latent heat fluxes is consistently attributable to ocean initial conditions (i.e., is due to variance between microensembles; Fig. 9b, which shows $\chi_{\text{OcnICs}} = \sigma^2_{F_{\text{LH}},\text{ocean}}/\sigma^2_{F_{\text{LH}}}$; note area between pink horizontal lines, which delineates the Southern Ocean). Furthermore, the fraction of ensemble variance in Southern Ocean latent heat fluxes attributable to ocean initial conditions fluctuates with time similarly to the upper Southern Ocean heat content: greatest from 1960 to 2000, weaker thereafter, and increasing again from 2050 to 2090 (cf. Figs. 9a,b). However, the fraction of ensemble variance attributable to ocean initial conditions for latent heat fluxes is substantially smaller than for upper ocean heat content: only between 10% and 15% of the ensemble variance in Southern Ocean latent heat fluxes, compared to 15%–25% for upper Southern Ocean heat content, is attributable to ocean initial conditions.

Similarly, surface temperature trends over the Southern Ocean also exhibit significant variance due to ocean initial conditions (Fig. 9c, which shows $\chi_{\text{OcnICs}} = \sigma^2_{dT_s/dt,\text{ocean}}/\sigma^2_{dT_s/dt}$; note area between pink horizontal lines) because upper ocean heat convergence impacts the ocean temperature tendency, dT_o/dt [recall Eq. (7)]. Like the ensemble variance in latent heat fluxes described above, the variance in Southern Ocean surface temperature trends also fluctuates with time similarly
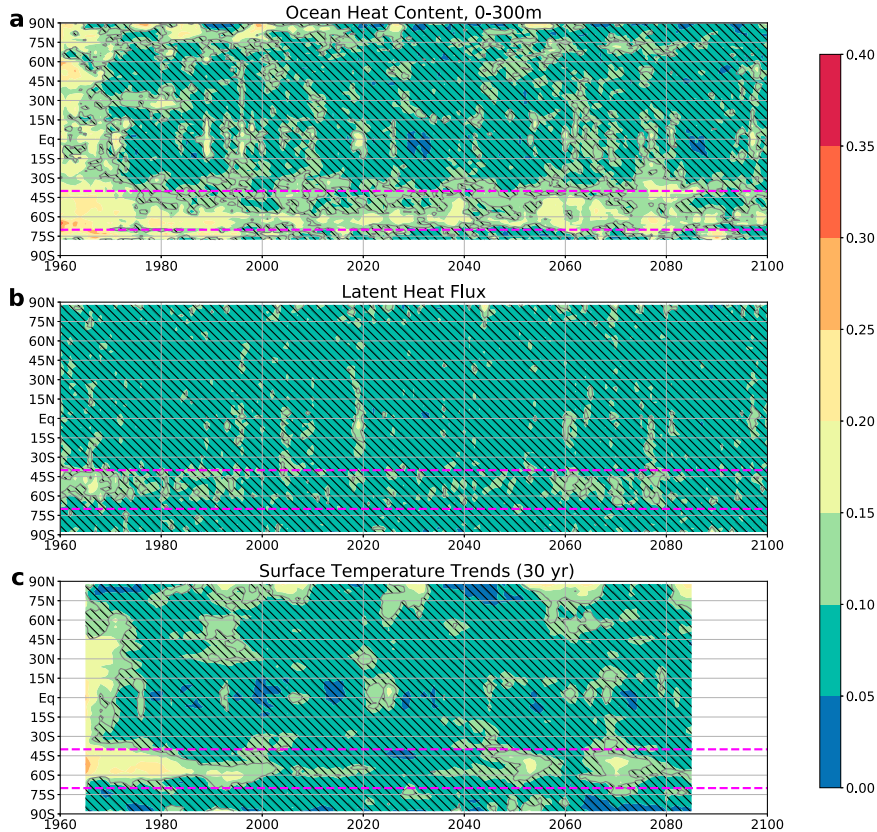
FIG. 9. Zonal-mean fraction of ensemble variance in (a) upper-300-m ocean heat content, (b) latent heat flux, and (c) 30-yr surface temperature trends, attributable to variance between microensembles ($\chi_{\text{OcnICs}} = \sigma^2_{X,\text{ocean}}/\sigma^2_X$) over the full 150 years of the ensemble. Hatched areas indicate that the fraction of ensemble variance attributable to ocean initial conditions is not statistically distinct from zero at $p < 0.1$ at more than 25% of the grid cells at that latitude. Dashed horizontal pink lines at 40° and 70°S delineate the Southern Ocean.

to the upper ocean heat content variance, and is also weaker in magnitude.

In Fig. 10, we examine surface flux anomalies (from 55°S to the pole) over four time periods in each microensemble, calculated as the difference between the microensemble mean and the full ensemble mean [i.e., $\overline{F}_{X,k}(t) - \overline{F}_X(t)$]. We find systematic differences between turbulent fluxes, both latent (F_{LH}; Fig. 10a) and sensible (F_{SH}; Fig. 10b), in microensembles with colder than average deep ocean temperatures (microensembles 1 and 2) compared to those with warmer than average deep ocean temperatures (microensembles 4 and 5). When deep ocean temperatures are anomalously cold, as in microensembles 1 and 2, both latent and sensible heat fluxes are anomalously low relative to the full ensemble mean over all time periods [Figs. 10a,b, dark and light blue markers; $\overline{F}_{X,1,2}(t) < \overline{F}_X(t)$]; conversely, when deep ocean temperatures are anomalously warm, as in microensembles 4 and 5, turbulent fluxes are anomalously high [Figs. 10a,b, pink and red markers; $\overline{F}_{X,4,5}(t) > \overline{F}_X(t)$]. The sign of these turbulent flux anomalies in each microensemble is consistent with the sign of the deep ocean temperature anomalies reported earlier [recall $\overline{\theta}_k(t) - \overline{\theta}(t)$ in Figs. 2, 5, and 6]: when warmer deep

ocean temperature anomalies advect into the upper ocean, we find ocean heat content and turbulent heat fluxes to be higher than the ensemble average (as in microensembles 4 and 5); on the other hand, when cooler deep ocean temperature anomalies advect into the upper ocean, we find that ocean heat content is lower than average and turbulent heat fluxes are weak (as in microensembles 1 and 2).

Differences in Southern Ocean turbulent fluxes between microensembles, attributable to deep ocean temperature differences, also impact the ocean heat uptake (OHU). The rate of deep ocean heat uptake is central to the forced transient climate response (Boé et al. 2009; Kuhlbrodt and Gregory 2012), and the Southern Ocean is the locale where most of this heat uptake occurs (Frölicher et al. 2015; Shi et al. 2018). The ocean heat uptake is computed as

$$\text{OHU} = R^\downarrow_{\text{SW+LW}} - F_{\text{SH}} - F_{\text{LH}}, \qquad (8)$$

where $R^\downarrow_{\text{SW+LW}}$ is the net (downward, shortwave plus longwave) radiative flux at the surface. In microensembles 1 and 2 where mean deep ocean temperatures are anomalously cool compared to the ensemble mean, turbulent heat fluxes over
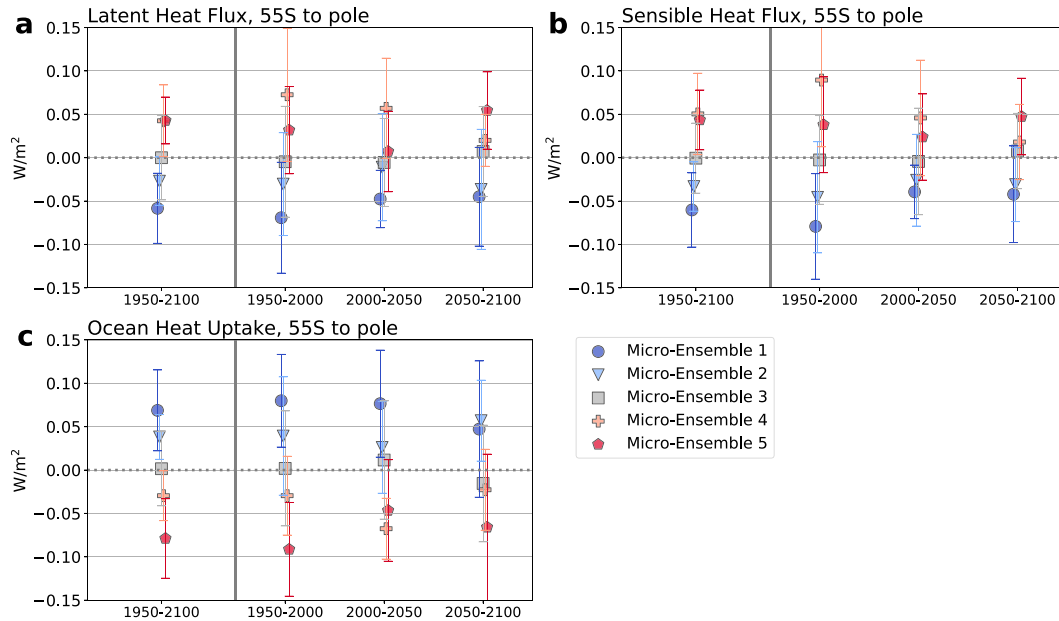
FIG. 10. Microensemble anomalies (in W m^{-2}) in (a) latent heat fluxes, (b) sensible heat fluxes, and (c) ocean heat uptake, all poleward of 55°S, in the CanESM2 large ensemble. We reckon both sensible and latent heat fluxes to be positive upward, while ocean heat uptake is positive downward. Anomalies for each microensemble are computed with respect to the mean of the full ensemble [i.e., as $\overline{X}_k(t) - \overline{X}(t)$] and are calculated over four time periods: the full 150 years of the experiment (1950–2100), from 1950 to 2000, from 2000 to 2050, and from 2050 to 2100. Over all time periods and for all quantities, the fraction of ensemble variance due to the ocean initial state is statistically significant at $p < 0.1$, with the exception of the sensible heat flux over years 2000–2050. Vertical bars indicate the standard deviation within each microensemble (i.e., $\sigma_{X,\mathrm{atmos},k}$ for the kth microensemble).

the Southern Ocean are weaker than the ensemble mean, and ocean heat uptake is greater than the ensemble mean over all time periods [Fig. 10c, dark and light blue markers; $\overline{\mathrm{OHU}}_{1,2}(t) > \overline{\mathrm{OHU}}(t)$]. Similarly, in microensembles 4 and 5 where mean deep ocean temperatures are anomalously warm compared to the ensemble mean, turbulent heat fluxes are more vigorous than the ensemble mean, and ocean heat uptake is weaker than the ensemble mean over all time periods [Fig. 10c, red and pink markers; $\overline{\mathrm{OHU}}_{4,5}(t) < \overline{\mathrm{OHU}}(t)$]. In other words, persistent cool anomalies in the deep ocean tend to augment Southern Ocean heat uptake with CO_2 forcing, while persistent warm anomalies in the deep ocean tend to suppress Southern Ocean heat uptake.

In CanESM2, the microensemble mean ocean heat uptake anomaly scales approximately one-to-one with the initial microensemble mean deep ocean temperature anomaly:

$$\frac{\overline{\mathrm{OHU}}_k(t) - \overline{\mathrm{OHU}}(t)}{\overline{T}_{\mathrm{deep},k}(t=1950) - \overline{T}_{\mathrm{deep}}(t=1950)} \sim -1 \text{ W m}^{-2}\text{ K}^{-1}. \quad (9)$$

For example, an initial mean deep ocean temperature anomaly of -0.1 K, as in microensemble 1, gives rise to approximately a 0.1 W m^{-2} mean anomaly in ocean heat uptake in microensemble 1 over the first 100 years of the experiment (i.e., from 1950 to 2000, and from 2000 to 2050; Fig. 10). We note that this scaling depends on the rate at which the ocean

meridional overturning upwells anomalies from the deep ocean, which varies substantially between global climate models (see, e.g., Behrens et al. 2016). Moreover, because deep ocean temperature differences between microensembles decrease with time (recall Fig. 5), we also expect this scaling to weaken over longer time scales.

Although it is clear that Southern Ocean heat uptake is sensitive to differences in deep ocean temperature between microensembles, we note that the ensemble range (i.e., the total ensemble spread, which is attributable to both atmospheric microperturbations and ocean initial condition differences) becomes substantially smaller over time relative to the forced response. Over years 1950–2000, the ensemble range in Southern Ocean heat uptake is of similar magnitude to the forced change: both are approximately 0.5 W m^{-2}. Over years 2000–2050, the ensemble range in Southern Ocean heat uptake decreases slightly to approximately 0.4 W m^{-2}, but greenhouse gas forcing has now increased ocean heat uptake over this region to 1.7 W m^{-2}. By years 2050–2100, the ensemble range is only a small fraction of the forced response in Southern Ocean heat uptake: the ensemble range is still approximately 0.4 W m^{-2}, but the forced change over the region has increased to 3.8 W m^{-2}, so uncertainty due to internal variability is only about 10% of the forced response. Thus, although ensemble spread (due to internal variability stemming from both macro- and microinitialization) contributes to
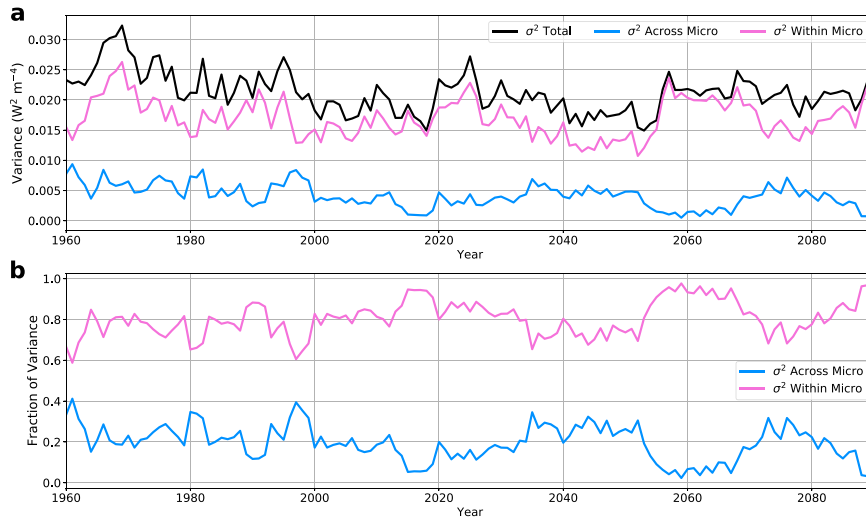
FIG. 11. Ensemble variance in ocean heat uptake poleward of 55°S: (a) total ensemble variance over the full 150 years of the experiment (black line), partitioned into the variance between microensembles ($\sigma^2_{\text{OHU,ocean}}$; blue line) and within microensembles ($\sigma^2_{\text{OHU,atmos}}$; purple line) and (b) fraction of the total ensemble variance between microensembles (blue line) and within microensembles (purple line).

uncertainty in Southern Ocean heat uptake over centennial time scales, it is likely that other sources of uncertainty (including that due to model physics and emissions scenario) are responsible for most of the uncertainty in the forced response over such long time scales (Hawkins and Sutton 2009). Indeed, the intermodel variance in ocean heat uptake at year 2100 in CMIP5-participating ESMs (subject to RCP8.5 forcing) is approximately 15 W m^{-2} at 60°S (Lyu et al. 2020), a range that is over an order of magnitude larger than the total variance in the CanESM2 ensemble.

In Fig. 11, we examine the variance in Southern Ocean heat uptake (from 55°S to the pole, as in Fig. 10c) between microensembles ($\sigma^2_{\text{OHU,ocean}}$; blue lines) and within microensembles ($\sigma^2_{\text{OHU,atmos}}$; purple lines). The total variance in the ocean heat uptake appears to decrease slightly over the first several decades, but thereafter remains relatively constant (Fig. 11a, black line). This suggests greater ensemble variance attributable to ocean initial conditions at the beginning of the experiment (approximately 30% over the first 50 years; Fig. 11b), and less ensemble variance attributable to ocean initial conditions near the end of the experiment (approximately 20% over the final 50 years). We note that the fraction of the ensemble variance in ocean heat uptake attributable to ocean initial conditions does not dwindle to zero because deep ocean temperature differences between microensembles continue to persist even at year 2100. Given the modest rate of Southern Ocean upwelling (of order 10^9 kg s^{-1}; recall Fig. 7) and the enormous volume of the deep ocean (of order 10^8 km^3), these deep ocean temperature anomalies can be expected to persist for over 10^3 years. As long as these deep ocean temperature anomalies exist, we expect that they will continue to impact surface fluxes over the Southern Ocean, albeit more modestly with time as their magnitude declines.

4. Discussion

In this study, we have used the CanESM2 large ensemble to answer a simple, but important, question: How much do varying ocean initial conditions impact variance in ESM large ensembles? To answer this, we have harnessed the macro-/microstructure of the CanESM2 large ensemble, first of its kind among full-complexity climate models, to separate ensemble variance due to ocean initial conditions from that due to atmospheric microperturbations. We find that deep ocean potential temperature anomalies associated with different ocean initial conditions, generated by deep ocean drift in the preindustrial control, persist for at least 150 years following model initialization. These anomalies primarily impact surface climate over the Southern Ocean as they upwell to the surface along the ascending branch of the lower cell of the ocean meridional overturning circulation. In turn, some ensemble variance in Southern Ocean heat content (from the surface to 300 m depth), turbulent heat fluxes, temperature trends, and ocean heat uptake is attributable to ocean initial conditions. In other words, using a range of ocean states to initialize a large ensemble increases uncertainty in how the Southern Ocean evolves, which is arguably the region that is most consequential for determining the pace of climate change. Although these impacts on surface climate are localized to the Southern Ocean and modest in magnitude, they are persistent over the full 150 years of the ensemble, and suggest that uncertainties in Southern Ocean surface climate due to uncertainties in ocean initial conditions can be expected to persist over at least 150 years and likely longer.

Most striking is the strength of the relationship between mean deep ocean temperature anomalies ($\overline{T}_{\text{deep},k} - \overline{T}_{\text{deep}}$) and mean Southern Ocean heat uptake anomalies in a given microensemble ($\overline{\text{OHU}_k} - \overline{\text{OHU}}$): we find that a 1 K anomaly

in deep ocean temperatures in a microensemble, relative to the full ensemble mean, would result in a -1 W m^{-2} anomaly in Southern Ocean heat uptake in that microensemble relative to full ensemble mean [recall Eq. (9)]. We expect that this relationship is model dependent, as the rate of upwelling of deep ocean temperature anomalies by the ocean meridional overturning circulation will determine the magnitude of the upper ocean heat content anomaly due to these deep ocean anomalies and, therefore, their impact on surface turbulent fluxes. Furthermore, the time scales over which deep ocean temperature anomalies persist, and continue to impact surface fluxes over the Southern Ocean, also depends on this same model-dependent rate of upwelling of deep ocean anomalies: models with a more vigorous meridional circulation will more rapidly dissipate any deep ocean temperature anomalies, while models with a weaker circulation will tend to have more persistent deep ocean temperature anomalies. Nevertheless, insofar as representation of ocean temperatures in climate models remains imperfect (see, e.g., Pohlmann et al. 2009; Smith et al. 2013; Yeager et al. 2018), we expect that there will be some irreducible uncertainty in the Southern Ocean surface energy budget over some time scale in all models. Such uncertainty further increases uncertainty in the transient climate response, as Southern Ocean processes determine the rate of deep ocean heat uptake and, therefore, the rate at which the globe warms in response to anthropogenic greenhouse gas emissions.

Our findings suggest that the Southern Ocean is the primary locale where persisting deep ocean anomalies continue to impact the surface climate over centennial (and longer) time scales. Previous studies have also pointed to the Southern Ocean as being a key site where deep and intermediate-depth ocean processes impact surface climate, through upwelling (Lumpkin and Speer 2007; Talley 2013; Tamsitt et al. 2017) or internal variability (Latif et al. 2013; Behrens et al. 2016; Zhang et al. 2019). Because the Southern Ocean is a central player in global heat and carbon uptake, which together govern how the climate system evolves, deep and intermediate-depth Southern Ocean processes that govern the rate of uptake also have the potential to impact secular climate trends over long time scales (see, e.g., Morrison et al. 2013; Marshall and Zanna 2014; Exarchou et al. 2015).

Surprisingly, we do not find that deep ocean temperature anomalies impact the Northern Hemisphere oceans, particularly the Arctic, over such long time scales. We submit that this is because deep ocean temperature anomalies in the Arctic basin do not have a ready pathway to upwell to the surface, as ocean density stratification is particularly strong under perennial sea ice cover (due to the presence of the cold halocline; see Aagaard et al. 1981). Furthermore, deep and intermediate convection in the North Atlantic tends to bring atmospheric anomalies to depth (where they flow equatorward in the deep branch of the upper cell; Peixoto and Oort 1992; Buckley and Marshall 2016), rather than bringing deep ocean anomalies up to the surface as occurs in the Southern Ocean. This behavior highlights the unique features of the Southern Ocean, particularly the upwelling branch contained therein, which closes the oceanic meridional overturning

circulation (Marshall and Speer 2012) and transports anomalies from the deep ocean to the surface.

Our analysis of the CanESM2 large ensemble corroborates the results of Hawkins et al. (2016), who also showed that varying ocean initial conditions increased variance in a large ensemble, albeit in one utilizing an Earth system model of intermediate complexity (EMIC), not a full ESM. While Hawkins et al. (2016) predominantly focus on the North Atlantic, and how initializing the model in different phases of the Atlantic multidecadal oscillation impacts Northern Hemisphere surface climate over multidecadal time scales, our work suggests that it is in the Southern Ocean where the impact of ocean initial conditions on ensemble variance persists over centennial time scales. We hypothesize that this difference may be due to the substantial multidecadal periodicity in the strength of the Atlantic meridional overturning circulation in the EMIC utilized by Hawkins et al. (2016). Because CanESM2 does not display such regular multidecadal variability in the strength of the global overturning circulation [as described in Behrens et al. (2016)], the impact of ocean initial conditions in our large ensemble likely depends less on the phase of coupled modes of variability, and more on the persistence of deep ocean temperatures.

Because temperature anomalies associated with ocean initial conditions can contribute substantially to ensemble variance in surface climate, potentially over very long time scales in the Southern Ocean as shown in this study, we suggest that it would be prudent to consider which ocean states are used to initialize a large ensemble. Our results indicate that an ensemble generated from a wide sampling of ocean initial states is necessary for generating maximum ensemble variance, if that were the goal. This sampling may be generated by ocean drift, as in the CanESM2 large ensemble, or through multidecadal or centennial time scale internal variability in a long control run. However, the precise way to sample ocean initial conditions in order to generate maximum, yet plausible, ensemble variance remains unexplored. Only a few studies have quantified internal variability in deep ocean heat content in models and observations (see, e.g., Santer et al. 1995; Häkkinen et al. 2013; Palter et al. 2014; Palmer et al. 2017). Moreover, ocean drift may yield ocean states that vary greatly over the course of a long control run (Sen Gupta et al. 2012; Hobbs et al. 2016), but it is unclear what subset of these drifting states it is most appropriate to sample from. On the other hand, a more limited set of ocean initial states may be preferable if some aspect of the ocean state is well constrained, such as the phase of the Atlantic multidecadal oscillation or the Pacific decadal oscillation, for example, or the total (or regional) ocean heat content. We suggest that the choice of ocean initial states is an important component of ensemble design, and this choice should reflect the goals of the ensemble.

Before concluding it is important to acknowledge that while variance in the ocean initial state continues to generate ensemble variance in the Southern Ocean surface energy budget over long time scales, the impacts of different ocean initial conditions on multidecadal and centennial time scale trends are relatively small over the rest of the globe in the CanESM2 large ensemble. Indeed, the impact of different ocean initial
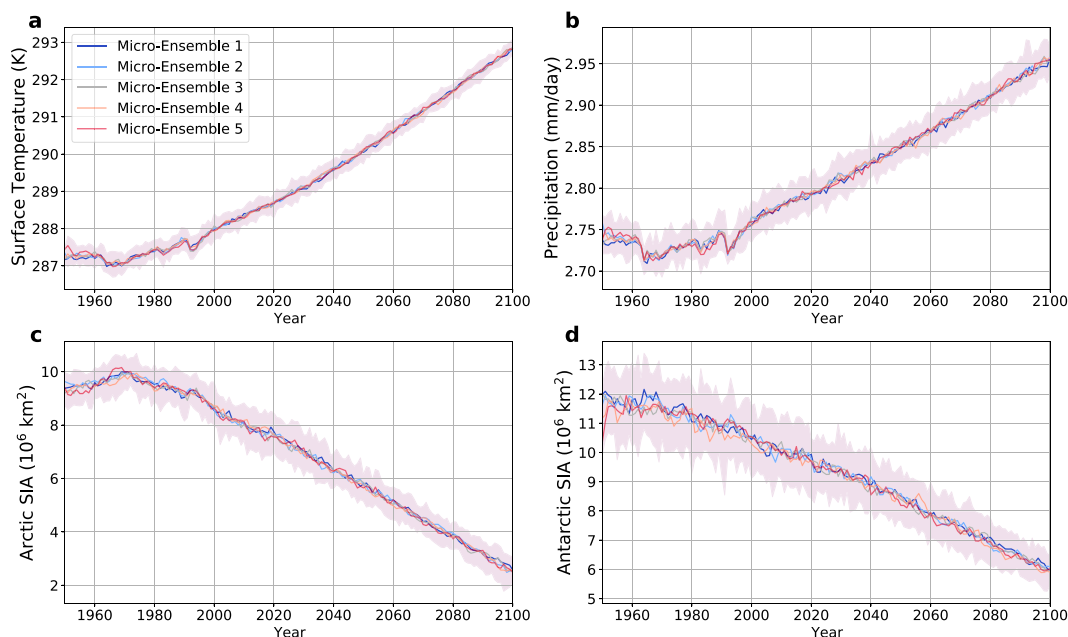
Fig. 12. Evolution of annual mean (a) global surface temperature (in K), (b) global precipitation (in mm day^{-1}), (c) Arctic sea ice area (in 10^6 km^2), and (d) Antarctic sea ice area (in 10^6 km^2) in the CanESM2 large ensemble. Color-coded lines show the microensemble means, and the shaded envelopes indicate the range of the annual mean in the full ensemble.

conditions on the global mean surface temperature and precipitation is not discernible beyond the first decade following ensemble initialization (Figs. 12a,b, respectively); even Arctic and Antarctic sea ice area show little sensitivity to ocean initial conditions beyond the first several decades following model initialization (Figs. 12c,d, respectively). Even over the Southern Ocean, where ocean initial conditions continue to impact surface fluxes over long time scales, we do not find systematic impacts of these on local atmospheric circulation features, such as jet position, the westerly wind maximum, and sea level pressure. We therefore conclude that because the variance attributable to ocean initial conditions is low over much of the upper ocean, apart from the Southern Ocean, and because the atmosphere is highly effective at generating variability, it is possible that centennial time scale projections of most quantities may be robust to the choice of the ocean initial state. We also must note that over such long time scales, uncertainty due to ensemble spread (whether arising from macro- or microinitialization) is small compared to the magnitude of the forced response [as evident in Fig. 12; also see Deser et al. (2012) and Kay et al. (2015)] and, for most quantities, is generally smaller than other sources of uncertainty (including uncertainties in model physics and future emissions scenario; see Hawkins and Sutton 2009).

Finally, we conclude with some caveats of the analysis we have presented here. First, our results rely on a large ensemble that utilizes a single global climate model, the CanESM2. As we discuss above, it is likely that some of our findings are model dependent. This includes the magnitude of the relationship between deep ocean temperatures and Southern

Ocean heat uptake, and how the phasing of coupled variability modes affects model evolution [recall differences between the CanESM2 large ensemble and that of Hawkins et al. (2016), as discussed above]. Furthermore, we point out that the creators of the CanESM2 large ensemble did not endeavor to maximize ensemble variance by choosing a range of ocean initial conditions from which to branch their microensembles. Because of deep ocean drift in the CanESM2 preindustrial control run, they sampled a range of deep ocean potential temperature variance larger than that possible with internal variability alone in this model, but within the range of uncertainty in deep ocean temperatures before the modern observational era (see Crowley et al. 2003; Abraham et al. 2013; Cheng et al. 2017; Zanna et al. 2019). Since this large ensemble was one of convenience, rather than one of design, the fraction of ensemble variance attributable to ocean initial conditions reported here should be interpreted as neither an upper bound nor a lower bound of this quantity. Further study will be necessary to understand exactly how large these upper and lower bounds in ensemble variance might be, given uncertainties in the ocean state. Despite these caveats, we expect that as long as there are uncertainties in reckoning the ocean state, these will likely contribute to irreducible uncertainty for future climate projections, especially over the Southern Ocean.

large ensemble simulations, and the Canadian Sea Ice and Snow Evolution (CanSISE) Network for proposing the simulations. The work of LMP is funded by an award from the U.S. National Science Foundation to Columbia University. All authors are grateful for helpful feedback and critiques from four anonymous reviewers.

REFERENCES

Aagaard, K., L. K. Coachman, and E. Carmack, 1981: On the halocline of the Arctic Ocean. *Deep-Sea Res.*, **28A**, 529–545, https://doi.org/10.1016/0198-0149(81)90115-1.

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

Arora, V. K., and Coauthors, 2011: Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases. *Geophys. Res. Lett.*, **38**, L05805, https://doi.org/10.1029/2010GL046270.

Athanasiadis, P. J., S. Yeager, Y.-O. Kwon, A. Bellucci, D. W. Smith, and S. Tibaldi, 2020: Decadal predictability of North Atlantic blocking and the NAO. *npj Climate Atmos. Sci.*, **3**, 20, https://doi.org/10.1038/s41612-020-0120-6.

Behrens, E., G. Rickard, O. Morgenstern, T. Martin, A. Osprey, and M. Joshi, 2016: Southern Ocean deep convection in global climate models: A driver for variability of subpolar gyres and Drake Passage transport on decadal timescales. *J. Geophys. Res. Oceans*, **121**, 3905–3925, https://doi.org/10.1002/2015JC011286.

Bellenger, H., E. Guilyardi, J. Leloup, M. Lengaigne, and J. Vialard, 2014: ENSO representation in climate models: From CMIP3 to CMIP5. *Climate Dyn.*, **42**, 1999–2018, https://doi.org/10.1007/s00382-013-1783-z.

Bellucci, A., and Coauthors, 2013: Decadal climate predictions with a coupled OAGCM initialized with oceanic reanalyses. *Climate Dyn.*, **40**, 1483–1497, https://doi.org/10.1007/s00382-012-1468-z.

Boé, J., A. Hall, and X. Qu, 2009: Deep ocean heat uptake as a major source of spread in transient climate change simulations. *Geophys. Res. Lett.*, **36**, L22701, https://doi.org/10.1029/2009GL040845.

Buckley, M. W., and J. Marshall, 2016: Observations, inferences, and mechanisms of the Atlantic meridional overturning circulation: A review. *Rev. Geophys.*, **54**, 5–63, https://doi.org/10.1002/2015RG000493.

Cavalieri, D., C. Parkinson, P. Gloersen, and H. Zwally, 1996: Updated yearly: Sea ice concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS passive microwave data, version 1. NASA National Snow and Ice Data Center Distributed Archive Center, accessed 2021, https://nsidc.org/data/nsidc-0051/versions/1.

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

Chikamoto, Y., and Coauthors, 2013: An overview of decadal climate predictability in a multimodel ensemble by climate model MIROC. *Climate Dyn.*, **40**, 1201–1222, https://doi.org/10.1007/s00382-012-1351-y.

Christian, J. R., and Coauthors, 2010: The global carbon cycle in the Canadian Earth System Model (CanESM1): Preindustrial control simulation. *J. Geophys. Res.*, **115**, G03014, https://doi.org/10.1029/2008JG000920.

Crowley, T. J., S. K. Baum, K.-Y. Kim, G. C. Hegerl, and W. T. Hyde, 2003: Modeling ocean heat content changes during the last millennium. *Geophys. Res. Lett.*, **30**, 1932, https://doi.org/10.1029/2003GL017801.

Deser, C., A. S. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate change projections: The role of internal variability. *Climate Dyn.*, **38**, 527–546, https://doi.org/10.1007/s00382-010-0977-x.

——, ——, M. A. Alexander, and B. V. Smoliak, 2014: Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *J. Climate*, **27**, 2271–2296, https://doi.org/10.1175/JCLI-D-13-00451.1.

——, and Coauthors, 2020: Insights from Earth system model initial-condition large ensembles and future prospects. *Nat. Climate Change*, **10**, 277–286, https://doi.org/10.1038/s41558-020-0731-2.

Ding, Q., and Coauthors, 2017: Influence of high-latitude atmospheric circulation changes on summertime Arctic sea ice. *Nat. Climate Change*, **7**, 289–295, https://doi.org/10.1038/nclimate3241.

Dunstone, N. J., D. M. Smith, and R. Eade, 2011: Multi-year predictability of the tropical Atlantic atmosphere driven by the high-latitude North Atlantic Ocean. *Geophys. Res. Lett.*, **38**, L14701, https://doi.org/10.1029/2011GL047949.

England, M., A. Jahn, and L. Polvani, 2019: Nonuniform contribution of internal variability to recent Arctic sea ice loss. *J. Climate*, **32**, 4039–4053, https://doi.org/10.1175/JCLI-D-18-0864.1.

Exarchou, E., T. Kuhlbrodt, J. M. Gregory, and R. S. Smith, 2015: Ocean heat uptake processes: A model intercomparison. *J. Climate*, **28**, 887–908, https://doi.org/10.1175/JCLI-D-14-00235.1.

Flato, G., and Coauthors, 2014: Evaluation of climate models. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 741–866.

Frölicher, T. L., J. L. Sarmiento, D. J. Paynter, J. P. Dunne, J. P. Krasting, and M. Winton, 2015: Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *J. Climate*, **28**, 862–886, https://doi.org/10.1175/JCLI-D-14-00117.1.

Ganachaud, A., and C. Wunsch, 2000: Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature*, **408**, 453–457, https://doi.org/10.1038/35044048.

Gasparin, F., M. Hamon, E. Rémy, and P.-Y. Le Traon, 2020: How deep Argo will improve the deep ocean in an ocean reanalysis. *J. Climate*, **33**, 77–94, https://doi.org/10.1175/JCLI-D-19-0208.1.

Gebbie, G., and P. Huybers, 2019: The Little Ice Age and 20th-century deep Pacific cooling. *Science*, **363**, 70–74, https://doi.org/10.1126/science.aar8413.

Gent, P. R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.*, **20**, 150–155, https://doi.org/10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2.

Griffies, S. M., and K. Bryan, 1997: Predictability of North Atlantic multidecadal climate variability. *Science*, **275**, 181–184, https://doi.org/10.1126/science.275.5297.181.

Hagos, S. M., L. R. Leung, J.-H. Yoon, J. Lu, and Y. Gao, 2016: A projection of changes in landfalling atmospheric river frequency and extreme precipitation over western North America from the large ensemble CESM simulations. *Geophys. Res. Lett.*, **43**, 1357–1363, https://doi.org/10.1002/2015GL067392.

Häkkinen, S., P. B. Rhines, and D. L. Worthen, 2013: Northern North Atlantic sea surface height and ocean heat content variability. *J. Geophys. Res. Oceans*, **118**, 3670–3678, https://doi.org/10.1002/jgrc.20268.

Hawkins, E., and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bull. Amer. Meteor. Soc.*, **90**, 1095–1108, https://doi.org/10.1175/2009BAMS2607.1.

——, R. S. Smith, J. M. Gregory, and D. A. Stainforth, 2016: Irreducible uncertainty in near-term climate projections. *Climate Dyn.*, **46**, 3807–3819, https://doi.org/10.1007/s00382-015-2806-8.

Hobbs, W., M. D. Palmer, and D. Monselesan, 2016: An energy conservation analysis of ocean drift in the CMIP5 global coupled models. *J. Climate*, **29**, 1639–1653, https://doi.org/10.1175/JCLI-D-15-0477.1.

Jahn, A., J. E. Kay, M. M. Holland, and D. M. Hall, 2016: How predictable is the timing of a summer ice-free Arctic? *Geophys. Res. Lett.*, **43**, 9113–9120, https://doi.org/10.1002/2016GL070067.

Kang, S. M., C. Deser, and L. M. Polvani, 2013: Uncertainty in climate change projections of the Hadley circulation: The role of internal variability. *J. Climate*, **26**, 7541–7554, https://doi.org/10.1175/JCLI-D-12-00788.1.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1.

Khairoutdinov, M., and Y. Kogan, 2000: A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Mon. Wea. Rev.*, **128**, 229–243, https://doi.org/10.1175/1520-0493(2000)128<0229:ANCPPI>2.0.CO;2.

Kirchmeier-Young, M. C., F. W. Zwiers, and N. P. Gillett, 2017: Attribution of extreme events in Arctic sea ice extent. *J. Climate*, **30**, 553–571, https://doi.org/10.1175/JCLI-D-16-0412.1.

Kuhlbrodt, T., and J. M. Gregory, 2012: Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. *Geophys. Res. Lett.*, **39**, L18608, https://doi.org/10.1029/2012GL052952.

Large, W. G., J. C. McWilliams, and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, **32**, 363–403, https://doi.org/10.1029/94RG01872.

Latif, M., and N. S. Keenlyside, 2011: A perspective on decadal climate variability and predictability. *Deep-Sea Res. II*, **58**, 1880–1894, https://doi.org/10.1016/j.dsr2.2010.10.066.

——, T. Martin, and W. Park, 2013: Southern Ocean sector centennial climate variability and recent decadal trends. *J. Climate*, **26**, 7767–7782, https://doi.org/10.1175/JCLI-D-12-00281.1.

Li, J., and H. W. Barker, 2005: A radiation algorithm with correlated-k distribution. Part I: Local thermal equilibrium. *J. Atmos. Sci.*, **62**, 286–309, https://doi.org/10.1175/JAS-3396.1.

Lovenduski, N. S., G. A. McKinley, A. R. Fay, K. Lindsay, and M. C. Long, 2016: Partitioning uncertainty in ocean carbon uptake projections: Internal variability, emission scenario, and model structure. *Global Biogeochem. Cycles*, **30**, 1276–1287, https://doi.org/10.1002/2016GB005426.

Lumpkin, R., and K. Speer, 2007: Global ocean meridional overturning. *J. Phys. Oceanogr.*, **37**, 2550–2562, https://doi.org/10.1175/JPO3130.1.

Lyu, K., X. Zhang, J. A. Church, and Q. Wu, 2020: Processes responsible for the Southern Hemisphere Ocean heat uptake and redistribution under anthropogenic warming. *J. Climate*, **33**, 3787–3807, https://doi.org/10.1175/JCLI-D-19-0478.1.

Ma, X., K. von Salzen, and J. Cole, 2010: Constraints on interactions between aerosols and clouds on a global scale from a combination of MODIS-CERES satellite data and climate simulations. *Atmos. Chem. Phys.*, **10**, 9851–9861, https://doi.org/10.5194/acp-10-9851-2010.

Maher, N., and Coauthors, 2019: The Max Planck Institute Grand Ensemble: Enabling the exploration of climate system variability. *J. Adv. Model. Earth Syst.*, **11**, 2050–2069, https://doi.org/10.1029/2019MS001639.

Marshall, D. P., and L. Zanna, 2014: A conceptual model of ocean heat uptake under climate change. *J. Climate*, **27**, 8444–8465, https://doi.org/10.1175/JCLI-D-13-00344.1.

Marshall, J., and K. Speer, 2012: Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nat. Geosci.*, **5**, 171–180, https://doi.org/10.1038/ngeo1391.

Meehl, G. A., and Coauthors, 2014: Decadal climate prediction: An update from the trenches. *Bull. Amer. Meteor. Soc.*, **95**, 243–267, https://doi.org/10.1175/BAMS-D-12-00241.1.

——, J. M. Arblaster, C. M. Bitz, C. T. Chung, and H. Teng, 2016: Antarctic sea-ice expansion between 2000 and 2014 driven by tropical Pacific decadal climate variability. *Nat. Geosci.*, **9**, 590–595, https://doi.org/10.1038/ngeo2751.

Mochizuki, T., and Coauthors, 2012: Decadal prediction using a recent series of MIROC global climate models. *J. Meteor. Soc. Japan*, **90A**, 373–383, https://doi.org/10.2151/jmsj.2012-A22.

Morrison, A. K., O. A. Saenko, A. M. Hogg, and P. Spence, 2013: The role of vertical eddy flux in Southern Ocean heat uptake. *Geophys. Res. Lett.*, **40**, 5445–5450, https://doi.org/10.1002/2013GL057706.

Palmer, M. D., and Coauthors, 2017: Ocean heat content variability and change in an ensemble of ocean reanalyses. *Climate Dyn.*, **49**, 909–930, https://doi.org/10.1007/s00382-015-2801-0.

Palter, J. B., S. M. Griffies, B. L. Samuels, E. D. Galbraith, A. Gnanadesikan, and A. Klocker, 2014: The deep ocean buoyancy budget and its temporal variability. *J. Climate*, **27**, 551–573, https://doi.org/10.1175/JCLI-D-13-00016.1.

Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate*. American Institute of Physics, 520 pp.

Pohlmann, H., J. H. Jungclaus, A. Köhl, D. Stammer, and J. Marotzke, 2009: Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic. *J. Climate*, **22**, 3926–3938, https://doi.org/10.1175/2009JCLI2535.1.

Rintoul, S. R., 1991: South Atlantic interbasin exchange. *J. Geophys. Res.*, **96**, 2675–2692, https://doi.org/10.1029/90JC02422.

Rodwell, M. J., D. P. Rowell, and C. K. Folland, 1999: Oceanic forcing of the wintertime North Atlantic oscillation and European climate. *Nature*, **398**, 320–323, https://doi.org/10.1038/18648.

Santer, B. D., U. Mikolajewicz, W. Brüggemann, U. Cubasch, K. Hasselmann, H. Höck, E. Maier-Reimer, and T. M. L. Wigley, 1995: Ocean variability and its influence on the detectability of greenhouse warming signals. *J. Geophys. Res.*, **100**, 10 693–10 725, https://doi.org/10.1029/95JC00683.

Screen, J. A., C. Deser, I. Simmonds, and R. Tomas, 2014: Atmospheric impacts of Arctic sea-ice loss, 1979–2009: Separating forced change from atmospheric internal variability. *Climate Dyn.*, **43**, 333–344, https://doi.org/10.1007/s00382-013-1830-9.

Sen Gupta, A., L. C. Muir, J. N. Brown, S. J. Phipps, P. J. Durack, D. Monselesan, and S. E. Wijffels, 2012: Climate drift in the CMIP3 models. *J. Climate*, **25**, 4621–4640, https://doi.org/10.1175/JCLI-D-11-00312.1.

——, N. C. Jourdain, J. N. Brown, and D. Monselesan, 2013: Climate drift in the CMIP5 models. *J. Climate*, **26**, 8597–8615, https://doi.org/10.1175/JCLI-D-12-00521.1.

Shi, J.-R., S.-P. Xie, and L. D. Talley, 2018: Evolving relative importance of the Southern Ocean and North Atlantic in anthropogenic ocean heat uptake. *J. Climate*, **31**, 7459–7479, https://doi.org/10.1175/JCLI-D-18-0170.1.

Simpson, I. R., S. G. Yeager, K. A. McKinnon, and C. Deser, 2019: Decadal predictability of late winter precipitation in Western Europe through an ocean–jet stream connection. *Nat. Geosci.*, **12**, 613–619, https://doi.org/10.1038/s41561-019-0391-x.

Singh, H. A., L. M. Polvani, and P. J. Rasch, 2019: Antarctic sea ice expansion, driven by internal variability, in the presence of increasing atmospheric CO_2. *Geophys. Res. Lett.*, **46**, 14 762–14 771, https://doi.org/10.1029/2019GL083758.

Smith, D. M., R. Eade, and H. Pohlmann, 2013: A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. *Climate Dyn.*, **41**, 3325–3338, https://doi.org/10.1007/s00382-013-1683-2.

Stainforth, D. A., M. R. Allen, E. R. Tredger, and L. A. Smith, 2007: Confidence, uncertainty and decision-support relevance in climate predictions. *Philos. Trans. Roy. Soc.*, **365A**, 2145–2161, https://doi.org/10.1098/rsta.2007.2074.

Sutton, R., and P.-P. Mathieu, 2002: Response of the atmosphere–ocean mixed-layer system to anomalous ocean heat-flux convergence. *Quart. J. Roy. Meteor. Soc.*, **128**, 1259–1275, https://doi.org/10.1256/003590002320373283.

Swart, N. C., J. C. Fyfe, E. Hawkins, J. E. Kay, and A. Jahn, 2015: Influence of internal variability on Arctic sea-ice trends. *Nat. Climate Change*, **5**, 86–89, https://doi.org/10.1038/nclimate2483.

Talley, L. D., 2013: Closure of the global overturning circulation through the Indian, Pacific, and Southern Oceans: Schematics and transports. *Oceanography*, **26**, 80–97, https://doi.org/10.5670/oceanog.2013.07.

Tamsitt, V., and Coauthors, 2017: Spiraling pathways of global deep waters to the surface of the Southern Ocean. *Nat. Commun.*, **8**, 172, https://doi.org/10.1038/s41467-017-00197-0.

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

Thomas, J. L., D. W. Waugh, and A. Gnanadesikan, 2015: Southern Hemisphere extratropical circulation: Recent trends and natural variability. *Geophys. Res. Lett.*, **42**, 5508–5515, https://doi.org/10.1002/2015GL064521.

von Salzen, K., and Coauthors, 2013: The Canadian Fourth Generation Atmospheric Global Climate Model (CanAM4). Part I: Representation of physical processes. *Atmos.–Ocean*, **51**, 104–125, https://doi.org/10.1080/07055900.2012.755610.

Von Storch, H., and F. W. Zwiers, 2002: *Statistical Analysis in Climate Research*. Cambridge University Press, 496 pp.

Wagner, T. J. W., and I. Eisenman, 2015: How climate model complexity influences sea ice stability. *J. Climate*, **28**, 3998–4014, https://doi.org/10.1175/JCLI-D-14-00654.1.

Wettstein, J. J., and C. Deser, 2014: Internal variability in projections of twenty-first-century Arctic sea ice loss: Role of the large-scale atmospheric circulation. *J. Climate*, **27**, 527–550, https://doi.org/10.1175/JCLI-D-12-00839.1.

Yang, D., and O. A. Saenko, 2012: Ocean heat transport and its projected change in CanESM2. *J. Climate*, **25**, 8148–8163, https://doi.org/10.1175/JCLI-D-11-00715.1.

Yeager, S. G., and J. I. Robson, 2017: Recent progress in understanding and predicting Atlantic decadal climate variability. *Curr. Climate Change Rep.*, **3**, 112–127, https://doi.org/10.1007/s40641-017-0064-z.

——, and Coauthors, 2018: Predicting near-term changes in the Earth system: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model. *Bull. Amer. Meteor. Soc.*, **99**, 1867–1886, https://doi.org/10.1175/BAMS-D-17-0098.1.

Yim, B. Y., M. Kwon, H. S. Min, and J.-S. Kug, 2015: Pacific decadal oscillation and its relation to the extratropical atmospheric variation in CMIP5. *Climate Dyn.*, **44**, 1521–1540, https://doi.org/10.1007/s00382-014-2349-4.

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

Zhang, L., T. L. Delworth, W. Cooke, and X. Yang, 2019: Natural variability of Southern Ocean convection as a driver of observed climate trends. *Nat. Climate Change*, **9**, 59–65, https://doi.org/10.1038/s41558-018-0350-3.