Geophysical Research Letters

RESEARCH LETTER
10.1029/2020GL091220

Special Section:
Community Earth System Model version 2 (CESM2) Special Collection

Assessment of Equilibrium Climate Sensitivity of the Community Earth System Model Version 2 Through Simulation of the Last Glacial Maximum

Jiang Zhu[1], Bette L. Otto-Bliesner[1], Esther C. Brady[1], Christopher J. Poulsen[2], Jessica E. Tierney[3], Marcus Lofverstrom[3], and Pedro DiNezio[4]

[1]Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [2]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA, [3]Department of Geosciences, the University of Arizona, Tucson, AZ, USA, [4]Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder, CO, USA

Key Points:

- Community Earth System Model version 2 simulates an Last Glacial Maximum (LGM) global temperature at least 5°C colder than the proxy estimate, indicating its equilibrium climate sensitivity is too high
- The large LGM cooling is caused by a strong shortwave cloud feedback in the new atmosphere model
- The shortwave cloud feedback in LGM simulations is connected to that in abrupt 4×CO_2 simulations from the present-day climate

Supporting Information:

- Supporting Information S1

Correspondence to:
J. Zhu,
jiangzhu@ucar.edu

Citation:
Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Poulsen, C. J., Tierney, J. E., Lofverstrom, M., & DiNezio, P. (2021). Assessment of equilibrium climate sensitivity of the Community Earth System Model version 2 through simulation of the Last Glacial Maximum. *Geophysical Research Letters*, *48*, e2020GL091220. https://doi.org/10.1029/2020GL091220

Received 9 OCT 2020
Accepted 24 DEC 2020

© 2021. The Authors.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Abstract The upper end of the equilibrium climate sensitivity (ECS) has increased substantially in the latest Coupled Model Intercomparison Projects phase 6 with eight models (as of this writing) reporting an ECS > 5°C. The Community Earth System Model version 2 (CESM2) is one such high-ECS model. Here we perform paleoclimate simulations of the Last Glacial Maximum (LGM) using CESM2 to examine whether its high ECS is realistic. We find that the simulated LGM global mean temperature decrease exceeds 11°C, greater than both the cooling estimated from proxies and simulated by an earlier model version (CESM1). The large LGM cooling in CESM2 is attributed to a strong shortwave cloud feedback in the newest atmosphere model. Our results indicate that the high ECS of CESM2 is incompatible with LGM constraints and that the projected future warming in CESM2, and models with a similarly high ECS, is thus likely too large.

Plain Language Summary Equilibrium climate sensitivity (ECS) is one of the most important metrics in climate science. It measures the amount of global warming over hundreds of years after a doubling of the atmospheric CO_2 concentration. An ECS range of 1.5°C–4.5°C has been consistently supported by climate models over the past 40 years. However, this has changed in the latest generation of climate models with eight (as of this writing) showing an ECS > 5°C. Such a high ECS implies that future warming will be much greater than previously thought for the same amount of greenhouse gas emissions, making it more challenging for human and natural systems to adapt. This study examines whether the ECS in one "high-ECS" model—Community Earth System Model version 2 (CESM2)—is realistic. Our approach is to perform a paleoclimate simulation of the culmination of the last glacial period, which was colder and had a lower atmospheric CO_2 level than today. We find that the magnitude of cooling in the CESM2 simulation is much larger than supported by the observational evidence, indicating the model's ECS is too large. We further find that the high ECS of CESM2 is caused by the treatment of clouds in the model.

1. Introduction

Equilibrium climate sensitivity (ECS) is the response of global mean surface temperature (GMST) to the radiative forcing caused by a doubling of the atmospheric CO_2 concentration. ECS is one of the most important metrics for projecting future climate and for crafting mitigation policies and adaptation plans. Since the 1960s, a tremendous amount of research has been conducted to quantify and understand ECS using climate models (e.g., Manabe & Wetherald, 1967). Despite these efforts, ECS remains highly variable among climate models with a large range of 2.1°C–4.7°C in the Coupled Model Intercomparison Projects, phase 5 (CMIP5) (Flato et al., 2013). In the latest generation of CMIP6 models, ECS ranges from 2.7°C to 5.7°C with more than 16 models having values larger than 4.7°C (Zelinka et al., 2020). Determination of whether these high model-based estimates of ECS are realistic is critically important for understanding the impact of future climate change on human and natural systems.

Past climates provide important "out-of-sample" evaluation of ECS, as climate feedback processes, especially cloud processes, are parameterized and tuned in climate models to reproduce the instrumental

record that has a narrow range of variations in climate forcing and response (Hourdin et al., 2017; Schmidt et al., 2017; Tierney et al., 2020a; Zhu et al., 2020). The Last Glacial Maximum (LGM; 21,000 before present [21 ka]) offers a prime opportunity to inform ECS, as it represents a quasiequilibrium climate state much different from today and has relatively well-known climate forcings and a high spatial coverage of proxy temperatures (CLIMAP Project Members, 1976; Kageyama et al., 2018; Tierney et al., 2020b; Waelbroeck et al., 2009). The LGM has been used to validate climate models and to constrain climate feedback processes since the 1980s (e.g. Manabe & Broccoli, 1985).

In this study, we assess whether the high ECS (>5°C) in CESM2 is supported by LGM constraints. We compare CESM2 simulated LGM cooling with proxy reconstructions and a simulation using its predecessor, the CESM1 (also referred as CESM1(CAM5)). We then use a forcing-feedback analysis to diagnose the origin of the differences between the CESM2 and CESM1 simulations.

2. Models and Experiments

CESM2 is the latest and most comprehensive Earth system model in the CESM series and is participating in both the CMIP6 and the Paleoclimate Modeling Intercomparison Project phase 4 (PMIP4) (Bacmeister et al., 2020; Danabasoglu et al., 2020; Feng et al., 2020; Gettelman et al., 2019; Meehl et al., 2020; Otto-Bliesner et al., 2020). CESM2 consists of component models of the atmosphere, ocean, land, sea-ice, and rivers. Substantial science and infrastructure improvements have been made from CESM1 to CESM2 (see Danabasoglu et al. (2020) for details), including updates to the cloud-related parameterizations in the Community Atmosphere Model version 6 (CAM6). These include a new higher-order turbulence closure scheme with a unified description of processes in the cloudy turbulent layers, changes to the two-moment stratiform microphysics scheme including a new capability to predict the mass and number concentrations of rain and snow, modifications of the mixed phase ice nucleation scheme, and many others (Gettelman et al., 2019). These changes in cloud parameterizations from CAM5 to CAM6 are the primary reasons for the increased ECS in CESM2 from CESM1 (~5.3°C vs. ~4.0°C) (Gettelman et al., 2012, 2019).

The CESM2 LGM simulation in this study was designed to follow, as closely as possible, the experimental configurations for the "low-top" (also referred as CESM2(CAM6)) simulations of preindustrial and future climates for the CMIP6 (Bacmeister et al., 2020; Danabasoglu et al., 2020). Similar to the CESM2 CMIP6 simulations, the LGM simulation has a horizontal resolution of 0.9° × 1.25° (latitude × longitude) in the atmosphere and land models, and a nominal 1° horizontal resolution in the ocean and sea ice models. Unlike the CESM2 CMIP6 simulations, the LGM was run without active biogeochemical modules in the land or ocean due to the lack of boundary and initial conditions for the LGM vegetation and marine biogeochemical processes. Similarly, we used an older version of the river transport module due to the insufficient knowledge of the LGM river hydrography that is required by the new river model. To examine the impact of this configuration, we performed a new preindustrial simulation (PI) with these customizations and found that the mean climate state and ECS changed little from the standard CESM2 CMIP6 simulations (see Text S1 and Table S1).

The LGM boundary conditions were implemented following the PMIP4 protocol (Kageyama et al., 2017). Specifically, greenhouse gas (GHG) concentrations were 190 ppmv, 375 ppb, and 200 ppb for CO_2, CH_4, and N_2O, respectively. Earth orbital parameters were fixed at 21 ka values. Land ice sheets (LISs) were prescribed using the ICE-6G reconstruction at 21 ka (Peltier et al., 2015) and comprised of changes in land surface properties (e.g., albedo), surface elevation, and land-sea distribution due to the presence of LISs and the associated effect of a lowered sea level. Surface elevation changes over the LIS and shelf-exposure regions included a resolved component of the grid-box mean elevation and an "unresolved" subgrid-scale component that was used for surface drag parameterizations (Lauritzen et al., 2015; see also Text S1). The LGM simulation used preindustrial aerosol emissions and vegetation cover, as well as a present-day vegetation phenology from satellite observations (Text S1).

Ocean temperature and salinity of the CESM2 LGM simulation were branched from an equilibrated LGM simulation using CESM1 (Table 1; DiNezio et al., 2016; Zhu & Poulsen, 2020a; Zhu et al., 2017), which in turn was initialized from a CCSM4 LGM simulation (Brady et al., 2013). The CESM2 simulation was

Table 1
List of CESM2 and CESM1 Experiments, Initial Condition, Simulation Length (years), the Effective Radiative Forcing (F_{eff}; W m^{-2}), Changes in Global Mean Surface Temperature (ΔGMST; °C), TOA Radiation Imbalance (ΔN; W m^{-2}), Effective Climate Feedback Parameter (λ_{eff}; W m^{-2} K^{-1}), and the Shortwave Cloud Feedback (λ_{sw_cld}; W m^{-2} K^{-1}).

	Initial condition	Length	F_{eff}	ΔGMST	ΔN	λ_{eff}	λ_{sw_cld}
CESM2 LGM	CESM1 LGM	500	−5.2	−11.3	−0.20	−0.48	0.74
CESM1 LGM	CCSM4 LGM	1,800	−6.0	−6.8	0.06	−0.88	0.36
CESM2 4×CO_2	CESM2 PI	1,000	9.4	11.0	0.92	−0.77	0.87
CESM1 4×CO_2	CESM1 PI	1,000	8.0	7.7	0.70	−0.95	0.48

All analysis was done for the averages over the last 50 years of each simulation. In addition to the LGM simulations, results from the abrupt 4×CO_2 simulations using CESM2 and CESM1 are also listed. CESM, Community Earth System Model; GMST, global mean surface temperature; LGM, Last Glacial Maximum.

integrated for an additional 500 years. The simulation has not reached equilibrium and has a small cooling trend of 0.07°C per 100 years in GMST and a top-of-atmosphere (TOA) radiation imbalance of approximately −0.2 W m^{-2} over the last 100 years of the simulation. Nonetheless for the purpose of this study, in which we show that the LGM simulation is too cold, the trends do not change our major results and conclusions (see below). To better understand the CESM2 LGM cooling, we performed additional simulations with both atmosphere-only and slab ocean configurations to quantify the LGM effective radiative forcing following the methodology in Zhu and Poulsen (2020a) (Text S2). The effective radiative forcing and feedback in CESM2 LGM simulation are compared with those in CESM1, as well as those in abrupt 4×CO_2 simulations using both models under the present-day climate (Bacmeister et al., 2020; Zhu et al., 2019). To illustrate the role of the new physical parameterizations in CAM6 in changing the climate response to the LGM forcing, we performed a parallel set of PI and LGM simulations using the CAM5 physical parameterizations within CESM2 (Text S3 and Table S1).

To benchmark the LGM simulations, we employed a recent synthesis of LGM sea-surface temperatures (SSTs) inferred from geochemical proxies (Mg/Ca, $U_{37}^{K'}$, and TEX_{86}), which used Bayesian proxy system models to systematically accommodate uncertainties in calibration and various environmental influences (Tierney & Tingley, 2014, 2018; Tierney et al., 2019, 2020b). Foraminiferal $\delta^{18}O$ SSTs relied on unrealistic assumptions regarding the $\delta^{18}O$ composition of surface seawater (Tierney et al., 2020b) and therefore were not used in this study. In addition, we compare the model simulated LGM cooling in GMST and tropical SST to the proxy-only estimates in Tierney et al., 2020b, which are 5.6°C (4.6°C–6.8°C; 95% confidence interval) and 2.6°C (2.3°C–3.0°C, 95% CI; averaged over 30°S–30°N), respectively. We note that these estimates of global and regional LGM cooling agree with several other independent studies (Bereiter et al., 2018; Friedrich & Timmermann, 2020; Holden et al., 2010; Snyder, 2016; Von Deimling et al., 2006) but are greater than some early estimates (Annan & Hargreaves, 2013; Schmittner et al., 2011; Shakun et al., 2012).

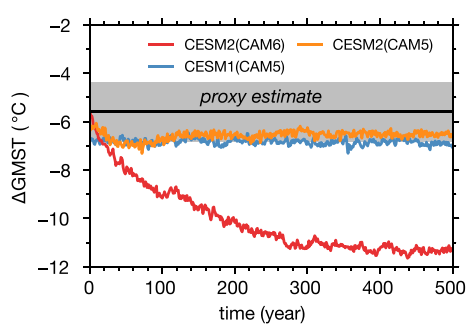

Figure 1. Model-data comparison of the GMST response (ΔGMST) in the LGM simulations. The black line and the gray shading indicate the mean and the 95% confidence intervals respectively for the proxy estimates compiled by Tierney et al. (2020b). GMST, global mean surface temperature; LGM, Last Glacial Maximum.

3. Results

3.1. Excessive LGM Cooling in CESM2

The CESM2 LGM simulation cooled rapidly from a starting ΔGMST (LGM−PI) of approximately −6°C to more than −11°C after 500 model years (Figure 1). CESM2 GMST is at least 5°C colder than the median of proxy estimates and at least 4°C colder than the upper limit of proxy uncertainty. The tropical mean ΔSST in CESM2 is −7°C, also much larger in magnitude than the proxy estimate of −2.6°C. In comparison, CESM1 ΔGMST and tropical mean ΔSST are −6.8°C and −3.6°C, respectively, which are near the upper end of the proxy uncertainty range.

The large discrepancy between CESM2 and proxy temperatures is also evident at individual proxy-SST sites (Figure 2). CESM2 LGM SST cools

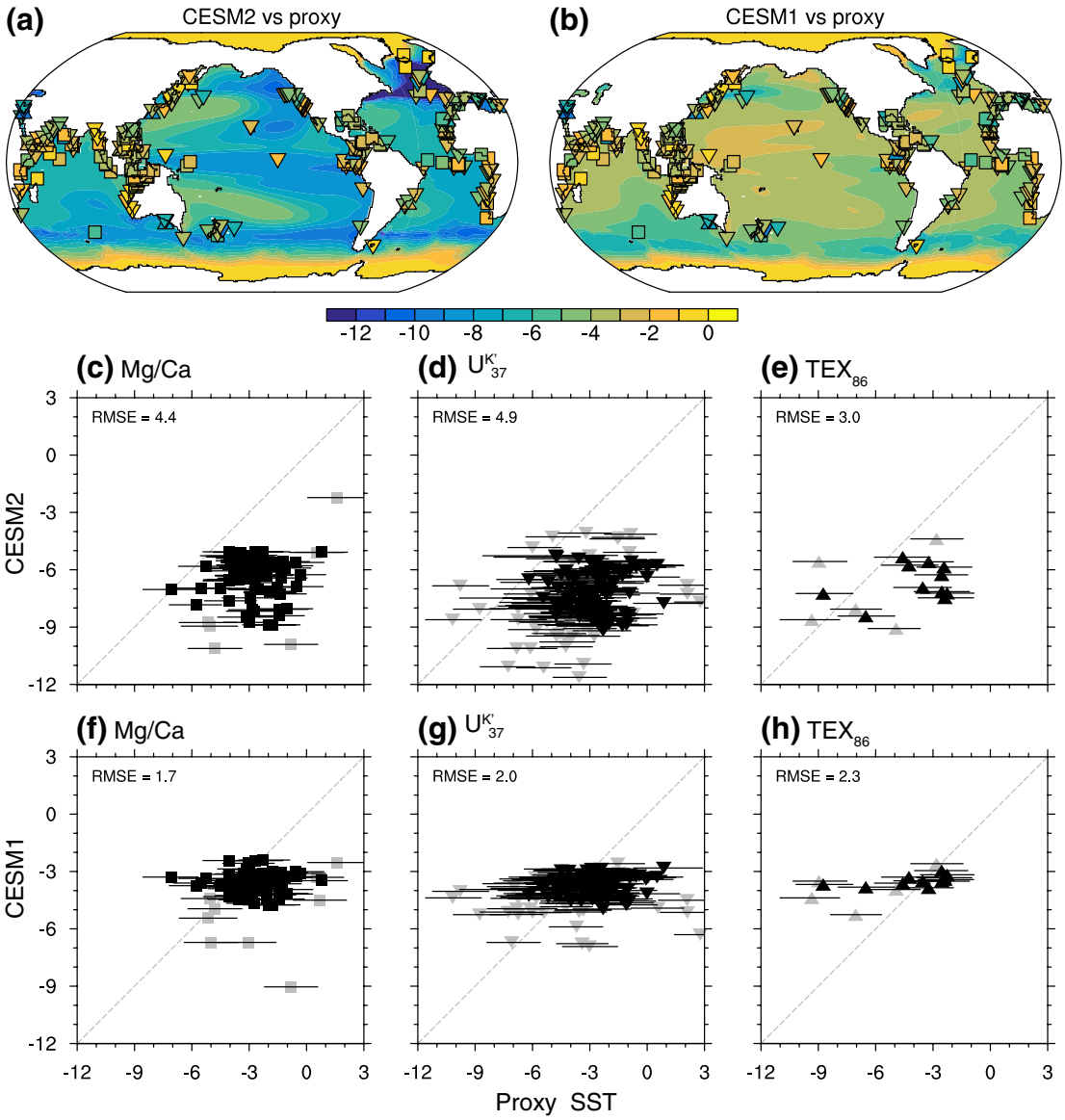

Figure 2. Model-data comparison of the LGM ΔSST (units: °C). (a) Spatial distribution of the CESM2 (shading) and proxy (markers) ΔSST. (b) As (a), but for CESM1. Scatter plots of the CESM2 versus proxy ΔSST at individual proxy sites are shown for (c) Mg/Ca, (d) $U_{37}^{K'}$, and (e) TEX_{86}, respectively. (f)–(h) As (c)–(e), but for the comparison in CESM1. In all plots, Mg/Ca, $U_{37}^{K'}$, and TEX_{86} SSTs are denoted as squares, downward pointing triangles, and upward pointing triangles, respectively. In (c)–(h), root-mean-squared error (RMSE) between model and data is listed; markers are color-coded black for tropical (30°S–30°N) records and gray for extratropical records. CESM, Community Earth System Model; LGM, Last Glacial Maximum; SST, sea-surface temperature.

more than 8.0°C in the eastern equatorial Pacific, eastern subtropical oceans, northern North Atlantic, and the Southern Ocean. The CESM2 ΔSST is greater in magnitude than the median of proxy estimates in 79 of 80 Mg/Ca, 142 of 147 $U_{37}^{K'}$, and 13 of 16 TEX_{86} records. CESM2 ΔSSTs fall within the uncertainty range in less than 10% of the proxy records (24 of 243 records). Root-mean-squared-error (RMSE) is 4.4°C and 4.9°C in Mg/Ca and $U_{37}^{K'}$ records, respectively. In contrast, CESM1 ΔSSTs are much smaller in magnitude and have values that fall within the uncertainty range for approximately 60% of the proxy records. RMSE in CESM1 ΔSST (~1.7°C–2°C) is much smaller than that in CESM2 and only slightly larger than the mean standard error in the proxies (~1.5°C; arithmetic mean of the standard error of all available SST records).

The cooling of the LGM climate in CESM2 compared to CESM1 is primarily attributed to the update of the atmosphere component from CAM5 to CAM6. We demonstrate this with an additional set of LGM and

PI simulations (CESM2(CAM5)), in which the same boundary conditions and model configurations as those in the CESM2 simulations are used with the exception that the CAM5 physical parameterization package, the default option in CESM1, is implemented. In CESM2(CAM5), the LGM ΔGMST decreases from −6.0°C to −6.8°C during the first 90 years, and then increases gradually to approximately −6.5°C after 500 model years (Figure 1). The CESM2(CAM5) ΔGMST is thus smaller in magnitude than that in CESM2 and agrees better with CESM1 (and proxy estimates), suggesting that differences between CAM6 and CAM5 predominantly contribute to the large magnitude of LGM cooling in CESM2. At individual sites, CESM2(CAM5) ΔSST also agrees better with proxy data compared to CESM2 (RMSE = ∼1.5°C; not shown). The smaller magnitude of the LGM ΔGMST in CESM2(CAM5) compared to CESM1 (−6.5°C vs. −6.8°C) is primarily due to lower snow and ice albedos in the new version of the land model (Lawrence et al., 2019).

We note that the above comparison of proxy and CESM2(CAM6) temperatures is for the average of the last 50 years (year 451–500) of the simulation, during which the TOA radiation imbalance is approximately −0.20 W m^{-2} (Table 1) and the model surface temperature is still decreasing. The model-data disagreement would be even more prominent if the CESM2 LGM simulation was extended further.

3.2. Effective Forcing and Feedbacks in CESM2

To understand the LGM cooling in CESM2, we first quantify the effective radiative forcing (F_{eff}) following the methodology in Zhu and Poulsen (2020a). F_{eff} includes the instantaneous radiative forcing and adjustments from the atmosphere and surface, and is more closely related to surface warming in the forcing-feedback framework (Sherwood et al., 2015). LGM F_{eff} comprises contributions from GHG and LIS, which are obtained using both "fixed-SST" simulations in an atmosphere-only configuration and slab ocean simulations (SOM) with thermodynamic atmosphere-ocean coupling but inactive ocean dynamics (Text S2 and Table S2). Simulations with prescribed sea-surface conditions provide changes in TOA radiation (F_{fsst}) and land surface temperature (ΔT_{lnd}) in response to a forcing agent (GHG or LIS); SOM simulations provide estimates of the climate sensitivity parameter associated with the forcing. The effective radiative forcing of GHG or LIS is estimated by correcting the effects of ΔT_{lnd} on F_{fsst} using the climate sensitivity parameter (Text S2).

The LGM effective radiative forcing is −5.2 W m^{-2} (Table 1) with contributions of −3.1 and −2.1 W m^{-2} from GHG and LIS, respectively (Table S2). The magnitude of F_{eff} in CESM2 is smaller than in CESM1 (−6.0 W m^{-2}) due primarily to the lower albedo of snow and ice in the new land model (Lawrence et al., 2019). A weaker LGM forcing in CESM2 than in CESM1 implies that climate feedbacks in CESM2 must be stronger in order to explain the larger temperature response.

Accordingly, the effective climate feedback parameter (λ_{eff}) in CESM2 is −0.48 W m^{-2} K^{-1} and larger than that in CESM1 by 0.40 W m^{-2} K^{-1} (Table 1). Here λ_{eff} is calculated as

$$\lambda_{\mathrm{eff}} = \left(\Delta N - F_{\mathrm{eff}}\right) / \Delta\mathrm{GMST}$$

where ΔN and ΔGMST are the final 50-year mean TOA radiation imbalance and LGM global cooling in the coupled simulation (Table 1). A less negative λ_{eff} in CESM2 indicates that the net climate feedback is less efficient at damping the surface cooling; that is, CESM2 is more sensitive to the LGM forcing than CESM1. We note that the LGM simulations are not initialized from the unperturbed preindustrial state and that ice sheets have nonradiative pathways to change surface temperature (Zhu & Poulsen, 2020a), which prevent us from implementing the traditional "Gregory method" (Gregory et al., 2004) to estimate the effective radiative forcing and feedback.

We attribute the larger climate feedback in CESM2 to a stronger shortwave cloud feedback than in CESM1. This is consistent with previous studies that show that the shortwave cloud feedback is mostly responsible for the ECS increase from CESM1 to CESM2 and that other climate feedbacks (such as water vapor and albedo feedbacks) depend less on atmospheric physical parameterizations and stay largely unchanged between model versions (Gettelman et al., 2019; Zhu & Poulsen, 2020b). To illustrate this for the LGM,

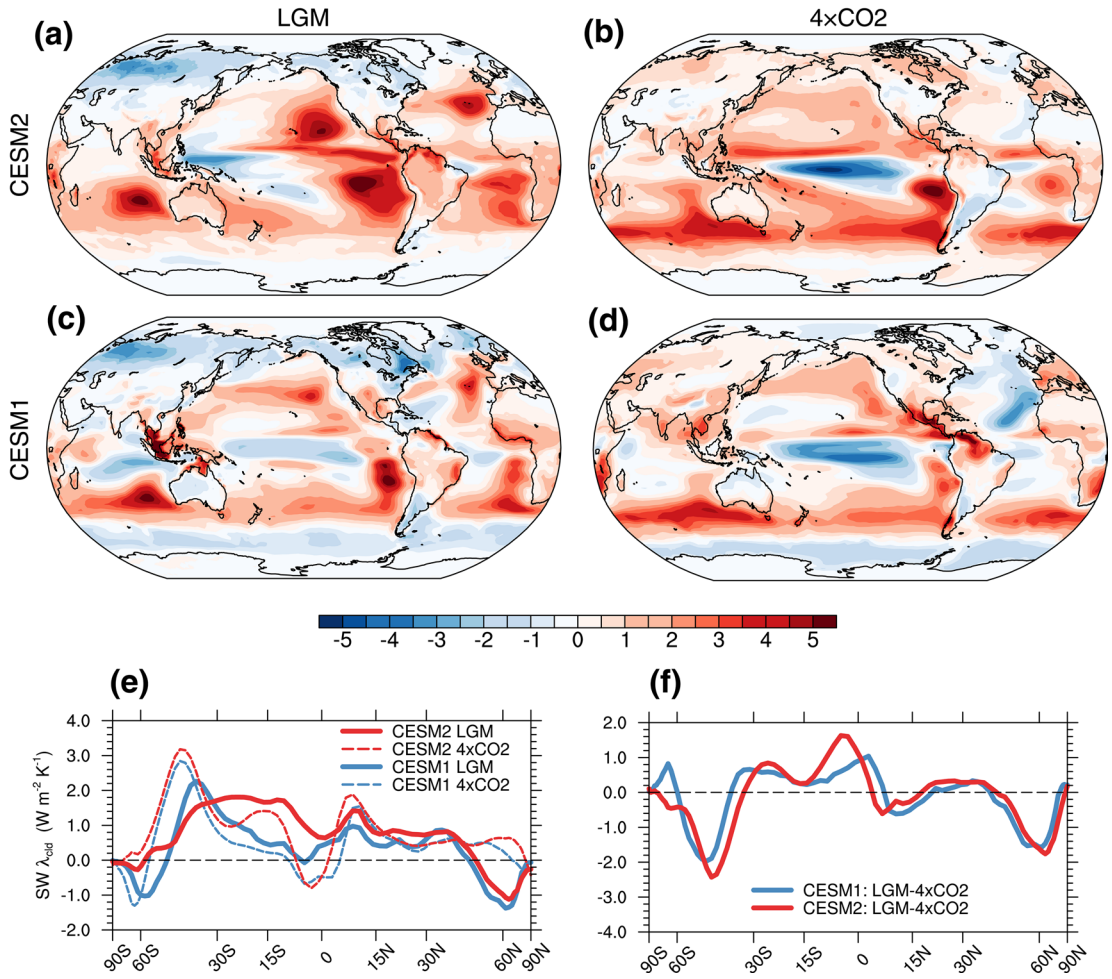

Figure 3. The shortwave cloud feedback (units: W m^{-2} K^{-1}) in (a) CESM2 and (c) CESM1 LGM simulations. (b) and (d) As (a) and (c), but for the abrupt 4×CO_2 simulations. (e) The zonal mean shortwave cloud feedback in each simulation. (f) Difference in the zonal mean shortwave cloud feedback between LGM and 4×CO_2 simulations in CESM2 and CESM1.

we calculate the shortwave cloud feedback parameter (λ_{sw_cld}) in the LGM simulations using the approximate partial radiative perturbation method (Taylor et al., 2007). Global mean λ_{sw_cld} is 0.74 W m^{-2} K^{-1} in CESM2, which is 0.38 W m^{-2} K^{-1} larger than that in CESM1 (Table 1 and Figure 3). This result suggests that the stronger shortwave cloud feedback explains the vast majority of the larger effective climate feedback (0.40 W m^{-2} K^{-1}) in CESM2 and therefore the greater LGM cooling.

In response to the LGM forcing, CESM2 λ_{sw_cld} exhibits a similar spatial pattern as CESM1 but with more positive values at almost all latitudes (Figure 3). Both CESM2 and CESM1 show maximum cloud feedbacks over the stratocumulus decks in the eastern subtropical ocean basins with larger values and spatial extent in CESM2. At ~15°S, the zonal mean λ_{sw_cld} in CESM2 is ~2 W m^{-2} K^{-1} larger than in CESM1 (Figure 3e). In the Southern Ocean (~60°S), both models show minimum λ_{sw_cld} with values close to zero in CESM2 and −1 W m^{-2} K^{-1} in CESM1. Over the Laurentide and Eurasian ice sheets, both models have negative λ_{sw_cld} of similar magnitude, which primarily reflects the presence of LGM ice sheets that occupy the space containing air masses and clouds in the PI simulation (Zhu & Poulsen, 2020a). In mid-latitudes (~30°N/S–40°N/S), λ_{sw_cld} in CESM2 is smaller than in CESM1, likely reflecting differences in the extent of equatorward sea-ice expansion and shifts of storm tracks. An equatorward shift of storm tracks associated with global cooling moves cloud systems toward lower latitudes with higher insolation, which causes additional cooling (a positive feedback). A stronger global cooling in CESM2 (ΔGMST of −11.3°C vs. −6.8°C in CESM1) leads to

a greater equatorward shift of storm tracks and a smaller λ_{sw_cld} in the Southern Hemisphere mid-latitudes (~40°S–30°S) than CESM1, as well as a larger λ_{sw_cld} to the north.

4. Discussion: The Connection to 4×CO_2 Simulations

The λ_{sw_cld} in LGM simulations is closely connected to that in abrupt 4×CO_2 simulations, which largely determines intermodel differences in ECS under the present-day climate (Bacmeister et al., 2020; Gettelman et al., 2019). Compared to CESM1, CESM2 global mean λ_{sw_cld} is larger by ~0.4 W m^{-2} K^{-1} in both the LGM and 4×CO_2 simulations (Table 1). For both simulations, the intermodel difference in λ_{sw_cld} is most prominent over the Southern Ocean and the Southern Hemisphere subtropics (red vs. blue lines in Figure 3e), likely reflecting a Southern Ocean origin of the different model behavior (Bacmeister et al., 2020) that impacts the lower latitudes through ocean and atmosphere processes (Zhu & Poulsen, 2020a). These results suggest that the cloud processes that give rise to the higher ECS in CESM2 (in the 4×CO_2 simulation) act in a similar way to produce the larger LGM cooling.

Comparing LGM and 4×CO_2 simulations, both CESM2 and CESM1 exhibit broadly consistent regional differences in λ_{sw_cld} (Figure 3f). Over the northern subtropical Pacific and Atlantic, λ_{sw_cld} is larger in LGM than 4×CO_2 simulations (Figures 3a–3d), which is a characteristic feature associated with the southward shift of storm tracks (Lofverstrom, 2020; Lofverstrom et al., 2016) and cloud systems (Zhu & Poulsen, 2020a) due to LGM ice-sheet forcing. In the southern subtropical oceans, λ_{sw_cld} is also stronger in LGM simulations, likely due to a strengthening of the tropical easterlies in a cold climate (e.g. Williams & Bryan, 2006). A stronger surface wind increases the boundary layer turbulence and the associated latent heat flux, leading to more low clouds with surface cooling and therefore a stronger λ_{sw_cld} than simulations with a warmer climate and weaker winds (Bretherton et al., 2013). This wind-driven mechanism could be further amplified by the SST pattern effect through changing the lower tropospheric stability and low clouds (Armour et al., 2013; Rose et al., 2014; Zhou et al., 2016). Over the middle to high latitudes, λ_{sw_cld} is weaker in LGM than 4×CO_2 simulations, likely due to an increase of cloud ice fraction with the background cooling and the more negative feedback associated with the phase transitioning between cloud ice and liquid (Frey & Kay, 2018; Mitchell et al., 1989; Tan et al., 2016; Zhu & Poulsen, 2020b).

The close connection in global and regional λ_{sw_cld} between LGM and 4×CO_2 simulations support the use of LGM constraints to inform cloud feedbacks and ECS. The LGM constraints, therefore, suggest that some of the new cloud-related parameterizations, including schemes of the stratiform microphysics, shallow convection and moist turbulence, and the mixed phase ice nucleation (Gettelman et al., 2019), likely produces unrealistic response to large variations in external forcings. Examining each of these new schemes in the LGM context is an ongoing study and beyond the scope of this study.

5. Conclusions

In this study, we performed a CESM2 simulation of the LGM and found that the simulated global cooling exceeds −11°C, which is at least 5°C colder than the median of the latest proxy-based estimates (Tierney et al., 2020b). LGM tropical mean ΔSST is −7°C in CESM2, which is also much colder than the proxy data estimate (−2°C – −3°C). At individual proxy SST sites, CESM2 LGM is on average 4°C–5°C colder than the proxy. In comparison, global and regional cooling in CESM1 (predecessor to CESM2) is much smaller and agrees better with proxies. The larger LGM cooling in CESM2 is attributed to the new physical parameterizations in the atmosphere model (CAM6), which simulates a stronger shortwave cloud feedback than its predecessor (CAM5 in CESM1).

The discrepancy between CESM2 simulated LGM cooling and proxy data cannot be explained by uncertainties in either LGM temperature reconstructions or climate forcings. If earlier proxy syntheses are used for the comparison (CLIMAP Project Members, 1976; Waelbroeck et al., 2009), the model-data discrepancy would be more prominent, as these earlier studies suggested more modest LGM cooling (Tierney et al., 2020b). If realistic distributions of dust aerosols and vegetation cover were available and implement-

ed, the CESM2 simulated LGM would be colder and even more difficult to reconcile with the proxy data (Köhler et al., 2010; Sherwood et al., 2020; Tierney et al., 2020b).

Our study suggests that the ECS in CESM2 is too high and closely related to the large model-data discrepancy in the LGM cooling. This result indicates either that the shortwave cloud feedback in CESM2 is too strong or that current climate models are missing important physical processes that counter the shortwave cloud feedback and stabilize the climate. Whichever is true, the projected future warming using high-ECS models, such as CESM2, is likely overestimated. This conclusion from constraints using a past cold climate is the same as our previous studies using past extreme warm climate of the Early Eocene (Zhu et al., 2019, 2020). Our findings are also consistent with independent studies, which show that the warming trends in high-ECS models including CESM2 are too large than that in the instrumental record (Brunner et al., 2020; Nijsse et al., 2020; Tokarska et al., 2020). Here, we emphasize the unique strength of the paleoclimate constraint as a true "out-of-sample" test and its advantage of being independent on the transient climate change and internal variability.

Data Availability Statement

Computing and data storage resources, including the Cheyenne supercomputer (doi:10.5065/D6RX99HX), were provided by the Computational and Information Systems Laboratory (CISL) at NCAR. Relevant CESM simulation data and results on the cloud feedback analysis are available in the Zenodo repository (https://doi.org/10.5281/zenodo.4075597).

Acknowledgments
The authors thank the Editor, Hui Su, and Deepak Chandan and an anonymous reviewer for their constructive comments. This work was supported by Heising-Simons Foundation grant #2016-15 to C. Poulsen and J. Tierney, Heising-Simons Foundation grant #2016-12 and National Science Foundation grant 2002397 to C. Poulsen, and National Science Foundation grant AGS-1602301 to J. Tierney. CESM model code is available through the National Center for Atmospheric Research GitHub repository (https://github.com/ESCOMP/CESM/releases/tag/cesm2.1.1). The CESM project is supported primarily by the National Science Foundation (NSF). This material is based upon work supported by the National Center for Atmospheric Research, which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977. The authors thank M. Zelinka for making available the updated forcing-feedback analysis of CMIP6 models.

References

Annan, J. D., & Hargreaves, J. C. (2013). A new global reconstruction of temperature changes at the Last Glacial Maximum. *Climate of the Past*, *9*(1), 367–376. https://doi.org/10.5194/cp-9-367-2013

Armour, K. C., Bitz, C. M., & Roe, G. H. (2013). Time-Varying climate sensitivity from regional feedbacks. *Journal of Climate*, *26*(13), 4518–4534. https://doi.org/10.1175/jcli-d-12-00544.1

Bacmeister, J. T., Hannay, C., Medeiros, B., Gettelman, A., Neale, R., Fredriksen, H. B., et al. (2020). CO2 increase experiments using the CESM: Relationship to climate sensitivity and comparison of CESM1 to CESM2. *Journal of Advances in Modeling Earth Systems*, *12*(11), e2020MS002120. https://doi.org/10.1029/2020MS002120

Bereiter, B., Shackleton, S., Baggenstos, D., Kawamura, K., & Severinghaus, J. (2018). Mean global ocean temperatures during the last glacial transition. *Nature*, *553*(7686), 39–44. https://doi.org/10.1038/nature25152

Brady, E. C., Otto-Bliesner, B. L., Kay, J. E., & Rosenbloom, N. (2013). Sensitivity to glacial forcing in the CCSM4. *Journal of Climate*, *26*(6), 1901–1925. https://doi.org/10.1175/JCLI-D-11-00416.1

Bretherton, C. S., Blossey, P. N., & Jones, C. R. (2013). Mechanisms of marine low cloud sensitivity to idealized climate perturbations: A single-LES exploration extending the CGILS cases. *Journal of Advances in Modeling Earth Systems*, *5*(2), 316–337. https://doi.org/10.1002/jame.20019

Brunner, L., Pendergrass, A. G., Lehner, F., Merrifield, A. L., Lorenz, R., & Knutti, R. (2020). Reduced global warming from CMIP6 projections when weighting models by performance and independence. *Earth System Dynamics*, *11*(4), 995–1012. https://doi.org/10.5194/esd-11-995-2020

CLIMAP Project Members. (1976). The Surface of the Ice-Age Earth. *Science*, *191*(4232), 1131–1137. https://doi.org/10.1126/science.191.4232.1131

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019ms001916

DiNezio, P. N., Timmermann, A., Tierney, J. E., Jin, F. F., Otto-Bliesner, B., Rosenbloom, N., et al. (2016). The climate response of the Indo-Pacific warm pool to glacial sea level. *Paleoceanography*, *31*(6), 866–894. https://doi.org/10.1002/2015PA002890

Feng, R., Otto-Bliesner, B. L., Brady, E. C., & Rosenbloom, N. (2020). Increased climate response and Earth system sensitivity from CCSM4 to CESM2 in Mid-Pliocene simulations. *Journal of Advances in Modeling Earth Systems*, *12*(8), e2019MS002033. https://doi.org/10.1029/2019ms002033

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., et al. (2013). Evaluation of climate models. In T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change* (pp. 741–866). Cambridge: Cambridge University Press.

Frey, W. R., & Kay, J. E. (2018). The influence of extratropical cloud phase and amount feedbacks on climate sensitivity. *Climate Dynamics*, *50*(7), 3097–3116. https://doi.org/10.1007/s00382-017-3796-5

Friedrich, T., & Timmermann, A. (2020). Using Late Pleistocene sea surface temperature reconstructions to constrain future greenhouse warming. *Earth and Planetary Science Letters*, *530*, 115911. https://doi.org/10.1016/j.epsl.2019.115911

Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., et al. (2019). High climate sensitivity in the Community Earth System Model Version 2 (CESM2). *Geophysical Research Letters*, *46*, 8329–8337. https://doi.org/10.1029/2019GL083978

Gettelman, A., Kay, J. E., & Shell, K. M. (2012). The Evolution of climate sensitivity and climate feedbacks in the community atmosphere model. *Journal of Climate*, *25*(5), 1453–1469. https://doi.org/10.1175/JCLI-D-11-00197.1

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., et al. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*(3), L03205. https://doi.org/10.1029/2003GL018747

Holden, P. B., Edwards, N. R., Oliver, K. I. C., Lenton, T. M., & Wilkinson, R. D. (2010). A probabilistic calibration of climate sensitivity and terrestrial carbon change in GENIE-1. *Climate Dynamics*, *35*(5), 785–806. https://doi.org/10.1007/s00382-009-0630-8

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., et al. (2017). The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, *98*(3), 589–602. https://doi.org/10.1175/bams-d-15-00135.1

Kageyama, M., Albani, S., Braconnot, P., Harrison, S. P., Hopcroft, P. O., Ivanovic, R. F., et al. (2017). The PMIP4 contribution to CMIP6 – Part 4: Scientific objectives and experimental design of the PMIP4-CMIP6 Last Glacial Maximum experiments and PMIP4 sensitivity experiments. *Geoscientific Model Development*, *10*(11), 4035–4055. https://doi.org/10.5194/gmd-10-4035-2017

Kageyama, M., Braconnot, P., Harrison, S. P., Haywood, A. M., Jungclaus, J. H., Otto-Bliesner, B. L., et al. (2018). The PMIP4 contribution to CMIP6 – Part 1: Overview and over-arching analysis plan. *Geoscientific Model Development*, *11*(3), 1033–1057. https://doi.org/10.5194/gmd-11-1033-2018

Köhler, P., Bintanja, R., Fischer, H., Joos, F., Knutti, R., Lohmann, G., et al. (2010). What caused Earth's temperature variations during the last 800,000 years? Data-based evidence on radiative forcing and constraints on climate sensitivity. *Quaternary Science Reviews*, *29*(1), 129–145. https://doi.org/10.1016/j.quascirev.2009.09.026

Lauritzen, P. H., Bacmeister, J. T., Callaghan, P. F., & Taylor, M. A. (2015). NCAR_Topo (v1.0): NCAR global model topography generation software for unstructured grids. *Geoscientific Model Development*, *8*(12), 3975–3986. https://doi.org/10.5194/gmd-8-3975-2015

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., et al. (2019). The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4245–4287. https://doi.org/10.1029/2018ms001583

Lofverstrom, M. (2020). A dynamic link between high-intensity precipitation events in southwestern North America and Europe at the Last Glacial Maximum. *Earth and Planetary Science Letters*, *534*, 116081. https://doi.org/10.1016/j.epsl.2020.116081

Lofverstrom, M., Caballero, R., Nilsson, J., & Messori, G. (2016). Stationary wave reflection as a mechanism for zonalizing the Atlantic Winter Jet at the LGM. *Journal of the Atmospheric Sciences*, *73*(8), 3329–3342. https://doi.org/10.1175/jas-d-15-0295.1

Manabe, S., & Broccoli, A. J. (1985). A comparison of climate model sensitivity with data from the Last Glacial Maximum. *Journal of the Atmospheric Sciences*, *42*(23), 2643–2651. https://doi.org/10.1175/1520-0469(1985)042<2643:ACOCMS>2.0.CO;2

Manabe, S., & Wetherald, R. T. (1967). Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *Journal of the Atmospheric Sciences*, *24*(3), 241–259. https://doi.org/10.1175/1520-0469(1967)024%3C0241:TEOTAW%3E2.0.CO

Meehl, G. A., Arblaster, J. M., Bates, S., Richter, J. H., Tebaldi, C., Gettelman, A., et al. (2020). Characteristics of future warmer base states in CESM2. *Earth and Space Science*, *7*(9), e2020EA001296. https://doi.org/10.1029/2020ea001296

Mitchell, J. F. B., Senior, C. A., & Ingram, W. J. (1989). CO2 and climate: A missing feedback? *Nature*, *341*(6238), 132–134. https://doi.org/10.1038/341132a0

Nijsse, F. J. M. M., Cox, P. M., & Williamson, M. S. (2020). Emergent constraints on transient climate response (TCR) and equilibrium climate sensitivity (ECS) from historical warming in CMIP5 and CMIP6 models. *Earth System Dynamics*, *11*(3), 737–750. https://doi.org/10.5194/esd-11-737-2020

Otto-Bliesner, B. L., Brady, E. C., Tomas, R. A., Albani, S., Bartlein, P. J., Mahowald, N. M., et al. (2020). A comparison of the CMIP6 midHolocene and lig127k simulations in CESM2. *Paleoceanography and Paleoclimatology*, *35*(11), e2020PA003957. https://doi.org/10.1029/2020PA003957

Peltier, W. R., Argus, D. F., & Drummond, R. (2015). Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model. *Journal of Geophysical Research: Solid Earth*, *120*(1), 450–487. https://doi.org/10.1002/2014JB011176

Rose, B. E. J., Armour, K. C., Battisti, D. S., Feldl, N., & Koll, D. D. B. (2014). The dependence of transient climate sensitivity and radiative feedbacks on the spatial pattern of ocean heat uptake. *Geophysical Research Letters*, *41*(3), 1071–1078. https://doi.org/10.1002/2013GL058955

Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J. C., Hannay, C., et al. (2017). Practice and philosophy of climate model tuning across six US modeling centers. *Geoscientific Model Development*, *10*(9), 3207–3223. https://doi.org/10.5194/gmd-10-3207-2017

Schmittner, A., Urban, N. M., Shakun, J. D., Mahowald, N. M., Clark, P. U., Bartlein, P. J., et al. (2011). Climate sensitivity estimated from temperature reconstructions of the Last Glacial Maximum. *Science*, *334*(6061), 1385–1388. https://doi.org/10.1126/science.1203513

Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., et al. (2012). Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature*, *484*(7392), 49–54. https://doi.org/10.1038/nature10915

Sherwood, S. C., Bony, S., Boucher, O., Bretherton, C., Forster, P. M., Gregory, J. M., et al. (2015). Adjustments in the forcing-feedback framework for understanding climate change. *Bulletin of the American Meteorological Society*, *96*(2), 217–228. https://doi.org/10.1175/BAMS-D-13-00167.1

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., et al. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *58*, e2019RG000678. https://doi.org/10.1029/2019RG000678

Snyder, C. W. (2016). Evolution of global temperature over the past two million years. *Nature*, *538*(7624), 226–228. https://doi.org/10.1038/nature19798

Tan, I., Storelvmo, T., & Zelinka, M. D. (2016). Observational constraints on mixed-phase clouds imply higher climate sensitivity. *Science*, *352*(6282), 224. https://doi.org/10.1126/science.aad5300

Taylor, K. E., Crucifix, M., Braconnot, P., Hewitt, C. D., Doutriaux, C., Broccoli, A. J., et al. (2007). Estimating shortwave radiative forcing and response in climate models. *Journal of Climate*, *20*(11), 2530–2543. https://doi.org/10.1175/JCLI4143.1

Tierney, J. E., Malevich, S. B., Gray, W., Vetter, L., & Thirumalai, K. (2019). Bayesian calibration of the Mg/Ca paleothermometer in planktic foraminifera. *Paleoceanography and Paleoclimatology*, *34*(12). 2005–2030. https://doi.org/10.1029/2019PA003744

Tierney, J. E., Poulsen, C. J., Montañez, I. P., Bhattacharya, T., Feng, R., Ford, H. L., & Zhang, Y. G. (2020a). Past climates inform our future. *Science*, *370*(6517), eaay3701. https://doi.org/10.1126/science.aay3701

Tierney, J. E., & Tingley, M. P. (2014). A Bayesian, spatially-varying calibration model for the TEX86 proxy. *Geochimica et Cosmochimica Acta*, *127*, 83–106. https://doi.org/10.1016/j.gca.2013.11.026

Tierney, J. E., & Tingley, M. P. (2018). BAYSPLINE: A new calibration for the alkenone paleothermometer. *Paleoceanography and Paleoclimatology*, *33*(3), 281–301. https://doi.org/10.1002/2017pa003201

Tierney, J. E., Zhu, J., King, J., Malevich, S. B., Hakim, G. J., & Poulsen, C. J. (2020b). Glacial cooling and climate sensitivity revisited. *Nature*, *584*(7822), 569–573. https://doi.org/10.1038/s41586-020-2617-x

Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., et al. (2020). Past warming trend constrains future warming in CMIP6 models. *Science Advances*, *6*(12), eaaz9549. https://doi.org/10.1126/sciadv.aaz9549

Von Deimling, T. S., Ganopolski, A., Held, H., & Rahmstorf, S. (2006). How cold was the Last Glacial maximum? *Geophysical Research Letters*, *33*(14), L14709–L14709. https://doi.org/10.1029/2006GL026484

Waelbroeck, C., Paul, A., Kucera, M., Rosell-Melé, A., Weinelt, M., Schneider, R., et al. (2009). Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum. *Nature Geoscience*, *2*(2), 127–132. https://doi.org/10.1038/ngeo411

Williams, G. P., & Bryan, K. (2006). Ice Age Winds: An aquaplanet model. *Journal of Climate*, *19*(9), 1706–1715. https://doi.org/10.1175/JCLI3766.1

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*(1), e2019GL085782. https://doi.org/10.1029/2019GL085782

Zhou, C., Zelinka, M. D., & Klein, S. A. (2016). Impact of decadal cloud variations on the Earth's energy budget. *Nature Geoscience*, *9*(12), 871–874. https://doi.org/10.1038/ngeo2828

Zhu, J., Liu, Z., Brady, E., Otto-Bliesner, B., Zhang, J., Noone, D., et al. (2017). Reduced ENSO variability at the LGM revealed by an isotope-enabled Earth system model. *Geophysical Research Letters*, *44*(13), 6984–6992. https://doi.org/10.1002/2017GL073406

Zhu, J., & Poulsen, C. J. (2020a). LGM climate forcing and ocean dynamical feedback and their implications for estimating climate sensitivity. *Climate of the Past Discussions*, 1–24. https://doi.org/10.5194/cp-2020-86

Zhu, J., & Poulsen, C. J. (2020b). On the increase of climate sensitivity and cloud feedback with warming in the community atmosphere models. *Geophysical Research Letters*, *47*(18), e2020GL089143. https://doi.org/10.1029/2020gl089143

Zhu, J., Poulsen, C. J., & Otto-Bliesner, B. L. (2020). High climate sensitivity in CMIP6 model not supported by paleoclimate. *Nature Climate Change*, *10*(5). 1–2. https://doi.org/10.1038/s41558-020-0764-6

Zhu, J., Poulsen, C. J., & Tierney, J. E. (2019). Simulation of Eocene extreme warmth and high climate sensitivity through cloud feedbacks. *Science Advances*, *5*(9), eaax1874. https://doi.org/10.1126/sciadv.aax1874