AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems®

COMMENTARY
10.1029/2022MS003389

Key Points:
- Paleoclimates provide valuable tests for the latest generation of climate models
- Zhu et al., 2022, https://doi.org/10.1029/2021ms002776 provide an example of the paleoclimate benefits for climate model development
- Future iterations of CMIP will ideally incorporate paleoclimate tests during model development

Correspondence to:
N. Burls,
nburls@gmu.edu

Citation:
Burls, N., & Sagoo, N. (2022). Increasingly sophisticated climate models need the out-of-sample tests paleoclimates provide. *Journal of Advances in Modeling Earth Systems*, *14*, e2022MS003389. https://doi.org/10.1029/2022MS003389

Received 12 SEP 2022
Accepted 5 DEC 2022

Increasingly Sophisticated Climate Models Need the Out-Of-Sample Tests Paleoclimates Provide

Natalie Burls[1] and Navjit Sagoo[2]

[1]Department of Atmospheric, Oceanic, and Earth Sciences, Center for Ocean-Land-Atmosphere Studies, George Mason University, Fairfax, VA, USA, [2]Department of Meteorology, Stockholm University, Stockholm, Sweden

Abstract

Climate models are becoming increasingly sophisticated as climate scientists continually work to improve the realism with which the processes influencing Earth's climate are represented. One example is the treatment of cloud microphysics: as complexity is added to cloud microphysical schemes, Earth's energy budget can respond to changes in climate forcings, such as carbon dioxide or aerosols, in new ways. This increase in degrees of freedom has illuminated larger spread in climate sensitivity across the latest generation of climate models participating Coupled Model Intercomparison Project, Phase 6, with more high climate sensitivity models (Zelinka et al., 2020, https://doi.org/10.1029/2019gl085782). Whilst the historical record gives us just over a century of data to apply toward climate sensitivity constraints (e.g., Nijsse et al., 2020, https://doi.org/10.5194/esd-11-737-2020), the ocean is still taking up much of the heat trapped by anthropogenic greenhouse gas emissions and the climate system is far from equilibrium which limits our understanding how climate sensitivity might change in response to long-term forced climate change. Here we discuss the valuable tests that paleoclimate reconstructions can provide the latest generation of climate models, as demonstrated by the recent study of Zhu et al., 2022, https://doi.org/10.1029/2021ms002776. Their study provides an example of the benefits for climate model development when climate models are confronted with simulating climates very different from today. Ideally the climate model development stage under future iterations of CMIP will involve such tests as an effort to constrain global climate sensitivity and the regional patterns of climate, such as polar amplification and subtropical aridification.

Plain Language Summary

In each successive generation of climate models the representation of climate processes becomes more realistic and more complex. The more sophisticated characterization of cloud processes has resulted in some models warming much more in response to increasing atmospheric carbon dioxide concentrations due to stronger cloud feedbacks; referred to as having a higher climate sensitivity. The Community Earth System Model version 2 (CESM2) is one such model. When used to simulate the Last Glacial Maximum (LGM), CESM2 has a global temperature around 5°C colder than surface temperature reconstructions and previous generations of this model. By investigating the updated cloud scheme, Zhu et al., 2022, https://doi.org/10.1029/2021ms002776 were able to identify and modify issues with the cloud microphysical scheme that were contributing to this problem. They subsequently produced a new version of CESM2 which can simulate the LGM more accurately without degrading the simulation of the 20th century climate. The Zhu et al., 2022, https://doi.org/10.1029/2021ms002776 study is an example of how a paleoclimate can be used to identify issues with cloud schemes not traditionally recognized in the model development and validation process. Incorporating paleoclimate simulations earlier in the model development and validation process may help constrain cloud feedbacks and subsequently climate sensitivity.

Main Text

Since their inception in the 1950s, weather forecast models have steadily advanced in what has been termed a "quiet revolution" (Bauer et al., 2015). Decades of routine tests and iterative testing and development underpin this revolution: every day that a weather forecast is made, the skill of this forecast is assessed in the following days as the weather is observed and recorded. Large disagreements between forecasts and observations lead to an exploration of the processes that cause this deviation, and how it can be improved and minimized, resulting in iterative model development. Decades of this process have resulted in skilled models that allow us to predict the very chaotic nature of weather with a degree of accuracy that enables society to make a range of cost, time and lifesaving decisions (Lazo et al., 2009).

Climate prediction deals with forecasting how the statistics of weather will change over longer timescales: weeks, seasons, years, decades, and centuries. Climate models simulate the interaction between each component of the climate system, namely the atmosphere, land, ocean, and cryosphere. Climate models are used to explore both interactions between the separate components and also how the influence of changing boundary conditions such as insolation, greenhouse gas concentrations, atmospheric aerosols, and land surface changes, lead to climate variability, and change. Much like the enterprise of weather prediction, using the historical observational record we are able to evaluate these models against sub-seasonal to seasonal forecasts (e.g., Pegion et al., 2019), decadal forecasts (e.g., Meehl et al., 2014), and 20th century climate change (e.g., Golaz et al., 2013). However, as we move to longer timescales it becomes harder to present these models with out-of-sample tests using observational data.

Given their limited temporal nature, observational records provide limited tests for model evaluation on centennial plus timescales, particularly in terms of assessing the role of internal variability versus the response to external forcing (e.g., Zhou et al., 2016). Whilst the historical record gives us just over a century of data to apply toward climate sensitivity constraints (e.g., Nijess et al., 2020), the climate system is far from equilibrium which has consequences for estimating how much warming will ultimately occur (Sherwood et al., 2020). Evaluating climate models on centennial-plus timescales is of critical importance as we prepare for the future: one integral piece is understanding how the climate feedbacks associated with observed short-term variability compare to the feedbacks associated with long-term forced climate change (Sherwood et al., 2020). The paleoclimate record is essential for filling this gap in model evaluation.

Paleoclimate reconstructions allow us to explore how the climate system responds to large climate forcings over centennial and longer timescales. We can evaluate how well the critical processes involved in climate variability and climate change on these timescales are represented in climate models (Tierney, Poulsen, et al., 2020). Paleoclimate reconstructions are imperfect, and come with their own unique limitations (e.g., Hollis et al., 2019), however, they provide a diverse range of out-of-sample climates that can be used to evaluate climate models. There is tremendous potential for a unified approach to the prediction of climate variability and change on centennial plus timescales which utilizes paleoclimate. The recently published study by Zhu et al. (2022), which uses the Last Glacial Maximum (LGM) to better constrain Equilibrium Climate Sensitivity (ECS) in the Community Earth System Model version 2 (CESM2), is at the forefront of what can be achieved with closer collaboration between the climate modeling and paleoclimate communities.

Climate sensitivity is a key metric for understanding how susceptible the Earth's climate is to anthropogenic forcing. Constraining climate sensitivity, particularly the upper boundary, has proved difficult with the range of ECS values for a doubling of CO_2 under modern boundary conditions estimated as 2.3–4.7°C (90% confidence) by Sherwood et al. (2020) and the latest Intergovernmental Panel on Climate Change (IPCC) estimating 2–5°C as very likely and 2.5–4°C as likely (Forster et al., 2021). This spread has been largely attributed to uncertainty in the strength of cloud feedbacks for example, (Boucher et al., 2013; Cess et al., 1990; Dufresne & Bony, 2008; Vial et al., 2013; Zelinka et al., 2017), particularly tropical and low-cloud feedback strengths (Bony & Dufresne, 2005; Sherwood et al., 2014; Webb et al., 2006; Wetherald & Manabe, 1988), and the lack of observational constraints on centennial-plus timescales is hindering our ability to constrain climate sensitivity (Sherwood et al., 2020).

The processes that determine the radiative forcing associated with clouds are complex as they span spatial scales from that of a cloud droplet to as large as an ocean basin. As a result, all processes acting on spatial scales smaller than the ~100 km horizontal scale typically resolved by the grid cell of a climate model, need to be parameterized in a climate model. In an effort to increase their realism, and to address the uncertainty associated with cloud feedbacks in climate models, the treatment of clouds within climate models has become increasingly sophisticated over time. This has increased the degrees of freedom in how cloud radiative properties can respond to global warming, or cooling in the case of LGM, and is the likely cause of the larger inter-model spread in cloud feedback strength and climate sensitivity in the sixth, and most recent, iteration of the Coupled Model Intercomparison Project Phase 6 (CMIP6) (Zelinka et al., 2020). The CESM2 model used in Zhu et al. (2022) is one of the CMIP models which has a high climate sensitivity, >5°C, which is outside of the IPCC's very likely range. This high climate sensitivity is attributed to strongly positive cloud feedbacks within its atmospheric component, the Community Atmosphere Model version 6 (e.g., Bjordal et al., 2020; Gettelman et al., 2019; Zhu et al., 2021).

As a consequence of its high sensitivity, Zhu et al., 2021 found that CESM2 simulates LGM global-mean cooling that exceeds 11°C which is outside the 4.6°C–6.8°C 95% confidence range indicated by proxy data and greater than that in previous versions of the model. In a recent follow up paper, Zhu et al. (2022) investigate the mechanisms driving the strong positive cloud feedback and strong LGM cooling in CESM2. They iteratively substitute the parameterization schemes affecting cloud properties in CESM2 with previous versions within their LGM configuration. This leads the authors to identify two microphysical issues that contribute to the strongly positive cloud feedbacks in CESM2: (a) a limit on cloud ice number in the ice nucleation scheme, that affects the treatment of mixed-phase clouds and (b) an issue related to the sub-stepping time step of the cloud microphysics scheme. Removing the limiter and increasing the sub-stepping improves the representation of cloud ice number, which in turn weakens the shortwave cloud feedback. The authors use these findings to develop a paleoclimate calibrated version of CESM2 that has an ECS of ∼4°C and is able simulate a more realistic LGM climate without degrading its ability to capture 20th century climate.

The representation of cloud microphysical processes such as precipitation efficiency, cloud phase and aerosols in general circulation models continues to grow in sophistication, complemented by efforts like "DYAMOND" (Stevens et al., 2019) that aim to explicitly simulate many of these processes at high resolution allowing for simpler and more accurate cloud parameterizations. It is not clear whether through these scientific advancements there will eventually be some convergence in the nature and strength of cloud feedbacks across models. Therefore a complementary paleoclimatic approach, such as that of Zhu et al., 2022, is needed as it provides an additional constraint on the influence of cloud microphysics. Applying the LGM constraint to CESM2 resulted in a 40% reduction in the SW cloud feedback and a 20% reduction in aerosol-cloud interactions. The Zhu et al. (2022) study is complemented by those of Feng et al., 2020 and Zhu et al., 2020 whom assessed CESM2's skill in simulating the warm Pliocene and Eocene climates, respectively, with both studies concluding that CESM2's climate sensitivity is likely too high due to cloud feedbacks. While there is no direct proxy for clouds, they exert a dominant control on large-scale sea surface temperature (SST) patterns (e.g., Burls & Fedorov, 2014; Erfani & Burls, 2019; Fedorov et al., 2015) and model differences in meridional energy transport (Donohoe & Battisti, 2012; Stone, 1978). Perturbed parameter ensembles in which uncertain cloud and convective parameters are perturbed in different climates (e.g., Ramos et al., 2022; Sagoo et al., 2013) can be used to determine the best suite of parameters and/or geographical locations with high surface temperature sensitivity to these parameters. Quantifying uncertainty in these parameters is critical for constraining ECS.

Paleoclimates have the potential to place robust additional constraints on modern climate models by providing out-of-sample tests. Efforts such as the Paleoclimate Modelling Intercomparison Project, PMIP, (Kageyama et al., 2018) which targets the last millennium (Jungclaus et al., 2017), two interglacials (the mid-Holocene and Last Interglacial, Otto-Bliesner et al., 2017) and the LGM (Kageyama et al., 2021); the Pliocene Model Intercomparision Project, PlioMIP, (Haywood et al., 2020); DeepMIP-Eocene (Lunt et al., 2021); and more recently DeepMIP-Miocene (Burls et al., 2021), have long recognized the utility of simulating paleoclimates as both a tool for the interpretation of reconstructions and a test for our climate models. Leveraging these paleoclimate modeling intercomparison activities, the latest IPCC report highlights that both high-ECS and low-ECS models struggle to correctly simulate multiple paleo time periods (Forster et al., 2021, Figure 7.19). There is however room for more coordination, for modeling centers to routinely simulate targeted paleoclimate intervals as a part of the model development cycle, facilitated by community-lead efforts. There is also the scope to go beyond simply using global mean temperature reconstructions as a climate sensitivity constraint by evaluating the ability of climate models to simulate the reconstructed patterns of warming. Polar amplified warming is a robust feature of past warm climates such as the Pliocene, Miocene, and Eocene yet climate model simulations targeting these intervals have historically not captured the full extent of polar amplified warming seen in the proxy data when run to near equilibrium. For example, climate models tend to underestimate the amount of warmth reconstructed in proxy records in the Southern Ocean and North Atlantic during the Miocene (Burls et al., 2021) and Eocene (Lunt et al., 2021).

Paleoclimate data is not without its challenges and there are outstanding issues for the LGM and other paleoclimates that must be resolved, specifically if we are to move beyond global mean comparisons. The LGM is oft-cited as a useful paleoclimate as it is a recent climate with an abundance of surface temperature proxy data and high-fidelity records of greenhouse gas and dust aerosol deposition data (Lambert et al., 2008; Petit et al., 1999). However, there are currently multiple LGM SST reconstructions (Annan et al., 2022; Paul et al., 2021; Tierney, Zhu, et al., 2020) with substantial differences. Correctly reconstructing SST patterns in the tropical Pacific is particularly important as the radiative response globally is very sensitive to details of Pacific SST changes (Andrews & Webb, 2018) and yet Pacific SSTs are weakly constrained in all three reconstructions due to a scarcity in data. Furthermore, aerosol-climate interactions such as that of mineral dust is poorly constrained at the LGM (Albani et al., 2018; Lambert & Albani, 2021; Sagoo & Storelvmo, 2017), with large discrepancies in dust emissions, radiative effects and total dust forcing in the few models that do incorporate the dust cycle into their simulations. Robust and accurate reconstructions of SST and sea ice patterns, and forcing estimates with uncertainty quantification are essential for the LGM and other paleoclimates.

A number of high ECS (>5°C) CMIP6/PMIP4 models planned to simulate the LGM, however, of the five studies currently available (Kageyama et al., 2021) the ECS range in the published studies is limited to 2.1–3.6°C when CESM2 (5.6°C) is excluded. Many high ECS models (EC-Earth, HadGEM and IPSL) have been delayed, or no longer plan to simulate the LGM due to issues with running the model (personal communications). Understanding why simulating the LGM has been so challenging for these models is important. Is it due to limitations in the models that is, numerical instabilities, issues with tuning and parameterizations, the models exhibiting runaway cooling, or simply the inherent challenges associated with implementing paleoclimate boundary conditions within models designed to simulate the historical period? If simulating paleoclimates such as the LGM climate are inconsistent with high ECS models it may be used as a line of evidence in ruling out high ECS.

There is no one ideal paleoclimate interval for testing climate models, and the choice of interval depends largely on the climate phenomena of interest. For example, the last millennium has proven to be a useful time period for evaluating the fidelity of the response to volcanic forcing (e.g., Sigl et al., 2015; Soden et al., 2002; Zanchettin et al., 2016). The more recent last millennium, mid-Holocene, and last interglacial have a plethora of data, making them particularly well suited to addressing questions surrounding internal modes of climate variability. When it comes to constraining cloud feedbacks however, their signal (forced climate change) to noise (internal climate variability) ratio is smaller than that of earlier climates such as the LGM, Pliocene, Miocene and Eocene. Large ice-sheets introduce additional uncertainty at the LGM (e.g., Ullman et al., 2014), as well as the issue of state-dependence (e.g., Kohler et al., 2015; Friedrich et al., 2016; von der Heydt et al., 2016; Anagnostou et al., 2020). Going further back to warmer climates such as the Pliocene, Miocene and Eocene not only is the signal-to-noise ratio large, but these climates are potential analogs of middle or high-end future scenarios respectively. However, one needs to contend with a paucity of proxy data and new uncertainties in boundary conditions such as continental configuration, orography and land surface type. Subsequently there are trade-offs between any of the paleoclimates selected, and careful thought needs to be given toward which paleoclimate interval can best constrain the simulated climate processes of interest. For ECS, a recent study using an emergent constraints approach, which infers statistical relationships between two variables of the climate system within an ensemble of climate models, has re-evaluated all generations of PMIP data and finds that the Pliocene provides a better constraint on ECS than the LGM despite there being lower-fidelity reconstructions and increased uncertainty that comes from using much older paleoclimate data (Renoult et al., 2022).

We propose that paleoclimate simulations need to be involved at an earlier stage in the model development and validation process, rather than in parallel with future scenarios (Figure 1). Running paleoclimate simulations earlier in this process (Figure 1b) would allow the findings to be fed back into model development and we speculate that validating with both warm and cold paleoclimates could help reduce the spread in the future ensemble (Figure 1d). This shift from the current practice (Figure 1a) comes however with the caveat that any paleoclimates used in the model development stage (Figure 1b) would no longer be able to provide an out-of-sample test. This caveat could be addressed by withholding select periods in Earth's history from the model development phase. For example, taking the Zhu et al. (2022) LGM calibrated version of CESM2 and using it to simulate the Pliocene or Eocene (something currently being undertaken). If the LGM calibrated version of CESM2 does a better job of simulating Pliocene and Eocene reconstructions this adds further confidence.

In short, there are opportunities for both climate modelers and paleoclimatologists if we are to fully harness the utility of paleoclimates in climate model development. We need strategic and coordinated community efforts to produce reliable reconstructions of surface temperature and boundary conditions that fill the gaps and allow for the generation of robust estimates of surface temperature patterns during intervals targeted by model-data comparisons. This is by no means a novel suggestion and there is a history of coordination amongst the paleoclimate "data" and "modeling" communities, but there is still plenty scope for targeting select time intervals and regions of greatest process and proxy uncertainty. To this end, the community could establish a paleo equivalent of the "obs4mips" initiative that makes observational products over the instrumental record more accessible for climate model intercomparison and evaluation. On the climate modeling side we need modeling centers to

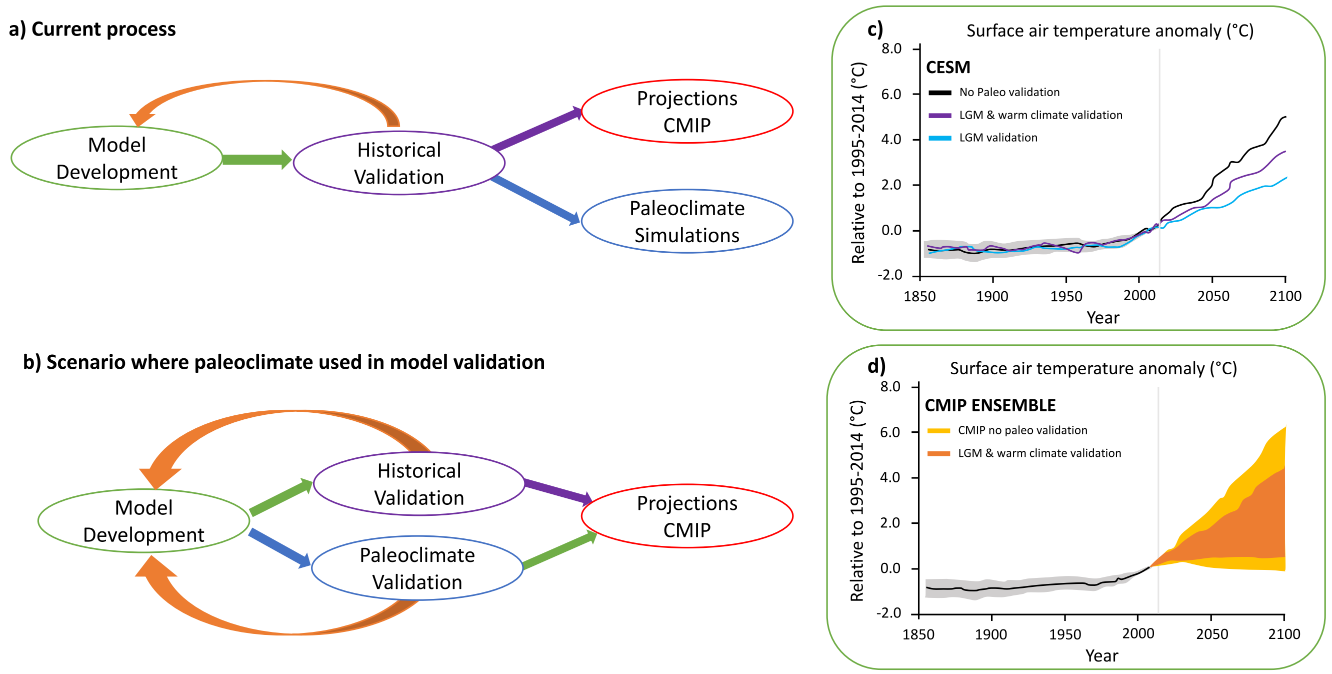

Figure 1. (a) Schematic of current model development and validation process. (b) Schematic showing paleoclimate validation impacting model developments with (c) the different potential trajectories of future scenarios for a single model that validates with the Last Glacial Maximum (LGM) and the LGM + warm climates (e.g., Pliocene, Miocene and Eocene) and (d) how this could potentially impact CMIP ensemble of future scenarios.

undertake the routine simulation of the LGM and select warm climate intervals targeted by proxy reconstructions efforts. A holistic understanding of when and why models are not able to simulate paleoclimates will be valuable.

Using a first-generation climate model, the recent Nobel prize laureate Suki Manabe and collaborators predicted in 1975 that the surface warming in response to elevated atmospheric CO_2 concentrations would be polar amplified (Manabe & Wetherald, 1975). This prediction has since been verified with the strongest historical surface warming seen over the Arctic, but projections of the fate of the Arctic differ widely across climate models due to different model feedback strengths. If we are to be similarly successful in predicting the patterns of future warming, and additionally the magnitude of change, in 2060, 2100 and beyond we need to thoroughly confront climate models with the paleoclimate record.

Data Availability Statement

Data were not used, nor created for this research.

Acknowledgments
The authors thank Thorsten Mauritsen and three anonymous reviewers for helpful feedback on this commentary. NJB acknowledges supported through funding from the National Science Foundation (NSF) Award # 1844380 and NS was supported through funding from the European Research Council (ERC) Grant agreement 770765.

References

Albani, S., Balkanski, Y., Mahowald, N., Winckler, G., Maggi, V., & Delmonte, B. (2018). Aerosol-climate interactions during the last glacial maximum. *Current Climate Change Reports*, *4*(2), 99–114. https://doi.org/10.1007/s40641-018-0100-7

Anagnostou, E., John, E. H., Babila, T. L., Sexton, P. F., Ridgwell, A., Lunt, D. J., et al. (2020). Proxy evidence for state-dependence of climate sensitivity in the Eocene greenhouse. *Nature Communications*, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-17887-x

Andrews, T., & Webb, M. J. (2018). The dependence of global cloud and lapse rate feedbacks on the spatial structure of tropical Pacific warming. *Journal of Climate*, *31*(2), 641–654. https://doi.org/10.1175/jcli-d-17-0087.1

Annan, J., Hargreaves, J., & Mauritsen, T. (2022). A new global climate reconstruction for the last glacial maximum. *Climate of the Past Discussions*, *18*(8), 1–20. https://doi.org/10.5194/cp-18-1883-2022

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*(7567), 47–55. https://doi.org/10.1038/nature14956

Bjordal, J., Storelvmo, T., Alterskjær, K., & Carlsen, T. (2020). Equilibrium climate sensitivity above 5°C plausible due to state-dependent cloud feedback. *Nature Geoscience*, *13*(11), 718–721. https://doi.org/10.1038/s41561-020-00649-1

Bony, S., & Dufresne, J.-L. (2005). Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, *32*(20), L20806. https://doi.org/10.1029/2005gl023851

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., et al. (2013). Clouds and aerosols. In T. F. Stocker, D. Qin, G. K. Plattner, M. M. M. B. Tignor, S. K. Allen, & J. Boschung (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. 571–657). Cambridge University Press.

Burls, N. J., Bradshaw, C. D., De Boer, A. M., Herold, N., Huber, M., Pound, M., et al. (2021). Simulating Miocene warmth: Insights from an opportunistic Multi-Model ensemble (MioMIP1). *Paleoceanography and Paleoclimatology*, *36*(5), e2020PA004054. https://doi.org/10.1029/2020pa004054

Burls, N. J., & Fedorov, A. V. (2014). What controls the mean east–west sea surface temperature gradient in the equatorial Pacific: The role of cloud albedo. *Journal of Climate*, *27*(7), 2757–2778. https://doi.org/10.1175/jcli-d-13-00255.1

Cess, R. D., Potter, G. L., Blanchet, J. P., Boer, G. J., Del Genio, A. D., Deque, M., et al. (1990). Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *Journal of Geophysical Research*, *95*(D10), 601–615. https://doi.org/10.1029/jd095id10p16601

Donohoe, A., & Battisti, D. S. (2012). What determines meridional heat transport in climate models? *Journal of Climate*, *25*(11), 3832–3850. https://doi.org/10.1175/jcli-d-11-00257.1

Dufresne, J.-L., & Bony, S. (2008). An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models. *Journal of Climate*, *21*(19), 5135–5144. https://doi.org/10.1175/2008jcli2239.1

Erfani, E., & Burls, N. J. (2019). The strength of low-cloud feedbacks and tropical climate: A CESM sensitivity study. *Journal of Climate*, *32*(9), 2497–2516. https://doi.org/10.1175/jcli-d-18-0551.1

Fedorov, A. V., Burls, N. J., Lawrence, K. T., & Peterson, L. C. (2015). Tightly linked zonal and meridional sea surface temperature gradients over the past five million years. *Nature Geoscience*, *8*(12), 975–980. https://doi.org/10.1038/ngeo2577

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

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., et al. (2021). The Earth's energy budget, climate feedbacks, and climate sensitivity. In P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, et al. (Eds.), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V*. (pp. 923–1054). Cambridge University Press. https://doi.org/10.1017/9781009157896.009

Friedrich, T., Timmermann, A., Tigchelaar, M., Timm, O. E., & Ganopolski, A. (2016). Nonlinear climate sensitivity and its implications for future greenhouse warming. *Science Advances*, *2*(11), e1501923. https://doi.org/10.1126/sciadv.1501923

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*(14), 8329–8337. https://doi.org/10.1029/2019gl083978

Golaz, J.-C., Horowitz, L. W., & Levy, H. (2013). Cloud tuning in a coupled climate model: Impact on 20th century warming. *Geophysical Research Letters*, *40*(10), 2246–2251. https://doi.org/10.1002/grl.50232

Haywood, A. M., Tindall, J. C., Dowsett, H. J., Dolan, A. M., Foley, K. M., Hunter, S. J., et al. (2020). The Pliocene Model Intercomparison Project Phase 2: Large-scale climate features and climate sensitivity. *Climate of the Past*, *16*(6), 2095–2123. https://doi.org/10.5194/cp-16-2095-2020

Hollis, C. J., Dunkley Jones, T., Anagnostou, E., Bijl, P. K., Cramwinckel, M. J., Cui, Y., et al. (2019). The DeepMIP contribution to PMIP4: Methodologies for selection, compilation, and analysis of latest Paleocene and early Eocene climate proxy data, incorporating version 0.1 of the DeepMIP database. *Geoscientific Model Development*, *12*(7), 3149–3206. https://doi.org/10.5194/gmd-12-3149-2019

Jungclaus, J. H., Bard, E., Baroni, M., Braconnot, P., Cao, J., Chini, L. P., et al. (2017). The PMIP4 contribution to CMIP6–Part 3: The last millennium, scientific objective, and experimental design for the PMIP4 past1000 simulations. *Geoscientific Model Development*, *10*(11), 4005–4033. https://doi.org/10.5194/gmd-10-4005-2017

Kageyama, M., Braconnot, P., Harrison, S. P., Haywood, A. M., Jungclaus, J. H., BetteOtto-Bliesner, 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

Kageyama, M., Harrison, S. P., Kapsch, M. L., Lofverstrom, M., Lora, J. M., Mikolajewicz, U., et al. (2021). The PMIP4 Last Glacial Maximum experiments: Preliminary results and comparison with the PMIP3 simulations. *Climate of the Past*, *17*(3), 1065–1089. https://doi.org/10.5194/cp-17-1065-2021

Köhler, P., de Boer, B., von der Heydt, A. S., Stap, L. B., & van de Wal, R. S. (2015). On the state dependency of the equilibrium climate sensitivity during the last 5 million years. *Climate of the Past*, *11*(12), 1801–1823. https://doi.org/10.5194/cp-11-1801-2015

Lambert, F., & Albani, S. (2021). Mineral dust in PMIP simulations: A short review. *Past Global Changes Magazine*, *29*(2), 86–87.

Lambert, F., Delmonte, B., Petit, J.-R., Bigler, M., Kaufmann, P. R., Hutterli, M. A., et al. (2008). Dust-climate couplings over the past 800,000 years from the EPICA Dome C ice core. *Nature*, *452*(7187), 616–619. https://doi.org/10.1038/nature06763

Lazo, J. K., Morss, R. E., & Demuth, J. L. (2009). 300 billion served: Sources, perceptions, uses, and values of weather forecasts. *Bulletin of the American Meteorological Society*, *90*(6), 785–798. https://doi.org/10.1175/2008bams2604.1

Lunt, D. J., Bragg, F., Chan, W.-L., Hutchinson, D. K., Ladant, J.-B., Morozova, P., et al. (2021). DeepMIP: Model intercomparison of early Eocene climatic optimum (EECO) large-scale climate features and comparison with proxy data. *Climate of the Past*, *17*(1), 203–227. https://doi.org/10.5194/cp-17-203-2021

Manabe, S., & Wetherald, R. T. (1975). The effects of doubling the CO_2 concentration on the climate of a general circulation model. *Journal of the Atmospheric Sciences*, *32*(1), 3–15. https://doi.org/10.1175/1520-0469(1975)032<0003:teodtc>2.0.co;2

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., et al. (2014). Decadal climate prediction: An update from the trenches. *Bulletin of the American Meteorological Society*, *95*(2), 243–267. https://doi.org/10.1175/bams-d-12-00241.1

Nijsse, F. J., 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., Braconnot, P., Harrison, S. P., Lunt, D. J., Abe-Ouchi, A., Albani, S., et al. (2017). The PMIP4 contribution to CMIP6–Part 2: Two interglacials, scientific objective and experimental design for Holocene and Last Interglacial simulations. *Geoscientific Model Development*, *10*(11), 3979–4003. https://doi.org/10.5194/gmd-10-3979-2017

Paul, A., Mulitza, S., Stein, R., & Werner, M. (2021). A global climatology of the ocean surface during the Last Glacial Maximum mapped on a regular grid (GLOMAP). *Climate of the Past*, *17*(2), 805–824. https://doi.org/10.5194/cp-17-805-2021

Pegion, K., Kirtman, B. P., Becker, E., Collins, D. C., LaJoie, E., Burgman, R., et al. (2019). The subseasonal experiment (SubX): A multimodel subseasonal prediction experiment. *Bulletin of the American Meteorological Society*, *100*(10), 2043–2060. https://doi.org/10.1175/bams-d-18-0270.1

Petit, J.-R., Jouzel, J., Raynaud, D., Barkov, N. I., Barnola, J.-M., Basile, I., et al. (1999). Climate and atmospheric history of the past 420, 000 years from the Vostok ice core, Antarctica. *Nature*, *399*(6735), 429–436. https://doi.org/10.1038/20859

Ramos, R. D., LeGrande, A. N., Griffiths, M. L., Elsaesser, G. S., Litchmore, D. T., Tierney, J. E., et al. (2022). Constraining clouds and convective parameterizations in a climate model using paleoclimate data. *Journal of Advances in Modeling Earth Systems*, *14*(8), e2021MS002893. https://doi.org/10.1029/2021ms002893

Renoult, M., Sagoo, N., Zhu, J., & Mauritsen, T. (2022). Causes of the weak emergent constraint on climate sensitivity at the Last Glacial Maximum. *Climate of the Past Discussions*, 1–55.

Sagoo, N., & Storelvmo, T. (2017). Testing the sensitivity of past climates to the indirect effects of dust. *Geophysical Research Letters*, *44*(11), 2017GL072584. https://doi.org/10.1002/2017GL072584

Sagoo, N., Valdes, P., Flecker, R., & Gregoire, L. J. (2013). The early Eocene equable climate problem: Can perturbations of climate model parameters identify possible solutions? *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *371*(2001), 20130123. https://doi.org/10.1098/rsta.2013.0123

Sherwood, S. C., Bony, S., & Dufresne, J.-L. (2014). Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, *505*(7481), 37–42. https://doi.org/10.1038/nature12829

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*(4), e2019RG000678. https://doi.org/10.1029/2019rg000678

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., et al. (2015). Timing and climate forcing of volcanic eruptions for the past 2, 500 years. *Nature*, *523*(7562), 543–549. https://doi.org/10.1038/nature14565

Soden, B. J., Wetherald, R. T., Stenchikov, G. L., & Robock, A. (2002). Global cooling after the eruption of Mount Pinatubo: A test of climate feedback by water vapor. *Science*, *296*(5568), 727–730. https://doi.org/10.1126/science.296.5568.727

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., et al. (2019). DYAMOND: The DYnamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Progress in Earth and Planetary Science*, *6*(1), 1–17. https://doi.org/10.1186/s40645-019-0304-z

Stone, P. H. (1978). Constraints on dynamical transports of energy on a spherical planet. *Dynamics of Atmospheres and Oceans*, *2*(2), 123–139. https://doi.org/10.1016/0377-0265(78)90006-4

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

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

Ullman, D. J., LeGrande, A. N., Carlson, A. E., Anslow, F. S., & Licciardi, J. M. (2014). Assessing the impact of Laurentide Ice Sheet topography on glacial climate. *Climate of the Past*, *10*(2), 487–507. https://doi.org/10.5194/cp-10-487-2014

Vial, J., Dufresne, J.-L., & Bony, S. (2013). On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Climate Dynamics*, *41*(11–12), 3339–3362. https://doi.org/10.1007/s00382-013-1725-9

von der Heydt, A. S., Dijkstra, H. A., van de Wal, R. S., Caballero, R., Crucifix, M., Foster, G. L., et al. (2016). Lessons on climate sensitivity from past climate changes. *Current Climate Change Reports*, *2*(4), 148–158. https://doi.org/10.1007/s40641-016-0049-3

Webb, M. J., Senior, C. A., Sexton, D. M. H., Ingram, W. J., Williams, K. D., Ringer, M. A., et al. (2006). On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dynamics*, *27*(1), 17–38. https://doi.org/10.1007/s00382-006-0111-2

Wetherald, R. T., & Manabe, S. (1988). Cloud feedback processes in a general circulation model. *Journal of the Atmospheric Sciences*, *45*(8), 1397–1416. https://doi.org/10.1175/1520-0469(1988)045<1397:cfpiag>2.0.co;2

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., et al. (2016). The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP): Experimental design and forcing input data for CMIP6. *Geoscientific Model Development*, *9*(8), 2701–2719. https://doi.org/10.5194/gmd-9-2701-2016

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

Zelinka, M. D., Randall, D. A., Webb, M. J., & Klein, S. A. (2017). Clearing clouds of uncertainty. *Nature Climate Change*, *7*(10), 674–678. https://doi.org/10.1038/nclimate3402

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., Otto-Bliesner, B. L., Brady, E. C., Gettelman, A., Bacmeister, J. T., Neale, R. B., et al. (2022). LGM paleoclimate constraints inform cloud parameterizations and equilibrium climate sensitivity in CESM2. *Journal of Advances in Modeling Earth Systems*, *14*(4), e2021MS002776. https://doi.org/10.1029/2021ms002776

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*(3), e2020GL091220. https://doi.org/10.1029/2020gl091220

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), 378–379. https://doi.org/10.1038/s41558-020-0764-6