Paleoclimate and Historical Perspectives on Modern Climate Sensitivity

Vincent Cooper

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Kyle Armour, Chair

Gregory Hakim, Chair

Cecilia Bitz

Dennis Hartmann

Program Authorized to Offer Degree:
Atmospheric and Climate Science

University of Washington

Abstract

Paleoclimate and Historical Perspectives on Modern Climate Sensitivity

Vincent Cooper

Co-Chairs of the Supervisory Committee:
Kyle Armour
Department of Atmospheric and Climate Science and School of Oceanography

Gregory Hakim
Department of Atmospheric and Climate Science

Determining modern climate sensitivity, i.e., the global surface warming from doubling preindustrial concentrations of CO_2, is an urgent task as it controls how much the planet will warm from greenhouse-gas emissions. The upper bound on estimates of climate sensitivity has been highly uncertain for decades, but paleoclimates now provide a strong constraint. In this dissertation, we combine proxy data from paleoclimate data assimilation with atmospheric general circulation models to show that the climate sensitivity inferred from paleoclimates is systematically higher than the climate sensitivity that applies to modern warming from CO_2. This difference in climate sensitivity arises because (a) ice sheets, topography, and vegetation changes drive atmospheric stationary waves that alter the spatial patterns of sea-surface temperature (SST) over distant oceans during both the cold Last Glacial Maximum and the warm Pliocene; and (b) these paleoclimate SST patterns are associated with amplifying cloud feedbacks that make past climates more sensitive than the modern climate. Accounting for these differences between climates leads to a substantial reduction ($\sim$1.0°C) in the upper bound on modern climate sensitivity compared to recent community assessments, such as IPCC AR6 (Forster et al., 2021).

The leading role of spatial patterns of temperature change in determining climate sensitivity also compels a re-evaluation of the historical climate record (c. 1850–present). Previous studies have identified major discrepancies in radiative feedbacks due to differences in the patterns of sea-surface temperature across instrumental datasets. These discrepancies result from statistical infilling of the

expansive gaps between sparse SST observations over the global oceans. In this dissertation, we use coupled data assimilation, which optimally combines observational and dynamical constraints from all climate fields simultaneously, to reconstruct monthly and globally resolved SST, near-surface air temperature, sea ice, and sea-level pressure over 1850–2023. The reconstruction provides a novel and internally consistent perspective on coupled climate variability and recent trends, which informs investigation of radiative feedbacks in the historical record.

Chapter 1 introduces the research topics addressed in this dissertation. Chapter 2 quantifies Last Glacial Maximum pattern effects and their impacts on modern climate sensitivity. Chapter 3 quantifies Pliocene pattern effects and provides stronger constraints on both modern climate sensitivity and 21st-century warming. Chapter 4 presents a reconstruction of the historical climate record (1850–2023) using linear inverse models and coupled data assimilation. Chapter 5 reviews the conclusions of the dissertation.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I am deeply grateful to my advisors, Kyle Armour, Greg Hakim, and Cecilia Bitz. Kyle encouraged me to pursue big ideas and, in so many ways, has shaped and inspired the future of my research. Greg has been an ardent, unwavering supporter of me and my goals; our deep dives into statistical thinking have been illuminating and fun for both of us. CC taught me how to be a scientist and the art and importance of designing my own experiments, and I will always be grateful to her for enabling my leap from finance to atmospheric science. I feel incredibly lucky to have worked with and learned so much from all of them. I thank Dennis Hartmann, who has been a mentor and friend, sharing wisdom and many good laughs throughout my graduate years. I appreciate Armeen Taeb for generously serving as the GSR on my committee.

I am thankful for the UW faculty, especially D. Battisti, A. Donohoe, A. Swann, E. Blanchard-Wrigglesworth, G. Roe, A. Turner, and R. Wood, for their guidance and many fruitful discussions. I am particularly grateful to Cristian Proistosescu for mentorship and helping me navigate my research interests and to Lettie Roach as I started graduate school. I have enjoyed learning from many collaborators, especially J. Tierney, T. Andrews, D. Amrhein, N. Burls, Y. Dong, J. Zhu, M. Osman, M. Dvorak, and R. Feng. My research groups and graduate cohorts have made graduate school fun and collaborative. I appreciate the camaraderie of my officemates, including Joe Robinson, Henry Hua, Adam Sokol, Aodhan Sweeney, Lily Zhang, and Brittany Dygert.

My family and friends have always been an incredible source of joy, love, and support. I am especially thankful for Canyon Woodward, who encouraged me to take the big risk and pursue science, and for further inspiration from friends Camilla Gibson and Sam Tifft. I am thankful for my fun and supportive family, including parents Kim, Terry, and Dave, and siblings Patrick and sister-in-law Helena, Anna, John, and Lisa. My niece Miriam and nephews Martin, Francis, and Laurence always make me smile, and I am proud to be their uncle. Most of all, I am grateful to Ursula—I am infinitely lucky to have her in my life and to have such an incredible braintrust.

Chapter 1

INTRODUCTION

Earth's response to greenhouse-gas forcing can be characterized by the equilibrium climate sensitivity (ECS), defined as the global-mean change in near-surface air temperature due to a doubling of carbon dioxide ($2xCO_2$) from pre-industrial levels (Charney et al., 1979). Earth's ECS is difficult to constrain due to uncertainty in the radiative feedbacks that govern Earth's energy balance (Roe and Baker, 2007; Knutti and Hegerl, 2008), but constraining ECS is essential for projecting future changes not only in surface temperature but also in the myriad climate features that depend on surface temperature. Progress in constraining ECS has been slow over the past 40 years, and climate models have not converged on a best estimate (Zelinka et al., 2020).

Recent community assessments of feedbacks and ECS have focused on three lines of evidence: "process understanding" of feedbacks, observed changes over the historical record (1850–present), and proxy evidence from the paleoclimate record. Sherwood et al. (2020) developed a Bayesian framework for quantitatively combining these lines of evidence, leading to a substantial narrowing of uncertainty compared to previous assessments. A primary motivation for this dissertation is constraining radiative feedbacks and modern ECS with evidence from paleoclimates and the historical record, leveraging recent advances in understanding past patterns of sea-surface temperature (SST) and their connections to radiative feedbacks.

Using paleoclimates to constrain modern-day ECS requires accounting for how climate feedbacks change across different climate states. The standard assumption is that colder climates are less sensitive (i.e., have more-negative feedbacks) than warmer states (Sherwood et al., 2020; Forster et al., 2021; PALAEOSENS Project Members, 2012; Köhler et al., 2015; von der Heydt et al., 2016; Friedrich et al., 2016; Rohling et al., 2018). However, the simple assumption that feedbacks change with *global-mean* temperature does not account for how feedbacks depend on changing *spatial patterns* of sea-surface temperature (SST), a phenomenon known as the SST "pattern effect" (Senior and Mitchell, 2000; Armour et al., 2013; Andrews et al., 2015; Zhou et al., 2016; Stevens

et al., 2016; Ceppi and Gregory, 2017; Andrews and Webb, 2018; Dong et al., 2019; Fueglistaler, 2019).

A robust understanding of the SST pattern effect has been developed in the context of recent warming. Over the past century, SSTs have warmed more in the tropical west Pacific and less in the east Pacific and Southern Ocean (Dong et al., 2019; Andrews et al., 2018, 2022). SST changes in tropical regions of deep convection (e.g., the west Pacific) produce strongly negative (stabilizing) feedbacks, whereas SST changes in regions with reflective low clouds (e.g., the east Pacific) or sea ice produce relatively positive (destabilizing) feedbacks (Zhou et al., 2016; Dong et al., 2019; Andrews and Webb, 2018; Fueglistaler, 2019; Ceppi and Gregory, 2017; Zhou et al., 2017). This historical pattern of SST trends is expected to reverse in the future as the tropical east Pacific and Southern Ocean eventually warm at higher rates, producing more-positive feedbacks and a more-sensitive climate (Ceppi and Gregory, 2017; Armour et al., 2016; Dong et al., 2020). Accounting for pattern effects causes the historical record to become a weak constraint on high values of ECS (Sherwood et al., 2020; Forster et al., 2021; Andrews et al., 2018, 2022), leaving paleoclimate evidence, especially from the Last Glacial Maximum (LGM), as a leading constraint on the ECS upper bound (Sherwood et al., 2020).

With the advent of paleoclimate data assimilation (Hakim et al., 2016), spatially complete reconstructions of SST and sea-ice concentration (SIC) now exist for the leading two paleoclimate periods used as ECS constraints (Sherwood et al., 2020): the cold LGM (19–23 thousand years ago) and the Pliocene's mid-Piacenzian warm period (3.0–3.3 million years ago). Data assimilation (DA) uses the spatial covariance from climate models to spread information from sparse observations (e.g., ocean sediments, ice cores, and pollen records) across the spatially complete estimate of the climate state. Importantly, uncertainty in the prior and in the observations is propagated to the posterior state. In this dissertation, we address whether the recent estimates of paleoclimate SST patterns for the LGM (Tierney et al., 2020; Osman et al., 2021; Amrhein et al., 2018; Annan et al., 2022) and Pliocene (Tierney et al., 2025b; Annan et al., 2024) lead to stronger or weaker constraints on modern-day ECS.

While DA has been used to reconstruct paleoclimates, applying DA to the recent historical record (1850–present) is a research frontier that will be addressed in this dissertation. Recent studies have emphasized that various historical SST and SIC datasets, which employ a variety

of statistical infilling methods, have trends that differ in their spatial patterns and hence their radiative feedbacks over the historical record (Modak and Mauritsen, 2023; Lewis and Mauritsen, 2021; Andrews et al., 2022, 2018; Fueglistaler and Silvers, 2021; Forster et al., 2021). Differences in the relative pace of warming between the warmest (deep convective) and coldest (subsiding) regions of the tropics can have a strong impact on cloud feedbacks and climate sensitivity (Fueglistaler, 2019; Hartmann and Dygert, 2022), hence past patterns of change across the tropics have been highlighted as a key uncertainty in understanding feedbacks over the historical record (Fueglistaler and Silvers, 2021). Alternatively, uncertainty in the magnitude of variability and change across the Southern Ocean, especially in Antarctic sea ice (Roach et al., 2020; Edinburgh and Day, 2016), may be underappreciated: Andrews et al. (2018) found that SST/SIC datasets with substantially different preindustrial concentrations of Antarctic sea ice produced global-mean net radiative feedbacks that differed by approximately 0.5 W m^{-2} K^{-1}. Other recent studies have highlighted the Southern Ocean and (its teleconnections to the tropics) as a major influence on global feedbacks and climate sensitivity (Kang et al., 2023c,a; Hartmann, 2022; Dong et al., 2022a,b). These findings demonstrate a need for better quantification of the uncertainty in historical feedbacks arising from uncertainty in patterns of past temperature change.

While there are numerous existing datasets for historical SST (Huang et al., 2017; Hurrell et al., 2008; Titchner and Rayner, 2014; Cowtan and Way, 2014; Rayner et al., 2003; Hirahara et al., 2014; Vaccaro et al., 2021) and sea ice (Rayner et al., 2003; Hurrell et al., 2008; Titchner and Rayner, 2014; Walsh et al., 2017; Brennan and Hakim, 2022), none have used coupled data assimilation (i.e., updating the atmosphere and ocean simultaneously by accounting for covariance across components) to incorporate observations of: SST from ships of opportunity (Kennedy et al., 2019), 2-m air temperature (T) from terrestrial stations (Osborn et al., 2021), and sea-level pressure (SLP) from marine data (Freeman et al., 2017). We aim to incorporate these SST, SLP, and T observations in a new, monthly reconstruction of historical SST and SIC, quantifying uncertainty in radiative feedbacks over the historical record.

This dissertation includes three chapters representing three publications. Of the studies presented in the following chapters, Chapter 2 is published, Chapter 3 is in review, and Chapter 4 is in re-review. A publication related to the M.S. degree (Cooper et al., 2022) is not included in the dissertation. Chapter 2 quantifies the pattern effect in the LGM and implications for modern

ECS. Chapter 3 quantifies pattern effects for the Pliocene, synthesizes the Pliocene results with the LGM, and provides stronger constraints on modern climate sensitivity and 21st-century warming. Chapter 4 reconstructs the historical climate record over 1850–2023, thus providing constraints on SST and sea ice that will be used to estimate historical feedbacks and pattern effects in a separate study (in preparation).

Chapter 2

LAST GLACIAL MAXIMUM PATTERN EFFECTS REDUCE CLIMATE SENSITIVITY ESTIMATES

2.1   Abstract

Here we show that the Last Glacial Maximum (LGM) provides a stronger constraint on equilibrium climate sensitivity (ECS)—the global-mean warming from increasing greenhouse-gas concentrations—after temperature patterns are accounted for. Feedbacks governing ECS depend on spatial patterns of surface temperature ("pattern effects"); hence, using the LGM to constrain future warming requires quantifying how temperature patterns produce different feedbacks during LGM cooling compared to modern-day warming. Combining data-assimilation LGM reconstructions with atmospheric models, we show that the climate is more sensitive to LGM forcing because ice sheets amplify temperature changes in the extratropics, where feedbacks are destabilizing. After accounting for LGM pattern effects, we find a best estimate (median) for modern-day ECS of 2.4 °C, with a 66% likely range of 1.7–3.5 °C (1.4–5.0 °C, 5 to 95%), based on LGM evidence alone. Combining the LGM with other lines of evidence, the combined best estimate becomes 2.9 °C, with a 66% likely range of 2.4–3.5 °C (2.1–4.1 °C, 5 to 95%), substantially narrowing uncertainty compared to recent assessments.

2.2   Introduction

Equilibrium climate sensitivity (ECS) is the steady-state response of global-mean near-surface air temperature to doubling atmospheric CO_2 above pre-industrial levels. ECS is a focus of climate

policy and projections because it governs Earth's long-term response to anthropogenic greenhouse gas changes (Sherwood et al., 2020; Forster et al., 2021). Recently, the World Climate Research Programme's 2020 climate sensitivity assessment, hereafter "WCRP20" (Sherwood et al., 2020), updated the 66% likely range for ECS to 2.6–3.9°C (2.3–4.7°C, 5–95%) with a central estimate of 3.1°C, which informed the likely range of 2.5–4.0°C (2.0–5.0°C, very likely) and central estimate of 3°C in the Intergovernmental Panel on Climate Change Sixth Assessment Report ("IPCC AR6") (Forster et al., 2021). This narrowing of uncertainty compared to previous assessments was achieved by quantitatively combining evidence from process understanding of climate feedbacks, observations over the historical record (1870–present), and paleoclimate reconstructions of past cold and warm periods. Of these lines of evidence, paleoclimate data from the Last Glacial Maximum (LGM), approximately 21,000 years ago, provide a leading constraint on the upper bound of ECS (Sherwood et al., 2020; Forster et al., 2021; Tierney et al., 2020).

Using paleoclimate data to constrain modern-day ECS requires accounting for how climate feedbacks change across different climate states (Sherwood et al., 2020; Forster et al., 2021; Manabe and Bryan, 1985; PALAEOSENS Project Members, 2012; Köhler et al., 2015; von der Heydt et al., 2016; Friedrich et al., 2016; Rohling et al., 2018). The standard assumption is that colder climates are less sensitive (i.e., have more-negative feedbacks) than warmer states (Sherwood et al., 2020; Forster et al., 2021; PALAEOSENS Project Members, 2012; Köhler et al., 2015; von der Heydt et al., 2016; Friedrich et al., 2016; Rohling et al., 2018). However, the simple assumption that feedbacks change with global-mean temperature does not account for how feedbacks depend on changing spatial patterns of sea-surface temperature (SST), a phenomenon known as the SST "pattern effect" (Armour et al., 2013; Zhou et al., 2016; Dong et al., 2019; Andrews and Webb, 2018; Fueglistaler, 2019; Ceppi and Gregory, 2017).

A robust understanding of the SST pattern effect has been developed in the context of recent warming. Over the past century, SSTs have warmed more in the tropical west Pacific and less in the east Pacific and Southern Ocean (Dong et al., 2019; Andrews et al., 2018, 2022). SST changes in tropical regions of deep convection (e.g., the west Pacific) produce strongly negative (stabilizing) feedbacks, whereas SST changes in regions with reflective low clouds (e.g., the east Pacific) or sea ice produce relatively positive (destabilizing) feedbacks (Zhou et al., 2016; Dong et al., 2019; Andrews and Webb, 2018; Fueglistaler, 2019; Ceppi and Gregory, 2017; Zhou et al., 2017). This

transient pattern of SST trends is expected to reverse in the future as the tropical east Pacific and Southern Ocean eventually warm at higher rates, producing more-positive feedbacks and a more-sensitive climate at equilibrium (Ceppi and Gregory, 2017; Armour et al., 2016; Dong et al., 2020). Accounting for this transient pattern effect causes the historical record to become a weak constraint on high values of ECS (Sherwood et al., 2020; Forster et al., 2021; Andrews et al., 2018, 2022; Proistosescu and Huybers, 2017), leaving the LGM as a leading constraint on the ECS upper bound (Sherwood et al., 2020).

However, pattern effects have not been accounted for in LGM evidence for modern-day ECS (Sherwood et al., 2020; Forster et al., 2021; Tierney et al., 2020; PALAEOSENS Project Members, 2012; Renoult et al., 2023). Importantly, if the spatial pattern of SST change in equilibrium at the LGM differs from the pattern of future warming, then the climate feedbacks governing climate sensitivity will differ as well. Continental ice sheets are responsible for approximately half of the total LGM forcing (Tierney et al., 2020; Zhu and Poulsen, 2021; Braconnot and Kageyama, 2015) and drive distinct climate responses from changes in topography, albedo, and sea level (Zhu and Poulsen, 2021; Manabe and Broccoli, 1985; Cook and Held, 1988; Lee et al., 2015; DiNezio et al., 2018; Roberts et al., 2019; Amaya et al., 2022), suggesting that patterns of SST change at the LGM may differ substantially from those in response to a modern-day doubling of CO_2. Previous work acknowledged this possibility (Sherwood et al., 2020; Forster et al., 2021) but did not account for LGM pattern effects because no quantification had yet been made. A key question is: would accounting for LGM pattern effects strengthen or weaken constraints on modern-day ECS?

Here we provide the first quantification of the LGM pattern effect and its uncertainty by leveraging two recent advances. First, with the advent of paleoclimate data assimilation (Hakim et al., 2016), spatially complete reconstructions of SST and sea ice now exist for the LGM (Tierney et al., 2020; Osman et al., 2021; Annan et al., 2022; Amrhein et al., 2018), including estimated uncertainties. Second, recent progress in quantifying pattern effects (Andrews et al., 2018, 2022) provides methods using atmospheric general circulation models (AGCMs) to link SST patterns to climate feedbacks. These advances present an opportunity to compare SST changes at the LGM with those expected under anthropogenic CO_2 forcing and to quantify resulting differences in climate feedbacks and sensitivity. To assess the robustness of our results, we use five AGCMs (sampling uncertainty in how feedbacks relate to SST patterns) and four reconstructions (Tierney et al., 2020;

Osman et al., 2021; Annan et al., 2022; Amrhein et al., 2018) of the LGM (sampling uncertainty in SST patterns).

2.3 Dependence of modern-day ECS on pattern effects

ECS and climate feedbacks are connected through the standard model of global-mean energy balance:

$$\Delta N = \lambda \Delta T + \Delta F, \tag{2.1}$$

where N is the top-of-atmosphere radiative imbalance; λ is the net climate feedback (negative for stable climates); T is the near-surface air temperature; and F is the "effective" radiative forcing, i.e., the change in net downward radiative flux after adjustments to imposed perturbations but excluding radiative responses to changing surface temperature (Sherwood et al., 2020; Forster et al., 2021). Differences (Δ) are relative to an equilibrium reference state, e.g., the pre-industrial period.

When the forcing is a CO_2-doubling ($2\times CO_2$) of pre-industrial values, and the climate system reaches equilibrium ($\Delta N = 0$), the resulting ΔT is referred to as the ECS:

$$\text{ECS} = -\frac{\Delta F_{2\times}}{\lambda_{2\times}}, \tag{2.2}$$

where $\Delta F_{2\times}$ is the effective radiative forcing, and $\lambda_{2\times}$ is the net feedback for $2\times CO_2$. More-negative values of $\lambda_{2\times}$ indicate a less-sensitive climate (lower ECS).

Here we aim to quantify the difference in feedbacks ($\Delta\lambda$) operating in the modern climate under $2\times CO_2$ ($\lambda_{2\times}$) and at the LGM (λ_{LGM}):

$$\Delta\lambda = \lambda_{2\times} - \lambda_{\text{LGM}}. \tag{2.3}$$

Following recent research on pattern effects in the historical record (Sherwood et al., 2020; Andrews et al., 2018, 2022), we estimate $\lambda_{2\times}$ and λ_{LGM} using AGCM simulations with SST and sea-ice concentration (SIC) prescribed as surface boundary conditions. We further evaluate the contributions to $\Delta\lambda$ from pattern effects and global-mean temperature changes between the LGM and $2\times CO_2$.

To infer the modern-day ECS from LGM evidence, Equations (2.2) and (2.3) can be combined (Sherwood et al., 2020; Andrews et al., 2018) to yield

$$\text{ECS} = -\frac{\Delta F_{2\times}}{\lambda_{\text{LGM}}^* + \Delta\lambda}, \tag{2.4}$$

where λ^*_{LGM} is the estimate of the unadjusted LGM feedback (determined using Eq. (2.1) applied to that state), which we take from previous assessments (Sherwood et al., 2020; Forster et al., 2021; Tierney et al., 2020), and $\Delta\lambda$ is estimated from our AGCM simulations. The value of $\Delta\lambda$ depends on spatial patterns of LGM SST and SIC anomalies, for which we use state-of-the-art reconstructions (Tierney et al., 2020; Osman et al., 2021; Annan et al., 2022; Amrhein et al., 2018) based on data assimilation.

2.4   Using data-assimilation reconstructions to quantify pattern effects

Similar to Bayesian statistics, paleoclimate data assimilation (Hakim et al., 2016) begins with a "prior" estimate of the climate state from model ensembles. Proxy data provide indirect climate observations that update the prior, balancing relative error in the prior and the observations. This results in a "posterior" state estimate, constrained by observations and accounting for uncertainty in priors and data. Since the posterior is sensitive to priors (Amrhein et al., 2020; Parsons et al., 2021), proxies, and methods, we sample this uncertainty by using multiple reconstructions.

Figure 2.1 shows the four SST reconstructions (see Materials and Methods) we use to quantify the LGM pattern effect. All four reconstructions have a prominent common feature: amplified extratropical cooling in both the North Pacific and North Atlantic Oceans. While the LGM reconstructions differ in other regions that are important for climate feedbacks, e.g., the tropical Pacific (Zhou et al., 2016; Dong et al., 2019; Andrews and Webb, 2018; Fueglistaler, 2019; Ceppi and Gregory, 2017) and Southern Ocean (Armour et al., 2016; Rose et al., 2014; Kang and Xie, 2014), their robust agreement in the northern extratropics proves to be essential for the LGM pattern effect. The zonally consistent maximum near 40°N in SST anomalies at the LGM is in strong contrast to the near-equilibrium response to modern-day $2\times CO_2$ (Fig. 2.1F, Fig. 2.S1) as simulated by climate models in LongRunMIP (Rugenstein et al., 2019) (see Materials and Methods), suggesting the potential for feedbacks to differ between LGM and $2\times CO_2$ climates.

Using data-constrained patterns to quantify how LGM feedbacks compare to feedbacks in $2\times CO_2$ is an advance over past comparisons (all based on models), which have produced conflicting results (Renoult et al., 2023; Zhu and Poulsen, 2021; Crucifix, 2006; Yoshimori et al., 2011; Stap et al., 2019; Shakun, 2017; Hopcroft and Valdes, 2015) (see Text S1). While our method overcomes the problem of unconstrained SST patterns from coupled atmosphere–ocean simulations
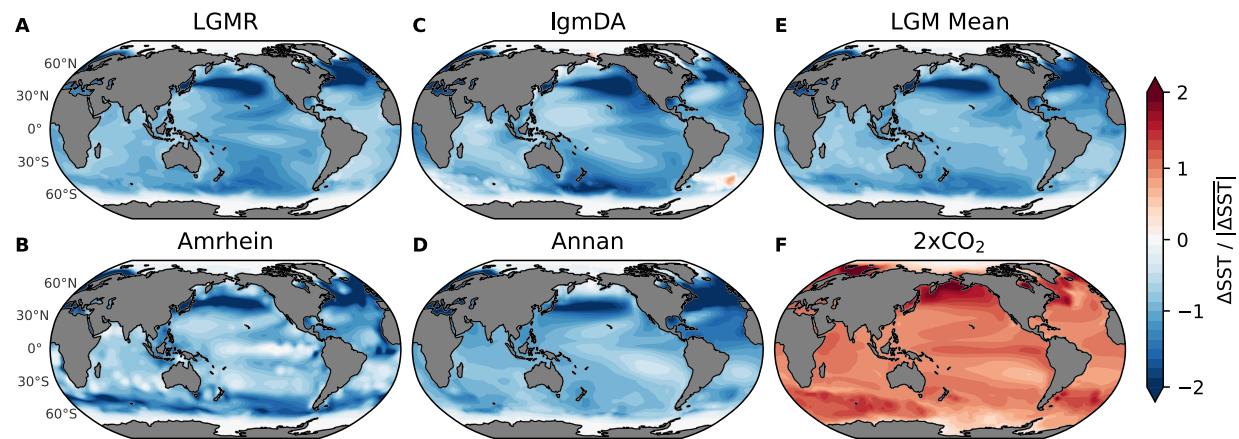
Figure 2.1: Patterns of sea-surface-temperature (SST) anomalies from data assimilation at the Last Glacial Maximum (LGM) compared to modern-day doubling of CO_2 ($2\times CO_2$). LGM reconstructions include (A) Last Glacial Maximum Reanalysis (LGMR) (Osman et al., 2021), (B) Amrhein et al. (2018), (C) lgmDA (Tierney et al., 2020), (D) Annan et al. (2022), and (E) the mean of the four LGM patterns. (F) Pattern of the multi-model mean from near-equilibrium $2\times CO_2$ simulations in LongRunMIP (Rugenstein et al., 2019), initialized from pre-industrial control. To show SST patterns, local SST anomalies are divided by absolute values of global-mean SST anomalies (consistent with feedbacks being radiative responses divided by temperature anomalies). All panels show annual means. LGM reconstructions are infilled to modern coastlines (see Materials and Methods).

of the LGM, we still rely on AGCMs to estimate feedbacks and their uncertainties.

We calculate net feedbacks using AGCMs with prescribed SST and SIC boundary conditions. We first conduct AGCM simulations with a "baseline" pattern representing the pre-industrial climate, for which we use SST and SIC in the Late Holocene (mean of 0–4,000 years ago) from the Last Glacial Maximum Reanalysis (LGMR) (Osman et al., 2021). We then perform AGCM simulations with SST and SIC boundary conditions (see Materials and Methods) from $2\times CO_2$ in LongRunMIP (Rugenstein et al., 2019) and the four LGM reconstructions (Tierney et al., 2020; Osman et al., 2021; Annan et al., 2022; Amrhein et al., 2018) (SST in Fig. 2.1; SIC in Fig. 2.S2).

Finally, we calculate global-mean ΔN and ΔT in each $2\times CO_2$ and LGM simulation relative to the baseline, which yields net feedbacks as $\lambda = \Delta N/\Delta T$ using Eq. (2.1). All forcings are held constant ($\Delta F = 0$) at modern-day levels across our AGCM simulations; therefore, all changes in simulated top-of-atmosphere radiation and feedbacks can be attributed solely to SST/SIC differences (see Materials and Methods).

We find that $\lambda_{2\times}$ is more negative (stabilizing) than λ_{LGM}, indicating that the climate system is more sensitive to LGM forcing than to $2\times CO_2$ (Fig. 2.2). We use the LGMR pattern (Fig. 2.1A) in five AGCMs (CAM4, CAM5, CAM6, GFDL-AM4, and HadGEM3-GC3.1-LL) to evaluate uncertainty from atmospheric model physics, and we use all four LGM reconstructions (Fig. 2.1A–D) in CAM4 and CAM5 to evaluate uncertainty from LGM patterns.

This approach is supported by the result that AGCMs tend to reproduce observed relationships between SSTs and top-of-atmosphere radiation when observed SST patterns are prescribed (Allan et al., 2014; Loeb et al., 2020). The LGM pattern effect, $\Delta\lambda$ in Eq. (2.3), is negative across all five AGCMs and all four LGM reconstructions. The five AGCMs produce a mean $\Delta\lambda = -0.40$ W m^{-2} K^{-1} (Fig. 2.2B; detailed results in Tables 2.S1–S2). We also evaluate uncertainty in the $2\times CO_2$ pattern but find that this is of secondary importance (see Materials and Methods; Figs. 2.S3–S4).

Our main result is that the climate is more sensitive to LGM forcing than it is to modern-day $2\times CO_2$ forcing ($\Delta\lambda < 0$), implying lower estimates of modern-day ECS by Eq. (2.4), and this finding is robust despite uncertainties in atmospheric physics and LGM reconstructions.
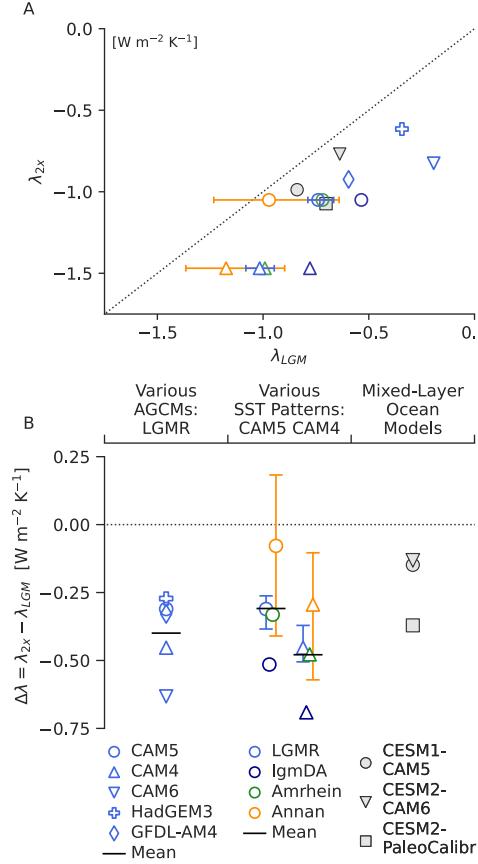
Figure 2.2: Last Glacial Maximum (LGM) and $2\times CO_2$ climate feedbacks and LGM pattern effect ($\Delta\lambda$). Different atmospheric general circulation models (AGCMs), all using the LGMR pattern for the LGM, are indicated by symbols; different LGM patterns (in CAM5 and CAM4) are indicated by colors. Error bars for Annan and LGMR represent 1st and 4th quartiles of ensemble members (Materials and Methods); central values indicate ensemble mean. For comparison with AGCM results using LGM data assimilation, the following feedbacks (in mixed-layer ocean coupled to AGCM) from previous studies are also included: CESM1-CAM5 (Zhu and Poulsen, 2021), CESM2-CAM6 (Zhu et al., 2021), and CESM2-PaleoCalibr (Zhu et al., 2022) (modified version of CAM6). (A) Scatter plot of $2\times CO_2$ feedbacks, λ_{2x}, versus LGM feedbacks, λ_{LGM}, with $\lambda_{2x} = \lambda_{LGM}$ shown as dotted line. (B) LGM pattern effect, $\Delta\lambda = \lambda_{2x} - \lambda_{LGM}$, using feedbacks shown in panel A, with $\Delta\lambda = 0$ shown as dotted line. Note that $\Delta\lambda$ includes SST pattern effects and contributions from temperature dependence.

2.5 Discussion

2.5.1 Physical mechanisms driving LGM pattern effects

For comparison with our feedbacks in AGCMs driven by LGM reconstructions, we examine previously published results (Zhu and Poulsen, 2021) from AGCMs coupled to mixed-layer "slab" oceans (Fig. 2.2), which allow SST changes in response to imposed forcings but exclude changes in ocean dynamics (Bitz et al., 2012). These mixed-layer-model versions of CESM1-CAM5 (Zhu and Poulsen, 2021), CESM2-CAM6 (Zhu et al., 2021), and CESM2-PaleoCalibr (Zhu et al., 2022) (using a modified CAM6), which differ from our AGCM experiments by including forcings from ice sheets and greenhouse gases, also produce $\Delta\lambda < 0$.

Although disagreements in simulated SST patterns compared to proxy data suggest that free-running coupled models cannot reliably estimate the value of $\Delta\lambda$, the coupled models point to mechanisms driving $\Delta\lambda$ that are consistent with the reconstructions and our AGCM simulations. In this section, we begin by reviewing simulations in coupled models that demonstrate the physical mechanisms linking patterns of forcing, SST response, and climate feedbacks.

First, compare zonal-mean patterns of effective radiative forcing and SST changes from CESM1-CAM5 simulations (Zhu and Poulsen, 2021) under three forcing scenarios: $2\times CO_2$ forcing, LGM forcing (ice sheets and greenhouse gases), and LGM ice-sheet forcing alone (including coastline changes). The localized ice-sheet forcing causes the amplified SST response in the northern extratropics at the LGM compared to $2\times CO_2$ (Fig. 2.3A–C).

Explaining the Northern Hemisphere's response to LGM ice sheets has been a focus of previous studies, which found that amplified SST cooling in the northern extratropics is associated with changes in atmospheric stationary waves, driven by changes in ice-sheet albedo and topography (Zhu and Poulsen, 2021; Roberts et al., 2019; Amaya et al., 2022; Roe and Lindzen, 2001). Differences in SST responses between LGM and $2\times CO_2$ persist at quasi-equilibrium in a fully coupled (atmosphere–ocean GCM) version of CESM1-CAM5 (Fig. 2.3C; Fig. 2.S5). Comparing the fully coupled model's response (Fig. 2.3C) to LGM forcing with the data-assimilation patterns (Fig. 2.3D) we use to quantify pattern effects supports the finding that LGM ice sheets amplify SST cooling in the northern extratropics (Zhu and Poulsen, 2021; Roberts et al., 2019; Amaya et al., 2022), but this cooling pattern is more pronounced in proxy reconstructions. The amplified cooling

14
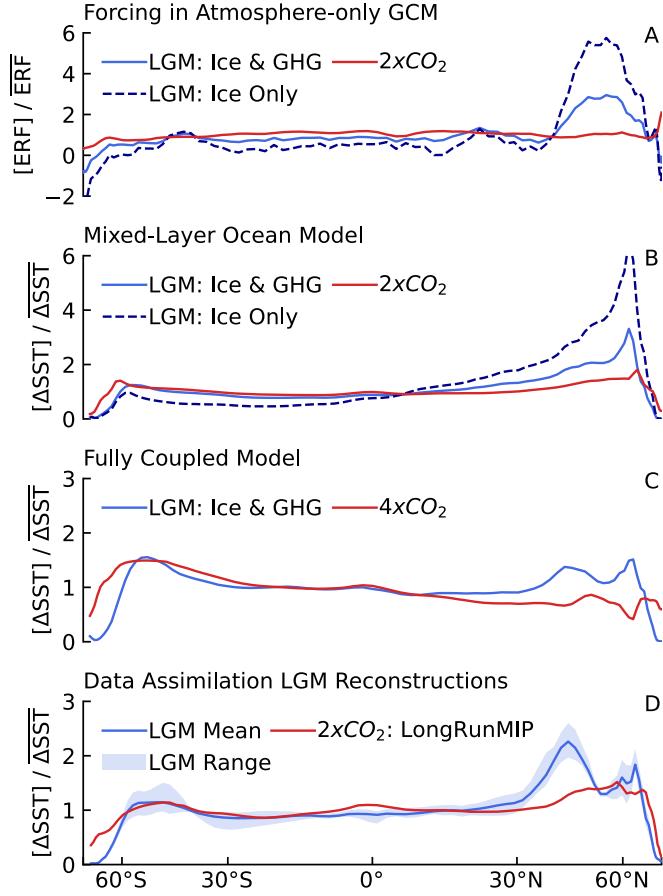
Figure 2.3: Zonal-mean patterns of effective radiative forcing (ERF) and sea-surface-temperature (SST) anomalies. All anomalies are normalized through division by global-mean anomalies. (A–C) Model simulations in CESM1-CAM5 from Zhu & Poulsen (Zhu et al., 2021). (A) ERF directly from three fixed-SST simulations using atmospheric general circulation model with LGM greenhouse-gas (GHG) and ice-sheet (Ice) forcing, $2\times CO_2$, and LGM ice-sheet forcing alone (Zhu et al., 2021) (including coastline changes). (B) Equilibrium SST patterns, corresponding to panel A, in coupled mixed-layer ocean model. (C) Quasi-equilibrium SST patterns from fully coupled atmosphere-ocean model, comparing LGM forcings (Zhu et al., 2021) with abrupt-$4\times CO_2$ forcing (Zhu et al., 2019); no long-run $2\times CO_2$ simulation is available. Note vertical-axis scales. (D) Mean and range of SST patterns from four data-assimilation reconstructions (Tierney et al., 2020; Osman et al., 2021; Annan et al., 2022; Amrhein et al., 2018) of the LGM compared to $2\times CO_2$ multi-model mean from LongRunMIP (Rugenstein et al., 2019).

of extratropical SST, caused by ice-sheet forcing, causes the LGM feedback to be less stabilizing than the feedback induced by CO_2 forcing alone.

Decomposing λ from our AGCM simulations into component feedbacks (Fig. 2.S6), including results from direct model output and from radiative kernels (see Materials and Methods), shows that shortwave cloud feedbacks are responsible for much of the negative value of $\Delta\lambda$ and for much of the spread across AGCMs. The combined feedback from changes in lapse rate and water vapor also contributes to negative values of $\Delta\lambda$. While shortwave clear-sky feedbacks from sea ice and snow are also more positive for the LGM, cloud masking strongly damps the impact of those LGM feedbacks.

Accounting for cloud masking (Soden et al., 2008; Raghuraman et al., 2023), feedbacks from surface albedo are more positive in 2×CO_2, i.e., contribute a positive $\Delta\lambda$, offsetting the negative total $\Delta\lambda$. Overall, our results align with the previous studies focused on the historical record that emphasize cloud and lapse-rate feedbacks in pattern effects (Zhou et al., 2016; Andrews and Webb, 2018; Ceppi and Gregory, 2017; Dong et al., 2020).

Spatial distributions of feedbacks (Fig. 2.S7) clarify the connection between ice-sheet forcing, SST response, and cloud feedbacks. Where the SST cooling from LGM ice sheets is amplified in the North Pacific and North Atlantic, positive shortwave cloud feedbacks are prominent due to increases in reflective low clouds (Zhou et al., 2016; Andrews and Webb, 2018; Ceppi and Gregory, 2017; Dong et al., 2019; Fueglistaler, 2019; Zhou et al., 2017; Amaya et al., 2022). Compared to 2×CO_2 simulations, LGM reconstructions have relatively small SST anomalies in tropical ascent regions (Fig. 2.S1) where feedbacks are most negative (Zhou et al., 2016; Andrews and Webb, 2018; Dong et al., 2019; Fueglistaler, 2019; Zhou et al., 2017; Kang and Xie, 2014). However, tropical patterns at the LGM differ across reconstructions, adding to the uncertainty in the LGM pattern effect.

Despite these differences in the tropics, all four reconstructions produce a negative pattern effect due to the robust amplification of cooling in the northern extratropics. The role of the northern extratropics illustrates that pattern effects are not always dominated by the tropical Pacific, distinguishing the LGM pattern effect from the well-studied pattern effect of the historical period. In summary, the LGM SST pattern produces a less-negative global climate feedback compared to the 2×CO_2 SST pattern and $\Delta\lambda < 0$.

2.5.2   Separating pattern effects from temperature dependence of feedbacks

While our explanation for feedback differences between LGM and $2\times CO_2$ forcing focuses on SST pattern differences, we also estimate how $\Delta\lambda$ is affected by global-mean temperature within our AGCM simulations. Our main AGCM simulations (Fig. 2.2), which determine our estimate of total $\Delta\lambda$, include not only the impact of SST patterns on feedbacks (pattern effects) but also differences in feedbacks caused by other asymmetries between LGM cooling and modern-day warming under $2\times CO_2$ forcing (temperature dependence). We consider that

$$\Delta\lambda \approx \Delta\lambda_{\text{PatternOnly}} + \Delta\lambda_T, \qquad (2.5)$$

where $\Delta\lambda_{\text{PatternOnly}}$ is the feedback change due to different patterns of SST anomalies and $\Delta\lambda_T$ is the feedback change due to different global-mean temperatures (T). Recent community assessments (Sherwood et al., 2020; Forster et al., 2021) assume warmer climates are more sensitive ($\Delta\lambda_T > 0$) (PALAEOSENS Project Members, 2012; Köhler et al., 2015; von der Heydt et al., 2016; Friedrich et al., 2016; Rohling et al., 2018; Yoshimori et al., 2011), which is at odds with the total $\Delta\lambda < 0$ we find for the LGM in AGCMs and coupled models (Fig. 2.2).

To separate pattern effects from temperature dependence, we perform additional "pattern-only" simulations in CAM4, CAM5, and CAM6 using the LGMR and $2\times CO_2$ patterns. For these simulations, we multiply local SST anomalies by constant scaling factors to yield global-mean ΔSST $= -0.5$ K with constant baseline SIC (see Materials and Methods). SST scaling preserves spatial patterns of anomalies but forces global-mean ΔT to be small and equal across simulations, i.e., $\Delta\lambda_T \approx 0$ in the pattern-only simulations. We then repeat the feedback calculations, computing $\Delta\lambda_{\text{PatternOnly}}$ as in Eq. (2.3).

We estimate the temperature dependence $\Delta\lambda_T$ as the residual difference between the main and pattern-only AGCM simulations, rearranging Eq. (2.5) to $\Delta\lambda_T \approx \Delta\lambda - \Delta\lambda_{\text{PatternOnly}}$ (see Materials and Methods). We note that ice-albedo contributions to $\Delta\lambda$ could arise from SST patterns or temperature dependence, but our partitioning of $\Delta\lambda$ treats sea ice as part of $\Delta\lambda_T$.

The magnitude and sign of $\Delta\lambda_T$ is found to be model-dependent, in agreement with recent multi-model assessments (Renoult et al., 2023; Bloch-Johnson et al., 2021), but $\Delta\lambda_T$ appears to be positive and directionally consistent with standard assumptions (Sherwood et al., 2020; Forster et al., 2021) for feedback temperature dependence. However, $\Delta\lambda_{\text{PatternOnly}}$ is negative and larger

than $\Delta\lambda_T$ such that total $\Delta\lambda < 0$ in each AGCM (Fig. 2.S8, Table 2.S3). These results suggest that total $\Delta\lambda$ for the LGM is mostly attributable to SST pattern effects, and $\Delta\lambda_T$ plays a smaller role over this range of climates.

Recent assessments (Sherwood et al., 2020; Forster et al., 2021) considered $\Delta\lambda_T$ for the LGM but did not account for the larger, opposing term, $\Delta\lambda_{\mathrm{PatternOnly}}$. The substantial LGM pattern effect found here motivates revising the LGM evidence for modern-day ECS.

2.5.3   Climate sensitivity accounting for LGM pattern effects

Constraining modern-day ECS with paleoclimate evidence requires accounting for how forcings and feedbacks differ in paleoclimates relative to the modern-day $2\times CO_2$ scenario (Sherwood et al., 2020; Forster et al., 2021; PALAEOSENS Project Members, 2012). LGM inferences of ECS begin with applying Eq. (2.1) to the LGM in equilibrium, estimating the unadjusted LGM feedback as $\lambda^*_{\mathrm{LGM}} = -\sum \Delta F/\Delta T$.

Effective radiative forcings (ΔF) include not only CO_2 but also ice sheets (including sea level) and, depending on the timescale chosen for ECS (Sherwood et al., 2020; Forster et al., 2021; Tierney et al., 2020; PALAEOSENS Project Members, 2012), additional changes that have distinct impacts at the LGM: vegetation, dust, N_2O, and CH_4 (see Materials and Methods). Finally, λ^*_{LGM} must be adjusted for differences in feedbacks ($\Delta\lambda$) relative to those operating in modern-day $2\times CO_2$, following Eq. (2.4).

Our results suggest that the LGM feedback is more positive than the $2\times CO_2$ feedback because of the LGM ice-sheet forcing and resulting SST pattern. Failing to account for this difference in feedbacks would lead to the inference of higher values of modern-day ECS from the LGM, e.g., Hansen et al. (2023). Some past studies using fully coupled models have considered these feedback differences indirectly by applying an "efficacy" adjustment (Yoshimori et al., 2009) to the LGM forcings.

The efficacy framework has led to disparate results for multiple reasons: changes in how forcing is quantified (Crucifix, 2006; Yoshimori et al., 2011; Kageyama et al., 2021) before ERF became standard (Forster et al., 2021); the lack of data constraints on SST patterns simulated by fully coupled models (Renoult et al., 2023; Hopcroft and Valdes, 2015; Kageyama et al., 2021); and

the behavior of intermediate-complexity models with simplified cloud feedbacks (Stap et al., 2019; Shakun, 2017). Because efficacy is equivalent to the ratio of feedbacks $\lambda_{2\times}/\lambda_{\mathrm{LGM}}$ (Zhou et al., 2023; Webb et al., 2017), our results could be framed as a median LGM-forcing efficacy of 1.7 (see Materials and Methods; Tables 2.S1–S2), consistent with recent studies that find LGM-forcing efficacy greater than 1 using ERF and fully coupled models (Zhu and Poulsen, 2021; Zhu et al., 2021, 2022).

However, the pattern-effect framework we use replaces the need for forcing efficacy (Zhou et al., 2023) (see Text S1), aligns with modern AGCM methods of quantifying feedbacks (Webb et al., 2017) and ERF (Pincus et al., 2016), and incorporates data from the latest reconstructions of the LGM.

To demonstrate the impact of LGM pattern effects, we follow methods in WCRP20 (Sherwood et al., 2020) and focus on the 150-year timescale of climate sensitivity (S) applicable to modern warming (Sherwood et al., 2020; Forster et al., 2021) (see Materials and Methods). We use WCRP20 because that assessment uniquely allows updates of individual parameters and quantitatively combines lines of evidence, but our results would have the same directional impact on other assessments (Forster et al., 2021; Tierney et al., 2020).

We use forcing values from WCRP20 to estimate the unadjusted LGM feedback, λ_{LGM}^* in Eq. (2.4). However, given emerging evidence (Forster et al., 2021; Tierney et al., 2020; Osman et al., 2021; Seltzer et al., 2021; Liu et al., 2023) after WCRP20, we report results using a global temperature anomaly for the LGM of $\Delta T_{\mathrm{LGM}} = -6\pm1$ K in addition to WCRP20's value of -5 ± 1 K. We implement our key finding by updating the LGM $\Delta\lambda$, which includes LGM pattern effects for the first time. We assign a Normal distribution to $\Delta\lambda$, $\mathcal{N}(\mu = -0.37, \sigma = 0.23)$ W m^{-2} K^{-1}, reflecting spread across AGCMs and SST reconstructions (see Materials and Methods). Our assessment of $\Delta\lambda$ and its uncertainty relies on AGCMs to estimate feedbacks from prescribed SST/SIC patterns. We include additional uncertainty tests in Figs. 2.S4 and 2.S9, demonstrating that our general conclusions hold if the assumed σ for $\Delta\lambda$ is doubled.

Accounting for the LGM pattern effect reduces climate sensitivity inferred from the LGM evidence (Fig. 2.4). With $\Delta T_{\mathrm{LGM}} \approx -6$ K, maximum likelihood for S from the LGM evidence alone becomes 2.0 K (change of -1.3 K). Assuming a prior that is uniform in S from 0–20 K (see Materials and Methods) for the LGM evidence alone (Table 2.S4), we find a posterior median for modern-day
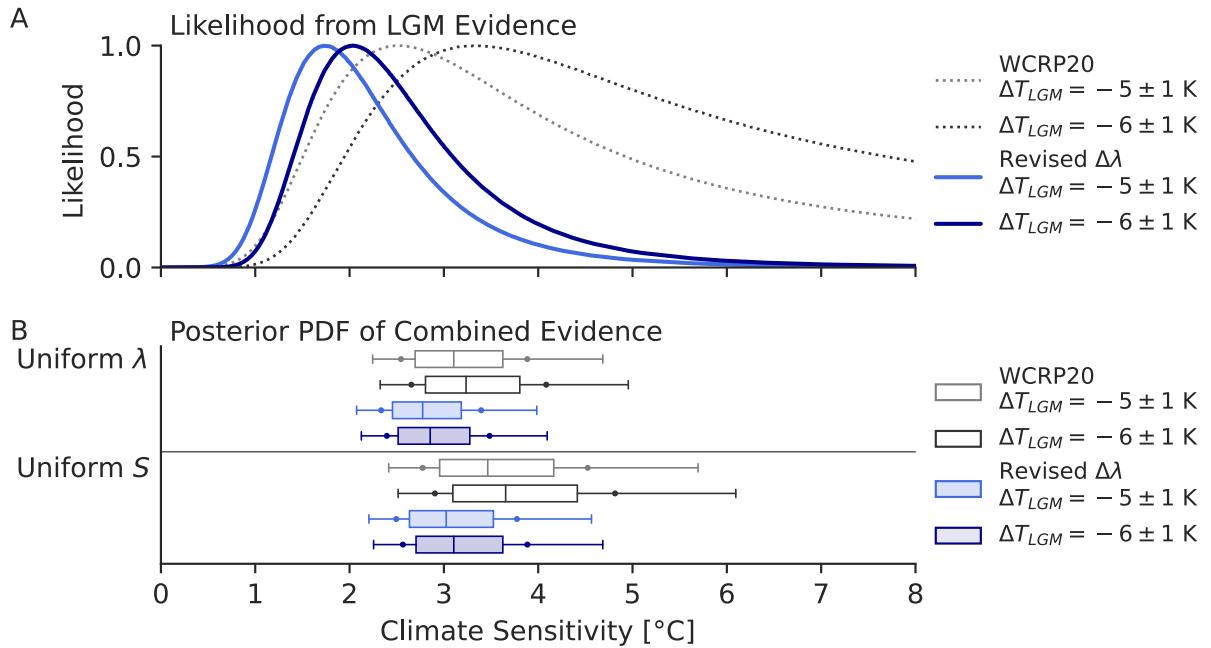
Figure 2.4: Inference of modern-day climate sensitivity including the LGM pattern effect. Results from WCRP20 (Sherwood et al., 2020) with no LGM pattern effects and original assumption of $\Delta T_{\mathrm{LGM}} \sim N(\mu = -5,\ \sigma = 1)$ °C (gray) and with revised $\Delta T_{\mathrm{LGM}} \sim N(-6,\ 1)$ °C (black) based on IPCC AR6 (Forster et al., 2021). Revised climate sensitivity including LGM pattern effects from this study (light and dark blue) assuming $\Delta \lambda \sim N(\mu = -0.37,\ \sigma = 0.23)$ W m^{-2} K^{-1}. Climate sensitivity shown is effective sensitivity (S) representing 150-year response, as in WCRP20 (Sherwood et al., 2020). (A) Likelihood functions for S based on only the LGM line of evidence. (B) Posterior PDF after combining LGM with other lines of evidence, assuming a uniform-λ prior (upper panel) or a uniform-S prior (lower panel). Outlier lines indicate 5–95th percentiles, dots indicate 66% likely range, and box indicates 25–75th percentiles and median.

ECS of 2.4°C, 66% likely range 1.7–3.5 K (1.4–5.0 K, 5–95%).

Combining the updated LGM evidence with existing likelihoods for the other lines of evidence (process understanding, historical record, and Pliocene) yields revised Bayesian probability distributions for the two priors in WCRP20: uniform in λ (WCRP20's "Baseline") and uniform in S (a robustness test).

The impact of the LGM pattern effect on the combined evidence is most pronounced on the upper bound of S, which has been notoriously difficult to constrain (Knutti and Hegerl, 2008). Assuming $\Delta T_{\mathrm{LGM}} \approx -6 \pm 1$ K, the median and 66% range from combining lines of evidence for S becomes 2.9 K (2.4–3.5 K) with a uniform-λ prior or 3.1 K (2.6–3.9 K) with a uniform-S prior. Corresponding 5–95% ranges are 2.1–4.1 K with uniform-λ and 2.3–4.7 K with uniform-S.

Accounting for pattern effects in $\Delta\lambda$ for the LGM thus reduces the central estimate of modern-day ECS by approximately 0.5 K and reduces the 66% range's upper bound by 0.6 K and 0.9 K for the uniform-λ and uniform-S priors, respectively, indicating substantially stronger constraints than WCRP20 (Sherwood et al., 2020) even after allowing for more glacial cooling. While the qualitative assessment in IPCC AR6 (Forster et al., 2021) cannot be quantitatively updated, these results suggest stronger constraints on modern-day ECS than assessed there, as well.

Accounting for LGM pattern effects—enabled by recent advances in LGM SST reconstruction using paleoclimate data assimilation and in quantifying pattern effects using atmospheric models—provides a tighter upper bound on modern-day ECS. While each line of evidence will surely evolve as scientific understanding improves, the results presented here demonstrate that pattern effects must be accounted for when inferring modern-day climate sensitivity from paleoclimate periods that are substantially affected by non-CO_2 forcing.

2.6   Materials and Methods

2.6.1   Data-assimilation reconstructions of the LGM

We use four LGM reconstructions to quantify the LGM pattern effect, sampling uncertainty across data assimilation methods and model priors (Amrhein et al., 2020; Parsons et al., 2021). Osman et al. (2021) produced the time-dependent Last Glacial Maximum Reanalysis ("LGMR") spanning the past 24,000 years; the SST and SIC fields that represent the LGM in their reanalysis are

time means spanning 19,000–23,000 years ago. Tierney et al. (2020) produced the state estimate "lgmDA" dataset.

Both the LGMR and lgmDA use priors from isotope-enabled simulations in iCESM1.2 and iCESM1.3 with assimilation of seasonal and annual SST proxies in an ensemble Kalman filter; there are differences in the proxy databases and methods between the two reconstructions. Annan et al. (2022) also used an ensemble Kalman filter but with a multi-model prior, including 19 ensemble members from a wide array of climate models spanning PMIP2 (launched in 2002) to PMIP4 (launched in 2017); they assimilated annual SST proxies and land-temperature proxies, and applied an adjustment to the prior ensemble to pre-center the prior around available proxy data.

Amrhein et al. (2018) fit the MITgcm ocean model to seasonal and annual SST proxies (Waelbroeck et al., 2009) using least-squares with Lagrange multipliers by adjusting prior atmospheric fields from a CCSM4 LGM simulation (Brady et al., 2013). While these approaches employ a diversity of DA methods, versions of CESM1-CAM5 form the prior for two of the reconstructions (Tierney et al., 2020; Osman et al., 2021), and the prior covariances could be biased by model errors. Moreover, archived proxy data are geographically inhomogeneous with strong preferences for the Northern Hemisphere and tropics; additional data could lead to greater SST agreement across reconstructions outside of the Northern Hemisphere.

2.6.2   Simulations with atmospheric general circulation models

SST/SIC boundary conditions (BCs) for the LGM, Late Holocene baseline, and $2\times CO_2$ are prepared to maintain constant forcing, i.e., $\Delta F = 0$ in Eq. (2.1), across simulations. Topography is held constant; that is, the LGM ice sheets are not present in AGCM simulations because their impact is already included as a forcing, and we are isolating feedbacks from changing SST/SIC. For the LGM and Late Holocene datasets, we adjust for differences relative to modern coastlines using kriging and extrapolation in polar regions. Details of sea-level adjustments are provided in Text S3.

The $2\times CO_2$ BC is the multi-model mean of 200 years from the end of six $2\times CO_2$ simulations, initialized from pre-industrial control states, in LongRunMIP (Rugenstein et al., 2019): CESM1.0.4 (years 2300–2500), CNRM-CM6-1 (years 550–750), HadCM3L (years 500–700), MPI-ESM-1.2 (years 800–1000), GFDL-ESM2M (years 4300–4500), and MIROC3.2 (years 1803–2003).

These simulations are near equilibrium but only represent an estimate of the true equilibrium SST response to $2\times CO_2$.

The Late Holocene, defined as the climatological mean of 0–4,000 years ago in the LGMR (Osman et al., 2021), is used as the baseline SST/SIC for all feedback calculations. This baseline represents a long-term mean of the pre-industrial climate, constrained by assimilation of proxy data. After adjusting for modern sea level, the four LGM BCs and the $2\times CO_2$ BC for SST are prepared by adding the SST anomalies from each of the four reconstructions to the Late Holocene baseline SST. Due to nonlinear behavior of sea ice, the LGM and $2\times CO_2$ BCs for SIC are not added to the baseline as anomalies but rather are used directly (Fig. 2.S2).

We run simulations with the Late Holocene baseline, $2\times CO_2$, and LGMR in each of five AGCMs. We run simulations with all four of the LGM reconstructions (LGMR, lgmDA, Amrhein, Annan) in CAM4 and CAM5, sampling the spread in LGM feedbacks from different reconstructions in two AGCMs with distinct relationships linking SST patterns to radiative feedbacks based on Green's functions (Dong et al., 2019; Zhou et al., 2017). Spin-up/analysis period/climatological forcing for each AGCM is 5 yr/25 yr/2000 (CESM1.2.2.1-CAM4 (Neale et al., 2013), CESM1.2.2.1-CAM5 (Neale et al., 2012), and CESM2.1-CAM6 (Danabasoglu et al., 2020) at $1.9°\times 2.5°$ latitude-by-longitude resolution); 5 yr/25 yr/2014 (HadGEM3-GC3.1-LL (Williams et al., 2017) at N96, $\sim$135-km resolution); and 1 yr/30 yr/2001 (GFDL-AM4 (Held et al., 2019) at C96, $\sim$100-km resolution). Parent coupled models of the AGCMs considered here sample a wide range of climate sensitivities, from 2.95 K to 5.54 K, and the AGCMs span a wide range of pattern effects in the historical record, from 0.38 W m^{-2} K^{-1} to 0.84 W m^{-2} K^{-1} (Andrews et al., 2022).

To compute λ, we take global means over the analysis periods for net top-of-atmosphere radiative imbalance (N) and near-surface air temperature (T). Differences are taken relative to the Late Holocene baseline, yielding "effective" feedbacks (Rugenstein and Armour, 2021) as $\lambda = \Delta N/\Delta T$ for LGM and $2\times CO_2$ simulations, given that $\Delta F = 0$ in Eq. (2.1) by design.

To evaluate the impact of uncertainty in the $2\times CO_2$ pattern, we also consider existing simulations of abrupt-$4\times CO_2$ with 150-yr regressions (Gregory, 2004) of ΔN versus ΔT, denoted as $\lambda_{4\times}(150\ \text{yr})$, to estimate $\lambda_{2\times}$ (results in Figs. 2.S3–S4, Tables 2.S1–S2). Results are consistent using either method of estimating $\lambda_{2\times}$. To compute $\Delta\lambda$ using $\lambda_{4\times}(150\ \text{yr})$, we apply a timescale adjustment (ζ) to reconcile feedbacks from equilibrium paleoclimate data with the feedback that

applies to 150-year "effective climate sensitivity."

2.6.3 Pattern-only simulations separating pattern and temperature dependence

Feedback differences can be attributed to differences in SST patterns and in global-mean near-surface air temperature (Sherwood et al., 2020), such that

$$\Delta\lambda \approx \Delta\lambda_{\text{PatternOnly}} + \Delta\lambda_T. \tag{2.6}$$

To separate pattern and temperature impacts on $\Delta\lambda$, we conduct additional "pattern-only" simulations in CAM4, CAM5, and CAM6 with the LGMR and $2\times CO_2$ patterns. For these simulations, we multiply local SST anomalies by constant scale factors, k, which are determined for each pattern so that the global-mean ΔSST is reduced to -0.5 K for both simulations. The constant scale factor for a given pattern of anomalies is calculated from the global-mean ΔSST as

$$k = \frac{-0.5 \text{ K}}{\Delta\text{SST}_{\text{global}}},$$

and scaled patterns are then created as

$$\Delta\text{SST}_{\text{scaled}} = k\Delta\text{SST}$$

at each grid cell. We hold SIC constant at the Late Holocene baseline.

SST scaling preserves the spatial pattern of anomalies but forces global-mean ΔT to be small enough that feedback changes due to temperature dependence are negligible ($\Delta\lambda_T \approx 0$). We repeat the feedback calculations, computing

$$\Delta\lambda_{\text{PatternOnly}} \approx \lambda_{2\times}^{(-0.5\text{K})} - \lambda_{\text{LGM}}^{(-0.5\text{K})}$$

as in Eq. (2.3). While there is no existing method that directly isolates temperature dependence in AGCM simulations, the temperature dependence can be approximated as the residual difference between our main and pattern-only simulations, rearranging Eq. (2.5) to

$$\Delta\lambda_T \approx \Delta\lambda - \Delta\lambda_{\text{PatternOnly}}.$$

In this framework, feedback changes due to sea ice are included in temperature dependence.

We employ this pattern-scaling method because it aligns with intuition for pattern effects captured by Green's functions (Dong et al., 2019; Zhou et al., 2017). We do not use Green's functions to calculate the pattern-only feedbacks, but we briefly discuss the Green's functions framework here to explain the pattern-only AGCM simulations. In the linear framework of Green's functions:

$$\Delta N = \sum_j \left(\frac{\partial N}{\partial \mathrm{SST}_j} \Delta \mathrm{SST}_j \right) + \epsilon_N, \tag{2.7}$$

$$\Delta T = \sum_j \left(\frac{\partial T}{\partial \mathrm{SST}_j} \Delta \mathrm{SST}_j \right) + \epsilon_T, \tag{2.8}$$

where j represents each grid cell, $\Delta \mathrm{SST}_j$ represents the full SST anomaly at grid cell j, $\partial N/\partial \mathrm{SST}_j$ represents the global-mean top-of-atmosphere radiative response to a unit increase in local SST at grid cell j, $\partial T/\partial \mathrm{SST}_j$ similarly represents the response of global-mean near-surface air temperature, and ϵ represents changes that are independent of SST.

Because the feedback $\lambda = \Delta N/\Delta T$, constant scale factors applied as $k\Delta\mathrm{SST}$ appear in the feedback calculation as

$$\lambda = \frac{k\Delta N}{k\Delta T}$$

if $\epsilon_N = \epsilon_T = 0$ and SST patterns determine λ. In this case where SST patterns are the sole control on λ, scale factors cancel and have no effect on feedbacks or pattern effects.

By comparing feedbacks from scaled pattern-only simulations with feedbacks from simulations with full SST anomalies, we quantify feedback changes that cannot be explained by SST patterns, which we attribute to feedback dependence on global-mean temperature. For example, temperature dependence could arise from $\partial N/\partial \mathrm{SST}_j$ changing with global-mean temperature or from sea ice appearing at lower latitudes as temperature decreases.

2.6.4 Feedback decomposition using model fields and radiative kernels

Net λ is calculated from changes in top-of-atmosphere radiation (ΔN) divided by changes in global-mean temperature (ΔT). ΔN can be separated into shortwave clear-sky (SWcs), longwave clear-sky (LWcs), and cloud radiative effect (CRE):

$$\Delta N = \Delta N_{\mathrm{SWcs}} + \Delta N_{\mathrm{LWcs}} + \Delta N_{\mathrm{CRE}}. \tag{2.9}$$

Each component of the radiation is available from AGCM output, and dividing all terms by ΔT yields feedbacks for each component, which sum to the net feedback. The total clear-sky feedback is the sum of shortwave and longwave components. Because CRE is calculated as all-sky radiation (N) minus clear-sky radiation, CRE is affected by changes in non-cloud variables.

With radiative kernels (Soden et al., 2008; Shell et al., 2008), feedbacks can be decomposed into contributions from temperature, moisture, and surface albedo. Cloud feedbacks can be estimated by controlling for changes in non-cloud variables, and feedbacks from changing surface albedo can be adjusted to account for overlying cloud cover, which we do here following past studies (Soden et al., 2008). Radiative kernels are linearized around a specific climate in a specific model, however, and are prone to errors when applied to different climates and models.

We use CAM5 kernels (Pendergrass et al., 2018), convolving them with the monthly mean climatology of anomalies in each AGCM simulation to produce feedbacks in Figures 2.S6–S7, and zonal means in Figures 2.S12–S22 (described in Text S5). HadGEM3-GC3.1-LL is not included in kernel analysis due to model-output limitations. GFDL-AM4's $2\times CO_2$ simulation has error in the kernel-derived clear-sky feedback equal to 15.6% of the actual feedback, exceeding the 15% threshold commonly used as a test of clear-sky linearity (Ceppi and Gregory, 2017; Shell et al., 2008); all other simulations have clear-sky feedback errors less than 10%.

Residuals shown in Figure 2.S6 are based on total (all-sky) radiation:

$$\lambda_{\text{Residual}} = \lambda_{\text{Net}} - \sum \lambda_j, \tag{2.10}$$

where λ_{Net} is the net feedback from model output, and $\sum \lambda_j$ is the sum of each of the following kernel-derived feedbacks: Planck, lapse rate, water vapor, surface albedo, shortwave cloud, and longwave cloud.

2.6.5 Bayesian estimate of modern-day climate sensitivity

We follow methods (Sherwood et al., 2020) and code (Webb, 2020) provided by WCRP20 for calculating climate sensitivity, but we provide a summary of relevant methods here. Equilibrium climate sensitivity (ECS) is the steady-state change in global-mean temperature (T) from a doubling of CO_2, traditionally with ice sheets and vegetation assumed fixed. When inferring climate sensitivity that is relevant to modern warming from paleoclimate evidence, changes in the paleoclimate radia-

tive budget that are distinct from feedback processes in modern-day $2\times CO_2$ are treated as forcings; this is typically accomplished by separating "slow" timescale changes as forcings (e.g., ice sheets) from "fast" timescale changes as feedbacks (PALAEOSENS Project Members, 2012). WCRP20 applies this framework by focusing on "effective" climate sensitivity (S), i.e., the 150-year system response.

Relative to WCRP20, our key update only affects $\Delta\lambda$ for the LGM. However, given evidence (Forster et al., 2021; Tierney et al., 2020; Osman et al., 2021; Seltzer et al., 2021; Liu et al., 2023) published after WCRP20 showing LGM cooling centered around $-6°C$ instead of $-5°C$, we report our main results using both assumptions for ΔT_{LGM} (Fig. 2.4; Fig. 2.S4).

To estimate S, we use a modified version of WCRP20's energy balance for the LGM,

$$\Delta T_{LGM} = \frac{-(-0.57\Delta F_{2\times} + \Delta F')}{\left(\frac{\lambda_{2\times}}{1+\zeta} - \Delta\lambda\right)}, \tag{2.11}$$

which determines $\lambda_{2\times}$ and $S = -\Delta F_{2\times}/\lambda_{2\times}$. We substitute our $\Delta\lambda$, which includes pattern and temperature dependence.

Other than testing a colder ΔT_{LGM}, the parameters are unchanged from WCRP20 with the following Normal distributions: modern-day forcing from $2\times CO_2$: $\Delta F_{2\times} \sim \mathcal{N}(\mu = 4.0, \sigma = 0.3)$ W m^{-2}; total non-CO$_2$ LGM forcing: $\Delta F' \sim \mathcal{N}(-6.15, 2)$ W m^{-2} (consisting of -3.2 from ice sheets, -1.1 from vegetation, -1.0 from dust aerosols, -0.28 from N$_2$O, and -0.57 from CH$_4$); timescale transfer parameter: $\zeta \sim \mathcal{N}(0.06, 0.2)$; LGM temperature anomaly: $\Delta T_{LGM} \sim \mathcal{N}(-5, 1)$ °C, or revised $\mathcal{N}(-6, 1)$ °C. In WCRP20, $\Delta\lambda = \Delta\lambda_T = -\alpha\Delta T_{LGM}/2$, with $\alpha \sim \mathcal{N}(\mu = 0.1, \sigma = 0.1)$ W m^{-2} K^{-2}.

Quantification of non-CO$_2$ effective radiative forcing from ice sheets (including sea level), dust and other aerosols, vegetation, and other greenhouse gases represents substantial uncertainty. As noted in Zhu et al. (2021), estimates of the effective radiative forcing for each component of LGM forcing still need to be constrained, and the uncertainty in radiative effects—especially due to dust/aerosols (Kok et al., 2023; Mahowald et al., 2024) and vegetation changes—may be underestimated in WCRP20. Future paleoclimate research on dust and other aerosols (Sagoo and Storelvmo, 2017; Albani and Mahowald, 2019; Albani et al., 2018) and vegetation (Prentice et al., 2011; Bartlein et al., 2011) could improve the estimates used here and in paleoclimate modeling (Kageyama et al., 2017; Schmidt et al., 2014). Recent assessments (Sherwood et al., 2020; Forster et al., 2021; ?) dis-

cuss how dust and other aerosols, vegetation, and non-CO_2 greenhouse gases also act as feedbacks on fast timescales, and some studies (Tierney et al., 2020; Hansen et al., 2023) have calculated a version of climate sensitivity that assumes equivalency in these feedbacks (and in feedbacks from SST patterns) between the LGM and modern-day CO_2, leading to higher values of ECS (Tierney et al., 2020). In the IPCC AR6 (Forster et al., 2021) framework for modern-day ECS, these biogeophysical and non-CO_2 biogeochemical changes are presented as feedbacks (central value of -0.01 W m^{-2} K^{-1}). However, AR6 does not address how to account for the LGM's distinct dust/aerosol and vegetation changes when estimating modern-day ECS from LGM evidence, and this accounting should be a topic of future research.

From the AGCM results in this study, we incorporate pattern effects in $\Delta\lambda$ of Eq. (2.6), assigning a revised $\Delta\lambda \sim \mathcal{N}(-0.37, 0.23)$ W m^{-2} K^{-1}. The revised distribution for $\Delta\lambda$ in our study is based on propagating uncertainty, estimated as spread across AGCMs and LGM reconstructions. To combine uncertainty, we assume that within CAM6, GFDL-AM4, and HadGEM3, the spread in $\Delta\lambda$ from different LGM reconstructions would be the same as in CAM4 and CAM5. We add the differences in $\Delta\lambda$ from each pattern in CAM4 and CAM5, where differences are computed relative to $\Delta\lambda$ using the LGMR pattern, to the results from the remaining three AGCMs. The effect is to treat errors as arising independently in reconstructions and AGCMs. We include $\Delta\lambda$ from extreme-quartile simulations using ensemble members from Annan and LGMR as part of the combined sample. There are 8 simulations from CAM4 and 8 from CAM5 that determine spread from LGM patterns. Note that the spread from LGM patterns is similar between CAM4 and CAM5 (Fig. 2.2).

With the combined sample, we perform bootstrap resampling (described in Text S4) with 10^5 iterations and a sample size of 19 (equal to the number of actual AGCM simulations). The mean over all iterations is $\overline{\Delta\lambda} = -0.37$ (95% range: -0.47 to -0.26) W m^{-2} K^{-1}, and mean sample standard deviation $= 0.23$ (95% range: 0.15 to 0.31) W m^{-2} K^{-1}, which informs our assigned μ and σ, respectively. In Figure 2.S4, we include an uncertainty test by doubling σ to 0.46 W m^{-2} K^{-1}.

Using the same bootstrap method, we calculate forcing efficacy (Hansen et al., 2005) for the LGM, which is equivalent to the ratio of feedbacks $\lambda_{2\times}/\lambda_{\text{LGM}}$, to have a median value of 1.7 (95% range: 1.5 to 2.0), mean value of 2.1 (95% range: 1.6 to 2.6), and sample standard deviation of 1.1 (95% CI: 0.6 to 1.4). Efficacy is strongly affected by division of small values of λ_{LGM}, hence CAM6

becomes an outlier in the efficacy calculation. We report the median in the main text to reduce the outlier impact.

Calculations for LGM likelihoods and Bayesian probability density functions (PDFs) for S follow the Monte Carlo methods in WCRP20 (Sherwood et al., 2020; Webb, 2020). Likelihoods are independent of the prior, but combining the likelihoods with a prior is required to create posterior PDFs that combine lines of evidence. We show results for both priors in WCRP20: the Uniform$(-10, 10)$ W m^{-2} K^{-1} prior on λ (their "Baseline") and the Uniform$(0, 20)$ °C prior on S (robustness test, using a prior that is more conservative regarding the possibility of high climate sensitivity).

2.7 Appendix: Supplemental information

2.7.1 Text S1. Forcing efficacy and pattern effects

In this section, we briefly consider the relationship between "efficacy" and pattern effects, which has been investigated in a recent study (Zhou et al., 2023). Hansen et al. (2005) defined forcing "efficacy" to be the global temperature response per unit forcing relative to the temperature response to CO_2 forcing. Forcing efficacy could also be viewed as translating one unit of forcing by a non-CO_2 agent, e.g., ice sheets, into the equivalent amount of CO_2 forcing which would cause the same global-mean ΔT. While past research on forcing efficacy has considered that different forcings have different temperature impacts (Hansen et al., 2005), analyses using the efficacy framework for the LGM have produced disparate results (Renoult et al., 2023; Zhu et al., 2021; Stap et al., 2019; Shakun, 2017; Hopcroft and Valdes, 2015; Yoshimori et al., 2011), possibly due to simplified physics of intermediate-complexity models (Stap et al., 2019; Shakun, 2017). Because of these results, WCRP20 inflates uncertainty on LGM forcings.

Efficacy, ϵ, can be equivalently framed as a ratio of radiative feedbacks (Richardson et al., 2019; Zhou et al., 2023), e.g., $\epsilon_{\text{IceSheet}} = \frac{\lambda_{2x}}{\lambda_{\text{IceSheet}}}$. The negative LGM pattern effect $\Delta\lambda = \lambda_{2x} - \lambda_{\text{LGM}}$, $\Delta\lambda < 0$, which we find in AGCM simulations using data-assimilation reconstructions for the LGM, is consistent with an LGM efficacy greater than 1. The efficacy of ice sheets is greater than 1 in the following model-only studies with mixed-layer oceans coupled to atmospheric general circulation models: CESM1-CAM5 (Zhu et al., 2021), CESM2 (Zhu et al., 2021), and CESM2-PaleoCalibr (Zhu et al., 2022) (Chapter 2 SI Appendix, Text S2). Some intermediate-complexity models (Stap et al., 2019; Shakun, 2017), however, have reported ice-sheet efficacy less than 1.

The pattern effect, combined with temperature dependence, can equivalently explain forcing efficacy (Zhou et al., 2023). We use the pattern-effect framework rather than efficacy because it allows for quantification of feedback changes in AGCMs using observational constraints on SST patterns from data assimilation and has strong theoretical underpinnings (Dong et al., 2019; Andrews et al., 2018; Zhou et al., 2023). The pattern-effect framework is oriented around the climate feedback, λ, which is the key uncertain parameter for climate sensitivity. We follow methods in WCRP20 (Sherwood et al., 2020) to account for $\Delta\lambda$ for the LGM in estimates of modern-day climate sensitivity. We refer readers to Zhou et al. (2023) for further explanation of the connection

30

between efficacy and pattern-effect frameworks.

2.7.2   Text S2. LGM pattern effects in coupled models

Simulations with mixed-layer ocean models coupled to AGCMs (known as slab ocean models (Bitz et al., 2012), "SOM" hereafter) in CESM1-CAM5 (Zhu et al., 2021), CESM2.1-CAM6 (Zhu et al., 2021), and CESM2-PaleoCalibr (Zhu et al., 2022) illustrate pattern effects in coupled models. Note that feedbacks from ocean dynamics are excluded in the SOM, and models' SST/SIC patterns are not constrained by proxy data, hence we use the SOM only to support interpretation of the LGM pattern effect. Feedbacks in SOM simulations are calculated as $\lambda = \frac{\Delta \text{ERF}}{\Delta T}$, where the effective radiative forcing (ERF) is determined from introducing forcings in separate simulations in the corresponding AGCMs (keeping SST/SIC fixed at pre-industrial values), and ΔT is the equilibrium change in global-mean near-surface air temperature in the SOM (also known as reference-height temperature, or "TREFHT" in CESM name conventions). The ERF is affected by changes in land-surface temperatures, which are not held constant in AGCM simulations due to practical limitations, and an adjustment (Zhu et al., 2021; Hansen et al., 2005) to the ERF can be made to account for land changes—see Zhu et al. (2021) for methods.

This adjustment, which is based on a climate sensitivity parameter (Zhu et al., 2021), can also be applied to estimate an "adjusted ERF" for LGM ice sheets, although it is difficult to assess the validity of the adjustment for ice-sheet forcing, which affects not only land temperatures but also topography. Radiative kernels based on modern climate would typically be used to validate the ERF adjustment (Zhu et al., 2021), but they cannot be applied with LGM topography. Figure 2.S11 of the SI Appendix shows feedbacks from coupled models using both ERF and adjusted ERF. Note that these values do not affect our quantification of $\Delta \lambda$ for ECS calculations, which comes from AGCM simulations.

2.7.3   Text S3. Preparation of SST/SIC boundary conditions

SST and SIC boundary conditions (BCs) for the LGM, Late Holocene baseline, and 2×CO$_2$ are prepared to enable consistent calculation of the net feedback (λ) that is applicable to a modern-day doubling of CO$_2$. When changing the surface BCs in AGCM simulations to compute λ, $\Delta F = 0$

in Eq. 1 only if there are no changes in land-sea distribution or ice sheets. For the LGM and Late Holocene datasets, we adjust for differences in land-sea distribution, determined from Peltier et al. (2015) and Argus et al. (2014), compared to present day using kriging and extrapolation near coastlines in polar regions. While sea-level changes must be neutralized to preserve $\Delta F = 0$ in the AGCM simulations, infilling SST over the Sunda Shelf represents a notable uncertainty (DiNezio and Tierney, 2013; DiNezio et al., 2018). The alternative option, holding all forcings constant at LGM rather than modern values, would require changing modern topography to include LGM ice sheets and inherit sea level of the LGM. Those changes could introduce more uncertainty in estimates of λ that are relevant to future warming. Here we only consider the framework with constant modern-day forcings.

For SST, kriging is performed across overlapping subset regions of radius $\approx$ 3000 km spaced around the globe. Results for overlapping subset regions are merged using inverse-distance weighting from the center of each subset region. Kriging results are retained only where no pre-existing SST value exists in a dataset. Over polar regions and inland waters, inverse-distance extrapolation populates the SST field.

For SIC, all values are first required to be no less than the ice-sheet fraction at that location, i.e., modern seas that were covered by ice sheets at the LGM, such as the Hudson Bay, are assigned a minimum SIC that equals the LGM ice fraction at 21,000 years ago (Peltier et al., 2015; Argus et al., 2014). For modern seas which were land but not ice sheet at the LGM, SIC is populated based on the SST. This step uses the SIC formula from the CAM boundary condition protocol (Hurrell et al., 2008), where $\text{SIC} = 100\%$ if $\text{SST} < -1.8°\text{C}$, $\text{SIC} = 0\%$ if $\text{SST} > 4.97°\text{C}$, and otherwise $\text{SIC} = 0.729 - \left(\frac{\text{SST}+1.8}{9.328}\right)^{1/3}$. Gaussian smoothing is applied to the result, reducing any sharp boundaries caused by the infilling. The SIC formula above is also applied to maintain internally consistent values of SST and SIC (Hurrell et al., 2008) in the Late Holocene baseline. See Chapter 2's SI Appendix, Text S4, for uncertainty tests regarding sea ice.

The Annan dataset includes only annual SST and no reconstruction of SIC. Because SIC is required in all AGCMs, we assign the SIC from Amrhein to the Annan data. In a CAM4 test using the LGMR SIC with Annan SSTs (instead of the Amrhein SIC), $\Delta\lambda$ is marginally more negative (λ_{LGM} changes by < 0.1 $\text{W}\,\text{m}^{-2}\,\text{K}^{-1}$). This result suggests that uncertainty from assigning a SIC reconstruction to Annan SSTs is small compared to uncertainty in the SST reconstruction.

We assign the Amrhein SIC for the Annan SST in our main results because this choice is more conservative in that it reduces the magnitude of the mean LGM pattern effect. For consistency, the Annan SST is assigned the annual cycle from the Amrhein data for SST/SIC.

For the $2\times CO_2$ BC, we use output from LongRunMIP (Rugenstein et al., 2019) simulations of abrupt and transient-1% yr^{-1} doubling of CO_2. We use the mean of 200 years of output from the following six models to create a multi-model mean SST/SIC BC: CESM1.0.4 (years 2300–2500), CNRM-CM6-1 (years 550–750), HadCM3L (years 500–700), MPI-ESM-1.2 (years 800–1000), GFDL-ESM2M (years 4300–4500), and MIROC3.2 (years 1803–2003). HadCM3L results use years 500–700 due to an output error in the pre-industrial control run after year 700. All LongRunMIP results are regridded to a standard $1.9° \times 2.5°$ (latitude $\times$ longitude) grid. For SIC, monthly output is available, and we compute a 200-year climatology for each model and then a multi-model-mean climatology. For SST, annual output is available for each model and monthly output from MIROC3.2. We compute the 200-year mean SST anomaly for each model and then apply the annual cycle from MIROC3.2 to the multi-model mean. We also show results in Chapter 2's SI Appendix, Fig. 2.S3–S4, which do not use the LongRunMIP-$2\times CO_2$ BC and instead use 150-year regressions of abrupt-$4\times CO_2$ from parent coupled models corresponding to each AGCM used in this study, thereby sampling uncertainty in warming patterns because the 150-year regressions are produced from different models' warming patterns.

BCs are regridded to the $1.9° \times 2.5°$ (latitude $\times$ longitude) grid used for CAM4, CAM5, and CAM6. HadGEM3-GC31-LL regrids to N96 (resolution of approximately 135 km), and GFDL-AM4 regrids to a C96 cubed sphere (resolution of approximately 100 km).

For the "pattern-only" simulations with SST anomalies normalized to $-0.5\,K$, we make the following changes to the LGM and $2\times CO_2$ BCs. For the LGM, we use the LGMR SST. For $2\times CO_2$, we use the LongRunMIP SST. We compute the global-mean ΔSST for both datasets as $\overline{\Delta SST}$, and we multiply all local SST anomalies by the scale factor $-0.5/\overline{\Delta SST}$. This scaling causes the resulting global-mean ΔSST to become $-0.5\,K$, but the spatial pattern of the SST anomalies is unchanged. We use $-0.5\,K$ for both the LGM and $2\times CO_2$ so that there is no cooling–warming asymmetry, and ΔT is small enough that temperature dependence of λ is negligible (i.e., $\Delta\lambda_T \approx 0$, and $\Delta\lambda \approx \Delta\lambda_{\text{PatternOnly}}$). ΔT is still large enough that we can compute $\lambda = \Delta N/\Delta T$ without requiring an excessively long simulation to overcome noise in the denominator. We use the baseline

SIC (Late Holocene) in all of the pattern-only simulations so there are no changes in sea ice, so this set of simulations also serves to check whether $\Delta\lambda$ is attributable to SIC rather than SST changes.

To examine whether the pattern-only results are sensitive to the scaling method of separating pattern effects, we tested an alternative subtraction method in CAM4 (using the LGMR pattern for the LGM and the LongRunMIP pattern for $2\times CO_2$). We ran alternative pattern-only simulations with global-mean SST anomalies set to zero by subtracting the global mean at all locations. These experiments produced consistent results for $\Delta\lambda_{\text{PatternOnly}}$ compared to scaling.

An additional simulation was run in HadGEM3-GC3.1-LL with SIC held constant at the Late Holocene baseline while the SST field is varied with the full value of anomalies, using the LongRunMIP-$2\times CO_2$ and LGMR patterns of SST. Results from this simulation are shared in Chapter 2's SI Appendix, Text S4.

This concludes the preparation steps for the main simulations (BCs from four data-assimilation reconstructions for the LGM, one Late Holocene, and one $2\times CO_2$) and the "pattern-only" simulations (two additional BCs: LGMR and LongRunMIP-$2\times CO_2$ scaled to $-0.5\,\text{K}$). The final adjustment to each BC follows the standard boundary-condition protocol for CAM, known as "bcgen." This process ensures that SIC and SST are plausibly bounded (e.g., SIC between 0 and 1), and it transfers the monthly climatology to mid-month values which can be linearly interpolated in an AGCM.

2.7.4   Text S4. Uncertainty of $\Delta\lambda$

Figures 2.S12–S22 of the Chapter 2 SI Appendix show zonal means (indicated by brackets as $[\lambda]$) of the global-mean feedbacks that appear in Figure 2.S6 of the Chapter 2 SI Appendix. The net feedback, clear-sky shortwave (SW), clear-sky longwave (LW), and cloud radiative effect are calculated directly from model output. The remaining feedbacks are from radiative kernel decomposition (Materials and Methods) using CAM5 kernels. GFDL-AM4's $2\times CO_2$ simulation has error in the kernel-derived clear-sky feedback equal to 15.6% of the actual feedback, exceeding the 15% threshold commonly used as a test of clear-sky linearity (Ceppi and Gregory, 2017; Shell et al., 2008; Zelinka et al., 2020); all other simulations have clear-sky feedback errors less than 10%. Total cloud feedback is also shown as the sum of kernel-derived SW and LW components.

Each of the zonal-mean figures consists of: (A) In CAM5, mean and range of feedbacks across four LGM reconstructions and $2\times CO_2$ from LongRunMIP. (B) In CAM5, mean and range of the difference in feedbacks ($\Delta\lambda = \lambda_{2x} - \lambda_{\mathrm{LGM}}$) across four LGM reconstructions from results in panel A. (C) Feedbacks across various AGCMs, using the LGMR reconstruction of the LGM and $2\times CO_2$ from LongRunMIP. (D) Mean and range of $\Delta\lambda$ across various AGCMs from results in panel C. Note that HadGEM3 is not included in the kernel-derived feedbacks due to limited availability of model output.

2.7.5 Supplemental Figures
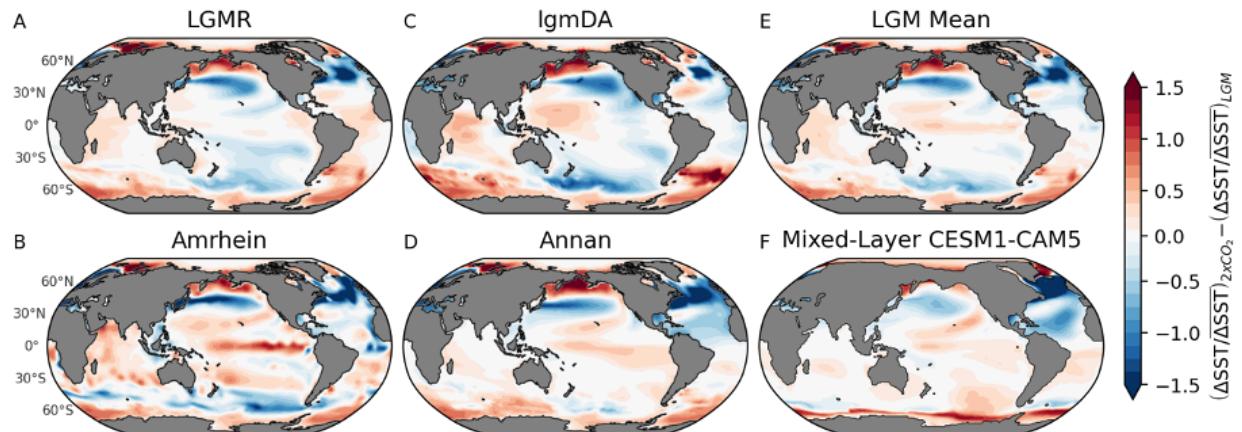
Figure 2.S1: Differences in LGM sea-surface temperature (SST) patterns compared to $2\times CO_2$ reference pattern. All local anomalies are normalized through division by the global-mean anomaly, then differences between the $2\times CO_2$ pattern and LGM pattern are taken. Red regions indicate where SST anomalies are relatively more amplified in $2\times CO_2$, while blue regions indicate where SST anomalies are relatively more amplified at the LGM. (A–E) LGM patterns corresponding to Fig. 2.1A–E, and $2\times CO_2$ reference pattern is Fig. 2.1F from LongRunMIP-$2\times CO_2$. (F) In CESM1-CAM5 (Zhu et al., 2021) mixed-layer ocean model without data assimilation, difference between $2\times CO_2$ and LGM patterns (shown in Fig. 2.S5C–D).
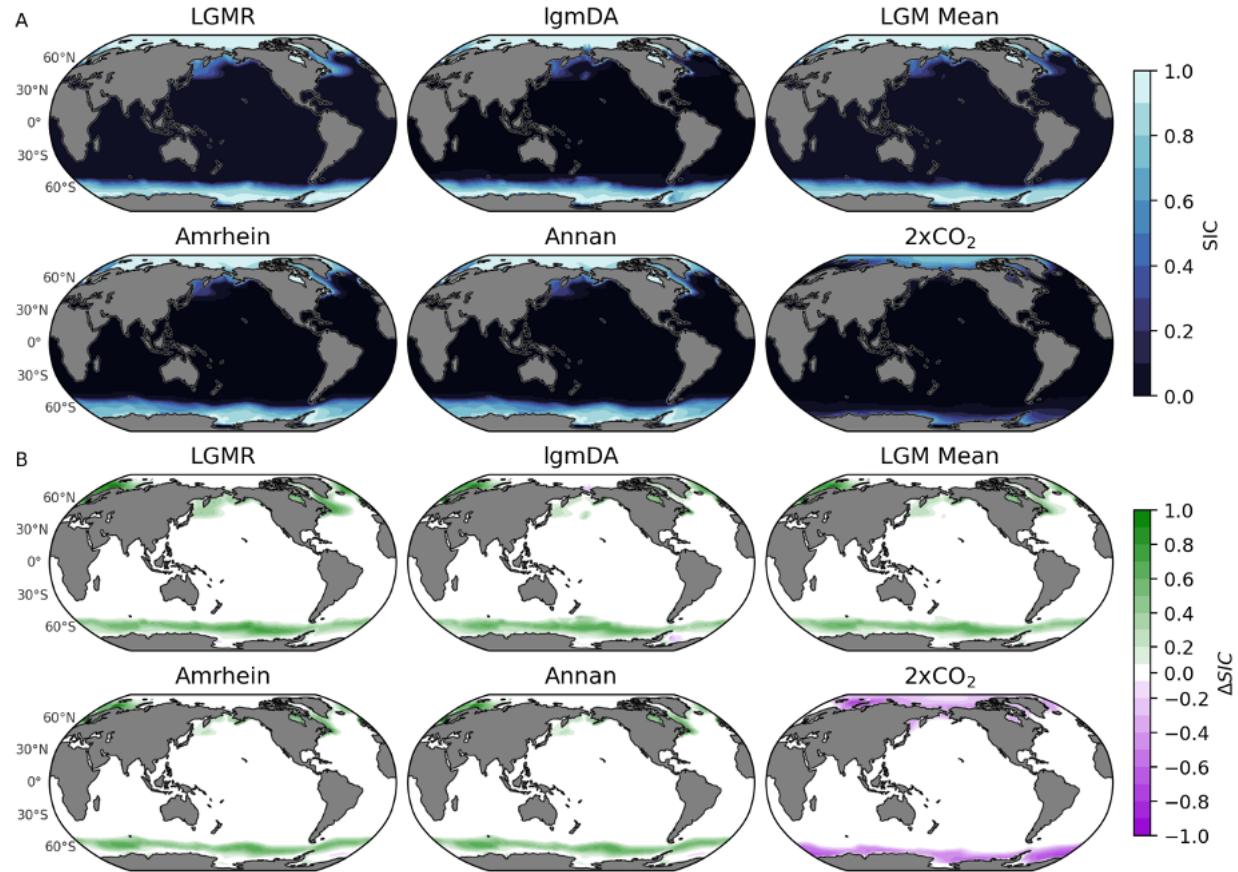
Figure 2.S2: Sea-ice concentration (SIC) from data-assimilation reconstructions of the Last Glacial Maximum (LGM) compared to 2×CO$_2$. (A) SIC from LGM Reanalysis (LGMR) (Osman et al., 2021), Amrhein (Amrhein et al., 2018), lgmDA (Tierney et al., 2020), Annan (Annan et al., 2022) (assigned SIC from Amrhein); mean of three LGM reconstructions (LGMR, Amrhein, and lgmDA); and multi-model mean from near-equilibrium simulations of 2×CO$_2$ in LongRunMIP (Rugenstein et al., 2019), where each of six models is averaged over the final 200 years of simulation. (B) Difference in sea-ice concentration relative to Late Holocene baseline (LGMR reconstruction). All panels show annual mean. Reconstructions are infilled to modern coastlines (Materials and Methods).
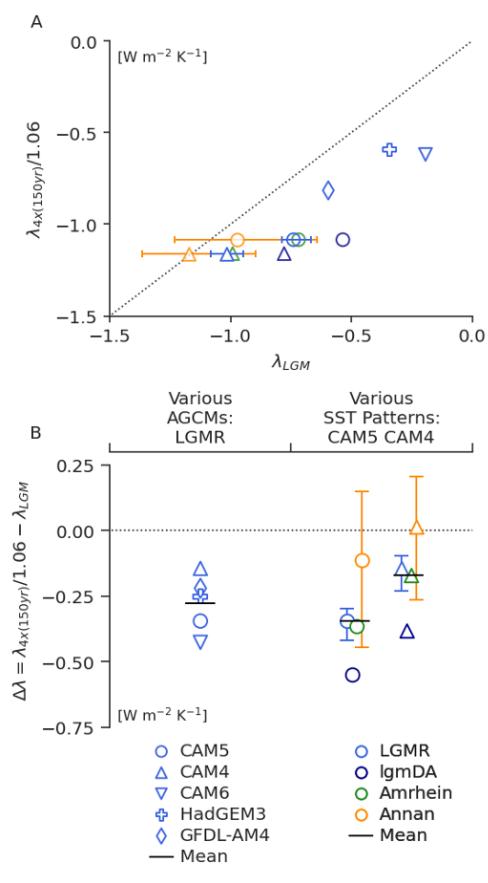
Figure 2.S3: LGM pattern effect ($\Delta\lambda$) based on LGM climate feedbacks in AGCMs and CO_2 climate feedbacks from 150-year regression of abrupt-4×CO_2 in coupled models. Similar to Fig. 2.2, except λ_{2x} is replaced by $\lambda_{4x}^{(150 \text{ yr})}/1.06$, the feedback from regression in abrupt-4×CO_2 simulations using parent coupled models corresponding to each AGCM; a timescale adjustment of $1/1.06$ is applied based on the WCRP20 central estimate (Sherwood et al., 2020) to make 150-year 4×CO_2 feedbacks comparable with λ_{LGM} equilibrium feedbacks. Different models (all using the LGMR pattern for the LGM) are indicated by symbols. Different LGM patterns (in CAM5 and CAM4) are indicated by colors. (A) Scatter plot of 4×CO_2 feedbacks (including adjustment factor of $1/1.06$) versus LGM feedbacks, with $\lambda_{4x}^{(150 \text{ yr})}/1.06 = \lambda_{\text{LGM}}$ shown as dashed line. (B) LGM pattern effect, $\Delta\lambda = \lambda_{4x}^{(150 \text{ yr})}/1.06 - \lambda_{\text{LGM}}$, using feedbacks shown in (A), with $\Delta\lambda = 0$ shown as dashed line. Note that $\Delta\lambda$ includes SST pattern effects and contributions from temperature dependence.
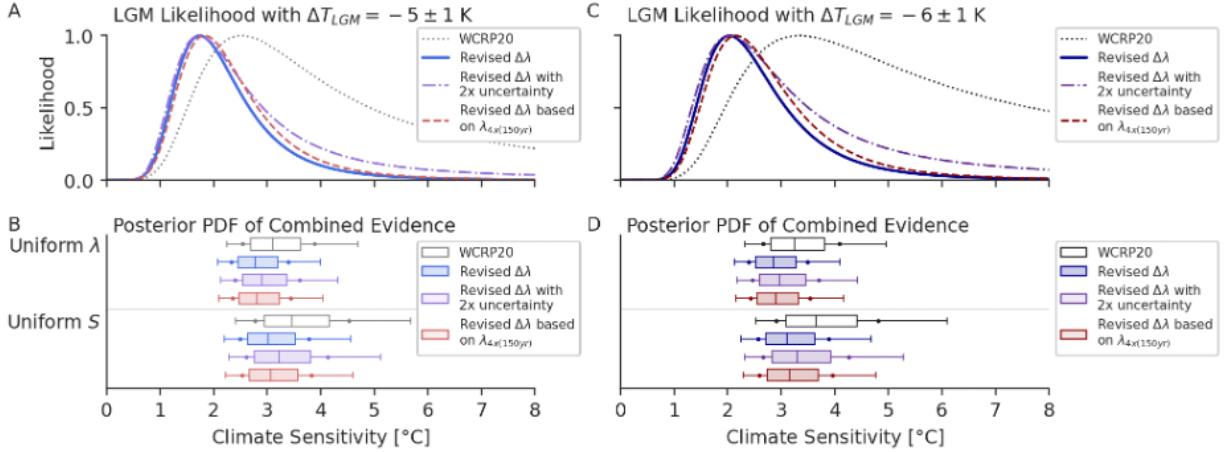
Figure 2.S4: Uncertainty tests for modern-day climate sensitivity including LGM pattern effects. Following Fig. 2.4, showing WCRP20 original (Sherwood et al., 2020) LGM $\Delta T_{\text{LGM}} \sim \mathcal{N}(\mu = -5, \ \sigma = 1)$ K in the left column and revised LGM $\Delta T_{\text{LGM}} \sim \mathcal{N}(-6, \ 1)$ K based on IPCC AR6 (Forster et al., 2021) in the right column, including two uncertainty tests. Results from WCRP20 (Sherwood et al., 2020) with no LGM pattern effect (gray and black) and our base assumption (light and dark blue) for revised $\Delta\lambda \sim \mathcal{N}(-0.37, \ 0.23)$ $\text{W m}^{-2}\,\text{K}^{-1}$ from Fig. 2.4 are repeated here for comparison. The first uncertainty test (light and dark purple) increases the σ assumption by a factor of two: $\Delta\lambda \sim \mathcal{N}(-0.37, \ 0.46)$ $\text{W m}^{-2}\,\text{K}^{-1}$. The second uncertainty test (light and dark red) concerns the $2\times\text{CO}_2$ pattern and feedback: a different distribution, $\Delta\lambda \sim \mathcal{N}(-0.27, \ 0.20)$ $\text{W m}^{-2}\,\text{K}^{-1}$, is assigned based on results shown in Fig. 2.S3 using $\lambda_{4x}^{(150 \text{ yr})}/1.06$, the feedback derived from 150-year regressions of abrupt-$4\times\text{CO}_2$ using parent coupled models corresponding to each AGCM, including a timescale-adjustment factor of $1/1.06$ from WCRP20's central estimate (Sherwood et al., 2020). Climate sensitivity shown is effective sensitivity (S) from 150-year response, as in WCRP20 (Sherwood et al., 2020). (A) Likelihood functions for S based on only the LGM line of evidence. (B) Posterior PDF after combining LGM with other lines of evidence in WCRP20 (Sherwood et al., 2020), assuming a uniform-λ prior (upper panel) or a uniform-S prior (lower panel). Outlier lines indicate 5–95th percentiles, dots indicate 66% likely range, and box indicates 25–75th percentiles and median.
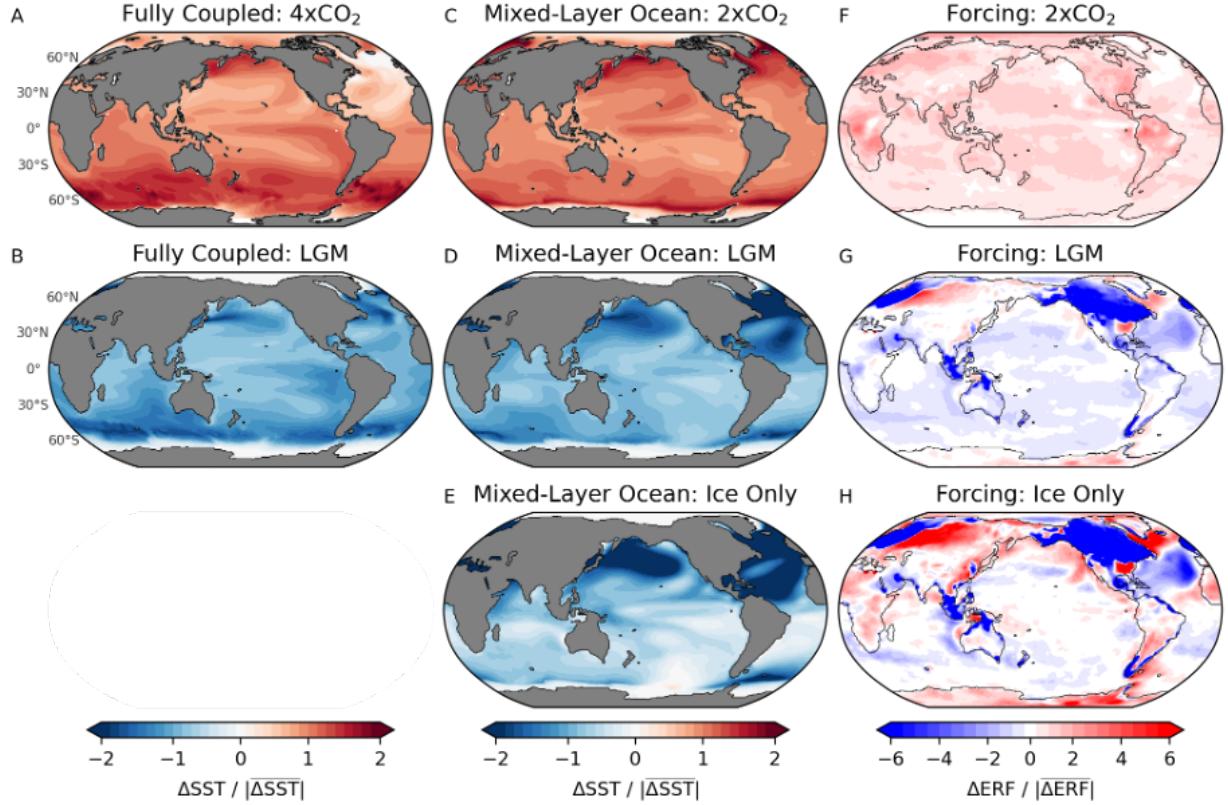
Figure 2.S5: Spatial patterns of sea-surface temperature (SST) response and effective radiative forcing (ERF) in CESM1-CAM5 model simulations from Zhu & Poulsen (Zhu et al., 2021). Spatial patterns here are shown as zonal means in Fig. 2.3. All local anomalies are normalized through division by the absolute value of the global-mean anomaly. (A–B) SST patterns in quasi-equilibrium from fully coupled atmosphere–ocean model with LGM ice-sheet and greenhouse-gas forcings (Zhu et al., 2021) compared to abrupt-4×CO$_2$ forcing (Zhu et al., 2019). (C–E) Equilibrium SST patterns from mixed-layer ocean model coupled to CAM5, including a simulation with only LGM ice-sheet forcing (Zhu et al., 2021). (F–H) ERF patterns from corresponding AGCM simulations in CAM5.
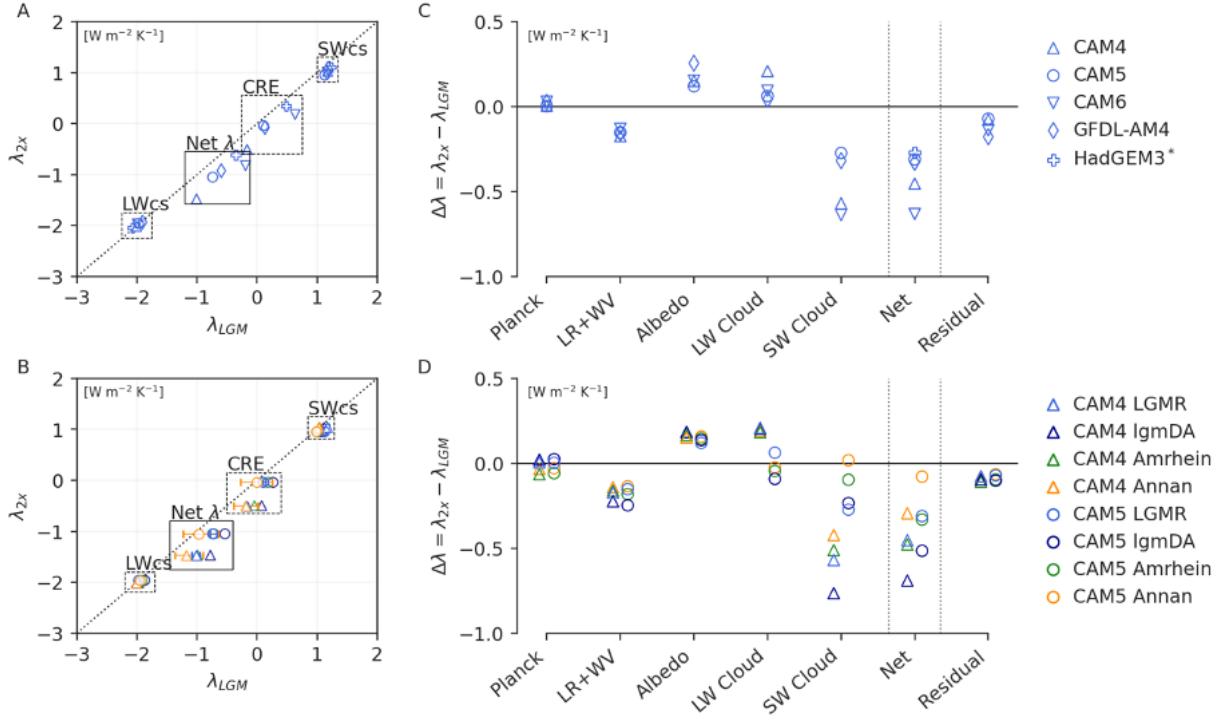
Figure 2.S6: Feedback decomposition of Last Glacial Maximum (LGM) and $2\times CO_2$ climate feedbacks in atmospheric general circulation models (AGCMs). Left column uses direct model outputs in scatter plots of $2\times CO_2$ feedbacks (λ_{2x}) versus LGM feedbacks (λ_{LGM}), with $\lambda_{2x} = \lambda_{\mathrm{LGM}}$ denoted by dashed line. Cloud radiative effect (CRE), shortwave clear-sky (SWcs), longwave clear-sky (LWcs), and net feedbacks are shown. (A) Results from various AGCMs, all using the LGMR reconstruction for the LGM. (B) Results from various LGM reconstructions in CAM4 and CAM5, with different reconstructions indicated by colors. Right column shows decomposition of $\Delta\lambda$ using CAM5 radiative kernels, with residual equal to the net feedback in models minus the sum of kernel-derived feedbacks. (C) Results from various AGCMs (note that only net λ is available for HadGEM3). (D) Results from various LGM reconstructions in CAM4 and CAM5. Lapse rate and water vapor feedbacks are combined (LR+WV) given their anti-correlation across models (Soden and Held, 2006). Note that $\Delta\lambda$ includes SST pattern effects and contributions from temperature dependence.
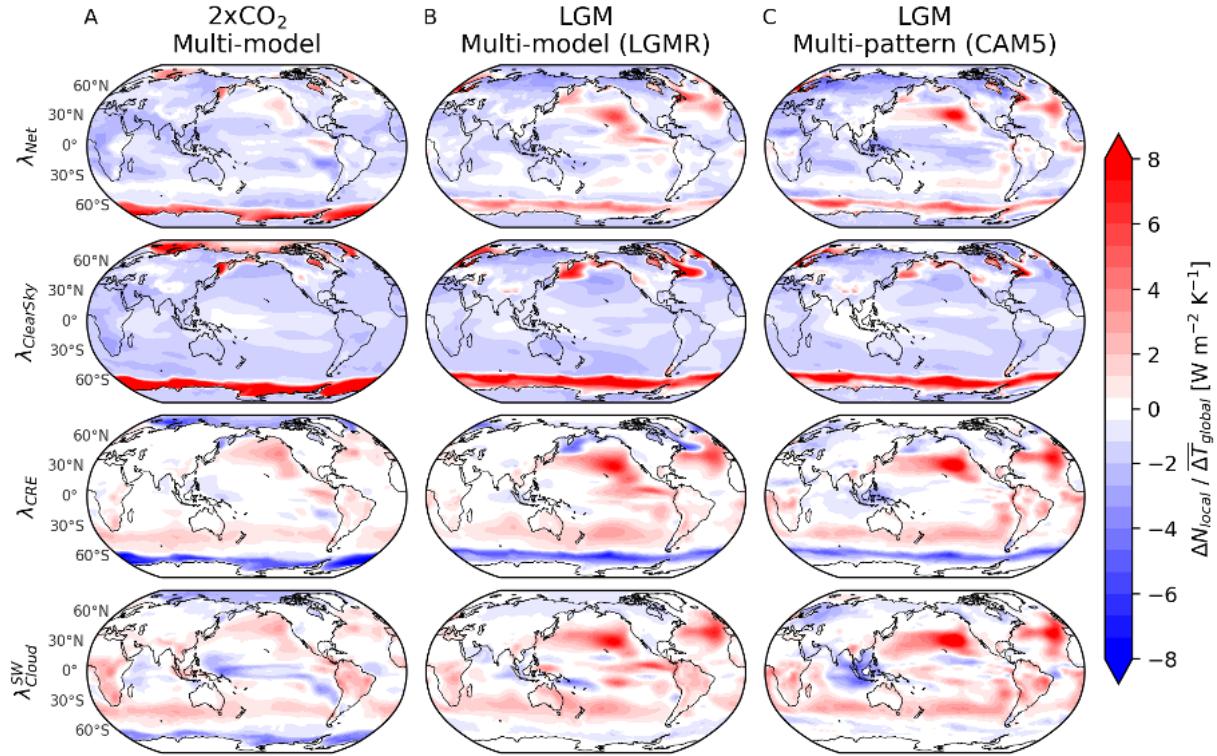
Figure 2.S7: Spatial decomposition of Last Glacial Maximum (LGM) and $2\times CO_2$ local climate feedbacks in atmospheric general circulation models (AGCMs). Local feedbacks represent local change in top-of-atmosphere radiation (ΔN_{local}) divided by global-mean change in near-surface air temperature (ΔT_{global}); global integrals of the local feedbacks equal the global-mean feedbacks. Top row shows net feedback (λ_{Net}) from total all-sky changes in ΔN, second row shows $\lambda_{ClearSky}$ from changes in ΔN attributable to clear-sky radiation, third row shows cloud radiative effects (λ_{CRE}); rows 1–3 use direct model output. Fourth row shows radiative-kernel estimates of shortwave cloud feedbacks (λ_{Cloud}^{SW}). (A) $2\times CO_2$ multi-model mean based on five AGCM simulations using LongRunMIP (Rugenstein et al., 2019) pattern. (B) LGM multi-model mean based on five AGCM simulations using LGMR (Osman et al., 2021) pattern. (C) LGM multi-pattern mean in CAM5 using four LGM reconstructions. Note that radiative-kernel results for λ_{Cloud}^{SW} exclude HadGEM3 due to output limitations.
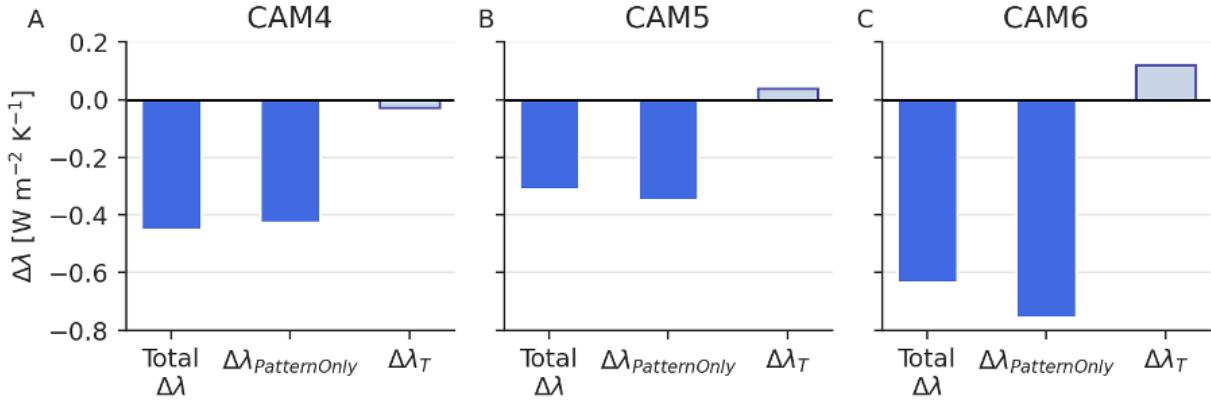
Figure 2.S8: Separating pattern and temperature dependence of feedback changes as total $\Delta\lambda \approx \Delta\lambda_{\text{PatternOnly}} + \Delta\lambda_T$. First column shows total $\Delta\lambda = \lambda_{2x} - \lambda_{\text{LGM}}$ from Figure 2.2, calculated in main simulations with full SST anomalies and SIC for $2\times CO_2$ and LGM (using LGMR reconstruction). Second column shows pattern-only simulations with global-mean ΔSST scaled to -0.5 K, where $\Delta\lambda_{\text{PatternOnly}} \approx \lambda_{2x}^{(-0.5\text{K})} - \lambda_{\text{LGM}}^{(-0.5\text{K})}$. Third column shows temperature dependence, $\Delta\lambda_T$, approximated as the residual difference between the main and pattern-only simulations, $\Delta\lambda_T \approx \Delta\lambda - \Delta\lambda_{\text{PatternOnly}}$. Results in (A) CAM4, (B) CAM5, and (C) CAM6.
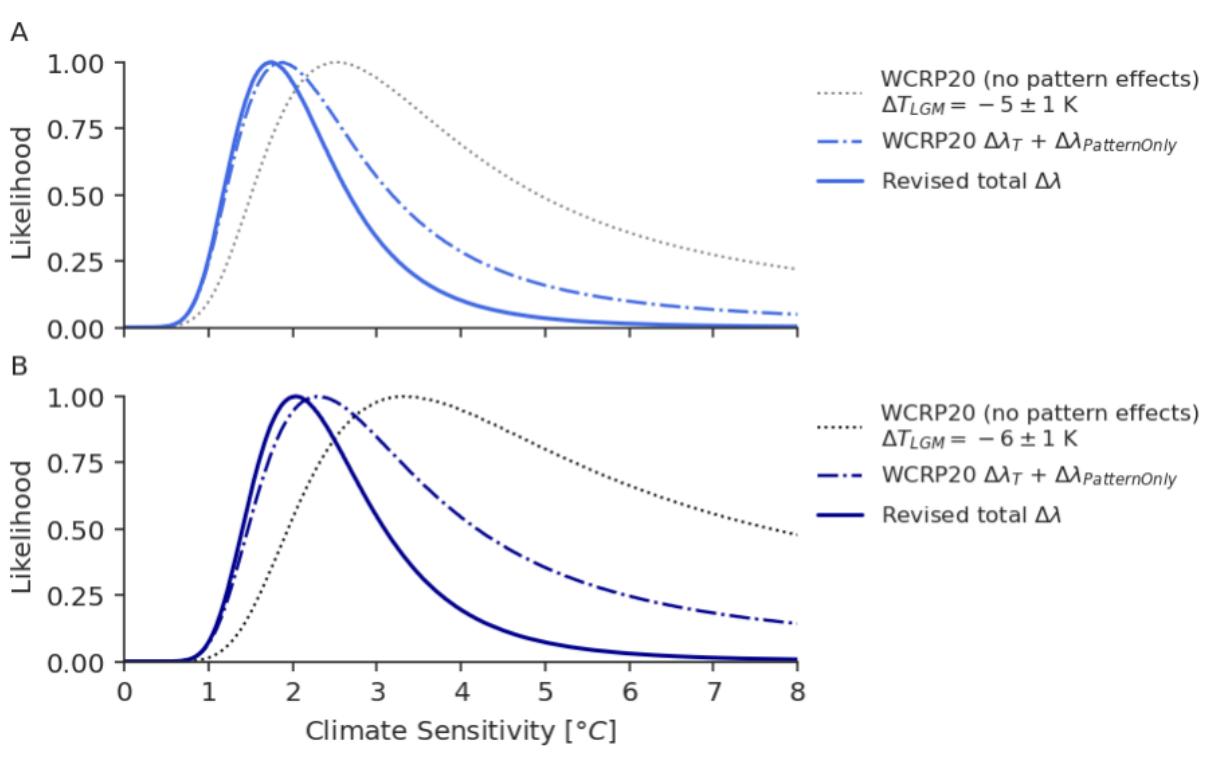
Figure 2.S9: Likelihoods for LGM line of evidence with separate updates for SST pattern effects and temperature dependence of feedbacks. (Dotted) WCRP20 LGM likelihood (Sherwood et al., 2020), which includes an estimate of $\Delta\lambda_T$ for the LGM but no adjustment for pattern effects. (Dash-dot) Revised likelihood using WCRP20 estimate of $\Delta\lambda_T$ but including feedback changes from SST patterns based on pattern-only simulations in this study, assuming $\Delta\lambda_{\text{PatternOnly}} \sim \mathcal{N}(\mu = -0.51,\ \sigma = 0.23)\ \text{W m}^{-2}\,\text{K}^{-1}$. (Solid) Revised likelihood using total revised $\Delta\lambda$ from this study, as shown in Fig. 2.4, which includes both pattern effects and temperature dependence, assuming $\Delta\lambda \sim \mathcal{N}(-0.37,\ 0.23)\ \text{W m}^{-2}\,\text{K}^{-1}$. (A) All likelihoods assume $\Delta T_{\text{LGM}} \sim \mathcal{N}(-5,\ 1)\ \text{K}$ as in original WCRP20 results (Sherwood et al., 2020). (B) All likelihoods assume $\Delta T_{\text{LGM}} \sim \mathcal{N}(-6,\ 1)\ \text{K}$, using the updated central estimate from IPCC AR6 (Forster et al., 2021).
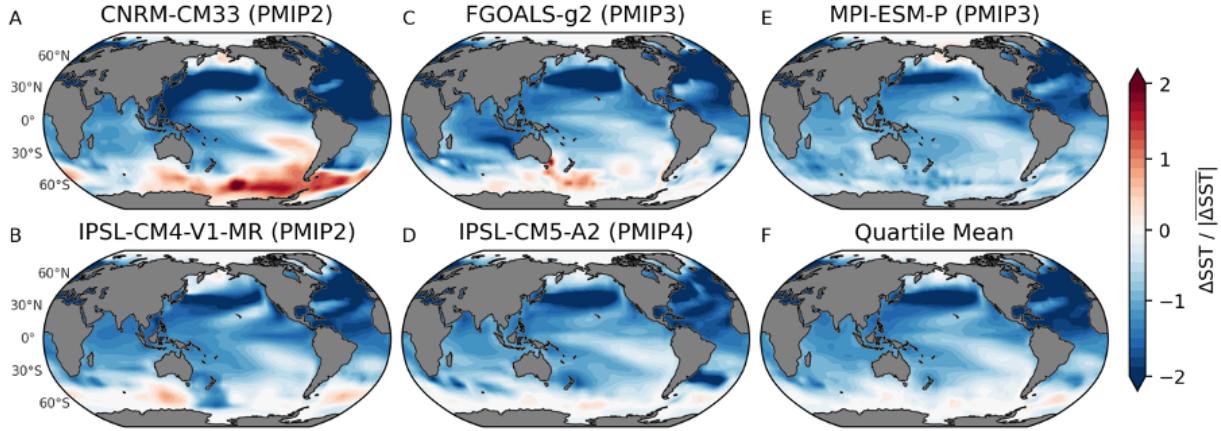
Figure 2.S10: Patterns of SST anomalies from Annan (Annan et al., 2022) ensemble members in the quartile with strongest negative climate feedback (λ). 19 ensemble members are ranked by estimated λ, which is produced from CAM5 Green's functions (Zhou et al., 2017), and 5 members shown comprise the quartile with most-negative estimated λ. (A–E) Data-assimilation posterior SST using model priors specified in subtitles. (F) Pattern of the quartile-mean SST. To show SST patterns, local SST anomalies are normalized into patterns through division by the absolute value of the global-mean SST anomaly (consistent with feedbacks being radiative responses divided by global-mean temperature anomalies). All panels show annual means. LGM reconstructions are infilled to modern coastlines (Materials and Methods).
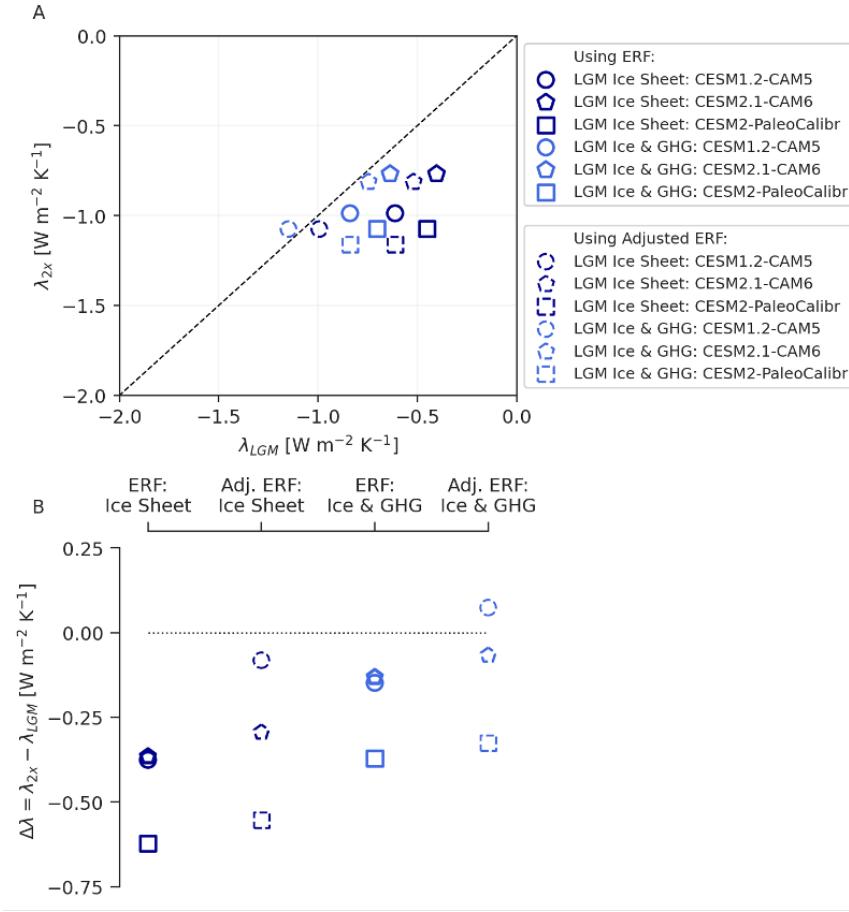
Figure 2.S11: Feedbacks and $\Delta\lambda$ using either effective radiative forcing (ERF) or adjusted ERF from previously published simulations in mixed-layer ocean models. (A) Scatter plot of λ_{2x} vs. λ_{LGM} in mixed-layer ocean models; λ_{LGM} is shown for simulations using only the LGM ice-sheet forcing (dark blue), which includes LGM sea-level changes, and for simulations using LGM ice-sheet forcing and greenhouse-gas (GHG) forcings (royal blue). Dashed markers indicate corresponding results using "adjusted ERF" to calculate feedbacks. (B) $\Delta\lambda$ based on feedbacks shown in panel A. Note that in LGM simulations using CESM2.1-CAM6 (Zhu et al., 2021) and CESM2-PaleoCalibr (Zhu et al., 2022), the LGM ice-sheet forcing and GHG forcing are applied in separate simulations, and their sums are shown as LGM Ice & GHG. This linearity assumption was validated in CESM1-CAM5 (Zhu et al., 2021).
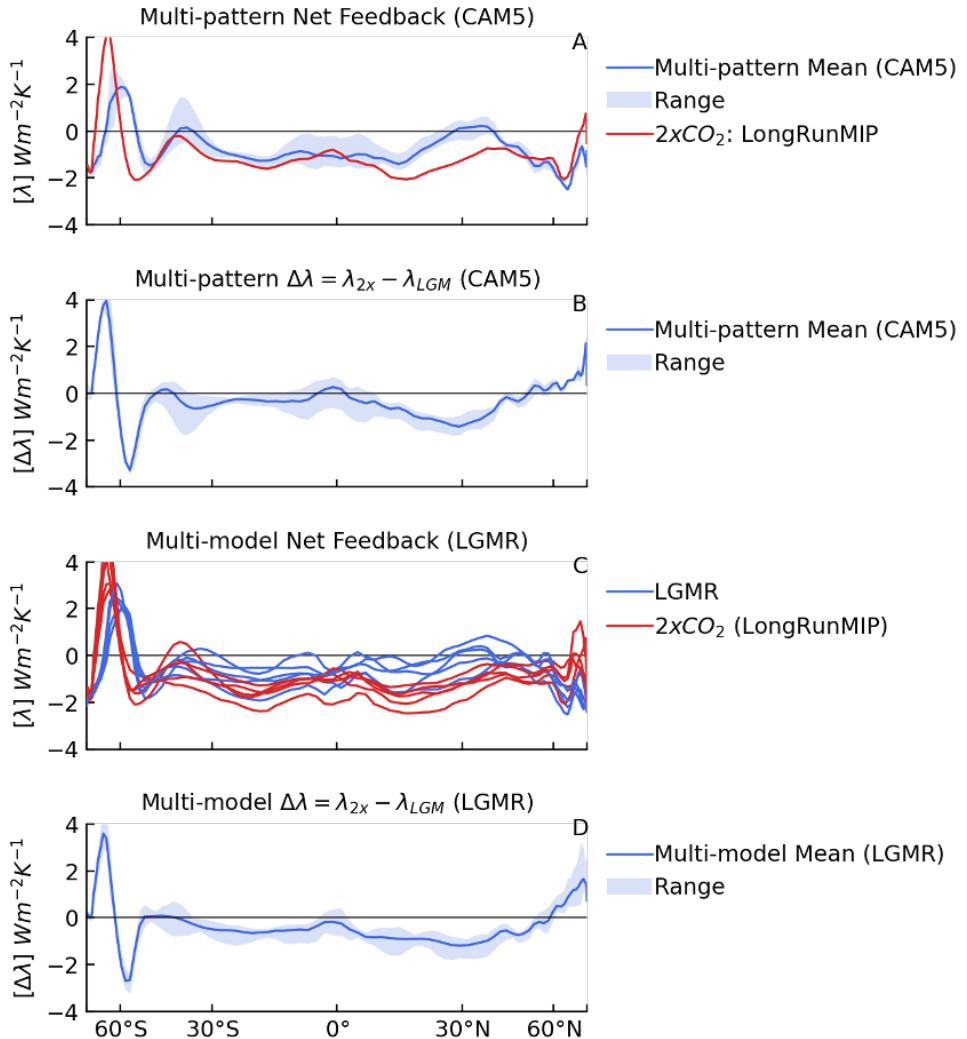
Figure 2.S12: Zonal-mean net feedback and $\Delta\lambda$. (A) In CAM5, mean and range of feedbacks across four LGM reconstructions and $2\times CO_2$ from LongRunMIP. (B) In CAM5, mean and range of the difference in feedbacks ($\Delta\lambda = \lambda_{2x} - \lambda_{\mathrm{LGM}}$) across four LGM reconstructions from results in (A). (C) Feedbacks across various AGCMs, using the LGMR reconstruction of the LGM and $2\times CO_2$ from LongRunMIP. (D) Mean and range of $\Delta\lambda$ across various AGCMs from results in (C). Note that HadGEM3 is not included in the kernel-derived feedbacks due to limited model output.
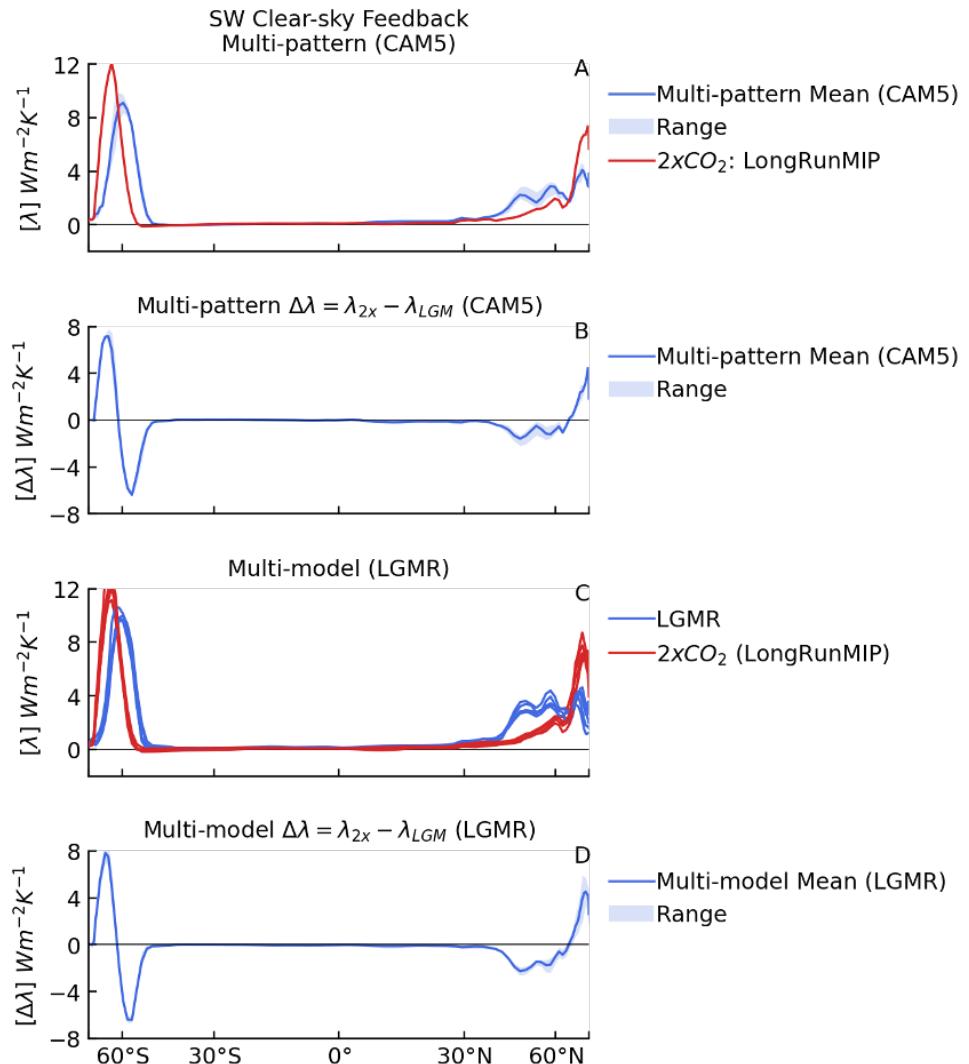
Figure 2.S13: Zonal-mean net feedback and $\Delta\lambda$. See caption of Figure 2.S12.
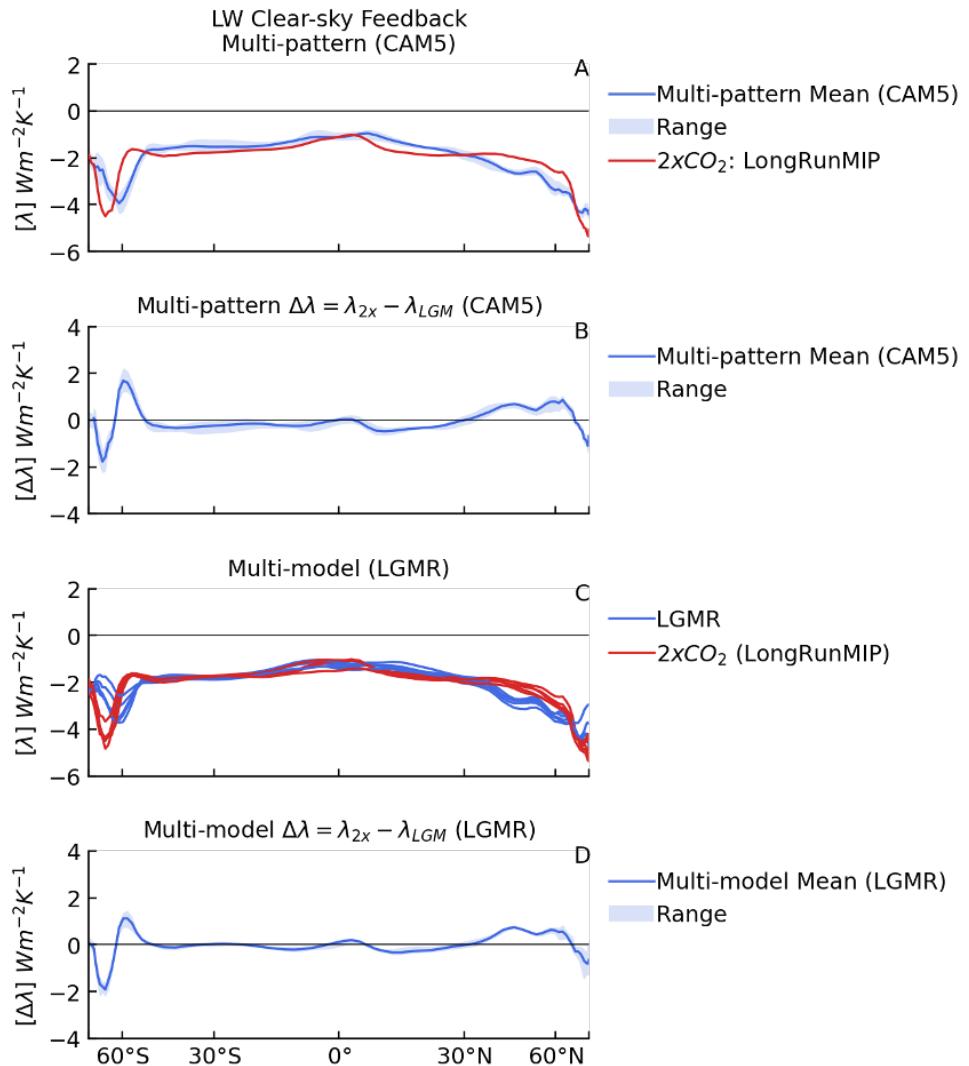
Figure 2.S14: Zonal-mean longwave clear-sky feedback and $\Delta\lambda$. See caption of Figure 2.S12.
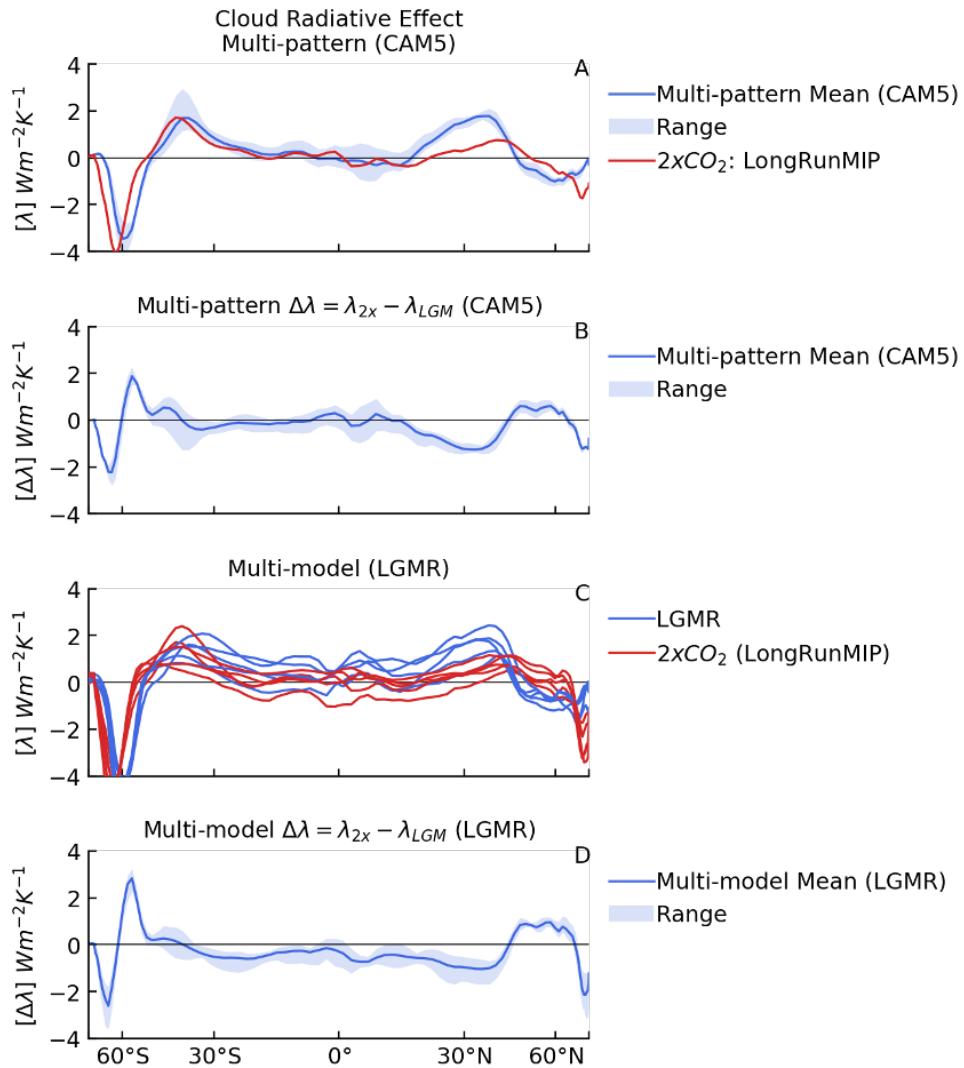
Figure 2.S15: Zonal-mean feedback from cloud radiative effect (CRE) and $\Delta\lambda$. See caption of Figure 2.S12.
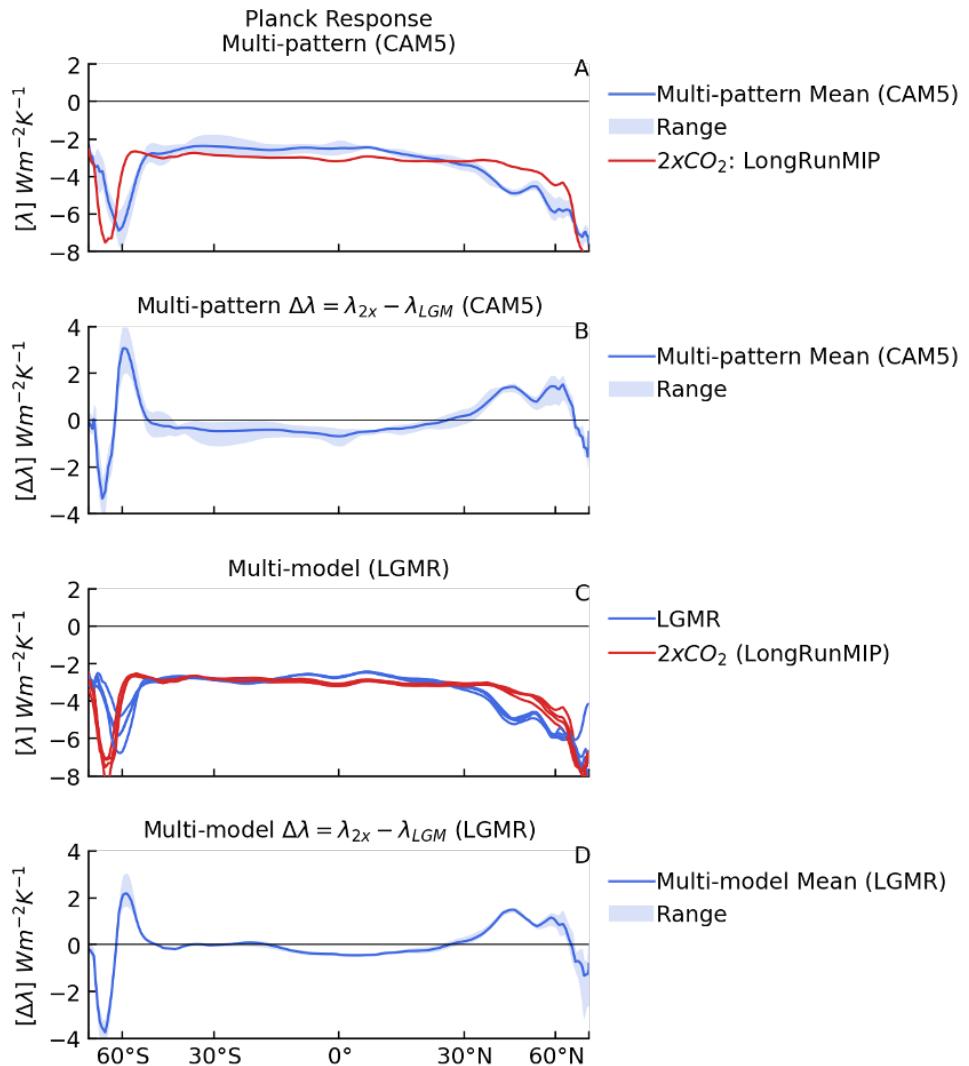
Figure 2.S16: Zonal-mean planck feedback and $\Delta\lambda$. See caption of Figure 2.S12.
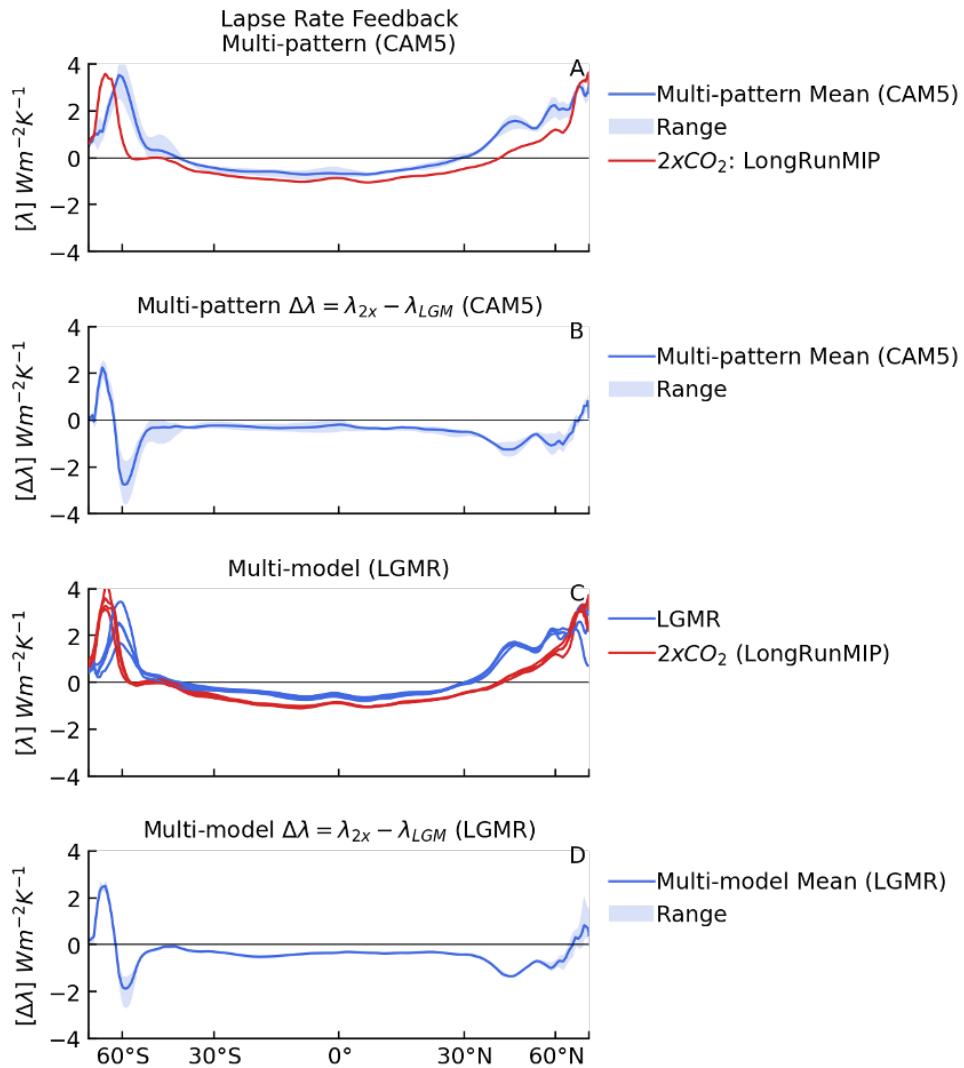
Figure 2.S17: Zonal-mean lapse rate feedback and $\Delta\lambda$. See caption of Figure 2.S12.
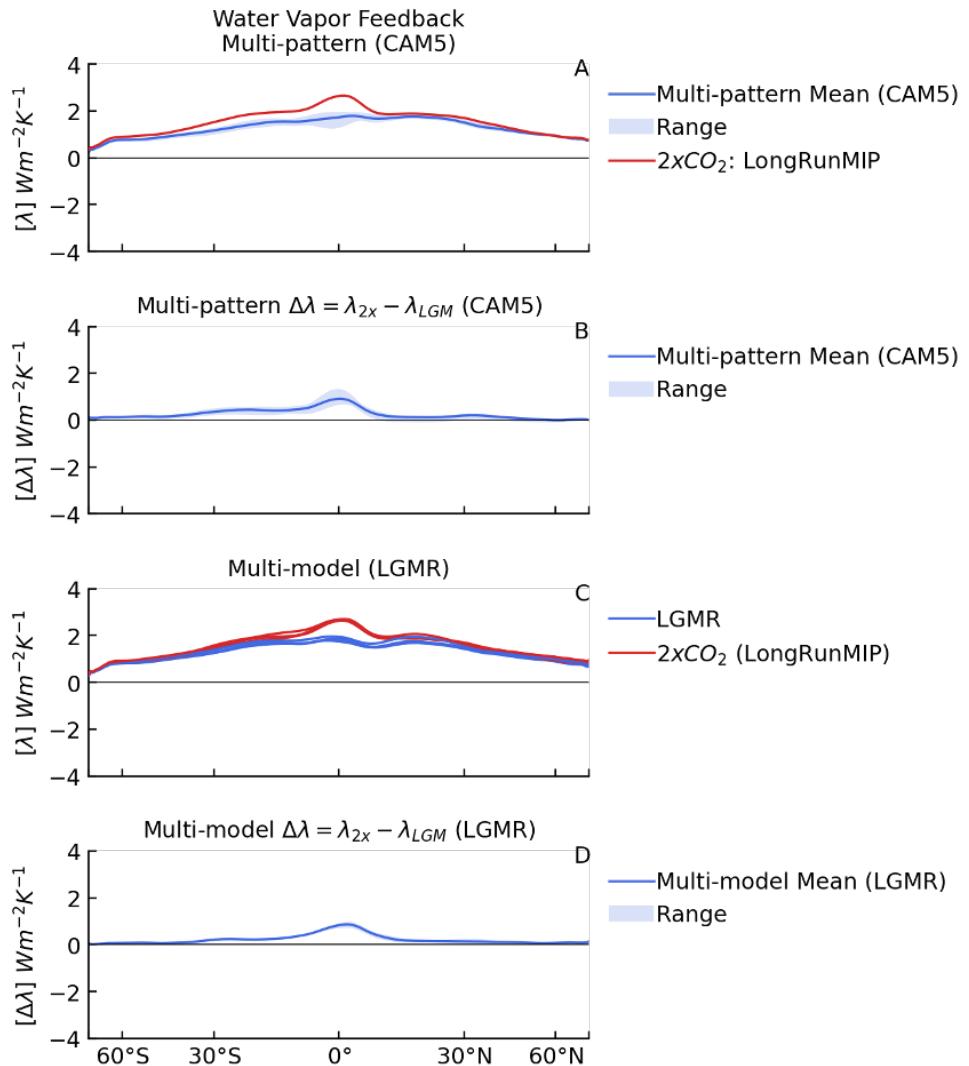
Figure 2.S18: Zonal-mean water vapor feedback and $\Delta\lambda$. See caption of Figure 2.S12.
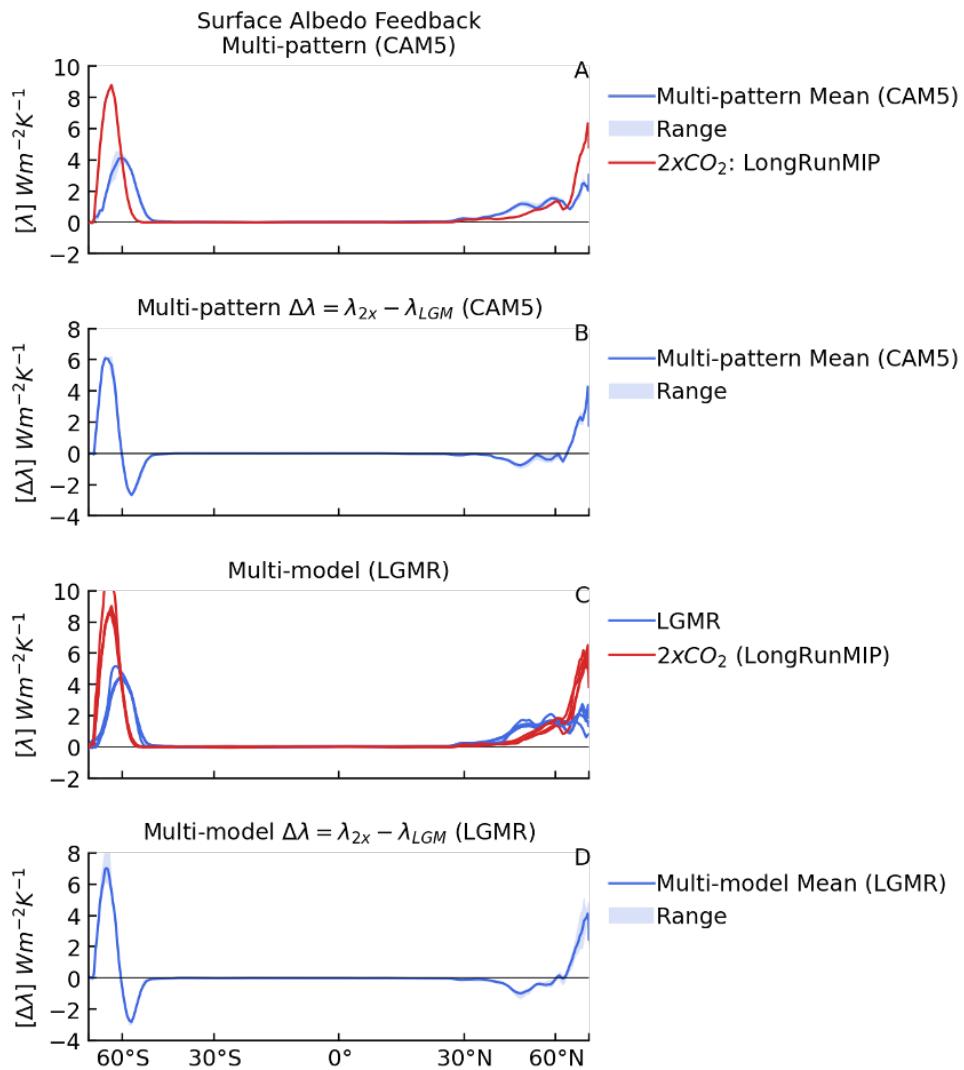
Figure 2.S19: Zonal-mean surface albedo feedback and $\Delta\lambda$. See caption of Figure 2.S12.
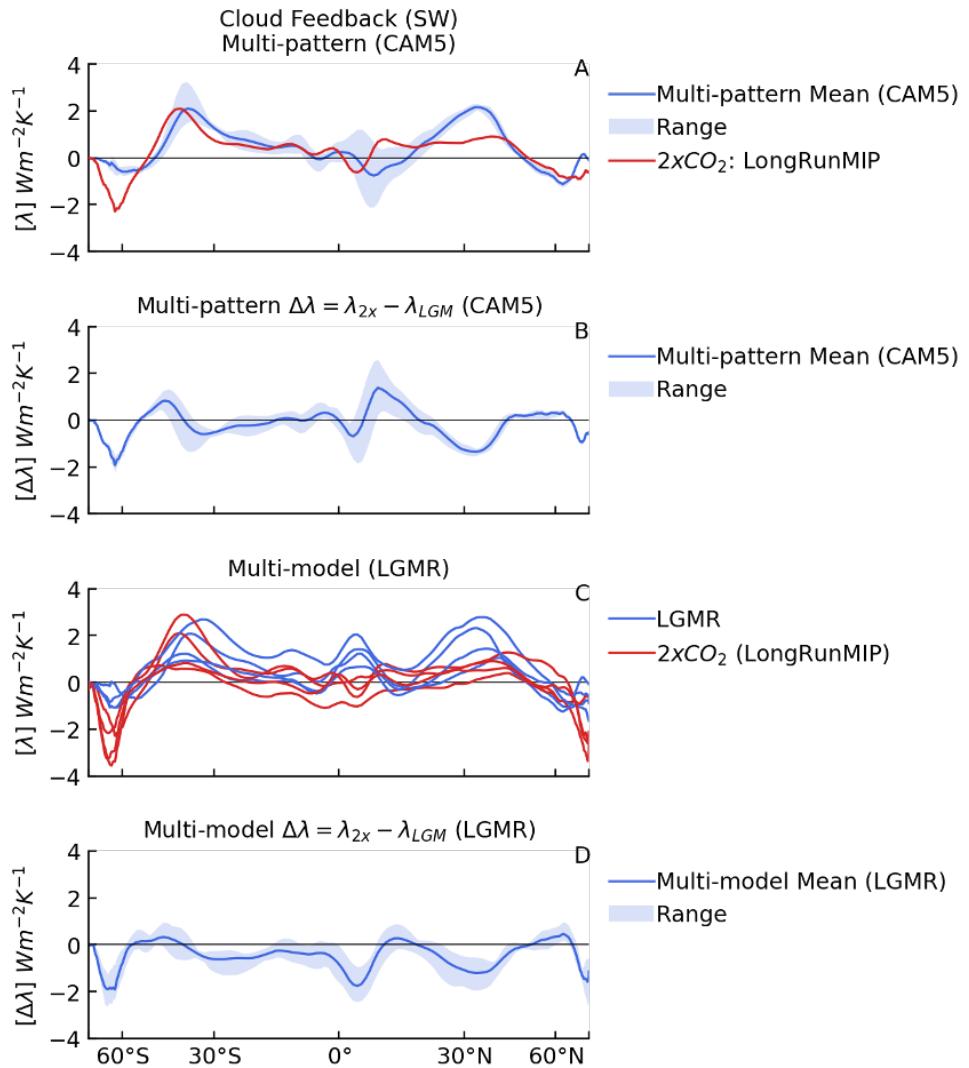
Figure 2.S20: Zonal-mean shortwave cloud feedback and $\Delta\lambda$. See caption of Figure 2.S12.
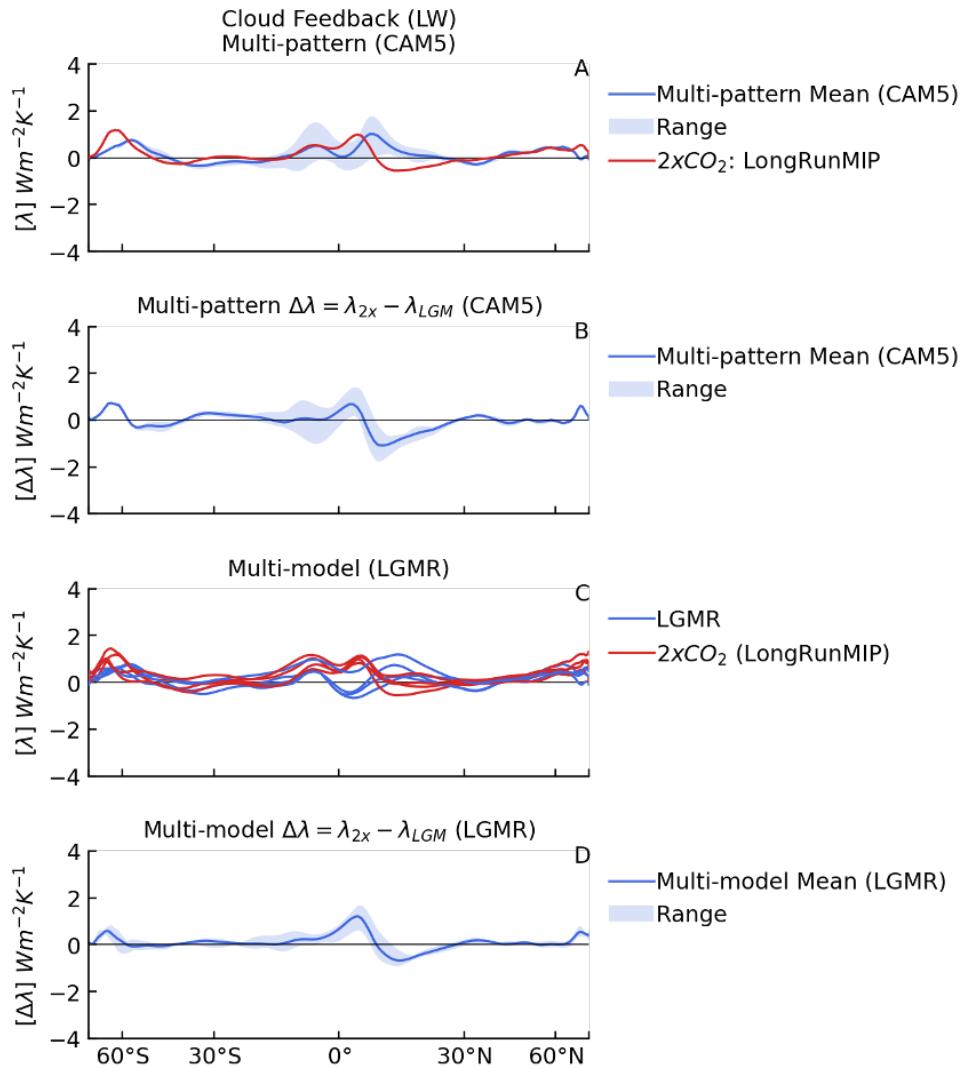
Figure 2.S21: Zonal-mean longwave cloud feedback and $\Delta\lambda$. See caption of Figure 2.S12.
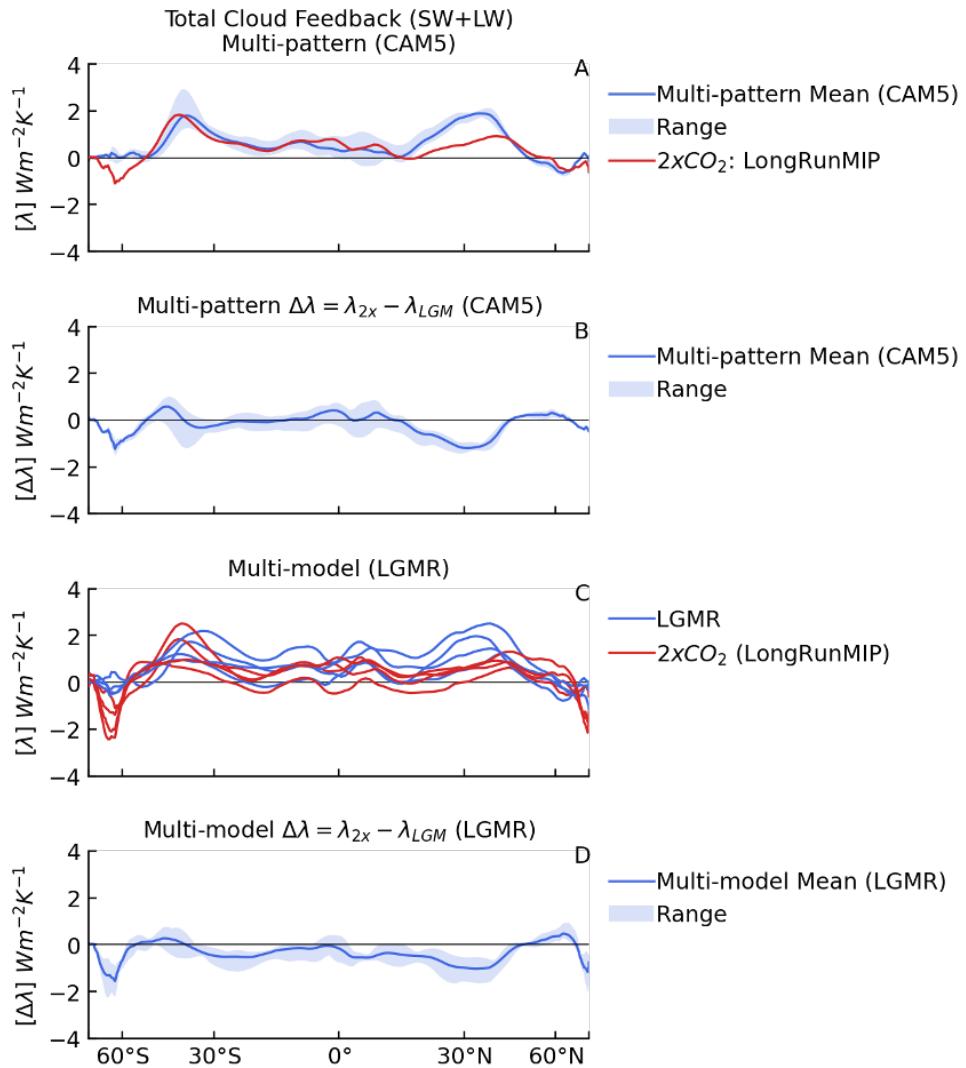
Figure 2.S22: Zonal-mean total cloud feedback and $\Delta\lambda$. See caption of Figure 2.S12.

2.7.6 Supplemental Tables

| $[Wm^{-2}K^{-1}]$	$\Delta\lambda = \lambda_{2x} - \lambda_{LGM}$	λ_{2x} LongRunMIP	λ_{LGM} LGMR	$\Delta\lambda = \lambda_{4x(150yr)}/(1+\zeta) - \lambda_{LGM}$	$\lambda_{4x(150yr)}$	$\varepsilon = \lambda_{2x}/\lambda_{LGM}$ [unitless]
CAM4	−0.45	−1.47	−1.02	−0.14	−1.23	1.44
CAM5	−0.31	−1.05	−0.74	−0.35	−1.15	1.42
CAM6	−0.63	−0.83	−0.19	−0.43	−0.66	4.37
GFDL-AM4	−0.33	−0.92	−0.60	−0.22	−0.86	1.53
HadGEM3-GC3.1-LL	−0.27	−0.62	−0.34	−0.25	−0.63	1.82
Mean	**−0.40**	**−0.98**	**−0.58**	**−0.28**	**−0.91**	**2.12**
Std. Dev.	*0.15*	*0.32*	*0.32*	*0.11*	*0.28*	*1.27*

Table 2.S1: LGM pattern effect and climate feedbacks in various AGCMs. LGM pattern effect ($\Delta\lambda$) calculated as the difference in net feedbacks (λ) from $2\times CO_2$ and LGM. λ_{2x} is calculated in AGCM simulations with LongRunMIP (Rugenstein et al., 2019) $2\times CO_2$ pattern of SST/SIC. λ_{LGM} is calculated in AGCM simulations with LGMR (Osman et al., 2021) pattern. Alternative values for $\Delta\lambda$ are shown using 150-year regression of abrupt-$4\times CO_2$ from coupled models corresponding to each AGCM (Andrews et al., 2022). ζ is assumed to be 0.06 based on WCRP20's central estimate (Sherwood et al., 2020). Efficacy, ϵ, shown in right column. Note that CAM6 is an outlier in efficacy calculations.

	$\Delta\lambda=\lambda_{2x}-\lambda_{\text{LGM}}$ $Wm^{-2}K^{-1}$	λ $Wm^{-2}K^{-1}$	$\Delta\overline{\text{SST}}$ K	$\Delta\overline{\text{T}}$ K	$\Delta\overline{\text{N}}$ Wm^{-2}	$\Delta\lambda=\lambda_{4x(150yr)}/(1+\zeta)-\lambda_{\text{LGM}}$ $Wm^{-2}K^{-1}$	$\varepsilon=\lambda_{2x}/\lambda_{\text{LGM}}$
CAM4							
LGMR	−0.45	−1.02	−3.79	−5.06	5.14	−0.14	1.44
lgmDA	−0.69	−0.78	−3.14	−4.16	3.24	−0.38	1.88
Amrhein	−0.48	−0.99	−2.21	−3.38	3.36	−0.17	1.48
Annan	−0.29	−1.17	−2.18	−3.36	3.95	0.01	1.26
Mean$_{\text{CAM4}}$	**−0.48**	**−0.99**	**−2.83**	**−3.99**	**3.92**	**−0.17**	**1.52**
StdDev$_{\text{CAM4}}$	*0.16*	*0.16*	*0.78*	*0.80*	*0.87*	*0.16*	*0.26*
2xCO$_2$	—	*−1.47*	*2.35*	*3.08*	*−4.52*	—	—
CAM5							
LGMR	−0.31	−0.74	−3.79	−5.15	3.81	−0.35	1.42
lgmDA	−0.51	−0.54	−3.14	−4.24	2.27	−0.55	1.94
Amrhein	−0.33	−0.72	−2.21	−3.40	2.44	−0.37	1.46
Annan	−0.09	−0.97	−2.18	−3.38	3.28	−0.11	1.08
Mean$_{\text{CAM5}}$	**−0.31**	**−0.74**	**−2.83**	**−4.05**	**2.95**	**−0.34**	**1.48**
StdDev$_{\text{CAM5}}$	*0.18*	*0.18*	*0.78*	*0.84*	*0.72*	*0.18*	*0.35*
2xCO$_2$	—	*−1.05*	*2.35*	*3.09*	*−3.24*	—	—
Mean$_{\text{CAM4&5}}$	**−0.39**	**−0.86**	**−2.83**	**−4.01**	**3.41**	**−0.26**	**1.50**
StdDev$_{\text{CAM4&5}}$	*0.21*	*0.21*	*0.72*	*0.76*	*0.90*	*0.18*	*0.29*

Table 2.S2: LGM pattern effect and climate feedbacks from various SST patterns. LGM pattern effect ($\Delta\lambda$) from net feedbacks (λ) in 2×CO$_2$ and with various LGM patterns of SST/SIC. λ_{2x} is calculated in AGCMs with LongRunMIP (Rugenstein et al., 2019) 2×CO$_2$ pattern of SST/SIC. λ_{LGM} is calculated in AGCM simulations with four LGM patterns. Global-mean anomalies for SST, near-surface air temperature (T), and top-of-atmosphere radiative imbalance (N) are shown for reference. Values for the LGM pattern effect are also shown using 150-year regression of abrupt-4×CO$_2$ from coupled models (Andrews et al., 2022). ζ is assumed to be 0.06 based on the WCRP20 central estimate (Sherwood et al., 2020). Efficacy, ϵ, is shown in the right column.

$Wm^{-2}K^{-1}$	$\lambda_{2x}^{-0.5K}$	$\lambda_{LGM}^{-0.5K}$	$\Delta\lambda_{Only}^{Pattern} =$ $\lambda_{2x}^{-0.5K} - \lambda_{LGM}^{-0.5K}$	$\Delta\lambda_T = \Delta\lambda - \Delta\lambda_{Only}^{Pattern}$	$\Delta\lambda = \Delta\lambda_{Only}^{Pattern} + \Delta\lambda_{T},$ $\Delta\lambda = \lambda_{2x} - \lambda_{LGM}$
CAM4	-1.98	-1.55	-0.42	-0.03	-0.45
CAM5	-1.59	-1.24	-0.35	0.04	-0.31
CAM6	-1.30	-0.55	-0.75	0.12	-0.63
Mean	$\mathbf{-1.63}$	$\mathbf{-1.12}$	$\mathbf{-0.51}$	$\mathbf{0.04}$	$\mathbf{-0.47}$

Table 2.S3: Climate feedbacks and temperature dependence from pattern-only simulations. $\Delta\lambda_{\mathrm{PatternOnly}}$ from pattern-only simulations, where LongRunMIP (Rugenstein et al., 2019) $2\times CO_2$ and LGMR (Osman et al., 2021) patterns of SST anomalies are scaled to global-mean ΔSST of -0.5 K. Feedback dependence on global-mean temperature ($\Delta\lambda_T$) is estimated as the residual between $\Delta\lambda$ in main simulations and $\Delta\lambda_{\mathrm{PatternOnly}}$, i.e., assuming $\Delta\lambda = \Delta\lambda_{\mathrm{PatternOnly}} + \Delta\lambda_T$. Note that total $\Delta\lambda = \lambda_{2x} - \lambda_{\mathrm{LGM}}$.

PDF from combined lines of evidence (units, K)	5th %	17th %	50th %	83rd %	95th %	Mean
Assuming $\Delta T_{LGM} \sim N(\mu = -5.0, \sigma = 1.0)$ K, as in WCRP20						
WCRP20 Baseline (uniform-λ prior)	2.3	2.6	3.1	3.9	4.7	3.2
… with Revised $\Delta\lambda_{LGM} \sim N(-0.37, 0.23)$ Wm^{-2}K^{-1}	2.1	2.3	2.8	3.4	4.0	2.9
… … and 2x uncertainty, $\Delta\lambda_{LGM} \sim N(-0.37, 0.46)$ Wm^{-2}K^{-1}	2.1	2.4	2.9	3.6	4.3	3.0
… … based on $\lambda_{4x150yr}$, $\Delta\lambda_{LGM} \sim N(-0.27, 0.20)$ Wm^{-2}K^{-1}	2.1	2.4	2.8	3.4	4.0	2.9
WCRP20 (uniform-*S* prior)	2.4	2.8	3.5	4.5	5.7	3.7
... with Revised $\Delta\lambda_{LGM} \sim N(-0.37, 0.23)$ Wm^{-2}K^{-1}	2.2	2.5	3.0	3.8	4.6	3.2
… … and 2x uncertainty, $\Delta\lambda_{LGM} \sim N(-0.37, 0.46)$ Wm^{-2}K^{-1}	2.3	2.6	3.2	4.1	5.1	3.4
… … based on $\lambda_{4x150yr}$, $\Delta\lambda_{LGM} \sim N(-0.27, 0.20)$ Wm^{-2}K^{-1}	2.2	2.5	3.1	3.8	4.6	3.2
Assuming $\Delta T_{LGM} \sim N(-6.0, 1.0)$ K						
WCRP20 Baseline (uniform-λ prior)	2.3	2.7	3.2	4.1	5.0	3.4
... with Revised $\Delta\lambda_{LGM} \sim N(-0.37, 0.23)$ Wm^{-2}K^{-1}	2.1	2.4	2.9	3.5	4.1	3.0
… … and 2x uncertainty, $\Delta\lambda_{LGM} \sim N(-0.37, 0.46)$ Wm^{-2}K^{-1}	2.2	2.5	3.0	3.7	4.4	3.1
… … based on $\lambda_{4x150yr}$, $\Delta\lambda_{LGM} \sim N(-0.27, 0.20)$ Wm^{-2}K^{-1}	2.2	2.4	2.9	3.5	4.2	3.0
WCRP20 (uniform-*S* prior)	2.5	2.9	3.7	4.8	6.1	3.9
... with Revised $\Delta\lambda_{LGM} \sim N(-0.37, 0.23)$ Wm^{-2}K^{-1}	2.3	2.6	3.1	3.9	4.7	3.3
… … and 2x uncertainty, $\Delta\lambda_{LGM} \sim N(-0.37, 0.46)$ Wm^{-2}K^{-1}	2.3	2.7	3.3	4.3	5.3	3.5
… … based on $\lambda_{4x150yr}$, $\Delta\lambda_{LGM} \sim N(-0.27, 0.20)$ Wm^{-2}K^{-1}	2.3	2.6	3.2	4.0	4.8	3.3
PDF from LGM evidence alone (uniform-*S* prior)	5th %	17th %	50th %	83rd %	95th %	Mean
Assuming $\Delta T_{LGM} \sim N(-5.0, 1.0)$ K as in WCRP20						
WCRP20	1.7	2.4	4.5	10.6	16.5	6.2
... with Revised $\Delta\lambda_{LGM} \sim N(-0.37, 0.23)$ Wm^{-2}K^{-1}	1.1	1.5	2.1	3.0	4.2	2.3
… … and 2x uncertainty, $\Delta\lambda_{LGM} \sim N(-0.37, 0.46)$ Wm^{-2}K^{-1}	1.2	1.5	2.4	4.7	10.0	3.4
Assuming $\Delta T_{LGM} \sim N(-6.0, 1.0)$ K						
WCRP20	2.3	3.4	6.8	13.9	18.0	8.2
... with Revised $\Delta\lambda_{LGM} \sim N(-0.37, 0.23)$ Wm^{-2}K^{-1}	1.4	1.7	2.4	3.5	5.0	2.7
… … and 2x uncertainty, $\Delta\lambda_{LGM} \sim N(-0.37, 0.46)$ Wm^{-2}K^{-1}	1.4	1.9	3.0	6.4	13.0	4.3

Table 2.S4: Summary Statistics for Posterior PDFs of Climate Sensitivity. Note: The posterior PDF from LGM evidence alone uses the uniform-*S* prior $(0, 20)$ K, hence the shape of the posterior PDF matches that of the LGM likelihood. Methods follow WCRP20.

Chapter 3

PALEOCLIMATE PATTERN EFFECTS HELP CONSTRAIN CLIMATE SENSITIVITY AND 21ST-CENTURY WARMING

This work is in review at *Proceedings of the National Academy of Sciences* as: Cooper, V., K. Armour, G. Hakim, J. Tierney, N. Burls, C. Proistosescu, T. Andrews, W. Dong, M. Dvorak, R. Feng, M. Osman, Y. Dong. Paleoclimate pattern effects help constrain climate sensitivity and 21st-century warming. *In review at Proceedings of the National Academy of Sciences.*

3.1   Abstract

Paleoclimates provide examples of past climate change that inform estimates of modern warming from greenhouse-gas emissions, known as Earth's climate sensitivity. However, differences between past and present climate change must be accounted for when inferring climate sensitivity from paleoclimate evidence (PALAEOSENS Project Members, 2012; Sherwood et al., 2020; Cooper et al., 2024). The Pliocene (5.3–2.6 Ma), a warm epoch with atmospheric CO_2 concentrations similar to today, is a potential analog for modern warming (Burke et al., 2018). Recent reconstructions indicate the Pliocene was 1°C warmer than previously thought (Tierney et al., 2025b; Annan et al., 2024), implying higher climate sensitivity (Tierney et al., 2025b), supported by reconstructions of more global cooling from reduced CO_2 at the Last Glacial Maximum (LGM; 19–23 thousand years ago) (Tierney et al., 2020; Osman et al., 2021; Seltzer et al., 2021). However, these same reconstructions indicate large-scale patterns of paleoclimate temperature change differ strongly from modern projections. Climate feedbacks and sensitivity depend on temperature patterns (e.g., Armour et al., 2013; Andrews et al., 2015; Gregory and Andrews, 2016; Ceppi and Gregory, 2017; Andrews and Webb, 2018), and such "pattern effects" must be accounted for when using paleoclimates to constrain modern climate sensitivity (Chapter 2) (Cooper et al., 2024). Here we combine data-assimilation reconstructions with atmosphere models to show Earth's climate is more sensitive to Pliocene forcing than modern CO_2 forcing. Pliocene ice sheets, topography, and vegetation alter patterns of

ocean warming and excite destabilizing cloud feedbacks, and LGM feedbacks are similarly amplified by the North American ice sheet. Accounting for paleoclimate pattern effects produces a best estimate (median) for modern climate sensitivity of $2.8°C$, 66% range $2.4 - 3.4°C$ (90% CI: $2.1 - 4.0°C$), substantially reducing uncertainty and narrowing projections of 21st-century warming.

3.2 Introduction

The paleoclimate record constitutes a series of natural experiments with fundamental insights into Earth's climate sensitivity. Using paleoclimate evidence to constrain the modern sensitivity to rising greenhouse-gas (GHG) concentrations requires accounting for differences in both climate forcings and feedbacks between the past and modern climates (PALAEOSENS Project Members, 2012; Sherwood et al., 2020; Cooper et al., 2024). A key driver of such feedback differences across past climates is variation in the spatial pattern of sea-surface temperature, i.e., "paleoclimate pattern effects" (Chapter 2) (Cooper et al., 2024). Pattern effects are variations in climate sensitivity and feedbacks that depend on spatial patterns of temperature change (e.g., Armour et al., 2013; Andrews et al., 2015; Gregory and Andrews, 2016; Ceppi and Gregory, 2017; Andrews and Webb, 2018), and they arise in paleoclimate contexts when non-CO_2 forcings (such as ice sheets, topography, and vegetation) affect large-scale temperature patterns. Paleoclimate pattern effects can have major impacts on estimates of modern climate sensitivity if non-CO_2 forcings strongly influence the temperature pattern, thereby producing climate feedbacks that differ from those that govern modern warming from GHG forcing (Chapter 2) (Cooper et al., 2024).

The Pliocene (5.3–2.6 Ma) is the closest analog to near-term warming (Burke et al., 2018). Its mid-Piacenzian warm period (c. 3.3–3.0 Ma), hereafter "Pliocene," is the most recent epoch with atmospheric concentrations of CO_2 (near 400 ppm) that are similar to present day (de la Vega et al., 2020). The magnitude of Pliocene warming thus provides an important constraint on the equilibrium climate sensitivity (ECS) of the modern climate, which is the steady-state response of global-mean near-surface air temperature to a doubling of atmospheric CO_2 from preindustrial levels (Sherwood et al., 2020; Forster et al., 2021). Previous assessments of paleoclimate proxies report approximately $3°C$ of global-mean Pliocene warming and an upper bound of $4°C$ relative to preindustrial conditions (Sherwood et al., 2020; Forster et al., 2021). However, recent reconstructions find a much warmer Pliocene, with central estimates of $4°C$ (Tierney et al., 2025b; Annan

et al., 2024). This revision to Pliocene warming suggests a much higher ECS of 4.8°C (Tierney et al., 2025b), implying increased likelihood of realizing the worst-case projections of 21st-century warming. But these globally resolved reconstructions tell us more than global means—they capture the spatial pattern of Pliocene warming, and this spatial information is essential to constraining modern ECS.

To infer modern ECS from Pliocene evidence, we must consider differences in both forcing and feedbacks between the Pliocene and present climate. The Pliocene has both elevated GHG levels (de la Vega et al., 2020; Hopcroft et al., 2020) as well as additional forcing from (i) reduced ice sheets over West Antarctica and Greenland, (ii) increased vegetation, especially over northern high latitudes, and (iii) changes in land-sea distribution (Salzmann et al., 2013; Dowsett et al., 2016; PALAEOSENS Project Members, 2012; Sherwood et al., 2020). Previous work found that the Pliocene's global-mean warming is mostly attributable to CO_2 (Lunt et al., 2012; Tierney et al., 2019; Burton et al., 2023). However, modeling studies show that the non-CO_2 forcings drive distinct climate responses especially at regional scales (Dvorak et al., 2025; Burton et al., 2023; Weiffenbach et al., 2023; Feng et al., 2022; Menemenlis et al., 2021; Lunt et al., 2012, 2010), and that Pliocene temperature patterns may differ substantially from those in response to modern CO_2 forcing (Dvorak et al., 2025), thereby producing different climate feedbacks. Accounting for such pattern effects in the Last Glacial Maximum (LGM), a cold period $19 - 23$ ka, led to stronger constraints on modern ECS (Chapter 2) (Cooper et al., 2024). The key question addressed here is: would accounting for Pliocene pattern effects also strengthen constraints on modern ECS?

We quantify Pliocene pattern effects by synthesizing proxy data with climate models, and we use these results to revise estimates of modern ECS and 21st-century warming. Spatially complete reconstructions of the Pliocene (Tierney et al., 2025b; Annan et al., 2024) from paleoclimate data assimilation (Hakim et al., 2016; Tierney et al., 2020; Osman et al., 2021) are used in numerical simulations with five atmospheric general circulation models (AGCMs) to quantify relationships between temperature patterns and climate feedbacks (e.g., Andrews et al., 2015; Cooper et al., 2024). We analyze differences between feedbacks in the Pliocene compared to modern warming from CO_2. We then combine our Pliocene results with an investigation of the LGM (Chapter 2) (Cooper et al., 2024), and we quantify the impacts of the feedback differences on estimates of modern ECS and projections of 21st-century warming.

3.3 Paleoclimate pattern effects and modern ECS

Modern ECS, climate feedbacks, and paleoclimate pattern effects are related through the global-mean energy balance,

$$\Delta N = \Delta F + \lambda \Delta T, \tag{3.1}$$

where ΔN is the change in top-of-atmosphere radiative balance; ΔF is the "effective" radiative forcing, i.e., the change in net downward radiative flux after atmospheric adjustments to imposed perturbations, excluding radiative responses to changing surface temperature (Forster et al., 2021); λ is the net climate feedback (negative for stable climates); and ΔT is the change in near-surface air temperature. All values are global means, and differences (Δ) are relative to the preindustrial baseline. When the forcing is a doubling of preindustrial CO_2 concentrations (2xCO_2), and the climate reaches equilibrium ($\Delta N = 0$), the resulting ΔT is the modern ECS:

$$ECS = -\Delta F_{2xCO_2}/\lambda_{2xCO_2}, \tag{3.2}$$

where ΔF_{2xCO_2} is the effective radiative forcing and λ_{2xCO_2} is the net feedback from modern CO_2 doubling. Increasingly negative values of λ indicate more-stable climates and lower ECS.

Paleoclimate pattern effects ($\Delta\lambda$) are quantified as the difference between λ_{2xCO_2} and a paleoclimate feedback, e.g., the Pliocene feedback (λ_{Plio}), due to differences in the spatial patterns of warming:

$$\Delta\lambda = \lambda_{2xCO_2} - \lambda_{Plio}. \tag{3.3}$$

$\Delta\lambda$ also can vary with global-mean temperature (e.g., Caballero and Huber, 2013; Sherwood et al., 2020; Cooper et al., 2024). However, this temperature dependence can be omitted for the Pliocene due to similar levels of global warming from Pliocene and 2xCO_2 forcings (Sherwood et al., 2020), and it is relatively small for LGM levels of global cooling (Cooper et al., 2024; Eisenman and Armour, 2024).

Modern ECS and λ_{2xCO_2} can be constrained by estimating λ_{Plio} and $\Delta\lambda$, then combining Equations 3.2 and 3.3:

$$ECS = -\Delta F_{2xCO_2}/(\lambda_{Plio} + \Delta\lambda). \tag{3.4}$$

$\Delta\lambda$ depends on spatial patterns of Pliocene temperature anomalies, for which we use state-of-the-art reconstructions from data assimilation (Tierney et al., 2025b; Annan et al., 2024) as boundary

64

conditions for simulations using five AGCMs, as described in the following section.

3.4   Pliocene pattern effects constrained by data assimilation

3.4.1   Patterns of Pliocene sea-surface temperature

In Fig. 3.1, we compare the projected sea-surface temperature (SST) anomalies from modern $2xCO_2$, based on the multi-model mean of quasi-equilibrium simulations in LongRunMIP (Rugenstein et al., 2019), with the various Pliocene reconstructions from "plioDA" (Tierney et al., 2025b) and Annan et al. (2024) that we use to quantify Pliocene pattern effects. The Pliocene patterns include the best estimates from plioDA (Tierney et al., 2025b) and Annan et al. (2024), as well as alternate plioDA reconstructions that test structural uncertainty and endmembers of the plioDA ensemble (Fig. 3.1; Fig. 3.S1–S4) (Methods). The reconstructions use paleoclimate data assimilation (Hakim et al., 2016; Tierney et al., 2020; Osman et al., 2021), which optimally combines dynamical constraints from model "priors" with proxy data. Data assimilation results depend on specific aspects of the methods, model priors (Amrhein et al., 2020), and observations.

To address reconstruction uncertainty, we analyze pattern effects across a wide range of possible Pliocene temperature patterns that use different assimilation methods, model priors, and subsets of proxy data. Focusing on sensitivity to the model prior, the "PlioMIP2 Prior" version of plioDA uses 14 PlioMIP2 simulations (Haywood et al., 2020) to inform its prior. The "Perturbed Cloud Prior" uses 21 simulations that are designed to capture Pliocene temperature gradients by substantially altering models' cloud physics (Burls and Fedorov, 2014; Erfani and Burls, 2019; Ford et al., 2022). Focusing on sensitivity to the proxy network, the "PlioVar Data" version restricts data to the KM5c interglacial (McClymont et al., 2020), and we also test endmembers of the plioDA ensemble (Fig. 3.S4) (Methods). Annan et al. (2024) and plioDA (Tierney et al., 2025b) have partially overlapping proxy networks, model priors (both best estimates include simulations from PlioMIP2), and assimilation methods (ensemble Kalman filter); however, there are substantial differences between the two reconstruction efforts in terms of the proxies included, model priors, and methods (e.g., forward modeling of proxies in plioDA) that lead to differences in their results (Tierney et al., 2025b) (Fig. 3.1b,f).

Despite the substantial uncertainty in the details of the Pliocene SST patterns shown in Fig. 3.1,
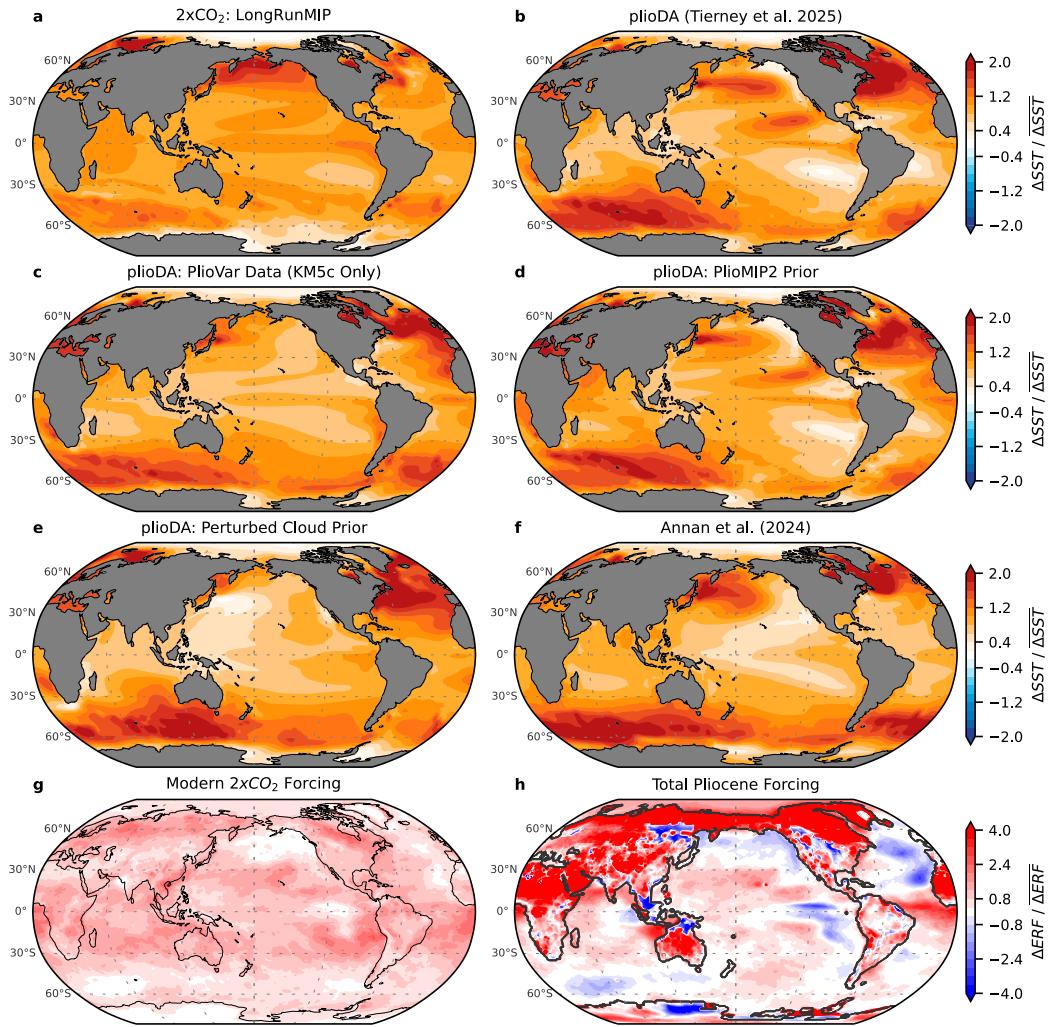
Figure 3.1: Patterns of sea-surface temperature (SST) anomalies and effective radiative forcing (ERF). **(a)** Multi-model mean of modern SST response to $2xCO_2$ in quasi-equilibrium simulations from LongRunMIP (Rugenstein et al., 2019). **(b–f)** Data-assimilation reconstructions from: **(b)** plioDA best estimate (Tierney et al., 2025b); alternate plioDA using **(c)** only the PlioVar proxy data representing the KM5c interglacial, **(d)** only the PlioMIP2 prior, or **(e)** only the perturbed-cloud prior; and **(f)** best estimate from Annan et al. (2024). ERF from **(g)** modern $2xCO_2$ and **(h)** Pliocene total forcing, including greenhouse gases, reduced Greenland and Antarctic ice sheets, sea level, and vegetation (Dvorak et al., 2025). All panels show annual-mean anomalies, and local values are divided by global means. Pliocene SSTs are infilled to modern coastlines.

the reconstructions all have two common features that distinguish the Pliocene from the modern response to 2xCO$_2$: the Pliocene has amplified SST warming in the Southern Ocean and the North Atlantic Ocean (Fig. 3.1; Fig. 3.S1). The distinct Pliocene warming pattern is driven by the distinct spatial pattern of Pliocene *forcing* (Fig. 3.1h) (Dvorak et al., 2025), which arises from the Pliocene's non-CO$_2$ forcings (changes in ice sheets, topography, and vegetation) and differs substantially from the relatively uniform forcing produced by CO$_2$ alone (Fig. 3.1g). The connection between the non-CO$_2$ Pliocene forcings and the SST patterns they produce has been demonstrated in coupled climate models (Dvorak et al., 2025), which we return to in the Discussion.

3.4.2 Quantifying feedbacks and pattern effects

We estimate the net climate feedback, λ, for each warming pattern in Fig. 3.1 using AGCM simulations with prescribed SST and sea-ice concentration (SIC) (Methods). Following Chapter 2 (Cooper et al., 2024), we begin with a control simulation using the preindustrial "baseline" pattern (Osman et al., 2021). We repeat the AGCM simulations, changing only the SST and SIC to the 2xCO$_2$ pattern from LongRunMIP (Fig. 3.1a) and to each of the Pliocene patterns (Fig. 3.1b–e; SIC in Fig. 3.S2–S4). We hold the forcings constant at modern levels across all simulations to isolate the radiative response to changes in surface temperature (Methods). For each simulation, we calculate ΔN and ΔT relative to the preindustrial baseline, and the net feedback is $\lambda = \Delta N / \Delta T$ from Eq. 3.1 with $\Delta F = 0$.

In Fig. 3.2, we compare $\lambda_{2\text{xCO}_2}$ with λ_{Plio} and quantify Pliocene pattern effects ($\Delta\lambda$). In all five AGCMs, λ_{Plio} is more positive (destabilizing) than $\lambda_{2\text{xCO}_2}$, which means that the climate system is more sensitive to Pliocene forcing than it is to modern 2xCO$_2$ forcing. We test whether this result is robust despite uncertainties in atmospheric model physics and Pliocene reconstructions by running the simulations in CAM4, CAM5, CAM6, GFDL-AM4, and HadGEM3-GC3.1-LL, and by testing three different Pliocene patterns (Fig 3.1B,D,F) in all five AGCMs. We test additional Pliocene patterns, including the 5th and 95th percentiles of the plioDA ensemble (Fig. 3.S4), in CAM4 and CAM5 (Methods). Despite the uncertainties in Pliocene SST patterns and atmospheric model physics, there is a clear Pliocene pattern effect with $\Delta\lambda < 0$ (Fig. 3.2b), albeit with uncertain magnitude.
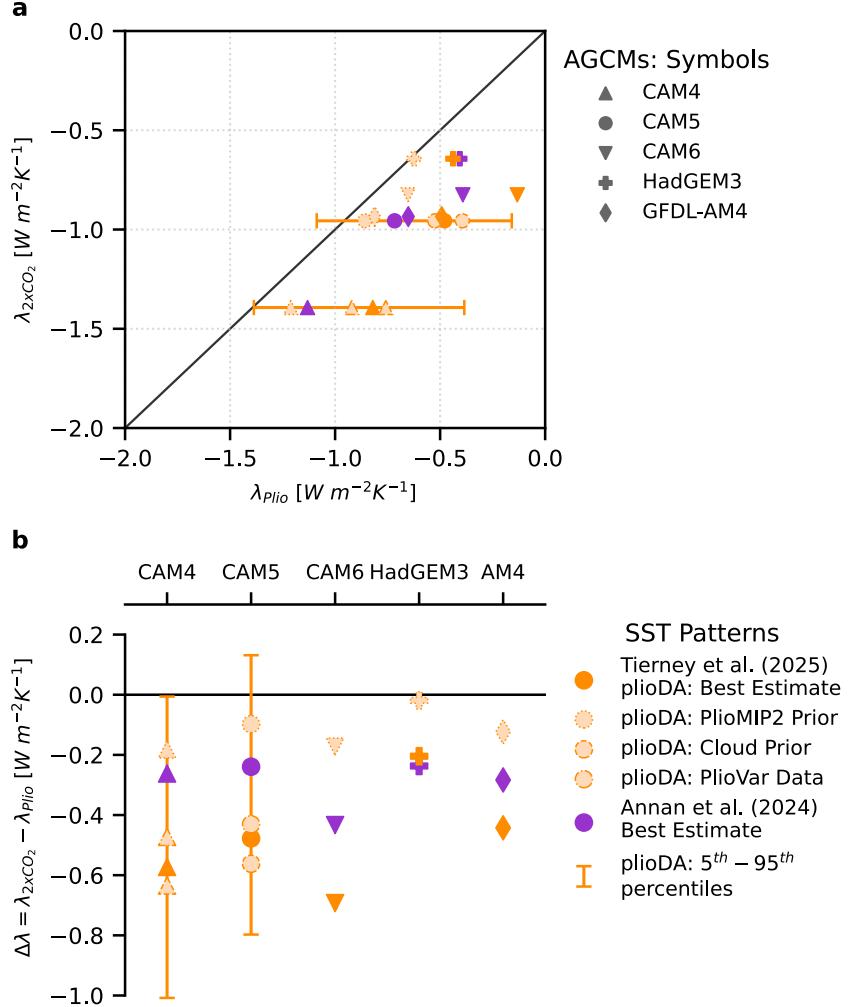
Figure 3.2: Net climate feedbacks (λ) and Pliocene pattern effect ($\Delta\lambda$). Note that each legend applies to both panels; different atmospheric general circulation models (AGCMs) are denoted by symbols, and different Pliocene warming patterns are denoted by colors and borders. **(a)** Scatter plot of λ_{2xCO_2} versus λ_{Plio} for each AGCM and Pliocene pattern, with $\lambda_{2xCO_2} = \lambda_{Plio}$ shown as solid line. **(b)** Pliocene pattern effect, $\Delta\lambda = \lambda_{2xCO_2} - \lambda_{Plio}$, using values in panel **a**. Error bars for plioDA represent endmembers of the ensemble reconstruction (Methods).

In summary, the Pliocene warming pattern excites more positive (destabilizing) climate feedbacks compared to the 2xCO$_2$ warming pattern ($\lambda_{Plio} > \lambda_{2xCO_2}$), i.e., the Pliocene pattern effect is negative ($\Delta\lambda < 0$). As will be shown below, the negative pattern effect indicates that positive feed-

backs amplifying Pliocene warming do not play an equivalent role in the modern climate's response to greenhouse-gas forcing. Accounting for this negative Pliocene pattern effect would lead to lower estimates of modern ECS and future warming (Eq. 3.4) (see Chapter 2; Cooper et al., 2024).

3.5   Discussion

3.5.1   Mechanisms responsible for Pliocene pattern effects

To diagnose the mechanisms contributing to more-positive climate feedbacks in the Pliocene, we first use radiative kernels to assess each component feedback within the AGCM simulations (Methods) (Soden et al., 2008). We find that the cloud feedback (λ_{cloud}), namely the shortwave component associated with low clouds, is the dominant driver of $\lambda_{\text{Plio}} > \lambda_{\text{2xCO}_2}$ (Fig. 3.S5–S6). The combined lapse-rate and water-vapor feedbacks make an additional contribution to more-positive λ_{Plio} (Fig. 3.S5). Next, we inspect the spatial distribution of the Pliocene's more-positive cloud feedbacks to understand their source.

In Fig. 3.3, we compare the spatial patterns of λ_{cloud} in the Pliocene versus 2xCO$_2$. The most pronounced differences are over the Southern Ocean (Indian sector) and the North Atlantic. The zonal mean of $\Delta\lambda_{\text{cloud}}$ (Fig. 3.3a) illustrates that the Pliocene's extratropical cloud feedbacks are responsible for $\lambda_{\text{Plio}} > \lambda_{\text{2xCO}_2}$, supported by extratropical lapse-rate feedbacks (Fig. 3.S9). Comparing Fig. 3.3's λ_{cloud} with Fig. 3.1's SST patterns (zonal mean SST in Fig. 3.S10), we see that the regions with amplified Pliocene SST anomalies are approximately collocated with the amplified Pliocene λ_{cloud}. That is, amplified SST anomalies in the extratropics are responsible for more-positive feedbacks in the Pliocene, which is consistent with a similar analysis of the Last Glacial Maximum (Chapter 2) (Cooper et al., 2024). When SST warming is strongly amplified in the extratropics compared to the SST warming in tropical regions of atmospheric deep convection (e.g., the west Pacific warm pool), tropospheric stability is decreased and low-cloud cover is reduced, which is a positive feedback (Ceppi and Gregory, 2017; Dong et al., 2019; Cooper et al., 2024). Past studies of the Pliocene broadly emphasize the zonal SST in the tropical Pacific and meridional temperature gradients (Wara et al., 2005; Fedorov et al., 2006; O'Brien et al., 2014; Brierley et al., 2015; Burls and Fedorov, 2017; Tierney et al., 2019, 2025b), while we find specifically that the amplification of warming in the North Atlantic and especially the Southern Ocean are the dominant
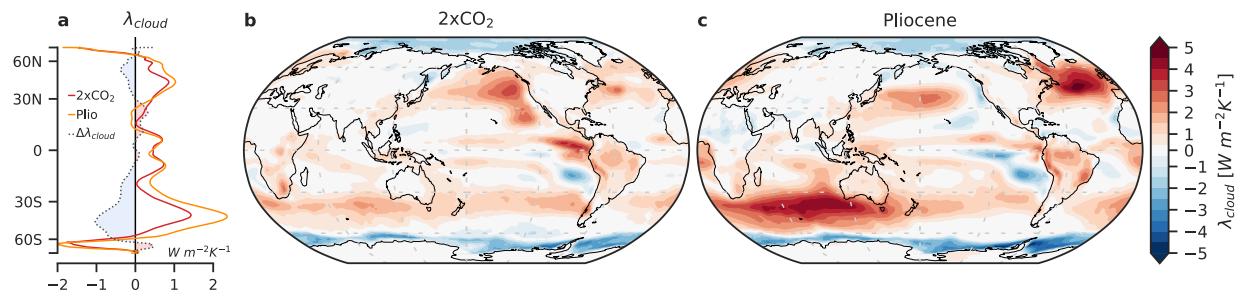
Figure 3.3: Cloud feedbacks from modern CO_2 forcing versus Pliocene warming. **(a)** Zonal means of panels **b**, **c**, and their difference, $\Delta\lambda_{cloud}$; negative values of $\Delta\lambda_{cloud}$ contribute to the negative Pliocene pattern effect. **(b–c)** Spatial distributions of cloud feedbacks, $\lambda_{cloud} = \Delta N_{local}/\overline{\Delta T}$, where ΔN_{local} is the local anomaly in top-of-atmosphere radiation attributable to cloud feedbacks (estimated with radiative kernels), and $\overline{\Delta T}$ is the global-mean T anomaly. Multi-model mean of **(b)** λ_{cloud} using the LongRunMIP $2xCO_2$ pattern and **(c)** multi-pattern mean λ_{cloud} from Pliocene patterns in Fig. 3.1b,d,f (plioDA best estimate (Tierney et al., 2025b), alternate plioDA using only the PlioMIP2 prior, and Annan et al. (2024) best estimate; these patterns were tested in all atmosphere models). All panels show multi-model means across atmosphere models.

features that distinguish Pliocene feedbacks from the modern response to $2xCO_2$.

The final and essential aspect of the mechanism is that amplified warming in the Southern Ocean and North Atlantic is due to non-CO_2 forcings (ice sheets, vegetation, and topography), as shown in Fig. 3.S11. This attribution has been illustrated by simulations in coupled climate models that separate the SST response to Pliocene CO_2 versus non-CO_2 forcings (e.g., Lunt et al., 2012; Haywood et al., 2020; Burton et al., 2023; Dvorak et al., 2025). Pliocene warming in the North Atlantic is amplified by the closure of ocean gateways (Bering Strait and Canadian Archipelago) through changes in the Atlantic Meridional Overturning Circulation (AMOC) (Weiffenbach et al., 2023), and it is further amplified by reductions in ice sheets (Menemenlis et al., 2021). Amplified warming in the Southern Ocean is associated with the reduced Antarctic Ice Sheet and topography through changes in ocean circulation (Weiffenbach et al., 2024; Dvorak et al., 2025). While amplified warming of the Southern Ocean appears in all reconstructions (Fig. 3.1), its magnitude is uncertain due to sparse proxy data, and this uncertainty makes a large contribution to our spread in $\Delta\lambda$ (Fig. 3.S8–S10). Compared to coupled models, both the North Atlantic and Southern Ocean SST features are even more pronounced in data-assimilation reconstructions constrained by paleoclimate proxies (Fig. 3.1) (Tierney et al., 2025b; Annan et al., 2024). Thus coupled models are essential for illustrating mechanisms of paleoclimate pattern effects, and incorporating observational constraints through data assimilation is key to producing reliable SST patterns and constraining $\Delta\lambda$.

While our comparison of the Pliocene versus modern $2xCO_2$ uses the LongRunMIP pattern (Rugenstein et al., 2019), we note that there is substantial uncertainty in the projected SST pattern from $2xCO_2$. However, because Pliocene and LGM pattern effects arise from how non-CO_2 forcings shape paleoclimate temperature patterns, we expect conclusions about $\Delta\lambda$ to be relatively insensitive to uncertainty in the SST pattern from CO_2 forcing. Furthermore, Wang et al. (2025) finds that the feedback uncertainty from CO_2-forced SST patterns is only 10% of the total feedback spread across different models. That result emphasizes the importance of using multiple atmospheric models to quantify $\Delta\lambda$ and that the feedback spread from $2xCO_2$ patterns is small compared to that arising from the Pliocene reconstructions. We test whether results are sensitive to the $2xCO_2$ pattern and find this uncertainty does not affect the conclusions (Methods).

In summary, non-CO_2 forcings from ice sheets, topography, and vegetation altered the spatial pattern of ocean warming, in turn producing positive cloud feedbacks in the extratropics that

strongly amplified global warming during the Pliocene (Fig. 3.3). Because of these amplifying feedbacks, more of the Pliocene warming was caused by non-CO_2 forcings than previously thought, meaning that less of the warming is attributable to elevated CO_2 alone. Since these amplifying feedbacks from non-CO_2 forcing do not play a role in the modern response to 2xCO_2 alone, we now show that accounting for the Pliocene pattern effect lowers estimates of modern ECS and reduces the likelihood of worst-case projections for 21st-century warming.

3.5.2 Modern climate sensitivity and 21st-century warming

To constrain modern ECS with paleoclimate evidence, we first infer climate feedbacks during a paleoclimate period from changes in Earth's energy budget, and then we account for differences relative to the modern response to 2xCO_2 (Sherwood et al., 2020; PALAEOSENS Project Members, 2012; Cooper et al., 2024). Measures of climate sensitivity depend on the timescale of interest, and we follow Sherwood et al. (2020), hereafter "SW20," in focusing on the 150-year timescale of "effective" climate sensitivity (S), and in treating slow paleoclimate feedbacks, e.g., ice sheets, as radiative forcings (PALAEOSENS Project Members, 2012).

First, we estimate λ_{Plio} by applying Equation 3.1 to the Pliocene (Methods). We update ΔT_{Plio} from SW20's values of 3.0 ± 1.0 °C (1σ) to plioDA's result of $\Delta T_{\mathrm{Plio}} = 4.1\pm0.6$ °C (1σ). We also update the non-GHG (greenhouse gas) effective radiative forcing to $\Delta F_{\mathrm{NonGHG}} = 1.7\pm1.0$ (1σ) W m^{-2} (Dvorak et al., 2025). Given that $\Delta F_{\mathrm{GHG}} \approx 2.2$ W m^{-2} (Sherwood et al., 2020; Dvorak et al., 2025), we have a central estimate of total $\Delta F_{\mathrm{Plio}} = 3.9$ W m^{-2} and $\lambda_{\mathrm{Plio}} \approx -1.0$ W m^{-2} K^{-1} (Methods).

The novel aspect of the modern ECS constraint in this study is the inclusion of paleoclimate pattern effects for the Pliocene ($\Delta\lambda$; Eq. 3.3 and 3.4) and the synthesis with pattern effects for the Last Glacial Maximum (Chapter 2) (Cooper et al., 2024). We combine uncertainty across SST patterns and atmospheric models (Fig. 3.2; Methods), which produces a central estimate for Pliocene pattern effects of $\Delta\lambda = -0.37\pm0.32$ (1σ) W m^{-2} K^{-1}. We adapt the Bayesian framework of SW20 to include Pliocene $\Delta\lambda$, following Chapter 2 (Cooper et al., 2024) (Methods).

In Fig. 3.4a, we show the S likelihoods from Pliocene evidence alone. For comparison, we include the original SW20 results and the likelihood with updated Pliocene global-mean ΔT and $\Delta F_{\mathrm{NonGHG}}$
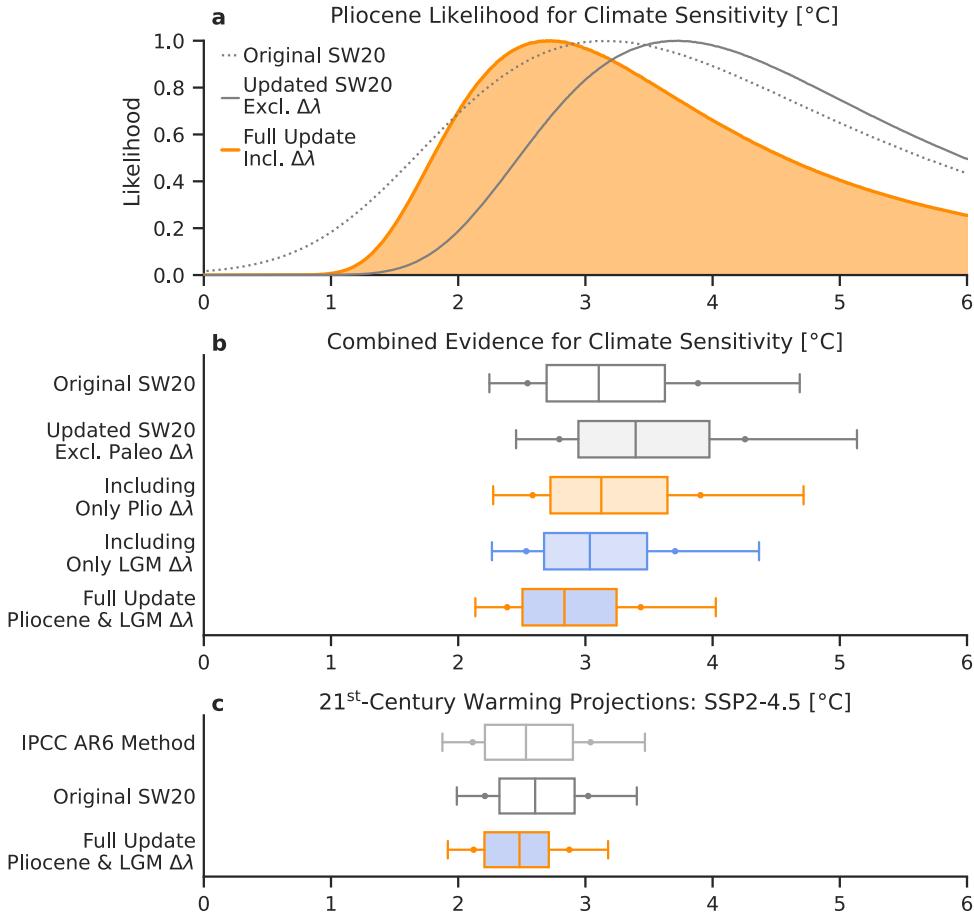
Figure 3.4: Modern climate sensitivity and 21st-century warming, accounting for paleoclimate pattern effects ($\Delta\lambda$). **(a)** Pliocene-only likelihoods (dotted) from SW20 (Sherwood et al., 2020); (gray) including updates to ΔT_{Plio} and ΔF_{Plio} but excluding pattern effects ($\Delta\lambda$); (orange) fully updated SW20 including $\Delta\lambda$. **(b)** Posterior probability density functions (PDFs) after combining lines of evidence: (gray, white fill) SW20, (gray) SW20 with updated paleoclimate ΔT and ΔF but excluding $\Delta\lambda$, (orange) including $\Delta\lambda$ only for the Pliocene, (blue) $\Delta\lambda$ only for the Last Glacial Maximum (LGM) (Cooper et al., 2024), and (orange, blue fill) Full Update including Pliocene and LGM $\Delta\lambda$. Panels **a**–**b** show effective climate sensitivity (S), as in SW20. **(c)** Projected global warming from the FaIR model (Smith et al., 2024), measured as mean anomaly over 2081–2100 relative to 1850–1900 mean, using climate sensitivity distributions from IPCC AR6 (Forster et al., 2021), SW20, and the Full Update in panel **b**. Line caps indicate 5th to 95th percentiles, dots indicate 66% *likely* range, box indicates 25th to 75th percentiles, and line indicates median.

but excluding Pliocene pattern effects. As seen in Fig. 3.4a, the updates from the *global-mean* information alone (excluding $\Delta\lambda$) suggest a much higher ECS (Tierney et al., 2025b). However, the *spatial information* in the Pliocene reconstructions—quantified as $\Delta\lambda$—has a larger and opposite impact. Including $\Delta\lambda$ shifts the maximum likelihood from 3.7°C to 2.7°C and substantially reduces the high tail of the distribution.

We now revise the best estimate for modern ECS by combining the Pliocene with the additional lines of evidence in SW20: the Last Glacial Maximum (LGM), the historical record (c. 1870–present), and process understanding (Methods) (Fig. 3.4b). We first show SW20's results, then we include paleoclimate updates only to global-mean quantities (i.e., excluding $\Delta\lambda$), which increases ECS substantially. We then include $\Delta\lambda$ from only the Pliocene or LGM (Cooper et al., 2024), and finally we combine our results for Pliocene and LGM $\Delta\lambda$ to provide a best estimate that fully accounts for paleoclimate pattern effects and their uncertainties. Once again, *global-mean* paleoclimate updates increase ECS, while the *spatial information* from pattern effects is more impactful and leads to much stronger overall constraints, particularly for the upper bound. The revised best estimate (median) for modern ECS becomes 2.8°C, with a 66% range of $2.4 - 3.4$°C (90% CI: $2.1 - 4.0$°C) (Fig. 3.4b; Table 3.S3). This range represents a major update to the upper bounds in SW20 (Sherwood et al., 2020) and the IPCC Sixth Assessment report (AR6) (Forster et al., 2021), while our lower bound confirms those assessments. For comparison with SW20's robustness tests, we find a 66% robust range of $2.6 - 3.8$°C (90% CI: $2.3 - 4.6$°C), which also represents a much stronger constraint compared to the 95th percentile of 5.7°C in SW20's robust range.

Importantly, our updates to modern ECS also reduce uncertainty in projections of 21st-century warming. Fig. 3.4c shows the 2081–2100 mean warming relative to 1850–1900 projected by the FaIR model (Smith et al., 2024), a climate emulator that produced projections for IPCC AR6, under the SSP2-4.5 emissions scenario (Forster et al., 2021). FaIR's large ensemble is calibrated to match the historical record through 2022 while sampling the full range of uncertainty in ECS (Smith et al., 2024). We first revise the FaIR ensemble's ECS distribution to match SW20, which produces a minor change (Methods). We then use our fully updated ECS distribution with the FaIR model (Fig. 3.4b), which yields a median of 2.5°C for end-of-century warming (relative to preindustrial) and substantially reduces uncertainty in the upper bound of warming projections,

with a 66% *likely* range of $2.1 - 2.9°C$ (90% CI: $1.9 - 3.2°C$).

Pliocene pattern effects arise from changes in ice sheets, vegetation, and topography that amplify SST warming in the extratropics, in turn leading to cloud feedbacks that further amplify global warming. Recent work on the Last Glacial Maximum also found that ice sheets amplify extratropical SST cooling, similarly leading to positive cloud feedbacks (Cooper et al., 2024). The modern climate feedback from CO_2 alone (in the absence of ice sheet, vegetation, and topography changes) is more stabilizing than the feedbacks associated with the Pliocene and LGM.

Updating *global mean* Pliocene and LGM temperatures based on the latest state-of-the-art reconstructions, while neglecting pattern effects, appears to suggest substantially higher estimates of climate sensitivity compared to SW20 (Sherwood et al., 2020) and IPCC AR6 (Forster et al., 2021). However, our results show that including *spatial information* from those same reconstructions leads to the opposite conclusion, such that paleoclimates now provide much stronger constraints on the modern climate's sensitivity to CO_2 and projected warming. We note that our 21st-century projections assume ice sheets will not be lost this century. An important corollary to our results is that a major shift in the modern warming pattern, e.g., caused by loss of the West Antarctic Ice Sheet (Dvorak et al., 2025; Weiffenbach et al., 2024; Lunt et al., 2010), could activate positive feedbacks on longer timescales in the modern climate similar to those that amplified global warming during the Pliocene.

3.6   Methods

3.6.1   AGCM simulations

Following Chapter 2 (Cooper et al., 2024), estimating paleoclimate $\Delta\lambda$ (Eq. 3.3) in AGCMs requires three simulations that differ only in their SST/SIC boundary conditions while all other forcings are constant at modern levels, similar to "amip-piForcing" simulations (Andrews, 2014; Gregory and Andrews, 2016).

The three categories of AGCM simulations are: **(a)** Preindustrial baseline, for which we use the Late Holocene $(0 - 4$ ka) (Osman et al., 2021); **(b)** $2xCO_2$, for which we use the multi-model mean of quasi-equilibrium $2xCO_2$ simulations in LongRunMIP (Rugenstein et al., 2019); **(c)** Pliocene, for which we use the various reconstructions described in the main text (Fig. 3.1; Fig. 3.S1–S3).

In CAM4 and CAM5, we also test the 5th and 95th percentiles of the plioDA ensemble (Fig. 3.S4); ensemble members are ranked by estimating λ_{Plio} with CAM4 Green's functions (Dong et al., 2019). SST/SIC boundary conditions are prepared as described in Chapter 2 (Cooper et al., 2024). We use plioDA's SIC for Annan et al. (2024), as no SIC is provided by the latter; this approach is supported by similar ΔT_{Plio} in both reconstructions.

For each AGCM, we compute anomalies in simulations **(b)** and **(c)** relative to **(a)**. Simulations are 30 years, and we analyze means over the final 25 years for CAM4 (2° resolution), CAM5.3 (2°), CAM6.0 (2°) (Danabasoglu et al., 2020), and HadGEM3-GC3.1-LL (N96, 135 km) (Williams et al., 2017), or the final 30 of 31 years for GFDL-AM4.0 (C96, 100 km) (Held et al., 2019). Results are included in Tables 3.S1–S2. As described in Chapter 2 (Cooper et al., 2024), we test sensitivity of $\Delta\lambda$ to the 2xCO$_2$ pattern by computing an alternate $\Delta\lambda^{\text{Alt}}_{\text{150yr}}$, which uses the 150-year regression of abrupt CO$_2$-forcing simulations in the parent coupled models corresponding to each AGCM instead of our λ_{2xCO_2}. Each coupled model produces a distinct warming pattern over the 150-year period, thus $\Delta\lambda^{\text{Alt}}_{\text{150yr}}$ samples uncertainty in CO$_2$-warming patterns. This test confirms our finding of $\Delta\lambda < 0$ (Table 3.S1–S2) and produces ECS constraints that agree with our main result within 0.1°C (Table 3.S3). We decompose λ into component feedbacks (Planck, lapse rate, water vapor, surface albedo, shortwave cloud, and longwave cloud) using CAM5 radiative kernels (Pendergrass et al., 2018), following Soden et al. (2008) (Fig. 3.S5–S8).

3.6.2 Constraining modern climate sensitivity

Modern climate sensitivity is the steady-state response of global-mean T to doubling preindustrial CO$_2$ concentrations, including only the feedbacks acting on an approximate 150-year timescale, i.e., assuming fixed ice sheets and vegetation. This metric, called "effective climate sensitivity" to distinguish it from true equilibrium, is termed S in SW20 (Sherwood et al., 2020) and hereafter. To infer S from Pliocene evidence, we build on SW20's equation of Pliocene energy balance by including the updates described below (distribution percentiles provided in Table 3.S3).

$$\Delta T_{\text{Plio}} = \frac{-\Delta F_{\text{CO}_2}\left(1 + f_{\text{CH}_4}\right) + \Delta F_{\text{NonGHG}}}{\frac{\lambda_{\text{2xCO}_2}}{1+\zeta} + \Delta\lambda} \tag{3.5}$$

(i) Our main update is incorporating Pliocene $\Delta\lambda$ as $\Delta\lambda \sim \mathcal{N}(\mu = -0.37, \sigma = 0.32)$ W m^{-2} K^{-1}. We estimate μ and σ for $\Delta\lambda$ by combining the spread across AGCMs and reconstructions using the

bootstrap approach in Chapter 2 (Cooper et al., 2024), with plioDA's best estimate as the reference value for differences in CAM4 and CAM5.

(ii) Pliocene forcing is updated based on the recent estimate of effective radiative forcing from non-GHG sources ($\Delta F_{\mathrm{NonGHG}}$), including ice sheets, vegetation, and land-sea distribution (Dvorak et al., 2025). We assign $\Delta F_{\mathrm{NonGHG}} \sim \mathcal{N}(1.7, 1.0)$ W m^{-2} K^{-1}, which assumes a 1σ uncertainty that approximately maintains the original SW20 uncertainty in total ΔF_{Plio}. For reference, total ΔF_{Plio} (numerator of Eq. 3.5) is 3.9 ± 1.2 (1σ) W m^{-2}, with $\Delta F_{\mathrm{CO_2}} \approx 2.2$ W m^{-2}. We note there is substantial uncertainty in the components of ΔF_{Plio}, which merit further study (Haywood et al., 2024; Hopcroft et al., 2020; Grant et al., 2019; Sagoo and Storelvmo, 2017; Dowsett et al., 2016; Dutton et al., 2015; Unger and Yue, 2014).

(iii) ΔT_{Plio} is updated from 3.0 ± 1.0°C (1σ) in SW20 to plioDA's constraint of $\Delta T_{\mathrm{Plio}} \sim \mathcal{N}(4.1, 0.6)$ °C (Tierney et al., 2025b), which is supported by the estimate in Annan et al. (2024) of 3.9 ± 1.1°C (1σ).

From SW20 (Sherwood et al., 2020), the remaining parameters in Equation 3.5 are: CO_2 forcing of $\Delta F_{\mathrm{CO_2}} = \Delta F_{\mathrm{2xCO_2}} \times \ln(\frac{[\mathrm{CO_2}]}{284\mathrm{ppm}})/\ln(2)$, where $[\mathrm{CO_2}] \sim \mathcal{N}(375, 25)$ ppm and $\Delta F_{\mathrm{2xCO_2}} \sim \mathcal{N}(4.0, 0.3)$ W m^{-2}; a scaling factor for methane and N$_2$O forcing, $1 + f_{CH_4}$, with $f_{CH_4} \sim \mathcal{N}(0.4, 0.1)$; and a timescale transfer factor between quasi-equilibrium and the 150-year S timescale, $1 + \zeta$, to account for feedbacks becoming more positive at longer timescales, with $\zeta \sim \mathcal{N}(0.06, 0.2)$ based on LongRunMIP (Rugenstein et al., 2019). Finally, modern climate sensitivity is $S = -\Delta F_{\mathrm{2xCO_2}}/\lambda_{\mathrm{2xCO_2}}$ (Sherwood et al., 2020).

We also use an alternate version of the $\Delta\lambda$ in **(i)** estimated by comparing our paleoclimate AGCM simulations with feedbacks from 150-year regression of abrupt CO_2-forcing simulations in the parent coupled models of each AGCM. Each coupled model produces a distinct warming pattern, thereby sampling uncertainty in the pattern of warming from CO_2. With $\lambda_{150\mathrm{yr}}^{\mathrm{CO_2}}$ representing the regression feedback, we estimate Pliocene $\Delta\lambda_{150\mathrm{yr}}^{\mathrm{Alt}} = \lambda_{150\mathrm{yr}}^{\mathrm{CO_2}} - \lambda_{\mathrm{Plio}}$, and we use the same bootstrap approach in **(i)** to find Pliocene $\Delta\lambda_{150\mathrm{yr}}^{\mathrm{Alt}} \sim \mathcal{N}(\mu = -0.44, \sigma = 0.40)$ W m^{-2} K^{-1}. Because $\Delta\lambda_{150\mathrm{yr}}^{\mathrm{Alt}}$ represents a comparison with the 150-year regression feedback rather than quasi-equilibrium simulations, the denominator of Equation 3.5 becomes $(\lambda_{\mathrm{2xCO_2}} + \Delta\lambda_{150\mathrm{yr}}^{\mathrm{Alt}})/(1 + \zeta)$ when using $\Delta\lambda_{150\mathrm{yr}}^{\mathrm{Alt}}$ instead of our standard $\Delta\lambda$. Note that the percentiles of the final S distribution agree within 0.1°C when using $\Delta\lambda_{150\mathrm{yr}}^{\mathrm{Alt}}$ (Table 3.S3).

There are advantages to our formulation of the Pliocene energy balance (Eq. 3.5) compared to SW20's Equation 23. First, the Pliocene is now consistent with the LGM, as ΔF_{NonGHG} is now added directly rather than estimated by multiplying ΔF_{CO_2} by a scale factor, $1 + f_{ESS}$, representing Earth system sensitivity (Lunt et al., 2010; PALAEOSENS Project Members, 2012). Second, f_{ESS} conflates forcings and feedbacks, and estimating f_{ESS} requires free-running coupled simulations that have inaccurate warming patterns (Dvorak et al., 2025). Instead of using f_{ESS}, our Equation 3.5 separately includes *effective radiative forcing*, ΔF_{NonGHG}, from AGCM simulations with paleoenvironmental boundary conditions informed by proxies for ice extent, vegetation, and topography (Dvorak et al., 2025; Zhu and Poulsen, 2021), and *paleoclimate pattern effects*, from AGCM simulations with SST/SIC patterns constrained by data assimilation (Cooper et al., 2024).

Climate sensitivity PDFs are summarized in Table 3.S3. We calculate likelihoods and PDFs for S using SW20's Bayesian framework (Sherwood et al., 2020). This framework quantitatively combines our findings with additional lines of evidence, and the methods can be continually developed in ongoing efforts (Marvel and Webb, 2025; Sherwood and Forest, 2024). Our findings would have the same directional impact on other assessments of ECS and modern warming (Forster et al., 2021; Kaufman and Masson-Delmotte, 2024).

In Fig. 3.4 and Table 3.S3, we show S with and without updates **(i), (ii),** and **(iii)**. For the LGM evidence in Fig. 3.4b, we include updated $\Delta T_{\text{LGM}} \sim \mathcal{N}(-6,1)$ °C and LGM $\Delta\lambda \sim \mathcal{N}(-0.37, 0.23)$ W m^{-2} K^{-1} (Cooper et al., 2024). We also use $\lambda_{150\text{yr}}^{\text{CO}_2}$ in Table 3.S1 to estimate LGM $\Delta\lambda_{150\text{yr}}^{\text{Alt}} \sim \mathcal{N}(\mu = -0.42, \sigma = 0.34)$ W m^{-2} K^{-1}. While SW20's framework generally assumes lines of evidence are independent, our estimates of Pliocene and LGM pattern effects are interrelated. We use the same AGCMs, and the reconstruction methods are partially overlapping. To account for the relationship between Pliocene and LGM $\Delta\lambda$ estimates, we identify pairs of estimates that use similar reconstruction methods and the same AGCM (Table 3.S4). From these pairs, we estimate the Pearson correlation (r) and covariance for $\Delta\lambda$ to be $r = 0.56$ and cov $= 0.0123$ [W m^{-2} K^{-1}]^2. For $\Delta\lambda_{150\text{yr}}^{\text{Alt}}$, we estimate $r = 0.87$ and cov $= 0.0562$ [W m^{-2} K^{-1}]^2. We account for the shared error covariance by drawing correlated values for LGM and Pliocene $\Delta\lambda$ from bivariate normal distributions. However, the S constraints are insensitive to the covariance, as our Full Update percentiles (Table 3.S3) change by less than 0.1°C if we assume zero covariance. This result aligns with the dependence tests in SW20, which also found relatively small impacts from codependencies

(Sherwood et al., 2020).

We include results corresponding to SW20's robustness test, which assumes a uniform prior on S from 0 to 20°C instead of the baseline prior of uniform λ from -10 to 10 W m^{-2} K^{-1}, in Table 3.S3. The robustness test yields a median of 3.1°C and 66% range of $2.6-3.8$°C (90% CI: $2.3-4.6$°C). As for our main result using the baseline prior, this represents a substantial narrowing of uncertainty compared to the robust ranges in SW20. For illustrative purposes, we also include posterior PDFs considering only the Pliocene evidence and assuming the uniform-S prior. The PDF from the Pliocene alone has a median of 3.8°C and 66% range of $2.4-7.2$°C (90% CI: $1.9-12.9$°C).

3.6.3 Projections of 21st-century warming

We analyze warming projections through 2100 under SSP2–4.5 (Forster et al., 2021) from the FaIR model v1.4.1, calibrated to match historical records as in IPCC AR6 but with updated constraints through 2022 (Smith et al., 2024). From FaIR, we have a large ensemble of global-mean temperatures from 1850–2100, and each member has an associated ECS. For each ensemble member, we compute the mean warming over 2081–2100 relative to the 1850–1900 mean. We then resample the ensemble with replacement to match the specified ECS distributions from SW20 and from our updated paleoclimate-constrained ECS. This resampling produces revised distributions of projected warming that are associated with the specified ECS distributions (Fig. 3.4).

3.7 Appendix: Supplemental information

SI Appendix Figures 3.S1–S11 and Tables 3.S1–S4.
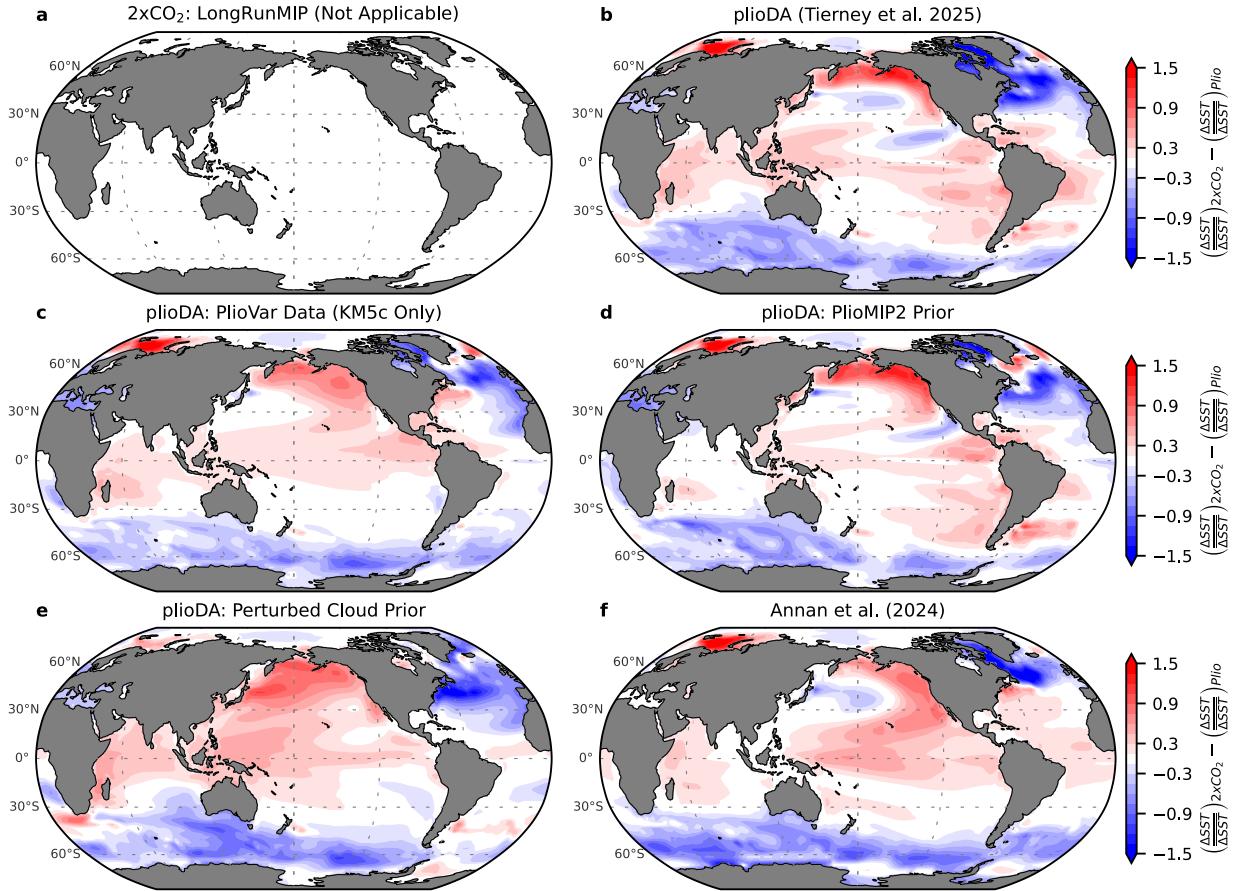
Figure 3.S1: Differences between the $2\text{x}CO_2$ pattern of sea-surface temperature (SST) anomalies and Pliocene patterns of SST anomalies. Panels correspond to Figure 3.1 of Main Text. Before taking the differences, each pattern's local anomalies are divided by its global-mean SST anomaly to emphasize the spatial patterns. Red regions indicate stronger relative amplification of warming in the LongRunMIP $2\text{x}CO_2$ pattern (Rugenstein et al., 2019), while blue regions indicate stronger relative amplification of Pliocene warming. See Figure 3.S10 for zonal-mean SST anomalies and pattern differences.
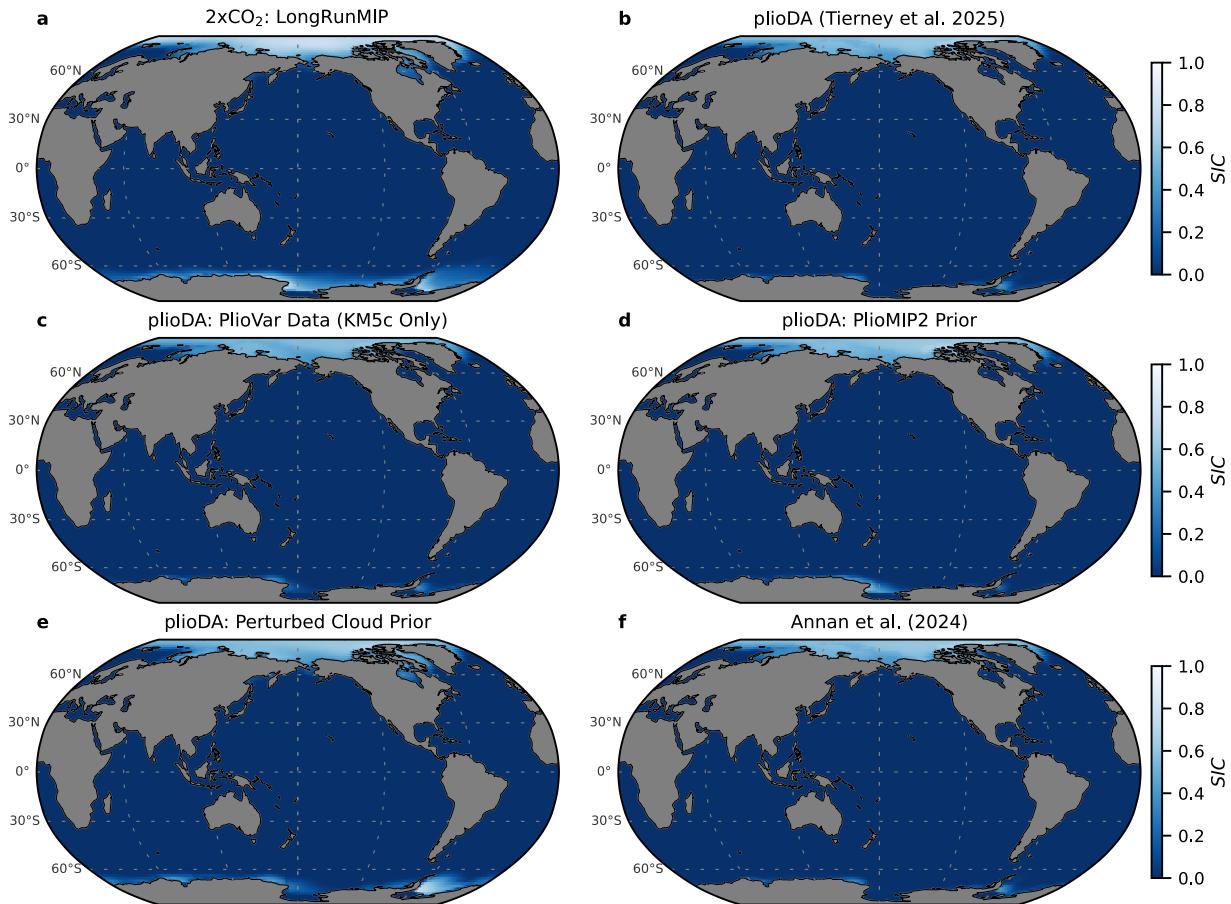
Figure 3.S2: Sea-ice concentration (SIC): LongRunMIP 2xCO$_2$ and Pliocene reconstructions. Panels show annual means. Note that plioDA sea ice is used for the Annan et al. (2024) reconstruction.
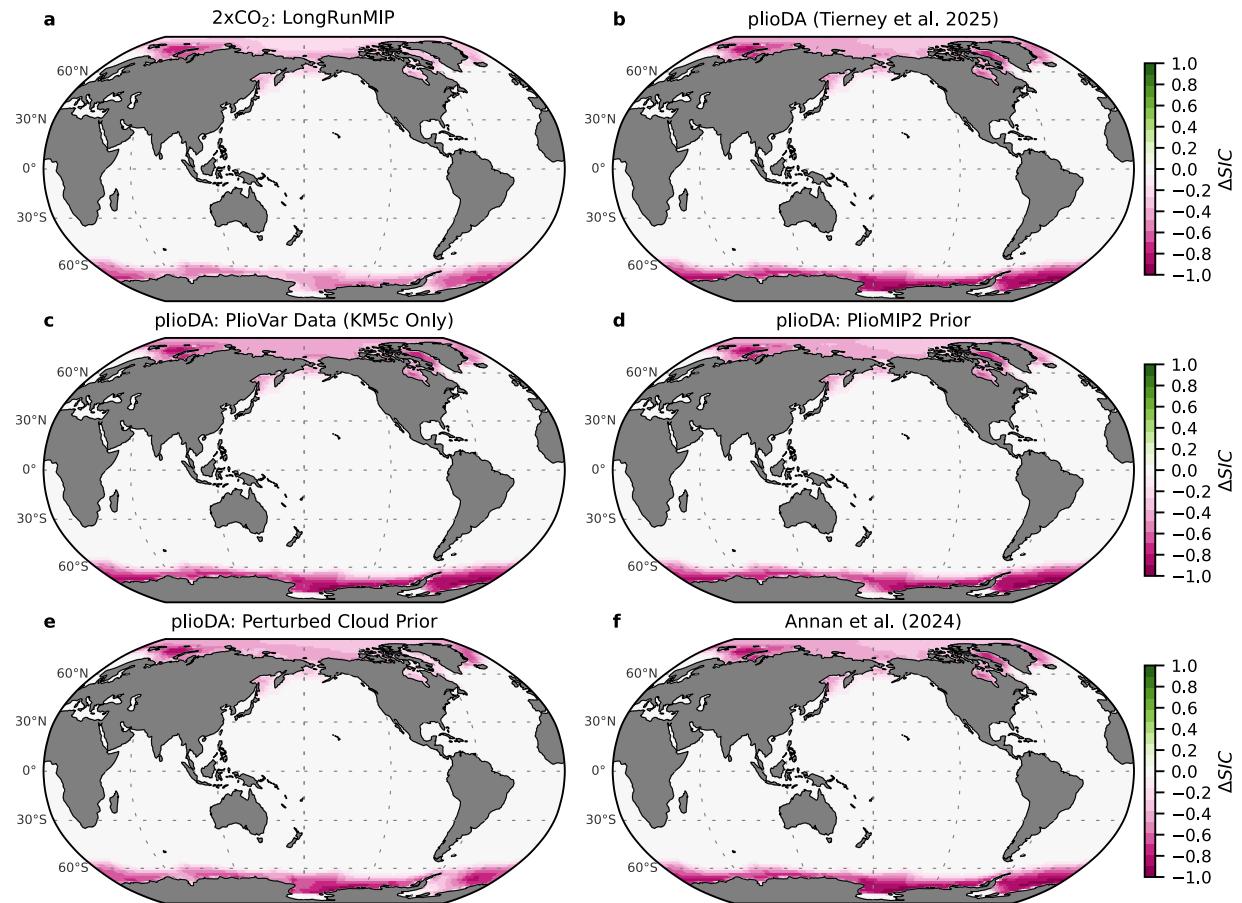
Figure 3.S3: Sea-ice concentration (SIC) anomalies: LongRunMIP 2xCO$_2$ and Pliocene reconstructions relative to preindustrial baseline. Panels show annual-mean differences relative to the preindustrial (Late Holocene) baseline (Osman et al., 2021). Note that plioDA sea ice is used for the Annan et al. (2024) reconstruction.
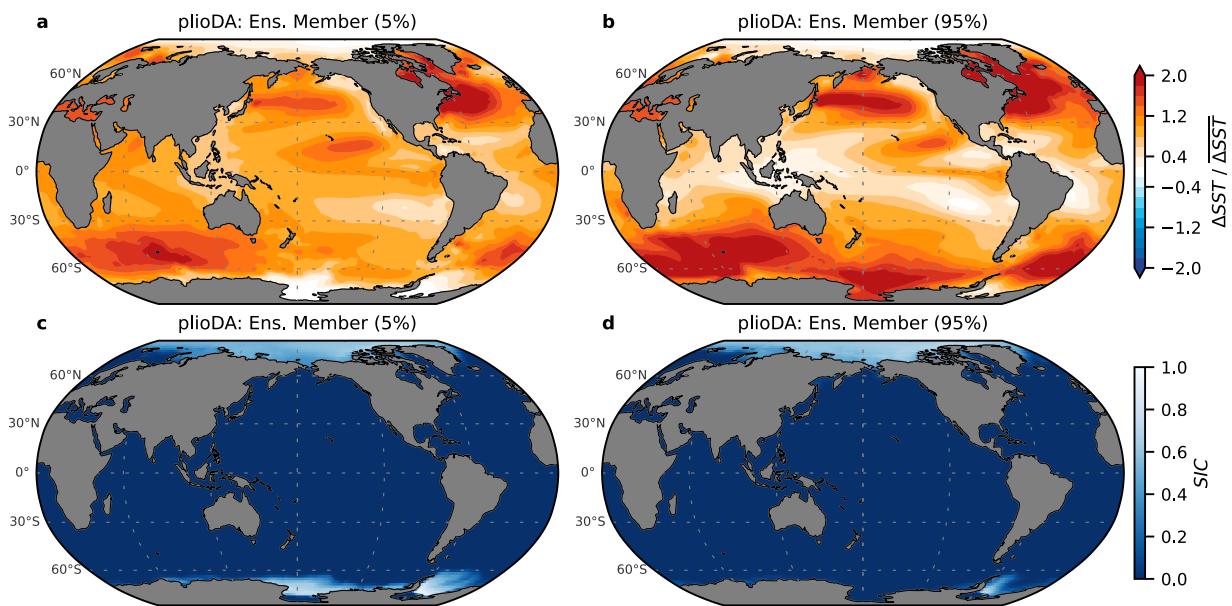
Figure 3.S4: 5th and 95th percentile ensemble members from plioDA reconstruction (Tierney et al., 2025b). **(a–b)** Sea-surface temperature (SST) anomalies and **(c–d)** sea-ice concentration (SIC) for ensemble members with the 5th percentile net feedback (more negative, stable climate) and 95th percentile net feedback (more positive, less stable climate). Ensemble members are ranked using CAM4 Green's functions (Dong et al., 2019).
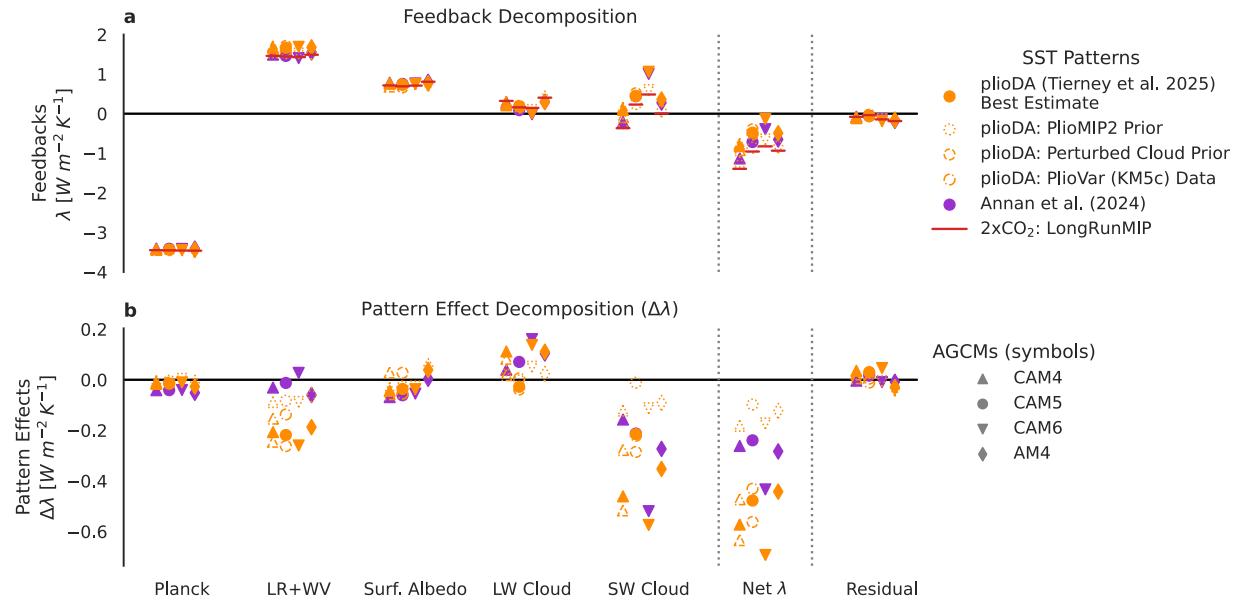
Figure 3.S5: Kernel decomposition of radiative feedbacks (λ). Note that each legend applies to both panels: different sea-surface temperature and sea ice patterns are distinguished by colors/borders, while the different atmospheric general circulation models (AGCMs) are distinguished by symbol shapes. **(a)** Decomposition of feedbacks using radiative kernels (Soden et al., 2008) from CAM5 (Pendergrass et al., 2018). LR+WV represents the lapse rate and water vapor feedbacks. **(b)** Pattern effects ($\Delta\lambda = \lambda_{2\mathrm{xCO}_2} - \lambda_{\mathrm{Plio}}$) for each component feedback in panel **a**.
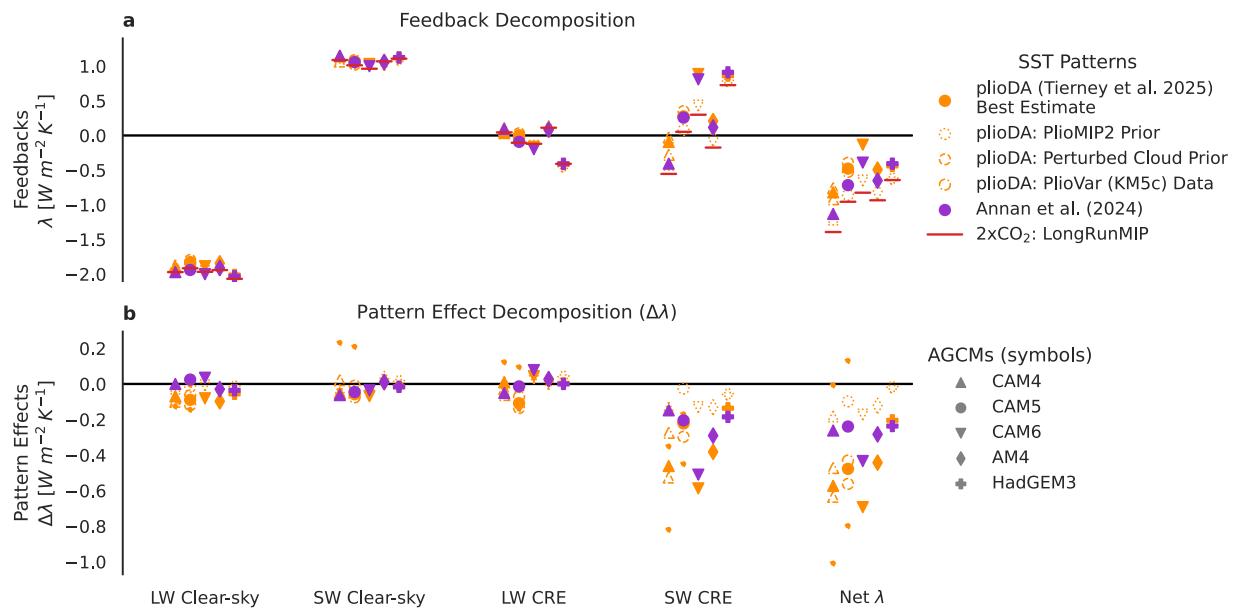
Figure 3.S6: Decomposition of radiative feedbacks (λ) from direct model outputs for clear-sky radiation and cloud radiative effects (CRE). Results are separated into longwave (LW) and short-wave (SW) components. **(a)** Decomposition of feedbacks, and **(b)** decomposition of pattern effects ($\Delta\lambda = \lambda_{2\text{xCO}_2} - \lambda_{\text{Plio}}$).
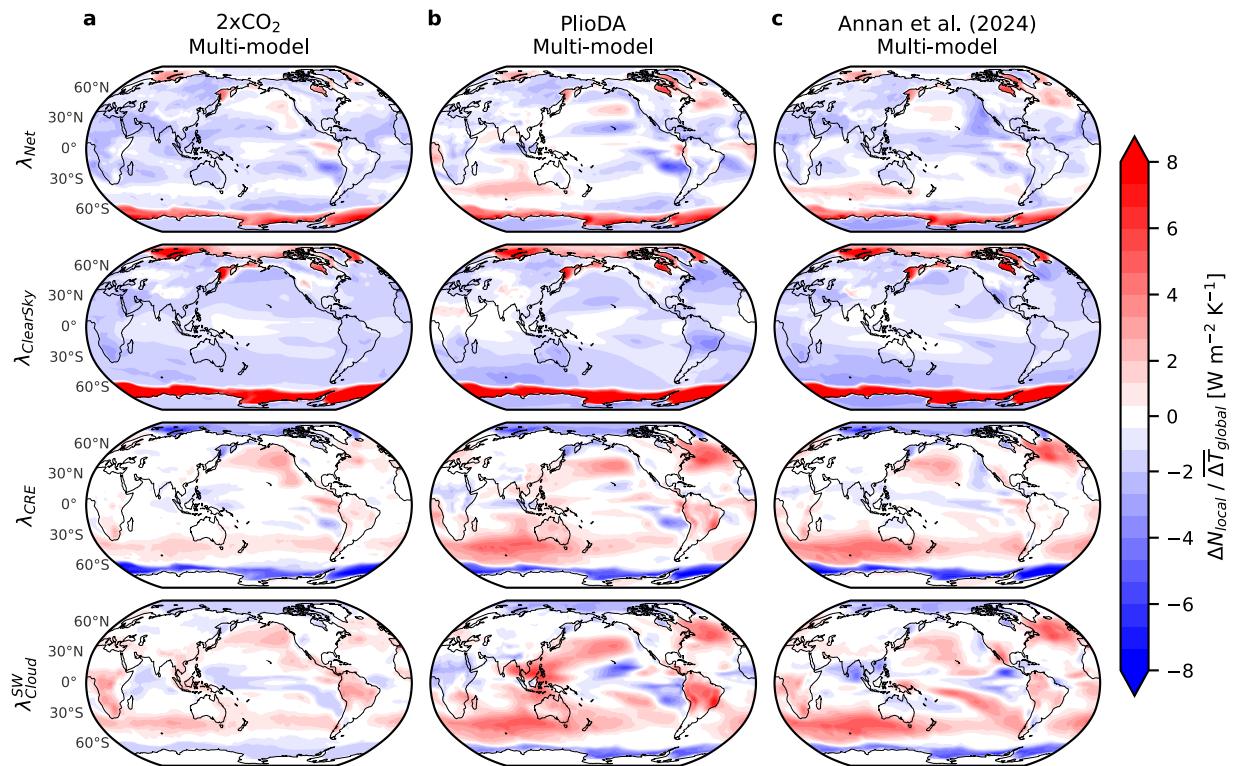
Figure 3.S7: Spatial pattern of local radiative feedbacks (λ). Local feedbacks are calculated as $\Delta N / \overline{\Delta T}$, where ΔN is the local anomaly in top-of-atmosphere radiation, and $\overline{\Delta T}$ is the global-mean anomaly in near-surface air temperature. Multi-model mean, including CAM4, CAM5, CAM6, and GFDL-AM4 from **(a)** LongRunMIP $2xCO_2$ (Rugenstein et al., 2019), **(b)** plioDA (Tierney et al., 2025b), and **(c)** (Annan et al., 2024).

Figure 3.S8–3.S9 (shown on following pages): Zonal mean of local radiative feedbacks (λ) and pattern effects, $\Delta\lambda = \lambda_{2\mathrm{xCO_2}} - \lambda_{\mathrm{Plio}}$. Local feedbacks are calculated as $\Delta N/\overline{\Delta T}$, where ΔN is the local anomaly in top-of-atmosphere radiation, and $\overline{\Delta T}$ is the global-mean anomaly in near-surface air temperature. **(a)** Feedbacks, λ, in CAM5 using various patterns of sea-surface temperature (SST) and sea ice, and **(b)** Pattern effects, $\Delta\lambda = \lambda_{2\mathrm{xCO_2}} - \lambda_{\mathrm{Plio}}$, in CAM5 corresponding to panel **a**. **(c–d)** Repeat of panels **a–b** with results from multiple models (CAM4, CAM5, CAM6, and GFDL-AM4) and a subset of SST and sea ice patterns.
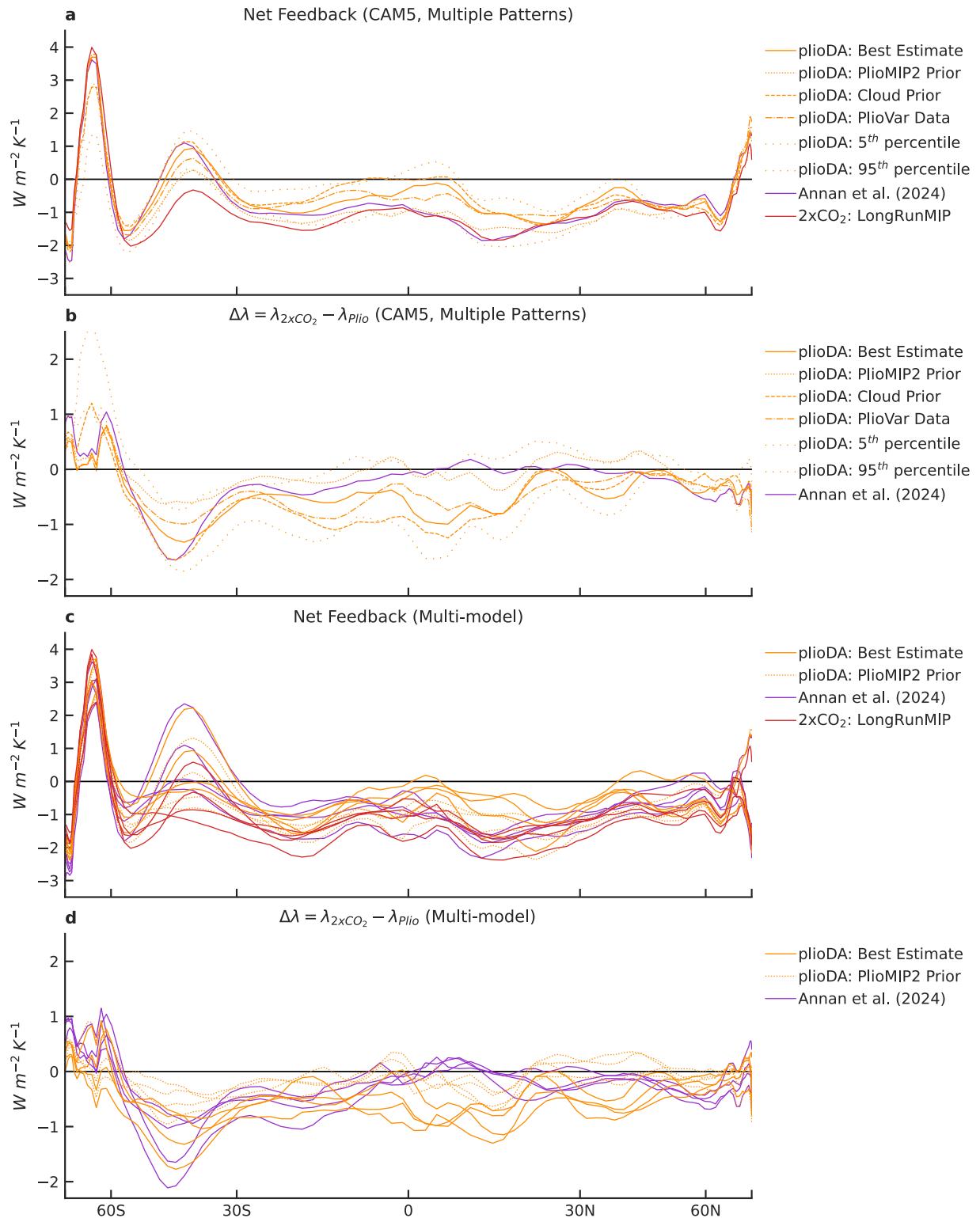
Figure 3.S8: See caption on preceding page.
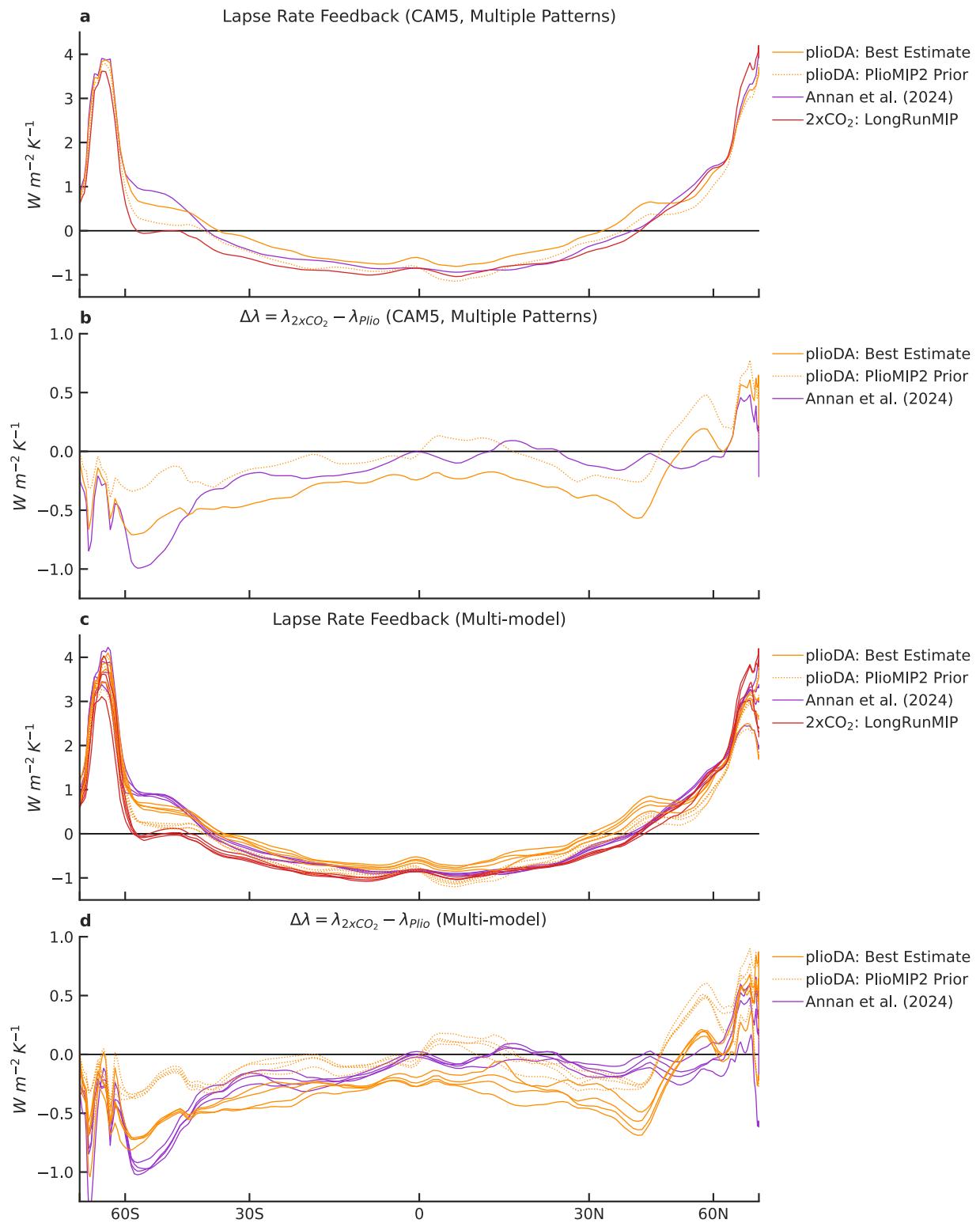
Figure 3.S9: See caption that precedes Figure 3.S8.
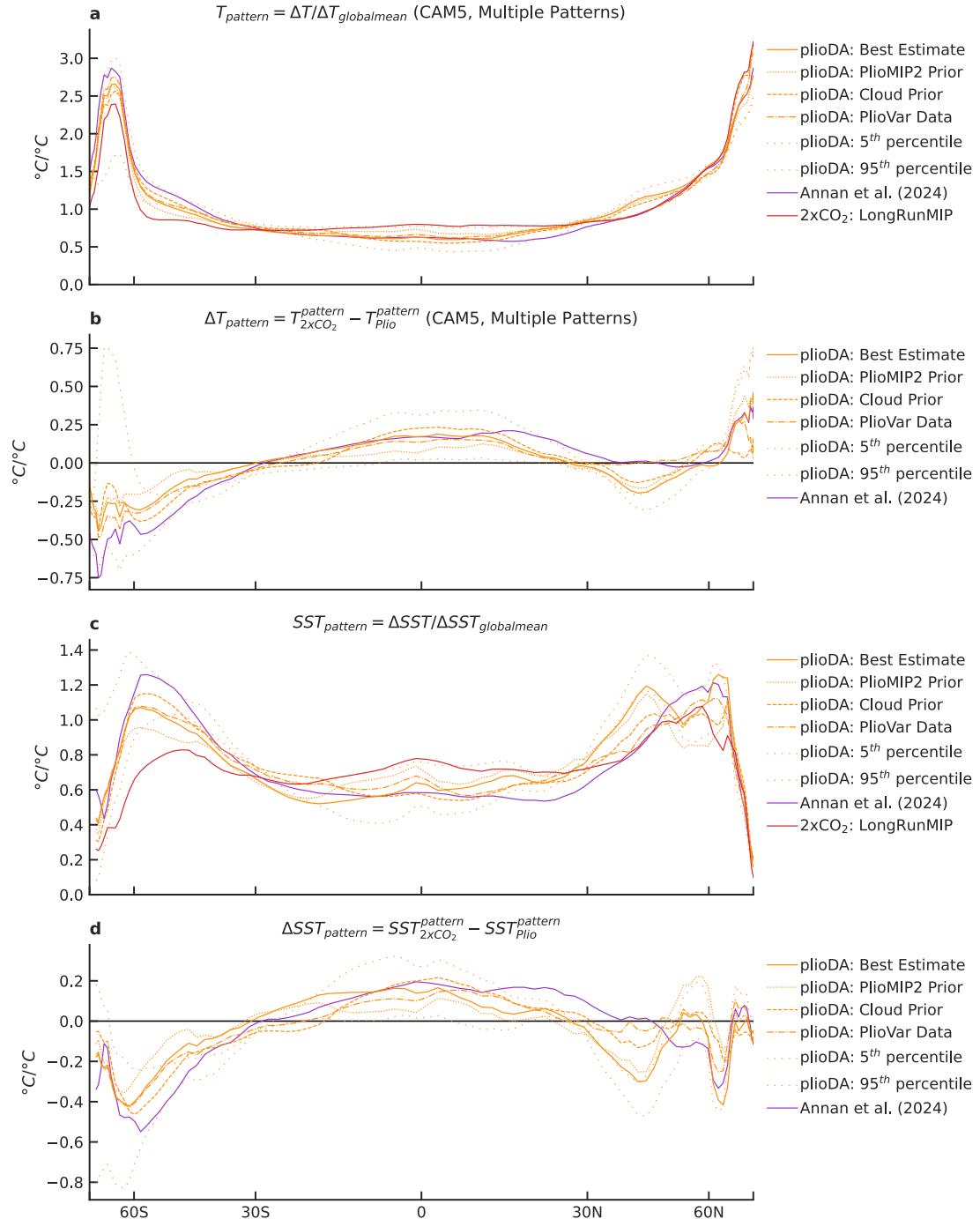
Figure 3.S10: Zonal-mean patterns of temperature anomalies. **(A)** Normalized T from various patterns and **(b)** differences versus LongRunMIP $2xCO_2$ pattern. **(c–d)** Repeats panels **a** and **b** for SST. Note that **a–b** show AGCM output from CAM5, whereas **c–d** show input boundary conditions for all AGCMs.
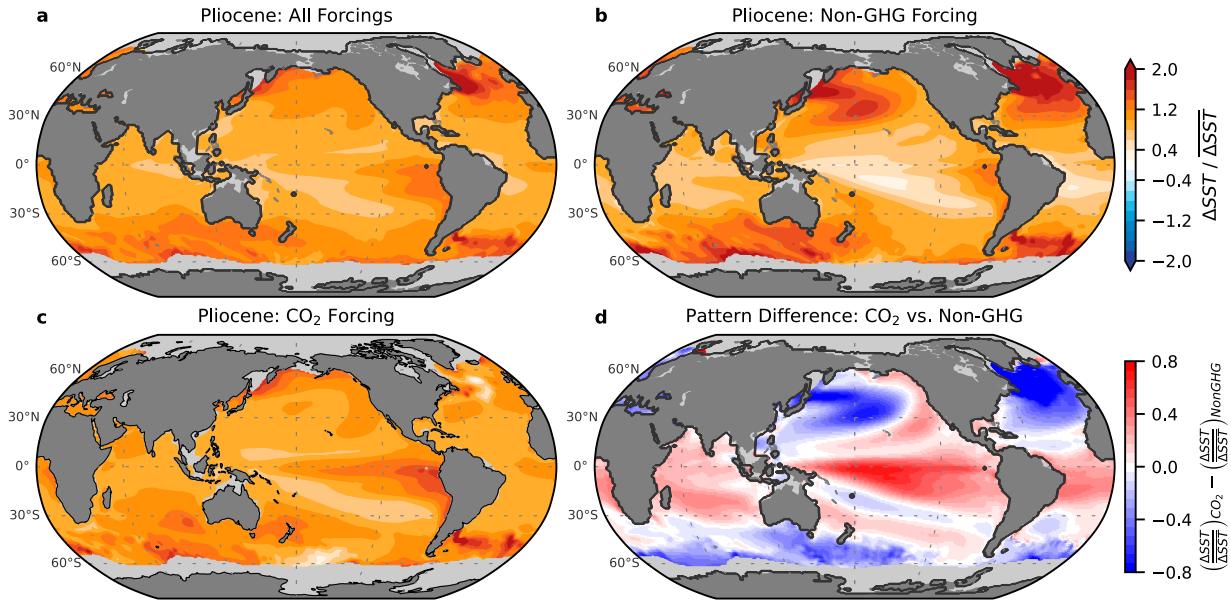
Figure 3.S11: Sea-surface temperature (SST) response to Pliocene forcings in CESM2.1. Results shown are from (Dvorak et al., 2025). **(a–c)** Patterns of SST anomalies (normalized by global-mean anomalies) relative to preindustrial control from **(a)** all Pliocene forcings, **(b)** Non-GHG forcings including ice sheets, vegetation, topography, and bathymetry, and **(c)** CO_2 concentration of 400 ppm, which accounts for both CO_2 and methane forcing. **(d)** Difference between SST response to CO_2 versus non-GHG forcing, represented as panel **c** minus panel **b**; red regions indicate stronger relative amplification of warming from CO_2, while blue regions indicate stronger relative amplification from non-GHG forcings. In all panels, regions of preindustrial sea ice are masked in light gray. The CESM2 simulations follow the PlioMIP2 protocol (Haywood et al., 2020; Feng et al., 2022).

Table 3.S1: All units are W m^{-2} K^{-1}. Pliocene pattern effects, $\Delta\lambda = \lambda_{2\text{xCO}_2} - \lambda_{\text{Plio}}$, from three patterns of reconstructed Pliocene SST and sea ice in various AGCMs (CAM4 coupled to CLM4.5, CAM5.3 coupled to CLM5.0, CAM6.0 coupled to CLM5.0, GFDL-AM4.0, and HadGEM3-GC3.1-LL). Alternate values for Pliocene pattern effects, $\Delta\lambda_{150\text{yr}}^{\text{Alt}} = \lambda_{150\text{yr}}^{\text{CO}_2} - \lambda_{\text{Plio}}$, are shown using 150-yr regression of abrupt-4xCO$_2$ simulations (abrupt-2xCO$_2$ is used for CESM2-CAM6 to avoid issues with the ice nucleation scheme and cloud microphysics timestep that impact the feedback diagnosed from the 4xCO$_2$ simulation (Zhu et al., 2022; Burls and Sagoo, 2022) from coupled models corresponding to each AGCM (Andrews et al., 2022).

Model	Pattern	$\Delta\lambda$	λ_{Plio}	$\lambda_{2\text{xCO}_2}$	$\Delta\lambda_{150\text{yr}}^{\text{Alt}}$	$\lambda_{150\text{yr}}^{\text{CO}_2}$
CAM4	plioDA	-0.57	-0.82	-1.39	-0.41	-1.23
CAM4	plioDA: PlioMIP2 Prior	-0.18	-1.21	-1.39	-0.02	-1.23
CAM4	Annan24	-0.26	-1.13	-1.39	-0.10	-1.23
CAM5	plioDA	-0.48	-0.48	-0.96	-0.67	-1.15
CAM5	plioDA: PlioMIP2 Prior	-0.10	-0.86	-0.96	-0.29	-1.15
CAM5	Annan24	-0.24	-0.72	-0.96	-0.43	-1.15
CAM6	plioDA	-0.69	-0.13	-0.83	-1.08	-1.21
CAM6	plioDA: PlioMIP2 Prior	-0.17	-0.65	-0.83	-0.56	-1.21
CAM6	Annan24	-0.43	-0.39	-0.83	-0.82	-1.21
GFDL-AM4	plioDA	-0.44	-0.49	-0.93	-0.37	-0.86
GFDL-AM4	plioDA: PlioMIP2 Prior	-0.12	-0.81	-0.93	-0.05	-0.86
GFDL-AM4	Annan24	-0.28	-0.65	-0.93	-0.21	-0.86
HadGEM3	plioDA	-0.20	-0.44	-0.64	-0.19	-0.63
HadGEM3	plioDA: PlioMIP2 Prior	-0.02	-0.62	-0.64	-0.01	-0.63
HadGEM3	Annan24	-0.24	-0.41	-0.64	-0.22	-0.63
CAM4	mean	-0.34	-1.05	-1.39	-0.18	-1.23
CAM5	mean	-0.27	-0.68	-0.96	-0.47	-1.15
CAM6	mean	-0.43	-0.39	-0.83	-0.82	-1.21
GFDL-AM4	mean	-0.28	-0.65	-0.93	-0.21	-0.86
HadGEM3	mean	-0.15	-0.49	-0.64	-0.14	-0.63
mean	Annan24	-0.29	-0.66	-0.95	-0.36	-1.02
mean	plioDA	-0.48	-0.47	-0.95	-0.54	-1.02
mean	plioDA: PlioMIP2 Prior	-0.12	-0.83	-0.95	-0.18	-1.02
mean	**mean**	-0.30	-0.65	-0.95	-0.36	-1.02
1σ	1σ	0.19	0.29		0.31	

Table 3.S2: Pliocene pattern effects, $\Delta\lambda = \lambda_{2\text{xCO}_2} - \lambda_{\text{Plio}}$, from various patterns of reconstructed Pliocene SST and sea ice in CAM4 and CAM5. Global-mean anomalies for SST, near-surface air temperature (T), and top-of-atmosphere radiative imbalance (N) are shown for reference. Alternate values for Pliocene pattern effects, $\Delta\lambda_{150\text{yr}}^{\text{Alt}} = \lambda_{150\text{yr}}^{\text{CO}_2} - \lambda_{\text{Plio}}$, are shown using 150-yr regression feedbacks (Table 3.S1).

Units		$\text{Wm}^{-2}\text{K}^{-1}$	$\text{Wm}^{-2}\text{K}^{-1}$	K	K	Wm^{-2}	$\text{Wm}^{-2}\text{K}^{-1}$
Model	Pattern	$\Delta\lambda$	λ	ΔSST	ΔT	ΔN	$\Delta\lambda_{150\text{yr}}^{\text{Alt}}$
CAM4	plioDA	-0.57	-0.82	3.00	3.90	-3.20	-0.41
CAM4	plioDA: PlioVar Data	-0.47	-0.92	2.89	3.78	-3.48	-0.31
CAM4	plioDA: PlioMIP2 Prior	-0.18	-1.21	2.94	3.86	-4.67	-0.02
CAM4	plioDA: Cloud Prior	-0.63	-0.76	2.83	3.68	-2.79	-0.47
CAM4	plioDA: 5%	-0.01	-1.39	3.96	4.88	-6.77	0.16
CAM4	plioDA: 95%	-1.01	-0.38	3.29	4.02	-1.55	-0.85
CAM4	Annan24	-0.26	-1.13	2.82	3.72	-4.21	-0.10
CAM4	**mean**	-0.45	-0.94	3.10	3.98	-3.81	-0.29
CAM4	**1σ**	0.33	0.33	0.41	0.41	1.65	0.33
CAM4	*2xCO2: LongRunMIP*		-1.39	2.35	3.16	-4.40	

Model	Pattern	$\Delta\lambda$	λ	ΔSST	ΔT	ΔN	$\Delta\lambda_{150\text{yr}}^{\text{Alt}}$
CAM5	plioDA	-0.48	-0.48	3.00	3.98	-1.90	-0.67
CAM5	plioDA: PlioVar Data	-0.43	-0.53	2.89	3.85	-2.02	-0.62
CAM5	plioDA: PlioMIP2 Prior	-0.10	-0.86	2.94	3.96	-3.40	-0.29
CAM5	plioDA: Cloud Prior	-0.56	-0.39	2.83	3.75	-1.48	-0.76
CAM5	plioDA: 5%	0.13	-1.09	3.96	4.99	-5.42	-0.06
CAM5	plioDA: 95%	-0.80	-0.16	3.29	4.10	-0.65	-0.99
CAM5	Annan24	-0.24	-0.72	2.82	3.78	-2.71	-0.43
CAM5	**mean**	-0.35	-0.60	3.10	4.06	-2.51	-0.55
CAM5	**1σ**	0.31	0.31	0.41	0.43	1.55	0.31
CAM5	*2xCO2: LongRunMIP*		-0.96	2.35	3.21	-3.07	

Table 3.S3: Posterior distributions of climate sensitivity (S). "Combined Evidence" assumes the Baseline Prior, $\lambda \sim \text{Unif}(-10, 10)$ W m^{-2} K^{-1}, and includes Process Understanding, Historical Evidence, and Paleoclimate Evidence from the Last Glacial Maximum (LGM) and Pliocene. The Robust Range also combines lines of evidence but assumes a Uniform S Prior, $S \sim \text{Unif}(0, 20)$ K (Sherwood et al., 2020). "Pliocene Only" considers only Pliocene evidence and assumes the Uniform S Prior. All uncertainties shown are 1σ values. Table structure is comparable to SW20's Table 10 (Sherwood et al., 2020).

Combined Evidence (Baseline Prior)	5th	17th	50th	83rd	95th	Mean	ΔT_{Plio}	ΔF^{Plio}_{NonGHG}	ΔT_{LGM}
SW20: Original	2.3	2.6	3.1	3.9	4.7	3.2	3.0 ± 1.0	f_{ESS}	-5 ± 1
+ Update ΔT_{LGM}	2.3	2.7	3.2	4.1	5.0	3.4	3.0 ± 1.0	f_{ESS}	-6 ± 1
+ Update ΔT_{Plio}	2.6	2.9	3.6	4.6	5.6	3.8	4.1 ± 0.6	f_{ESS}	-6 ± 1
+ Update ΔF^{Plio}_{NonGHG}	2.5	2.8	3.4	4.3	5.2	3.6	4.1 ± 0.6	1.7 ± 1.0	-6 ± 1
Include only LGM $\Delta\lambda$	2.3	2.6	3.0	3.7	4.4	3.2	4.1 ± 0.6	1.7 ± 1.0	-6 ± 1
Include only Pliocene $\Delta\lambda$	2.3	2.6	3.1	3.9	4.7	3.3	4.1 ± 0.6	1.7 ± 1.0	-6 ± 1
Full Update incl. Paleo $\Delta\lambda$	**2.1**	**2.4**	**2.8**	**3.4**	**4.0**	**2.9**	4.1 ± 0.6	1.7 ± 1.0	-6 ± 1
Alt. Update incl. Paleo $\Delta\lambda^{Alt}_{150yr}$	2.1	2.4	2.8	3.5	4.1	3.0	4.1 ± 0.6	1.7 ± 1.0	-6 ± 1

Combined, Robust Range (Unif. S Prior)	5th	17th	50th	83rd	95th	Mean	ΔT_{Plio}	ΔF^{Plio}_{NonGHG}	ΔT_{LGM}
SW20: Original Robust Range (Unif. S)	2.4	2.8	3.5	4.5	5.7	3.7	3.0 ± 1.0	f_{ESS}	-5 ± 1
+ Update ΔT, ΔF^{Plio}_{NonGHG} (Unif. S)	2.6	3.0	3.8	4.9	6.2	4.0	4.1 ± 0.6	1.7 ± 1.0	-6 ± 1
Full Update incl. Paleo $\Delta\lambda$ (Unif. S)	2.3	2.6	3.1	3.8	4.6	3.2	4.1 ± 0.6	1.7 ± 1.0	-6 ± 1
Alt. Update incl. Paleo $\Delta\lambda^{Alt}_{150yr}$ (Unif. S)	2.3	2.6	3.1	3.9	4.8	3.3	4.1 ± 0.6	1.7 ± 1.0	-6 ± 1

Pliocene Only (Unif. S Prior)	5th	17th	50th	83rd	95th	Mean	ΔT_{Plio}	ΔF^{Plio}_{NonGHG}
SW20: Original	1.6	2.4	4.0	6.8	10.1	4.7	3.0 ± 1.0	f_{ESS}
+ Update ΔT_{Plio}	2.9	3.8	5.6	8.6	12.3	6.3	3.0 ± 1.0	f_{ESS}
+ Update ΔF^{Plio}_{NonGHG}	2.5	3.2	4.7	7.4	11.2	5.4	4.1 ± 0.6	1.7 ± 1.0
Include Pliocene $\Delta\lambda$	1.9	2.4	3.8	7.2	12.9	5.0	4.1 ± 0.6	1.7 ± 1.0
Alt. Pliocene $\Delta\lambda^{Alt}_{150yr}$	1.8	2.4	3.8	8.3	14.8	5.3	4.1 ± 0.6	1.7 ± 1.0

Units in °C; ΔF units in W m^{-2}.

Table 3.S4: Paired estimates of Pliocene and LGM pattern effects, which use similar methods for data assimilation and the same AGCMs. The pairs are used to estimate the Pearson correlation and covariance between estimates of Pliocene and LGM pattern effects (Cooper et al., 2024). For the standard $\Delta\lambda, r = 0.56$ and cov $= 0.0123$ [W m^{-2} K^{-1}]^2. For $\Delta\lambda_{150\text{yr}}^{\text{Alt}}$, $r = 0.87$ and cov $= 0.0562$ [W m^{-2} K^{-1}]^2. Table units are W m^{-2} K^{-1}. LGM results use updated CESM2.1 $\lambda_{150\text{yr}}^{\text{Alt}}$ in Table 3.S1.

AGCM	Plio Pattern	LGM Pattern	$\Delta\lambda_{\text{Plio}}$	$\Delta\lambda_{\text{LGM}}$	$\Delta\lambda_{\text{Plio}}^{\text{Alt150}}$	$\Delta\lambda_{\text{LGM}}^{\text{Alt150}}$
CAM4	plioDA	LGMR	-0.57	-0.45	-0.41	-0.21
CAM5	plioDA	LGMR	-0.48	-0.31	-0.67	-0.41
CAM6	plioDA	LGMR	-0.69	-0.63	-1.08	-1.02
AM4	plioDA	LGMR	-0.44	-0.33	-0.37	-0.27
HadGEM3	plioDA	LGMR	-0.20	-0.27	-0.19	-0.29
CAM4	Annan	Annan	-0.57	-0.29	-0.10	-0.06
CAM5	Annan	Annan	-0.48	-0.09	-0.43	-0.18
CAM4	plioDA: Cloud Prior	LGMR	-0.63	-0.45	-0.47	-0.21
CAM5	plioDA: Cloud Prior	LGMR	-0.56	-0.31	-0.76	-0.41
CAM4	plioDA: Cloud Prior	lgmDA	-0.63	-0.69	-0.47	-0.45
CAM5	plioDA: Cloud Prior	lgmDA	-0.56	-0.51	-0.76	-0.61
CAM4	plioDA	lgmDA	-0.57	-0.69	-0.41	-0.45
CAM5	plioDA	lgmDA	-0.48	-0.51	-0.67	-0.61

Chapter 4

MONTHLY SEA-SURFACE TEMPERATURE, SEA ICE, AND SEA-LEVEL PRESSURE OVER 1850–2023 FROM COUPLED DATA ASSIMILATION

4.1 Abstract

Historical observations of Earth's climate underpin our knowledge and predictions of climate variability and change. However, the observations are incomplete and uncertain, and existing datasets based on these observations typically do not assimilate observations simultaneously across different components of the climate system, yielding inconsistencies that limit understanding of coupled climate dynamics. Here we use coupled data assimilation, which synthesizes observational and dynamical constraints across all climate fields simultaneously, to reconstruct globally resolved sea-surface temperature (SST), near-surface air temperature (T), sea-level pressure (SLP), and sea-ice concentration (SIC), over 1850–2023. We use a Kalman filter and forecasts from an efficient emulator (Linear Inverse Model; LIM) to assimilate observations of SST, land T, marine SLP, and satellite-era SIC. We account for model error by training LIMs on eight CMIP6 models, and we use the LIMs to generate eight independent reanalyses with 200 ensemble members, yielding 1600 total members. Key findings in the Tropics include post-1980 trends in the Walker circulation that are consistent with past variability, whereas the tropical SST contrast (the difference between warmer and colder SSTs) shows a distinct strengthening since 1975. ENSO amplitude exhibits substantial low-frequency variability and a local maximum in variance over 1875–1910. In polar regions, we find a muted cooling trend in the Southern Ocean post-1980 and substantial uncertainty. Changes in Antarctic sea ice are relatively small between 1850 and 2000, while Arctic sea ice declines by

0.5 ± 0.1 (1σ) million km^2 during the 1920s.

4.2 Significance statement

The key advance in our reconstruction is that the ocean, atmosphere, and sea ice are dynamically consistent with each other and with observations across all components, thus forming a true climate reanalysis. Existing climate datasets are typically derived separately for each component (e.g., atmosphere, ocean, and sea ice), leading to spurious trends and inconsistencies in coupled climate variability. We use coupled data assimilation to unify observations and coupled dynamics across components. We combine forecasts from climate models with observations from ocean vessels and weather stations to produce monthly state estimates spanning 1850–2023 and a novel quantification of globally resolved uncertainty. This reconstruction provides insights into historical variability and trends while motivating future efforts to reduce uncertainties in the climate record.

4.3 Introduction

The historical record (c. 1850–present) is central to our understanding of climate variability and Earth's response to anthropogenic forcings, but we have yet to fully extract the available information from instrumental data. Observations of sea-surface temperature (SST), near-surface air temperature (T), and sea-level pressure (SLP) from ships of opportunity and weather stations are noisy, sparse, and vary over time, which adds an incomplete-data problem (Schneider, 2001) to analyses of climate variability and change that cannot be avoided and should not be ignored.

To prepare observations for climate analysis, data sources must first be homogenized (e.g., Kent and Kennedy, 2021; Chan and Huybers, 2019; Chan et al., 2023; Karl et al., 2015; Hausfather et al., 2017; Cowtan et al., 2018), and then the missing values must be imputed. Imputation is typically statistical, employing pattern-based methods (including empirical orthogonal functions; EOFs) derived from recent decades or kriging, and does not involve dynamical constraints (e.g., Kaplan et al., 1998; Rohde et al., 2013; Cowtan and Way, 2014; Hirahara et al., 2014; Huang et al., 2017; Kadow et al., 2020; Vaccaro et al., 2021). Furthermore, when values are imputed for different climate fields, e.g., SST and SLP, there are no dynamical constraints ensuring that the coupled fields are physically consistent. Imputation and homogenization can have pronounced impacts on assessments of the climate sensitivity to increasing greenhouse gases (e.g., Sherwood et al., 2020;

Forster et al., 2021; Modak and Mauritsen, 2023), efforts to distinguish internal variability from forced climate change (e.g., Hegerl et al., 2019; Wills et al., 2020), understanding of atmosphere–ocean variability (e.g., Battisti et al., 2019), and evaluation of climate models (e.g., Wills et al., 2022; Simpson et al., 2025). Here we apply a different approach to solve the incomplete-data problem: we use coupled data assimilation to impose observational and dynamical constraints across all climate fields simultaneously, ensuring that the full state estimate is internally consistent.

SST patterns play a ubiquitous role in regulating climate variability (e.g., Bjerknes, 1969; Barsugli and Battisti, 1998; Alexander et al., 2002; Deser et al., 2010a; Newman et al., 2016; Czaja et al., 2019; Capotondi et al., 2023), radiative feedbacks (e.g., Armour et al., 2013; Andrews et al., 2015; Zhou et al., 2016; Ceppi and Gregory, 2017; Andrews and Webb, 2018; Dong et al., 2019; Fueglistaler, 2019; Andrews et al., 2022; Salvi et al., 2023; Cooper et al., 2024), and the hydrologic cycle (Hastenrath and Greischar, 1993; Xie et al., 2010; Hoerling et al., 2010; Chadwick et al., 2014; Lehner et al., 2018; Siler et al., 2019; Cook et al., 2022; Kuo et al., 2023; Seager et al., 2023). There are a variety of recent (c. 1980–present) climate phenomena tied to SSTs that seem either unprecedented or unremarkable depending on what we deem to be natural variability, and this interpretation of recent trends relies on the incomplete and brief instrumental record (e.g., Wunsch, 1999). In the tropical Pacific, the zonal SST gradient has strengthened (Solomon and Newman, 2012; Coats and Karnauskas, 2017; Lee et al., 2022; Watanabe et al., 2024), with cooling in the East Pacific and warming in the West Pacific that has coincided with an apparent strengthening of the Walker circulation post-1980 (e.g., L'Heureux et al., 2013; McGregor et al., 2014; Watanabe et al., 2023; Heede and Fedorov, 2023) and distinct tropospheric temperature trends (Flannaghan et al., 2014; Fueglistaler, 2019; Po-Chedley et al., 2021). At the poles, the Arctic has warmed rapidly since 1980 with substantial loss of sea ice (Dörr et al., 2023; England et al., 2021; Notz and SIMIP Community, 2020), while the Southern Ocean has cooled with an overall expansion of sea ice—until 2015, after which the Southern Ocean has shown rapid warming and sea ice loss (Fan et al., 2014; Stuecker et al., 2017; Fogt et al., 2022; Espinosa et al., 2024; Roach and Meier, 2024; Zhang and Li, 2023; Turner et al., 2022; Dong et al., 2023; Suryawanshi et al., 2023; Bonan et al., 2024).

A major challenge for the climate dynamics community is understanding the causes of these observed changes as well as the apparent yet debated inability of our state-of-the-art coupled climate models to replicate them (e.g., Wills et al., 2022; Dong et al., 2021; Rugenstein et al., 2023; Seager

et al., 2022; Olonscheck et al., 2020; Chung et al., 2019; Watanabe et al., 2021; Roach et al., 2020; Notz and SIMIP Community, 2020; Chemke et al., 2022; Kang et al., 2024). Progress on this endeavor requires robustly quantifying observational uncertainties and placing recent changes in historical context with reliable reconstructions of past climate changes. For example, are the post-1980 trends in tropical SST gradients, the Walker circulation, and polar climates unique over the historical record, or have such changes occurred often due to internal climate variability?

Existing SST datasets designed for climate analysis use a variety of statistical interpolation methods. These methods have been recently summarized in Modak and Mauritsen (2023) and Lewis and Mauritsen (2021) and described in detail in a review by Kent and Kennedy (2021), which also explains the extensive efforts to homogenize time-varying sources of in situ data. To assess the atmospheric response to SST and sea-ice concentration (SIC) over the historical record in atmospheric general circulation models (i.e., in AMIP-type simulations; Eyring et al., 2016; Webb et al., 2017), complete coverage and monthly resolution of SST/SIC is required. Combined SST/SIC datasets for this purpose include the 1870–2022 PCMDI/AMIP-II boundary condition (Hurrell et al., 2008) used as the standard for CMIP6, 1854–present NOAA ERSSTv5 (Huang et al., 2017), Met Office Hadley Centre's 1870–present HadISST1 (Rayner et al., 2003) and 1850–2010 HadISST2.1 (no longer maintained; Titchner and Rayner, 2014), and the Japanese Meteorological Agency's 1850–present COBE-SST2 (Hirahara et al., 2014). Kaplan et al. (1998) developed a landmark SST analysis using optimal interpolation, and since then the incomplete-data problem has been addressed using kriging (Cowtan and Way, 2014), Markov random graphs (Vaccaro et al., 2021), and machine learning (Kadow et al., 2020) to impute hybrid air-sea surface temperatures over land and ocean.

Figure 4.1 depicts the time-evolving observing network of in situ SST measurements in HadSST4 (Kennedy et al., 2019). As motivation for this study, we illustrate the spread (1σ) across existing datasets (HadISST1, HadISST2.1, ERSSTv5, COBE-SST2, and AMIP-II) in their preindustrial-baseline SST (mean anomaly over years 1870–1899) and the spread in their SST trends from 1900–1979 and 1980–2010. We separate the satellite era (c. 1980–present) from the earlier warming because of the variety of studies highlighting and questioning the peculiarity of recent SST trends (e.g., Fueglistaler and Silvers, 2021; Andrews et al., 2022; Lewis and Mauritsen, 2021). The spatial pattern of uncertainty is influenced by varying methods of imputation, homogenization of data
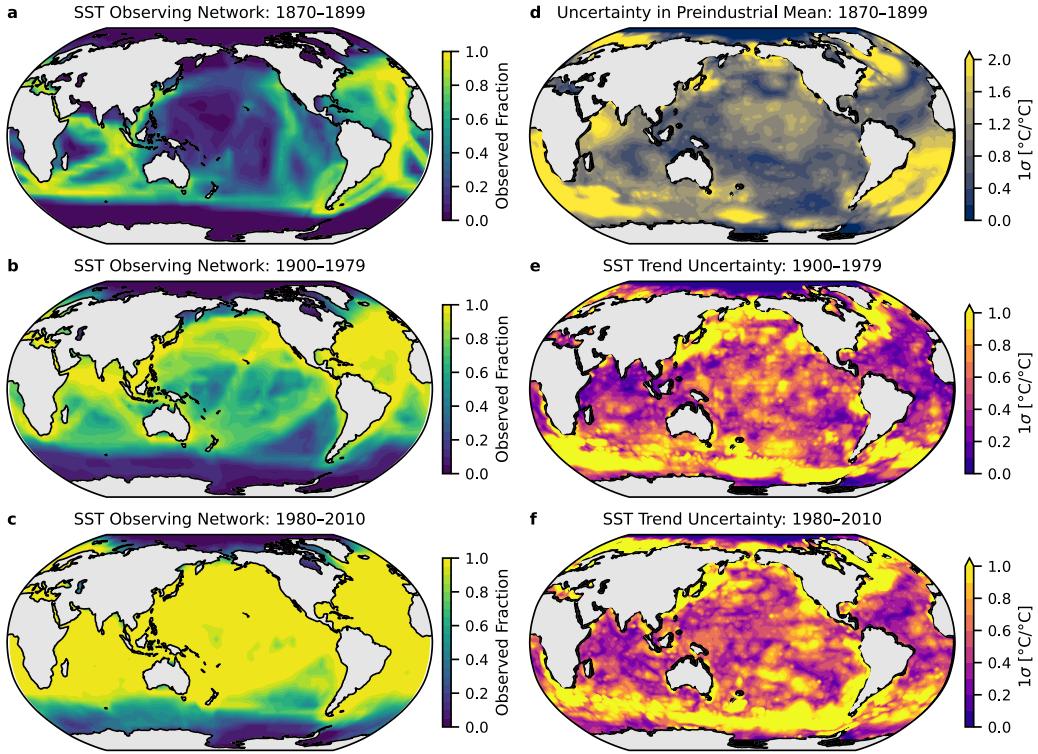
Figure 4.1: Historical observing network and SST uncertainty in pre-existing infilled datasets. **(a–c)** Fraction of months with in situ data for SST over three time periods in HadSST4, where 1.0 indicates data in every month during the period. **(d)** Illustration of systematic uncertainty in normalized pattern of preindustrial-mean SST anomalies across existing infilled datasets, calculated as the sample standard deviation (1σ) of the 1870–1899 mean anomalies across HadISST1, HadISST2.1, ERSSTv5, PCMDI/AMIP-II, and COBE-SST2, relative to their 1961–1990 climatologies; local anomalies are divided by global-mean anomalies (60°S–60°N) to highlight uncertainty in spatial patterns. **(e–f)** Illustration of systematic uncertainty in patterns of SST trends, calculated as the 1σ of local trends across the same datasets in panel **d**; local SST trends are first divided by the global-mean SST trends (60°S–60°N) to highlight uncertainty in the patterns, and local values greater than 1.0 indicate that the local 1σ is greater than the global-mean trend. Note different colorbars in panels **d–f**.

sources, and representativeness error in using point observations as estimates of grid-scale means. Even after 1980, the data coverage over the Southern Ocean and southeast Pacific is notably far from complete, and the inter-dataset differences in those regions are substantial in recent decades (Figure 4.1c,f).

Atmospheric reanalyses address the incomplete-data problem with data assimilation, which uses a weather model's dynamics to constrain the atmospheric state. Data assimilation (DA) broadly describes the collection of methods that synthesize model forecasts with sparse and noisy observations, producing posterior analyses and uncertainties that are subject to the dynamical constraints of the model. DA is computationally intensive, hence existing reanalyses only assimilate atmospheric observations and only apply dynamical constraints to the atmospheric component, meaning that the SST and SIC boundary conditions are prescribed a priori in, for example, ERA5 (Hersbach et al., 2020; Soci et al., 2024), JRA-55 and JRA-3Q (Kobayashi et al., 2015; Kosaka et al., 2024), NOAA-CIRES-DOE's 20th Century Reanalysis (Compo et al., 2011; Slivinski et al., 2019), NCEP/NCAR Reanalysis (Kalnay et al., 1996), MERRA-2 (Gelaro et al., 2017), and ModE-RA (Franke et al., 2017; Valler et al., 2024).

Coupled atmosphere–ocean reanalysis remains a frontier and formidable challenge in climate research. ECMWF's coupled DA program, CERA-20C (Laloyaux et al., 2018), is now inactive, and ECMWF no longer hosts its output. NCEP's CFSR made a major advance (Saha et al., 2010), which assimilated observations into atmosphere and ocean components separately, and incorporated coupling by using a coupled model during the forecast step—this process is known as "weakly coupled" DA. The UFS-Replay dataset (NOAA, 2024), employs a weakly coupled "replay" approach (Orbe et al., 2017), in which the coupled UFS model is nudged toward the existing ERA5 atmospheric reanalysis and ORAS5 ocean reanalysis (Zuo et al., 2019). In this study, we will use "strongly coupled" DA, which (i) ensures that the coupled atmosphere-ocean-ice state is internally consistent and (ii) synthesizes observational and dynamical constraints across each component simultaneously.

To circumvent the computational obstacles associated with DA in fully coupled models, lightweight DA methods have been developed primarily for paleoclimate reconstruction, as reviewed by Tierney et al. (2025a). The "offline" DA method uses a static, uninformed prior from pre-existing model output (e.g., Hakim et al., 2016; Franke et al., 2017; Steiger et al., 2014, 2018; Samakinwa et al.,

2021; Tierney et al., 2020; Osman et al., 2021; Annan et al., 2022; Smerdon et al., 2023; Valler et al., 2024; Tierney et al., 2025b). "Online" methods use a time-evolving prior that is informed by the previous initial conditions produced by data assimilation, thus retaining memory of past observations. Online DA requires integrating a forecast model after each assimilation step, which is the main computational bottleneck.

Data-driven approaches that emulate climate models can overcome the computational bottleneck. The linear inverse model (LIM) has been tested in annual-mean DA with proxies over the last millennium (Perkins and Hakim, 2021) and for subseasonal forecasting (Hakim et al., 2022). LIMs have been applied to study dynamics and predictability of the El Niño–Southern Oscillation (ENSO) (e.g., Penland and Sardeshmukh, 1995; Shin et al., 2021; Vimont et al., 2014; Lou et al., 2020; Kido et al., 2023), meridional modes (Vimont, 2012), global surface temperatures (Newman, 2013), SSTs in the North Atlantic (e.g., Zanna, 2012) and North Pacific (e.g., Newman, 2007; Newman et al., 2016; Zhao et al., 2024), the Pacific-North American pattern (Henderson et al., 2020), hydroclimate (Coats et al., 2020; Tseng et al., 2021), and sea ice (Brennan et al., 2023). LIMs are computationally efficient, enabling coupled assimilation of observations across Earth system components, e.g., pressure observations in the atmosphere and SST observations in the ocean can each inform both SST and SLP in coupled DA. Combining LIMs with data assimilation presents an opportunity to constrain and quantify uncertainty in the historical climate record.

Here we use coupled data assimilation to reconstruct monthly and globally resolved SST, near-surface air temperature (T), SLP, and SIC over 1850–2023. The novelty of our approach compared to past reanalyses is that we constrain all climate fields simultaneously with (i) coupled dynamics and (ii) observations across climate components. Our DA method is made computationally tractable by efficient emulators (LIMs), which are trained on eight CMIP6 models and capture the essential dynamics for monthly reanalysis. We combine forecasts from LIMs with a Kalman filter to produce a coupled reconstruction with time-varying uncertainty quantification. Section 4.4 describes methods and data, including LIMs, data assimilation, validation with an out-of-sample pseudo-reconstruction, observations, and comparison datasets. Section 4.5 presents the historical reconstruction. Section 4.6 discusses the implications of the results for interpreting climate variability and change and the caveats of the method. Section 4.7 presents the conclusions.

4.4 Methods and data

In this section, we describe the reconstruction method, validation, and data sources. The reconstruction of monthly means consists of (i) a monthly forecast, for which we use LIMs that emulate eight CMIP6 models, and (ii) data assimilation in every month, for which we use the classic Kalman filter (Kalman, 1960; Kalnay, 2003). We validate the method with a pseudo-reconstruction of a climate model's 1850–2014 historical simulation (MPI-ESM1-2-HR), from which we draw observations that mimic the true observing network.

4.4.1 Linear inverse models

Anomalies around an equilibrium state in the nonlinear climate system can be approximated as a stochastically forced, linear dynamical system (e.g., Hasselmann, 1976; Penland and Sardeshmukh, 1995; Penland, 1996):

$$\frac{d\mathbf{x}}{dt} = \mathbf{L}\mathbf{x} + \mathbf{S}\eta, \tag{4.1}$$

where $\mathbf{x}$ is a state vector of N principal components of SST, T, SLP, and SIC, $\mathbf{L}$ is an $N \times N$ linear operator representing the deterministic dynamics, and $\mathbf{S}\eta$ approximates the unresolvable nonlinear dynamics as stochastic forcing with an $N \times N$ noise-amplitude matrix, $\mathbf{S}$, and a vector, η, of independent, Gaussian white noise with unit variance and length N.

LIMs typically assume stationary statistics, but Shin et al. (2021) extend the LIM framework to include monthly variations in the dynamics. The monthly, or "cyclostationary" LIM, has been applied to ENSO (Shin et al., 2021; Vimont et al., 2022; Kido et al., 2023). We build on this recent work and use cyclostationary LIMs to model global SST, T, SLP, and SIC. We use the fixed-phase approach (Shin et al., 2021; OrtizBeviá, 1997) to train the 12 $\mathbf{L}_j$ operators in the cyclostationary LIM, where j indicates the month:

$$\mathbf{L}_j = \tau^{-1} \log[\mathbf{C}_j(\tau)\mathbf{C}_j(0)^{-1}], \quad \text{for } j = 1, 2, ..., 12. \tag{4.2}$$

$\mathbf{C}_j(\tau)$ and $\mathbf{C}_j(0)$ are the τ-lag and zero-lag covariance matrices of $\mathbf{x}$ for month j, and $\tau = 1$ month in all of the following equations. The stochastic amplitude matrices, $\mathbf{S}_j$, are estimated from the fluctuation-dissipation relation of Equation (4.1) (Penland and Matrosova, 1994),

$$\frac{d\mathbf{C}_j(0)}{dt} = \mathbf{L}_j\mathbf{C}_j(0) + \mathbf{C}_j(0)\mathbf{L}_j^T + \mathbf{Q}_j, \tag{4.3}$$

where $\mathbf{Q}_j = \mathbf{S}_j\mathbf{S}_j^T$. We follow Shin et al. (2021) in estimating the cyclostationary $\mathbf{Q}_j$ as

$$\mathbf{Q}_j = \frac{\mathbf{C}_{j+1}(0) + \mathbf{C}_{j-1}(0)}{2\Delta t} - [\mathbf{L}_j\mathbf{C}_j(0) + \mathbf{C}_j(0)\mathbf{L}_j^T], \tag{4.4}$$

with $\Delta t = 1$ month. Before computing $\mathbf{L}_j$ and $\mathbf{Q}_j$, we take the 3-month running means of $\mathbf{C}_j(\tau)$ and $\mathbf{C}_j(0)$, e.g., we estimate $\mathbf{C}_j(\tau) \approx \langle\mathbf{C}_{j-1}(\tau), \mathbf{C}_j(\tau), \mathbf{C}_{j+1}(\tau)\rangle$, where angle brackets denote an equal-weighted average (Shin et al., 2021). As in previous LIM studies (e.g., Penland, 1996), we remove any negative eigenvalues in $\mathbf{Q}_j$ and rescale remaining eigenvalues to conserve the original variance.

The LIM produces forecasts at lead $\tau = 1$ month from integrating (4.1) in time as

$$\mathbf{x}(t + \tau) = \mathbf{G}_j\mathbf{x}(t) + \mathbf{n}, \tag{4.5}$$

where $\mathbf{G}_j = \exp(\mathbf{L}_j\tau) = \mathbf{C}_j(\tau)\mathbf{C}_j(0)^{-1}$. The integrated stochastic term, $\mathbf{n}$, equals zero in a deterministic forecast, such as the prior-mean forecast in the Kalman filter as described below.

The forecast equation for the error covariance, assuming no correlation between error and state, is

$$\mathbf{P}(t + \tau) = \mathbf{G}_j\mathbf{P}(t)\mathbf{G}_j^T + \mathbf{N}_j(\tau). \tag{4.6}$$

We forecast the full covariance matrix with the LIM, instead of estimating it from ensemble members, because this approach is exact for a given LIM. It is equivalent to using an infinite ensemble. To solve for $\mathbf{N}_j(\tau)$, we extend the logic that applies to the stationary LIM (Hakim et al., 2022; Penland, 1989) for the cyclostationary case. Equation (4.6) must be valid for any month's initial condition, including $\mathbf{C}_j(0)$, from which the monthly forecast must arrive at $\mathbf{C}_{j+1}(0)$ because the statistics are cyclostationary, therefore:

$$\mathbf{N}_j(\tau) = \mathbf{C}_{j+1}(0) - \mathbf{G}_j\mathbf{C}_j(0)\mathbf{G}_j^T. \tag{4.7}$$

We train separate LIMs to emulate the following eight CMIP6 models: CESM2, GFDL-ESM4, HadGEM3-GC3.1-LL, SAM0-UNICON, UKESM1.0-LL, NorESM2-LM, EC-Earth3, and E3SM-2-0. Our selection of models is informed by Lou et al. (2023), which found that this subgroup performs best in an analog method for ENSO forecasting, although we make two changes: we remove HadGEM3-GC3.1-MM to prevent having two versions of HadGEM3, and we substitute E3SMv2.0 (Qin et al., 2024) for CIESM because of issues simulating sea ice in CIESM (Lin et al., 2020). For

training data, we use preindustrial-control simulations with historical (1850–2014) simulations appended (summary in Appendix A). LIMs are trained separately for each model using monthly mean anomalies, and each LIM has a minimum of 665 years of training data (500+ preindustrial and 165 historical years). While approximately 100 years of training data is sufficient for a Tropics-only cyclostationary LIM (Shin et al., 2021), global LIMs require a longer record. Thus long preindustrial simulations are essential for training, and we find that appending the historical simulations expands the footprint of regions with nonzero SIC variability in the training data, which improves reconstruction of SIC.

We regrid all training data bilinearly to 2° resolution (96 × 144 latitude-longitude grid). For consistency with observations, which are expressed as anomalies relative to a 1961–1990 climatology, we remove the mean and climatology in each gridcell calculated over 1961–1990 for each model. Separately for each model and state variable, we compute EOFs area-weighted by the square-root of the cosine of latitude for SST, T, SLP, Northern Hemisphere (NH) SIC, and Southern Hemisphere (SH) SIC. We retain approximately 85% of each field's variance in the truncated state. We form each model's standardized state vector from its principal components, $\mathbf{x}_k$, as:

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_{SST}/\sigma_{SST} \\ \mathbf{x}_T/\sigma_T \\ \mathbf{x}_{SLP}/\sigma_{SLP} \\ \mathbf{x}_{SIC_{NH}}/\sigma_{SIC_{NH}} \\ \mathbf{x}_{SIC_{SH}}/\sigma_{SIC_{SH}} \end{bmatrix},$$

where σ_k^2 is the retained variance after EOF truncation of field k. We use the standardized state vectors $\mathbf{x}$ to compute covariance matrices for each model, and we project into and out of the LIM basis by storing the EOFs and scale factors, σ_k, for each field. Each LIM is run independently in parallel through the data assimilation framework.

4.4.2 Data assimilation

Given a prior forecast of the state's monthly mean $\mathbf{x}_f$ and error covariance $\mathbf{P}_f$, we assimilate observations to produce the posterior analysis $\mathbf{x}_a$ and $\mathbf{P}_a$ using the Kalman filter:

$$\mathbf{x}_a = \mathbf{x}_f + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_f), \tag{4.8}$$

$$\mathbf{P}_a = [\mathbf{I} - \mathbf{KH}]\mathbf{P}_f, \tag{4.9}$$

$$\mathbf{K} = \mathbf{P}_f\mathbf{H}^T[\mathbf{HP}_f\mathbf{H}^T + \mathbf{R}]^{-1}, \tag{4.10}$$

where $\mathbf{K}$ is the Kalman gain, $\mathbf{y}$ is the vector of observations, $\mathbf{H}$ is the linear observation operator, and $\mathbf{R}$ is the observation error covariance. After solving (4.8–4.10) for a given month, we forecast the next month from (4.5), with $\mathbf{n} = 0$, and (4.6).

Our method is "strongly coupled online DA," where "strongly coupled" means that we assimilate observations concurrently across the atmosphere–ocean–ice system, and all fields influence each other through cross-component covariances. "Online" means that we use a forecast model with the previous assimilation step's initial conditions to inform the prior. Because this method uses the classic Kalman filter and propagates $\mathbf{P}_f$ exactly, we avoid the sample error and localization issues that arise when estimating $\mathbf{P}_f$ in an ensemble Kalman filter (Evensen, 1994; Houtekamer and Zhang, 2016). However, ensemble-member trajectories are needed to analyze statistics of temporal variability, and this variability must be constrained by dynamics rather than sampled independently (Emile-Geay et al., 2024). We solve this problem with a modified version of the ensemble Kalman filter, described subsequently, that has no impact on the mean or covariance, (4.8–4.9), but rather simply provides sample estimates from the posterior distribution.

We generate ensemble members using the perturbed-observations version of the ensemble Kalman filter (Houtekamer and Mitchell, 1998; Burgers et al., 1998), except we use the exact prior covariance forecast from the classic Kalman filter (4.6, 4.9). For each LIM, we initialize 200 ensemble members in January 1850 with random draws from a multivariate-normal distribution with covariance $\mathbf{C}_1(0)$. Each ensemble member is updated using (4.8), with $\mathbf{x}_f^n$ corresponding to ensemble member n in place of the ensemble mean, and $\mathbf{y}^n$ is a multivariate-normal random draw of the observations with mean $\mathbf{y}$ and covariance $\mathbf{R}$. After the assimilation, each $\mathbf{x}_a^n$ is advanced to the next month using (4.5). The noise term in (4.5), $\mathbf{n}$, becomes a random draw from $\mathbf{N}_j(\tau)$ in (4.7) for each ensemble member. Because our LIMs are built to forecast monthly means, we can draw from the distributions of the monthly statistics rather than stochastically integrating (Penland and Matrosova, 1994) each ensemble member. Spread across the ensemble therefore arises from each member having different initial conditions, different realizations of the stochastic term in the forecast model, and different observations due to the perturbations and the unique values of the

200-member ensemble of observations from HadSST4.

An additional benefit of the ensemble is that we can propagate temporally correlated observation errors that are associated with uncertainties in bias corrections. For example, HadSST4 (described below) provides a 200-member ensemble of monthly SST observations to represent temporally correlated errors (Kennedy et al., 2019). To incorporate these errors, we let $\mathbf{y}$ vary across the ensemble members, but each of our 200 ensemble members $\mathbf{x}^n$ is paired at every timestep with the corresponding ensemble member n from the HadSST4 ensemble.

4.4.3 Observations

We use four sources of observations corresponding to each of the four state variables (SST, T, SLP, SIC). All observations are anomalies relative to a 1961–1990 climatology, which is the period chosen by Kennedy et al. (2019) and Osborn et al. (2021).

SST observations are from HadSST4 version 4.0.1.0 (Kennedy et al., 2019), provided by the Met Office Hadley Centre on a $5° \times 5°$ grid. HadSST4 quality controls and corrects biases in the in situ measurements from ICOADS 3.0.0 (1850–2014) and ICOADS 3.0.1 (2015–present), the central database of ship records (Freeman et al., 2017). HadSST4 provides non-infilled data as monthly means spanning 1850–present, and ship coverage varies substantially over time (Figure 4.1). Measurement and sampling errors are provided for every gridcell and month, and error covariance matrices are provided that estimate the spatially correlated errors. We include these sources of error in $\mathbf{R}$. Temporally correlated errors from uncertain bias corrections are estimated with a 200-member ensemble of observations, and we account for these errors with our ensemble DA method, described in Section 4.4.2.

Observations of near-surface air temperature (T) over land are from CRUTEM5 version 5.0.2.0 (Osborn et al., 2021). The weather-station data is quality controlled, bias-corrected, and provided as non-infilled monthly means with error estimates on a $5° \times 5°$ grid. We include CRUTEM5's time-varying measurement and sampling errors in $\mathbf{R}$.

SLP observations are from ICOADS Enhanced Release 3.1 for 1850–2014 and Release 3.0.2, for 2015–2023 (Freeman et al., 2017), which only includes marine data. Ideally, we would assimilate terrestrial SLP observations, but the ISPD dataset (Cram et al., 2015) of surface pressures does

not have a homogenized product available that combines data at various elevations into a gridded dataset of monthly means. The lack of direct constraints on terrestrial SLP is a limitation of our reconstruction; hence, we focus our analysis on marine SLP. ICOADS marine SLP data are provided as monthly means on a 2° × 2° grid, along with the number of observations, n_{obs}, in each month and the intra-month standard deviation, s, of the observations in each gridcell. The baseline climatology for anomalies is from Hersbach et al. (2020). There are a large number of SLP observations due to the finer grid of ICOADS compared to HadSST4. We eliminate observations with $n_{\mathrm{obs}} < 5$, which are expected to have a low signal-to-noise ratio. For months that have data in more than 3000 gridcells, we mask up to 40% of the values between 25°S and 60°N using random sampling. These limits increase computational efficiency of the assimilation and maintain a reasonable balance between the number of SLP and SST observations, otherwise there would be approximately five times as many SLP as SST observations. Valler et al. (2024) also reduce the number of ICOADS observations of SLP in their atmospheric reanalysis and set a similar threshold of $n_{\mathrm{obs}} = 10$ per gridcell. Past studies identified a bias in ICOADS SLP data before 1870, which is discussed in Slivinski et al. (2019), Freeman et al. (2017), and Allan and Ansell (2006). NOAA 20CRv3 performed a bias correction of the pre-1870 SLP observations, so we substitute the 1850–1870 SLP from ICOADS with the collocated values from NOAA 20CRv3. ICOADS does not provide an estimate of measurement and sampling errors which comprise the diagonal terms in $\mathbf{R}$. As described in Appendix B, we estimate $\mathbf{R}$ from the intramonth spread in individual observations and the variance across neighboring observations.

Sea ice observations are provided by the NOAA/NSIDC Climate Data Record (CDR) of Passive Microwave Sea Ice Concentration, Version 4 from 11/1978–09/2023 and Near-Real-Time, Version 2, for 10/2023–12/2023 (Meier et al., 2021b,a). We coarsen the observations from 25 km to 2° resolution. At each timestep in the assimilation with satellite data, we use a subset of the available data, which has nearly complete coverage of the polar regions. We retain all observations with SIC ranging from 0.01 to 0.98 and 40% of the remaining observations using random sampling. For measurement and sampling errors that form the diagonal terms in $\mathbf{R}$, we use the provided standard deviations of daily values, but we set the minimum error to 0.01. As described in Appendix B for SLP, these intramonth standard deviations approximate the monthly mean error. For SIC, they are calculated across both the NASA Team and Bootstrap algorithms, sampling the systematic error

across data-processing methods. Errors are small in open water and pack ice but are often between 0.3 and 0.5 in partial ice cover. We do not have satellite data for sea ice from 1961–1978, but we require a full climatology from 1961–1990 to calculate the SIC anomalies relative to a baseline that is consistent with the HadSST4 anomalies. The mean of the eight models used for LIM training agrees well with observations over the satellite era (SI of Roach et al., 2020; Notz and SIMIP Community, 2020), so we combine the multi-model mean of the eight historical simulations from 1/1961–11/1978 with the satellite data from 12/1978–12/1990, and we use the merged climatology from 1961–1990 as the reference for SIC anomalies. Because solutions to (4.8) are not restricted to SIC between zero and one, we use the climatology in postprocessing to ensure that SIC is between zero and one.

4.4.4 Validation: Pseudo-reconstruction of an out-of-sample model

To test our method, we mimic the real reconstruction problem and attempt to reconstruct the 1850–2014 historical simulation from a climate model. Our target model is MPI-ESM1-2-HR, ensemble member r1i1p1f1 (Mauritsen et al., 2019). We have chosen MPI-ESM1-2-HR because it is a difficult test of the method given that, unlike nearly all other models, it has cooling in the Southern Ocean from 1980–2014. It also has a low-bias in Antarctic sea ice (Roach et al., 2020) and substantially different ENSO statistics and radiative feedbacks (Bloch-Johnson et al., 2024) compared to the models used for LIMs and priors in the data assimilation. The pseudo-reconstruction's target is out-of-sample because MPI-ESM1-2-HR is not used for LIM training; the dynamics of the target model are unknown to our eight forecast models.

We draw pseudo-observations from the target simulation at the same times and locations where real observations are available for SST, T, SLP, and SIC. Random errors are added to the pseudo-observations by sampling from the real observation errors in $\mathbf{R}$. Note that real observations also have biases and unknown, unquantified errors which make the real reconstruction more challenging than this test. On the other hand, the LIMs used as model priors are selected based on their ability to collectively emulate reality rather than the target model of the pseudo-reconstruction.

Figure 4.2 shows timeseries representing climate variability from the pseudo-reconstruction. The ensemble mean is calculated as the grand mean across all 1600 ensemble members (8 LIMs
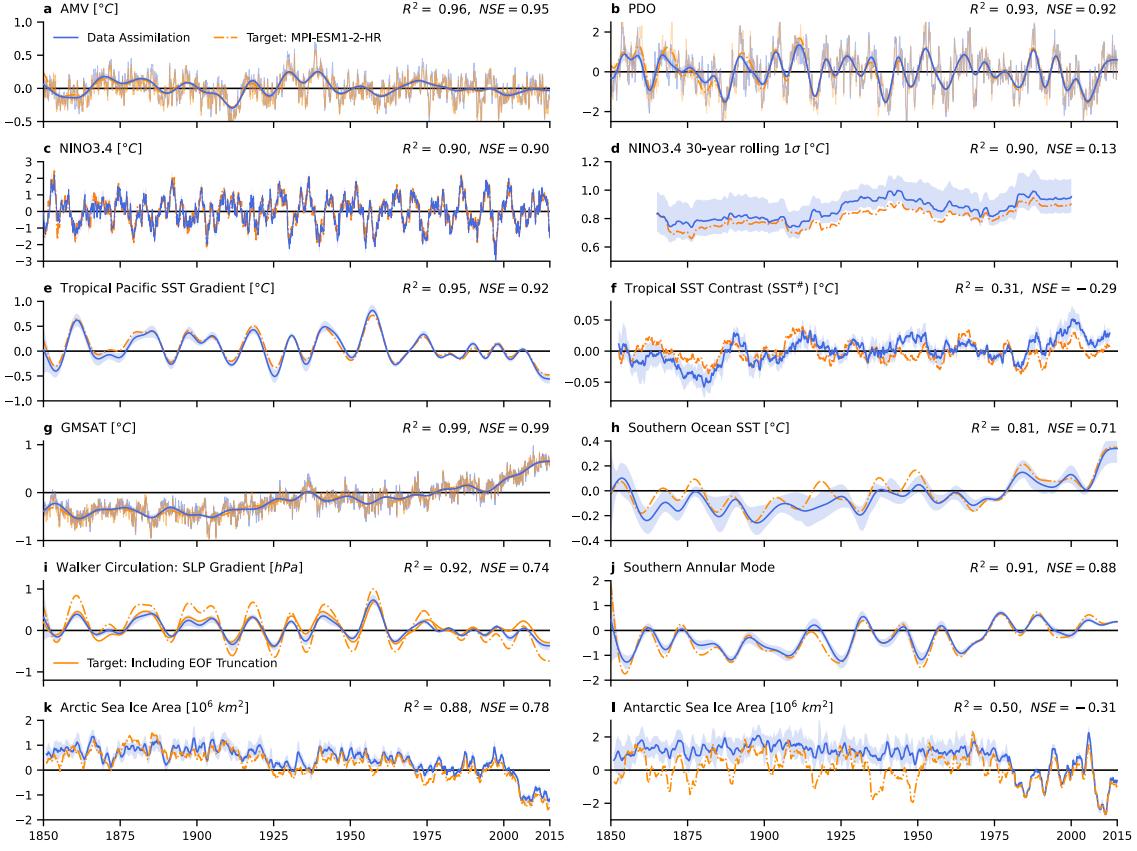
Figure 4.2: Validation by pseudo-reconstruction: timeseries. (Orange) Values from the target model, the 1850–2014 historical simulation from MPI-ESM1-2-HR. (Blue) Data assimilation result, showing mean of 1600 ensemble members; shading denotes 17$^{\text{th}}$ and 83$^{\text{rd}}$ percentiles. **(a)** Atlantic Multidecadal Variability with 10-yr low-pass filter and monthly values as thin lines. **(b)** Pacific Decadal Oscillation with 6-yr low-pass filter and monthly values as thin lines. **(c)** Monthly Nino3.4 with 30-yr running mean removed. **(d)** Rolling 30-yr standard deviation of Nino3.4 in panel **c**. **(e)** Zonal gradient of tropical Pacific SST with 10-yr low-pass filter. **(f)** Tropical SST contrast, SST$^{\#}$, 5-yr running mean. **(g)** Global-mean near-surface air temperature (GMSAT) with 10-yr low-pass filter and monthly values in thin lines. **(h)** Zonal mean of Southern Ocean SST (50°–70°S) with 10-yr low-pass filter. **(i)** Walker circulation, i.e., zonal SLP gradient across tropical Pacific, with 10-yr low-pass filter. **(j)** Southern Annular Mode with 10-yr low-pass filter. **(k)** Arctic and **(l)** Antarctic sea ice area, showing 12-month running mean. R^2 and Nash-Sutcliffe Efficiency (NSE) are based on the filtered metrics (see Supplemental Figure 4.S5 for unfiltered monthly results). Calculation of metrics is described in Section 4.4.4.

$\times$ 200 members), and the ensemble shading spans the 17th–83rd percentiles. We note that any one of the eight LIM-DA systems may not have a posterior distribution that spans the true state by itself. However, the grand ensemble of posterior distributions from all eight LIM-DA systems, which includes the spread from model error, generally spans the target (Figure 4.2). Because the grand ensemble represents eight separate DA systems, its distribution is non-Gaussian.

The metrics in Figure 4.2 are calculated as follows, with anomalies representing the departures from the 1961–1990 climatological annual cycle unless stated otherwise:

- Atlantic multidecadal variability (AMV) is the monthly mean SST anomaly in the North Atlantic (0°–60°N, 80°W–0°W) minus the global mean; the mean of the index from 1900–1970 is removed before plotting (Trenberth and Shea, 2006).

- The Pacific Decadal Oscillation (PDO) is the leading EOF of the monthly mean SST anomaly in the North Pacific (20°–70°N) after removing the global mean (Newman et al., 2016).

- Nino3.4 is the monthly mean SST anomaly over 170°W to 120°W and 5°S–5°N, with the 30-yr running mean removed.

- The zonal SST gradient in the tropical Pacific is the mean SST anomaly in the west (80°E–150°E) minus the east (160°W–80°W), spanning 5°S–5°N (e.g., Heede and Fedorov, 2023).

- $\text{SST}^{\#}$, which denotes the tropical SST contrast, is the mean of the warmest 30% of all tropical SSTs (30°S–30°N) minus the mean tropical SST, and the 1961–1990 mean is removed (Fueglistaler, 2019). We note that $\text{SST}^{\#}$ requires actual SSTs, not just anomalies. To estimate $\text{SST}^{\#}$ for the pseudo-reconstruction, we add reconstructed anomalies to the target model's 1961–1990 climatology, which assumes outside knowledge of the target's climatology. For the actual reconstruction of $\text{SST}^{\#}$, we add the reconstructed anomalies to the 1961–1990 climatology from HadISST2.1's ensemble mean. We show the 5-year running mean of $\text{SST}^{\#}$ for consistency with Fueglistaler and Silvers (2021).

- Southern Ocean SST is the zonal-mean SST anomaly from 50°–70°S (Doddridge and Marshall, 2017).

- Global-mean near-surface air temperature (GMSAT) is the global-mean T anomaly.

- The Walker circulation, measured by the zonal SLP gradient, is the mean SLP anomaly in the west Pacific (130°E–150°E) minus the central-east Pacific (160°W–120°W), spanning 5°S–5°N (e.g., Heede and Fedorov, 2023).

- The Southern Annular Mode (SAM) is the standardized zonal-mean SLP anomaly at 40°S ± 2° minus the standardized zonal-mean SLP anomaly at 65°S ± 2° (Gong and Wang, 1999); the reference period for standardization is 1961–1990, and each month is standardized separately.

- Sea ice area is the sum of the products of SIC and gridcell area; a common land mask is used when comparing ice area across various SIC datasets.

Most large-scale metrics are reconstructed with accuracy. We assess performance by the Pearson correlation (R), the fraction of variance explained (R^2), and the Nash-Sutcliffe Efficiency (NSE),

$$\text{NSE} = 1 - \frac{\sum (x_i - \hat{x}_i)^2}{\sum (x_i - \bar{x})^2},$$

which accounts for the relative phasing of the target timeseries (x_i) versus the reconstructed timeseries $(\hat{x}_i)$, the signal amplitude, and bias. The NSE has an upper bound equal to one and can become negative from biases in the mean or amplitude of variability (Nash and Sutcliffe, 1970). We find $R^2 > 0.80$ for the AMV, PDO, Nino3.4, the 30-year rolling 1σ of Nino3.4, the zonal SST gradient in the tropical Pacific, GMSAT, Southern Ocean SST, the Walker circulation (zonal SLP gradient), the SAM, and Arctic ice area. The tropical SST contrast, $SST^{\#}$, has the lowest R^2 at 0.31.

The reconstruction of the Walker circulation has a damped amplitude compared to the target, which is due to the EOF truncation of SLP in the LIM training. We show an additional version of the target model's Walker circulation, which is calculated after truncating the target's SLP into the leading 30 EOFs. Truncation is expected to affect tropical SLP because the variance in tropical SLP is low compared to the variance at higher latitudes, but truncation does not appear to have a substantial influence on other metrics.

Antarctic sea ice has $R^2 = 0.50$ and is biased high in the reconstruction before 1979. The reason for this bias is that the target model is biased low relative to the multi-model mean of the LIMs and

relative to the satellite record (Roach et al., 2020). There are decadal periods of abrupt ice loss in the target model which are not captured in the reconstruction. These ice-loss events are associated with brief warming episodes in Southern Ocean SST (Figure 4.2h), which are also not detected in the reconstruction. While we do not know whether such Antarctic ice-loss events happen in nature, we note that our pseudo-reconstruction of MPI-ESM1-2-HR does not capture its ice-loss events when observations are very sparse. Reasons for this deficiency could be (i) the LIMs used as model priors are too different from the target model, and sparse observations cannot overcome those differences, and/or (ii) the ice-loss events do not covary with available observations, and even a perfect model would be unable to reconstruct them from the data. Despite missing these decadal warmings, the lower-frequency variability in Southern Ocean SST and the SAM is captured by the reconstruction.

Figure 4.3 shows the pattern of trends in annual-mean SST for 1900–1979 and 1980–2014. Local trends are divided by the global-mean trend to emphasize the patterns. We also show the reconstruction's ensemble spread (1σ) in trend patterns, which highlights regions of elevated uncertainty. Note that SST is defined in all ocean gridcells at all times, even when SIC is 100%, so there are no missing values in the SST field. It is important to recall that observations in the Southern Ocean and southeast Pacific are sparse even after 1980 (Figure 4.1c), which is evident in our uncertainty quantification.

To further illustrate the uncertainty, we show trends from individual ensemble members (Figure 4.3c,g). These ensemble members show more cooling in the Southern Ocean than is seen in the ensemble mean. The key point, which is relevant to the next section on the real reconstruction, is that our DA framework is capable of reconstructing cooling over the Southern Ocean, even though the models used to train the LIMs do not show post-1980 cooling over the Southern Ocean in their historical simulations. Model biases can often be overcome if there are enough observations, and the LIM dynamics allow for cooling trends in the Southern Ocean. However, due to poor data coverage and quality in the Southern Ocean, SST trends in this region should be interpreted with caution.

Figure 4.4 shows trends in annual-mean SLP for 1900–1979 and 1980–2014. We only assimilate marine SLP observations, hence terrestrial SLP is expected to deviate from the target model. Large-scale patterns are consistent, but the errors in the magnitude of trends are substantial, especially
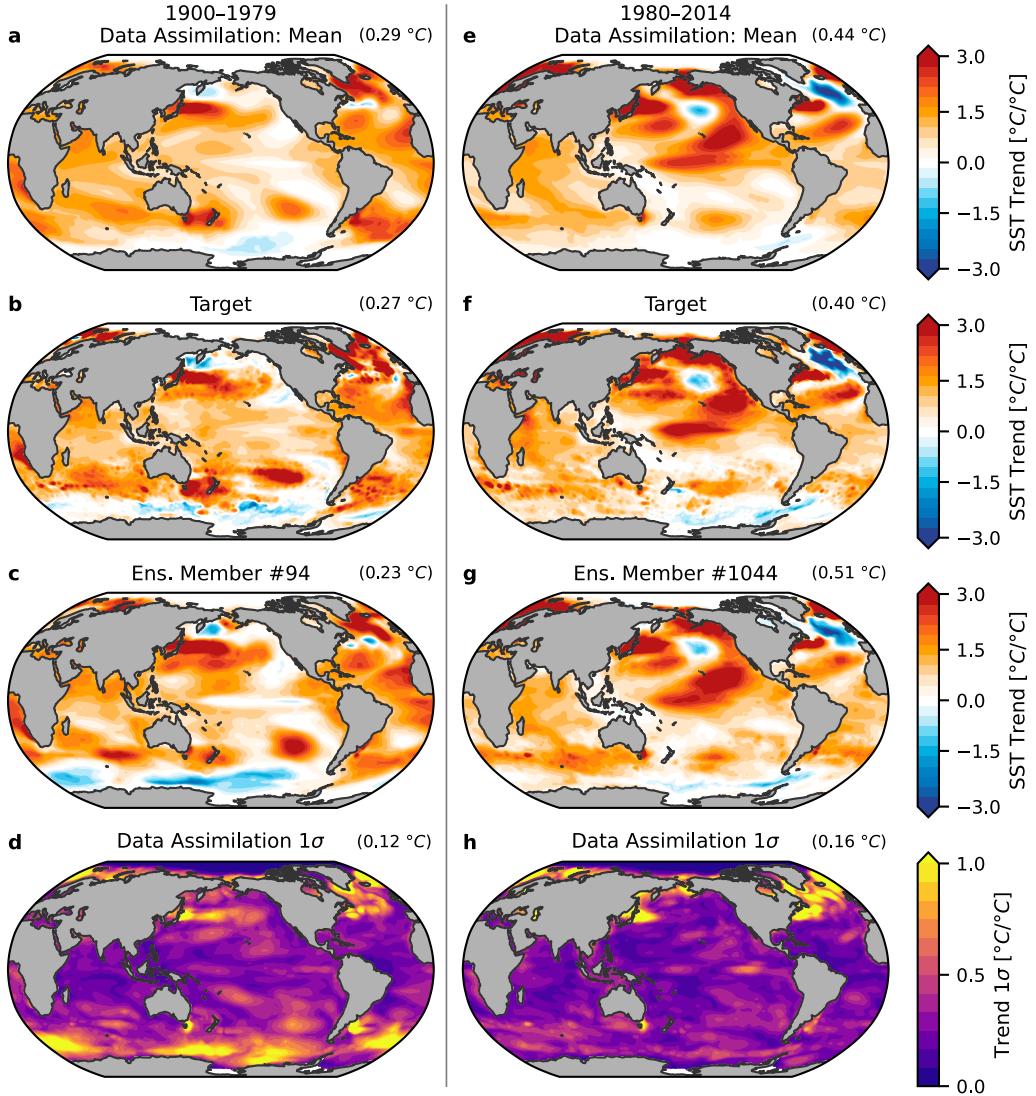
Figure 4.3: Validation by pseudo-reconstruction: SST trends. **(a)** Normalized 1900–1979 ensemble mean of trends in the annual mean from data assimilation; local trends are divided by the global-mean trend to show SST patterns; upper-right indicates the global-mean trend before normalization, scaled by the number of years to show trend in °C per 80 years. **(b)** Repeats panel **a** but showing trends in the pseudo-reconstruction's target model, MPI-ESM1-2-HR's historical simulation. **(c)** Repeats panel **a** but shows an individual member from ensemble data assimilation. **(d)** Uncertainty in results from data assimilation, calculated as the sample standard deviation (1σ) across 1600 ensemble members' normalized trends; values greater than 1.0 indicate that local 1σ is greater than the global-mean trend; upper-right shows the global-mean of the 1σ in local trends before normalization. **(e–f)** Repeats panels **a–d** for 1980–2014.
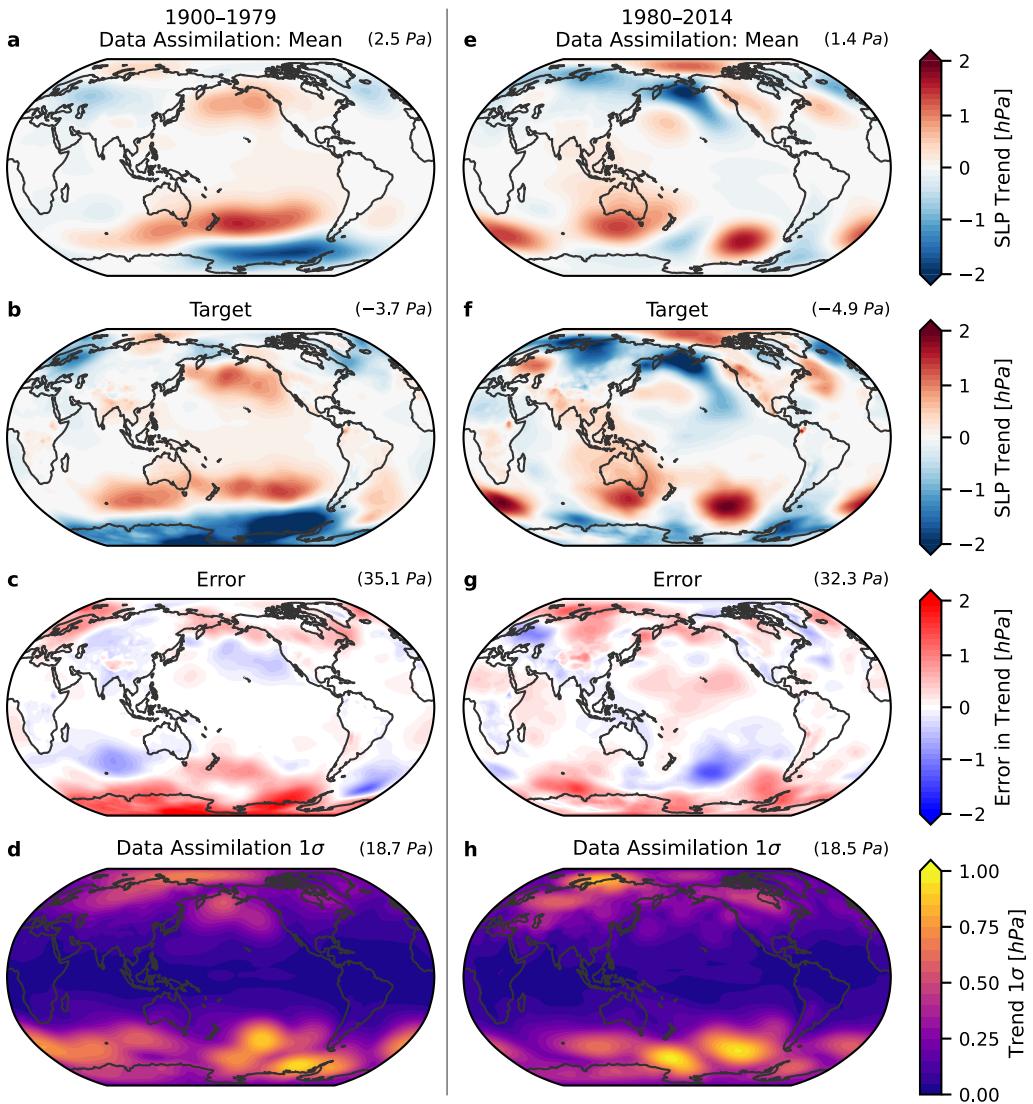
Figure 4.4: Validation by pseudo-reconstruction: trends in sea-level pressure (SLP). **(a)** 1900–1979 ensemble mean of trends in the annual mean from data assimilation, scaled by the number of years to show trends in hPa per 80 years; upper-right indicates the global-mean trend in Pa per 80 years. **(b)** Repeats panel **a** but showing trends in the pseudo-reconstruction's target model, MPI-ESM1-2-HR's historical simulation. **(c)** Error, shown as mean reconstruction minus target; RMSE shown in upper right. **(d)** Uncertainty in results from data assimilation, calculated as the sample standard deviation (1σ) across trends from 1600 ensemble members; upper-right shows the global mean of the 1σ in local trends. **(e–f)** Repeats panels **a–d** for 1980–2014.

over the Southern Ocean. Sparse observations and the unique physics of the target model compared to the forecast models results in considerable uncertainty. The uncertainty indicates that many ensemble members have local trends that differ substantially from the target model, and therefore accurately capturing the trend pattern requires considering the mean across the ensemble.

For additional validation, we show the spatial distribution of correlation and Nash-Sutcliffe efficiency for multiple time periods in Supplemental Figures 4.S1–S2. In Supplemental Figures 4.S3–S4, we also show the correlation and Nash-Sutcliffe efficiency when using only one LIM instead of the multi-model mean of eight LIMs, which illustrates the major improvements from using multiple models in the reconstruction (Amrhein et al., 2020; Parsons et al., 2021). Additionally, we show monthly breakdowns of R^2 and Nash-Sutcliffe efficiency calculated without low-pass filtering the monthly resolved results for each of the metrics in Figure 4.2 (see Supplemental Figure 4.S5). The results show some seasonal variation in skill, which depends on the metric considered.

4.4.5   Observation validation from Desroziers statistics and HadSST4 comparison

For the reconstruction using real observations (Section 4.5), we also evaluate performance using the Desroziers statistics of the DA system (Desroziers et al., 2005) as described in Slivinski et al. (2021),

$$\text{RMSE}_{\text{actual}} = \left(\frac{1}{N_{\text{obs}}} \sum_{j=1}^{N_{\text{obs}}} (\mathbf{y}_j - [\mathbf{H}\mathbf{x}_f]_j)^2 \right)^{1/2}, \tag{4.11}$$

$$\text{RMSE}_{\text{exp}} = \left(\frac{1}{N_{\text{obs}}} \sum_{j=1}^{N_{\text{obs}}} \left(\mathbf{R}_j + [\mathbf{H}\mathbf{P}_f\mathbf{H}^T]_j \right) \right)^{1/2}, \tag{4.12}$$

where j is the observation index for each observation in a given month, $\mathbf{y}_j$ is observation j, $[\mathbf{H}\mathbf{x}_f]_j$ is the forecast prior mean of observation j, and the corresponding $\mathbf{R}_j$ and $[\mathbf{H}\mathbf{P}_f\mathbf{H}^T]_j$ in Equation (4.12) are the observation and forecast errors associated with observation j. $\text{RMSE}_{\text{actual}}$ is related to the innovations from Equation (4.8) and compares the forecasts with observations that have not yet been assimilated, while RMSE_{exp} is related to the innovation covariance in Equation (4.9). The results described below are shown in Supplemental Figure 4.S6 and illustrate the calibration of the DA system.

If the calibration ratio $\text{RMSE}_{\text{actual}}/\text{RMSE}_{\text{exp}} \approx 1$, the system is well calibrated (Slivinski et al.,

2021; Houtekamer and Mitchell, 1998). We group the data into $20°N - 90°N$ (NH), $20°S - 20°N$ (Tropics), and $20°S - 90°S$ (SH), and we compute the calibration ratio using 30-year running means of the RMSE values, then we take the mean of the ratio over 1850–2023. For SST, we find calibration ratios of 1.2 (NH), 1.1 (Tropics), and 1.2 (SH). These ratios are close to 1 and confirm that the DA system is performing well for SST. Calibration ratios for SLP are 0.9 (NH), 0.7 (Tropics), and 0.8 (SH), indicating that the expected errors are larger than the actual errors due to excessive ensemble spread. For T, calibration ratios are 1.1 (NH), 1.6 (Tropics), and 1.3 (SH). Although tropical SST is very well calibrated with a ratio of 1.1, the expected errors are larger than $\text{RMSE}_{\text{actual}}$ for tropical SLP (ratio 0.7) and smaller than $\text{RMSE}_{\text{actual}}$ for tropical air temperatures over land (ratio 1.6). Overall, the Desroziers statistics suggest the DA system is well calibrated, especially for SST.

To illustrate observation validation at specific locations, we show sample timeseries comparing the assimilated HadSST4 observations with the real DA results (from Section 4.5) at seven ocean locations (Supplemental Figures 4.S7–S14). Overall, the results show that errors in the reconstruction relative to the observations are in good agreement with observation error. When outliers appear in the data, their influence is limited by the DA prior and the other observations that are simultaneously assimilated. Supplemental Figures 4.S8–S14 also illustrate the time-varying observation density and uncertainty at various locations, reinforcing the summary calibration results in that the reconstructed ensemble mean and spread are consistent with the assimilated observations and their errors.

4.4.6 Comparison data

We include a variety of datasets for comparison with our reconstruction. For SST, we focus on datasets which are globally complete and have monthly resolution. We include PCMDI/AMIP-II (Hurrell et al., 2008), which was used for CMIP6's AMIP simulations, NOAA ERSSTv5 (Huang et al., 2017), HadISST1 (Rayner et al., 2003), HadISST2.1 (no longer maintained; Titchner and Rayner, 2014), and COBE-SST2 (Hirahara et al., 2014). The statistical infilling in these products is briefly described by Modak and Mauritsen (2023) and Lewis and Mauritsen (2021), with more detail in Kent and Kennedy (2021). All products are regridded to the 2° resolution of our reconstruction.

For SLP, we show gridded reanalyses from ERA5 (1950–present) (Hersbach et al., 2020), NOAA

/ CIRES / DOE 20CRv3 (1836–2015) from Slivinski et al. (2019), and NCEP/NCAR (1948–present) from Kalnay et al. (1996), all regridded to 2° and monthly resolution. We also include an older product, HadSLP2 infilled (Allan and Ansell, 2006). HadSLP2 is no longer maintained, but it provides monthly means of SLP and its non-infilled product would be a companion to HadSST4 if updated. We include an offline-DA reconstruction of the Walker circulation using proxy data, labeled F23 (Falster et al., 2023). We include the SAM from multiple reconstructions using offline DA (O'Connor et al., 2021; Dalaiden et al., 2021; King et al., 2023) and regression (Fogt et al., 2009), labeled as O21, D21, K23, and F09.

For SIC, we show HadISST2.2 (Titchner and Rayner, 2014), HadISST1 (Rayner et al., 2003), and AMIP-II (Hurrell et al., 2008), which is largely based on HadISST1. The satellite record from NOAA/NSIDC CDR (Meier et al., 2021b,a) is shown from 11/1978–12/2023. We include the proxy-based reconstruction of Arctic SIC from Brennan and Hakim (2022), labeled BH22, which has annual rather than monthly resolution. We regrid all SIC data to 2° resolution. When comparing total anomalies in sea-ice area, we restrict the comparison to only include gridcells that have SIC data in every dataset. Otherwise, one dataset may have large anomalies where another dataset has missing values from different land masks, skewing the comparison.

For global-mean T (GMSAT), we compare with HadCRUT5 (Morice et al., 2021) and BEST (Rohde et al., 2013). Note that our reconstruction is of the near-surface air temperature, while the comparison datasets are hybrids of air temperature over land and SST over ocean.

Notably, various datasets can impact one another. The lower boundary condition in ERA5 is the SST from HadISST2 until 2007 and sea ice from HadISST2 until 1979 (Hersbach et al., 2020). NOAA 20CRv3 also uses HadISST2 sea ice over 1836–2015, HadISST2 SST after 1981, and SODAsi.3 SST adjusted to HadISST2 climatology before 1981 (Slivinski et al., 2019; Giese et al., 2016). An SST dataset, ERSSTv5, also uses HadISST2 sea ice to adjust its SST values in the Southern Ocean (Huang et al., 2017). These are examples of how uncertainty in one dataset can affect others.

4.5   Historical Reconstruction

In this section, we share the results of our reconstruction of SST, T, SLP, and SIC from coupled atmosphere–ocean data assimilation with linear inverse models. We show timeseries and spatial

trends of SST, SLP, and SIC, and the El Niño of 1877/1878.

4.5.1   Variability over 1850–2023

Figure 4.5 shows timeseries of the real reconstruction, as for the pseudo-reconstruction (Figure 4.2). The AMV and PDO are similar across datasets for most of the historical record, as described for the PDO in Newman et al. (2016), but PDO uncertainty is notably larger from 1850–1900.

Nino3.4 shows substantial inter-dataset spread before 1875, but the most interesting ENSO feature is the low-frequency evolution of ENSO variance in Figure 4.5d, measured by the 30-year rolling 1σ of Nino3.4. Recent studies have argued for increased ENSO variance with global warming (e.g., Cai et al., 2021, 2023), although other work suggests that ENSO variance could decrease with long-term warming (Callahan et al., 2021), and uncertainties in future ENSO variance have substantial implications for global-scale climate predictability (Amaya et al., 2025). In our results, ENSO variance was at local maximum over 1875–1900, decreased to a local minimum over 1930–1960, and subsequently trended higher to the present. Overall, Figure 4.5d suggests considerable centennial-scale power in ENSO variance.

Tropical SST gradients are diagnosed using two measures. The Pacific zonal SST gradient (Figure 4.5e), shows that the magnitude of the strengthening trend from 1980–2023 is not clearly distinguishable from past variability, such as the weakening from 1875–1905. The long-term strengthening trend since 1900 has also been a focus of many studies (e.g., Cane et al., 1997; Karnauskas et al., 2009; Deser et al., 2010b; Solomon and Newman, 2012; Coats and Karnauskas, 2017; Seager et al., 2022; Lee et al., 2022), but 1900–1905 has the weakest zonal gradient during the historical record, and the gradient in 1890 is comparable to 2023. However, the $SST^{\#}$ metric (Figure 4.5f), representing the contrast between the warmest tropical SSTs and the tropical mean (Fueglistaler, 2019; Fueglistaler and Silvers, 2021), shows a consistent strengthening from 1975–present. The persistent 1975–2023 trend in $SST^{\#}$ may indeed be distinct compared to the variability before 1975, but further investigation is needed.

The Pacific Walker circulation (zonal SLP gradient) appears to be dominated by stationary decadal variability over the full historical record (Figure 4.5i). Our reconstruction does not show a trend toward weakening of the Walker circulation over the 20th century (Vecchi et al., 2006; Tokinaga
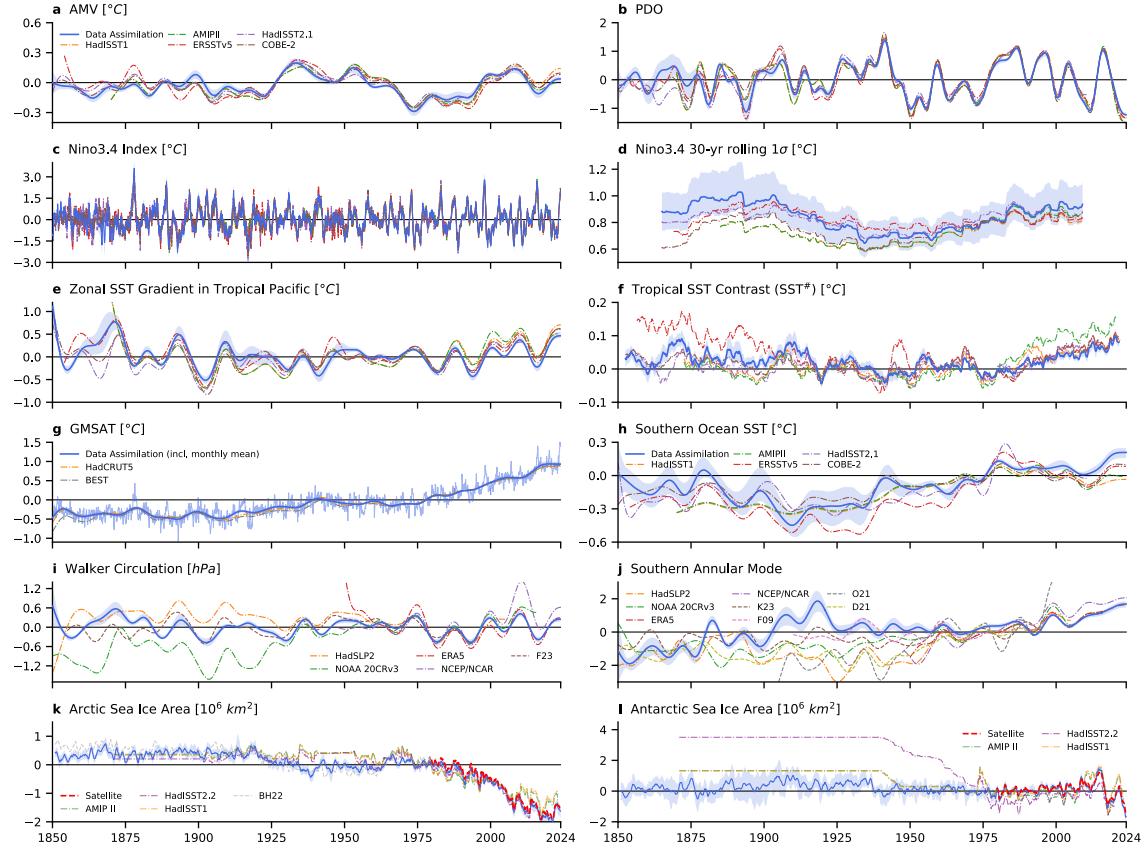
Figure 4.5: Climate variability over 1850–2023. (Blue) Data assimilation results, showing mean of 1600 ensemble members; shading denotes 17th and 83rd percentiles. Note that panel **a**'s legend applies to SST datasets in panels **a–f**; re-used line colors in SLP, T, and SIC panels do not necessarily indicate consistency with SST datasets. **(a)** Atlantic Multidecadal Variability (SST) with 10-yr low-pass filter. **(b)** Pacific Decadal Oscillation (SST) with 6-yr low-pass filter. **(c)** Monthly SST in Nino3.4 region with 30-yr running mean removed. **(d)** Rolling 30-yr standard deviation of Nino3.4 in panel **c**. **(e)** Zonal gradient of tropical Pacific SST with 10-yr low-pass filter. **(f)** Tropical SST contrast, SST$^{\#}$, 5-yr running mean. **(g)** Global-mean near-surface air temperature (GMSAT) with 10-yr low-pass filter and monthly values from data assimilation as thin line. **(h)** Zonal mean of Southern Ocean SST (50°–70°S) with 10-yr low-pass filter. **(i)** Walker circulation, i.e., zonal SLP gradient across tropical Pacific, with 10-yr low-pass filter. **(j)** Southern Annular Mode (SLP) with 10-yr low-pass filter. **(k)** Total Arctic and **(l)** Antarctic sea ice area, showing 12-month running mean, and comparison satellite data from NOAA/NSIDC CDR. Calculation of metrics is described in Section 4.4.4; comparison data is summarized in Section 4.4.6.

et al., 2012), and the strengthening from c. 1979–2014 (e.g., Chung et al., 2019; L'Heureux et al., 2013; Watanabe et al., 2023, 2024) appears within the range of variability prior to 1975. Heede and Fedorov (2023) found large recent changes in the zonal SLP gradient in the NCEP/NCAR Reanalysis, but that product may be an outlier over 2005–2020 (Figure 4.5i).

Our reconstruction of the Southern Annular Mode (SAM) has relatively small ensemble spread relative to the spread across other products (Figure 4.5j). Notably, the pre-1980 disagreement across reanalyses and other reconstructions is larger than the decadal variability in any one product. Spurious trends in Southern Hemisphere SLP have been identified in reanalyses poleward of 60°S during the early twentieth century and c. 1950 due to the general paucity of data over much of the Southern Hemisphere (Schneider and Fogt, 2018; Fogt and Connolly, 2021; Laloyaux et al., 2018). Local observations in the SAM region are sparse throughout most of the historical record. Consequently, our SAM reconstruction is primarily constrained by remote observations of SLP, SST, and T, with the dynamics of the LIMs acting to connect those remote observations to the SAM region's SLP. Many studies have highlighted the positive trend in the SAM from c. 1980–present (e.g., Thompson and Solomon, 2002; Marshall, 2003; Polvani et al., 2011; Swart et al., 2015; Banerjee et al., 2020; Fogt and Marshall, 2020), but some datasets in Figure 4.5j show longer-term positive trends, possibly spanning the entire 20th century (O'Connor et al., 2021; Dalaiden et al., 2021; Slivinski et al., 2019; Allan and Ansell, 2006). Our results indicate that the recent trend only extends from approximately 1970–present, and the trends are most notable in DJF (Supplemental Figure 4.S16). There appears to be another prolonged positive trend from 1850–1920 in our reconstruction but not in any of the comparison data, and that the SAM trend aligns with SST cooling in the Southern Ocean over the same period. Brönnimann et al. (2024) analyzed newly digitized ship records from 1903–1916 and also find a positive SAM index and pronounced surface cooling over the Southern Ocean during the early 1900s.

Sea ice from data assimilation (Figure 4.5k,l) exhibits major differences compared to the HadISST and AMIP-II datasets, which have been used to assess the atmospheric response to SIC changes over the historical record. Over much of the historical record, these datasets have constant values at inferred climatologies. There are also differences in the satellite era due to uncertainties in data processing and discontinuities in satellite sources (e.g., Eisenman et al., 2014; Buckley et al., 2024), which are responsible for the spurious high values in Antarctic sea ice from 2009–2011 in HadISST1

and AMIP-II (Screen, 2011), evident in Figure 4.5l.

For Arctic sea ice, the main difference across datasets relates to the early 20th-century warming (Brönnimann, 2009; Hegerl et al., 2018). HadISST1 and AMIP-II do not have any signal of the early 20th-century warming in sea-ice area. Our reconstruction shows a loss of 0.5 ± 0.1 (1σ) million km^2 during the 1920s, measured by comparing the decadal means of the 1930s and 1910s. Note that this value should not be compared directly with other datasets unless land masks are consistently applied. The Brennan and Hakim (2022) reconstruction of annual means, using only proxy data with offline DA, agrees closely with our results.

Antarctic sea ice is a unique result compared to existing estimates. In stark contrast to the datasets used for CMIP6/DECK/AMIP/CFMIP (Eyring et al., 2016; Webb et al., 2017) and as boundary conditions in reanalyses (e.g., Slivinski et al., 2019; Hersbach et al., 2020), our reconstruction shows much less ice loss from the preindustrial to present conditions. AMIP-II, HadISST1, and HadISST2 are at the edge or outside of our likely range for the entire pre-1980 period. Note that HadISST2 is the ice boundary condition in ERA5 and NOAA 20CRv3 before 1979, and it is used to adjust SST in NOAA ERSSTv5.

In the early 20th century, we find a wide envelope of uncertainty in Antarctic ice area that spans the range over the satellite record until 2022. Our results show a local maximum c. 1910, consistent with the SH cooling reported by Brönnimann et al. (2024). We find greater Antarctic ice cover in the early 1960s compared to the 1980s (Fan et al., 2014), consistent with Goosse et al. (2024). However, our reconstruction shows a decrease throughout the 1970s (Supplemental Figure 4.S15) in contrast to the sharp drop in ice extent at the end of the 1970s reported by Goosse et al. (2024). Early single-channel satellite retrievals from ESMR suggest Antarctic ice cover may have been more extensive in the 1970s (Goosse et al., 2024; Kolbe et al., 2024), though the reliability of ESMR is debated (Titchner and Rayner, 2014; Kolbe et al., 2024). As evident in the ensemble spreads (Fig. 4.5l; Supplemental Figure 4.S15), the uncertainty before 1980 is substantial, and more work is needed to constrain Antarctic SIC. The preindustrial-mean ice area (1850–1900) does not appear clearly different from the satellite-era range until the ice loss of 2022–2023 (Roach and Meier, 2024; Espinosa et al., 2024; Zhang and Li, 2023; Fogt et al., 2022; Turner et al., 2022). Our results for preindustrial ice area are consistent with Edinburgh and Day (2016)'s analysis of ship records from the Heroic Age (1897–1917), who found ice expansion in the Weddell Sea but

comparable conditions to 1989–2014 in the other sectors.

Finally, we consider variability in Southern Ocean SST (zonal mean 50°–70°S). We find a large spread in our ensemble before 1950 and a larger disagreement across SST datasets, which persists from 1850 to 2023. We note two interesting results in Figure 4.5h. First, we find a long-term warming trend from 1910–2023, which is approximately half as large as the 1910–present warming trend in GMSAT. This is consistent with expectations, since Southern Ocean warming is muted by upwelling of deep water that has not yet experienced the global warming signal (Armour et al., 2016).

Second, we find a muted cooling of the Southern Ocean from 1980–2013, and slight warming from 1980–2023. The comparison datasets show 1980–2013 cooling that is mostly outside of our likely range. In situ observations are still sparse from 1980–2023 (Figure 4.1; Supplemental Figure 4.S14) and the data sources change dramatically over that period, possibly introducing spurious trends from homogenizing different data sources (Kennedy et al., 2019; Huang et al., 2019; Kent and Kennedy, 2021; Hausfather et al., 2017; Karl et al., 2015). We elaborate on Southern Ocean trends below and in the Discussion section.

The Southern Ocean cooling over recent decades is not unprecedented given that we find stronger cooling from 1880–1910. Brönnimann et al. (2024) report that this cooling is a real climatic phenomenon, not a data artifact. However, Sippel et al. (2024) suggest that biases in the bucket measurements of SST are responsible for a cold bias from 1910–1930. If SST-bucket biases are indeed responsible for this cooling trend, an explanation is still required for why the night-time marine air temperatures (Cornes et al., 2020) also show this 1910–1930 cooling trajectory (Figure 1 of Sippel et al., 2024).

4.5.2 Patterns of SST, SLP, and SIC trends

Figure 4.6 shows spatial patterns of SST trends separately for the gradual warming from 1900–1979 and the recent period of 1980–2023. We show our reconstruction and its uncertainty alongside comparison trends from NOAA ERSSTv5 and COBE-SST2. Despite similar global-mean trends from 1900–1979, there are substantial disagreements in the pattern of trends especially over the Southern Ocean and tropical Pacific. The post-1980 period is often viewed as having small uncertainty due
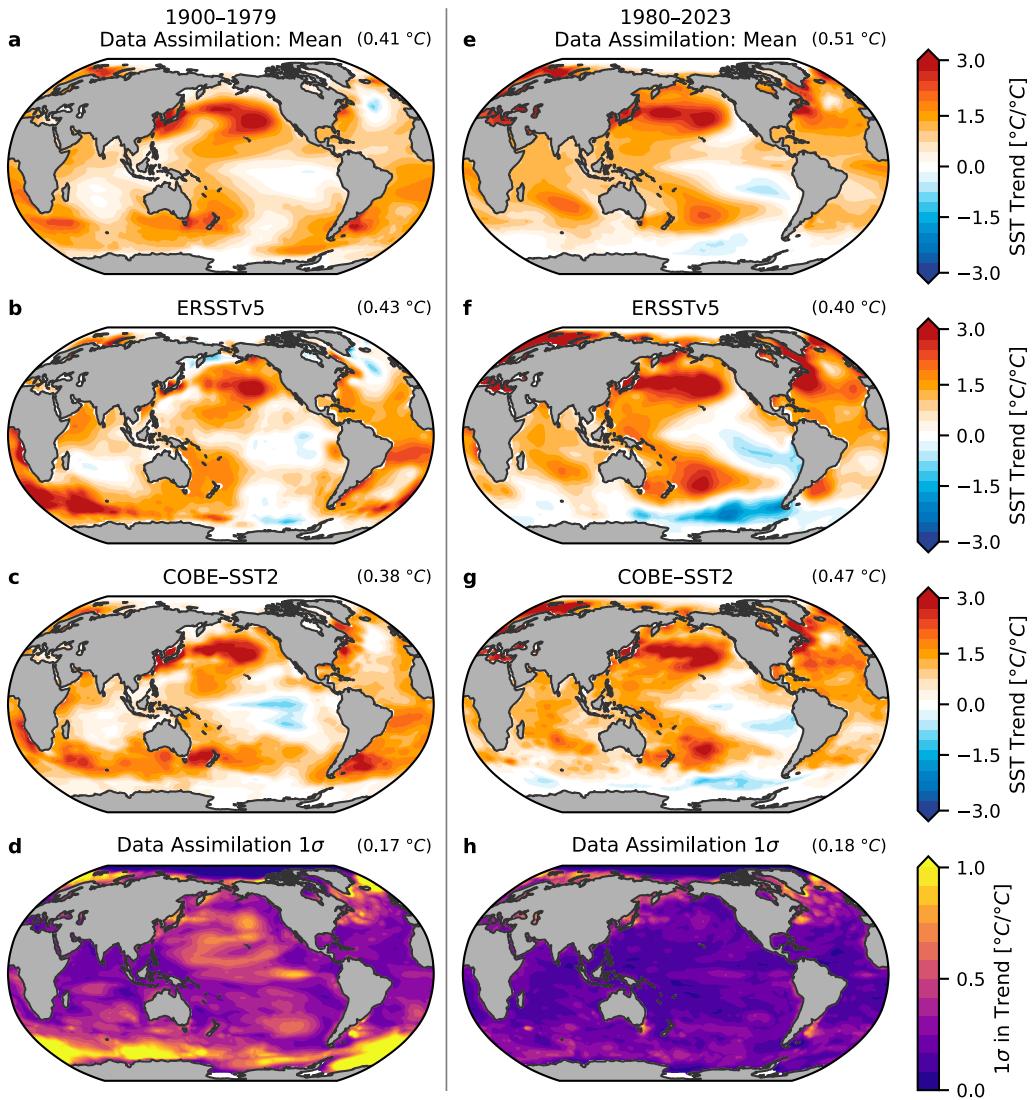
Figure 4.6: Historical patterns of SST trends. **(a)** Normalized 1900–1979 ensemble mean of trends in the annual mean from data assimilation; local trends are divided by the global-mean trend to show SST patterns; upper-right value is the global-mean trend before normalization, scaled by the number of years to show trend in °C per 80 years. **(b)** Repeats panel **a** but showing comparison data from NOAA ERSSTv5 and **(c)** COBE-SST2. **(d)** Uncertainty in results from data assimilation, calculated as the sample standard deviation (1σ) across 1600 ensemble members' normalized trends; values greater than 1.0 indicate that local 1σ is greater than the global-mean trend; upper-right value is the global-mean of the 1σ in local trends before normalization. **(e–f)** Repeats panels **a–d** for 1980–2023.
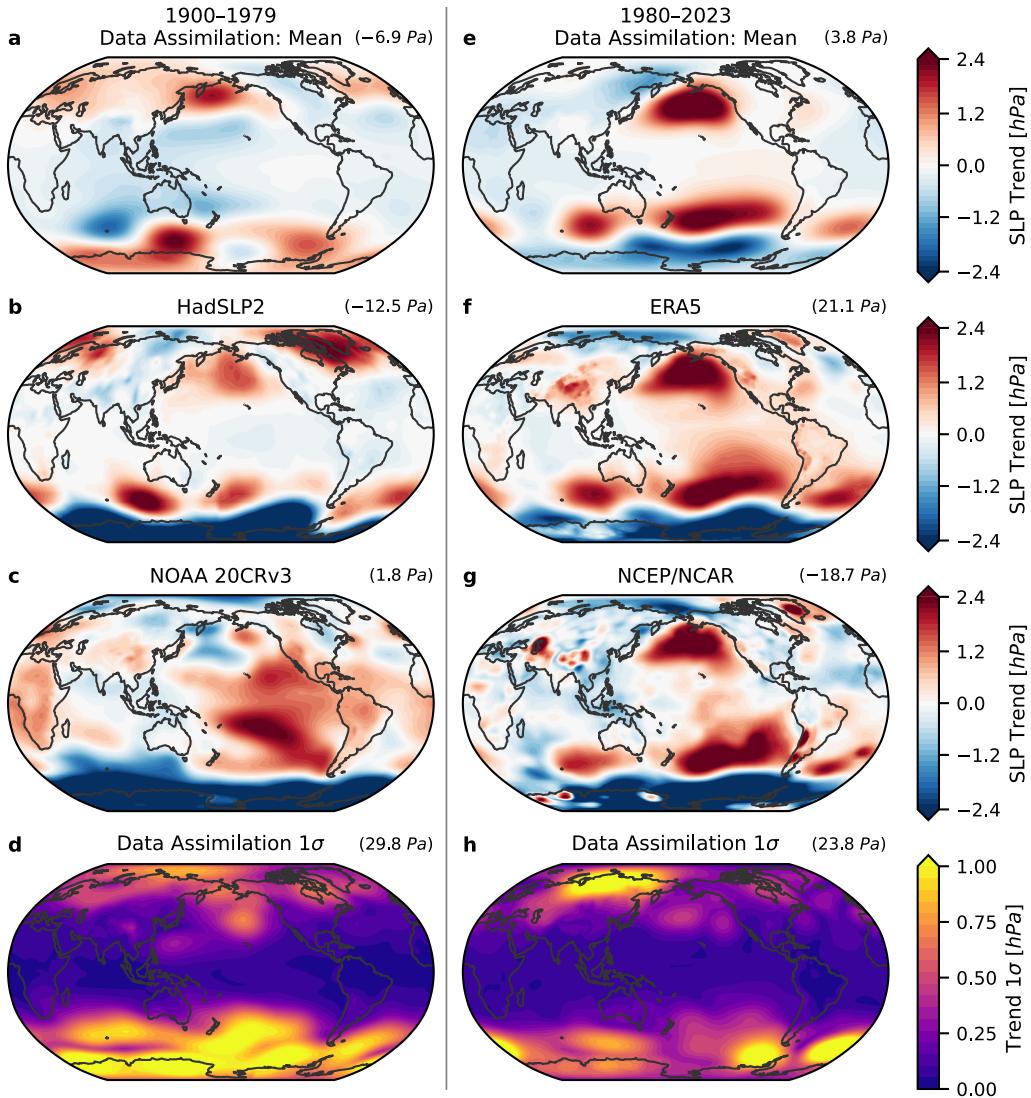
Figure 4.7: Historical trends in sea-level pressure (SLP). **(a)** 1900–1979 ensemble mean of trends in the annual mean from data assimilation, scaled by the number of years to show trends in hPa per 80 years; upper-right indicates the global-mean trend in Pa per 80 years. **(b)** Repeats panel **a** but showing comparison datasets HadSLP2 and **(c)** NOAA 20CRv3. **(d)** Uncertainty in results from data assimilation, calculated as the sample standard deviation (1σ) across local trends from 1600 ensemble members; upper-right shows the global mean of the 1σ in local trends. **(e–f)** Repeats panels **a–d** for 1980–2023, with comparison reanalyses from **(f)** ERA5 and **(g)** NCEP/NCAR.

to observation density (Figure 4.1), but the inter-dataset disagreements in Figure 4.6e–g suggest there are nontrivial uncertainties in large-scale SST gradients. The southeast Pacific and Southern Ocean regions, which have strong impacts on global climate variability and radiative feedbacks (e.g., Dong et al., 2022a; Kang et al., 2023c,a; Espinosa and Zelinka, 2024), have the worst observation coverage (Figure 4.1).

Figure 4.7 shows spatial patterns of SLP trends for 1900–1979 and 1980–2023 from our reconstruction and comparison datasets. Note that our reconstruction only assimilates marine SLP observations, so we expect it to differ from reanalyses over land regions. From 1900–1979, there are many large-scale differences between our reconstruction, HadSLP2, and NOAA 20CRv3. The comparison datasets show strong negative trends in SLP over Antarctica and most of the Southern Ocean during both time periods, whereas we find positive trends over 1900-1979. In this region, regression-based reconstructions find positive trends in the early 20th century, also in contrast to the negative trends in existing reanalyses (Fogt et al., 2019; Fogt and Connolly, 2021; Fogt et al., 2024). Schneider and Fogt (2018) and Laloyaux et al. (2018) highlight problems with the atmospheric circulation in the Southern Hemisphere in multiple reanalyses and how those problems create spurious climate signals. The key problem identified in ERA-20C is that the assumed error is too small for pressure observations. This is one reason why we ensure our SLP observation error is not too small, as described in Appendix B.

Over 1980–2023, our SLP trends over the global oceans largely align with ERA5, albeit with weaker positive trends in the central and eastern Pacific (Figure 4.7e,f). ERA5 has a substantial trend in global-mean SLP, which increases by 21.1 Pa $(44 \text{ yr})^{-1}$ from 1980–2023, and removing this trend would improve agreement with our reconstruction in many regions. NCEP/NCAR has a substantial and opposite trend of -18.7 Pa $(44 \text{ yr})^{-1}$. Our reconstruction has a much smaller 1980–2023 trend in global-mean SLP of 3.8 Pa $(44 \text{ yr})^{-1}$ (Figure 4.7e) and similarly small trends from 1900–1979 and also in the pseudo-reconstruction experiment (Figure 4.4). Once again, our reconstruction highlights uncertainty over the Southern Ocean, especially the Amundsen Sea Low and the Atlantic sector.

Figure 4.8 shows trends in Arctic SIC over 1900–1979, during the early 20th-century warming from 1920–1935, and for the recent loss from 1980–2023. We compare with HadISST2, which is the pre-satellite boundary condition used in ERA5 and NOAA 20CRv3, and with the NOAA/NSIDC
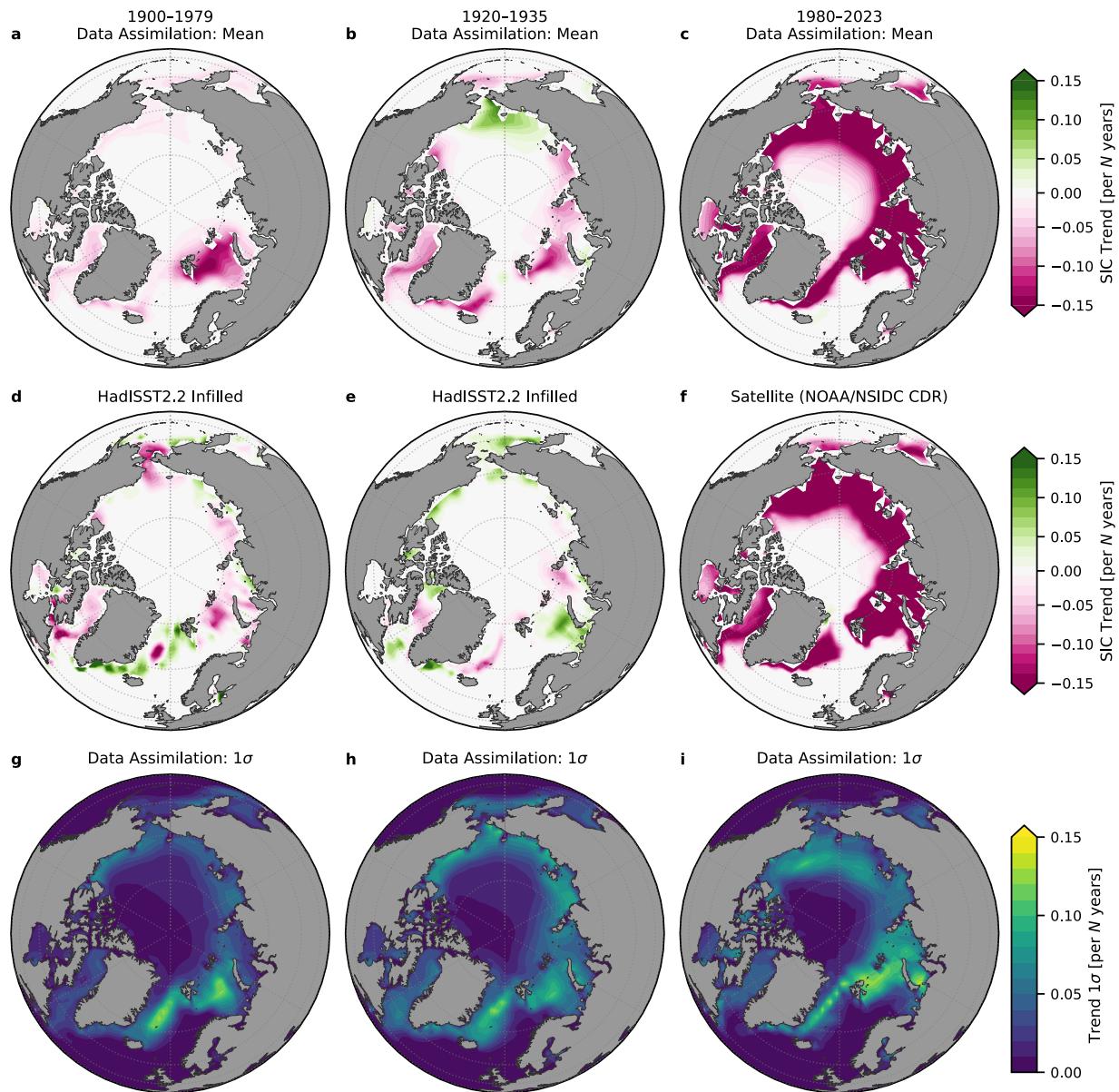
Figure 4.8: Historical trends in Arctic sea-ice concentration (SIC). **(a–c)** Ensemble mean of trends from data assimilation, scaled by the number of years in each period to show trends in SIC per N years. **(d–f)** Repeats panels **a–c** but showing comparison datasets, with infilled HadISST2.2 in panels **d–e** and satellite data from NOAA/NSIDC CDR in panel **f**. **(g–i)** Uncertainty in results from data assimilation, calculated as local standard deviation (1σ) across 1600 ensemble members, corresponding to time periods in panels **a–c**. Note that SIC is bounded from 0 to 1.
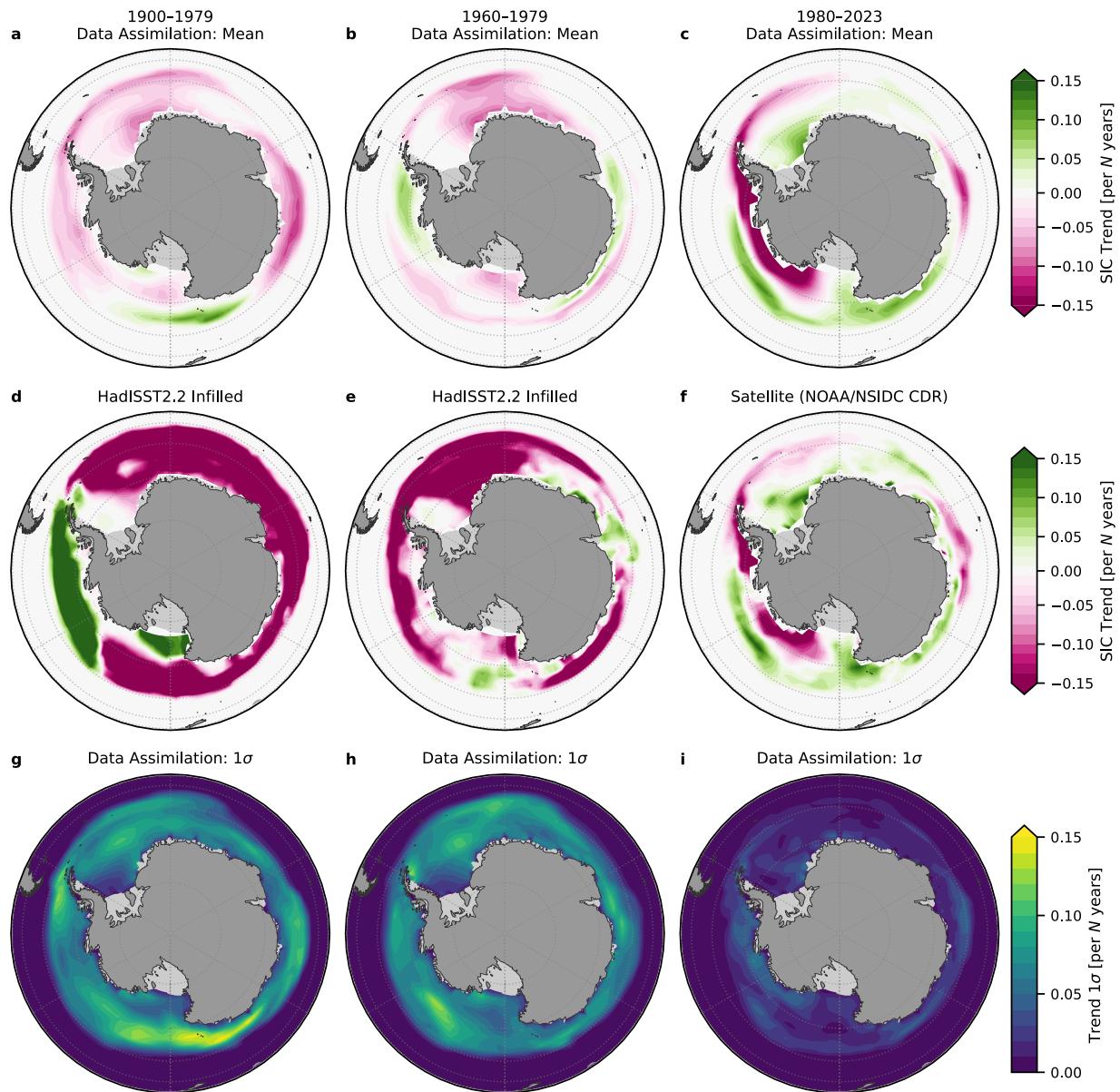
Figure 4.9: Historical trends in Antarctic sea-ice concentration (SIC). **(a–c)** Ensemble mean of trends from data assimilation, scaled by the number of years in each period to show trends in SIC per N years. **(d–f)** Repeats panels **a–c** but showing comparison datasets, with infilled HadISST2.2 in panels **d–e** and satellite data from NOAA/NSIDC CDR in panel **f**. **(g–i)** Uncertainty in results from data assimilation, calculated as local standard deviation (1σ) across 1600 ensemble members, corresponding to time periods in panels **a–c**. Note that SIC is bounded from 0 to 1.

satellite data that we assimilate. From 1900–1979, we find ice loss in the Barents Sea between Svalbard and Russia. From 1920–1935, we find ice loss around most of the Arctic, partially offset by gains poleward of the Bering Strait. HadISST2 does not have this 1920–1935 ice loss. From 1980–2023, our ice loss looks very similar to the satellite record, but it does not match exactly because of uncertainty in the satellite data, the influence of non-SIC observations, and the particularities of our LIM and DA methods.

Figure 4.9 shows trends in Antarctic SIC from 1900–1979, during the 1960–1979 period of ice loss hypothesized by Fan et al. (2014), and from 1980–2023, a period with steady but small growth and then recent rapid loss (e.g., Stuecker et al., 2017). Our reconstruction of 1900–1979 shows some ice loss alongside the Southern Ocean SST warming, but we find a lesser magnitude and a different pattern compared to HadISST2. If sea ice has a relationship with the atmospheric circulation (e.g., Kohyama and Hartmann, 2016), the HadISST2 boundary condition may impact the circulation in ERA5 and NOAA 20CRv3. From 1960–1979, we find ice loss in the Atlantic sector, which mostly aligns with the pattern in HadISST2 but with a substantially different magnitude. We see a minor gain of ice in the Bellingshausen Sea, where HadISST2 shows large loss.

4.5.3   El Niño in 1877

The extreme El Niño that began in 1877, which is the largest event in the historical record, is an instructive comparison case for infilled datasets. Observations are sparse but the signal is large. Recent reconstructions of hybrid air/sea-surface temperature also focused on this event (Vaccaro et al., 2021; Kadow et al., 2020) to illustrate how different the imputed values can be for different datasets.

Figure 4.10 shows the onset of El Niño in July 1877. We show the ensemble spread in our reconstructed SST and land T, the observations of SST and station temperatures, and two comparison datasets. ERSSTv5 depicts the center of action in the coastal-eastern Pacific, whereas the central Pacific is most notable in HadISST1. Our ensemble mean displays some commonalities with each dataset, but we find higher confidence in the east Pacific El Niño (Figure 4.10b), and we note that our method leverages teleconnections with observations of SLP and land temperatures to constrain the SSTs. Our results also show large uncertainties in the central Pacific and the coastal-eastern
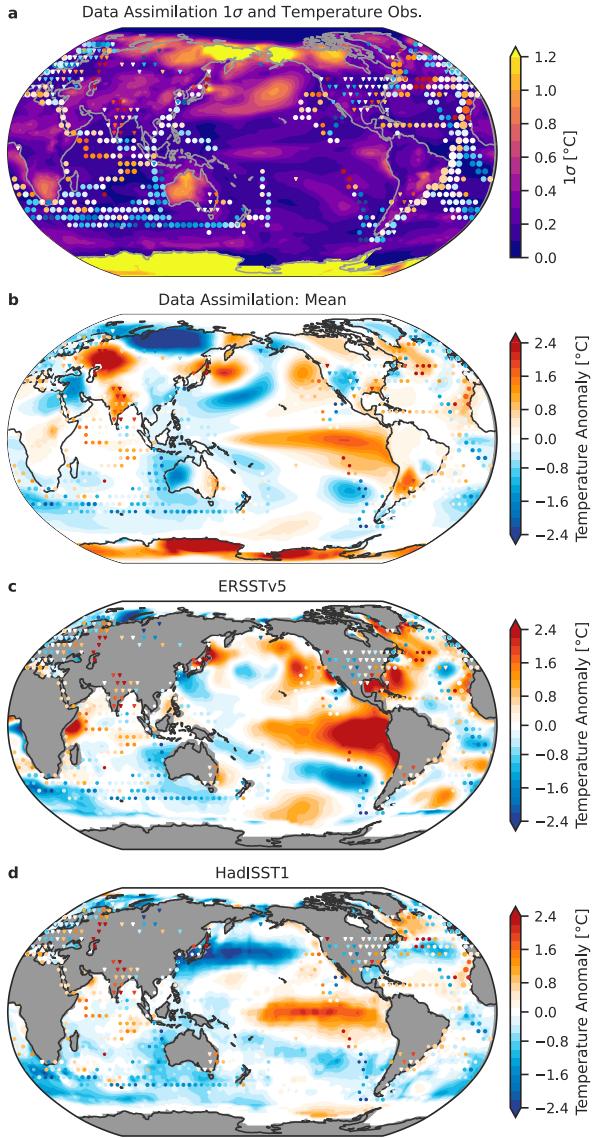
Figure 4.10: El Niño in July 1877: reconstruction, observations, and uncertainty. **(a)** Contours show uncertainty in the data assimilation, calculated as the sample standard deviation (1σ) across the local anomalies in SST and near-surface air temperatures (T) over land for the 1600 ensemble members; scattered dots show anomalies in SST from HadSST4, with size inversely proportional to error, while triangles show land T from CRUTEM5; T and SST points use colorbar from panels **b–d**. **(b)** Contours show ensemble mean of SST anomalies and land T from data assimilation, with HadSST4 and CRUTEM5 observations. **(c–d)** Repeats panel **b** but with comparison SST datasets, NOAA ERSSTv5 and HadISST1.

Pacific (Figure 4.10a); i.e., uncertainty in the type of ENSO (e.g., Newman et al., 2011; Karnauskas, 2013; Capotondi et al., 2015). There are also large differences across datasets in the North Pacific. Around the Southern Ocean in ERSSTv5, the influence of the HadISST2 sea ice is evident as a ring of cold anomalies. This results from the expansion of Antarctic sea ice in HadISST2 (Figure 4.5l).

4.6 Discussion

With coupled DA, we provide a dynamically and observationally constrained perspective on coupled variability and trends over the historical record. These results suggest it may be worth revisiting assessments of forced versus internal variability and climate-model biases using this internally consistent reconstruction. Many studies have characterized post-1980 trends, but placing those changes in the context of the longer record may help disentangle the mechanisms and causes of both variability and trends. Several large-scale model biases, including those in the Southern Ocean and the Tropics, now appear less drastic than previously estimated, suggesting climate models may perform better than indicated by comparison with earlier datasets (e.g., Wills et al., 2022; Simpson et al., 2025).

4.6.1 Tropical trends

The zonal SST gradient and Walker circulation in the tropical Pacific has been a focus of many discussions of forced versus internal variability (Vecchi and Soden, 2007; DiNezio et al., 2009; Coats and Karnauskas, 2017; Kohyama et al., 2017; Seager et al., 2019; Lee et al., 2022; Kang et al., 2023b; Watanabe et al., 2024; Jiang et al., 2024). In our results, the 1979–2014 strengthening trend in the Walker circulation (Pacific zonal SLP gradient) does not appear distinct from variability over the historical record. The Pacific zonal SST gradient has a more notable trend from 1980–present, but it is difficult to convincingly say that the trend is outside of the range of natural variability.

Over the full twentieth century, we do not find a long-term weakening of the Walker circulation (Tokinaga et al., 2012; Vecchi et al., 2006) nor a clear strengthening of the zonal SST gradient (Coats and Karnauskas, 2017; Seager et al., 2022) that is distinct from past variability. If the recent trend is a forced response to global warming from CO_2 (e.g, Clement et al., 1996; Seager et al., 2019), that trend is not yet distinct from past variability in the reconstruction.

However, our results indicate that there is a peculiar trend from c. 1975–present in the strengthening of the SST contrast between the warmest SSTs and the mean SST over the entire Tropics ($SST^{\#}$; Fueglistaler and Silvers, 2021; Fueglistaler, 2019). Fueglistaler and Silvers (2021) questioned whether the recent trend in $SST^{\#}$ could be due to data artifacts in the SST record or purely coincidence, i.e., a rare occurrence of variability during the satellite record. Data artifacts are still a possible influence, but the dynamical constraints in our method reduce the likelihood of that explanation, especially considering that the SST is also informed by SLP observations and station temperatures. Further analysis of paleoclimate proxy data in the Tropics (e.g., Deutsch et al., 2014; Sanchez et al., 2020, 2021) could help assess the role of possible data artifacts and the range of natural variability in SST contrasts.

4.6.2 Southern annular mode

The positive trend in the SAM (c. 1980–present) has been associated with stratospheric ozone depletion, CO_2 forcing, natural variability, and other factors (Doddridge and Marshall, 2017; Polvani et al., 2021; Bitz and Polvani, 2012; Seviour et al., 2016; Thomas et al., 2015; Thompson et al., 2011; England et al., 2016; Fogt and Marshall, 2020; Banerjee et al., 2020). Efforts to determine what has caused the SAM trend have been complicated by recent results, included in Figure 4.5j, depicting a positive trend over the entire twentieth century (O'Connor et al., 2021; Dalaiden et al., 2021; King et al., 2023; Slivinski et al., 2019). Our findings, which show no trend from 1925–1970, then a prolonged positive trend from 1970–present, are consistent with a trend onset that is associated with stratospheric ozone depletion (Thompson and Solomon, 2002; Fogt et al., 2009; Polvani et al., 2011; Thompson et al., 2011). The regression-based reconstruction of F09 is in general agreement with our results over 1920–1970, showing no significant SAM trend until a positive trend emerges in DJF around 1970 (Fogt et al., 2009). We also find that DJF has the strongest SAM trend over 1970–present (Supplemental Figure 4.S16). Another large positive trend over 1850–1920 warrants further investigation into possible drivers and the role of data quality, particularly given the sparse and imperfect SLP observations in the early record. We note that Brönnimann et al. (2024) also report positive SAM in the early 1900s in newly digitized ship data, supporting our results.

4.6.3   Southern Ocean cooling

Studies of the post-1980 cooling in the Southern Ocean typically use SSTs from NOAA ERSST, the latest of which is Version 5 (Huang et al., 2017). Even when nudging a climate model (CESM1) to ERA reanalysis winds, the model does not reproduce the Southern Ocean SST cooling from ERSST (Blanchard-Wrigglesworth et al., 2021; Dong et al., 2022a). Therefore, it seems that the winds alone cannot explain the SST cooling over the Southern Ocean (Dong et al., 2023), and other explanations have been proposed (e.g., Zhang et al., 2019; Haumann et al., 2020; Dong et al., 2022b; Swart et al., 2023; Schmidt et al., 2023).

Pacemaker experiments, which nudge a coupled climate model's SST in the Southern Ocean to match an infilled SST dataset (typically NOAA ERSST), have been used to investigate how SST cooling of the Southern Ocean affects global climate, radiative feedbacks, and the atmospheric circulation (Zhang et al., 2021; Kang et al., 2023c,a). The Southern Ocean cooling has also been proposed as a driver of cooling in the tropical east Pacific (Dong et al., 2022a), possibly forced by the ozone hole (Hartmann, 2022) or other means (Watanabe et al., 2024). Kang et al. (2024) leverage the pacemaker experiments, but they also highlight the importance of regional-scale discrepancies in SST trends for the atmospheric circulation and uncertainty in post-1979 trends across reanalyses in the Southern Hemisphere.

In our results, we find much less cooling over the Southern Ocean compared to NOAA ERSSTv5. While more work is needed before definitive conclusions can be made about which reconstruction is more accurate, we compare the non-infilled SST dataset that we use to inform our data assimilation, HadSST4, with the non-infilled SST data from ERSSTv5 and from a recent product that has undergone extensive bias corrections (DCENT, Chan et al., 2024). Then we also compare our results with 1980–2023 trends in other infilled SST datasets.

Figure 4.11a compares the non-infilled anomalies in the southeast-Pacific sector of the Southern Ocean (latitudes 50°S–70°S and longitudes 70°W–140°W). HadSST4 and DCENT show similar trajectories, but they have a substantial offsets relative to ERSSTv5. This suggests that not only the infilling but also the homogenization of time-varying data sources affects trends in this region. Kennedy et al. (2019) show the major transition from bucket measurements to drifting buoys between 1980 and 2005, and Huang et al. (2019) find substantial differences in SST analyses from
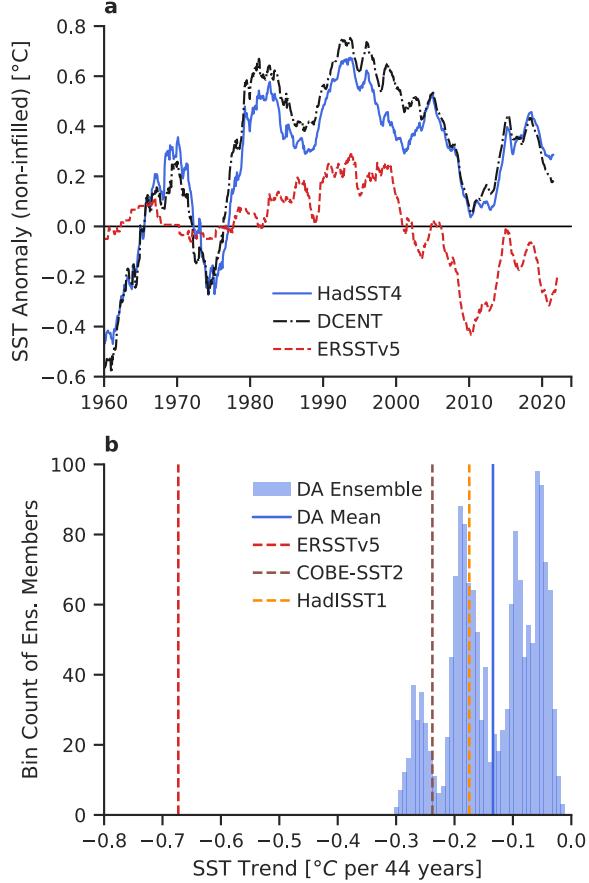
Figure 4.11: Recent evolution of Southern Ocean SST: comparing in situ data and infilled trends in the southeast-Pacific sector. Both panels analyze the mean of values in the region west of the Drake Passage, spanning latitudes 50°S to 70°S and longitudes 70°W to 140°W. **(a)** Comparison of non-infilled SST anomalies, illustrating differences from the homogenization of time-varying in situ sources; for visual clarity, 5-yr running mean is applied and the 1961–1979 mean is removed. **(b)** Infilled SST trends for 1980–2023 from data assimilation (DA), with 1600 ensemble members shown as histogram; the distribution is shaped by the eight distinct model priors. Vertical lines indicate the mean trend and comparison datasets.

2000–2016 when including drifting buoys and/or ARGO floats in NOAA ERSSTv5. ERSSTv5 has a detailed bias-correction procedure and consequently could provide the best estimate in this region. A key point is that the processing of time-varying data sources could have a spurious influence on what appear to be climate trends.

Figure 4.11b shows the distribution of 1980–2023 SST trends in the southeast-Pacific sector of the Southern Ocean (latitudes 50°S–70°S and longitudes 70°W–140°W). Our reconstruction shows a wide range of uncertainty, with possible trends ranging from $-0.3°C$ to $0.0°C$ $(44 \text{ yr})^{-1}$. Our distribution is shaped by the uncertainty in bias corrections from HadSST4 and by the eight LIMs used as priors in the assimilation. COBE-SST2 and HadISST1 are within our uncertainty range, but ERSSTv5 has a much larger trend of $-0.7°C$ $(44 \text{ yr})^{-1}$. Determining which of these trends is correct may be important for advancing understanding of the mechanisms driving Southern Ocean cooling. For example, nudging a climate model's winds to reanalysis may not explain the magnitude of cooling in ERSSTv5, but wind-nudging might be sufficient to explain all of the cooling in our reconstruction. Thus the result of weaker cooling in our reconstruction supports the notion that ozone depletion, through its influence on the SAM and surface winds, may be a key driver of the observed SST trends (Hartmann, 2022). The possibility that our reconstruction is closer to the true (but unknown) trend motivates revisiting investigations of Southern Ocean cooling, as well as its impacts on the tropical Pacific and global climate (e.g., Kang et al., 2023a,c).

4.6.4 *Radiative feedbacks and historical pattern effects*

The pattern effect on climate sensitivity, i.e., the dependence of radiative feedbacks on spatial patterns of SST and SIC anomalies (Armour et al., 2013; Andrews et al., 2015; Zhou et al., 2016; Ceppi and Gregory, 2017; Andrews and Webb, 2018; Fueglistaler, 2019; Dong et al., 2019, 2020; Cooper et al., 2024), has strong ties to the incomplete-data problem. The pattern effect over the historical record (Andrews et al., 2018, 2022; Marvel et al., 2018; Salvi et al., 2023; Armour et al., 2024) depends on what the SST patterns were in the past, and recent studies have revealed that differences across infilled SST datasets lead to disparate interpretations of the historical pattern effect (Fueglistaler and Silvers, 2021; Lewis and Mauritsen, 2021; Zhou et al., 2021), or possibly no pattern effect at all (Modak and Mauritsen, 2023).

Uncertainty in sea ice is typically omitted from studies of the pattern effect, but Andrews et al. (2018, SI) found that differences in sea ice between AMIP-II and HadISST2 change the shortwave clear-sky feedback by approximately 0.6 W m^{-2} K^{-1}. This change from sea ice alone is approximately the same magnitude as the total pattern effect over the historical record, as the mean pattern effect is 0.48 W m^{-2} K^{-1} using HadISST1 (Andrews et al., 2022). Constraining uncertainty in Antarctic sea ice is important for quantifying historical pattern effects.

We find many differences in the spatial patterns of SST and SIC anomalies relative to AMIP-II and HadISST1, which have been used to account for historical pattern effects and quantify variability in feedbacks over the historical record (Zhou et al., 2016; Andrews et al., 2018; Marvel et al., 2018; Dong et al., 2019; Gregory et al., 2020; Sherwood et al., 2020; Lewis and Mauritsen, 2021; Zhou et al., 2021; Andrews et al., 2022; Salvi et al., 2023; Modak and Mauritsen, 2023). Our reconstruction of monthly SST and SIC can be used as boundary conditions in atmospheric general circulation models to examine the implications for historical feedbacks, pattern effects, and climate sensitivity.

4.6.5 Future opportunities and caveats of the method

Future efforts to reconstruct the historical record could improve on our results in a variety of ways, and we list a few of them here:

- LIMs and DA: Future investigations could elaborate on optimizing the LIMs, their training data, and possibly consider machine-learning methods (e.g., Meng and Hakim, 2024). Our method uses climate models to train the LIMs, and therefore inherits some of the problems in climate models. We mitigate this effect by using eight different CMIP6 models to sample the range of systematic uncertainty and through DA. There are many varieties of DA that could improve on our results, including 4D-Var, quantile-conserving filtering, or multi-model Kalman filtering with a large ensemble generated by various LIMs (Kalnay, 2003; Houtekamer and Zhang, 2016; Anderson, 2022; Bach and Ghil, 2023). Our method assumes state variables can be approximated with Gaussian distributions, which appears to work reasonably well for SIC but could likely be improved in future studies.

- Pressure data: An update of non-infilled HadSLP2 (Allan and Ansell, 2006) would be helpful,

as no quality-controlled dataset of gridded monthly mean SLP with error estimates is currently available. ICOADS provides only marine data (Freeman et al., 2017) and does not include observation errors. Including terrestrial pressure data (Cram et al., 2015) could improve our reconstruction, but no gridded product exists, and elevation differences are a considerable source of error.

- Sea ice: There are many observations available before the satellite era (e.g., Walsh et al., 2019; Edinburgh and Day, 2016; Titchner and Rayner, 2014), but we do not have a current compilation of this data in a format that can be used in reconstructions. A dataset structured like HadSST4 or DCENT but with historical SIC observations would be helpful.

- SST: Ongoing efforts to digitize new data, quantify error, and correct the biases of existing data will continue to be critical (e.g., Brönnimann et al., 2019, 2024; Chan et al., 2019, 2023; Kent and Kennedy, 2021; Kennedy et al., 2019). For SST anomalies (also T and SLP), it would be helpful to use a climatological period that overlaps with satellite observations of SIC (i.e., post-1979).

4.7 Conclusions

The historical record is essential to our understanding of coupled climate dynamics and variability, but instrumental observations are sparse and noisy. Moreover, existing observational datasets are typically derived separately for each component of the climate system, leading to inconsistencies in coupled variability when they are combined.

In this study, we develop a method for climate reanalysis using strongly coupled data assimilation. The key advance of our method compared to past work is that we (i) ensure that the coupled atmosphere-ocean-ice state is internally consistent and (ii) synthesize observational and dynamical constraints across all components simultaneously. Using a Kalman filter, we combine monthly forecasts from linear inverse models (LIMs), which are trained on eight CMIP6 models to account for model error, with observations of SST, land temperature, marine sea-level pressure, and satellite-era sea ice.

We first validate the method through pseudo-reconstruction of an out-of-sample climate model,

then we present the actual reconstruction on a global $2° \times 2°$ grid with monthly resolution of SST, near-surface air temperature, sea-level pressure, and sea-ice concentration over 1850–2023. We also provide a novel quantification of the time-varying uncertainty in all fields and its spatial fingerprints.

In many ways, our results differ from comparison datasets regarding how recent trends (c. 1980–present) compare to past variability. The recent evolution of the Walker circulation appears consistent with past variability, but the SST contrast (SST$^{\#}$; warmest regions versus the tropical mean) exhibits a prolonged strengthening from 1975–present that appears distinct from past variability.

In the Southern Ocean, we find a weaker SST cooling post-1980 compared to the strong cooling in other estimates (namely ERSSTv5), which climate models have been unable to replicate. We emphasize the observational uncertainty over the Southern Ocean, which merits more attention due to sparse and problematic data even after 1980. The Southern Annular Mode appears well constrained but differs substantially from existing estimates before 1980. Antarctic sea ice also follows a different trajectory in our reconstruction compared to other estimates over the majority of the record (1850–1980). Our constraints on Antarctic sea ice are a key result, as we find much less ice loss over 1900–1980 compared to existing datasets, but with large uncertainty.

Our historical reconstruction is designed for climate analysis and is publicly available. We provide the grand mean of all 1600 ensemble members, the separate ensemble means for each of the eight model priors, and a subset of 200 fully gridded ensemble members. Our monthly SST and sea ice can also be used as boundary conditions in atmospheric general circulation models (i.e., in AMIP-type simulations). Through coupled data assimilation, this reconstruction improves constraints on coupled climate dynamics and variability, highlights key uncertainties in the historical record, and guides future investigations into coupled atmosphere–ocean–ice interactions.

4.8   *Data availability*

The reconstruction will be publicly available in a Zenodo repository, accessible through Cooper et al. (2025).

Observation data is available as follows: HadSST4.0.1.0 at `https://www.metoffice.gov.uk/hadobs/hadsst4/index.html` at CRUTEM.5.0.2.0, `https://www.metoffice.gov.uk/hadobs/crutem5/`

data/CRUTEM.5.0.2.0/download.html; ICOADS SLP at https://downloads.psl.noaa.gov/Datasets/icoads/2degree/enh/; NOAA/NSIDC CDRv4 at https://doi.org/10.7265/efmz-2t65 and NRTv2 at https://doi.org/10.7265/tgam-yv28.

SST and SIC comparison data is available as follows: PCMDI/AMIP-II at https://aims2.llnl.gov/ with specifier input4MIPs.CMIP6Plus.CMIP.PCMDI.PCMDI-AMIP-1-1-9; HadISST1 at https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html, HadISST2.1.0.0 SST at https://www.metoffice.gov.uk/hadobs/hadisst2/data/HadISST.2.1.0.0/index.html and HadISST2.2.0.0 SIC at https://www.metoffice.gov.uk/hadobs/hadisst2/data/download.html; COBE-SST2 at https://downloads.psl.noaa.gov/Datasets/COBE2/; ERSSTv5 from https://www.ncei.noaa.gov/products/extended-reconstructed-sst or https://doi.org/10.5065/JZ08-3W17. Brennan and Hakim (2022) is available at https://doi.org/10.5281/zenodo.5809703

GMSAT comparison data is available as follows: BEST at https://climate.metoffice.cloud/temperature.html; HadCRUT5 at https://www.metoffice.gov.uk/hadobs/hadcrut5/data/HadCRUT.5.0.2.0/download.html.

SLP comparison data is available as follows: NOAA/CIRES/DOE 20th Century Reanalysis (V3) and additional datasets listed here are provided by the NOAA PSL, Boulder, Colorado, USA, from their website at https://psl.noaa.gov; NCEP/NCAR Reanalysis at https://downloads.psl.noaa.gov//Datasets/ncep.reanalysis/Monthlies/surface/; ERA5 post-1979 at https://doi.org/10.5065/P8GT-0R61 and 1950–1978 back extension at https://doi.org/10.5065/JAXB-X906; HadSLP2 at https://www.metoffice.gov.uk/hadobs/hadslp2/data/download.html; King et al. (2023) at https://doi.org/10.5281/zenodo.8156908; O'Connor et al. (2021) at https://doi.org/10.5281/zenodo.5507607; Dalaiden et al. (2021) at https://doi.org/10.5281/zenodo.4770179; Fogt et al. (2009) at https://polarmet.osu.edu/ACD/sam/sam_recon.html; Falster et al. (2023) at https://doi.org/10.5281/zenodo.8280559.

CMIP6 output is available on NCAR Glade and at https://esgf-node.llnl.gov/search/cmip6/.

4.9 Appendix

4.9.1 Summary of LIM training data

Model	Total Years (piControl range)	Ens. Mem.	EOFs	Reference
CESM2	1166 (200–1200)	r1i1p1f1	408	Danabasoglu et al. (2020)
UKESM1.0	1754 (2250–3839)	r1i1p1f2	408	Sellar et al. (2019)
SAM0-UNICON	865 (1–700)	r1i1p1f1	306	Park et al. (2019)
GFDL-ESM4	665 (1–500)	r1i1p1f1	306	Dunne et al. (2020)
NorESM2-LM	666 (1600–2100)	r1i1p1f1	306	Seland et al. (2020)
EC-Earth3	1165 (2103–3102)	r2i1p1f1	408	Döscher et al. (2022)
HadGEM3-GC31-LL	2165 (1850–3849)	r1i1p1f1	408	Kuhlbrodt et al. (2018)
E3SM-2	665 (1–500)	r1i1p1f1	306	Qin et al. (2024)

Table 4.1: CMIP6 training data for 8 linear inverse models. All models with 408 EOFs have the following distribution across state variables: 108 SST, 108 T, 48 SLP, 72 Arctic SIC, 72 Antarctic SIC. Models with 306 EOFs have 92 SST, 84 T, 30 SLP, 50 Arctic SIC, 50 Antarctic SIC. Note that Total Years includes piControl plus 165 years of historical simulation (1850–2014).

4.9.2   Observation error for sea-level pressure

To estimate **R** for observations of monthly mean SLP, we apply a method similar to that in Kaplan et al. (2000). The intramonth standard deviation (s) provided by ICOADS is comprised of submonthly variability, measurement error, and representativeness error, thus providing an estimate of the observation error in the monthly mean (Leith, 1973). We take the local time-average of $s^2 \frac{n_{\mathrm{obs}}}{n_{\mathrm{obs}}-1}$ over the well observed period 1961–2023 to estimate the climatological error variance, σ^2, in the monthly mean for each gridcell, and we restrict the estimate to gridcells with $n_{\mathrm{obs}} > 30$ in a given month. Again using a similar approach to Kaplan et al. (2000), we then spatially smooth the resulting climatological maps of σ using a running-mean window of 12° latitude $\times$ 50° longitude equatorward of 52°N/S and a window of 18° latitude $\times$ 100° longitude poleward of 52°N/S. This results in 12 monthly 2° $\times$ 2° fields of the random measurement and sampling error, σ_{random}.

We then must assign a time-varying error, σ, to each monthly value of SLP. We start with the random error described above, then reduce the random error by the number of intramonth observations in a gridcell. To account for autocovariance and possible sampling errors even when n_{obs} is large, reduce n_{obs} to $n_{\mathrm{adjusted}} = n_{\mathrm{obs}}/2$, and we set the maximum of n_{adjusted} at 30 (Leith, 1973; Bretherton et al., 1999). We then consider the systematic component of the total error, $\sigma^2 = \sigma^2_{\mathrm{systematic}} + \sigma^2_{\mathrm{random}}/n_{\mathrm{adjusted}}$, as discussed in Kennedy (2014). We estimate $\sigma^2_{\mathrm{systematic}}$ from the variance across neighboring observations. The idea is that if neighboring observations consistently differ, the differences are from irreducible, systematic errors. Separately for each month from 1961–2023, we calculate the spatial variance across a running-mean window of 16° latitude $\times$ 32° longitude, restricting the calculation to gridcells with $n_{\mathrm{obs}} \geq 5$. We use the zonal mean of the climatology of this field to represent $\sigma^2_{\mathrm{systematic}}$. We make one adjustment by setting the minimum $\sigma_{\mathrm{systematic}}$ at 6 hPa south of 72°S, preventing the error from decreasing near the Antarctic coastline. The systematic error ranges from approximately 1 hPa on the equator to 7 hPa in polar regions, with a local maximum of 9.5 hPa over the Southern Ocean at 55°S.

4.9.3   Supplemental information

Supplemental Figures 4.S1–S17.
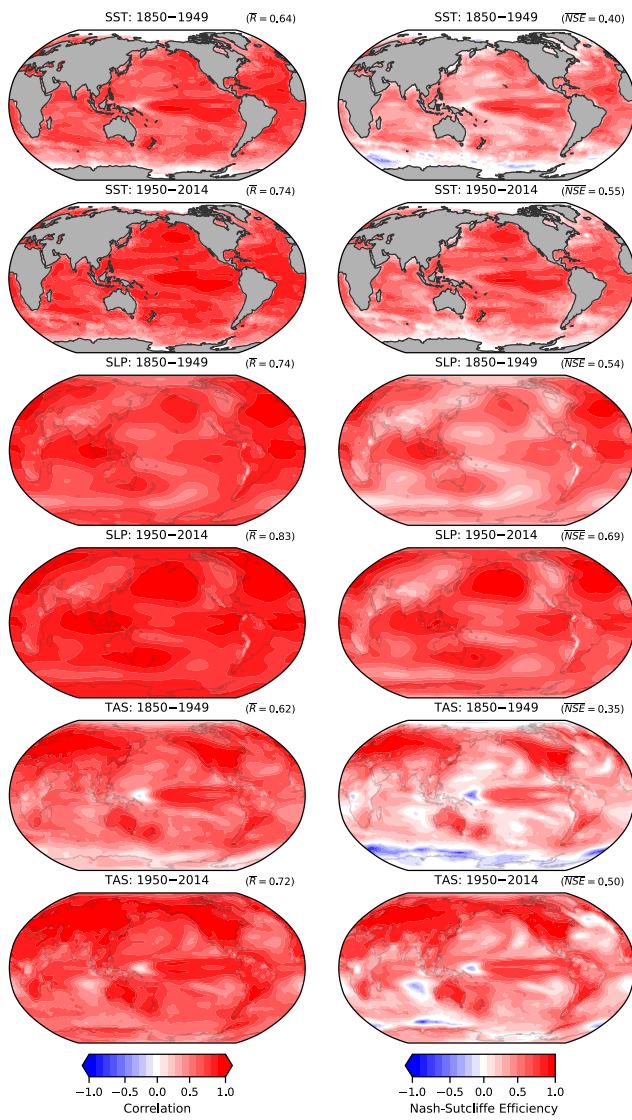
Figure 4.S1: Pseudo-reconstruction of MPI-ESM1-2-HR: correlation and Nash-Sutcliffe efficiency (NSE) for SST, SLP, and T. (Left) Local correlation (R) of ensemble-mean reconstruction with the target monthly values from MPI-ESM1-2-HR, separately for 1850–1949 and 1950–2014; global mean of local R is shown in upper right. (Right) Same as left column but showing the local NSE, with the global mean of the local NSE shown in upper right.
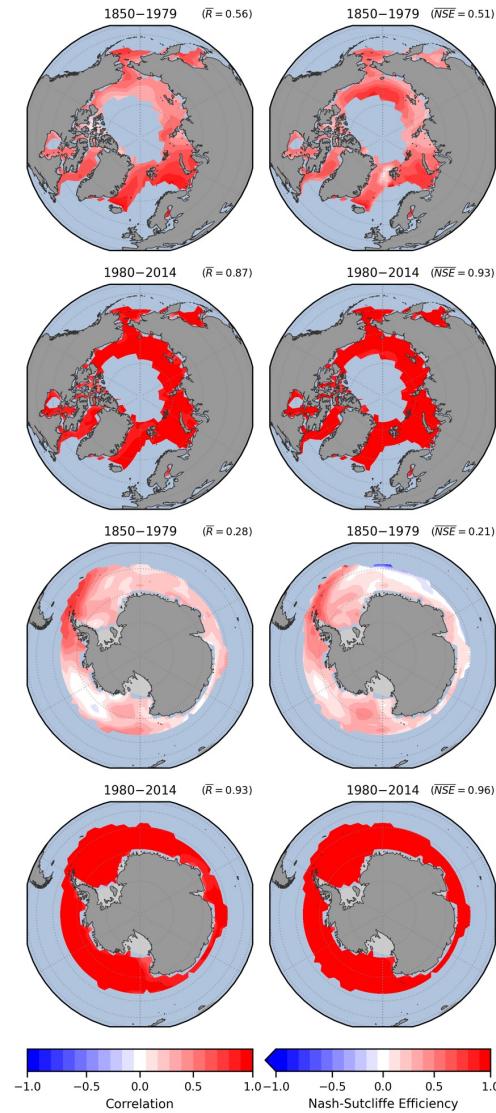
Figure 4.S2: Pseudo-reconstruction of MPI-ESM1-2-HR: correlation (R) and Nash-Sutcliffe efficiency for sea-ice concentration (SIC). (Left column) Local correlation of ensemble mean from the reconstruction with the true annual values from MPI-ESM1-2-HR's historical simulation, separately for 1850–1979 and 1980–2014, and the global-mean of local correlations is shown in upper right. (Right column) Same as left column but showing the Nash-Sutcliffe Efficiency (NSE), with the global mean of the local NSE shown in upper right.
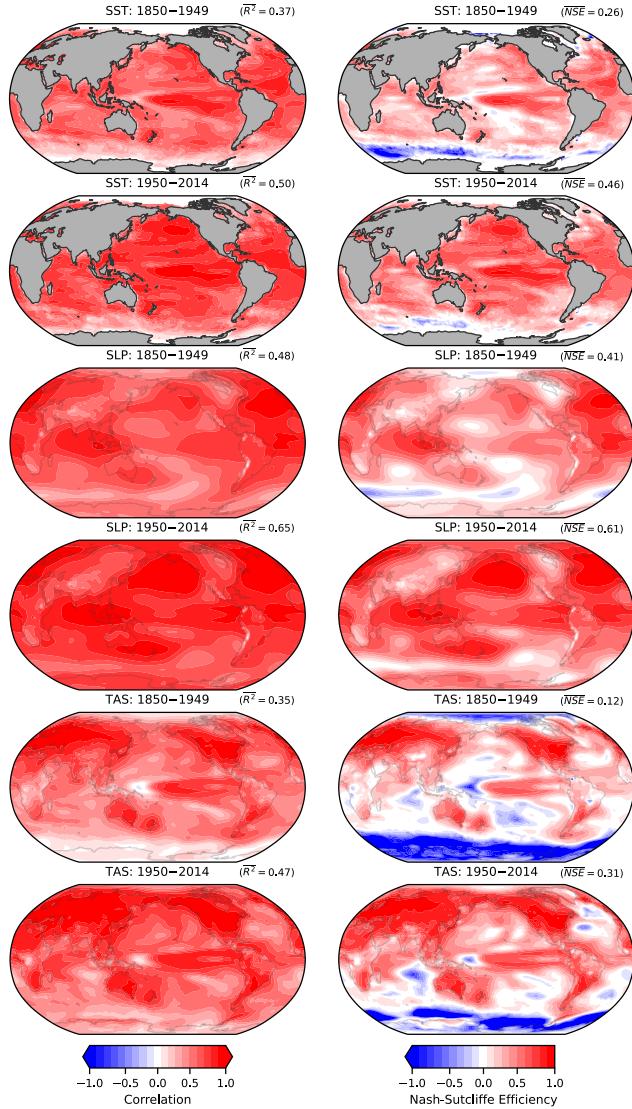
Figure 4.S3: Pseudo-reconstruction of MPI-ESM1-2-HR, using only a single model prior. Same as Figure 4.S1, but using only the LIM trained on CESM2 to produce the pseudo-reconstruction (i.e., excluding the other model priors). Compare with Figure 4.S1 to see the benefit of including multiple priors in the reconstruction.
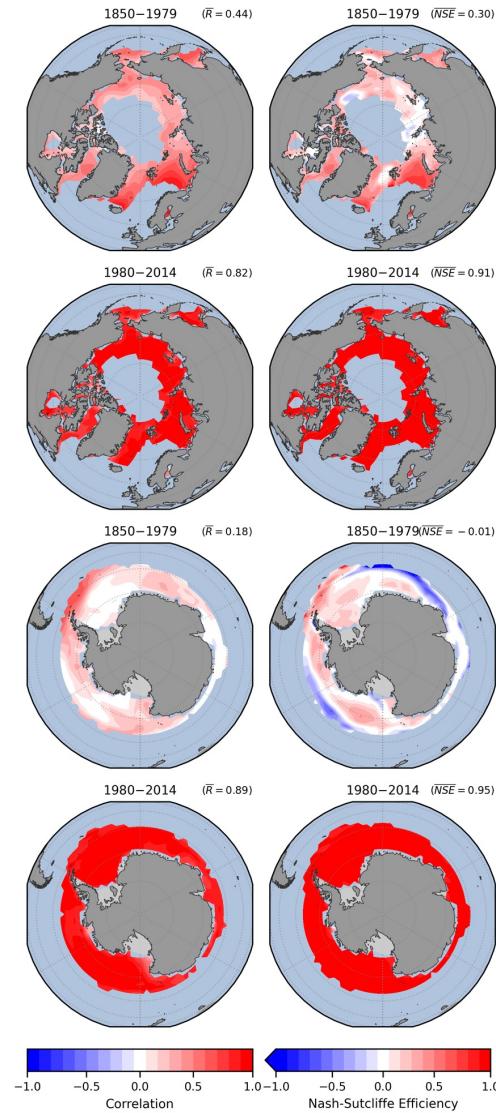
Figure 4.S4: Pseudo-reconstruction of MPI-ESM1-2-HR, using only a single model prior. Same as Figure 4.S2, but using only the linear inverse model trained on CESM2 to produce the pseudo-reconstruction (i.e., excluding the other model priors). Compare with Figure 4.S2 to see the benefit of including multiple priors in the reconstruction.
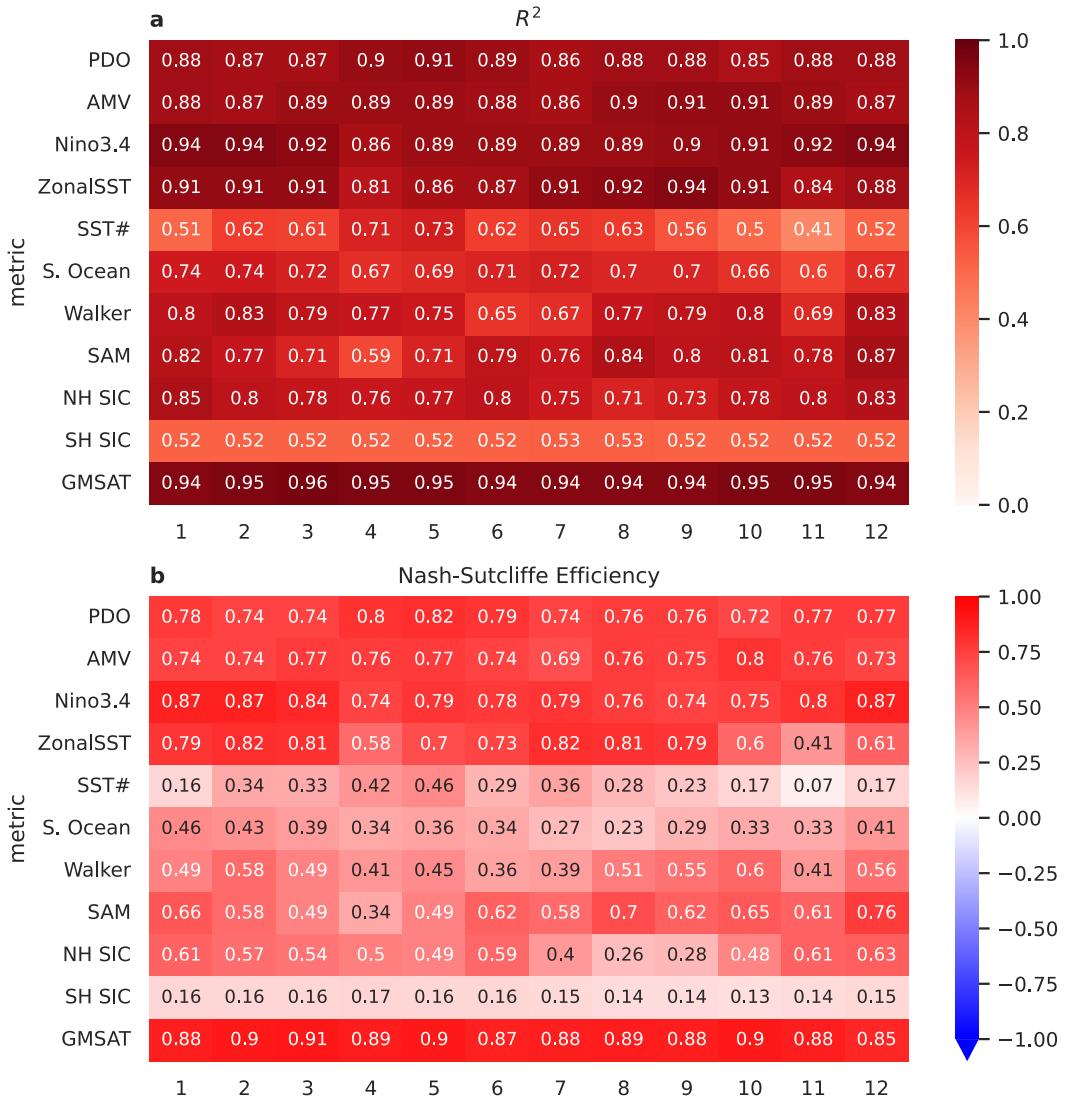
Figure 4.S5: Validation by pseudo-reconstruction of MPI-ESM1-2-HR: squared Pearson's correlation and Nash-Sutcliffe efficiency by month. Metrics on the vertical axis correspond to those in Figure 4.2 but the calculations are based on the monthly resolved data without any low-pass filtering. The horizontal axis represents months January–December. Calculation is over the full pseudo-reconstruction period (1850–2014).
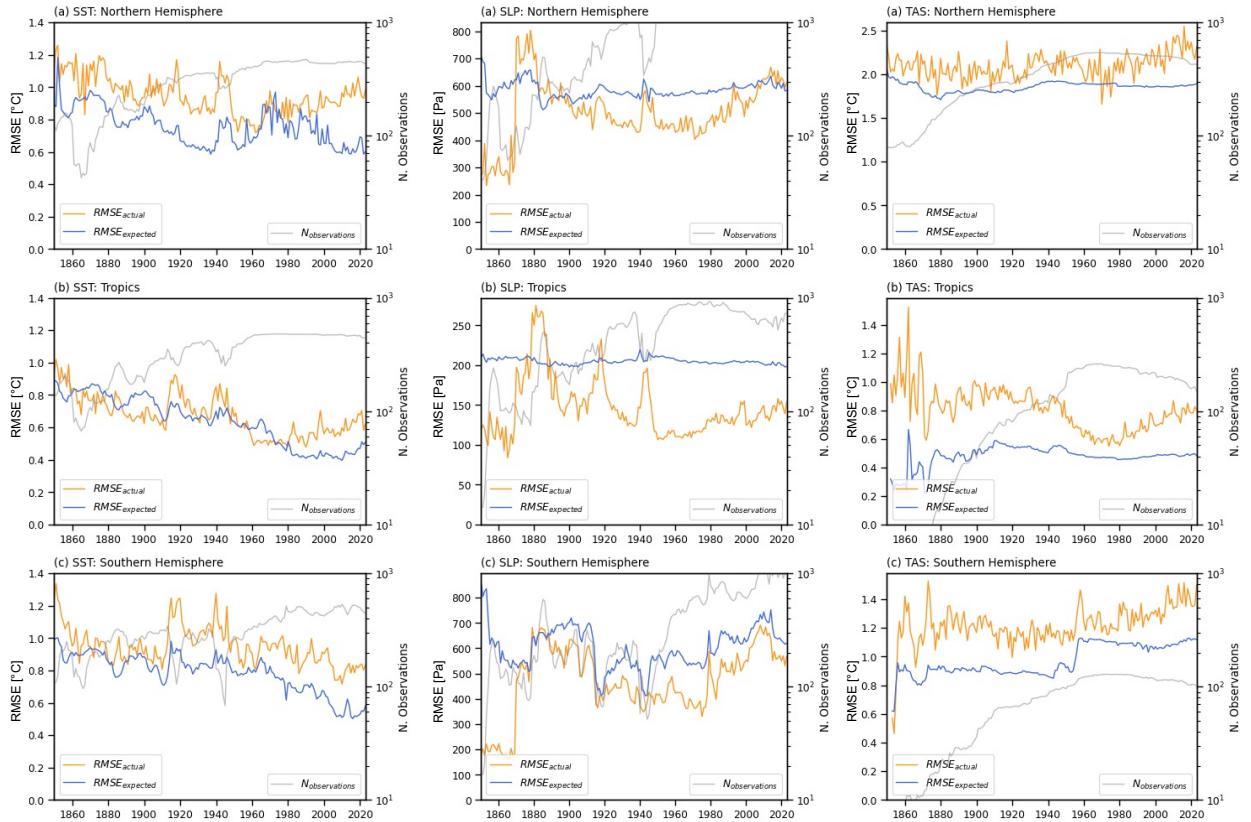
Figure 4.S6: Desroziers validation statistics. (Left column) Actual versus expected RMSE averaged for sea-surface temperature over **(a)** the Northern Hemisphere from 20°–90°N, **(b)** the Tropics from 20°S–20°N, and **(c)** the Southern Hemisphere from 20°–90°S. The annual-mean number of observations assimilated per month is shown in gray (right vertical axis). Note that the vertical axes differ between subplots. (Middle column) Same as left column for sea-level pressure. (Right column) Same as left column for near-surface air temperature. All time series show annual means.
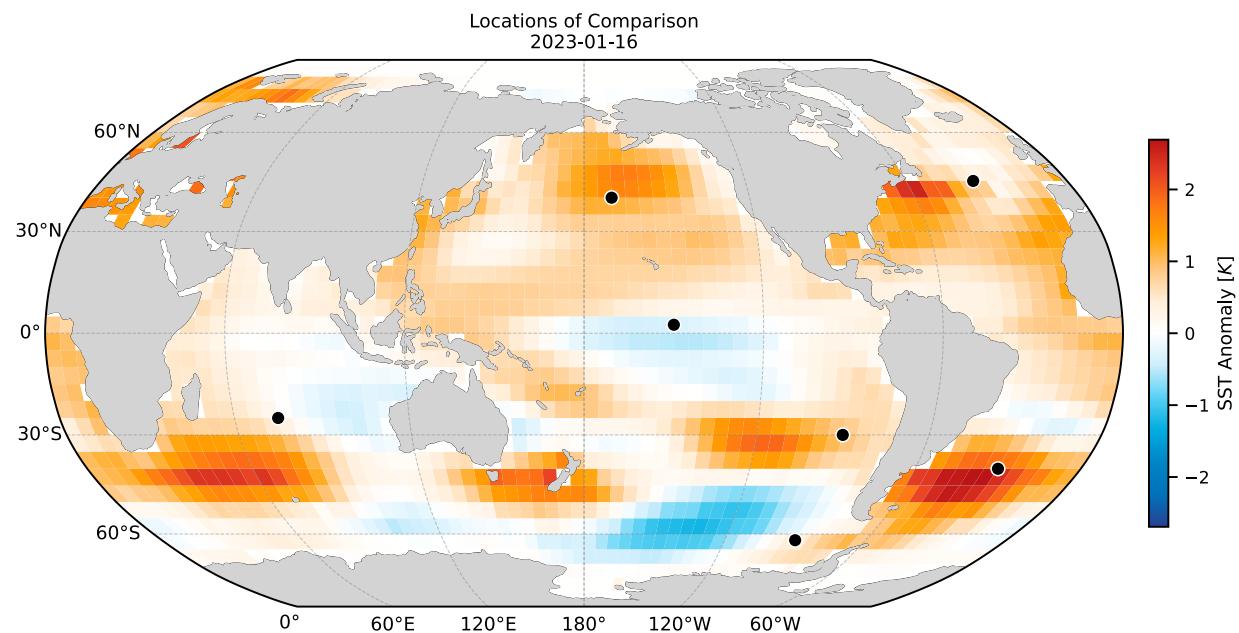
Figure 4.S7: Select point locations for comparing reconstruction to HadSST4 in following figures. Black dots indicate the locations shown in Figures 4.S8–S14. The SST plotted shows a single monthly mean, illustrating the January 2023 reconstruction anomaly relative to the 1961–1990 climatology.
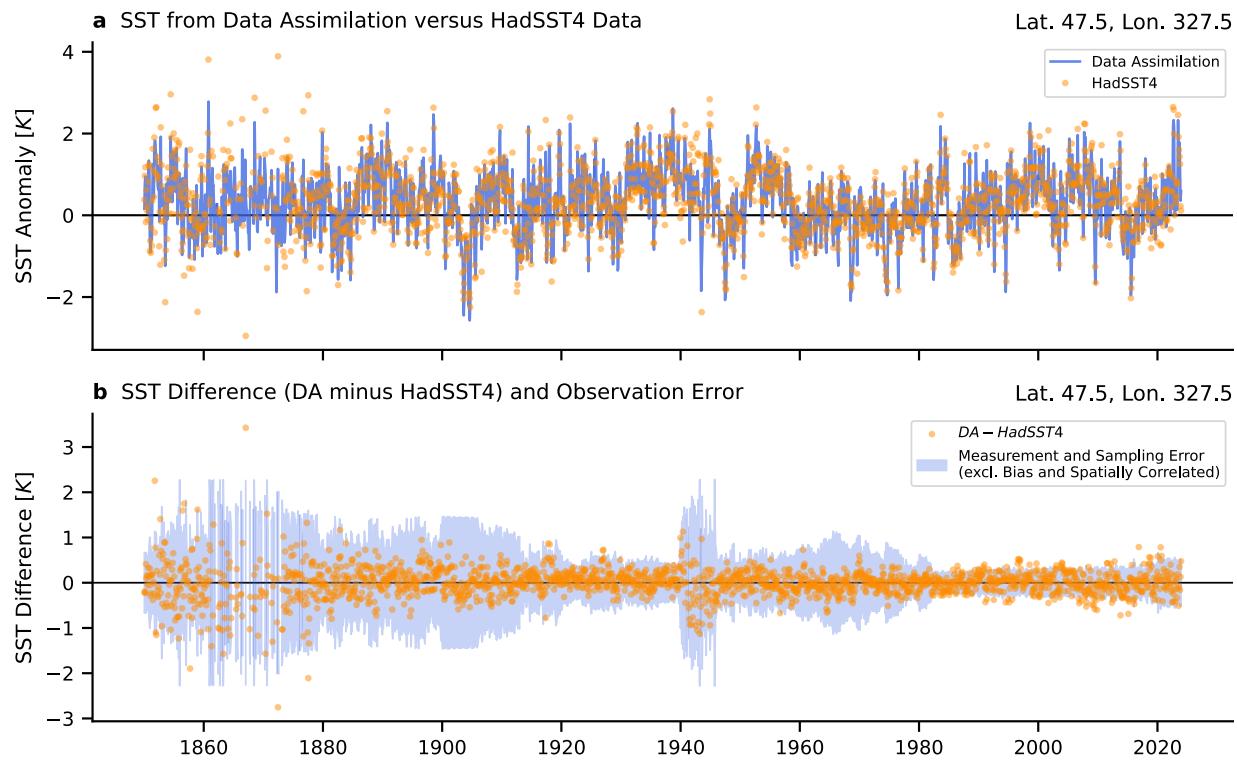
Figure 4.S8: North Atlantic. **(a)** Timeseries of SST anomalies in a single gridcell from the data assimilation (DA) ensemble-mean posterior versus the HadSST4 data. **(b)** Analysis residuals, i.e., the SST difference between the DA posterior and HadSST4 data shown, corresponding to panel **a**; Diagonal terms in **R**, i.e., the measurement and sampling error ($\pm 2\sigma$), shown as shading. DA result is regridded to the coarser 5° resolution of HadSST4 for the comparison.

a SST from Data Assimilation versus HadSST4 Data — Lat. -37.5, Lon. 332.5

b SST Difference (DA minus HadSST4) and Observation Error — Lat. -37.5, Lon. 332.5
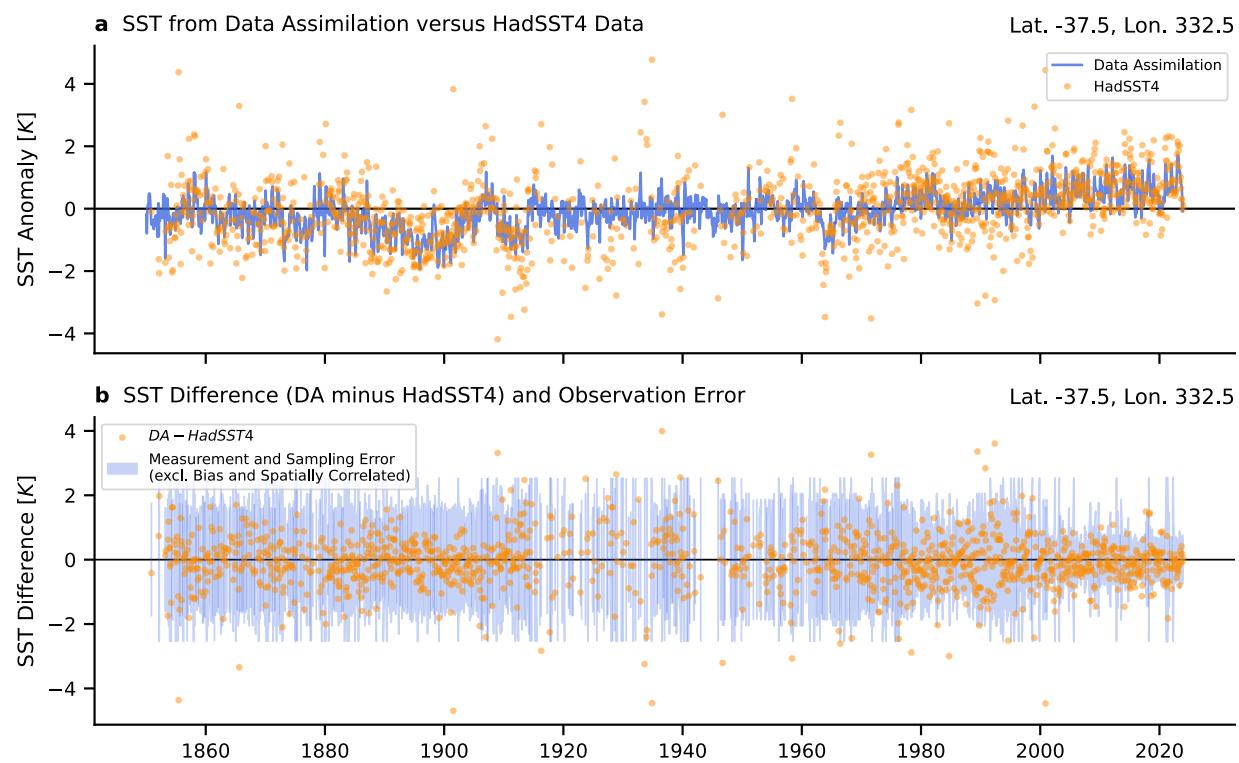
Figure 4.S9: South Atlantic. See caption of Figure 4.S8.
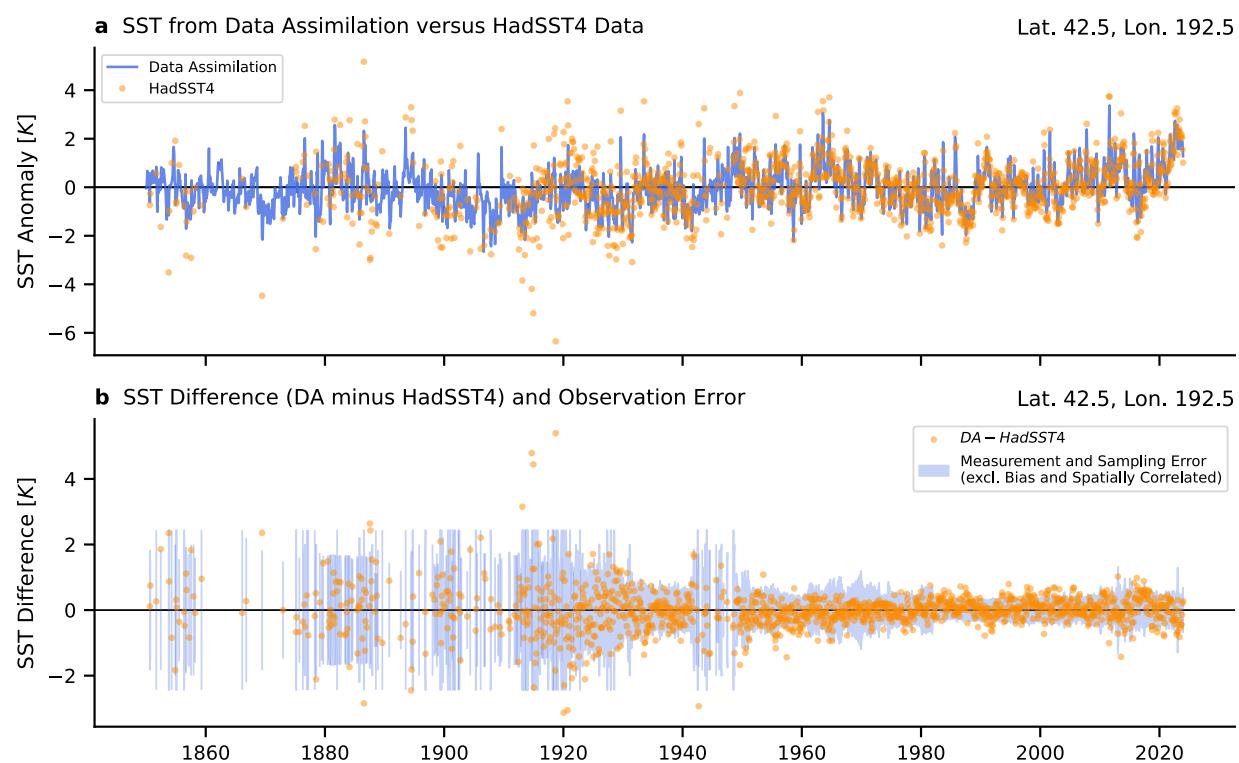
Figure 4.S10: North Pacific. See caption of Figure 4.S8.
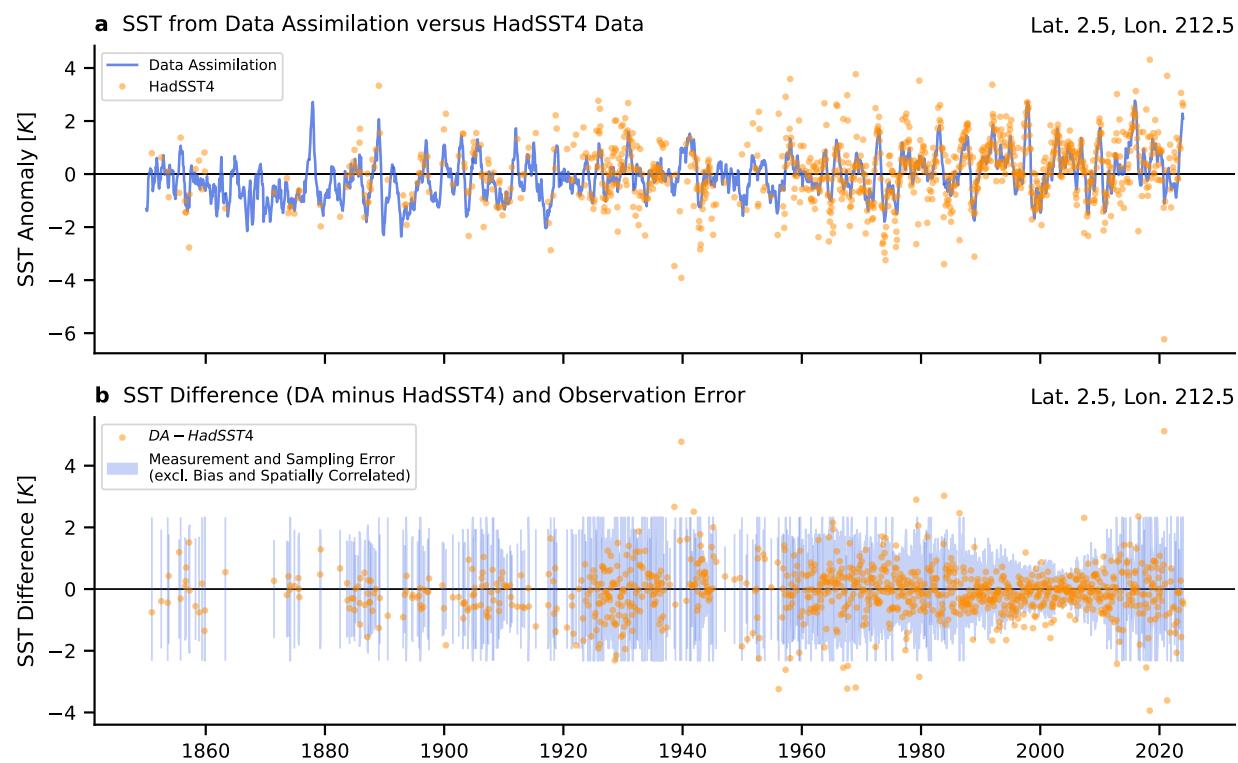
Figure 4.S11: North Pacific. See caption of Figure 4.S8.
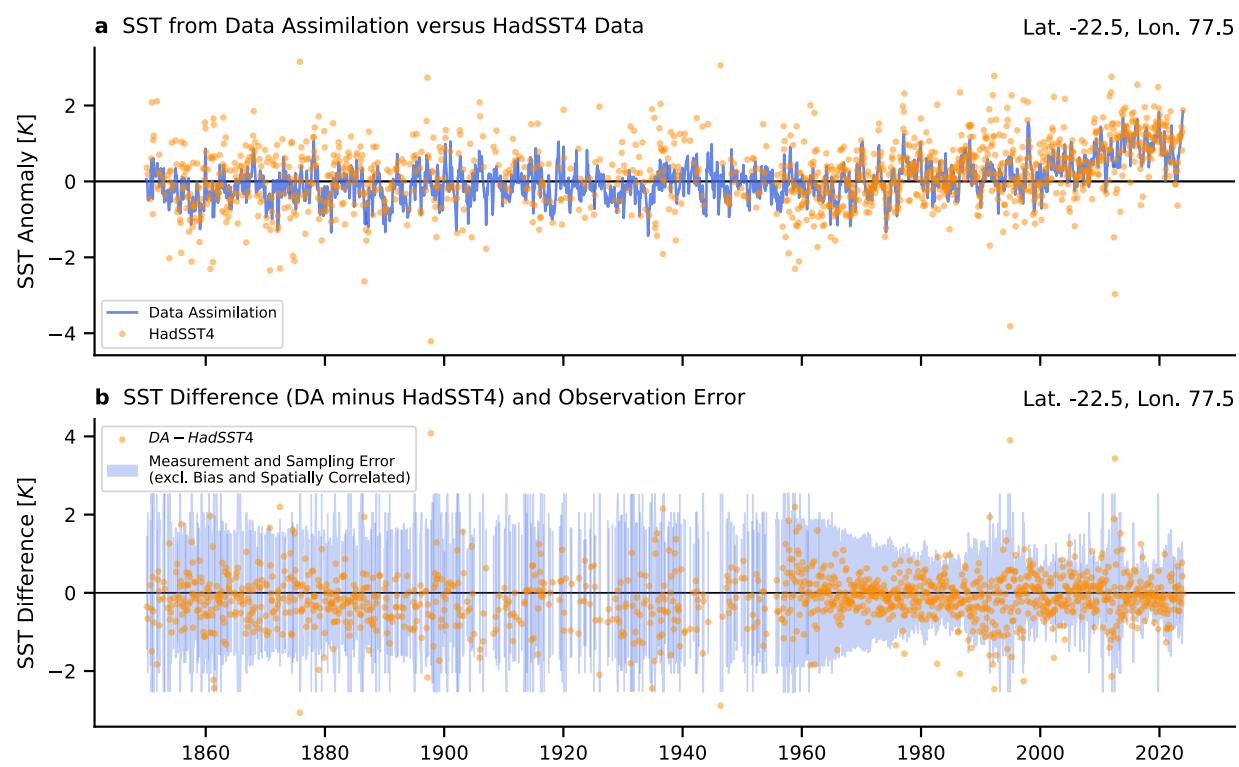
Figure 4.S12: Indian Ocean. See caption of Figure 4.S8.

a SST from Data Assimilation versus HadSST4 Data   Lat. -27.5, Lon. 272.5

b SST Difference (DA minus HadSST4) and Observation Error   Lat. -27.5, Lon. 272.5
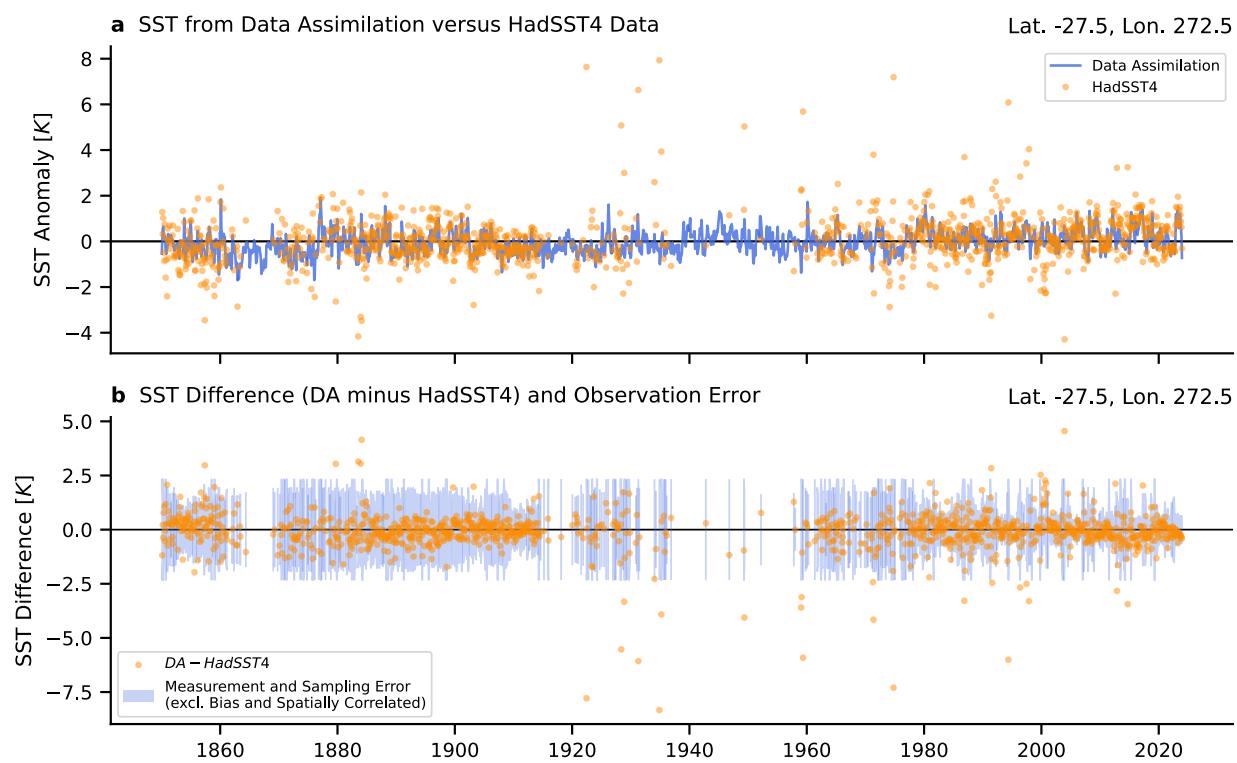
Figure 4.S13: Southeast Pacific. See caption of Figure 4.S8.
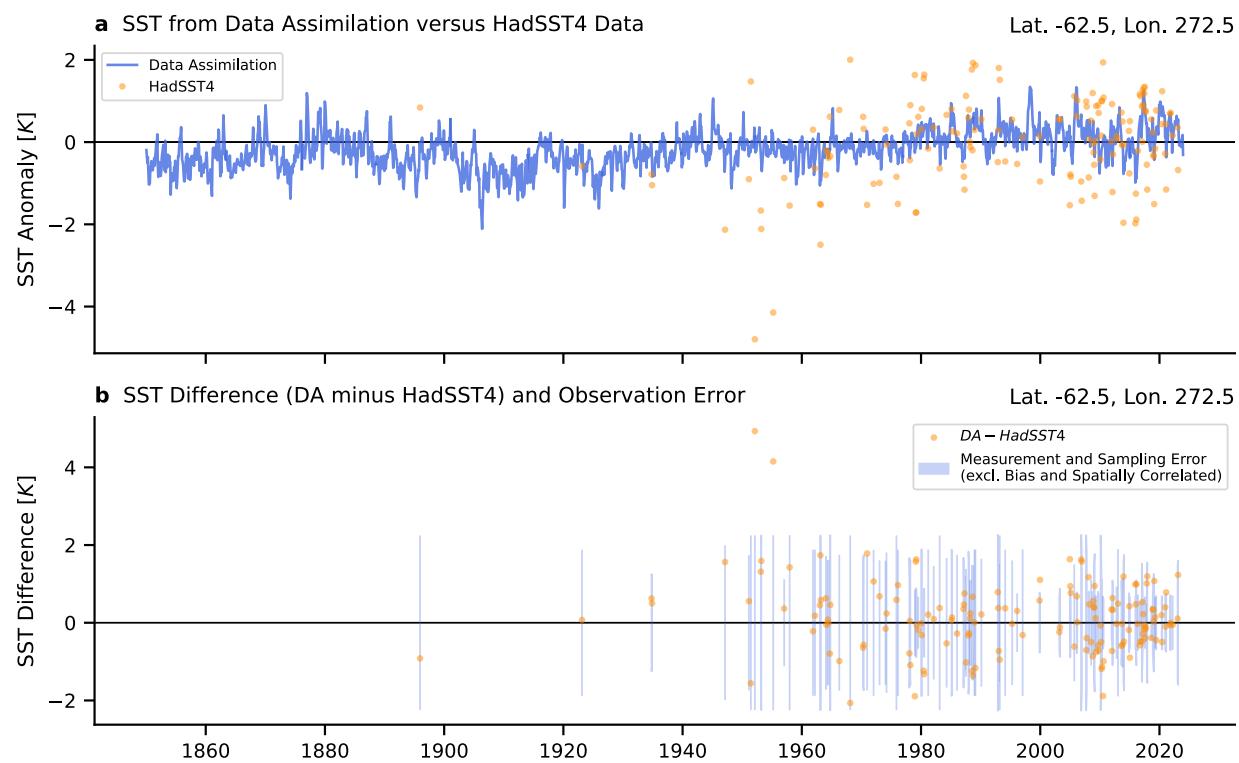
Figure 4.S14: Southern Ocean (SE Pacific sector). See caption of Figure 4.S8.
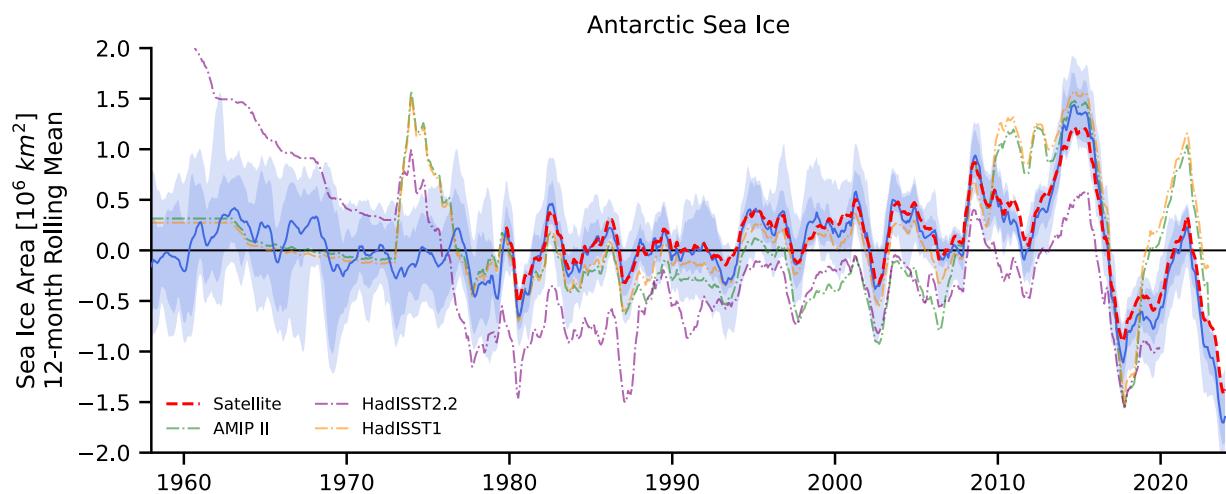
Figure 4.S15: Antarctic sea ice area, 1958–2023. As shown in Figure 4.5 but limited to recent decades. Data assimilation results are shown in blue (mean, 66% range, and 90% range). All timeseries are filtered with a 12-month running mean of the monthly means. Anomalies are relative to the 1961–1990 baseline.
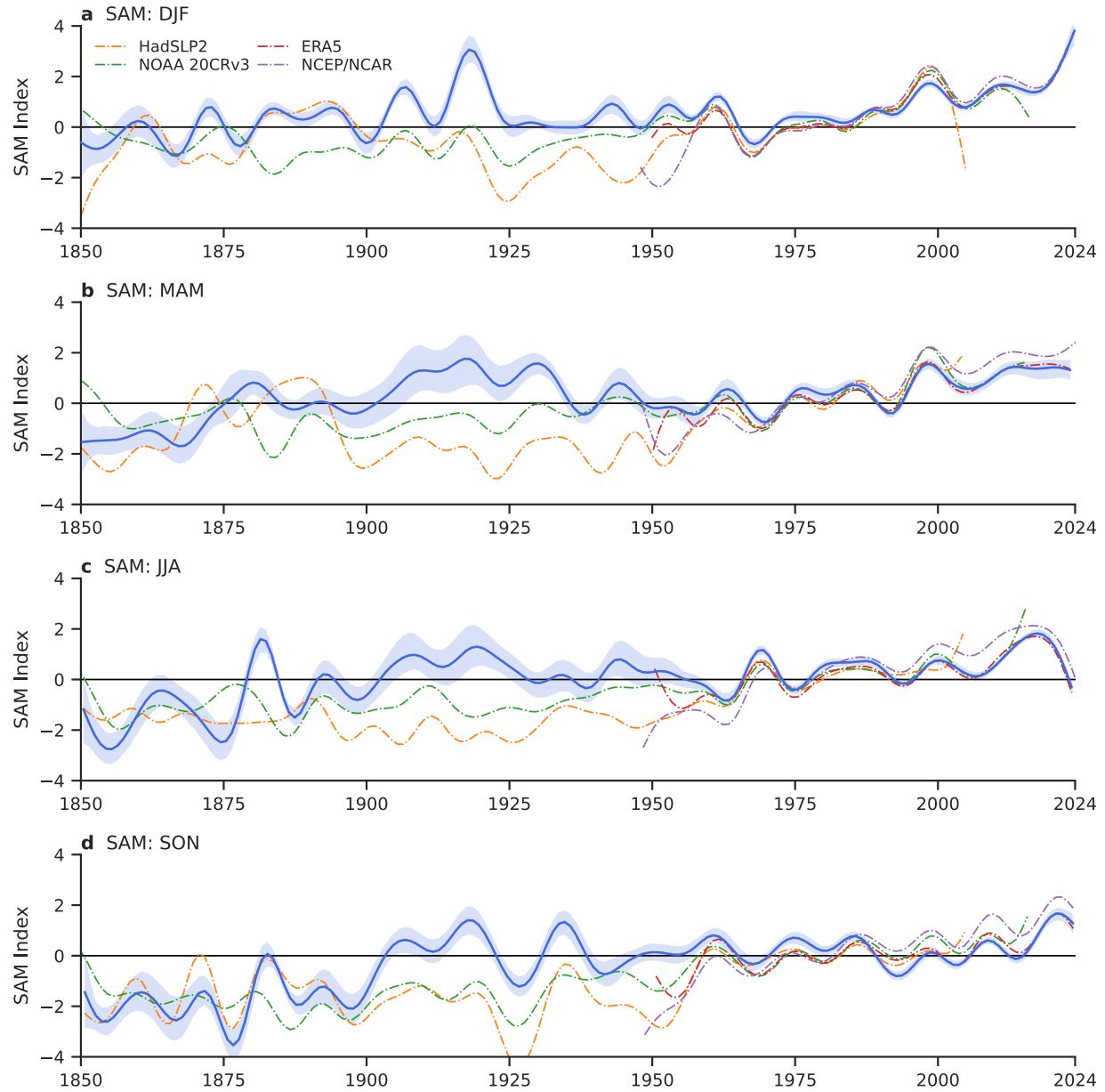
Figure 4.S16: Seasonal values of the Southern Annular Mode (SAM). As shown in Figure 4.5 but for separate seasons. Data assimilation results are shown in blue (mean, 66% range). All timeseries are filtered with a 10-year low-pass filter on the monthly means. Anomalies are relative to the 1961–1990 baseline.

a Arctic Sea Ice vs. Global-mean Temperature

**b** Antarctic Sea Ice vs. Global-mean Temperature
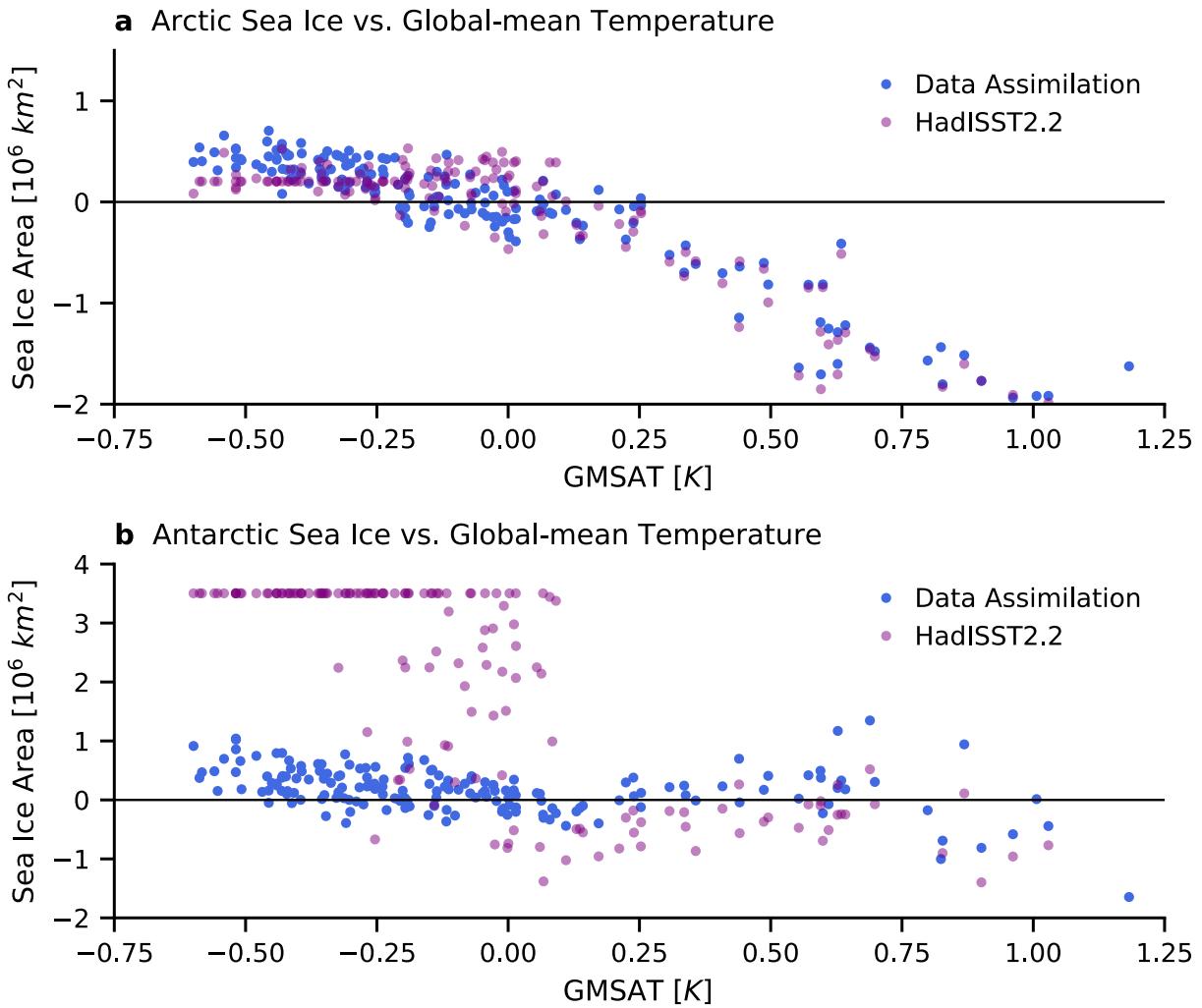
Figure 4.S17: Sea ice area versus global-mean near-surface air temperature (GMSAT). Using annual means of the results from data assimilation shown in Figure 4.5, **(a)** scatter plot of anomalies in total Arctic sea ice area versus anomalies in GMSAT, and **(b)** repeated for the Antarctic.

Chapter 5

CONCLUSIONS

In this dissertation, we have analyzed paleoclimates and the historical record to develop a better understanding of climate sensitivity and coupled variability. We found that paleoclimates provide a stronger constraint on climate sensitivity and future warming after accounting for paleoclimates' distinct patterns of temperature change and their associated radiative feedbacks. For the recent historical record (1850–present), we used coupled data assimilation to synthesize observational and dynamical constraints on climate variability and trends, thus providing a physically consistent reconstruction of Earth's surface climate over the past 170 years.

In Chapter 2, we investigated the relationship between radiative feedbacks and the spatial pattern of SST in the Last Glacial Maximum (LGM). We found that the LGM temperature pattern is associated with cloud feedbacks that amplify the temperature change. The SST pattern and its amplifying feedbacks are attributable to the LGM's massive North American ice sheet rather than the LGM's direct response reduced CO_2 levels. Accounting for such pattern effects in the LGM evidence leads to stronger constraints on modern climate sensitivity.

In Chapter 3, we investigated radiative feedbacks in the Pliocene, a past warm period with CO_2 levels similar to present day (approximately 400 ppm). We found that non-CO_2 forcings, such as ice sheets, topography, and vegetation, amplified Pliocene warming through their influence on SST patterns and cloud feedbacks over the North Atlantic and Southern Ocean. Because the Pliocene's non-CO_2 forcings play a large role in amplifying Pliocene warming, but they do not play a role in the modern response to CO_2 alone, modern climate sensitivity is lower than previously assessed from Pliocene evidence. We combine our analysis of the warm Pliocene with Chapter 2's investigation of the cold LGM, and we find that accounting for paleoclimate pattern effects produces a best estimate (median) for modern climate sensitivity of $2.8°C, 66\%$ range $2.4 - 3.4°C$ (90% CI: $2.1 - 4.0°C$), substantially reducing uncertainty and narrowing projections of 21$^{\text{st}}$-century warming.

In Chapter 4, we turned our attention to the historical record (1850–2023), inspired by the

uncertainty in SST patterns and sea ice that impacts historical constraints on climate sensitivity and variability. We presented a method combining *linear inverse models*, which we use to forecast Earth's surface climate (SST, near-surface air temperature, sea-level pressure, and sea ice) at monthly resolution, with *strongly coupled data assimilation*. We used this method to constrain the coupled climate state based on the dynamics of the models and the observations across the atmospheric and oceanic fields. In our reconstruction, we find that recent trends (post-1980) in the Walker circulation are more consistent with past variability, whereas the tropical SST contrast (the difference between warmer and colder SSTs) shows a distinct strengthening since 1975. ENSO amplitude exhibits substantial low-frequency variability and a local maximum in variance over 1875–1910. In the Southern Ocean, we find a muted cooling trend in SST with substantial uncertainty and that changes in Antarctic sea ice are relatively small between 1850 and 2000.

Together, the chapters of this thesis highlight the importance of past patterns of temperature change in understanding modern climate sensitivity and variability. Spatial information plays a pivotal role in constraining climate feedbacks and sensitivity, hence analyses with a limited focus only on global-mean changes lead to flawed inferences of modern climate sensitivity. The information on spatial patterns in this dissertation is a product of *data assimilation*, which is the core of the current revolution in combining models and data. Accounting for paleoclimate pattern effects leads to stronger constraints on future warming and especially on the persistently uncertain upper bound of modern climate sensitivity. Additionally, a new reconstruction and analysis using instrumental data (1850–2023) provides an internally consistent, dynamically constrained perspective on the historical climate record. We clarify many long-standing uncertainties surrounding the tropical Pacific and the Southern Ocean by synthesizing surface observations across the atmosphere and ocean. This dissertation demonstrates the value of paleoclimate and historical records in understanding modern and future climate.

BIBLIOGRAPHY

Albani, S. and Mahowald, N. M.: Paleodust Insights into Dust Impacts on Climate, J. Climate, 32, 7897–7913, https://doi.org/10.1175/JCLI-D-18-0742.1, 2019.

Albani, S., Balkanski, Y., Mahowald, N., Winckler, G., Maggi, V., and Delmonte, B.: Aerosol-Climate Interactions During the Last Glacial Maximum, Current Climate Change Reports, 4, 99–114, https://doi.org/10.1007/s40641-018-0100-7, 2018.

Alexander, M. A., Bladé, I., Newman, M., Lanzante, J. R., Lau, N.-C., and Scott, J. D.: The Atmospheric Bridge: The Influence of ENSO Teleconnections on Air–Sea Interaction over the Global Oceans, J. Climate, 15, 2205–2231, https://doi.org/10.1175/1520-0442(2002)015⟨2205:TABTIO⟩2.0.CO;2, 2002.

Allan, R. and Ansell, T.: A New Globally Complete Monthly Historical Gridded Mean Sea Level Pressure Dataset (HadSLP2): 1850–2004, J. Climate, 19, 5816–5842, https://doi.org/10.1175/JCLI3937.1, 2006.

Allan, R. P., Liu, C., Loeb, N. G., Palmer, M. D., Roberts, M., Smith, D., and Vidale, P.: Changes in global net radiative imbalance 1985–2012, Geophys. Res. Lett., 41, 5588–5597, https://doi.org/10.1002/2014GL060962, 2014.

Amaya, D. J., Seltzer, A. M., Karnauskas, K. B., Lora, J. M., Zhang, X., and DiNezio, P. N.: Air-sea coupling shapes North American hydroclimate response to ice sheets during the Last Glacial Maximum, Earth and Planetary Science Letters, 578, 117 271, https://doi.org/10.1016/j.epsl.2021.117271, 2022.

Amaya, D. J., Maher, N., Deser, C., Jacox, M. G., Alexander, M. A., Newman, M., Dias, J., and Lou, J.: Linking Projected Changes in Seasonal Climate Predictability and ENSO Amplitude, J. Climate, 38, 675–688, https://doi.org/10.1175/JCLI-D-23-0648.1, 2025.

Amrhein, D. E., Wunsch, C., Marchal, O., and Forget, G.: A Global Glacial Ocean State Estimate Constrained by Upper-Ocean Temperature Proxies, J. Climate, 31, 8059–8079, https://doi.org/10.1175/JCLI-D-17-0769.1, 2018.

Amrhein, D. E., Hakim, G. J., and Parsons, L. A.: Quantifying Structural Uncertainty in Paleoclimate Data Assimilation With an Application to the Last Millennium, Geophys. Res. Lett., 47, https://doi.org/10.1029/2020GL090485, 2020.

Anderson, J. L.: A Quantile-Conserving Ensemble Filter Framework. Part I: Updating an Observed Variable, Monthly Weather Review, 150, 1061–1074, https://doi.org/10.1175/MWR-D-21-0229.1, 2022.

Andrews, T.: Using an AGCM to Diagnose Historical Effective Radiative Forcing and Mechanisms of Recent Decadal Climate Change, J. Climate, 27, 1193–1209, https://doi.org/10.1175/JCLI-D-13-00336.1, 2014.

Andrews, T. and Webb, M. J.: The Dependence of Global Cloud and Lapse Rate Feedbacks on the Spatial Structure of Tropical Pacific Warming, J. Climate, 31, 641–654, https://doi.org/10.1175/JCLI-D-17-0087.1, 2018.

Andrews, T., Gregory, J. M., and Webb, M. J.: The Dependence of Radiative Forcing and Feedback on Evolving Patterns of Surface Temperature Change in Climate Models, J. Climate, 28, 1630–1648, https://doi.org/10.1175/JCLI-D-14-00545.1, 2015.

Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., Webb, M. J., Armour, K. C., Forster, P. M., and Titchner, H.: Accounting for Changing Temperature Patterns Increases Historical Estimates of Climate Sensitivity, Geophys. Res. Lett., 45, 8490–8499, https://doi.org/10.1029/2018GL078887, 2018.

Andrews, T., Bodas-Salcedo, A., Gregory, J. M., Dong, Y., Armour, K. C., Paynter, D., Lin, P., Modak, A., Mauritsen, T., Cole, J. N. S., Medeiros, B., Benedict, J. J., Douville, H., Roehrig, R., Koshiro, T., Kawai, H., Ogura, T., Dufresne, J., Allan, R. P., and Liu, C.: On the Effect of Historical SST Patterns on Radiative Feedback, J. Geophys. Res.: Atmospheres, 127, https://doi.org/10.1029/2022JD036675, 2022.

162

Annan, J. D., Hargreaves, J. C., and Mauritsen, T.: A new global surface temperature recon-struction for the Last Glacial Maximum, Climate of the Past, 18, 1883–1896, https://doi.org/10.5194/cp-18-1883-2022, 2022.

Annan, J. D., Hargreaves, J. C., Mauritsen, T., McClymont, E., and Ho, S. L.: Can we reliably reconstruct the mid-Pliocene Warm Period with sparse data and uncertain models?, Climate of the Past, 20, 1989–1999, https://doi.org/10.5194/cp-20-1989-2024, 2024.

Argus, D. F., Peltier, W. R., Drummond, R., and Moore, A. W.: The Antarctica component of postglacial rebound model ICE-6G_C (VM5a) based on GPS positioning, exposure age dat-ing of ice thicknesses, and relative sea level histories, Geophysical Journal International, 198, https://doi.org/10.1093/gji/ggu140, 2014.

Armour, K. C., Bitz, C. M., and Roe, G. H.: Time-Varying Climate Sensitivity from Regional Feedbacks, J. Climate, 26, 4518–4534, https://doi.org/10.1175/JCLI-D-12-00544.1, 2013.

Armour, K. C., Marshall, J., Scott, J. R., Donohoe, A., and Newsom, E. R.: Southern Ocean warming delayed by circumpolar upwelling and equatorward transport, Nature Geoscience, 9, 549–554, https://doi.org/10.1038/ngeo2731, 2016.

Armour, K. C., Proistosescu, C., Dong, Y., Hahn, L. C., Blanchard-Wrigglesworth, E., Pauling, A. G., Jnglin Wills, R. C., Andrews, T., Stuecker, M. F., Po-Chedley, S., Mitevski, I., Forster, P. M., and Gregory, J. M.: Sea-surface temperature pattern effects have slowed global warming and biased warming-based constraints on climate sensitivity, Proc. Natl. Acad. Sci. (USA), 121, e2312093 121, https://doi.org/10.1073/pnas.2312093121, 2024.

Bach, E. and Ghil, M.: A Multi-Model Ensemble Kalman Filter for Data Assimilation and Forecasting, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2022MS003123, 2023.

Banerjee, A., Fyfe, J. C., Polvani, L. M., Waugh, D., and Chang, K.-L.: A pause in Southern Hemi-sphere circulation trends due to the Montreal Protocol, Nature, 579, 544–548, https://doi.org/10.1038/s41586-020-2120-4, 2020.

Barsugli, J. J. and Battisti, D. S.: The basic effects of atmosphere-ocean thermal coupling on midlatitude variability, Journal of the Atmospheric Sciences, https://doi.org/10.1175/1520-0469(1998)055⟨0477:TBEOAO⟩2.0.CO;2, 1998.

Bartlein, P. J., Harrison, S. P., Brewer, S., Connor, S., Davis, B. A. S., Gajewski, K., Guiot, J., Harrison-Prentice, T. I., Henderson, A., Peyron, O., Prentice, I. C., Scholze, M., Seppä, H., Shuman, B., Sugita, S., Thompson, R. S., Viau, A. E., Williams, J., and Wu, H.: Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis, Climate Dynamics, 37, 775–802, https://doi.org/10.1007/s00382-010-0904-1, 2011.

Battisti, D. S., Vimont, D. J., and Kirtman, B. P.: 100 Years of Progress in Understanding the Dynamics of Coupled Atmosphere–Ocean Variability, Meteorological Monographs, 59, 1–8, https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0025.1, 2019.

Bitz, C. M. and Polvani, L. M.: Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model, Geophys. Res. Lett., 39, https://doi.org/10.1029/2012GL053393, 2012.

Bitz, C. M., Shell, K. M., Gent, P. R., Bailey, D. A., Danabasoglu, G., Armour, K. C., Holland, M. M., and Kiehl, J. T.: Climate Sensitivity of the Community Climate System Model, Version 4, J. Climate, 25, 3053–3070, https://doi.org/10.1175/JCLI-D-11-00290.1, 2012.

Bjerknes, J.: Atmospheric Teleconnections from the Equatorial Pacific, Monthly Weather Review, 97, 163–172, https://doi.org/10.1175/1520-0493(1969)097⟨0163:ATFTEP⟩2.3.CO;2, 1969.

Blanchard-Wrigglesworth, E., Roach, L. A., Donohoe, A., and Ding, Q.: Impact of Winds and Southern Ocean SSTs on Antarctic Sea Ice Trends and Variability, J. Climate, 34, 949–965, https://doi.org/10.1175/JCLI-D-20-0386.1, 2021.

Bloch-Johnson, J., Rugenstein, M., Stolpe, M. B., Rohrschneider, T., Zheng, Y., and Gregory, J. M.: Climate Sensitivity Increases Under Higher CO¡sub¿2¡/sub¿ Levels Due to Feedback Temperature Dependence, Geophys. Res. Lett., 48, e2020GL089 074, https://doi.org/10.1029/2020GL089074, 2021.

164

Bloch-Johnson, J., Rugenstein, M. A. A., Alessi, M. J., Proistosescu, C., Zhao, M., Zhang, B., Williams, A. I. L., Gregory, J. M., Cole, J., Dong, Y., Duffy, M. L., Kang, S. M., and Zhou, C.: The Green's Function Model Intercomparison Project (GFMIP) Protocol, Journal of Advances in Modeling Earth Systems, 16, https://doi.org/10.1029/2023MS003700, 2024.

Bonan, D. B., Dörr, J., Wills, R. C. J., Thompson, A. F., and Årthun, M.: Sources of low-frequency variability in observed Antarctic sea ice, The Cryosphere, 18, 2141–2159, https://doi.org/10.5194/tc-18-2141-2024, 2024.

Braconnot, P. and Kageyama, M.: Shortwave forcing and feedbacks in Last Glacial Maximum and Mid-Holocene PMIP3 simulations, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373, 20140 424, https://doi.org/10.1098/rsta.2014.0424, 2015.

Brady, E. C., Otto-Bliesner, B. L., Kay, J. E., and Rosenbloom, N.: Sensitivity to Glacial Forcing in the CCSM4, J. Climate, 26, 1901–1925, https://doi.org/10.1175/JCLI-D-11-00416.1, 2013.

Brennan, M. K. and Hakim, G. J.: Reconstructing Arctic Sea Ice over the Common Era Using Data Assimilation, J. Climate, 35, 1231–1247, https://doi.org/10.1175/JCLI-D-21-0099.1, 2022.

Brennan, M. K., Hakim, G. J., and Blanchard-Wrigglesworth, E.: Monthly Arctic Sea-Ice Prediction With a Linear Inverse Model, Geophys. Res. Lett., 50, e2022GL101 656, https://doi.org/10.1029/2022GL101656, 2023.

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., and Bladé, I.: The effective number of spatial degrees of freedom of a time-varying field, J. Climate, https://doi.org/10.1175/1520-0442(1999)012⟨1990:TENOSD⟩2.0.CO;2, 1999.

Brierley, C., Burls, N., Ravelo, C., and Fedorov, A.: Pliocene warmth and gradients, Nature Geoscience, 8, 419–420, https://doi.org/10.1038/ngeo2444, 2015.

Brönnimann, S.: Early twentieth-century warming, Nature Geoscience, 2, 735–736, https://doi.org/10.1038/ngeo670, 2009.

Brönnimann, S., Allan, R., Ashcroft, L., Baer, S., Barriendos, M., Brázdil, R., Brugnara, Y., Brunet, M., Brunetti, M., Chimani, B., Cornes, R., Domínguez-Castro, F., Filipiak, J., Founda, D., Herrera, R. G., Gergis, J., Grab, S., Hannak, L., Huhtamaa, H., Jacobsen, K. S., Jones, P., Jourdain, S., Kiss, A., Lin, K. E., Lorrey, A., Lundstad, E., Luterbacher, J., Mauelshagen, F., Maugeri, M., Maughan, N., Moberg, A., Neukom, R., Nicholson, S., Noone, S., Nordli, Ø., Ólafsdóttir, K. B., Pearce, P. R., Pfister, L., Pribyl, K., Przybylak, R., Pudmenzky, C., Rasol, D., Reichenbach, D., Řezníčková, L., Rodrigo, F. S., Rohr, C., Skrynyk, O., Slonosky, V., Thorne, P., Valente, M. A., Vaquero, J. M., Westcottt, N. E., Williamson, F., and Wyszyński, P.: Unlocking Pre-1850 Instrumental Meteorological Records: A Global Inventory, Bull. Amer. Meteor. Soc., 100, ES389–ES413, https://doi.org/10.1175/BAMS-D-19-0040.1, 2019.

Brönnimann, S., Brugnara, Y., and Wilkinson, C.: Early 20th century Southern Hemisphere cooling, Climate of the Past, 20, 757–767, https://doi.org/10.5194/cp-20-757-2024, 2024.

Buckley, E. M., Horvat, C., and Yoosiri, P.: Sea Ice Concentration Estimates from ICESat-2 Linear Ice Fraction. Part 1: Multi-sensor Comparison of Sea Ice Concentration Products, https://doi.org/10.5194/egusphere-2024-3861, 2024.

Burgers, G., Jan van Leeuwen, P., and Evensen, G.: Analysis Scheme in the Ensemble Kalman Filter, Monthly Weather Review, 126, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126⟨1719:ASITEK⟩2.0.CO;2, 1998.

Burke, K. D., Williams, J. W., Chandler, M. A., Haywood, A. M., Lunt, D. J., and Otto-Bliesner, B. L.: Pliocene and Eocene provide best analogs for near-future climates, Proc. Natl. Acad. Sci. (USA), 115, 13 288–13 293, https://doi.org/10.1073/pnas.1809600115, 2018.

Burls, N. and Sagoo, N.: Increasingly Sophisticated Climate Models Need the Out-Of-Sample Tests Paleoclimates Provide, Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2022MS003389, 2022.

Burls, N. J. and Fedorov, A. V.: What Controls the Mean East–West Sea Surface Temperature Gradient in the Equatorial Pacific: The Role of Cloud Albedo, J. Climate, 27, 2757–2778, https://doi.org/10.1175/JCLI-D-13-00255.1, 2014.

Burls, N. J. and Fedorov, A. V.: Wetter subtropics in a warmer world: Contrasting past and future hydrological cycles, Proc. Natl. Acad. Sci. (USA), 114, 12 888–12 893, https://doi.org/10.1073/pnas.1703421114, 2017.

Burton, L. E., Haywood, A. M., Tindall, J. C., Dolan, A. M., Hill, D. J., Abe-Ouchi, A., Chan, W.-L., Chandan, D., Feng, R., Hunter, S. J., Li, X., Peltier, W. R., Tan, N., Stepanek, C., and Zhang, Z.: On the climatic influence of CO2 forcing in the Pliocene, Climate of the Past, 19, 747–764, https://doi.org/10.5194/cp-19-747-2023, 2023.

Caballero, R. and Huber, M.: State-dependent climate sensitivity in past warm climates and its implications for future climate projections, Proc. Natl. Acad. Sci. (USA), 110, 14 162–14 167, https://doi.org/10.1073/pnas.1303365110, 2013.

Cai, W., Santoso, A., Collins, M., Dewitte, B., Karamperidou, C., Kug, J.-S., Lengaigne, M., McPhaden, M. J., Stuecker, M. F., Taschetto, A. S., Timmermann, A., Wu, L., Yeh, S.-W., Wang, G., Ng, B., Jia, F., Yang, Y., Ying, J., Zheng, X.-T., Bayr, T., Brown, J. R., Capotondi, A., Cobb, K. M., Gan, B., Geng, T., Ham, Y.-G., Jin, F.-F., Jo, H.-S., Li, X., Lin, X., McGregor, S., Park, J.-H., Stein, K., Yang, K., Zhang, L., and Zhong, W.: Changing El Niño–Southern Oscillation in a warming climate, Nature Reviews Earth & Environment, 2, 628–644, https://doi.org/10.1038/s43017-021-00199-z, 2021.

Cai, W., Ng, B., Geng, T., Jia, F., Wu, L., Wang, G., Liu, Y., Gan, B., Yang, K., Santoso, A., Lin, X., Li, Z., Liu, Y., Yang, Y., Jin, F.-F., Collins, M., and McPhaden, M. J.: Anthropogenic impacts on twentieth-century ENSO variability changes, Nature Reviews Earth & Environment, 4, 407–418, https://doi.org/10.1038/s43017-023-00427-8, 2023.

Callahan, C. W., Chen, C., Rugenstein, M., Bloch-Johnson, J., Yang, S., and Moyer, E. J.: Robust decrease in El Niño/Southern Oscillation amplitude under long-term warming, Nature Climate Change, 11, 752–757, https://doi.org/10.1038/s41558-021-01099-2, 2021.

Cane, M. A., Clement, A. C., Kaplan, A., Kushnir, Y., Pozdnyakov, D., Seager, R., Zebiak, S. E., and Murtugudde, R.: Twentieth-Century Sea Surface Temperature Trends, Science, 275, 957–960, https://doi.org/10.1126/science.275.5302.957, 1997.

Capotondi, A., Wittenberg, A. T., Newman, M., Di Lorenzo, E., Yu, J.-Y., Braconnot, P., Cole, J., Dewitte, B., Giese, B., Guilyardi, E., Jin, F.-F., Karnauskas, K., Kirtman, B., Lee, T., Schneider, N., Xue, Y., and Yeh, S.-W.: Understanding ENSO Diversity, Bull. Amer. Meteor. Soc., 96, 921–938, https://doi.org/10.1175/BAMS-D-13-00117.1, 2015.

Capotondi, A., McGregor, S., McPhaden, M. J., Cravatte, S., Holbrook, N. J., Imada, Y., Sanchez, S. C., Sprintall, J., Stuecker, M. F., Ummenhofer, C. C., Zeller, M., Farneti, R., Graffino, G., Hu, S., Karnauskas, K. B., Kosaka, Y., Kucharski, F., Mayer, M., Qiu, B., Santoso, A., Taschetto, A. S., Wang, F., Zhang, X., Holmes, R. M., Luo, J.-J., Maher, N., Martinez-Villalobos, C., Meehl, G. A., Naha, R., Schneider, N., Stevenson, S., Sullivan, A., van Rensch, P., and Xu, T.: Mechanisms of tropical Pacific decadal variability, Nature Reviews Earth & Environment, 4, 754–769, https://doi.org/10.1038/s43017-023-00486-x, 2023.

Ceppi, P. and Gregory, J. M.: Relationship of tropospheric stability to climate sensitivity and Earth's observed radiation budget, Proc. Natl. Acad. Sci. (USA), 114, 13 126–13 131, https://doi.org/10.1073/pnas.1714308114, 2017.

Chadwick, R., Good, P., Andrews, T., and Martin, G.: Surface warming patterns drive tropical rainfall pattern responses to CO ¡sub¿2¡/sub¿ forcing on all timescales, Geophys. Res. Lett., 41, 610–615, https://doi.org/10.1002/2013GL058504, 2014.

Chan, D. and Huybers, P.: Systematic Differences in Bucket Sea Surface Temperature Measurements among Nations Identified Using a Linear-Mixed-Effect Method, J. Climate, 32, 2569–2589, https://doi.org/10.1175/JCLI-D-18-0562.1, 2019.

Chan, D., Kent, E. C., Berry, D. I., and Huybers, P.: Correcting datasets leads to more homogeneous early-twentieth-century sea surface warming, Nature, 571, 393–397, https://doi.org/10.1038/s41586-019-1349-2, 2019.

Chan, D., Gebbie, G., and Huybers, P.: Global and Regional Discrepancies between Early-Twentieth-Century Coastal Air and Sea Surface Temperature Detected by a Coupled Energy-Balance Analysis, J. Climate, 36, 2205–2220, https://doi.org/10.1175/JCLI-D-22-0569.1, 2023.

168

Chan, D., Gebbie, G., Huybers, P., and Kent, E. C.: A Dynamically Consistent ENsemble of Temperature at the Earth surface since 1850 from the DCENT dataset, Scientific Data, 11, 953, https://doi.org/10.1038/s41597-024-03742-x, 2024.

Charney, J. G., Arakawa, A., Baker, D. J., Bolin, B., Dickinson, R. E., Goody, R. M., Leith, C. E., Stommel, H. M., and Wunsch, C. I.: Carbon dioxide and climate: a scientific assessment, National Academy of Sciences, Washington, DC, `http://scholar.google.com/scholar_lookup?&title=Carbon%20Dioxide%20and%20Climate%3A%20A%20Scientific%20Assessment&publication_year=1979&author=Charney%2CJ`, 1979.

Chemke, R., Ming, Y., and Yuval, J.: The intensification of winter mid-latitude storm tracks in the Southern Hemisphere, Nature Climate Change, 12, 553–557, https://doi.org/10.1038/s41558-022-01368-8, 2022.

Chung, E.-S., Timmermann, A., Soden, B. J., Ha, K.-J., Shi, L., and John, V. O.: Reconciling opposing Walker circulation trends in observations and model projections, Nature Climate Change, 9, 405–412, https://doi.org/10.1038/s41558-019-0446-4, 2019.

Clement, A. C., Seager, R., Cane, M. A., and Zebiak, S. E.: An Ocean Dynamical Thermostat, J. Climate, 9, 2190–2196, https://doi.org/10.1175/1520-0442(1996)009⟨2190:AODT⟩2.0.CO;2, 1996.

Coats, S. and Karnauskas, K. B.: Are Simulated and Observed Twentieth Century Tropical Pacific Sea Surface Temperature Trends Significant Relative to Internal Variability?, Geophys. Res. Lett., 44, 9928–9937, https://doi.org/10.1002/2017GL074622, 2017.

Coats, S., Smerdon, J. E., Stevenson, S., Fasullo, J. T., Otto-Bliesner, B., and Ault, T. R.: Paleoclimate Constraints on the Spatiotemporal Character of Past and Future Droughts, J. Climate, 33, 9883–9903, https://doi.org/10.1175/JCLI-D-20-0004.1, 2020.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D.,

and Worley, S. J.: The Twentieth Century Reanalysis Project, Quart. J. Roy. Meteor. Soc., 137, 1–28, https://doi.org/10.1002/qj.776, 2011.

Cook, B. I., Smerdon, J. E., Cook, E. R., Williams, A. P., Anchukaitis, K. J., Mankin, J. S., Allen, K., Andreu-Hayles, L., Ault, T. R., Belmecheri, S., Coats, S., Coulthard, B., Fosu, B., Grierson, P., Griffin, D., Herrera, D. A., Ionita, M., Lehner, F., Leland, C., Marvel, K., Morales, M. S., Mishra, V., Ngoma, J., Nguyen, H. T. T., O'Donnell, A., Palmer, J., Rao, M. P., Rodriguez-Caton, M., Seager, R., Stahle, D. W., Stevenson, S., Thapa, U. K., Varuolo-Clarke, A. M., and Wise, E. K.: Megadroughts in the Common Era and the Anthropocene, Nature Reviews Earth & Environment, 3, 741–757, https://doi.org/10.1038/s43017-022-00329-1, 2022.

Cook, K. H. and Held, I. M.: Stationary Waves of the Ice Age Climate, J. Climate, 1, 807–819, https://doi.org/10.1175/1520-0442(1988)001⟨0807:SWOTIA⟩2.0.CO;2, 1988.

Cooper, V. T., Roach, L. A., Thomson, J., Brenner, S. D., Smith, M. M., Meylan, M. H., and Bitz, C. M.: Wind waves in sea ice of the western Arctic and a global coupled wave-ice model, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 380, https://doi.org/10.1098/rsta.2021.0258, 2022.

Cooper, V. T., Armour, K. C., Hakim, G. J., Tierney, J. E., Osman, M. B., Proistosescu, C., Dong, Y., Burls, N. J., Andrews, T., Amrhein, D. E., Zhu, J., Dong, W., Ming, Y., and Chmielowiec, P.: Last Glacial Maximum pattern effects reduce climate sensitivity estimates, Science Advances, 10, 9461, https://doi.org/10.1126/sciadv.adk9461, 2024.

Cooper, V. T., Hakim, G. J., and Armour, K. C.: Monthly Sea-Surface Temperature, Sea Ice, and Sea-Level Pressure over 1850–2023 from Coupled Data Assimilation, Journal of Climate (in review), https://doi.org/10.31223/X5JH8K, 2025.

Cornes, R. C., Kent, E., Berry, D., and Kennedy, J. J.: CLASSnmat: A global night marine air temperature data set, 1880–2019, Geoscience Data Journal, 7, 170–184, https://doi.org/10.1002/gdj3.100, 2020.

Cowtan, K. and Way, R. G.: Coverage bias in the HadCRUT4 temperature series and its impact

on recent temperature trends, Quart. J. Roy. Meteor. Soc., 140, 1935–1944, https://doi.org/10.1002/qj.2297, 2014.

Cowtan, K., Rohde, R., and Hausfather, Z.: Evaluating biases in sea surface temperature records using coastal weather stations, Quart. J. Roy. Meteor. Soc., 144, 670–681, https://doi.org/10.1002/qj.3235, 2018.

Cram, T. A., Compo, G. P., Yin, X., Allan, R. J., McColl, C., Vose, R. S., Whitaker, J. S., Matsui, N., Ashcroft, L., Auchmann, R., Bessemoulin, P., Brandsma, T., Brohan, P., Brunet, M., Comeaux, J., Crouthamel, R., Gleason, B. E., Groisman, P. Y., Hersbach, H., Jones, P. D., Jónsson, T., Jourdain, S., Kelly, G., Knapp, K. R., Kruger, A., Kubota, H., Lentini, G., Lorrey, A., Lott, N., Lubker, S. J., Luterbacher, J., Marshall, G. J., Maugeri, M., Mock, C. J., Mok, H. Y., Nordli, Ø., Rodwell, M. J., Ross, T. F., Schuster, D., Srnec, L., Valente, M. A., Vizi, Z., Wang, X. L., Westcott, N., Woollen, J. S., and Worley, S. J.: The International Surface Pressure Databank version 2, Geoscience Data Journal, 2, 31–46, https://doi.org/10.1002/gdj3.25, 2015.

Crucifix, M.: Does the Last Glacial Maximum constrain climate sensitivity?, Geophys. Res. Lett., 33, L18 701, https://doi.org/10.1029/2006GL027137, 2006.

Czaja, A., Frankignoul, C., Minobe, S., and Vannière, B.: Simulating the Midlatitude Atmospheric Circulation: What Might We Gain From High-Resolution Modeling of Air-Sea Interactions?, Current Climate Change Reports, 5, 390–406, https://doi.org/10.1007/s40641-019-00148-5, 2019.

Dalaiden, Q., Goosse, H., Rezsöhazy, J., and Thomas, E. R.: Reconstructing atmospheric circulation and sea-ice extent in the West Antarctic over the past 200 years using data assimilation, Climate Dynamics, 57, 3479–3503, https://doi.org/10.1007/s00382-021-05879-6, 2021.

Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore,

J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS001916, 2020.

de la Vega, E., Chalk, T. B., Wilson, P. A., Bysani, R. P., and Foster, G. L.: Atmospheric CO2 during the Mid-Piacenzian Warm Period and the M2 glaciation, Scientific Reports, 10, 11 002, https://doi.org/10.1038/s41598-020-67154-8, 2020.

Deser, C., Alexander, M. A., Xie, S.-P., and Phillips, A. S.: Sea Surface Temperature Variability: Patterns and Mechanisms, Annual Review of Marine Science, 2, 115–143, https://doi.org/10.1146/annurev-marine-120408-151453, 2010a.

Deser, C., Phillips, A. S., and Alexander, M. A.: Twentieth century tropical sea surface temperature trends revisited, Geophys. Res. Lett., 37, https://doi.org/10.1029/2010GL043321, 2010b.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Quart. J. Roy. Meteor. Soc., 131, 3385–3396, https://doi.org/10.1256/qj.05.108, 2005.

Deutsch, C., Berelson, W., Thunell, R., Weber, T., Tems, C., McManus, J., Crusius, J., Ito, T., Baumgartner, T., Ferreira, V., Mey, J., and van Geen, A.: Centennial changes in North Pacific anoxia linked to tropical trade winds, Science, 345, 665–668, https://doi.org/10.1126/science.1252332, 2014.

DiNezio, P. N. and Tierney, J. E.: The effect of sea level on glacial Indo-Pacific climate, Nature Geoscience, 6, 485–491, https://doi.org/10.1038/ngeo1823, 2013.

DiNezio, P. N., Clement, A. C., Vecchi, G. A., Soden, B. J., Kirtman, B. P., and Lee, S. K.: Climate Response of the Equatorial Pacific to Global Warming, J. Climate, 22, 4873–4892, https://doi.org/10.1175/2009JCLI2982.1, 2009.

DiNezio, P. N., Tierney, J. E., Otto-Bliesner, B. L., Timmermann, A., Bhattacharya, T., Rosenbloom, N., and Brady, E.: Glacial changes in tropical climate amplified by the Indian Ocean, Science Advances, 4, https://doi.org/10.1126/sciadv.aat9658, 2018.

172

Doddridge, E. W. and Marshall, J.: Modulation of the Seasonal Cycle of Antarctic Sea Ice Extent Related to the Southern Annular Mode, Geophys. Res. Lett., 44, 9761–9768, https://doi.org/10.1002/2017GL074319, 2017.

Dong, Y., Proistosescu, C., Armour, K. C., and Battisti, D. S.: Attributing Historical and Future Evolution of Radiative Feedbacks to Regional Warming Patterns using a Green's Function Approach: The Preeminence of the Western Pacific, J. Climate, 32, 5471–5491, https://doi.org/10.1175/JCLI-D-18-0843.1, 2019.

Dong, Y., Armour, K. C., Zelinka, M. D., Proistosescu, C., Battisti, D. S., Zhou, C., and Andrews, T.: Intermodel Spread in the Pattern Effect and Its Contribution to Climate Sensitivity in CMIP5 and CMIP6 Models, J. Climate, 33, 7755–7775, https://doi.org/10.1175/JCLI-D-19-1011.1, 2020.

Dong, Y., Armour, K. C., Proistosescu, C., Andrews, T., Battisti, D. S., Forster, P. M., Paynter, D., Smith, C. J., and Shiogama, H.: Biased Estimates of Equilibrium Climate Sensitivity and Transient Climate Response Derived From Historical CMIP6 Simulations, Geophys. Res. Lett., 48, https://doi.org/10.1029/2021GL095778, 2021.

Dong, Y., Armour, K. C., Battisti, D. S., and Blanchard-Wrigglesworth, E.: Two-Way Teleconnections between the Southern Ocean and the Tropical Pacific via a Dynamic Feedback, J. Climate, 35, 6267–6282, https://doi.org/10.1175/JCLI-D-22-0080.1, 2022a.

Dong, Y., Pauling, A. G., Sadai, S., and Armour, K. C.: Antarctic Ice-Sheet Meltwater Reduces Transient Warming and Climate Sensitivity Through the Sea-Surface Temperature Pattern Effect, Geophys. Res. Lett., 49, https://doi.org/10.1029/2022GL101249, 2022b.

Dong, Y., Polvani, L. M., and Bonan, D. B.: Recent Multi-Decadal Southern Ocean Surface Cooling Unlikely Caused by Southern Annular Mode Trends, Geophys. Res. Lett., 50, https://doi.org/10.1029/2023GL106142, 2023.

Dörr, J. S., Bonan, D. B., Årthun, M., Svendsen, L., and Wills, R. C. J.: Forced and internal components of observed Arctic sea-ice changes, The Cryosphere, 17, 4133–4153, https://doi.org/10.5194/tc-17-4133-2023, 2023.

Döscher, R., Acosta, M., Alessandri, A., Anthoni, P., Arsouze, T., Bergman, T., Bernardello, R., Boussetta, S., Caron, L.-P., Carver, G., Castrillo, M., Catalano, F., Cvijanovic, I., Davini, P., Dekker, E., Doblas-Reyes, F. J., Docquier, D., Echevarria, P., Fladrich, U., Fuentes-Franco, R., Gröger, M., v. Hardenberg, J., Hieronymus, J., Karami, M. P., Keskinen, J.-P., Koenigk, T., Makkonen, R., Massonnet, F., Ménégoz, M., Miller, P. A., Moreno-Chamarro, E., Nieradzik, L., van Noije, T., Nolan, P., O'Donnell, D., Ollinaho, P., van den Oord, G., Ortega, P., Prims, O. T., Ramos, A., Reerink, T., Rousset, C., Ruprich-Robert, Y., Le Sager, P., Schmith, T., Schrödner, R., Serva, F., Sicardi, V., Sloth Madsen, M., Smith, B., Tian, T., Tourigny, E., Uotila, P., Vancoppenolle, M., Wang, S., Wårlind, D., Willén, U., Wyser, K., Yang, S., Yepes-Arbós, X., and Zhang, Q.: The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6, Geoscientific Model Development, 15, 2973–3020, https://doi.org/10.5194/gmd-15-2973-2022, 2022.

Dowsett, H., Dolan, A., Rowley, D., Moucha, R., Forte, A. M., Mitrovica, J. X., Pound, M., Salzmann, U., Robinson, M., Chandler, M., Foley, K., and Haywood, A.: The PRISM4 (mid-Piacenzian) paleoenvironmental reconstruction, Climate of the Past, 12, 1519–1538, https://doi.org/10.5194/cp-12-1519-2016, 2016.

Dunne, J. P., Horowitz, L. W., Adcroft, A. J., Ginoux, P., Held, I. M., John, J. G., Krasting, J. P., Malyshev, S., Naik, V., Paulot, F., Shevliakova, E., Stock, C. A., Zadeh, N., Balaji, V., Blanton, C., Dunne, K. A., Dupuis, C., Durachta, J., Dussin, R., Gauthier, P. P. G., Griffies, S. M., Guo, H., Hallberg, R. W., Harrison, M., He, J., Hurlin, W., McHugh, C., Menzel, R., Milly, P. C. D., Nikonov, S., Paynter, D. J., Ploshay, J., Radhakrishnan, A., Rand, K., Reichl, B. G., Robinson, T., Schwarzkopf, D. M., Sentman, L. T., Underwood, S., Vahlenkamp, H., Winton, M., Wittenberg, A. T., Wyman, B., Zeng, Y., and Zhao, M.: The GFDL Earth System Model Version 4.1 (GFDL-ESM 4.1): Overall Coupled Model Description and Simulation Characteristics, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/10.1029/2019MS002015, 2020.

Dutton, A., Carlson, A. E., Long, A. J., Milne, G. A., Clark, P. U., DeConto, R., Horton, B. P., Rahmstorf, S., and Raymo, M. E.: Sea-level rise due to polar ice-sheet mass loss during past warm periods, Science, 349, https://doi.org/10.1126/science.aaa4019, 2015.

174

Dvorak, M., Armour, K., Feng, R., Cooper, V., Zhu, J., Proistosescu, C., and Burls, N.: Mid-Pliocene climate forcing, sea-surface temperature patterns, and implications for modern-day climate sensitivity, Journal of Climate, https://doi.org/10.1175/JCLI-D-24-0410.1, 2025.

Edinburgh, T. and Day, J. J.: Estimating the extent of Antarctic summer sea ice during the Heroic Age of Antarctic Exploration, The Cryosphere, 10, 2721–2730, https://doi.org/10.5194/tc-10-2721-2016, 2016.

Eisenman, I. and Armour, K. C.: The radiative feedback continuum from Snowball Earth to an ice-free hothouse, Nature Communications, 15, 6582, https://doi.org/10.1038/s41467-024-50406-w, 2024.

Eisenman, I., Meier, W. N., and Norris, J. R.: A spurious jump in the satellite record: has Antarctic sea ice expansion been overestimated?, The Cryosphere, 8, 1289–1296, https://doi.org/10.5194/tc-8-1289-2014, 2014.

Emile-Geay, J., Hakim, G. J., Viens, F., Zhu, F., and Amrhein, D. E.: Temporal Comparisons Involving Paleoclimate Data Assimilation: Challenges & Remedies, J. Climate, -1, https://doi.org/10.1175/JCLI-D-24-0101.1, 2024.

England, M. R., Polvani, L. M., Smith, K. L., Landrum, L., and Holland, M. M.: Robust response of the Amundsen Sea Low to stratospheric ozone depletion, Geophys. Res. Lett., 43, 8207–8213, https://doi.org/10.1002/2016GL070055, 2016.

England, M. R., Eisenman, I., Lutsko, N. J., and Wagner, T. J.: The Recent Emergence of Arctic Amplification, Geophys. Res. Lett., 48, https://doi.org/10.1029/2021GL094086, 2021.

Erfani, E. and Burls, N. J.: The Strength of Low-Cloud Feedbacks and Tropical Climate: A CESM Sensitivity Study, J. Climate, 32, 2497–2516, https://doi.org/10.1175/JCLI-D-18-0551.1, 2019.

Espinosa, Z. I. and Zelinka, M. D.: The Shortwave Cloud-SST Feedback Amplifies Multi-Decadal Pacific Sea Surface Temperature Trends: Implications for Observed Cooling, Geophys. Res. Lett., 51, https://doi.org/10.1029/2024GL111039, 2024.

Espinosa, Z. I., Blanchard-Wrigglesworth, E., and Bitz, C. M.: Understanding the drivers and predictability of record low Antarctic sea ice in austral winter 2023, Communications Earth & Environment, 5, 723, https://doi.org/10.1038/s43247-024-01772-2, 2024.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research, 99, https://doi.org/10.1029/94jc00572, 1994.

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

Falster, G., Konecky, B., Coats, S., and Stevenson, S.: Forced changes in the Pacific Walker circulation over the past millennium, Nature, https://doi.org/10.1038/s41586-023-06447-0, 2023.

Fan, T., Deser, C., and Schneider, D. P.: Recent Antarctic sea ice trends in the context of Southern Ocean surface climate variations since 1950, Geophys. Res. Lett., 41, 2419–2426, https://doi.org/10.1002/2014GL059239, 2014.

Fedorov, A. V., Dekens, P. S., McCarthy, M., Ravelo, A. C., deMenocal, P. B., Barreiro, M., Pacanowski, R. C., and Philander, S. G.: The Pliocene Paradox (Mechanisms for a Permanent El Nino), Science, 312, 1485–1489, https://doi.org/10.1126/science.1122666, 2006.

Feng, R., Bhattacharya, T., Otto-Bliesner, B. L., Brady, E. C., Haywood, A. M., Tindall, J. C., Hunter, S. J., Abe-Ouchi, A., Chan, W.-L., Kageyama, M., Contoux, C., Guo, C., Li, X., Lohmann, G., Stepanek, C., Tan, N., Zhang, Q., Zhang, Z., Han, Z., Williams, C. J. R., Lunt, D. J., Dowsett, H. J., Chandan, D., and Peltier, W. R.: Past terrestrial hydroclimate sensitivity controlled by Earth system feedbacks, Nature Communications, 13, 1306, https://doi.org/10.1038/s41467-022-28814-7, 2022.

Flannaghan, T. J., Fueglistaler, S., Held, I. M., Po-Chedley, S., Wyman, B., and Zhao, M.: Tropical temperature trends in Atmospheric General Circulation Model simulations and the impact of

176

uncertainties in observed SSTs, J. Geophys. Res.: Atmospheres, 119, 327–340, https://doi.org/10.1002/2014JD022365, 2014.

Fogt, R. L. and Connolly, C. J.: Extratropical Southern Hemisphere Synchronous Pressure Variability in the Early Twentieth Century, J. Climate, 34, 5795–5811, https://doi.org/10.1175/JCLI-D-20-0498.1, 2021.

Fogt, R. L. and Marshall, G. J.: The Southern Annular Mode: Variability, trends, and climate impacts across the Southern Hemisphere, WIREs Climate Change, 11, https://doi.org/10.1002/wcc.652, 2020.

Fogt, R. L., Perlwitz, J., Monaghan, A. J., Bromwich, D. H., Jones, J. M., and Marshall, G. J.: Historical SAM Variability. Part II: Twentieth-Century Variability and Trends from Reconstructions, Observations, and the IPCC AR4 Models*, J. Climate, 22, 5346–5365, https://doi.org/10.1175/2009JCLI2786.1, 2009.

Fogt, R. L., Schneider, D. P., Goergens, C. A., Jones, J. M., Clark, L. N., and Garberoglio, M. J.: Seasonal Antarctic pressure variability during the twentieth century from spatially complete reconstructions and CAM5 simulations, Climate Dynamics, 53, 1435–1452, https://doi.org/10.1007/S00382-019-04674-8/METRICS, 2019.

Fogt, R. L., Sleinkofer, A. M., Raphael, M. N., and Handcock, M. S.: A regime shift in seasonal total Antarctic sea ice extent in the twentieth century, Nature Climate Change, 12, 54–62, https://doi.org/10.1038/s41558-021-01254-9, 2022.

Fogt, R. L., Dalaiden, Q., and O'Connor, G. K.: A comparison of South Pacific Antarctic sea ice and atmospheric circulation reconstructions since 1900, Climate of the Past, 20, 53–76, https://doi.org/10.5194/cp-20-53-2024, 2024.

Ford, H. L., Burls, N. J., Jacobs, P., Jahn, A., Caballero-Gill, R. P., Hodell, D. A., and Fedorov, A. V.: Sustained mid-Pliocene warmth led to deep water formation in the North Pacific, Nature Geoscience, 15, 658–663, https://doi.org/10.1038/s41561-022-00978-3, 2022.

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D., Lunt, D., Mauritsen, T., Palmer, M., Watanabe, M., Wild, M., and Zhang, H.: 2021: The Earth's energy budget,

climate feedbacks, and climate sensitivity, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J., Maycock, T., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., chap. 7, Cambridge Univ. Press, Cambridge, UK and New York, NY, https://doi.org/10.1017/9781009157896.009, 2021.

Franke, J., Brönnimann, S., Bhend, J., and Brugnara, Y.: A monthly global paleo-reanalysis of the atmosphere from 1600 to 2005 for studying past climatic variations, Scientific Data, 4, 170 076, https://doi.org/10.1038/sdata.2017.76, 2017.

Freeman, E., Woodruff, S. D., Worley, S. J., Lubker, S. J., Kent, E. C., Angel, W. E., Berry, D. I., Brohan, P., Eastman, R., Gates, L., Gloeden, W., Ji, Z., Lawrimore, J., Rayner, N. A., Rosenhagen, G., and Smith, S. R.: ICOADS Release 3.0: a major update to the historical marine climate record, International Journal of Climatology, 37, 2211–2232, https://doi.org/10.1002/joc.4775, 2017.

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

Fueglistaler, S.: Observational Evidence for Two Modes of Coupling Between Sea Surface Temperatures, Tropospheric Temperature Profile, and Shortwave Cloud Radiative Effect in the Tropics, Geophys. Res. Lett., 46, 9890–9898, https://doi.org/10.1029/2019GL083990, 2019.

Fueglistaler, S. and Silvers, L.: The Peculiar Trajectory of Global Warming, J. Geophys. Res.: Atmospheres, 126, 1–15, https://doi.org/10.1029/2020JD033629, 2021.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M.,

178

Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), J. Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Giese, B. S., Seidel, H. F., Compo, G. P., and Sardeshmukh, P. D.: An ensemble of ocean reanalyses for 1815–2013 with sparse observational input, J. Geophys. Res.: Oceans, 121, 6891–6910, https://doi.org/10.1002/2016JC012079, 2016.

Gong, D. and Wang, S.: Definition of Antarctic Oscillation index, Geophys. Res. Lett., 26, 459–462, https://doi.org/10.1029/1999GL900003, 1999.

Goosse, H., Dalaiden, Q., Feba, F., Mezzina, B., and Fogt, R. L.: A drop in Antarctic sea ice extent at the end of the 1970s, Communications Earth & Environment, 5, 628, https://doi.org/10.1038/s43247-024-01793-x, 2024.

Grant, G. R., Naish, T. R., Dunbar, G. B., Stocchi, P., Kominz, M. A., Kamp, P. J. J., Tapia, C. A., McKay, R. M., Levy, R. H., and Patterson, M. O.: The amplitude and origin of sea-level variability during the Pliocene epoch, Nature, 574, 237–241, https://doi.org/10.1038/s41586-019-1619-z, 2019.

Gregory, J. M.: A new method for diagnosing radiative forcing and climate sensitivity, Geophys. Res. Lett., 31, L03 205, https://doi.org/10.1029/2003GL018747, 2004.

Gregory, J. M. and Andrews, T.: Variation in climate sensitivity and feedback parameters during the historical period, Geophys. Res. Lett., 43, 3911–3920, https://doi.org/10.1002/2016GL068406, 2016.

Gregory, J. M., Andrews, T., Ceppi, P., Mauritsen, T., and Webb, M. J.: How accurately can the climate sensitivity to $\hbox{CO}_{2}$ be estimated from historical climate change?, Climate Dynamics, 54, 129–157, https://doi.org/10.1007/s00382-019-04991-y, 2020.

Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., Steiger, N., and Perkins, W. A.: The last millennium climate reanalysis project: Framework and first results, J. Geophys. Res.: Atmospheres, 121, 6745–6764, https://doi.org/10.1002/2016JD024751, 2016.

Hakim, G. J., Snyder, C., Penny, S. G., and Newman, M.: Subseasonal Forecast Skill Improvement From Strongly Coupled Data Assimilation With a Linear Inverse Model, Geophys. Res. Lett., 49, https://doi.org/10.1029/2022GL097996, 2022.

Hansen, J., Sato, M., Ruedy, R., Nazarenko, L., Lacis, A., Schmidt, G. A., Russell, G., Aleinov, I., Bauer, M., Bauer, S., Bell, N., Cairns, B., Canuto, V., Chandler, M., Cheng, Y., Del Genio, A., Faluvegi, G., Fleming, E., Friend, A., Hall, T., Jackman, C., Kelley, M., Kiang, N., Koch, D., Lean, J., Lerner, J., Lo, K., Menon, S., Miller, R., Minnis, P., Novakov, T., Oinas, V., Perlwitz, J., Perlwitz, J., Rind, D., Romanou, A., Shindell, D., Stone, P., Sun, S., Tausnev, N., Thresher, D., Wielicki, B., Wong, T., Yao, M., and Zhang, S.: Efficacy of climate forcings, J. Geophys. Res.: Atmospheres, 110, 1–45, https://doi.org/10.1029/2005JD005776, 2005.

Hansen, J. E., Sato, M., Simons, L., Nazarenko, L. S., Sangha, I., Kharecha, P., Zachos, J. C., von Schuckmann, K., Loeb, N. G., Osman, M. B., Jin, Q., Tselioudis, G., Jeong, E., Lacis, A., Ruedy, R., Russell, G., Cao, J., and Li, J.: Global warming in the pipeline, Oxford Open Climate Change, 3, https://doi.org/10.1093/oxfclm/kgad008, 2023.

Hartmann, D. L.: The Antarctic ozone hole and the pattern effect on climate sensitivity, Proc. Natl. Acad. Sci. (USA), 119, https://doi.org/10.1073/pnas.2207889119, 2022.

Hartmann, D. L. and Dygert, B. D.: Global Radiative Convective Equilibrium With a Slab Ocean: SST Contrast, Sensitivity and Circulation, J. Geophys. Res.: Atmospheres, 127, e2021JD036 400, https://doi.org/10.1029/2021JD036400, 2022.

Hasselmann, K.: Stochastic climate models: Part I. Theory, Tellus A: Dynamic Meteorology and Oceanography, 28, 473, https://doi.org/10.3402/tellusa.v28i6.11316, 1976.

Hastenrath, S. and Greischar, L.: Circulation mechanisms related to northeast Brazil rainfall anomalies, J. Geophys. Res.: Atmospheres, 98, 5093–5102, https://doi.org/10.1029/92JD02646, 1993.

Haumann, F. A., Gruber, N., and Münnich, M.: Sea-Ice Induced Southern Ocean Subsurface Warming and Surface Cooling in a Warming Climate, AGU Advances, 1, https://doi.org/10.1029/2019AV000132, 2020.

180

Hausfather, Z., Cowtan, K., Clarke, D. C., Jacobs, P., Richardson, M., and Rohde, R.: Assessing recent warming using instrumentally homogeneous sea surface temperature records, Science Advances, 3, https://doi.org/10.1126/SCIADV.1601207/SUPPL{_}FILE/1601207{_}SM.PDF, 2017.

Haywood, A., Tindall, J., Burton, L., Chandler, M., Dolan, A., Dowsett, H., Feng, R., Fletcher, T., Foley, K., Hill, D., Hunter, S., Otto-Bliesner, B., Lunt, D., Robinson, M., and Salzmann, U.: Pliocene Model Intercomparison Project Phase 3 (PlioMIP3) – Science plan and experimental design, Global and Planetary Change, 232, 104316, https://doi.org/10.1016/j.gloplacha.2023.104316, 2024.

Haywood, A. M., Tindall, J. C., Dowsett, H. J., Dolan, A. M., Foley, K. M., Hunter, S. J., Hill, D. J., Chan, W.-L., Abe-Ouchi, A., Stepanek, C., Lohmann, G., Chandan, D., Peltier, W. R., Tan, N., Contoux, C., Ramstein, G., Li, X., Zhang, Z., Guo, C., Nisancioglu, K. H., Zhang, Q., Li, Q., Kamae, Y., Chandler, M. A., Sohl, L. E., Otto-Bliesner, B. L., Feng, R., Brady, E. C., von der Heydt, A. S., Baatsen, M. L. J., and Lunt, D. J.: The Pliocene Model Intercomparison Project Phase 2: large-scale climate features and climate sensitivity, Climate of the Past, 16, 2095–2123, https://doi.org/10.5194/cp-16-2095-2020, 2020.

Heede, U. K. and Fedorov, A. V.: Colder Eastern Equatorial Pacific and Stronger Walker Circulation in the Early 21st Century: Separating the Forced Response to Global Warming From Natural Variability, Geophys. Res. Lett., 50, https://doi.org/10.1029/2022GL101020, 2023.

Hegerl, G. C., Brönnimann, S., Schurer, A., and Cowan, T.: The early 20th century warming: Anomalies, causes, and consequences, WIREs Climate Change, 9, https://doi.org/10.1002/wcc.522, 2018.

Hegerl, G. C., Brönnimann, S., Cowan, T., Friedman, A. R., Hawkins, E., Iles, C., Müller, W., Schurer, A., and Undorf, S.: Causes of climate change over the historical record, Environmental Research Letters, 14, 123006, https://doi.org/10.1088/1748-9326/AB4557, 2019.

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., Shevliakova, E., Winton, M., Zhao, M., Bushuk, M., Wittenberg, A. T., Wyman, B., Xiang, B., Zhang, R.,

Anderson, W., Balaji, V., Donner, L., Dunne, K., Durachta, J., Gauthier, P. P. G., Ginoux, P., Golaz, J., Griffies, S. M., Hallberg, R., Harris, L., Harrison, M., Hurlin, W., John, J., Lin, P., Lin, S., Malyshev, S., Menzel, R., Milly, P. C. D., Ming, Y., Naik, V., Paynter, D., Paulot, F., Ramaswamy, V., Reichl, B., Robinson, T., Rosati, A., Seman, C., Silvers, L. G., Underwood, S., and Zadeh, N.: Structure and Performance of GFDL's CM4.0 Climate Model, Journal of Advances in Modeling Earth Systems, 11, 3691–3727, https://doi.org/10.1029/2019MS001829, 2019.

Henderson, S. A., Vimont, D. J., and Newman, M.: The Critical Role of Non-Normality in Partitioning Tropical and Extratropical Contributions to PNA Growth, J. Climate, 33, 6273–6295, https://doi.org/10.1175/JCLI-D-19-0555.1, 2020.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quart. J. Roy. Meteor. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hirahara, S., Ishii, M., and Fukuda, Y.: Centennial-Scale Sea Surface Temperature Analysis and Its Uncertainty, J. Climate, 27, 57–75, https://doi.org/10.1175/JCLI-D-12-00837.1, 2014.

Hoerling, M., Eischeid, J., and Perlwitz, J.: Regional Precipitation Trends: Distinguishing Natural Variability from Anthropogenic Forcing, J. Climate, 23, 2131–2145, https://doi.org/10.1175/2009JCLI3420.1, 2010.

Hopcroft, P. O. and Valdes, P. J.: How well do simulated last glacial maximum tropical temperatures constrain equilibrium climate sensitivity?, Geophys. Res. Lett., 42, 5533–5539, https://doi.org/10.1002/2015GL064903, 2015.

Hopcroft, P. O., Ramstein, G., Pugh, T. A. M., Hunter, S. J., Murguia-Flores, F., Quiquet, A.,

Sun, Y., Tan, N., and Valdes, P. J.: Polar amplification of Pliocene climate by elevated trace gas radiative forcing, Proc. Natl. Acad. Sci. (USA), 117, 23 401–23 407, https://doi.org/10.1073/pnas.2002320117, 2020.

Houtekamer, P. L. and Mitchell, H. L.: Data Assimilation Using an Ensemble Kalman Filter Technique, Monthly Weather Review, 126, 796–811, https://doi.org/10.1175/1520-0493(1998)126⟨0796:DAUAEK⟩2.0.CO;2, 1998.

Houtekamer, P. L. and Zhang, F.: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation, Monthly Weather Review, 144, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1, 2016.

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S., and Zhang, H. M.: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons, J. Climate, 30, 8179–8205, https://doi.org/10.1175/JCLI-D-16-0836.1, 2017.

Huang, B., Liu, C., Ren, G., Zhang, H. M., and Zhang, L.: The Role of Buoy and Argo Observations in Two SST Analyses in the Global and Tropical Pacific Oceans, J. Climate, 32, 2517–2535, https://doi.org/10.1175/JCLI-D-18-0368.1, 2019.

Hurrell, J. W., Hack, J. J., Shea, D., Caron, J. M., and Rosinski, J.: A New Sea Surface Temperature and Sea Ice Boundary Dataset for the Community Atmosphere Model, J. Climate, 21, 5145–5153, https://doi.org/10.1175/2008JCLI2292.1, 2008.

Jiang, F., Seager, R., and Cane, M. A.: A climate change signal in the tropical Pacific emerges from decadal variability, Nature Communications, 15, 8291, https://doi.org/10.1038/s41467-024-52731-6, 2024.

Kadow, C., Hall, D. M., and Ulbrich, U.: Artificial intelligence reconstructs missing climate information, Nature Geoscience, 13, 408–413, https://doi.org/10.1038/s41561-020-0582-5, 2020.

Kageyama, M., Albani, S., Braconnot, P., Harrison, S. P., Hopcroft, P. O., Ivanovic, R. F., Lambert, F., Marti, O., Peltier, W. R., Peterschmitt, J.-Y., Roche, D. M., Tarasov, L., Zhang, X., Brady,

E. C., Haywood, A. M., LeGrande, A. N., Lunt, D. J., Mahowald, N. M., Mikolajewicz, U., Nisancioglu, K. H., Otto-Bliesner, B. L., Renssen, H., Tomas, R. A., Zhang, Q., Abe-Ouchi, A., Bartlein, P. J., Cao, J., Li, Q., Lohmann, G., Ohgaito, R., Shi, X., Volodin, E., Yoshida, K., Zhang, X., and Zheng, W.: 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, 4035–4055, https://doi.org/10.5194/gmd-10-4035-2017, 2017.

Kageyama, M., Abe-Ouchi, A., Obase, T., Ramstein, G., and Valdes, P. J.: Modeling the climate of the Last Glacial Maximum from PMIP1 to PMIP4, Past Global Changes Magazine, 29, 80–81, https://doi.org/10.22498/pages.29.2.80, 2021.

Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, 82, 35–45, https://doi.org/10.1115/1.3662552, 1960.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 2003.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, Bull. Amer. Meteor. Soc., 77, https://doi.org/10.1175/1520-0477(1996)077⟨0437:TNYRP⟩2.0.CO;2, 1996.

Kang, J. M., Shaw, T. A., Kang, S. M., Simpson, I. R., and Yu, Y.: Revisiting the reanalysis-model discrepancy in Southern Hemisphere winter storm track trends, npj Climate and Atmospheric Science, 7, 252, https://doi.org/10.1038/s41612-024-00801-3, 2024.

Kang, S. M. and Xie, S. P.: Dependence of Climate Response on Meridional Structure of External Thermal Forcing, J. Climate, 27, 5593–5600, https://doi.org/10.1175/JCLI-D-13-00622.1, 2014.

Kang, S. M., Ceppi, P., Yu, Y., and Kang, I.-S.: Recent global climate feedback controlled by Southern Ocean cooling, Nature Geoscience, https://doi.org/10.1038/s41561-023-01256-6, 2023a.

184

Kang, S. M., Shin, Y., Kim, H., Xie, S.-P., and Hu, S.: Disentangling the mechanisms of equatorial Pacific climate change, Science Advances, 9, https://doi.org/10.1126/sciadv.adf5059, 2023b.

Kang, S. M., Yu, Y., Deser, C., Zhang, X., Kang, I.-S., Lee, S.-S., Rodgers, K. B., and Ceppi, P.: Global impacts of recent Southern Ocean cooling, Proc. Natl. Acad. Sci. (USA), 120, https://doi.org/10.1073/pnas.2300881120, 2023c.

Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., and Rajagopalan, B.: Analyses of global sea surface temperature 1856–1991, J. Geophys. Res.: Oceans, 103, 18 567–18 589, https://doi.org/10.1029/97JC01736, 1998.

Kaplan, A., Kushnir, Y., and Cane, M. A.: Reduced Space Optimal Interpolation of Historical Marine Sea Level Pressure: 1854–1992*, J. Climate, 13, 2987–3002, https://doi.org/10.1175/1520-0442(2000)013⟨2987:RSOIOH⟩2.0.CO;2, 2000.

Karl, T. R., Arguez, A., Huang, B., Lawrimore, J. H., McMahon, J. R., Menne, M. J., Peterson, T. C., Vose, R. S., and Zhang, H. M.: Possible artifacts of data biases in the recent global surface warming hiatus, Science, 348, 1469–1472, https://doi.org/10.1126/SCIENCE.AAA5632/SUPPL{_}FILE/PAPV2.PDF, 2015.

Karnauskas, K. B.: Can we distinguish canonical El Niño from Modoki?, Geophys. Res. Lett., 40, 5246–5251, https://doi.org/10.1002/grl.51007, 2013.

Karnauskas, K. B., Seager, R., Kaplan, A., Kushnir, Y., and Cane, M. A.: Observed Strengthening of the Zonal Sea Surface Temperature Gradient across the Equatorial Pacific Ocean*, J. Climate, 22, 4316–4321, https://doi.org/10.1175/2009JCLI2936.1, 2009.

Kaufman, D. and Masson-Delmotte, V.: Opinion: Distribute paleoscience information across the next Intergovernmental Panel on Climate Change reports, Climate of the Past, 20, 2587–2594, https://doi.org/10.5194/cp-20-2587-2024, 2024.

Kennedy, J. J.: A review of uncertainty in in situ measurements and data sets of sea surface temperature, Rev. Geophys., 52, 1–32, https://doi.org/10.1002/2013RG000434, 2014.

Kennedy, J. J., Rayner, N. A., Atkinson, C. P., and Killick, R. E.: An Ensemble Data Set of Sea Surface Temperature Change From 1850: The Met Office Hadley Centre HadSST.4.0.0.0 Data Set, J. Geophys. Res.: Atmospheres, 124, 7719–7763, https://doi.org/10.1029/2018JD029867, 2019.

Kent, E. C. and Kennedy, J. J.: Historical Estimates of Surface Marine Temperatures, Annual Review of Marine Science, 13, 283–311, https://doi.org/10.1146/annurev-marine-042120-111807, 2021.

Kido, S., Richter, I., Tozuka, T., and Chang, P.: Understanding the interplay between ENSO and related tropical SST variability using linear inverse models, Climate Dynamics, 61, 1029–1048, https://doi.org/10.1007/s00382-022-06484-x, 2023.

King, J., Anchukaitis, K. J., Allen, K., Vance, T., and Hessl, A.: Trends and variability in the Southern Annular Mode over the Common Era, Nature Communications, 14, 2324, https://doi.org/10.1038/s41467-023-37643-1, 2023.

Knutti, R. and Hegerl, G. C.: The equilibrium sensitivity of the Earth's temperature to radiation changes, Nature Geoscience, 1, 735–743, https://doi.org/10.1038/ngeo337, 2008.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Kiyotoshi, T.: The JRA-55 reanalysis: General specifications and basic characteristics, Journal of the Meteorological Society of Japan, 93, https://doi.org/10.2151/jmsj.2015-001, 2015.

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

Kohyama, T. and Hartmann, D. L.: Antarctic Sea Ice Response to Weather and Climate Modes of Variability*, J. Climate, 29, 721–741, https://doi.org/10.1175/JCLI-D-15-0301.1, 2016.

Kohyama, T., Hartmann, D. L., and Battisti, D. S.: La Niña–like Mean-State Response to Global Warming and Potential Oceanic Roles, J. Climate, 30, 4207–4225, https://doi.org/10.1175/JCLI-D-16-0441.1, 2017.

186

Kok, J. F., Storelvmo, T., Karydis, V. A., Adebiyi, A. A., Mahowald, N. M., Evan, A. T., He, C., and Leung, D. M.: Mineral dust aerosol impacts on global climate and climate change, Nature Reviews Earth & Environment, 4, 71–86, https://doi.org/10.1038/s43017-022-00379-5, 2023.

Kolbe, W. M., Tonboe, R. T., and Stroeve, J.: Mapping of sea ice concentration using the NASA NIMBUS 5 Electrically Scanning Microwave Radiometer data from 1972–1977, Earth System Science Data, 16, 1247–1264, https://doi.org/10.5194/essd-16-1247-2024, 2024.

Kosaka, Y., Kobayashi, S., Harada, Y., Kobayashi, C., Naoe, H., Yoshimoto, K., Harada, M., Goto, N., Chiba, J., Miyaoka, K., Sekiguchi, R., Deushi, M., Kamahori, H., Nakaegawa, T., Tanaka, T. Y., Tokuhiro, T., Sato, Y., Matsushita, Y., and Onogi, K.: The JRA-3Q Reanalysis, Journal of the Meteorological Society of Japan, 102, https://doi.org/10.2151/jmsj.2024-004, 2024.

Kuhlbrodt, T., Jones, C. G., Sellar, A., Storkey, D., Blockley, E., Stringer, M., Hill, R., Graham, T., Ridley, J., Blaker, A., Calvert, D., Copsey, D., Ellis, R., Hewitt, H., Hyder, P., Ineson, S., Mulcahy, J., Siahaan, A., and Walton, J.: The Low-Resolution Version of HadGEM3 GC3.1: Development and Evaluation for Global Climate, Journal of Advances in Modeling Earth Systems, 10, 2865–2888, https://doi.org/10.1029/2018MS001370, 2018.

Kuo, Y., Kim, H., and Lehner, F.: Anthropogenic Aerosols Contribute to the Recent Decline in Precipitation Over the U.S. Southwest, Geophys. Res. Lett., 50, https://doi.org/10.1029/2023GL105389, 2023.

Laloyaux, P., de Boisseson, E., Balmaseda, M., Bidlot, J., Broennimann, S., Buizza, R., Dalhgren, P., Dee, D., Haimberger, L., Hersbach, H., Kosaka, Y., Martin, M., Poli, P., Rayner, N., Rustemeier, E., and Schepers, D.: CERA-20C: A Coupled Reanalysis of the Twentieth Century, Journal of Advances in Modeling Earth Systems, 10, 1172–1195, https://doi.org/10.1029/2018MS001273, 2018.

Lee, S., L'Heureux, M., Wittenberg, A. T., Seager, R., O'Gorman, P. A., and Johnson, N. C.: On the future zonal contrasts of equatorial Pacific climate: Perspectives from Observations, Simulations, and Theories, npj Climate and Atmospheric Science, 5, 82, https://doi.org/10.1038/s41612-022-00301-2, 2022.

Lee, S.-Y., Chiang, J. C. H., and Chang, P.: Tropical Pacific response to continental ice sheet topography, Climate Dynamics, 44, 2429–2446, https://doi.org/10.1007/s00382-014-2162-0, 2015.

Lehner, F., Deser, C., Simpson, I. R., and Terray, L.: Attributing the U.S. Southwest's Recent Shift Into Drier Conditions, Geophys. Res. Lett., 45, 6251–6261, https://doi.org/10.1029/2018GL078312, 2018.

Leith, C. E.: The Standard Error of Time-Average Estimates of Climatic Means, J. Appl. Meteor., 12, 1066–1069, https://doi.org/10.1175/1520-0450(1973)012⟨1066:TSEOTA⟩2.0.CO;2, 1973.

Lewis, N. and Mauritsen, T.: Negligible Unforced Historical Pattern Effect on Climate Feedback Strength Found in HadISST-Based AMIP Simulations, J. Climate, 34, 39–55, https://doi.org/10.1175/JCLI-D-19-0941.1, 2021.

Lin, Y., Huang, X., Liang, Y., Qin, Y., Xu, S., Huang, W., Xu, F., Liu, L., Wang, Y., Peng, Y., Wang, L., Xue, W., Fu, H., Zhang, G. J., Wang, B., Li, R., Zhang, C., Lu, H., Yang, K., Luo, Y., Bai, Y., Song, Z., Wang, M., Zhao, W., Zhang, F., Xu, J., Zhao, X., Lu, C., Chen, Y., Luo, Y., Hu, Y., Tang, Q., Chen, D., Yang, G., and Gong, P.: Community Integrated Earth System Model (CIESM): Description and Evaluation, Journal of Advances in Modeling Earth Systems, 12, e2019MS002 036, https://doi.org/10.1029/2019MS002036, 2020.

Liu, Z., Bao, Y., Thompson, L. G., Mosley-Thompson, E., Tabor, C., Zhang, G. J., Yan, M., Lofverstrom, M., Montanez, I., and Oster, J.: Tropical mountain ice core δ ¡sup¿18¡/sup¿ O: A Goldilocks indicator for global temperature change, Science Advances, 9, https://doi.org/10.1126/sciadv.adi6725, 2023.

Loeb, N. G., Wang, H., Allan, R. P., Andrews, T., Armour, K., Cole, J. N. S., Dufresne, J., Forster, P., Gettelman, A., Guo, H., Mauritsen, T., Ming, Y., Paynter, D., Proistosescu, C., Stuecker, M. F., Willén, U., and Wyser, K.: New Generation of Climate Models Track Recent Unprecedented Changes in Earth's Radiation Budget Observed by CERES, Geophys. Res. Lett., 47, https://doi.org/10.1029/2019GL086705, 2020.

Lou, J., O'Kane, T. J., and Holbrook, N. J.: A Linear Inverse Model of Tropical and South Pacific

188

Seasonal Predictability, J. Climate, 33, 4537–4554, https://doi.org/10.1175/JCLI-D-19-0548.1, 2020.

Lou, J., Newman, M., and Hoell, A.: Multi-decadal variation of ENSO forecast skill since the late 1800s, npj Climate and Atmospheric Science, 6, 89, https://doi.org/10.1038/s41612-023-00417-z, 2023.

Lunt, D. J., Haywood, A. M., Schmidt, G. A., Salzmann, U., Valdes, P. J., and Dowsett, H. J.: Earth system sensitivity inferred from Pliocene modelling and data, Nature Geoscience, 3, 60–64, https://doi.org/10.1038/ngeo706, 2010.

Lunt, D. J., Haywood, A. M., Schmidt, G. A., Salzmann, U., Valdes, P. J., Dowsett, H. J., and Loptson, C. A.: On the causes of mid-Pliocene warmth and polar amplification, Earth and Planetary Science Letters, 321-322, 128–138, https://doi.org/10.1016/j.epsl.2011.12.042, 2012.

L'Heureux, M. L., Lee, S., and Lyon, B.: Recent multidecadal strengthening of the Walker circulation across the tropical Pacific, Nature Climate Change, 3, 571–576, https://doi.org/10.1038/nclimate1840, 2013.

Mahowald, N. M., Li, L., Albani, S., Hamilton, D. S., and Kok, J. F.: Opinion: The importance of historical and paleoclimate aerosol radiative effects, Atmospheric Chemistry and Physics, 24, 533–551, https://doi.org/10.5194/acp-24-533-2024, 2024.

Manabe, S. and Broccoli, A. J.: The influence of continental ice sheets on the climate of an ice age, Journal of Geophysical Research, 90, 2167, https://doi.org/10.1029/JD090iD01p02167, 1985.

Manabe, S. and Bryan, K.: CO¡sub¿2¡/sub¿-induced change in a coupled ocean-atmosphere model and its paleoclimatic implications, Journal of Geophysical Research, 90, 11 689, https://doi.org/10.1029/JC090iC06p11689, 1985.

Marshall, G. J.: Trends in the Southern Annular Mode from Observations and Reanalyses, J. Climate, 16, 4134–4143, https://doi.org/10.1175/1520-0442(2003)016⟨4134:TITSAM⟩2.0.CO;2, 2003.

Marvel, K. and Webb, M.: Towards robust community assessments of the Earth's climate sensitivity, Earth System Dynamics, 16, 317–332, https://doi.org/10.5194/esd-16-317-2025, 2025.

Marvel, K., Pincus, R., Schmidt, G. A., and Miller, R. L.: Internal Variability and Disequilibrium Confound Estimates of Climate Sensitivity From Observations, Geophys. Res. Lett., 45, 1595–1601, https://doi.org/10.1002/2017GL076468, 2018.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenéz-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., von Storch, J., Tian, F., Voigt, A., Vrese, P., Wieners, K., Wilkenskjeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO_2, Journal of Advances in Modeling Earth Systems, 11, 998–1038, https://doi.org/10.1029/2018MS001400, 2019.

McClymont, E. L., Ford, H. L., Ling Ho, S., Tindall, J. C., Haywood, A. M., Alonso-Garcia, M., Bailey, I., Berke, M. A., Littler, K., Patterson, M. O., Petrick, B., Peterse, F., Christina Ravelo, A., Risebrobakken, B., De Schepper, S., Swann, G. E., Thirumalai, K., Tierney, J. E., Van Der Weijst, C., White, S., Abe-Ouchi, A., Baatsen, M. L., Brady, E. C., Chan, W. L., Chandan, D., Feng, R., Guo, C., Von Der Heydt, A. S., Hunter, S., Li, X., Lohmann, G., Nisancioglu, K. H., Otto-Bliesner, B. L., Richard Peltier, W., Stepanek, C., and Zhang, Z.: Lessons from a high-CO2 world: An ocean view from 3 million years ago, Climate of the Past, 16, 1599–1615, https://doi.org/10.5194/CP-16-1599-2020, 2020.

McGregor, S., Timmermann, A., Stuecker, M. F., England, M. H., Merrifield, M., Jin, F.-F., and

190

Chikamoto, Y.: Recent Walker circulation strengthening and Pacific cooling amplified by Atlantic warming, Nature Climate Change, 4, 888–892, https://doi.org/10.1038/nclimate2330, 2014.

Meier, W. N., Fetterer, F., Windnagel, A. K., and Stewart, J. S.: Near-Real-Time NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration (G10016, Version 2), Tech. rep., NSIDC: National Snow and Ice Data Center, Boulder, CO, https://doi.org/10.7265/tgam-yv28, 2021a.

Meier, W. N., Fetterer, F., Windnagel, A. K., and Stewart, J. S.: NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 4, Tech. rep., NSIDC: National Snow and Ice Data Center, Boulder, Colorado USA, https://doi.org/https://doi.org/10.7265/efmz-2t65, 2021b.

Menemenlis, S., Lora, J. M., Lofverstrom, M., and Chandan, D.: Influence of stationary waves on mid-Pliocene atmospheric rivers and hydroclimate, Global and Planetary Change, 204, 103 557, https://doi.org/10.1016/j.gloplacha.2021.103557, 2021.

Meng, Z. and Hakim, G. J.: Reconstructing the Tropical Pacific Upper Ocean Using Online Data Assimilation With a Deep Learning Model, Journal of Advances in Modeling Earth Systems, 16, https://doi.org/10.1029/2024MS004422, 2024.

Modak, A. and Mauritsen, T.: Better-constrained climate sensitivity when accounting for dataset dependency on pattern effect estimates, Atmospheric Chemistry and Physics, 23, 7535–7549, https://doi.org/10.5194/acp-23-7535-2023, 2023.

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., and Simpson, I. R.: An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set, J. Geophys. Res.: Atmospheres, 126, https://doi.org/10.1029/2019JD032361, 2021.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Neale, R. B., Gettelman, A., Park, S., Chen, C.-C., Lauritzen, P. H., Williamson, D. L., Conley, A. J., Kinnison, D., Marsh, D., Smith, A. K., Vitt, F., Garcia, R., Lamarque, J.-F., Mills, M., Tilmes, S., Morrison, H., Cameron-Smith, P., Collins, W. D., Iacono, M. J., Easter, R. C., Liu, X., Ghan, S. J., Rasch, P. J., and Taylor, M. A.: Description of the NCAR Community Atmosphere Model (CAM 5.0) (NCAR/TN-486+STR), Tech. rep., NCAR, https://doi.org/10.5065/wgtk-4g06, 2012.

Neale, R. B., Richter, J., Park, S., Lauritzen, P. H., Vavrus, S. J., Rasch, P. J., and Zhang, M.: The Mean Climate of the Community Atmosphere Model (CAM4) in Forced SST and Fully Coupled Experiments, J. Climate, 26, 5150–5168, https://doi.org/10.1175/JCLI-D-12-00236.1, 2013.

Newman, M.: Interannual to Decadal Predictability of Tropical and North Pacific Sea Surface Temperatures, J. Climate, 20, 2333–2356, https://doi.org/10.1175/JCLI4165.1, 2007.

Newman, M.: An Empirical Benchmark for Decadal Forecasts of Global Surface Temperature Anomalies, J. Climate, 26, 5260–5269, https://doi.org/10.1175/JCLI-D-12-00590.1, 2013.

Newman, M., Shin, S.-I., and Alexander, M. A.: Natural variation in ENSO flavors, Geophys. Res. Lett., 38, n/a–n/a, https://doi.org/10.1029/2011GL047658, 2011.

Newman, M., Alexander, M. A., Ault, T. R., Cobb, K. M., Deser, C., Di Lorenzo, E., Mantua, N. J., Miller, A. J., Minobe, S., Nakamura, H., Schneider, N., Vimont, D. J., Phillips, A. S., Scott, J. D., and Smith, C. A.: The Pacific Decadal Oscillation, Revisited, J. Climate, 29, 4399–4427, https://doi.org/10.1175/JCLI-D-15-0508.1, 2016.

NOAA: The Global Ensemble Forecast System (version 13) Replay dataset, `https://psl.noaa.gov/data/ufs_replay/`, 2024.

Notz, D. and SIMIP Community: Arctic Sea Ice in CMIP6, Geophys. Res. Lett., 47, e2019GL086 749, https://doi.org/10.1029/2019GL086749, 2020.

Olonscheck, D., Rugenstein, M., and Marotzke, J.: Broad Consistency Between Observed and Simulated Trends in Sea Surface Temperature Patterns, Geophys. Res. Lett., 47, e2019GL086 773, https://doi.org/10.1029/2019GL086773, 2020.

192

Orbe, C., Oman, L. D., Strahan, S. E., Waugh, D. W., Pawson, S., Takacs, L. L., and Molod, A. M.: Large-Scale Atmospheric Transport in GEOS-Replay Simulations, Journal of Advances in Modeling Earth Systems, 9, 2545–2560, https://doi.org/10.1002/2017MS001053, 2017.

OrtizBeviá, M. J.: Estimation of the cyclostationary dependence in geophysical data fields, J. Geophys. Res.: Atmospheres, 102, 13 473–13 486, https://doi.org/10.1029/97JD00243, 1997.

Osborn, T. J., Jones, P. D., Lister, D. H., Morice, C. P., Simpson, I. R., Winn, J. P., Hogan, E., and Harris, I. C.: Land Surface Air Temperature Variations Across the Globe Updated to 2019: The CRUTEM5 Data Set, J. Geophys. Res.: Atmospheres, 126, https://doi.org/10.1029/2019JD032352, 2021.

Osman, M. B., Tierney, J. E., Zhu, J., Tardif, R., Hakim, G. J., King, J., and Poulsen, C. J.: Globally resolved surface temperatures since the Last Glacial Maximum, Nature, 599, 239–244, https://doi.org/10.1038/s41586-021-03984-4, 2021.

O'Brien, C. L., Foster, G. L., Martínez-Botí, M. A., Abell, R., Rae, J. W. B., and Pancost, R. D.: High sea surface temperatures in tropical warm pools during the Pliocene, Nature Geoscience, 7, 606–611, https://doi.org/10.1038/ngeo2194, 2014.

O'Connor, G. K., Steig, E. J., and Hakim, G. J.: Strengthening Southern Hemisphere Westerlies and Amundsen Sea Low Deepening Over the 20th Century Revealed by Proxy-Data Assimilation, Geophys. Res. Lett., 48, e2021GL095 999, https://doi.org/10.1029/2021GL095999, 2021.

PALAEOSENS Project Members: Making sense of palaeoclimate sensitivity, Nature, 491, 683–691, https://doi.org/10.1038/nature11574, 2012.

Park, S., Shin, J., Kim, S., Oh, E., and Kim, Y.: Global Climate Simulated by the Seoul National University Atmosphere Model Version 0 with a Unified Convection Scheme (SAM0-UNICON), J. Climate, 32, 2917–2949, https://doi.org/10.1175/JCLI-D-18-0796.1, 2019.

Parsons, L. A., Amrhein, D. E., Sanchez, S. C., Tardif, R., Brennan, M. K., and Hakim, G. J.: Do Multi-Model Ensembles Improve Reconstruction Skill in Paleoclimate Data Assimilation?, Earth and Space Science, 8, e2020EA001 467, https://doi.org/10.1029/2020EA001467, 2021.

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

Pendergrass, A. G., Conley, A., and Vitt, F. M.: Surface and top-of-atmosphere radiative feedback kernels for CESM-CAM5, Earth Syst. Sci. Data, 10, 317–324, https://doi.org/10.5194/essd-10-317-2018, 2018.

Penland, C.: Random Forcing and Forecasting Using Principal Oscillation Pattern Analysis, Monthly Weather Review, 117, 2165–2185, https://doi.org/10.1175/1520-0493(1989)117⟨2165:RFAFUP⟩2.0.CO;2, 1989.

Penland, C.: A stochastic model of IndoPacific sea surface temperature anomalies, Physica D: Nonlinear Phenomena, 98, 534–558, https://doi.org/10.1016/0167-2789(96)00124-8, 1996.

Penland, C. and Matrosova, L.: A Balance Condition for Stochastic Numerical Models with Application to the El Niño-Southern Oscillation, J. Climate, 7, 1352–1372, https://doi.org/https://doi.org/10.1175/1520-0442(1994)007{\%}3C1352:ABCFSN{\%}3E2.0.CO;2, 1994.

Penland, C. and Sardeshmukh, P. D.: The Optimal Growth of Tropical Sea Surface Temperature Anomalies, J. Climate, 8, 1999–2024, https://doi.org/10.1175/1520-0442(1995)008⟨1999:TOGOTS⟩2.0.CO;2, 1995.

Perkins, W. A. and Hakim, G. J.: Coupled Atmosphere–Ocean Reconstruction of the Last Millennium Using Online Data Assimilation, Paleoceanography and Paleoclimatology, 36, https://doi.org/10.1029/2020PA003959, 2021.

Pincus, R., Forster, P. M., and Stevens, B.: The Radiative Forcing Model Intercomparison Project (RFMIP): experimental protocol for CMIP6, Geoscientific Model Development, 9, 3447–3460, https://doi.org/10.5194/gmd-9-3447-2016, 2016.

Po-Chedley, S., Santer, B. D., Fueglistaler, S., Zelinka, M. D., Cameron-Smith, P. J., Painter, J. F., and Fu, Q.: Natural variability contributes to model–satellite differences in tropical tropospheric warming, Proc. Natl. Acad. Sci. (USA), 118, https://doi.org/10.1073/pnas.2020962118, 2021.

194

Polvani, L. M., Waugh, D. W., Correa, G. J. P., and Son, S.-W.: Stratospheric Ozone Deple-
tion: The Main Driver of Twentieth-Century Atmospheric Circulation Changes in the Southern
Hemisphere, J. Climate, 24, 795–812, https://doi.org/10.1175/2010JCLI3772.1, 2011.

Polvani, L. M., Banerjee, A., Chemke, R., Doddridge, E. W., Ferreira, D., Gnanadesikan, A.,
Holland, M. A., Kostov, Y., Marshall, J., Seviour, W. J. M., Solomon, S., and Waugh, D. W.:
Interannual SAM Modulation of Antarctic Sea Ice Extent Does Not Account for Its Long-Term
Trends, Pointing to a Limited Role for Ozone Depletion, Geophys. Res. Lett., 48, https://doi.org/
10.1029/2021GL094871, 2021.

Prentice, I. C., Harrison, S. P., and Bartlein, P. J.: Global vegetation and terrestrial carbon
cycle changes after the last ice age, New Phytologist, 189, 988–998, https://doi.org/10.1111/j.
1469-8137.2010.03620.x, 2011.

Proistosescu, C. and Huybers, P. J.: Slow climate mode reconciles historical and model-based
estimates of climate sensitivity, Science Advances, 3, https://doi.org/10.1126/sciadv.1602821,
2017.

Qin, Y., Zheng, X., Klein, S. A., Zelinka, M. D., Ma, P., Golaz, J., and Xie, S.: Causes of Reduced
Climate Sensitivity in E3SM From Version 1 to Version 2, Journal of Advances in Modeling Earth
Systems, 16, https://doi.org/10.1029/2023MS003875, 2024.

Raghuraman, S. P., Paynter, D., Menzel, R., and Ramaswamy, V.: Forcing, Cloud Feedbacks, Cloud
Masking, and Internal Variability in the Cloud Radiative Effect Satellite Record, J. Climate, 36,
4151–4167, https://doi.org/10.1175/JCLI-D-22-0555.1, 2023.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent,
E. C., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine
air temperature since the late nineteenth century, J. Geophys. Res.: Atmospheres, 108, 4407,
https://doi.org/10.1029/2002JD002670, 2003.

Renoult, M., Sagoo, N., Zhu, J., and Mauritsen, T.: Causes of the weak emergent constraint on cli-
mate sensitivity at the Last Glacial Maximum, Climate of the Past, 19, 323–356, https://doi.org/
10.5194/cp-19-323-2023, 2023.

Richardson, T. B., Forster, P. M., Smith, C. J., Maycock, A. C., Wood, T., Andrews, T., Boucher, O., Faluvegi, G., Fläschner, D., Hodnebrog, Ø., Kasoar, M., Kirkevåg, A., Lamarque, J., Mülmenstädt, J., Myhre, G., Olivié, D., Portmann, R. W., Samset, B. H., Shawki, D., Shindell, D., Stier, P., Takemura, T., Voulgarakis, A., and Watson-Parris, D.: Efficacy of Climate Forcings in PDRMIP Models, J. Geophys. Res.: Atmospheres, 124, 12 824–12 844, https://doi.org/10.1029/2019JD030581, 2019.

Roach, L. A. and Meier, W. N.: Sea ice in 2023, Nature Reviews Earth & Environment, 5, 235–237, https://doi.org/10.1038/s43017-024-00542-0, 2024.

Roach, L. A., Dörr, J., Holmes, C. R., Massonnet, F., Blockley, E. W., Notz, D., Rackow, T., Raphael, M. N., O'Farrell, S. P., Bailey, D. A., and Bitz, C. M.: Antarctic Sea Ice Area in CMIP6, Geophys. Res. Lett., 47, https://doi.org/10.1029/2019GL086729, 2020.

Roberts, W. H., Li, C., and Valdes, P. J.: The Mechanisms that Determine the Response of the Northern Hemisphere's Stationary Waves to North American Ice Sheets, J. Climate, 32, 3917–3940, https://doi.org/10.1175/JCLI-D-18-0586.1, 2019.

Roe, G. H. and Baker, M. B.: Why Is Climate Sensitivity So Unpredictable?, Science, 318, 629–632, https://doi.org/10.1126/science.1144735, 2007.

Roe, G. H. and Lindzen, R. S.: The Mutual Interaction between Continental-Scale Ice Sheets and Atmospheric Stationary Waves, J. Climate, 14, 1450–1465, https://doi.org/10.1175/1520-0442(2001)014⟨1450:TMIBCS⟩2.0.CO;2, 2001.

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, A., Wurtele, J., Curry, J., Wickham, C., and Mosher, S.: An Overview Berkeley Earth Temperature Averaging Process, Geoinformatics & Geostatistics: An Overview, https://doi.org/10.4172/gigs.1000103, 2013.

Rohling, E. J., Marino, G., Foster, G. L., Goodwin, P. A., von der Heydt, A. S., and Köhler, P.: Comparing Climate Sensitivity, Past and Present, Annual Review of Marine Science, 10, 261–288, https://doi.org/10.1146/annurev-marine-121916-063242, 2018.

Rose, B. E. J., Armour, K. C., Battisti, D. S., Feldl, N., Rose, B. E. J., Armour, K. C., Battisti, D. S., Feldl, N., and Koll, D. D. B.: The dependence of transient climate sensitivity and radiative

feedbacks on the spatial pattern of ocean heat uptake, Geophys. Res. Lett., 41, 1071–1078, https://doi.org/10.1002/2013GL058955, 2014.

Rugenstein, M., Bloch-Johnson, J., Abe-Ouchi, A., Andrews, T., Beyerle, U., Cao, L., Chadha, T., Danabasoglu, G., Dufresne, J.-L., Duan, L., Foujols, M.-A., Frölicher, T., Geoffroy, O., Gregory, J., Knutti, R., Li, C., Marzocchi, A., Mauritsen, T., Menary, M., Moyer, E., Nazarenko, L., Paynter, D., Saint-Martin, D., Schmidt, G. A., Yamamoto, A., and Yang, S.: LongRunMIP: Motivation and Design for a Large Collection of Millennial-Length AOGCM Simulations, Bull. Amer. Meteor. Soc., 100, 2551–2570, https://doi.org/10.1175/BAMS-D-19-0068.1, 2019.

Rugenstein, M., Dhame, S., Olonscheck, D., Wills, R. J., Watanabe, M., and Seager, R.: Connecting the SST Pattern Problem and the Hot Model Problem, Geophys. Res. Lett., 50, https://doi.org/10.1029/2023GL105488, 2023.

Rugenstein, M. A. and Armour, K. C.: Three Flavors of Radiative Feedbacks and Their Implications for Estimating Equilibrium Climate Sensitivity, Geophys. Res. Lett., 48, https://doi.org/10.1029/2021GL092983, 2021.

Sagoo, N. and Storelvmo, T.: Testing the sensitivity of past climates to the indirect effects of dust, Geophys. Res. Lett., 44, 5807–5817, https://doi.org/10.1002/2017GL072584, 2017.

Saha, S., Moorthi, S., Pan, H.-L., Wu, X., Wang, J., Nadiga, S., Tripp, P., Kistler, R., Woollen, J., Behringer, D., Liu, H., Stokes, D., Grumbine, R., Gayno, G., Wang, J., Hou, Y.-T., Chuang, H.-y., Juang, H.-M. H., Sela, J., Iredell, M., Treadon, R., Kleist, D., Van Delst, P., Keyser, D., Derber, J., Ek, M., Meng, J., Wei, H., Yang, R., Lord, S., van den Dool, H., Kumar, A., Wang, W., Long, C., Chelliah, M., Xue, Y., Huang, B., Schemm, J.-K., Ebisuzaki, W., Lin, R., Xie, P., Chen, M., Zhou, S., Higgins, W., Zou, C.-Z., Liu, Q., Chen, Y., Han, Y., Cucurull, L., Reynolds, R. W., Rutledge, G., and Goldberg, M.: The NCEP Climate Forecast System Reanalysis, Bull. Amer. Meteor. Soc., 91, 1015–1058, https://doi.org/10.1175/2010BAMS3001.1, 2010.

Salvi, P., Gregory, J. M., and Ceppi, P.: Time-Evolving Radiative Feedbacks in the Historical Period, J. Geophys. Res.: Atmospheres, 128, https://doi.org/10.1029/2023JD038984, 2023.

Salzmann, U., Dolan, A. M., Haywood, A. M., Chan, W.-L., Voss, J., Hill, D. J., Abe-Ouchi, A., Otto-Bliesner, B., Bragg, F. J., Chandler, M. A., Contoux, C., Dowsett, H. J., Jost, A., Kamae, Y., Lohmann, G., Lunt, D. J., Pickering, S. J., Pound, M. J., Ramstein, G., Rosenbloom, N. A., Sohl, L., Stepanek, C., Ueda, H., and Zhang, Z.: Challenges in quantifying Pliocene terrestrial warming revealed by data–model discord, Nature Climate Change, 3, 969–974, https://doi.org/10.1038/nclimate2008, 2013.

Samakinwa, E., Valler, V., Hand, R., Neukom, R., Gómez-Navarro, J. J., Kennedy, J., Rayner, N. A., and Brönnimann, S.: An ensemble reconstruction of global monthly sea surface temperature and sea ice concentration 1000–1849, Scientific Data, 8, 261, https://doi.org/10.1038/s41597-021-01043-1, 2021.

Sanchez, S. C., Westphal, N., Haug, G. H., Cheng, H., Edwards, R. L., Schneider, T., Cobb, K. M., and Charles, C. D.: A Continuous Record of Central Tropical Pacific Climate Since the Mid-nineteenth Century Reconstructed From Fanning and Palmyra Island Corals: A Case Study in Coral Data Reanalysis, Paleoceanography and Paleoclimatology, 35, 1877–1878, https://doi.org/10.1029/2020PA003848, 2020.

Sanchez, S. C., Hakim, G. J., and Saenger, C. P.: Climate Model Teleconnection Patterns Govern the Niño-3.4 Response to Early Nineteenth-Century Volcanism in Coral-Based Data Assimilation Reconstructions, J. Climate, 34, 1863–1880, https://doi.org/10.1175/JCLI-D-20-0549.1, 2021.

Schmidt, G. A., Annan, J. D., Bartlein, P. J., Cook, B. I., Guilyardi, E., Hargreaves, J. C., Harrison, S. P., Kageyama, M., LeGrande, A. N., Konecky, B., Lovejoy, S., Mann, M. E., Masson-Delmotte, V., Risi, C., Thompson, D., Timmermann, A., Tremblay, L.-B., and Yiou, P.: Using palaeoclimate comparisons to constrain future projections in CMIP5, Climate of the Past, 10, 221–250, https://doi.org/10.5194/cp-10-221-2014, 2014.

Schmidt, G. A., Romanou, A., Roach, L. A., Mankoff, K. D., Li, Q., Rye, C. D., Kelley, M., Marshall, J. C., and Busecke, J. J. M.: Anomalous Meltwater From Ice Sheets and Ice Shelves Is a Historical Forcing, Geophys. Res. Lett., 50, https://doi.org/10.1029/2023GL106530, 2023.

Schneider, D. P. and Fogt, R. L.: Artifacts in Century-Length Atmospheric and Coupled Reanal-

yses Over Antarctica Due To Historical Data Availability, Geophys. Res. Lett., 45, 964–973, https://doi.org/10.1002/2017GL076226, 2018.

Schneider, T.: Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values, J. Climate, 14, 853–871, https://doi.org/10.1175/ 1520-0442(2001)014⟨0853:AOICDE⟩2.0.CO;2, 2001.

Screen, J. A.: Sudden increase in Antarctic sea ice: Fact or artifact?, Geophys. Res. Lett., 38, n/a–n/a, https://doi.org/10.1029/2011GL047553, 2011.

Seager, R., Cane, M., Henderson, N., Lee, D.-E., Abernathey, R., and Zhang, H.: Strengthening tropical Pacific zonal sea surface temperature gradient consistent with rising greenhouse gases, Nature Climate Change, 9, 517–522, https://doi.org/10.1038/s41558-019-0505-x, 2019.

Seager, R., Henderson, N., and Cane, M.: Persistent Discrepancies between Observed and Modeled Trends in the Tropical Pacific Ocean, J. Climate, 35, 4571–4584, https://doi.org/ 10.1175/JCLI-D-21-0648.1, 2022.

Seager, R., Ting, M., Alexander, P., Liu, H., Nakamura, J., Li, C., and Newman, M.: Ocean-forcing of cool season precipitation drives ongoing and future decadal drought in southwestern North America, npj Climate and Atmospheric Science, 6, 141, https://doi.org/10.1038/ s41612-023-00461-9, 2023.

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, Geoscientific Model Development, 13, 6165– 6200, https://doi.org/10.5194/gmd-13-6165-2020, 2020.

Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., de Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews,

T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahaan, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, Journal of Advances in Modeling Earth Systems, 11, 4513–4558, https://doi.org/10.1029/2019MS001739, 2019.

Seltzer, A. M., Ng, J., Aeschbach, W., Kipfer, R., Kulongoski, J. T., Severinghaus, J. P., and Stute, M.: Widespread six degrees Celsius cooling on land during the Last Glacial Maximum, Nature, 593, 228–232, https://doi.org/10.1038/s41586-021-03467-6, 2021.

Senior, C. A. and Mitchell, J. F. B.: The time-dependence of climate sensitivity, Geophys. Res. Lett., 27, 2685–2688, https://doi.org/10.1029/2000GL011373, 2000.

Seviour, W. J. M., Gnanadesikan, A., and Waugh, D. W.: The Transient Response of the Southern Ocean to Stratospheric Ozone Depletion, J. Climate, 29, 7383–7396, https://doi.org/10.1175/JCLI-D-16-0198.1, 2016.

Shakun, J. D.: Modest global-scale cooling despite extensive early Pleistocene ice sheets, Quaternary Science Reviews, 165, 25–30, https://doi.org/10.1016/j.quascirev.2017.04.010, 2017.

Shell, K. M., Kiehl, J. T., and Shields, C. A.: Using the Radiative Kernel Technique to Calculate Climate Feedbacks in NCAR's Community Atmospheric Model, J. Climate, 21, 2269–2282, https://doi.org/10.1175/2007JCLI2044.1, 2008.

Sherwood, S. C. and Forest, C. E.: Opinion: Can uncertainty in climate sensitivity be narrowed further?, Atmospheric Chemistry and Physics, 24, 2679–2686, https://doi.org/10.5194/acp-24-2679-2024, 2024.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka,

M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Rev. Geophys., 58, https://doi.org/10.1029/2019RG000678, 2020.

Shin, S.-I., Sardeshmukh, P. D., Newman, M., Penland, C., and Alexander, M. A.: Impact of Annual Cycle on ENSO Variability and Predictability, J. Climate, 34, 171–193, https://doi.org/10.1175/JCLI-D-20-0291.1, 2021.

Siler, N., Proistosescu, C., and Po-Chedley, S.: Natural Variability Has Slowed the Decline in Western U.S. Snowpack Since the 1980s, Geophys. Res. Lett., 46, 346–355, https://doi.org/10.1029/2018GL081080, 2019.

Simpson, I. R., Shaw, T. A., Ceppi, P., Clement, A. C., Fischer, E., Grise, K. M., Pendergrass, A. G., Screen, J. A., Wills, R. C. J., Woollings, T., Blackport, R., Kang, J. M., and Po-Chedley, S.: Confronting Earth System Model trends with observations, Science Advances, 11, 8035, https://doi.org/10.1126/sciadv.adt8035, 2025.

Sippel, S., Kent, E. C., Meinshausen, N., Chan, D., Kadow, C., Neukom, R., Fischer, E. M., Humphrey, V., Rohde, R., de Vries, I., and Knutti, R.: Early-twentieth-century cold bias in ocean surface temperature observations, Nature, 635, 618–624, https://doi.org/10.1038/s41586-024-08230-1, 2024.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., Allan, R., Yin, X., Vose, R., Titchner, H., Kennedy, J., Spencer, L. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Crouthamel, R., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Jourdain, S., Kaplan, A., Kubota, H., Blancq, F. L., Lee, T., Lorrey, A., Luterbacher, J., Maugeri, M., Mock, C. J., Moore, G. K., Przybylak, R., Pudmenzky, C., Reason, C., Slonosky, V. C., Smith, C. A., Tinz, B., Trewin, B., Valente, M. A., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński, P.: Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system, Quart. J. Roy. Meteor. Soc., 145, 2876–2908, https://doi.org/10.1002/qj.3598, 2019.

Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Whitaker, J. S., McColl, C., Allan, R. J., Brohan, P., Yin, X., Smith, C. A., Spencer, L. J., Vose, R. S., Rohrer, M., Conroy, R. P., Schuster,

D. C., Kennedy, J. J., Ashcroft, L., Brönnimann, S., Brunet, M., Camuffo, D., Cornes, R., Cram, T. A., Domínguez-Castro, F., Freeman, J. E., Gergis, J., Hawkins, E., Jones, P. D., Kubota, H., Lee, T. C., Lorrey, A. M., Luterbacher, J., Mock, C. J., Przybylak, R. K., Pudmenzky, C., Slonosky, V. C., Tinz, B., Trewin, B., Wang, X. L., Wilkinson, C., Wood, K., and Wyszyński, P.: An Evaluation of the Performance of the Twentieth Century Reanalysis Version 3, J. Climate, 34, 1417–1438, https://doi.org/10.1175/JCLI-D-20-0505.1, 2021.

Smerdon, J. E., Cook, E. R., and Steiger, N. J.: The Historical Development of Large-Scale Paleoclimate Field Reconstructions Over the Common Era, Rev. Geophys., 61, https://doi.org/10.1029/2022RG000782, 2023.

Smith, C., Cummins, D. P., Fredriksen, H.-B., Nicholls, Z., Meinshausen, M., Allen, M., Jenkins, S., Leach, N., Mathison, C., and Partanen, A.-I.: fair-calibrate v1.4.1: calibration, constraining, and validation of the FaIR simple climate model for reliable future climate projections, Geoscientific Model Development, 17, 8569–8592, https://doi.org/10.5194/gmd-17-8569-2024, 2024.

Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Villaume, S., Haimberger, L., Woollen, J., Buontempo, C., and Thépaut, J.: The ERA5 global reanalysis from 1940 to 2022, Quart. J. Roy. Meteor. Soc., 150, 4014–4048, https://doi.org/10.1002/qj.4803, 2024.

Soden, B. J. and Held, I. M.: An Assessment of Climate Feedbacks in Coupled Ocean–Atmosphere Models, J. Climate, 19, 3354–3360, https://doi.org/10.1175/JCLI3799.1, 2006.

Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., and Shields, C. A.: Quantifying Climate Feedbacks Using Radiative Kernels, J. Climate, 21, 3504–3520, https://doi.org/10.1175/2007JCLI2110.1, 2008.

Solomon, A. and Newman, M.: Reconciling disparate twentieth-century Indo-Pacific ocean temperature trends in the instrumental record, Nature Climate Change, 2, 691–699, https://doi.org/10.1038/nclimate1591, 2012.

Stap, L. B., Köhler, P., and Lohmann, G.: Including the efficacy of land ice changes in deriving

climate sensitivity from paleodata, Earth System Dynamics, 10, 333–345, https://doi.org/10. 5194/esd-10-333-2019, 2019.

Steiger, N. J., Hakim, G. J., Steig, E. J., Battisti, D. S., and Roe, G. H.: Assimilation of time-averaged pseudoproxies for climate reconstruction, J. Climate, 27, https://doi.org/ 10.1175/JCLI-D-12-00693.1, 2014.

Steiger, N. J., Smerdon, J. E., Cook, E. R., and Cook, B. I.: A reconstruction of global hydroclimate and dynamical variables over the Common Era, Scientific Data, 5, 180 086, https://doi.org/ 10.1038/sdata.2018.86, 2018.

Stevens, B., Sherwood, S. C., Bony, S., and Webb, M. J.: Prospects for narrowing bounds on Earth's equilibrium climate sensitivity, Earth's Future, 4, 512–522, https://doi.org/10.1002/ 2016EF000376, 2016.

Stuecker, M. F., Bitz, C. M., and Armour, K. C.: Conditions leading to the unprecedented low Antarctic sea ice extent during the 2016 austral spring season, Geophys. Res. Lett., 44, 9008–9019, https://doi.org/10.1002/2017GL074691, 2017.

Suryawanshi, K., Jena, B., Bajish, C. C., and Anilkumar, N.: Recent Decline in Antarctic Sea Ice Cover From 2016 to 2022: Insights From Satellite Observations, Argo Floats, and Model Reanalysis, Tellus, Series A: Dynamic Meteorology and Oceanography, 75, 193–212, https://doi.org/ 10.16993/TELLUSA.3222, 2023.

Swart, N. C., Fyfe, J. C., Gillett, N., and Marshall, G. J.: Comparing Trends in the Southern Annular Mode and Surface Westerly Jet, J. Climate, 28, 8840–8859, https://doi.org/10.1175/ JCLI-D-15-0334.1, 2015.

Swart, N. C., Martin, T., Beadling, R., Chen, J.-J., Danek, C., England, M. H., Farneti, R., Griffies, S. M., Hattermann, T., Hauck, J., Haumann, F. A., Jüling, A., Li, Q., Marshall, J., Muilwijk, M., Pauling, A. G., Purich, A., Smith, I. J., and Thomas, M.: The Southern Ocean Freshwater Input from Antarctica (SOFIA) Initiative: scientific objectives and experimental design, Geoscientific Model Development, 16, 7289–7309, https://doi.org/10.5194/gmd-16-7289-2023, 2023.

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

Thompson, D. W. J. and Solomon, S.: Interpretation of Recent Southern Hemisphere Climate Change, Science, 296, 895–899, https://doi.org/10.1126/science.1069270, 2002.

Thompson, D. W. J., Solomon, S., Kushner, P. J., England, M. H., Grise, K. M., and Karoly, D. J.: Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change, Nature Geoscience, 4, 741–749, https://doi.org/10.1038/ngeo1296, 2011.

Tierney, J. E., Haywood, A. M., Feng, R., Bhattacharya, T., and Otto-Bliesner, B. L.: Pliocene Warmth Consistent With Greenhouse Gas Forcing, Geophys. Res. Lett., 46, 9136–9144, https://doi.org/10.1029/2019GL083802, 2019.

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

Tierney, J. E., Judd, E. J., Osman, M. B., King, J. M., Truax, O. J., Steiger, N. J., Amrhein, D. E., and Anchukaitis, K. J.: Advances in Paleoclimate Data Assimilation, Annu. Rev. Earth Planet. Sci., 28, 55, https://doi.org/10.1146/annurev-earth-032320-064209, 2025a.

Tierney, J. E., King, J., Osman, M. B., Abell, J. T., Burls, N. J., Erfani, E., Cooper, V. T., and Feng, R.: Pliocene Warmth and Patterns of Climate Change Inferred From Paleoclimate Data Assimilation, AGU Advances, 6, https://doi.org/10.1029/2024AV001356, 2025b.

Titchner, H. A. and Rayner, N. A.: The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations, J. Geophys. Res.: Atmospheres, 119, 2864–2889, https://doi.org/10.1002/2013JD020316, 2014.

Tokinaga, H., Xie, S.-P., Timmermann, A., McGregor, S., Ogata, T., Kubota, H., and Okumura, Y. M.: Regional Patterns of Tropical Indo-Pacific Climate Change: Evidence of the Walker Circulation Weakening, J. Climate, 25, 1689–1710, https://doi.org/10.1175/JCLI-D-11-00263.1, 2012.

Trenberth, K. E. and Shea, D. J.: Atlantic hurricanes and natural variability in 2005, Geophys. Res. Lett., 33, 12 704, https://doi.org/10.1029/2006GL026894, 2006.

Tseng, K.-C., Johnson, N. C., Maloney, E. D., Barnes, E. A., and Kapnick, S. B.: Mapping Large-Scale Climate Variability to Hydrological Extremes: An Application of the Linear Inverse Model to Subseasonal Prediction, J. Climate, 34, 4207–4225, https://doi.org/10.1175/JCLI-D-20-0502.1, 2021.

Turner, J., Holmes, C., Caton Harrison, T., Phillips, T., Jena, B., Reeves-Francois, T., Fogt, R., Thomas, E. R., and Bajish, C. C.: Record Low Antarctic Sea Ice Cover in February 2022, Geophys. Res. Lett., 49, https://doi.org/10.1029/2022GL098904, 2022.

Unger, N. and Yue, X.: Strong chemistry-climate feedbacks in the Pliocene, Geophys. Res. Lett., 41, 527–533, https://doi.org/10.1002/2013GL058773, 2014.

Vaccaro, A., Emile-Geay, J., Guillot, D., Verna, R., Morice, C., Kennedy, J., and Rajaratnam, B.: Climate Field Completion via Markov Random Fields: Application to the HadCRUT4.6 Temperature Dataset, J. Climate, 34, 4169–4188, https://doi.org/10.1175/JCLI-D-19-0814.1, 2021.

Valler, V., Franke, J., Brugnara, Y., Samakinwa, E., Hand, R., Lundstad, E., Burgdorf, A.-M., Lipfert, L., Friedman, A. R., and Brönnimann, S.: ModE-RA: a global monthly paleo-reanalysis of the modern era 1421 to 2008, Scientific Data, 11, 36, https://doi.org/10.1038/s41597-023-02733-8, 2024.

Vecchi, G. A. and Soden, B. J.: Global Warming and the Weakening of the Tropical Circulation, J. Climate, 20, 4316–4340, https://doi.org/10.1175/JCLI4258.1, 2007.

Vecchi, G. A., Soden, B. J., Wittenberg, A. T., Held, I. M., Leetmaa, A., and Harrison, M. J.: Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing, Nature, 441, 73–76, https://doi.org/10.1038/nature04744, 2006.

Vimont, D. J.: Analysis of the Atlantic Meridional Mode Using Linear Inverse Modeling: Seasonality and Regional Influences, J. Climate, 25, 1194–1212, https://doi.org/10.1175/JCLI-D-11-00012.1, 2012.

Vimont, D. J., Alexander, M. A., and Newman, M.: Optimal growth of Central and East Pacific ENSO events, Geophys. Res. Lett., 41, 4027–4034, https://doi.org/10.1002/2014GL059997, 2014.

Vimont, D. J., Newman, M., Battisti, D. S., and Shin, S.-I.: The Role of Seasonality and the ENSO Mode in Central and East Pacific ENSO Growth and Evolution, J. Climate, 35, 3195–3209, https://doi.org/10.1175/JCLI-D-21-0599.1, 2022.

von der Heydt, A. S., Dijkstra, H. A., van de Wal, R. S. W., Caballero, R., Crucifix, M., Foster, G. L., Huber, M., Köhler, P., Rohling, E., Valdes, P. J., Ashwin, P., Bathiany, S., Berends, T., van Bree, L. G. J., Ditlevsen, P., Ghil, M., Haywood, A. M., Katzav, J., Lohmann, G., Lohmann, J., Lucarini, V., Marzocchi, A., Pälike, H., Baroni, I. R., Simon, D., Sluijs, A., Stap, L. B., Tantet, A., Viebahn, J., and Ziegler, M.: Lessons on Climate Sensitivity From Past Climate Changes, Current Climate Change Reports, 2, 148–158, https://doi.org/10.1007/s40641-016-0049-3, 2016.

Waelbroeck, C., Paul, A., Kucera, M., Rosell-Melé, A., Weinelt, M., Schneider, R., Mix, A. C., Abelmann, A., Armand, L., Bard, E., Barker, S., Barrows, T. T., Benway, H., Cacho, I., Chen, M. T., Cortijo, E., Crosta, X., De Vernal, A., Dokken, T., Duprat, J., Elderfield, H., Eynaud, F., Gersonde, R., Hayes, A., Henry, M., Hillaire-Marcel, C., Huang, C. C., Jansen, E., Juggins, S., Kallel, N., Kiefer, T., Kienast, M., Labeyrie, L., Leclaire, H., Londeix, L., Mangin, S., Matthiessen, J., Marret, F., Meland, M., Morey, A. E., Mulitza, S., Pflaumann, U., Pisias, N. G., Radi, T., Rochon, A., Rohling, E. J., Sbaffi, L., Schäfer-Neth, C., Solignac, S., Spero, H., Tachikawa, K., and Turon, J. L.: Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum, Nature Geoscience 2009 2:2, 2, 127–132, https://doi.org/10.1038/ngeo411, 2009.

Walsh, J. E., Fetterer, F., Scott stewart, J., and Chapman, W. L.: A database for depicting Arctic sea ice variations back to 1850, Geographical Review, 107, 89–107, https://doi.org/10.1111/j.1931-0846.2016.12195.x, 2017.

Walsh, J. E., Chapman, W. L., Fetterer, F., and Stewart, S.: Gridded Monthly Sea Ice Extent and Concentration, 1850 Onward, Version 2, Tech. rep., NSIDC: National Snow and Ice Data Center., Boulder, Colorado USA, https://doi.org/10.7265/jj4s-tq79, 2019.

Wang, C., Yang, W., Vecchi, G., Zhang, B., Soden, B. J., and Chan, D.: Diagnosing the Factors That Contribute to the Intermodel Spread of Climate Feedback in CMIP6, J. Climate, 38, 663–674, https://doi.org/10.1175/JCLI-D-23-0528.1, 2025.

Wara, M. W., Ravelo, A. C., and Delaney, M. L.: Permanent El Nino-Like Conditions During the Pliocene Warm Period, Science, 309, 758–761, https://doi.org/10.1126/science.1112596, 2005.

Watanabe, M., Dufresne, J.-L., Kosaka, Y., Mauritsen, T., and Tatebe, H.: Enhanced warming constrained by past trends in equatorial Pacific sea surface temperature gradient, Nature Climate Change, 11, 33–37, https://doi.org/10.1038/s41558-020-00933-3, 2021.

Watanabe, M., Iwakiri, T., Dong, Y., and Kang, S. M.: Two Competing Drivers of the Recent Walker Circulation Trend, Geophys. Res. Lett., 50, e2023GL105 332, https://doi.org/10.1029/2023GL105332, 2023.

Watanabe, M., Kang, S. M., Collins, M., Hwang, Y.-T., McGregor, S., and Stuecker, M. F.: Possible shift in controls of the tropical Pacific surface warming pattern, Nature, 630, 315–324, https://doi.org/10.1038/s41586-024-07452-7, 2024.

Webb, M.: Code and Data for WCRP Climate Sensitivity Assessment, Tech. rep., https://doi.org/10.5281/zenodo.3945275, 2020.

Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., Chepfer, H., Douville, H., Good, P., Kay, J. E., Klein, S. A., Marchand, R., Medeiros, B., Siebesma, A. P., Skinner, C. B., Stevens, B., Tselioudis, G., Tsushima, Y., and Watanabe, M.: The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6, Geoscientific Model Development, 10, 359–384, https://doi.org/10.5194/gmd-10-359-2017, 2017.

Weiffenbach, J. E., Baatsen, M. L. J., Dijkstra, H. A., Von Der Heydt, A. S., Abe-Ouchi, A., Brady, E. C., Chan, W.-L., Chandan, D., Chandler, M. A., Contoux, C., Feng, R., Guo, C., Han, Z., Haywood, A. M., Li, Q., Li, X., Lohmann, G., Lunt, D. J., Nisancioglu, K. H., Otto-Bliesner, B. L., Peltier, W. R., Ramstein, G., Sohl, L. E., Stepanek, C., Tan, N., Tindall, J. C., Williams, C. J. R., Zhang, Q., and Zhang, Z.: Unraveling the mechanisms and implications of a stronger

mid-Pliocene Atlantic Meridional Overturning Circulation (AMOC) in PlioMIP2, Clim. Past, 19, 61–85, https://doi.org/10.5194/cp-19-61-2023, 2023.

Weiffenbach, J. E., Dijkstra, H. A., von der Heydt, A. S., Abe-Ouchi, A., Chan, W.-L., Chandan, D., Feng, R., Haywood, A. M., Hunter, S. J., Li, X., Otto-Bliesner, B. L., Peltier, W. R., Stepanek, C., Tan, N., Tindall, J. C., and Zhang, Z.: Highly stratified mid-Pliocene Southern Ocean in PlioMIP2, Climate of the Past, 20, 1067–1086, https://doi.org/10.5194/cp-20-1067-2024, 2024.

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., Davis, P., Graham, T., Hewitt, H. T., Hill, R., Hyder, P., Ineson, S., Johns, T. C., Keen, A. B., Lee, R. W., Megann, A., Milton, S. F., Rae, J. G. L., Roberts, M. J., Scaife, A. A., Schiemann, R., Storkey, D., Thorpe, L., Watterson, I. G., Walters, D. N., West, A., Wood, R. A., Woollings, T., and Xavier, P. K.: The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations, Journal of Advances in Modeling Earth Systems, 10, 357–380, https://doi.org/10.1002/2017MS001115, 2017.

Wills, R. C., Battisti, D. S., Armour, K. C., Schneider, T., and Deser, C.: Pattern Recognition Methods to Separate Forced Responses from Internal Variability in Climate Model Ensembles and Observations, J. Climate, 33, 8693–8719, https://doi.org/10.1175/JCLI-D-19-0855.1, 2020.

Wills, R. C. J., Dong, Y., Proistosecu, C., Armour, K. C., and Battisti, D. S.: Systematic Climate Model Biases in the Large-Scale Patterns of Recent Sea-Surface Temperature and Sea-Level Pressure Change, Geophys. Res. Lett., 49, e2022GL100 011, https://doi.org/10.1029/2022GL100011, 2022.

Wunsch, C.: The Interpretation of Short Climate Records, with Comments on the North Atlantic and Southern Oscillations, Bull. Amer. Meteor. Soc., 80, 245–255, https://doi.org/10.1175/1520-0477(1999)080⟨0245:TIOSCR⟩2.0.CO;2, 1999.

Xie, S. P., Deser, C., Vecchi, G. A., Ma, J., Teng, H., and Wittenberg, A. T.: Global Warming Pattern Formation: Sea Surface Temperature and Rainfall, J. Climate, 23, 966–986, https://doi.org/10.1175/2009JCLI3329.1, 2010.

Yoshimori, M., Yokohata, T., and Abe-Ouchi, A.: A Comparison of Climate Feedback Strength between CO2 Doubling and LGM Experiments, J. Climate, 22, 3374–3395, https://doi.org/10.1175/2009JCLI2801.1, 2009.

Yoshimori, M., Hargreaves, J. C., Annan, J. D., Yokohata, T., and Abe-Ouchi, A.: Dependency of Feedbacks on Forcing and Climate State in Physics Parameter Ensembles, J. Climate, 24, 6440–6455, https://doi.org/10.1175/2011JCLI3954.1, 2011.

Zanna, L.: Forecast Skill and Predictability of Observed Atlantic Sea Surface Temperatures, J. Climate, 25, 5047–5056, https://doi.org/10.1175/JCLI-D-11-00539.1, 2012.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of Higher Climate Sensitivity in CMIP6 Models, Geophys. Res. Lett., 47, https://doi.org/10.1029/2019GL085782, 2020.

Zhang, C. and Li, S.: Causes of the record-low Antarctic sea-ice in austral summer 2022, Atmospheric and Oceanic Science Letters, 16, 100 353, https://doi.org/10.1016/j.aosl.2023.100353, 2023.

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

Zhang, X., Deser, C., and Sun, L.: Is There a Tropical Response to Recent Observed Southern Ocean Cooling?, Geophys. Res. Lett., 48, https://doi.org/10.1029/2020GL091235, 2021.

Zhao, Y., Sun, D., Di Lorenzo, E., Liu, G., and Wu, S.: Separate the Role of Southern and Northern Extra-Tropical Pacific in Tropical Pacific Climate Variability, Geophys. Res. Lett., 51, https://doi.org/10.1029/2024GL109466, 2024.

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

Zhou, C., Zelinka, M. D., and Klein, S. A.: Analyzing the dependence of global cloud feedback on the spatial pattern of sea surface temperature change with a Green's function approach, Journal

of Advances in Modeling Earth Systems, 9, 2174–2189, https://doi.org/10.1002/2017MS001096, 2017.

Zhou, C., Zelinka, M. D., Dessler, A. E., and Wang, M.: Greater committed warming after accounting for the pattern effect, Nature Climate Change 2021 11:2, 11, 132–136, https://doi.org/10.1038/s41558-020-00955-x, 2021.

Zhou, C., Wang, M., Zelinka, M. D., Liu, Y., Dong, Y., and Armour, K. C.: Explaining Forcing Efficacy With Pattern Effect and State Dependence, Geophys. Res. Lett., 50, https://doi.org/10.1029/2022GL101700, 2023.

Zhu, J. and Poulsen, C. J.: Last Glacial Maximum (LGM) climate forcing and ocean dynamical feedback and their implications for estimating climate sensitivity, Climate of the Past, 17, 253–267, https://doi.org/10.5194/cp-17-253-2021, 2021.

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

Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Poulsen, C. J., Tierney, J. E., Lofverstrom, M., and DiNezio, P.: Assessment of Equilibrium Climate Sensitivity of the Community Earth System Model Version 2 Through Simulation of the Last Glacial Maximum, Geophys. Res. Lett., 48, https://doi.org/10.1029/2020GL091220, 2021.

Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Gettelman, A., Bacmeister, J. T., Neale, R. B., Poulsen, C. J., Shaw, J. K., McGraw, Z. S., and Kay, J. E.: LGM Paleoclimate Constraints Inform Cloud Parameterizations and Equilibrium Climate Sensitivity in CESM2, Journal of Advances in Modeling Earth Systems, 14, e2021MS002 776, https://doi.org/10.1029/2021MS002776, 2022.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF operational ensemble reanalysis-analysis system for ocean and sea ice: A description of the system and assessment, Ocean Science, 15, https://doi.org/10.5194/os-15-779-2019, 2019.

ProQuest Number: 32115188

INFORMATION TO ALL USERS
The quality and completeness of this reproduction is dependent on the quality
and completeness of the copy made available to ProQuest.

Distributed by
ProQuest LLC a part of Clarivate (2025).
Copyright of the Dissertation is held by the Author unless otherwise noted.

ProQuest LLC
789 East Eisenhower Parkway
Ann Arbor, MI 48108 USA