A Modeling Perspective on State and Timescale Dependency of Surface Climate Variability

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Elisa Ziegler
aus Berlin

Tübingen
2024

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	03.02.2025
Dekan:	Prof. Dr. Thilo Stehle
1. Berichterstatter/in:	Prof. Dr. Kira Rehfeld
2. Berichterstatter/in:	Prof. Dr. Reinhard Drews

List of abbreviations

AOVGCM	Coupled Atmosphere-Ocean-Vegetation General Circulation Model
BP	before present, where present refers to 1950
CE	Common Era
CMIP	Coupled Model Intercomparison Project
EBM	Energy balance model
EMIC	Earth System Model of intermediate complexity
ENSO	El-Niño Southern Oscillation
ESM	Earth System Model
GCM	General Circulation Model
GHG	greenhouse gas (plural: GHGs)
GMP	global mean precipitation
GMST	global mean surface temperature
IPCC	Intergovernmental Panel on Climate Change
LGM	Last Glacial Maximum (23–19 kyr BP)
NAO	North Atlantic Oscillation
NH	Northern Hemisphere
PI	pre-industrial period, here: 1850–1899 CE
PMIP	Paleoclimate Modeling Intercomparison Project
PSM	proxy-system model
SH	Southern Hemisphere
SSP	shared socioeconomic pathway
yr	year

Summary

Global warming entails not only an increase in mean surface temperature, but widespread changes across all components and scales of the climate system. To understand these changes, it is necessary to consider the fluctuations around the mean and the distributions of climate variables, that is, their variability. How this variability depends on the mean climate state and evolves during periods of warming remains, in many respects, unclear. For regional surface temperature variability, models and observations provide conflicting evidence on decadal and longer timescales that suggests that models likely underestimate variability.

Here, we investigate how the variability of surface climate depends on the timescale and mean climate state. To this end, we examine the moments — standard deviation, skewness and kurtosis — of the distributions of surface temperature and precipitation, as well as their power spectra. We use an ensemble of transient simulations of the last 23,000 years, from the Last Glacial Maximum (LGM) to present-day to analyze the influence of forcings and model complexity on simulated variability. This ensemble thus covers Earth's most recent warming episode, the Last Deglaciation, during which Earth's global mean surface temperature warmed by about 4–7°C. We compare the variability in the Deglacial ensemble to that found in projections of possible future warming, as well as to reconstructions, reanalyses and direct observations.

Surface climate variability depends on timescale and background state. Annual to millennial variability is enhanced during the transitionary Deglaciation with respect to the quasi-stationary LGM and Holocene. For surface temperature, we find larger variability during the LGM than the Holocene, but the difference is smaller than in observations. For precipitation, the opposite holds, as variability is smaller during the LGM. There are large inter-model differences in skewness and kurtosis for both temperature and precipitation in the Deglacial ensemble and future projections, suggesting considerable uncertainty in the simulation of extreme conditions.

We differentiate the Deglacial simulations based on the complexity of the underlying model, from a two-dimensional Energy Balance Model (EBM) and Earth System Models of Intermediate Complexity (EMICs), to General Circulation (GCMs) and Earth System Models (ESMs). GCMs and ESMs demonstrate complex patterns of surface climate variability in space and time that change with the background state. EMICs, on the other hand, display less variability and more homogeneous patterns, while the EBM simulates at times extreme excursions in variability. Comparison to observational records shows that the complexity of GCMs and ESMs is necessary to adequately simulate variability. Beyond that, boundary conditions and applied forcings become more relevant. Variations in ice sheets, meltwater routing and volcanism affect variability on all examined timescales, from annual to millennial. The inclusion of transient volcanic forcing narrows the gap between simulated and reconstructed variability.

Overall, our results demonstrate that the ability of models to simulate variability of surface climate depends on model complexity, forcing protocol and boundary conditions. We identify requirements for simulating adequate levels of variability, from transient volcanic forcing to a minimal model complexity.

Zusammenfassung

Die globale Erwärmung hat nicht nur einen Anstieg der mittleren Oberflächentemperatur zur Folge, sondern auch weitreichende Veränderungen in allen Komponenten und auf allen Skalen des Klimasystems. Um diese zu verstehen, ist es notwendig die Fluktuationen um den Mittelwert und die Verteilungen von Klimavariablen zu verstehen, das heißt ihre Variabilität. Wie diese Variabilität vom mittleren Klimazustand abhängt und sich in Zeiten globaler Erwärmung entwickelt, ist in vielerlei Hinsicht unklar. In Hinblick auf die regionale Variabilität der Oberflächentemperatur widersprechen sich Modelle und Beobachtungen auf dekadischen und längeren Zeitskalen, da Modelle Variabilität wahrscheinlich unterschätzen.

Hier untersuchen wir, wie die Variabilität des Oberflächenklimas von Zeitskalen und dem mittlerem Klimazustand abhängt. Dazu betrachten wir die Momente — Standardabweichung, Schiefe und Wölbung — der Verteilungen von Oberflächentemperatur und Niederschlag, sowie ihre Leistungsdichtespektren. Wir verwenden ein Ensemble transienter Simulationen der letzten $23,000$ Jahre, vom letzten Glazialen Maximum (LGM) bis heute, um die Auswirkungen von Antriebskräften und Modellkomplexität auf die simulierte Variabilität zu analysieren. Dieses Ensemble deckt die letzte Erwärmung der Erde ab, die letzte Deglaziation, während der sich die globale Oberflächentemperatur der Erde um 4–7°C erhöht hat. Wir vergleichen die Variabilität des deglazialen Ensembles mit Projektionen möglicher zukünftiger Erwärmung, sowie mit Rekonstruktionen, Reanalysen und direkten Beobachtungen.

Die Variabilität des Oberflächenklimas hängt von Zeitskala und mittlerem Zustand ab. Jährliche bis tausendjährige Variabilität ist in der Übergangsphase Deglaziation im Vergleich zum quasistationären LGM und Holozän erhöht. Für Oberflächentemperatur finden wir größere Variabilität während des LGMs als im Holozän, aber der Unterschied ist kleiner als in Rekonstruktionen. Für Niederschlag trifft das Gegenteil zu. Es gibt große modellübergreifende Unterschiede in Schiefe und Wölbung für Temperatur und Niederschlag in Deglaziation und in Zukunftsprojektionen, was auf beträchtliche Unsicherheit in der Simulation von Extremen hindeutet.

Wir unterscheiden die Simulationen der Deglaziation nach der Komplexität des zugrundeliegenden Modells, von einem zweidimensionalen Energiebilanzmodell (EBM) und Erdsystemmodellen mittlerer Komplexität (EMICs) zu Globalen Zirkulationsmodellen (GCMs) und Erdsystemmodellen (ESMs). GCMs und ESMs zeigen vom mittleren Klima abhängende, komplexe räumliche und zeitliche Muster in der Variabilität des Oberflächenklimas. EMICs simulieren geringere Variabilität und homogenere Muster, während das EBM bisweilen extreme Exkursionen simuliert. Ein Vergleich zu Beobachtungsdaten zeigt, dass die Komplexität der GCMs und ESMs zur Simulation hinreichender Variabilität notwendig ist. Darüber hinaus werden Randbedingungen und Antriebskräfte wichtiger. Variationen in Eisschilden, Schmelzwasser und Vulkanismus beeinflussen Variabilität auf allen untersuchten Zeitskalen. Eine transiente vulkanische Antriebskraft verringert die Lücke zwischen simulierter und rekonstruierter Variabilität.

Insgesamt zeigen unsere Resultate, dass die Fähigkeit von Modellen die Variabilität des Oberflächenklimas zu simulieren, von deren Komplexität, verwendeten Antriebskräften und Randbedingungen abhängt. Wir identifizieren Voraussetzungen zur Simulation ausreichender Variabilität, von transienter vulkanischer Antriebskraft zu einer minimalen Modellkomplexität.

List of publications

Accepted publications

Publication 1 (P1): E. Ziegler & K. Rehfeld (2021): TransEBM v. 1.0: Description, tuning, and validation of a transient model of the Earth's energy balance in two dimensions, Geoscientific Model Development, 14. doi:10.5194/gmd-14-2843-2021

Publication 2 (P2): T. Laepple, **E. Ziegler**, N. Weitzel, R. Hébert, B. Ellerhoff, P. Schoch, B. Martrat, O. Bothe, E. Moreno-Chamarro, M. Chevalier, A. Herbert & K. Rehfeld (2023): Regional but not global temperature variability underestimated by climate models at supradecadal timescales. Nature Geoscience, 16. doi:10.1038/s41561-023-01299-9

Publication 3 (P3): E. Ziegler, N. Weitzel, J.-P. Baudouin, M.-L. Kapsch, U. Mikolajewicz, L. Gregoire, R. Ivanovic, P. J. Valdes, C. Wirths & K. Rehfeld (2025). Patterns of changing surface climate variability from the Last Glacial Maximum to present in transient model simulations. Climate of the Past, 21. doi:10.5194/cp-21-627-2025

Data and code availability

The code and data for **P1** – **P3** are available with the publications, as detailed in the respective statements on data and code availability. The repository `https://github.com/paleovar/simulatedSurfaceClimateVariability` (access upon request) contains all figures, as well as the code and data for all additional figures and analyses presented in this thesis.

Declaration of author contributions

P1: E. Ziegler & K. Rehfeld

On publication one, I was the main author. I implemented the extension of the EBM by Zhuang et al. (2017) for transient simulations. I designed and ran all simulations with the model, analyzed the data created all figures and wrote the manuscript. KR supervised all the work and together we conceptualized the study and revised the manuscript.

P2: T. Laepple, **E. Ziegler**, N. Weitzel, R. Hébert, B. Ellerhoff, P. Schoch, B. Martrat, O. Bothe, E. Moreno-Chamarro, M. Chevalier, A. Herbert & K. Rehfeld

On this publication, I am the second author with a significant contribution. I led the literature analysis and compiled the detailed results and evaluation of the evidence in the Supplementary Table 1. I wrote and revised the introductory section, with the exception of the first paragraph, which was written by RH. Based on the literature analysis, PS and I designed Figure 2 and Extended Data Figure 2, the code for which was written by PS and subsequently revised by me. BE and I together conceptualized and wrote the section on the "implications for climate projections and attribution efforts" and led the discussion of it with all co-authors. Together with BE, I wrote the glossary (Box 1). BE, RH and I wrote the methods. I further coordinated the study as a whole.

The project was initiated by KR and TL, who led the discussion. BE and RH created Fig. 1, RH and TL Fig. 3 and BE Fig. 4. NW and TL wrote the section on reconstruction deficiencies. NW, EMC and TL wrote the section on the consequences for the spatial structure, while RH led the writing of the concluding section. All authors reviewed the manuscript. NW, BE, RH, TL and I revised the manuscript in detail.

P3: E. Ziegler, N. Weitzel, J.-P. Baudouin, M.-L. Kapsch, U. Mikolajewicz, L. Gregoire, R. Ivanovic, P. J. Valdes, C. Wirths & K. Rehfeld

I am the first author of this study. I conceptualized the study under the supervision of KR and with input by NW. I carried out the analysis, visualized the results, wrote and revised the manuscript. KR, NW, MK, JB, RI and I reviewed the manuscript in detail. All authors discussed the analysis. ML and UM designed and created the MPI-ESM simulations r1–r7. LG, RI and PV designed and ran the two HadCM3B simulations. CW adapted TransEBM from **P1** and ran the Deglacial EBM simulation.

Table of Contents

1 Introduction

Earth's climate is an intricate system, its elements interconnected, interacting and constantly changing. As a result, even subtle changes in one part can reverberate far beyond their place and time of origin. The incoming radiation from the Sun, which is modified by the Earth's orbital configuration, tectonic shifts and volcanic activity drive the complex climatic history of Earth (Fig. 1.1). Taking global mean surface temperature (GMST) as an indicator[1], this history includes climates warmer than today, for example around the Paleocene-Eocene thermal maximum roughly 55 Myr ago, as well as colder climates (Marshall and Plumb, 2008; Zachos et al., 2008). At the Pliocene–Pleistocene transition about 2.58 Myr ago, the seaway separating North and South America closed. This marks the beginning of the Quaternary, a period of repeated cycling between glacials and interglacials (Head et al., 2008; Ogg and Pillans, 2008). Around that time in Earth's history, about 2.8 Myr ago, the genus homo emerged in East and South Africa (Dunsworth, 2010; Villmoare et al., 2015). Generally, global mean surface temperatures during the Quaternary were colder than pre-industrial (PI, 1850–1899) temperatures as Earth's climate went through prolonged cooling episodes into glacial states. These cold periods were followed by rapid warming events and relatively shorter interglacials during which GMST occasionally exceeded PI levels (Fig. 1.1). Throughout the Quaternary, GMST fluctuated by about 10°C and atmospheric carbon dioxide (CO_2) ranged from about 170 to 300 parts per million (ppm, Fig. 1.1, 1.2, Marshall and Plumb, 2008). The anatomically modern human, homo sapiens, then evolved around 200 kyr ago (Brown et al., 2012; Mounier and Mirazón Lahr, 2019)[2].

Within these ranges, climate fluctuated considerably across timescales, for example from the annual cycle to orbital variations on several ten thousands or even a hundred thousand years (Franzke et al., 2020; Huybers and Curry, 2006; Mitchell, 1976). For a specific timescale, climate can then be described as the average conditions, that is the trend or mean, and the fluctuations around it, the variability (Bradley, 2015). The fluctuations around the mean are captured in the overall shape of the distribution of a climatic parameter like surface temperature (Bradley, 2015). Climate variability usually changes alongside the average conditions (Brown et al., 2017; Rehfeld et al., 2020; Shao and Ditlevsen, 2016). For example, reconstructions from various climate archives suggest that global surface temperature variability decreased approximately by a factor of four between the Last Glacial Maximum (LGM) around 21 kyr ago and the current interglacial, the Holocene (Rehfeld et al., 2018). However, many questions remain regarding

[1] The term climate encompasses the whole climate system and all its observables, two of which are surface temperature and precipitation, on which much of this work focuses. GMST provides a commonly used and important indicator of the climate system's state that helps differentiate between different periods of Earth's climatic history.

[2] An earlier appearance of homo sapiens remains possible (Bergström et al., 2021) as fossils from 350 to 200 kyr ago are morphologically diverse (Bergström et al., 2021; Hublin et al., 2017; Stringer and Galway-Witham, 2017) and genetic divergences have been found during this time period as well (Schlebusch et al., 2017).

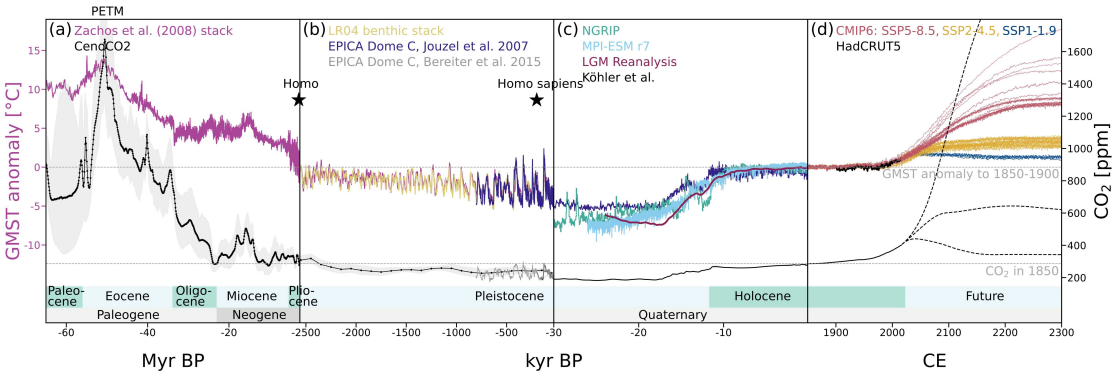


Figure 1.1: Reconstructed, observed and projected climate history of Earth spanning the last 65 Myr as captured by the global mean surface temperature (GMST) anomaly to pre-industrial (PI, 1850–1899 CE) levels and atmospheric CO_2.
(a) 65–2.56 Myr before present (BP) including the Paleogene and Neogene period with temperature data from Hansen et al. (2013) based on the stack by Zachos et al. (2008) and CO_2 data from the Cenozoic CO2 proxy integration project (CENCO2PIP) consortium (2023).
(b) Early Quaternary, 2560–30 kyr BP, GMST and CO_2 data from panel (a) as well as GMST reconstructions from Lisiecki and Raymo (2005) and EPICA Dome C (Jouzel et al., 2007) and reconstructed CO_2 from EPICA Dome C (Bazin et al., 2013; Bereiter et al., 2015; Monnin et al., 2001; Schwander et al., 2001).
(c) Last Glacial Maximum (LGM) to PI, 30 kyr BP – 1850 CE. Added to the GMST and CO_2 reconstructions from EPICA Dome C are temperature records from NGRIP (Andersen et al., 2004), the LGM reanalysis (Osman et al., 2021) and an MPI-ESM simulation (Kapsch et al., 2022, 2021, r7 in **P3**), as well as the CO_2 reconstruction by Köhler et al. (2017). The EPICA and NGRIP proxy records for local temperature were scaled according to GMST= $0.5 \times T_{local}$.
(d) 1850–2300 CE with temperature observations from HadCRUT5 (Morice et al., 2021), simulated temperatures according to all CMIP6 ensemble members extending to 2300 CE for low to high emission scenarios SSP 1-1.9, 2-4.5 and 5-8.5 (cf. Sec. 1.2.1). Observed and projected CO_2 forcings according to Meinshausen et al. (2017).

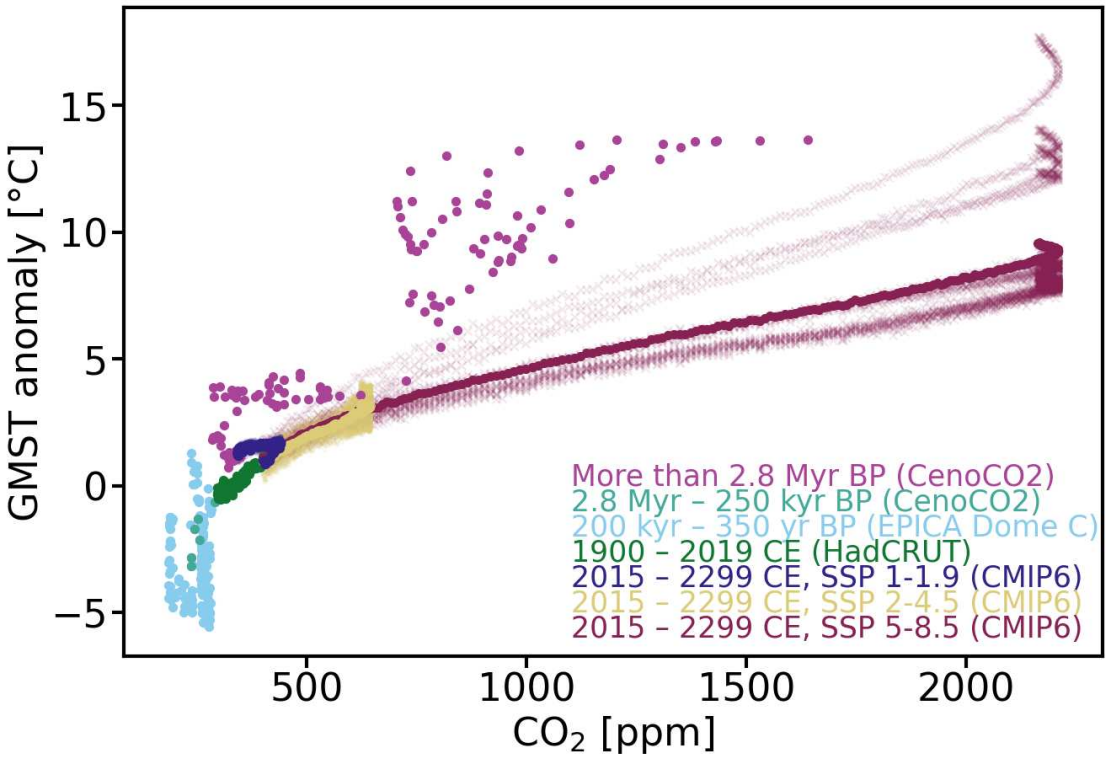


Figure 1.2: GMST anomaly change with atmospheric CO_2 in reconstructions and models. The anomaly is computed with respect to pre-industrial GMST (1850–1899 CE). Reconstructions of past CO_2 and GMST are from the Cenozoic CO2 proxy integration project (CENCO2PIP) consortium (2023) and EPICA Dome C (Bazin et al., 2013; Jouzel et al., 2007; Monnin et al., 2001; Schwander et al., 2001). The proxy records for local temperature data from EPICA Dome C were scaled to GMST using GMST= $0.5 \times T_{local}$ and then linearly interpolated to the temporal resolution of the CO_2 record. CO_2 after 1900 CE is the compiled observed and projected forcing used in CMIP6 and the SSP scenarios (Meinshausen et al., 2017). GMST from direct observations is from HadCRUT5 (Morice et al., 2021) and simulated future GMST is the ensemble mean of all CMIP6 simulations extending to 2300 CE for three scenarios as in Fig. 1.1 (cf. Sec. 1.2.2, 1.2.1). Individual ensemble members from CMIP6 shown in lighter shading.

the relationship between changes in the mean state and accompanying changes in variability and how variability depends on timescale and forcings (Franzke et al., 2020). Variability further varies in space. For Greenland, for example, ice cores suggest a decrease in surface temperature variability by a factor of 73 between LGM and Holocene, whereas it is only a factor of 1.6–2.8 in the tropics (Rehfeld et al., 2018). Thus, the changes in average climate throughout homo sapiens' history were also accompanied by substantial changes in variability.

In the near future, Earth is likely to leave the climate range of the past million years as humans have emitted about 695 ± 70GtC into the atmosphere between 1850 and 2022 Common Era (CE) (Friedlingstein et al., 2023). This has lead to a rise in atmospheric CO_2 from 286 ppm (Köhler et al., 2017) in 1850 to 417 ppm in 2022 (Friedlingstein et al., 2023). Atmospheric CO_2 levels of this magnitude last occurred at least 2 Myr ago (Canadell et al., 2021) and emissions continue to rise (Friedlingstein et al., 2023). As a greenhouse gas (GHG), CO_2 absorbs longwave radiation emitted from Earth's surface. This radiation gets re-emitted as electrons in the CO_2 molecules return from their excited to their ground state. Since re-emission occurs in all directions with equal probability, much of this radiation remains in the atmosphere. There it is again absorbed by other GHGs or reflected back to the surface, reducing the amount of longwave radiation escaping into space. As the concentrations of GHGs in the atmosphere increase, this effect becomes stronger, further reducing the amount of outgoing radiation. The resulting change in Earth's energy balance has taken Earth to a GMST $1.45 \pm 0.12\,^{\circ}$C above the PI mean (WMO, 2024). GMST was last at a similar level during the Last Interglacial (130–115 kyr ago, IPCC, 2021d). It is set to rise further depending on future emissions and human activities, except in the most ambitious mitigation and adaptation scenarios that employ negative emissions, which remove CO_2 from the atmosphere (IPCC, 2021c). Depending on the eventual emissions pathway, this could lead to a similar degree of warming by the end of the century as occurred during the Last Deglaciation (around 19–11 kyr ago), when GMST increased by about 4–7 °C (Annan et al., 2022; Gulev et al., 2021; Osman et al., 2021; Tierney et al., 2020).

Altering the Earth's energy balance and atmospheric composition affects all major components of the climate system (Masson-Delmotte et al., 2021). Impacts include the warming of the surface oceans, increasing ocean acidification and deoxygenation, changes in atmospheric circulation and storm systems, as well as changes to various extreme events and the modes of variability (IPCC, 2021c). The scope of these changes depends on warming levels and thus emission pathways (IPCC, 2021c). Variability, too, depends on mean warming and changes in variability have cascading effects in the climate system. For instance, variability determines the occurrence and frequency of extreme events (**P2**, Ionita et al., 2021; Katz and Brown, 1992; Schär et al., 2004; Simolo and Corti, 2022). Many applications require a reliable representation of variability in models such as simulated projections of future climate (Deser et al., 2012; Hawkins and Sutton, 2009), impact studies (Calel et al., 2020; Schwarzwald and Lenssen, 2022) and thus mitigation and adaptation efforts (Lempert et al., 2000). Attribution, too, relies on simulated internal variability to determine the likelihood of an extreme event not occurring in a world without anthropogenic climate change (Harrington et al., 2021; Hegerl and Zwiers, 2011; Van Oldenborgh et al., 2022). However, the representation of variability in climate models has been called into question as models are generally found to be too stable (Valdes, 2011) and simulate for example too little surface temperature variability in comparison to observations (**P2**). This holds in

particular for variability on decadal and longer timescales, on which observed and simulated variability disagree during the instrumental era (Laepple and Huybers, 2014a; Parsons et al., 2017) and in past climates (Hébert et al., 2022; Laepple and Huybers, 2014b; Rehfeld et al., 2018). Altogether, projections of future climate (Bethke et al., 2017; Deser et al., 2012) and impacts (Calel et al., 2020) likely lack variability. As such and in the face of anthropogenic warming, improving our understanding of variability and its representation in models has taken on increased urgency.

This thesis therefore presents an investigation of variability with a focus on its representation in climate models. We explore the dependence of variability on the mean climate state and timescale, from annual up to multi-millennial (**P2**, **P3**, Sec. 3.1.1, 3.1.2). To this end, variability in various climates is examined, from past climates like the Last Deglaciation (**P2**, **P3**) to potential future climate under various emission scenarios (Sec. 3.3). Examining potential reasons for discrepancies between modeled and observed temperature variability during the Holocene indicates that models lack variability, as the mismatch cannot be explained by potential sources of uncertainty in reconstructions and direct observations alone (**P2**). We expand the comparison of simulated and observed variability with the Deglacial ensemble (**P3**, Sec. 3.2). While investigating possible reasons for the lack of simulated variability, we examine the impact of forcings and forcing protocols and simulated variability (**P1**, **P3**, Sec. 1.1.2). As variability differs considerably between models and long transient simulations with state-of-the-art models require considerable computational resources, we explore variability in models of varying complexity to identify a minimal complexity needed for the representation of variability (**P1**, **P3**, Sec. 3.1.4). Lastly, we compare variability under past Deglacial warming to that under potential future warming for low to high emission scenarios (Sec. 3.3).

Earth's climate varies on all scales in space and time, from the hours necessary to form convection cells to the hundreds of millennia over which plate tectonics rearranges Earth's surface, from the scale of molecules absorbing radiation to the whole planet. As such, variability can be observed in all components of the climate system (Sec. 1.1.1). It arises both in response to external forcings (Sec. 1.1.2) and internal dynamics (Sec. 1.1.1). While variability has changed and will continue to change throughout the history of Earth's climate, much less is known about its evolution than that of the mean climate (Sec. 1.1.3). The analyses focus on climate model simulations (Sec. 1.2.1), but also draw on data from observations and reanalyses (Sec. 1.2.2). Here, we investigate variability using power spectra (Sec. 1.3.2) and the moments of the distributions (Sec. 1.3.1) of surface climate variables.

1.1 Climate and climate variability

The climate system is a complex dynamical system of interacting nonlinear and multi-scale processes and components. Continuously driven by the incoming radiation from the Sun, which depends on the strength of solar radiation and Earth's position in the Sun's orbit, tectonic processes and internal dynamics, it is an open and dissipative system. Its state, the climate[3] can

[3]Climate is often used more specifically to refer to aggregate weather, that is the long-term — commonly 30yr — mean and variability of meteorological variables.

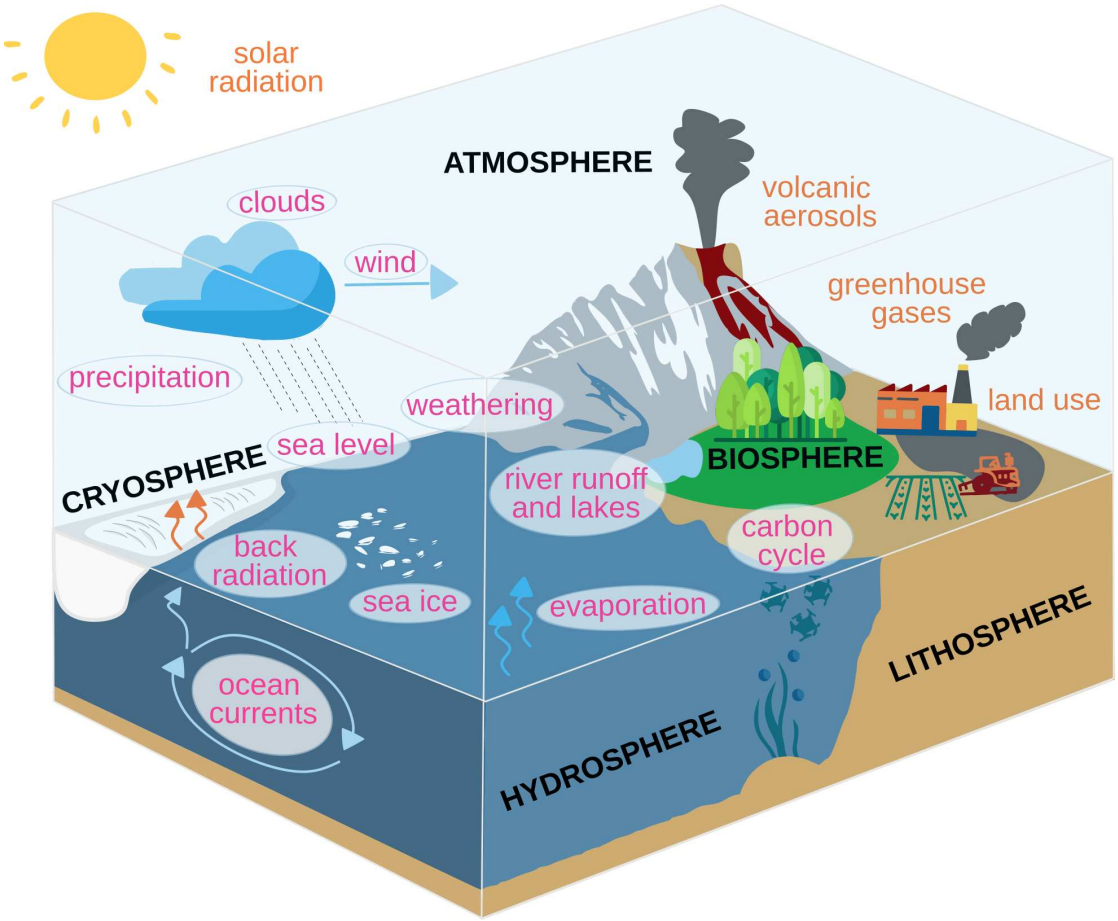


Figure 1.3: Compartments (black), forcings (orange) and processes (pink) in the climate system. Figure adapted from Ellerhoff (2023), courtesy of Beatrice Ellerhoff.

be described by various observables like surface temperature or precipitation, which evolve in time. The climate of a certain period and timescale is then the average state of the climate system as characterized by the averages of climatic variables and their variability. Often mean climate states refer to periods of relatively stable GMST. However, this purported stability can hide considerable variations, for example regionally or in other variables and components of the climate system not reflected in GMST changes.

1.1.1 Internal variability of the climate system

The climate system is generally described as consisting of five interacting, connected and overlapping compartments: atmosphere, hydrosphere, cryosphere, lithosphere and biosphere (Fig. 1.3). The **atmosphere** is the gaseous envelope of Earth kept in place by gravity. It is quick to react to disturbances and provides the fastest connection between geographically distant regions (Hartmann, 2016). Weather patterns, for example, develop on hourly to daily timescales

(Baede et al., 2002). Its presence and composition maintain temperatures suitable for life and the presence of liquid water (Hartmann, 2016). The composition of the atmosphere has changed considerably throughout Earth's history. Today, dry air contains mostly molecular nitrogen (N_2 at 78% volume mixing ratio), oxygen (O_2, 21%) and argon (Ar, 1%, Hartmann, 2016). These barely interact with incoming solar radiation and not at all with outgoing longwave radiation emitted from Earth's surface. As such, they have a limited impact on Earth's energy balance (Baede et al., 2002). However, there are numerous trace gases, altogether making up less than 1% of the atmospheric composition. Among them, a number of so-called greenhouse gases absorb and emit infrared, that is longwave, radiation, most notably water vapor (H_2O), CO_2, methane (CH_4), nitrous oxide (N_2O) and ozone (O_3), affecting the energy balance and raising Earth's surface temperature as a result (Hartmann, 2016). Some of them can also absorb incoming solar shortwave radiation (Baede et al., 2002). Through this greenhouse effect, increasing concentrations in GHGs lead to rising temperatures. Warmer temperatures further increase the atmosphere's water holding capacity (Hartmann, 2016). As water vapor is the most potent GHG, this can lead to a positive feedback reinforcing initial warming (Hartmann, 2016). Trends in atmospheric temperature change with height, that is in the lowest layer, the troposphere, temperature decreases with height, whereas in the stratosphere above it, the opposite is the case. The different layers are affected differently by anthropogenic GHG emissions. Warming is generally observed near the surface and in the troposphere, while the stratosphere has been found to cool (IPCC, 2021d).

The atmosphere not only plays a critical role in the overall energy balance of Earth, but also in the transport of heat and moisture. Circulation in the atmosphere works to reduce the meridional temperature gradient resulting from the latitudinal and seasonal patterns of incoming solar radiation, as well as the temperature gradients with altitude. This redistribution of energy happens from scales as small as millimeters to tens of thousand of kilometers (Hartmann, 2016). Thunderstorms and other relatively small-scale phenomena tend to transport energy upward, while large-scale phenomena like extra-tropical cyclones carries heat, momentum and energy from the tropics towards the poles (Hartmann, 2016). Circulation patterns get affected by many factors like differences in surface properties, instabilities or Earth's rotation. As dampening due to friction is low, energy can get transferred across scales as instabilities get amplified non-linearly (Hartmann, 2016). In particular in the mid-latitudes, atmospheric circulation patterns vary on timescales of hours to seasons and dominate surface temperature and precipitation variability on these timescales (Hartmann, 2016).

The **hydrosphere** comprises all surface and sub-surface liquid water. The oceans dominate Earth as they contain about 97% of all water on Earth and cover about 71% of its surface (Hartmann, 2016). Through evaporation at their surfaces they act as a source of moisture in the hydrological cycle. Most of the remaining water, at 2.2% is frozen and considered part of the cryosphere, with a further 0.7% available as groundwater, 0.013% stored as soil moisture and about the same amount found in rivers and lakes (Hartmann, 2016). Only about 0.0009% of all water is in the atmosphere (Hartmann, 2016). The transport of heat and nutrients is determined by density differences due to temperature and salinity, that is the amount of dissolved salts like chloride, sodium or calcium (Hartmann, 2016). At the top of the ocean, there is a well-mixed surface layer of near constant temperature, the depth of which can change during

the year (Hartmann, 2016). Below the mixed layer, temperature decreases with depth until close to the freezing point of ocean water in the thermocline (Hartmann, 2016). The deep ocean below the thermocline is characterized by roughly uniform temperatures (Hartmann, 2016). While the mixed layer responds to disturbances on timescales from days to months, the thermocline takes years to centuries and the deep ocean decades/centuries to multiple millennia (Marshall and Plumb, 2008; Ruddiman, 2008). The mixed layer integrates small variations such that sea surface temperature variability increases with timescale (Hasselmann, 1976). The oceans play a crucial role as a reservoir of heat and carbon, buffering much of the anthropogenic carbon emissions as they have taken up about 91% of the excess heat and carbon (IPCC, 2021d). This has led to a warming of the surface ocean, increased stratification of the surface ocean, ocean acidification and ocean deoxygenation, impacting ecosystems like coral reefs (IPCC, 2021c).

As for the atmosphere, circulation patterns in the ocean contribute to climate variability across timescales. On sub-annual scales, eddy transport likely drives temperature variability (Klavans et al., 2017). The North Atlantic and the North Pacific, on the other hand, are major drivers of Northern Hemisphere temperature variability on multidecadal timescales (Steinman et al., 2015). One of the largest sources of oceanic variability, mostly on centennial scales, is The Atlantic Meridional Overturning Circulation (AMOC) (Bakker et al., 2022; Knight et al., 2005). Driven by temperature and density differences, the upper ocean layers transport relatively fresh, warm water poleward with a compensating southward flow of colder, saltier water in the deep Atlantic ocean. The AMOC thus contributes to the thermohaline circulation in the ocean. Its reorganization has been linked to abrupt millennial warming events during glacials, as the AMOC cycled through periods of lessened and heightened activity (Sévellec and Fedorov, 2015).

The **cryosphere** covers all ice, that includes the ice sheets, ice shelves, sea ice, glaciers and permafrost (Goosse et al., 2010a). Despite making up only 2% of all water available on Earth, it is the largest freshwater reserve, as it contains 80% of all freshwater (Hartmann, 2016). Most of it makes up the ice sheets covering Antarctica (89%) and Greenland (8.6%) (Hartmann, 2016). Their growth during glacial periods and shrinking when the climate warms influences sea level and salinity and thus ocean circulation. They further impact Earth's radiative balance through their high albedo, which results in ice covered areas reflecting more sunlight. Thus, the larger the surface area that ice covers the more it alters the surface energy balance and contributes to decreasing surface temperatures. Melting ice, on the other hand, creates a positive feedback loop, as the darker land revealed by melting ice absorbs more energy (Goosse et al., 2018). Converted into heat, the resulting warming further promotes ice melting (ice-albedo feedback) (Goosse et al., 2018). Sea ice further insulates seawater from atmosphere above, restricting the exchange of heat, mass, momentum and chemical components between ocean and atmosphere (Divine and Dick, 2006; Hartmann, 2016). Within the cryosphere, sea ice responds fastest to changes, on timescales of weeks to years (Bradley, 2015; Marshall and Plumb, 2008; Ruddiman, 2008). The characteristic timescales of mountain glaciers, on the other hand, range from decadal to centennial and those of land ice stretch from centennial to tens or hundreds of thousands of years (Bradley, 2015; Marshall and Plumb, 2008; Ruddiman, 2008). As a result of current anthropogenic warming, glaciers are expected to melt for decades or even centuries to come (IPCC, 2021d).

Changes in the cryosphere add substantial variability to the climate system. As sea ice cover

decreases, so does daily to multi-decadal temperature variability (Collow et al., 2019; Dai and Deng, 2021; Huntingford et al., 2013). This decrease implies a narrowing of the probability distribution, which further results in a reduction in the frequency and amplitude of both cold and warm extremes (Collow et al., 2019; Dai and Deng, 2021). (Olonscheck et al., 2021) paint a more nuanced picture, their results showing an initial increase in interannual temperature variability in higher latitudes as seasonal variations in sea ice become more pronounced. Once sea ice disappears throughout the year, though, they, too, find decreasing variability. In many areas, except the Arctic, simulations have further shown that the decrease in sea ice variability is proportional to that of mean sea ice volume and area (Olonscheck and Notz, 2017). As the sea ice disappears, the meridional temperature gradient weakens, which might explain the resulting reduction in variability with warming (Collow et al., 2019; Ellerhoff et al., 2022; Huntingford et al., 2013). This connection between sea ice and atmospheric temperature variability runs both ways, as temperature variability is also the dominant driver of sea ice variability (Olonscheck et al., 2019).

The **lithosphere** is the uttermost layer of rocks of Planet Earth. As such it comprises the oceanic and continental crust as well as Earth's upper mantle. The continental land surface covers 29% of Earth's surface, but its configuration and topography are crucial in determining climate on Earth (Hartmann, 2016; Stocker, 2011). In its current configuration, 68% of the land mass is in the northern hemisphere, influencing among others circulation patterns in atmosphere and ocean (Hartmann, 2016). Incoming solar radiation gets partially reflected at the land surface, depending on the local albedo of the surface and partially heats the surface. The land surface itself then emits longwave radiation into the atmosphere. While changes spread within hours to months on the land surface (Ruddiman, 2008), many processes of the lithosphere have far longer characteristic timescales. For example, the weathering of rocks, a major player in the carbon cycle, unfolds its influence on timescales of tens of thousands of years to multi-millennia (Bradley, 2015; Marshall and Plumb, 2008).

All living organisms form the **biosphere**, including the vegetation and organisms on land and in the oceans. Vegetation, in particular, modifies the albedo of the surface and plays a major role in the hydrological and carbon cycles. Characteristic timescales of vegetation, as well as the biosphere overall, cover a wide range, as vegetation can react to extreme conditions within hours or days whereas the evolution of new species can take multiple millennia (Bradley, 2015; Marshall and Plumb, 2008; Ruddiman, 2008).

All these compartments are complex in their own rights and interact with each other. For example, atmosphere and ocean exchange gases, momentum and heat at the ocean's surface, the availability of soil moisture is crucial for vegetation, and the flow of freshwater into the oceans as ice sheets melt modifies ocean circulation patterns. The interactions of the processes within and between the compartments, often non-linear in nature, introduce substantial variability into the climate system across all temporal and spatial scales (Ghil, 2002).

A notable expression of this so-called **internal variability** are the modes of variability. These are recurring patterns of climate variability in space and time (IPCC, 2021a). Among them is for example the El-Niño Southern Oscillation (ENSO) in the Pacific: it is a coupled oceanic (EN) and atmospheric (SO) mode that dominates interannual variability. Its effects are strongest in the

tropics, but it affects global climate. Two phases exist that it cycles through with semi-regularity every 2–7 yrs. El-Niño is a phase of warm sea surface temperatures in the eastern equatorial Pacific, which turn colder than usual during the La Niña phase. In the atmosphere, corresponding patterns of surface pressure exist, which affects winds and rainfall. Further modes of variability are for example the North Atlantic Oscillation (NAO) and the Northern and Southern Annular Modes. The modes of variability show a clear co-varying structure in space, which varies in time without an external driving force (Braconnot et al., 2012). For surface temperatures, increasing internal variability has been linked to stronger covariance between temperatures across larger regions as modes of variability, for example, have an intrinsic spatial structure (Bakker et al., 2022; PAGES2k-PMIP3 group, 2015). On the other hand, an increase in local internal variability might also disrupt such teleconnections. The modes additionally change with the mean state (Brown et al., 2017; Rehfeld et al., 2020).

While variability can thus be generated internally in the climate system (internal variability), it can also be driven externally (externally-forced variability).

1.1.2 Drivers of climate variability

The climate system is continuously driven by external forces. Among them is the pattern of incoming insolation, which depends both on the strength of solar irradiation and Earth's orbit relative to the Sun. Additionally, processes within Earth's core lead among others to volcanic activity and the tectonic processes that rearrange Earth's surface.

Volcanic activity directly affects climate as an eruption catapults various gases and ash into the stratosphere (Timmreck, 2012). For the climatic response gases containing sulfur are particularly relevant. Turned into sulfate aerosols they scatter short- and absorb longwave radiation, leading to an overall decrease of solar radiation reaching the Earth's surface (Hopcroft et al., 2018; Timmreck, 2012). This leads to a cooling of surface temperatures that persists for several years (Lovejoy and Varotsos, 2016; Schurer et al., 2014; Swingedouw et al., 2015; Timmreck, 2012). For repeated or particularly strong volcanic eruptions, this cooling can persist for decades (Berdahl and Robock, 2013; Timmreck, 2012). Besides the strength of an eruption, its timing and location determine its climatic impacts (Stevenson et al., 2017; Timmreck, 2012). Tropical eruptions, for example, likely affect both hemispheres unlike those in mid- and high latitudes (Timmreck, 2012) and their effects depend on the eruption season (Toohey et al., 2011).

Further impacts can be observed across the climate system, as volcanic eruptions disrupt the hydrological cycle (Liu et al., 2016; Stevenson et al., 2017; Trenberth and Dai, 2007), generally reducing precipitation (Gu and Adler, 2011; Iles and Hegerl, 2014, 2015), affect the cryosphere (Berdahl and Robock, 2013; Zanchettin et al., 2013) and influence atmosphere (Robock and Mao, 1992; Santer et al., 2015; Stevenson et al., 2017), as well as ocean dynamics and circulation patterns (Ding et al., 2014; Swingedouw et al., 2015; Zanchettin et al., 2013). Through feedbacks and internal dynamics, perturbations can remain beyond the residence time of the aerosols in the atmosphere, as salinity anomalies or sea ice can build up leading to persistent cool summers and temperature anomalies in the deep ocean that can endure up to centennial timescales (Berdahl and Robock, 2013; Ding et al., 2014; Miller et al., 2012; Swingedouw et al., 2015;

Timmreck, 2012). However, there is a large spread in simulated responses to volcanic forcings (Swingedouw et al., 2017; Zanchettin et al., 2016). Additionally, the effects of volcanic forcing on decadal and longer timescales are uncertain due to a lack of longterm observational records and large inter-model differences (Swingedouw et al., 2017; Timmreck, 2012; Zanchettin et al., 2016).

Impacts of volcanic activity on variability can also be observed across timescales, from daily (Wang et al., 2021) to annual (Bethke et al., 2017; Schurer et al., 2014) and interannual (Le et al., 2016), up to decadal (Bethke et al., 2017; Hegerl and Zwiers, 2011; Hegerl et al., 2007; Schurer et al., 2014) and multidecadal timescales (Otterå et al., 2010; Schurer et al., 2014). For daily timescales these impacts include changes to surface temperature and precipitation extremes, as eruptions generally lead to less hot or wet extremes and more cold extremes, but with regional differences (Wang et al., 2021). Impacts on the longer timescales have been confirmed for both strong (Timmreck, 2012) and weak (Ellerhoff et al., 2022) volcanic forcing as short-term effects are integrated by non-linear feedbacks to longer timescales. During Glacials, volcanic forcing has even been suggested as a driver of millennial variability (Baldini et al., 2015). Changes in atmospheric (Lehner et al., 2016; Maher et al., 2015; Swingedouw et al., 2017; Zanchettin et al., 2015) and oceanic modes (Swingedouw et al., 2015) further highlight volcanism's significance for climate variability.

There is some evidence that the climate's response to volcanic eruption can depend on the mean climate state at the time of eruption (Hopcroft et al., 2018; Muthers et al., 2015; Timmreck, 2012). This holds for the surface temperature response to volcanic forcing, which shows a dependence on the initial state, especially that of the ocean, at the time of eruption (Zanchettin et al., 2013). Fasullo et al. (2017) found the cooling response to the 1815 Tambora eruption enhanced when applying that eruption in a simulation of a warmer future climate. Simulations of the same eruption by Hopcroft et al. (2018) showed the opposite with reduced cooling in a warmer future, confirming the uncertainties in simulated responses to volcanic forcing (Zanchettin et al., 2016). For surface temperature variability, Ellerhoff et al. (2022) found no state-dependency globally. Locally, they found only modest differences, mostly on interannual scales and towards higher latitudes. A state-dependent response to volcanic forcing has further been found for several components of the climate system, e.g., sea ice (Berdahl and Robock, 2013; Zanchettin et al., 2013), stratospheric ozone (Muthers et al., 2015), as well as atmosphere (Muthers et al., 2015) and ocean dynamics (Swingedouw et al., 2015; Zanchettin et al., 2013).

The inclusion of volcanic forcing thus has widespread consequences across states. However, it is often misrepresented, e.g., assumed to be constant, when simulating past or future climates, largely due to uncertainty in reconstructions (Ellerhoff et al., 2022; Kageyama et al., 2018; O'Neill et al., 2016). However, along with the inclusion of time-varying solar irradiance, volcanic forcing reduces model-reconstruction mismatches between LGM and PI simulations (Ellerhoff et al., 2022). Additionally, volcanic forcing adds significant variability on interannual to decadal scales in projections of future climate, suggesting that variability is underestimated in projections without volcanic forcing (Bethke et al., 2017).

Simplified, variations in incoming solar radiation can be described as a solar constant with an 11yr-cycle and can be described as such in models. In fact, the intensity of incoming radiation

has increased by about 30% throughout Earth's history (Hartmann, 2016) such that the total solar irradiance at the top of the atmosphere today is 1361 W/m^2 (Wild et al., 2013). Both direct observations and reconstructions show that the strength of irradiation varies across timescales, from minutes to the order of 10^9 yr (Goosse et al., 2010a; IPCC, 2021b; Marshall and Plumb, 2008). On decadal timescales, a cycle in the Sun's activity with an average length of 11 years is associated with the increase and waning of the amount of sunspots[4]. As a result, the solar irradiation varies by about 0.08% over one cycle (Bradley, 2015). Based on reconstructions from fossils, this cycle is thought to have existed for at least the past 300 Myr (Luthardt and Rößler, 2017). Various longer additional cycles in solar activity have been proposed, but longer term changes in solar irradiation remain poorly understood (Luthardt and Rößler, 2017; Marshall and Plumb, 2008; Ruddiman, 2008). However, including realistic solar forcing in simulations increases surface temperature variability from annual to multi-decadal timescales (Ellerhoff et al., 2022).

The Earth's orbital configuration can be described by its eccentricity, obliquity, and precession. These determine the seasonal and meridional distribution of solar radiation and change over time according to so-called Milankovitch cycles. Expanding the equations describing them yields the dominant periodicities of these cycles (e.g., Berger, 1978; Laskar et al., 2004). Eccentricity describes the elliptical shape of the orbit around the Sun, resulting from the gravitational force of the other planets of the Solar System on Earth. The orbits elliptical shape causes the distance between the Earth and Sun to vary throughout the year. This affects the intensity of the seasons as Earth receives relatively more radiation when closer to the Sun (Bradley, 2015). As the orbit is only slightly elliptical and the distance between Earth and Sun differs only by about 3% of the mean throughout the year, the effect on the received radiation is small (Marshall and Plumb, 2008). For eccentricity, the dominant cycles are a quasi-periodic one of about 100 kyr, the result of four cycles of similar strength with periodicities between 95 and 131 kyr, and another with a periodicity of about 400 kyr (Marshall and Plumb, 2008; Ruddiman, 2008). As a result, eccentricity varies between close to a circle to slightly elliptical. At present, it is at the lower end of this range and further decreasing and thus becoming less elliptical (Ruddiman, 2008).

Additionally, the position of perihelion changes with a period of about 21.7 kyr as a result of climatic precession (Bradley, 2015). Climatic precession describes the combined effect of the precession of Earth's axis and the precession of Earth's orbit, which change the timing of the perihelion relative to the equinoxes (Bradley, 2015). Earth's axis precesses like a top. This change in the orientation of the rotation axis relative to fixed stars causes the time between solstices to be about 20 minutes less than a full orbit around the Sun (Britannica, 2024). After about 26 kyr this adds up to a full year, completing one cycle (North and Kim, 2017). Earth's orbit, too, precesses such that the major axis of the orbit rotates around the Sun, thus also changing the position of the perihelion. At present, Earth is closest to the Sun in early January, such that radiation is enhanced during the Northern Hemisphere winter and Southern Hemisphere summer, leading to an enhanced (subdued) seasonal contrast in the Southern (Northern) Hemisphere

[4]The sun spots are a result of changes in the magnetic field of the Sun, which actually exhibit a cycle of about 22 yr length (Babcock, 1961; Leighton, 1969). This is because the poles of the magnetic field switch about every 11 yr. For those 11 yr, the difference in rotation speed between the Sun's equator and poles deforms the magnetic field and creates more and more sunspots. With the switch of the poles, the sunspots become more regular again. Thus, there is an effective 11 yr cycle in radiation received on Earth.

(Ruddiman, 2008). Eccentricity modulates precession such that differences due to precession are larger when eccentricity is large (Bradley, 2015).

Lastly, Earth's rotation axis is tilted with respect to an axis perpendicular to the orbital plane. This so-called obliquity can range between 22.1–24.5° with a 41 kyr periodicity (North and Kim, 2017). Currently, the tilt is 23.5° and decreasing (Ruddiman, 2008). Due to the tilt, seasons exist and daytime length varies throughout the year. Obliquity has a symmetric effect in both hemispheres that grows towards higher latitudes: for a larger tilt seasonality increases as the summer (winter) hemisphere is turned more (less) directly towards the Sun. Overall, changes in orbit affect the relative seasonal distributions of insolation rather than annual mean insolation (Bradley, 2015). The influence of the orbital parameters on temperature variability depends on latitude with obliquity dominant in high, precession in mid- and eccentricity in low latitudes (Laepple and Lohmann, 2009; Westerhold et al., 2020). The background state further alters how temperature variability responds to changes in orbit as for example the presence of ice sheets can amplify the response (Westerhold et al., 2020).

Plate tectonics act on even longer timescales of millions of years and longer (Bradley, 2015). They change the position of the continents, the closing or opening of ocean straits and uplift mountain ranges. Humans have become another driving force, their effects summarized as anthropogenic forcings. These include, but are not limited to, burning of fossil fuels and the resulting emission of greenhouse gases and aerosols as well as widespread land use changes.

All other changes in the climate system happen in response to the described forcings and the non-linear dynamics of the climate system itself. However, due to the complexity of the system, not all of the natural processes can be included in climate models with their full complexity or are even understood well enough to reproduce past developments of climate. For example, the inclusion of a carbon cycle that reproduces reconstructed past changes in atmospheric carbon has proved challenging (Willeit et al., 2023). Similarly, the dynamic simulation of ice sheets and their interactions with the climate is complicated (Mikolajewicz et al., 2025; Muntjewerf et al., 2020). As such, climate models might use further inputs as forcings, like the development of atmospheric greenhouse gases or ice sheet extent and height over time as further described in Sec. 1.2.1. Together, the, often non-linear interactions between forcings and the internal dynamics of the climate system give rise to the mean climate state and its variability and drove its evolution throughout Earth's history.

1.1.3 Past to future climate variability

Throughout its 4.55 Gyr history (Brent Dalrymple, 2001; Manhes et al., 1980), Earth and its climate have undergone massive transformations. Molten and hot at the time of its formation, Earth subsequently cooled down, developed a magnetic field, and the formation of the Moon stabilized its axis of rotation (Laskar et al., 1993; Nichols et al., 2024). Earth formed a solid crust with high volcanic activity, a primordial atmosphere and first reservoirs of liquid water that eventually would become oceans (Ge et al., 2023; Johnson et al., 2022; Smithies et al., 2021; Wilde et al., 2001). The evolution of Earth's surface, atmosphere and oceans eventually allowed the first forms of life to appear between 4.3 and 3.8 Myr ago (Bell et al., 2015; Dodd et al.,

2017). Since then, life evolved into more complex forms, species appeared and disappeared, influenced by Earth's changing climate and influencing it in turn.

In many respects, conditions on Earth have been drastically different to those during humanity's existence for most of Earth's history (Fig. 1.1, 1.2). The primordial atmosphere was for example almost devoid of oxygen, something which only changed about 2.4–2.3 Gyr ago with the Great Oxidation Event, which increased atmospheric oxygen levels such that they were permanently above 0.001% of today's level (Gumsley et al., 2017; Ligrone, 2019; Poulton et al., 2021). In a second oxidation event, atmospheric oxygen levels rose to about 3% 850–540 Myr ago, although the evolution of atmospheric oxygen levels in between the events remains debated (Ligrone, 2019; Liu et al., 2021; Lyons et al., 2014). Similarly, the current continental configuration has only been present since about 14 Myr ago (Hartmann, 2016). While some changes, like the slow movement of the tectonic plates that rearranges and creates continents, happen smoothly, Earth's history has also been marked by times of drastic upheaval and abrupt climatic changes.

For most of Earth's history, its climate was warm enough for Earth to be free of major ice sheets. Without ice sheets and in a warmer climate, sea level was as much as 170 m higher than at present (Müller et al., 2008). Atmospheric concentrations of greenhouse gases were elevated, but do not alone explain reconstructed GMST (Bowen et al., 2015; Marshall and Plumb, 2008; Poulsen et al., 2015; Wade et al., 2019; Zeebe et al., 2016). Other forcings, feedbacks and differences in the Earth System contributed to these warm climates, such as continental drift, silicate rock weathering or a lower atmospheric mass modifying Earth's energy balance (Chen et al., 2021; Poulsen et al., 2015; Wade et al., 2019). Among these warmer climates, two states can be differentiated: warmhouse climates with a GMST about 5°C warmer than today[5] and hothouse climates, during which GMST was about 10°C warmer (Westerhold et al., 2020). Westerhold et al. (2020) found that surface temperature variability was larger during hothouse than warmhouse climate, largely due to the occurrence of hyperthermals, transient warming events during which carbon is rapidly released into the atmosphere.

While warm climates dominate Earth's history, glaciations have repeatedly occurred (Hartmann, 2016; Hoffman et al., 2017; Hofmann and Bindeman, 2023; Young et al., 1998). In particular, the presence of ice sheets on Earth define the more recent climate history as Earth gradually cooled since about 50 Myr ago and transitioned into a coolhouse climate, characterized by the existence of some ice cover mostly around Antarctica (Westerhold et al., 2020). Until about 34 Myr ago ice was still rare, at which point ice sheets started rapidly forming and expanding in East Antarctica (Lear and Lunt, 2016; Past Interglacials Working Group of PAGES, 2016). There was significant variability superimposed on the general cooling trend, e.g., due to the state-dependent response to changes in orbital forcing (Westerhold et al., 2020). As the cooling persisted, ice sheets eventually appeared also in the Northern Hemisphere about 2.58 Myr ago, marking the beginning of the ongoing Quaternary period (Fig. 1.1, Westerhold et al., 2020). The Quaternary is an icehouse climate, characterized by the presence of ice sheets in both hemispheres [6].

[5] Here, Westerhold et al. (2020) refer to the mean temperature of the current interglacial, the Holocene, which they calculate to be 14.15°C, as today's temperature.

[6] Depending on the authors, 'greenhouse climates' is sometimes used to describe any state other than icehouse climates (e.g., Marshall and Plumb, 2008) or as a summary term for warmhouse and coolhouse climates (Barral

In an icehouse climate, Earth cycles through glacial and interglacial conditions, which correspond to waxing and waning ice sheets, respectively. During glacials, sea level has been as much as 120 m lower than today, with ice sheets extending into mid-latitudes (Miller et al., 2020), whereas sea level during interglacials is similar to today's (IPCC, 2021b; Past Interglacials Working Group of PAGES, 2016). At the beginning of the Quaternary, the glacial-interglacial cycles had a period of about 41 kyr, the same as obliquity (Westerhold et al., 2020). At the mid-Pleistocene transition between 1.25 and 0.7 Myr ago, a time when ice sheets expanded from one cycle to the next, this quasi-periodicity transitioned to cycles of roughly 100 kyr length accompanied by a higher amplitude in GMST change (Clark et al., 2006; Shackleton and Opdyke, 1973; Westerhold et al., 2020). These match the periodicity of eccentricity. However, while orbital changes in insolation pace the cycles (Hays et al., 1976), no change in insolation caused the transition itself and eccentricity especially is, without amplifying feedbacks, not strong enough to force the 100 kyr temperature variability (Clark et al., 2006). Besides being stronger in amplitude as a result of more severe glaciations, the 100 kyr cycles are also asymmetric with surface temperature exhibiting a sawtooth pattern of about 10 kyr of rapid warming followed by gradual cooling (Brook and Buizert, 2018; Crucifix, 2012; Shackleton and Opdyke, 1973). Glacial conditions were then the most pronounced just before the glacial termination. Temperature were the coldest temperatures, ice sheet extent the largest ice sheet and CO_2 levels at their lowest (Marshall and Plumb, 2008).

The glacial-interglacial cycles exhibit significant millennial- and sub-millennial-scale variability (Crucifix, 2012; Dansgaard et al., 1993; McManus et al., 1999). Prominent among them are Dansgaard-Oeschger events, which show abrupt warming of 8–15°C within decades in the Northern Hemisphere, most prominently Greenland, followed by multiple centuries of slow cooling (Dansgaard et al., 1993; Grootes et al., 1993; Huber et al., 2006). While many causal mechanisms have been proposed, a definite explanation remains elusive in the absence of a clear driving force (Baldini et al., 2015).

From Last Glacial Maximum to the Holocene: the Last Deglaciation

Between 23 and 19 kyr BP, glacial condition of the Last Glacial Period last reached a maximum (IPCC, 2021b). This time of quasi-equilibrium conditions referred to as the **Last Glacial Maximum** (LGM), ice sheet cover was at its largest extent (Clark et al., 2009; Gowan et al., 2021) and, conversely, global mean sea level at its lowest (IPCC, 2021b; Lambeck et al., 2014; Yokoyama et al., 2018). At 21 kyr BP global mean sea level was about 130 m lower than at present (Lambeck et al., 2014). Locally, the timing can differ such that the largest ice sheet extents and heights in different places occurred between 29 and 16 kyr BP (IPCC, 2021b). According to reconstructions of past surface climate, GMST during the LGM was about 4.5–7.0°C lower than at present (Annan et al., 2022; Osman et al., 2021; Shakun and Carlson, 2010; Tierney et al., 2020). Models further suggest drier climate conditions during the LGM in comparison to today across much of the globe (**P3**). Such cold and dry glacial conditions were favored by a decrease in Northern Hemisphere summer insolation due to changes in the orbital configuration.

et al., 2017; Kidder and Worsley, 2010).

Changes in northern summer insolation, amplified by feedback processes, then triggered the termination of the LGM and triggered a deglaciation (Clark et al., 2009; Hays et al., 1976).

The **Last Deglaciation** is the most recent episode of significant global warming and marks the transition from the Last Glacial to the current interglacial (Annan and Hargreaves, 2013; Gulev et al., 2021; Osman et al., 2021; Shakun et al., 2012; Tierney et al., 2020). It lasted roughly from 19 to 10.65 kyr BP. Over the course of the Deglaciation the Laurentide ice sheet over North America and the Fennoscandian ice sheet over northwestern Eurasia retreated and global sea levels rose. Atmospheric greenhouse gas concentration increased, with CO_2 rising from about 193 to 271 ppm and methane doubling (Gulev et al., 2021; Köhler et al., 2017). The Last Deglaciation started in the Northern Hemisphere, where ice sheets began melting around 19 kyr BP, causing an initial sharp rise in sea levels, followed by near constantly rising sea levels until the onset of the Deglaciation in the Southern Hemisphere (Clark et al., 2009; Lambeck et al., 2014). Melting in the Southern Hemisphere, mostly of parts of the West Antarctic ice sheet, began around 15–14 kyr BP, leading to a period of fast sea level rise by as much as 40 $\frac{\text{mm}}{\text{yr}}$ during what is called meltwater pulse 1A (MWP-1A) (Lambeck et al., 2014). Rates of sea level and warming differed throughout the Deglaciation and were disrupted by abrupt millennial scale events, especially in the Northern Hemisphere (Lambeck et al., 2014). Such events were the Bølling-Allerød (14.7 – 12.9 kyr BP) with abrupt warming in the Northern and cooling in the Southern Hemisphere and the following Younger Dryas (12.9–11.4 kyr BP) a period of marked cooling in the Northern Hemisphere to almost glacial conditions (Marshall and Plumb, 2008).

There is large uncertainty regarding the evolution of the ice sheets and consequently in meltwater and sea level (Abe-Ouchi et al., 2015; Lambeck et al., 2014; Stokes et al., 2015). For example, estimates of the contribution of the melting of the Antarctic ice sheet to global sea level rise vary between 10 and 3 m (Lambeck et al., 2014; Nakada et al., 2000; Philippon et al., 2006; Pollard and DeConto, 2009; Whitehouse et al., 2012). This uncertainty carries over to climate model simulations of the Deglaciation as timing and volume of meltwater releases crucially affect the evolution of simulated climate (Snoll et al., 2024).

Around 11.7 kyr BP, proxies worldwide indicate a shift from dry, cold glacial conditions to a warm and wet climate (Walker et al., 2009). As such, this is defined as the beginning of the **Holocene**, the current quasi-equilibrium interglacial (Walker et al., 2009). However, significant ice sheet retreat and sea level rise occurred until about 8.2 kyr BP and continued at smaller rate until about 2.5 kyr BP (Lambeck et al., 2014). After that, sea level remained approximately constant until 150–100 yr ago (Lambeck et al., 2014). While GMST during the Holocene was comparatively stable in comparison to the Deglaciation, this perceived stability masks considerable regional climate variations like the regional cooling of the Little Ice Age in the North Atlantic region during the 16th to 19th century or the Saharan Greening, the most recent African Humid period during the mid-Holocene.

Holocene multi-centennial surface temperature variance is generally smaller than LGM variance, globally by about a factor of 4, although with large zonal variations (Rehfeld et al., 2018). In particular, Greenlandic ice cores suggest a decrease in variance by a factor of 73, in contrast to the tropics, where proxies find a particularly small difference of about 1.6–2.8 (Rehfeld et al., 2018). Overall, the state-dependency in variability is larger in the Northern than the Southern

Hemisphere (Rehfeld et al., 2018). Shao and Ditlevsen (2016) similarly find the Holocene to be more stable than the LGM. LGM temperature variability is also larger than Holocene variability in equilibrium simulations on interannual scales, but to a lesser extent than suggested by reconstructions (Rehfeld et al., 2018; Shi et al., 2022). This difference in variability between the colder LGM and warmer Holocene is not related to external forcings (Prange et al., 2023), but likely due to a difference in the meridional temperature gradient. This gradient was larger during the LGM and decreases with warming due to polar amplification and is found to correlate with temperature variability (Rehfeld et al., 2018; Shi et al., 2022). Changes in variability with warming further show a land-ocean contrast as variability generally increases over land, but decreases over the oceans as mean temperatures rise (Rehfeld et al., 2020).

Present-day climate and future projections

Present-day climate shows a clear, ongoing warming trend, which is projected to continue and driven by anthropogenic greenhouse gas emissions, in particular of CO_2 (IPCC, 2021c). The increase in atmospheric greenhouse gas concentrations and the subsequent increase in GMST and warming of the troposphere has widespread consequences across all compartments of the climate system: Sea level is rising, global precipitation increasing, atmospheric circulation patterns change and the mid-latitude storm tracks for example shift poleward, sea ice and glaciers are retreating and vegetation zones are moving (Divine and Dick, 2006; IPCC, 2021c). Additionally, extreme conditions are changing (Seneviratne et al., 2021). Notably, hot extremes continue to become more frequent and intense over most land areas, while cold extremes become less frequent and severe (Seneviratne et al., 2021). Across most land areas, the frequency and intensity of high precipitation events is increasing, while agricultural and ecological droughts are becoming more common in some regions as evapotranspiration increases with warming (Seneviratne et al., 2021).

Many of these trends are projected to continue in the **future**. As the increase in atmospheric greenhouse concentrations is faster than in the reconstructed records of past CO_2 variations (IPCC, 2021c), the resulting warming and many concurrent changes to the climate system occur faster than in the past climate record. As such, the response of slower components of the climate system like the deep ocean is lagging behind (IPCC, 2021c). While the response of the climate system depends on anthropogenic greenhouse gas emissions as well as other anthropogenic forcings, some changes will continue even if emissions halt and would lead to a halt in GMST rise (IPCC, 2021d). Especially changes of the oceans and sea ice as well as sea level rise will continue for centuries to millennia (IPCC, 2021d).

The future evolution of the climate system depends on the eventual emissions pathway. Climate model simulations therefore assume these future emissions pathways based on possible socio-economic developments. These scenarios consider not only GHG emissions, but all anthropogenic drivers of climate change like land use changes or the emission of aerosols and chemically reactive gases. To this end, the Shared Socioeconomic Pathways (SSPs) were implemented in CMIP6 and range from low emission pathways compatible with sustainable development to the continued emissions of GHG as countries continue to rely on fossil fuels to drive their economies.

Current and projected future warming affects variability, the effect of which varies by region and depends on timescale. Globally, temperature variance is projected to decrease under future warming from intra-seasonal (Rehfeld et al., 2020) to annual (Rehfeld et al., 2020; Shi et al., 2024) and decadal timescales (Brown et al., 2017; Rehfeld et al., 2020). However, there is a land-sea contrast, as variance has been found to increase over land and decrease over the oceans (Rehfeld et al., 2020). In particular, variance increases with warming over tropical and subtropical land on monthly (Bathiany et al., 2018; Holmes et al., 2016; Shi et al., 2024), annual (Holmes et al., 2016; Olonscheck et al., 2021; Rehfeld et al., 2020; Shi et al., 2024) and decadal timescales (Brown et al., 2017; Rehfeld et al., 2020). This has been linked to changes in vegetation cover and drying soil (Bathiany et al., 2018; Shi et al., 2024), as well as shifts in circulation pattern (Bathiany et al., 2018) and increased cloud cover (Shi et al., 2024). In the mid- and high latitudes, on the other hand, temperature variability is projected to decrease under future warming on monthly (Bathiany et al., 2018), annual (Boer, 2009; Olonscheck et al., 2021; Rehfeld et al., 2020) and up to decadal timescales (Rehfeld et al., 2020). Sea ice loss (Bathiany et al., 2018; Olonscheck et al., 2021) and the reduced meridional temperature gradient (Bathiany et al., 2018; Holmes et al., 2016) have been suggested as the mechanisms responsible for this decreasing temperature variability.

For precipitation variability, there is mostly an increase with warming from daily (Pendergrass et al., 2017) and monthly (Pendergrass et al., 2017; Wetherald, 2009) to annual (Boer, 2009; Pendergrass et al., 2017; Rehfeld et al., 2020; Wetherald, 2009) and decadal timescales (Boer, 2009; Pendergrass et al., 2017; Rehfeld et al., 2020). This increase is stronger over land areas (Pendergrass et al., 2017) and happens especially in areas, where mean precipitation increases as well (Räisänen, 2002; Wetherald, 2009).

For many climatic variables, such as temperature, precipitation, specific humidity or wind fields, distributions are not normal, as they can be skewed or have heavy or light tails (Perron and Sura, 2013; Ruff and Neelin, 2012; Skelton et al., 2020; Tamarin-Brodsky et al., 2022; Volodin and Yurova, 2013). The shapes of the tails of distributions especially affect the likelihood of extremes (McKinnon et al., 2016; Ruff and Neelin, 2012; Simolo and Corti, 2022). Under current and future warming, changes in the tails of the distributions have been reported in observations and simulations, e.g., an increasingly strong warm tail under warming in Northern Hemisphere winter temperatures (Tamarin-Brodsky et al., 2020), but evidence remains mixed (McKinnon et al., 2016; Skelton et al., 2020; Tamarin-Brodsky et al., 2020).

1.2 Evidence on climate variability

1.2.1 Climate Modeling

Climate models describe the state of the climate system and its evolution in time with equations following laws of physics and geobiochemistry (Stocker, 2011). However, the large range of scales, processes and interactions constituting the climate system cannot possibly be resolved or known in their entirety. As such, models need to make simplifications and approximate the state of the system numerically at discrete grid points and points in time (Goosse et al., 2010a). Any

phenomena below the scale of the grid are included using parameterizations. These parameterizations have to be tuned, often such that they produce a realistic present-day 30 yr-climatology, usually for 1960–1989 CE, with respect to important climate variables (Hourdin et al., 2017). Different choices about the processes and components included, as well as about sub-grid-scale parameterizations result in a large variety of models. These can range from 0-dimensional point models approximating Earth's global energy balance to **Earth System Models** (ESMs) that can resolve 3-dimensional space at high resolutions.

Simple models do not attempt to model the whole climate system and instead focus on simplified representations of key processes to compute a limited set of variables. Among these are energy balance and radiative transport models. **Energy balance models** (EBMs) compute temperature based on the fluxes of incoming and outgoing radiation. Depending on the spatial dimensions modeled, they further simulate heat transport to derive the temperature field in space.

Different compartments of the climate system are often modeled separately and the resulting models then coupled together or run on their own using boundary conditions for input from the other compartments. Atmosphere and ocean models solve equations of motion in a fluid based on the conservation of momentum, mass and energy as well as thermodynamic equations for the evolution of climatic variables. Models use simplified versions of the full equations of motions called primitive equations. Further equations describe processes such as changes in salinity in the ocean, while turbulence and clouds are parameterized. To improve the representation of the climate system, atmosphere models can further include aerosols or atmospheric chemistry, while ocean models might simulate biogeochemical processes. Land surface models require- at minimum equations describing the surface moisture, heat balance and snow cover. The spatial resolution limits the topographic features that can be resolved, resulting in a smoothed land surface, which is particularly noticeable in mountainous regions.

Coupled together such that they exchange heat, momentum and moisture at certain intervals, these models form **Atmosphere-Ocean General Circulation Models** (AOGCMs or often just GCMs). Usually, an AOGCM also contains an interactive sea ice component, describing the thermodynamic growth and melting of sea ice cover. State-of-the-art sea ice components further contain a dynamical component modeling sea ice transport and deformation. Dynamic vegetation models group vegetation into so-called plant functional types and model their response to and influence on climate. General circulation models with dynamical vegetation can be referred to as **Atmosphere-Ocean-Vegetation GCMs** (AOVGCMs). An ESM can further couple a dynamic carbon cycle. While many climate models include ice sheets as a boundary condition, ice sheet models have been coupled to ESMs and other climate models (e.g., Mikolajewicz et al., 2025). Many other processes have been investigated and included in climate models besides this brief overview, such as a dynamic methane cycle (e.g., Kleinen et al., 2020).

While modern climate models can share a common ancestor such that different atmospheric models share for example the code of their dynamical cores, the implementation of components and their coupling varies from model to model. As such multi-model ensembles can capture some of the uncertainty around current knowledge of the climate system. Depending on the representation of the compartments, some processes internal to the climate system will be modeled as boundary conditions. For example, many paleoclimate simulations provide reconstructed

changes in greenhouse gases, as well as changes in ice sheets, sea level and the land-sea mask as external forcings to reduce the computational resources required. Another option, for more resource efficient simulations are **models of intermediate complexity (EMICs)**. These simplify or exclude some of the compartments and can thus allow for longer integrations or large ensembles sampling the range of internal variability.

Here, we consider simulations from various climate models, which we categorize based on the complexity of the individual model components due to their relevance for simulated variability. For atmosphere and ocean models, we evaluate the complexity of the model equations, that is whether the model uses primitive equations or simplified representations such as a 0D point model or slab ocean. The complexity of the radiation schemes can vary as well, as models use a number of bands that cover a range of wavelengths instead of modeling individual absorption lines in the spectrum of radiation (Hartmann, 2016). As such, the number of bands both for the shortwave and longwave radiation differs, which we use as an indication of the complexity of the radiation scheme.

While a thermodynamic component of sea ice evolution is common to all but the most simple models, dynamic sea ice can be missing, for example in EMICs. And coupled ice-sheet-climate models remain rare as they significantly increase the computational costs of running a model (Mikolajewicz et al., 2025). For simulating effects of vegetation prescribed surface types remain in use, although many models now include a terrestrial carbon cycle, dynamic vegetation or even a nitrogen cycle. Modeling of individual plant types remains the prerogative of specialized vegetation models, though. Land surface models differ in particular to how land hydrology is handled, ranging from single layer bucket models to complex river routing schemes. Lastly, significant strides have been made regarding the inclusion of natural aerosols and processes of atmospheric chemistry.

Additionally, the spatial resolution contributes to model complexity and in particular to the resources required for running a model. This includes the latitudinal and longitudinal extent of grid boxes, as well as the amount of layers in the atmosphere and ocean. Spatial resolution can differ a lot, including between different components of a coupled model or regionally within one compartment. For paleoclimate simulations, spatial resolution tends to be coarse, commonly varying between T21 and T42[7] (cf. Table 1.1) at most. In contrast, high resolution global models can simulate changes even at a scale of kilometers, allowing small scale processes like convection to be resolved instead of parameterized (Ding et al., 2022; Guichard and Couvreux, 2017; Kendon et al., 2021).

Many studies suggest that various climatic features improve with an increasing horizontal and vertical resolution. However, degree and regions of improvements vary widely between models. At times, findings can also be contradictory, as Klavans et al. (2017) and Hack et al. (2006) find a marginal change or very small decrease in GMST with increasing horizontal resolution, while Otto-Bliesner et al. (2006) and Kirtman et al. (2012) demonstrate an increase in GMST. While a broad decrease in sea ice accompanies such an increased GMST (Kirtman et al., 2012),

[7] In the naming conventions of model resolutions, T stands for triangular and refers to the grid truncation for a spectral model.

Table 1.1: Common model resolutions types and their grid definitions (National Center for Atmospheric Research Staff (Eds), 2017). The side length of the grid boxes are given at the equator.

Resolution	latitude points	longitude points	grid box side length [km]	grid box side length [°]
T21	32	64	625	5.61
T31	48	96		
T42	64	128	310	2.79
T63	96	192	210	1.88
T85	128	256	155	1.39
T106	160	320	125	1.12
T255	256	512	60	0.54

an increasing horizontal resolution has also been found to reduce Arctic sea ice biases in either direction as Arctic circulation patterns improve in comparison to observations (Gent et al., 2010; Hack et al., 2006). Slight to substantial improvements in atmospheric circulation patterns with increasing resolution have been repeatedly reported (e.g., Collier and Zhang, 2007; Gent et al., 2010; Hack et al., 2006; Marti et al., 2010; Roeckner et al., 2006). Synoptic-scale phenomena like tropical cyclones improve especially at very high resolutions like T213 (Kobayashi and Sugi, 2004), but only if vertical resolution is high enough as well (Roeckner et al., 2006). Similarly, eddy-resolving models show substantial reductions of biases in ocean models (Delworth et al., 2012). Linked to improved circulation patterns are improvements in the modes of variability to various degrees, for example for ENSO (Delworth et al., 2012; Deser et al., 2006; Guilyardi et al., 2004; Marti et al., 2010) and NAO (Marti et al., 2010).

For daily to decadal timescales, Klavans et al. (2017) find increased temperature variability in tropical regions, but decreased variability in mid- and high latitudes with increasing horizontal resolution (T31, T42 and T85) in CCSM3. This results in a decrease in global temperature variability, except on daily timescales. Decadal sea ice variability has been reported to better match observations at higher resolutions (Gent et al., 2010). Improvements are by no means universal (Collier and Zhang, 2007; Hack et al., 2006) and often development of the higher resolution model versions were accompanied by improvements in various parts and parameterizations of the model. As such, the degree to which improvements can be linked to changes in resolution alone and the role of resolution in removing model biases remain uncertain if resolutions are not high enough to allow previously parameterized sub-grid-scale processes to be resolved. For long-term transient simulations of paleoclimate, especially, models have to be run at relatively low resolution — usually no more than T42 —, such that most of these improvements are unlikely to be observed. Overall, we classify models by the complexity of the individual compartments to analyze how complexity affects simulated variability.

When accounting for mismatches between simulated and actualized anthropogenic forcings, climate models have in hindsight been very successful in projecting changes in GMST (Hausfather et al., 2020). Simulated global variability, too, agrees well with observations and reconstructions from annual to millennial timescales (Askjær et al., 2022; Ellerhoff and Rehfeld, 2021;

Laepple and Huybers, 2014a; Neukom et al., 2019; Zhu et al., 2019). Regionally, however, simulated variability is lower than that in reconstructions on decadal and longer timescales (e.g., **P2**, Bühler et al. (2021); Ellerhoff and Rehfeld (2021); Franke et al. (2013); Fredriksen and Rypdal (2016); Hébert et al. (2022); Ljungqvist et al. (2019)). Models further simulate smaller changes in variance between LGM and Holocene than found in proxy reconstructions (Rehfeld et al., 2018). This potential lack of variability in models is compounded by the fact that known sources of external variability like volcanism are omitted in many simulations of past and future climate. As a result, it remains unclear to what degree models lack variability and for which reasons.

To evaluate simulated variability, we use ensembles of the Deglaciation (**P3**) as well as of the historical period and potential future pathways (Sec. 3.3. The **Deglacial ensemble** is described in detail in **P3** and consists of fifteen transient simulations from eight models. All of the simulations cover at least the Deglaciation and most cover the LGM and run to present-day. The ensemble includes a simulation from an EBM to investigate the linear response of the climate system to the employed forcings as well as sensitivity experiments that vary by meltwater protocol, ice sheet reconstruction and the inclusion of volcanic forcing.

We complement the investigation of variability since the LGM with an analysis of variability in three future scenarios. These scenarios are from the framework of shared socioeconomic pathways (SSPs) (O'Neill et al., 2017, 2016) employed in CMIP6 (Eyring et al., 2016) and the Intergovernmental Panel on Climate Change's sixth assessment report (IPCC, 2021c). The framework differentiates five categories of SSPs that vary by mitigation and adaptation requirements such that they encompass a variety of pathways such as sustainable development (SSP1) or fossil-fueled development (SSP5). Within the boundaries established by these categories several scenarios of socioeconomic development were created. Based on the scenarios, integrated assessment models simulated possible future emission and land use changes that would hit certain targets. Such scenario targets are for example certain longterm forcing levels or the policy targets of at most 2°C or preferably 1.5°C warming by the end of the century set by international community in the Paris Agreement (O'Neill et al., 2016).

To cover the range of possible future scenarios, we chose three scenarios: SSP5-8.5 exemplifies the upper end of emission scenarios with a radiative forcing of 8.5 W m^{-2} by the end of the century. SSP 2-4.5 is an intermediate forcing scenario that extrapolates historical developments. Lastly, SSP 1-1.9 represents the lower end of emissions scenarios and would likely be compatible with the 1.5°C warming limit of the Paris Agreement. As we are interested in changes of variability, in particular in the long term, we consider only those CMIP6 simulations that run until 2300 CE. This limits the amount of available simulations, especially for SSP 1.1-9 and 2-4.5, for which only simulations from three configurations of the NASA-GISS model (E2.1-G, E2.2-G and E2.1-H) are available. These versions differ by the coupled ocean model – GISS Ocean (G) or Hycom (H) – and by atmospheric resolution – higher in the 2.2 version –, as well as some related differences in tuning and parameterizations (Kelley et al., 2020; Orbe et al., 2020). For SSP 5-8.5, runs from six more models are available: CanESM5 (Swart et al., 2019), CESM2-WACCM (Danabasoglu et al., 2020), IPSL-CM6A-LR (Boucher et al., 2020), MIROC-ES2L (Hajima et al., 2020), UK ESM 1.0 LL (Sellar et al., 2019, 2020) and MRI-ESM2.0 (Yukimoto et al., 2019). In addition to the CMIP6 simulations, we analyze variability in the continuations

according to the three SSP scenarios of the Deglacial MPI-ESM ch4 simulation used in **P3** (Kleinen et al., 2021, 2023, 2020).

1.2.2 Observational products

For all but the most recent centuries, knowledge about past climate is reconstructed by proxy as even the longest record of direct instrumental measurements of surface temperature, the Hadley Centre Central England Temperature (HadCET) record, only starts in 1659 CE (Parker et al., 1992). Additionally, global coverage by satellite observations has only started to be available at times since 1979 (Hartmann, 2016). As such, direct instrumental measurements of Earth's climate span only a small fraction of Earth's history (Bradley, 2015). Instead, tracing Earth's climatic history relies on reconstructions of past climate based on various climate archives, like ice cores, tree rings, lake sediments or speleothems. Climate archives were exposed to past climatic conditions, which affected their properties. They reflect climate conditions in the form of measurable proxy properties and can thus be used to reconstruct climatic variables of interest (Bradley, 2015; Evans et al., 2013). Methods for measuring and dating proxies, as well as converting a proxy property, such as oxygen isotopic composition into a variable like surface temperature, introduce significant uncertainty in the reconstruction (**P2**, Anchukaitis and Smerdon, 2022; Christiansen and Ljungqvist, 2017; Esper et al., 2004). Furthermore, records become more and more sparse in space and time further back in Earth's history and archives have limitations with respect to the temporal resolution they can yield (Bradley, 2015; Westerhold et al., 2020).

For comparison to observation-based products, we use both datasets from the instrumental period as well as reconstructions of climate since the LGM. In particular, we use retrospective analysis (reanalysis) products, which assimilate observations of atmosphere and ocean using state-of-the art forecasting models. Reanalyses can be differentiated according to the time period they cover and thus, the observational data available for assimilation. Slivinski et al. (2019) distinguish modern from historical reanalysis. Modern reanalysis usually begin at earliest in 1950 CE and make use of upper air and satellite observations available since then. Historical reanalyses, on the other hand, tend to be at least a century long, using near-surface observations that cover the whole period. Here, we further use a paleo reanalysis, which assimilated proxy reconstructions. In all cases, issues can arise due to gaps in coverage and biases in different observational records.

Our work relies on the following datasets: the LGM reanalysis LGMR by Osman et al. (2021), which covers the past $24,000$ years, the ERA5 (Hersbach et al., 2023, 2020) and NOAA 20th century reanalysis v3 (NOAA 20 CR v3), as well as the observational dataset HadCRUT5 (Morice et al., 2021) for comparison to present-day and historical variability. The ERA5 reanalysis by the ECMWF (Hersbach et al., 2023, 2020) describes the global evolution of the atmosphere, land surface and ocean waves from 1950–2023 CE at high resolution in space and time. The reanalysis uses a numerical weather prediction model for medium-range forecasts, IFS (Integrated Forecast System) model cycle CY41r2, to assimilate various observational datasets, e.g., ground-based and satellite measurements. This procedure is executed for one control and

nine perturbed ensemble members. As a historical reanalysis, NOAA 20 CR v3 (Compo et al., 2011; Slivinski et al., 2019) covers the period 1806–2015 CE. The data assimilation uses only surface-based observations to constrain simulations with the US National Centers for Environmental Prediction Global Forecast System (NCEP GFS). Here, we use surface temperature and precipitation of ERA5 and NOAA 20 CR v3 to put the model simulations in the context of modern climate. Lastly, HadCRUT5 is a dataset that consists of sea surface temperatures measured from ships and by buoys, as well as near and surface air temperature measurements from weather stations and covers 1850–2022 CE (Morice et al., 2021). However, there are increasing gaps in coverage further back in time, especially over and around Antarctica.

1.3 Statistical understanding of climate variability

A description of the climate system in its entirety, that is of its state at a given point in time with all its processes determining its evolution in time, is impossible. Instead climate science makes use of statistical methods to describe the state and its evolution in the phase space. Measurements or model experiments sample this phase space of the (simulated) climate system. Any variable describing the state of the climate system and its evolution, such as surface temperature or precipitation, can be seen as the result of a random process and itself a random variable (Chatfield, 2016; Hasselmann, 1976; von Storch and Zwiers, 1999). The sampling of a random variable yields a **probability distribution**, describing the probability for the variable to take on a certain value during the sample period. Conversely, climate data is ordered in time and thus forms a **timeseries**.

1.3.1 Statistical description of the shapes of probability distributions

The shape of a climate variable's probability distribution provides information about the variable in question and the processes influencing it. Fundamental properties to consider in the analysis of a probability distribution's shape are

1. the **location**, that is the central tendency of the data, measured for example by the **mean** μ, the median or the mode,

2. the **spread** of the distribution, commonly estimated using the variance and its square root **standard deviation** σ,

3. the **(a)symmetry** of the distribution described by its **skewness** s,

4. the **weight of the tails**, that is its so-called **kurtosis** k'.

Mean, variance, skewness and kurtosis are the first four **moments** of a distribution (Fig. 1.4) and described in greater detail in **P3**. In general, the moments of a distribution describe how

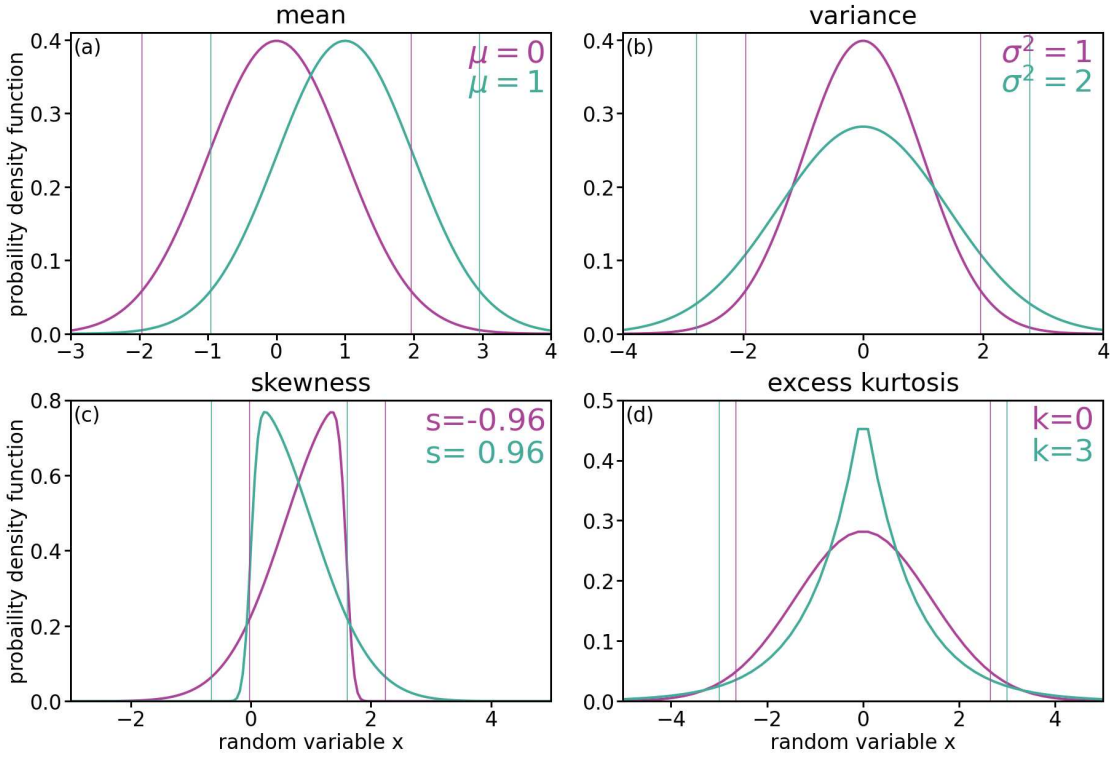


Figure 1.4: Exemplary distributions and accompanying changes in moments: (a) mean, (b) standard deviation, (c) skewness and (d) kurtosis. Differences in moments for higher (green) and lower (pink) values relative to each other. Vertical lines indicate changes in 95% confidence intervals for a two-tailed test. Figure adapted from **P3**.

probability mass is distributed. The generalized nth moment μ_n of a distribution with sample size N about point A is defined via the expected value of random variable X as:

$$\mu'_n = \mathbb{E}[(X - A)^n] = \frac{1}{N} \sum_{i=1}^{N} (X_i - A)^n, \tag{1.1}$$

where X_i are independent realizations of X.

Variations on that definition exist, by giving a specific point for A, e.g., the origin 0 or the mean. Table 1.2 summarizes the various definitions and highlights the ones used within this work.

Mean and standard deviation provide a full description for a **normal** distribution, a special case of an unimodal, symmetric distribution. Many statistical methods require the data to be normal. However, various climate variables are locally and regionally non-normal (Franzke et al., 2020; Perron and Sura, 2013). If normality is violated, higher order moments are necessary to provide additional information about the shape, in particular about the tails of the distribution. Skewness is then positive (negative), if the weight of the distribution is shifted towards higher (lower) values in comparison to a normal distribution, with a long low (high) valued tail (Fig. 1.4c). As standardized kurtosis is $k' = 3$ for a normal distribution, it is often replaced by the **excess kurtosis** $k = k' - 3$ (cf. Filliben and Heckert, 2024), which we will refer to as kurtosis for simplicity. For a positive (negative) excess kurtosis, the tails of the distribution are heavier (thinner) than those of a normal distribution (Fig. 1.4d). As many studies rely only on mean and standard deviation, we will sometimes use higher order moments as a summary term for skewness and kurtosis in this work.

1.3.2 Frequency components of a timeseries: spectral analysis

A timeseries is the realization of a stochastic process, which can be described as a sequence of random variables $X(t)$ at points in time (Chatfield, 2016; von Storch and Zwiers, 1999). A timeseries can consist of fluctuations that reoccur at various frequencies. These can be analyzed using autocovariance function γ in the time domain or its Fourier transform, the power spectral density S (PSD) in the frequency domain (cf. **P3**). The autocovariance provides a measure of the dependency between different points in time t_1 and t_2 of a timeseries as

$$\gamma(t_1, t_2) = \mathbb{E}[(X(t_1) - \mu(t_1))(X(t_2) - \mu(t_2))]. \tag{1.2}$$

Many statistical methods require the underlying stochastic process to be stationary or at least weakly stationary. A stochastic process $X(t)$ is stationary if all its stochastic properties are independent of time. It is weakly stationary if its mean is independent of time and its autocovariance function depends only on the lag $\tau = t_2 - t_1$ between two points in time t_1 and t_2, such that

$$\gamma(\tau) = \mathbb{E}[(X(t_1) - \mu)(X(t_1 + \tau) - \mu)]. \tag{1.3}$$

For $t_1 = t_2$ and thus lag $\tau = 0$, this yields the variance, implying that the variance of a weakly stationary process is time-independent (Chatfield, 2016).

Table 1.2: Possible definitions of the moments of a distribution. Definitions used in this work are bold.

n	Moment	generalized moment μ'_n about point A	raw moment μ_n $(A = 0)$	central moment m_n $(A = \mu)$	unbiased estimator of m_n	standardized moment
n	nth	$\mu'_n = \mathbb{E}[(X - A)^n]$	$\mu_n = \mathbb{E}[X^n]$	$m_n = \mathbb{E}[(X - \mu)^n]$	$\frac{N}{N-1} m_n$	$\frac{m_n}{(\sigma^2)^n} = \frac{m_n}{m_2^{n/2}}$
1	Mean	$\mathbb{E}[(X - A)]$	$\boldsymbol{\mu_1 \equiv \mu = \mathbb{E}[X]}$	0	0	0
2	Variance	$\mathbb{E}[(X - A)^2]$	$\mathbb{E}[(X)^2]$	$\sigma'^2 \equiv m_2 = \mathbb{E}[(X - \mu)^2]$	$\boldsymbol{\sigma^2 = \frac{N}{N-1} m_2}$	1
3	Skewness	$\mathbb{E}[(X - A)^3]$	$\mathbb{E}[(X)^3]$	$\mathbb{E}[(X - \mu)^3]$	$\frac{N}{N-1} m_3$	$\boldsymbol{s \equiv \frac{m_3}{m_2^{3/2}}}$
4	Kurtosis	$\mathbb{E}[(X - A)^4]$	$\mathbb{E}[(X)^4]$	$\mathbb{E}[(X - \mu)^4]$	$\frac{N}{N-1} m_4$	$k' \equiv \frac{m_4}{m_2^{4/2}}$

Fourier transforming the autocovariance function decomposes the timeseries into its frequency components:

$$S(\omega) = \mathcal{F}(\gamma(\tau))(\omega) \tag{1.4}$$

$$= \int_{-\infty}^{\infty} \gamma(\tau) \exp^{-i\omega\tau} \, d\tau. \tag{1.5}$$

In this work, we assume that the sampled stochastic process is weakly stationary and ergodic. Processes are weakly stationary if their mean and autocovariance are time-independent and if they have a finite second moment for all times (von Storch and Zwiers, 1999). They are said to be ergodic if the sampling in time of one realization of the process is equivalent to sampling random independent realizations (von Storch and Zwiers, 1999).

The PSD describes how fluctuations at the various frequencies ω contribute to the overall variance. Peaks represent (quasi-)oscillatory components in the timeseries and are broadened due to dampening (von Storch and Zwiers, 1999). For many climate variables, the background of the spectrum, on the other hand, is continuous and scales according to a power law (e.g., Ellerhoff and Rehfeld, 2021; Franzke et al., 2020; Huybers and Curry, 2006; Nilsen et al., 2016) such that

$$S(\tau) \propto \tau^{-\beta} \tag{1.6}$$

with scaling coefficient β (Ellerhoff and Rehfeld, 2021; Franzke et al., 2020). The larger the scaling coefficient in a system, the larger is the persistence or memory of it as statistical properties are maintained across scales (Ellerhoff and Rehfeld, 2021; Franzke et al., 2020). It has been suggested that for temperature as well as many other atmospheric variables of the climate system, several scaling regimes with distinct power law exponents exist (Huybers and Curry, 2006; Lovejoy and Varotsos, 2016). However, at least for the Holocene, the possibility of only one scaling regime cannot be rejected (Nilsen et al., 2016).

1.4 Tackling gaps in the understanding of surface climate variability

Most studies that investigate climate variability focus on surface or near-surface temperature variability during the observational period (cf. Sec. 1.1.3). As such, they cover short-term variability, mostly from daily to interannual timescales. Beyond the observational period, proxy-based reconstructions allow the analysis of variability on decadal and longer timescales. Due to computational constraints, studies of simulated variability still often rely on equilibrium or short transient simulations, such that long timescales remain inaccessible to study. For paleoclimates, research tends to concentrate on recent periods for which observations are more abundant and quasi-equilibrium climate states. As a result most studies of long-term variability focus on the past millennium (cf. **P2**). Explorations of variability since the LGM or whole Holocene remain few and transitionary periods like the Last Deglaciation in particular remain largely unexplored (cf. Sec. 1.1.3).

The reliable simulation of variability is crucial to cover the full range of potential future climate and thus, for example, of extremes (Blanusa et al., 2023; Ionita et al., 2021; Katz and Brown, 1992; Schär and Jendritzky, 2004). An accurate representation of variability in models also provides the basis for attribution studies (Harrington et al., 2021; Hegerl and Zwiers, 2011) and the quantification of impacts of climate change (Calel et al., 2020; Schwarzwald and Lenssen, 2022). While studies of future variability under various emission scenarios exist (cf. Sec. 1.1.3), these are limited in their scope by the short lengths of the simulations. Further, comparisons of simulated with observed variability suggest that models regionally underestimate variability on decadal and longer timescales (**P2**). In this context, projections are unlikely to represent the full range of variability and thus possible future climate conditions. As such, considering past climate can extend our understanding of climate variability on longer timescales, under various mean states as well as during climatic transitions.

This work tackles the question of how variability evolves in a changing climate, in particular as it is simulated by climate models. To this end, we consider an ensemble of Deglacial simulations from models of various complexities (**P1**, **P3**). Relying on the moments as well as the spectra of surface temperature and precipitation, we analyze how variability depends on timescale (**P3**, Sec. 3.1.1), background state (**P3**, Sec. 3.1.2), forcings (**P3**, Sec. 3.1.3) and model complexity (**P3**, Sec. 3.1.4). We then explore the regional model-observation mismatch on decadal and longer timescales and possible causes (**P2**, **P3**, Sec. 3.2.1), as well as compare simulated variability since the pre-industrial period with observations and reanalyses (Sec. 3.2.2). Lastly, we contrast how variability changed with warming during the Deglaciation with results from ensembles of future scenarios (Sec. 3.3).

2 Research objectives

Surface climate variability, expressed in the statistical properties of the distributions of surface temperature and precipitation, depends on temporal and spatial scale, mean state and forcing. In three publications (**P1–P3**) and additional analyses in chapter 3 (**AA**), we investigate the representation of this variability in models on annual to millennial timescales in models and how it compares to observations with the following research objectives (**O**):

O1: To determine the dependence of climate variability on temporal and spatial scale
Approach: Analyze moments of simulated distributions and spectra of surface temperature and precipitation on annual to millennial, local to global scales.
Results: We confirm that model simulations and reconstructions of surface temperature variability from the Holocene to present-day agree globally, but find local disagreement as simulations generally show lower levels of variability than models, especially on decadal and longer timescales (**P2**). Surface temperature shows a stronger dependence on timescale than precipitation (**P3**). While models agree regarding variance changes in space and time, there are large inter-model differences when it comes to skewness and kurtosis (**P3**, **AA**).

O2: To analyze the dependence of variability on the mean climate state.
Approach: Compare simulated variability in the quasi-equilibrium states of the LGM and Holocene to the transitionary Deglaciation.
Results: We find variability to be larger during the LGM for temperature and during the Holocene for precipitation, a difference that increases with timescale (**P3**). The Deglaciation and future warming stand out as a periods of high variability in comparison to the LGM and Holocene and larger changes in variability are associated with stronger and faster warming (**P3**, **AA**).

O3: To isolate the effects of forcings and boundary conditions on simulated variability.
Approach: Analyze simulated variability in sensitivity experiments employing different ice sheet reconstructions, methods of meltwater distribution and with or without volcanic forcing.
Results: We show that meltwater forcing, ice sheet changes and volcanism significantly affect simulated climate and its variability (**P1**, **P3**, **AA**). Their effects on variability are larger for surface temperature than precipitation (**P3**). Effects exist not only on the forcings' characteristic timescales, but across the investigated timescales, from annual to millennial (**P3**).

O4: To identify the sufficient and necessary complexity of climate models for the representation of variability.
Approach: Establish a model hierarchy and compare simulated variability with respect to the models' ranking in that hierarchy.
Results: We establish a hierarchy of climate models and sort them into four categories, ESMs,

AOVGCMs, EMICs and EBMs, using eight different axes (**P1**, **P3**). Low and intermediate complexity models lack state-dependency and have spatial patterns that are either too homogeneous or too extreme (**P3**). With the minimal necessary level of complexity of GCMs, however, differences in simulated variability start depending more on forcing and model setup than model complexity (**P3**).

O5: To evaluate how simulated variability differs between warming in the past and in future projections
Approach: Analyze moments of surface temperature and precipitation in low to high emission future scenarios until 2300 CE and compare with results in simulations of the past 25 kyr
Results: We find that past and future warming show strong resemblances in mean and standard deviation, but that inter-model differences are large with respect to the tails of distributions (**AA**). However, in high emission scenarios, similarities to Deglacial warming can break down and abrupt changes in variability happen more often (**AA**).

3 Results and Discussion

Analyzing distributions and spectra of surface temperature and precipitation in an ensemble of simulations since the LGM, we show how simulated variability depends on timescale, location and spatial scale (Sec. 3.1.1). Simulated variability further depends on the background state, as a comparison of variability between the LGM and Holocene demonstrates (Sec. 3.1.2). Turning to the Last Deglaciation, we further investigate the expression of variability in a transitionary climate in contrast to that in such quasi-equilibrium states (Sec. 3.1.2). We highlight commonalities, but also find notable differences, which can relate to differences in forcings and boundary conditions (Sec. 3.1.3), as well as model complexity (Sec. 3.1.4).

In Sec. 3.2, we contrast our analysis of simulated variability with results from observations. For long-term variability from the LGM to Holocene, we turn to the LGMR as well as results from the literature to explore mismatches between models and observations and possible reasons (Sec. 3.2.1). Reanalyses and records from the observational period serve as a point of comparison for present-day annual variability (Sec. 3.2.2). Altogether, the results provide a perspective on challenges in comparisons of models with observations and we discuss potential options for resolving these issues (Sec. 3.2.3).

Lastly, we extend our analysis of simulated surface variability to future projections (Sec. 3.3). In ensembles of three future scenarios, we investigate how simulated future variability differs under past and future warming.

3.1 Simulated variability of surface temperature and precipitation on annual to millennial timescales

To evaluate simulated variability of surface temperature and precipitation, we use an ensemble of simulations of the Last Deglaciation (Sec. 1.2.1). All of the simulations encompass the Last Deglaciation and most cover the LGM to present-day. As such, they are long enough to allow the study of variability on annual to millennial timescales. In **P3**, we analyze the ensemble by separating the detrended timeseries of surface temperature and precipitation into LGM, Deglaciation and Holocene. We then compute the annual, decadal and centennial moments of the distributions, as well as global and regional power spectra for each period. Further details on the methodology can be found in **P3** and the results will be summarized and discussed below. Here, we add to the analysis of **P3** by first considering the evolution of the annual moments over time (Fig. 3.1, 3.2 in Sec. 3.1.1). To this end, we compute the annual moments over 100 yr sliding windows from the LGM to present-day. Additionally, we compare the centennial moments in three of the simulations of ¸**P3** in further detail (Fig. 3.3, 3.4 in Sec. 3.1.4).

In addition to the analysis of the dependence of simulated variability on timescale and background state, we investigate the influence of forcings and model complexity. To this end, the ensemble contains simulations that vary only by ice sheet reconstruction, meltwater protocol and the in- or exclusion of volcanic forcing. The ensemble further includes simulations from models of varying complexity enabling us to analyze the relationship between model complexity and simulated variability.

3.1.1 Dependency of simulated surface climate variability on scales in time and space

Examining the shape of the distributions of surface temperature and precipitation, we find that the moments change from annual to centennial timescales (**P3**). For surface temperature, the standard deviation decreases towards longer timescales, while absolute skewness and kurtosis tend to increase (**P3**, Fig. S1 and S2). These changes are particularly strong during the Deglaciation. As such, temperature distributions become less normal for the longer timescales investigated here.

Simulated temperature variability tends to be larger towards higher latitudes and small in the tropics (**P3**). This can be seen most clearly for standard deviation and in the spectra, while skewness and kurtosis show this trend only to a lesser degree. An exception is the MPI-ESM simulation that includes volcanic forcing (r7), which simulates considerable non-normality and increased variability in the tropics[1]. Some further differences exist between land and ocean areas as standard deviation tends to be higher over land than over ocean, while the higher order moments can be particularly large over the ocean.

Over time and with warming, standard deviation of surface temperature decreases from the LGM to the Holocene, especially in high latitudes (Fig. 3.1). While the decrease in standard deviation happens relatively smoothly, skewness and kurtosis show that abrupt shifts in temperature distributions do happen within decades or centuries and affect the tails (e.g., Fig. 3.1 l, n, o). Such shifts mostly occur during the Deglaciation (Fig. S6), but can happen during LGM and Holocene as well, in particular with the inclusion of volcanic forcing (cf. Fig. S5 l and S7 l).

The temperature spectra show increasing power towards longer timescales, often with a scale break around multi-decadal to centennial scales globally and in some regions (**P3**). If a scale break is present, the spectra tend to flatten on longer, centennial to millennial, timescales. The spectra also confirm that surface temperature variability tends to be larger at high latitudes.

Distributions of precipitation, similar to those of surface temperature, narrow towards longer timescales as standard deviation decreases (**P3**). The largest standard deviation can be found in the tropics in relation to the intertropical convergence zone, a band of strong convection, heavy precipitation and low pressure near the equator. With increasing latitude, standard deviation decreases. On annual scales, precipitation distributions are mostly positively skewed. This

[1] During the Holocene, MPI-ESM r1–r6 additionally exhibit high tropical variability in contrast to the rest of the results. This reflects a peculiarity of the model, in particular in its representation of ENSO (Ellerhoff and Rehfeld, 2021), rather than any physical processes in the climate system. As such, we consider this simulated tropical Holocene variability an unphysical outlier and exclude it in the discussion of our results going forward.

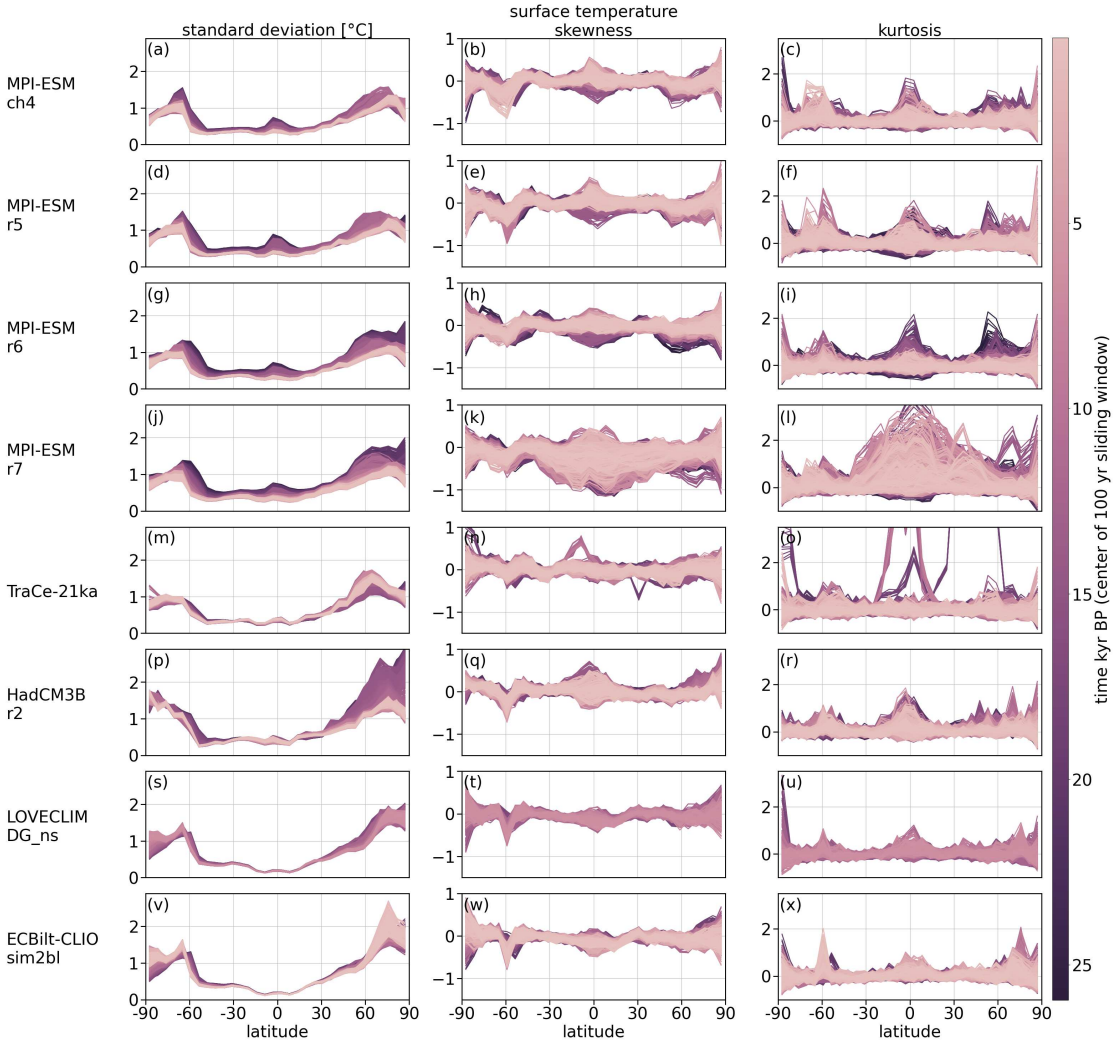


Figure 3.1: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of surface temperature across latitudes since the LGM. Moments are computed over 100 yr long rolling windows. Time is given as the middle of that rolling window.

skew then decreases towards longer timescales. The degree of this decrease is state-dependent as the quasi-equilibrium LGM and Holocene behave differently to the transitory Deglaciation. For the Deglaciation, skewness turns negative in the Northern Hemisphere and parts of the tropics on decadal and centennial scales. For the LGM and Holocene, on the other hand, precipitation distributions skewness tends to be close to zero almost everywhere on decadal and centennial scales. Kurtosis is mostly positive on annual scales, especially in the tropics and Antarctica. During the Deglaciation, such heavy tails prevail on long timescales, whereas LGM and Holocene tend to show little kurtosis. As a result LGM and Holocene precipitation distributions are normal on long timescales. In general, precipitation variability is largest in the tropics across moments and in the regional spectra.

Changes in standard deviation of precipitation over time show a meridional and state-dependency as it decreases in the tropics, but slightly increases in the mid- and high latitudes during the Deglaciation (Sec. 3.2). During the LGM and Holocene, standard deviation remains comparatively stable. Some models show changes in the tropics, but without any common trends. The higher order moments can change abruptly, especially during the beginning of the Deglaciation and Holocene (Fig. S9, S10). During the Holocene, skewness tends to increase in the subtropics in the Northern hemisphere, while it decreases or remains the same elsewhere. Kurtosis similarly changes mostly in the subtropics. During the LGM and Deglaciation, it decreases in the subtropics of the Southern Hemisphere, while it increases in the subtropics of the Northern Hemisphere during the Holocene .

The decrease in standard deviation of surface temperature throughout the Deglaciation coincides with the reduction of sea ice cover (**P3**). Our results confirm the crucial role of sea ice in creating local variability on decadal and longer timescales in line with previous studies Collow et al. (2019); Dai and Deng (2021); Ellerhoff et al. (2022); Huntingford et al. (2013). Along the decrease in standard deviation, our results show that the higher order moments increase towards longer timescales, as surface temperature distributions become less normal. This is in contrast to expectation based on the Central Limit Theorem, which would suggest that fluctuations tend toward normality on longer timescales. While this is the case for stationary systems such as equilibrium climate model simulations, changes in boundary conditions and abrupt shifts occur in transient simulations. As such, even after removing the multi-millennial trend, the simulation data can only be assumed to be weakly stationary as fluctuations on shorter than millennial timescales remain. Therefore, distributions can be more normal on annual than centennial timescales, as we find especially during the Deglaciation.

With warming, the higher order moments shift, at times, abruptly. As skewness can be an early warning signal for abrupt changes (Guttal and Jayaprakash, 2008; He et al., 2013; Skelton et al., 2020), these shifts in skewness in the Deglacial ensemble would merit investigation in the future, but was out of scope for this work. In particular, establishing the timing between changes in skewness and changes in the mean state could clarify whether the shifts signal abrupt events.

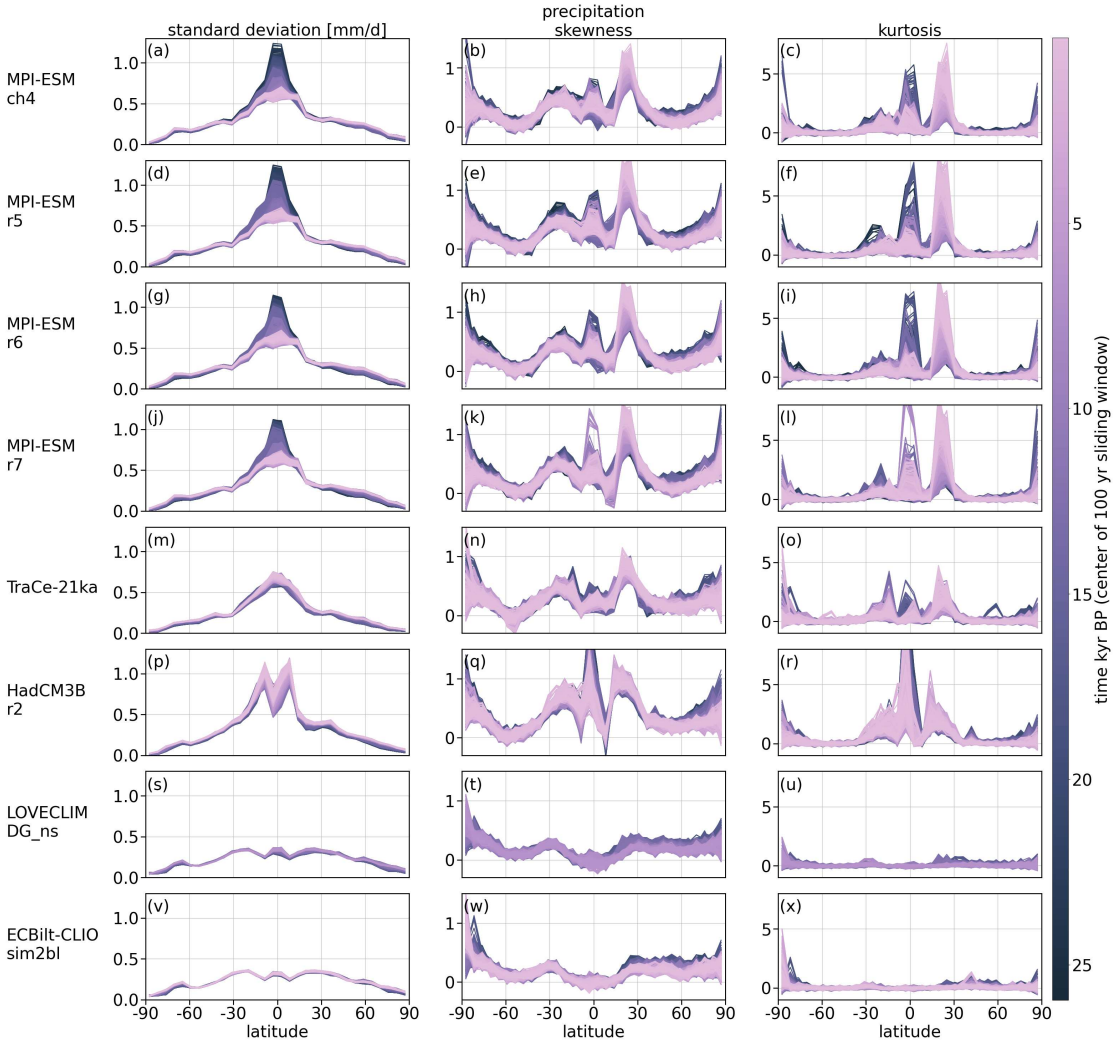


Figure 3.2: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of precipitation across latitude since the LGM. Moments are computed over 100 yr long rolling windows and time is given as the middle of that rolling window.

3.1.2 State-dependency of simulated surface climate variability

Simulated variability depends on the background state (**P3**). This state-dependency manifests in different ways depending on the aspect of variability under consideration. Standard deviation and power spectra of surface temperature show that variability is usually largest during the Deglaciation and the smallest during the Holocene. In particular, almost all simulations show that LGM variability is larger than Holocene variability, although to varying degrees regionally. This state-dependency is relatively small on annual timescales, but increases with timescale such that on decadal and longer timescales the Deglaciation stands out as a period of high temperature variability in comparison to the LGM and Holocene. State-dependency of temperature further tends to increase towards higher latitudes, especially in the Northern Hemisphere and areas close to the sea ice edge. For skewness and kurtosis, no such clear patterns emerge as state-dependency varies between models and forcing protocols (**P3**). Generally, absolute moments of temperature tend to be larger during the Deglaciation, excluding an erroneous Holocene tropical signal in some of the MPI-ESM simulations (Ellerhoff and Rehfeld, 2021). During the Deglaciation the more complex models, GCMs and ESMs, show a bipolar skewness pattern between the hemispheres. However, which hemisphere shows predominantly positive and which negative skewness depends on meltwater forcing and employed ice sheets.

Similarly, precipitation variance and power spectra also demonstrate that the Deglaciation is a high variability period in comparison to LGM and Holocene (**P3**). Comparing LGM and Holocene, the LGM with its comparatively dry mean state has the smaller precipitation variance with few exceptions. The polar regions in particular are an area of decreased precipitation variance in comparison to the Holocene. State-dependency of precipitation remains similar from annual to centennial timescales (Fig. S14, **P3**). Beyond centennial timescales, it increases, especially for the Deglaciation (**P3**). As for temperature, skewness and kurtosis of precipitation show heterogeneous patterns, with the absolute values largest during the Deglaciation. The Deglaciation also has more areas of significant skewness and kurtosis. For Deglacial precipitation, the bipolar pattern between the hemispheres exist, but to a lesser extent than for temperature, as it is mostly confined to the tropics.

Overall, the Deglaciation stands out as a period of high variability in comparison to the LGM and Holocene. As Deglacial variability has rarely been studied and no other study investigated the tails of distributions of surface temperature and precipitation, further work is needed to clarify potential patterns in tail behavior under warming. An investigation of distributions using a large collection of reconstructions from multiple proxy types could illuminate these points, but was outside the scope of this work. Section 3.2 will, however, provide a comparison of simulated variability to various observation-based products such as a reanalysis of the last 25 kyr.

Our findings of generally larger temperature variance during the LGM in comparison to the Holocene are in line with results from Rehfeld et al. (2020, 2018) Shao and Ditlevsen (2016) and Shi et al. (2022). The global LGM-to-Holocene temperature variance ratio in the ensemble is generally equal to or larger than that found for PMIP3 equilibrium simulations across timescales (**P3**). We find that the state dependency increases with timescale and is larger at high latitudes. Rehfeld et al. (2020) similarly found larger differences between LGM and present.day

at high latitudes. Unlike Rehfeld et al. (2020), we find no clear land-sea contrast where variability increases over land with warming, but decreases over the ocean. An in-depth comparison of surface climate variability changes over land and the ocean with warming considering the changing land-sea masks of the ensemble members could further clarify this. The pattern of smaller precipitation variance during the LGM agree with results for PMIP3 simulations (Rehfeld et al., 2020), including the strong decrease in precipitation variance over Northern North America during the LGM in comparison to the Holocene (**P3**) and present-day (Rehfeld et al., 2020).

3.1.3 Forcing dependency of simulated surface climate variability

Forced variability is a crucial component of simulated variability. In particular, since models have to rely on forcings to represent internal climate processes that are not implemented in the model and for example apply reconstruction of ice sheets instead of directly simulating their evolution. To investigate the dependency of simulated variability on ice sheet reconstruction, meltwater protocol and volcanic forcing, we analyze a set of sensitivity experiments (cf. **P3**, Sec. 1.2.1).

Regarding the influence of ice sheet reconstruction, we consider four simulations from two setups of MPI-ESM. For both model setups, one simulation employs GLAC1-D (Briggs et al., 2014; Tarasov et al., 2012), the other ICE6G (Peltier et al., 2015). The reconstructions differ with respect to height, extent and spatio-temporal evolution of the ice sheets with glacial ice sheet extent larger and height lower in GLAC1-D. Further, ICE6G has an original temporal resolution of 500 yr, while that of GLAC1-D is 100 yr. The reconstructed evolution of the ice sheets entails releases of meltwater. The PMIP3 Deglacial protocol offers three options for dealing with the meltwater in the simulations: 1) the melt-routed scenario, in which river-routing schemes distribute the meltwater, 2) the melt-uniform scenario, in which the same meltwater flux gets applied everywhere globally, and 3) the no-melt scenario, in which no meltwater is introduced despite the ice sheets melting (Ivanovic et al., 2016). For MPI-ESM, the ensemble contains one simulation using ICE6G for each of these scenarios. For HadCM3B, there is a melt-routed and a melt-uniform simulation. Lastly, the ensemble includes two MPI-ESM simulations that differ only in the inclusion of volcanic forcing in one of them.

Differences between ice sheet reconstructions used as boundary conditions influence simulated variability on timescales from annual to millennial, not only on their characteristic timescale, on which their impacts are strongest (**P3**). Instead, their impacts on the climate system through their influence on e.g., circulation patterns, affect processes acting on shorter timescales. Overall, changes in the ice sheets have a far stronger influence on the moments of surface temperature compared to those of precipitation. For the Deglacial ensemble of **P3**, the GLAC1-D reconstruction is often associated with higher variability than ICE6G, especially on longer than annual timescales, with some, regionally confined, exceptions. During the Deglaciation there is a notable bipolar pattern on centennial scales: Surface temperature distributions are skewed towards locally lower temperatures in the Northern Hemisphere and towards higher temperatures in the Southern Hemisphere, where the Antarctic ice sheet melts less by comparison to the shrinking

ice sheets in the North. In particular, the locations of meltwater pulses in the melt-routed scenarios are associated with regions of negative temperature skewness and positive skewness in the other hemisphere.

The differences in time step between the GLAC1-D and ICE6G reconstructions might contribute to the observed differences in variability. However, the spectra of both surface temperature and precipitation show larger variability in the Northern Hemisphere mid- and high latitudes for the GLAC1-D simulations from centennial to millennial timescales, including at the underlying timescales. As such, the different time steps cannot alone explain the differences in simulated variability, which instead likely stem from the properties of the reconstructed ice sheets.

The ice sheet reconstructions have considerable associated uncertainty, especially regarding ice sheet elevation and extent (Abe-Ouchi et al., 2015; Ivanovic et al., 2016; Stokes et al., 2015), but also their evolution over time, and both provide plausible boundary conditions. Due to their slow response and depending on future emissions, melting of ice sheets could continue to increase and continue for centuries to come, resulting in a corresponding rise in sea levels, with widespread consequences, among them potentially, changes in variability (Fox-Kemper et al., 2021). As shown in **P3**, the choice between them affects not only the mean climate, but also its variability and the simulation of extremes. As such reducing the uncertainty of ice sheet reconstructions remains vital for paleoclimate simulations during Glacial periods. However, using ice sheet reconstructions as a boundary conditions means that the simulated climate lacks interactions between ice sheets and climate, as the ice sheets cannot develop freely in interaction with the climate system as a whole. Coupled ice-sheet-climate models, on the other hand, allow for the interaction and co-evolution of climate and ice sheets. Their usage could thus provide insights into the relationship between ice sheets and variability.

Meltwater forcing is a source of considerable variability in a warming climate, as seen in the comparison of simulations with and without meltwater forcing (**P3**). Simulations with meltwater forcing show increased variability on annual to millennial timescales in all their moments and power spectra. Scenarios with routed meltwater further generate more variability than those with a globally homogeneous distribution of meltwater. Places of meltwater injection, like the North Atlantic during the Deglaciation, exhibit with increased standard deviation, negative skewness and positive kurtosis. The impact of meltwater forcing increases with timescale and is stronger for surface temperature than precipitation. Differences in associated meltwater forcing likely contribute significantly to the observed differences between ice sheet reconstructions. To separate the effects of meltwater forcing from changes in ice sheet extent and elevation, another no-melt simulation would be needed using GLAC1-D.

Volcanism strongly affects the variability of surface temperature, but not so much precipitation (**P3**). Its impact is largest on annual timescales (**P3**, Fig. 3.1), with a decrease towards longer timescales. From annual to centennial timescales, the inclusion of volcanic forcing leads to increased temperature variability, in agreement with findings from Ellerhoff et al. (2022). These impacts of volcanic forcing are particularly strong in the tropics and the northern mid-latitudes. In the tropics, the EBM simulation, as the only other run with volcanic forcing besides MPI-ESM r7, comes closest to that MPI-ESM run when considering the power in the regional spectra. Tropical variability will thus be particularly impacted by a lack of volcanic forcing in

simulations. Comparisons of simulated with reconstructed variability show that the mismatch is particularly large in the tropics (Laepple and Huybers, 2014a,b), as models underestimate variability. Our work suggests that the inclusion of volcanic forcing is a step towards closing that gap, although it is not necessarily sufficient. Instead, models seem to have weaknesses in simulating tropical dynamics on decadal and longer timescales. Cloud feedbacks are particularly relevant to tropical dynamics, but are still poorly understood and a large source of uncertainty (Bretherton, 2015; Heinze et al., 2019; Ingram and Bushell, 2021; Klein et al., 2017). The inclusion of volcanic forcing is however clearly necessary as focusing on tropical dynamics alone carries a risk of overtuning.

While volcanic forcings acts most upon short, annual, timescales, differences between simulations that only differ in the inclusion of volcanic forcing, persist at least until millennial timescales. This imprint of volcanic forcing on long timescales highlights how forcings with short characteristic timescales interact with components with longer characteristic timescales. With some regional exceptions, volcanic forcing changes the shapes of the distribution such that they are wider, with long and heavy low temperature tails (more negative skew and higher kurtosis). This suggests an increase in the likelihood of locally extreme low temperature conditions. For future projections, Bethke et al. (2017) found a similar change in temperature distributions with the inclusion of volcanic forcing, although their change in standard deviation is substantially larger. This underlines the importance of a realistic forcing in past and future simulations, especially when investigating extremes. For future projections, however, CMIP6 protocols for volcanic forcing have been found lacking as they do not sample from the full range of possible eruptions, instead employing constant volcanic forcing after 2024 CE (cf., Bethke et al., 2017; O'Neill et al., 2016).

The inclusion of volcanic forcing leads to a different response depending on the mean state from local to global, annual to millennial scales (**P3**). Many studies have found such state-dependent differences in the response to volcanic forcing for time periods from the past millennium to future simulations (e.g., Muthers et al., 2015; Zanchettin et al., 2013). However, they suggest at times opposite directions of change, from a stronger (Fasullo et al., 2017) to a weaker response in a warmer climate state (Hopcroft et al., 2018). Ellerhoff et al. (2022), on the other hand, found no state-dependency when comparing global temperature variability in response to volcanic forcing during LGM and PI and only small local differences, mostly on interannual scales. Their sensitivity experiments consisted of 1000 yr long HadCM3 simulations using LGM or PI boundary conditions with transient solar and volcanic forcing, in contrast to our fully transient simulations. This suggests that transient solar and volcanic forcing alone do not suffice to capture the full response of the climate system to time-varying volcanic forcing and that transient simulations should be used whenever possible. However, the Deglacial ensemble contains only two simulations that incorporate volcanic eruptions, one of them the EBM that can only capture the linear climate response to external forcing. As such, more experiments are needed to confirm these results, for example by comparing a transient simulation to one where only volcanic forcing varies and including with other climate models.

3.1.4 Dependency of simulated surface climate variability on model complexity

As outlined in Sec. 1.2.1, we classify the complexity of a model based on the the complexity of the individual compartments. For the Deglacial ensemble (**P3**) models then fall into one of four categories: EBM, EMIC, GCM, ESM. The ensemble contains eight simulations from an ESM, four from a GCM, two from an EMIC and one from an EBM.

Evaluating simulated variability according to the complexity of the model, we find that there is a minimal complexity required for a model to be able to represent patterns of variability adequately. This minimal complexity is that of the GCM category. Only GCMs and ESMs are able to simulate the variability of surface climate with some degree of accuracy. Among GCMs and ESMs, on the other hand, the differences of the individual models and simulation protocols matter more than the complexity of the models.

The EBM can capture global mean climate and, in broad strokes, regional climate on annual and longer timescales (**P1**). On monthly timescales, however, limitations of the simulated seasonal cycle and a lack of seasonality in sea ice forcing come to the fore and limit the EBM's abilities. For variability, the EBM struggles even at the global scale (**P3**). Its simulated standard deviation becomes more similar to the other models towards longer timescales, while regional and local differences remain (**P3**, Fig. 3.3). The spectra similarly show similarities globally and in the tropics, where the EBM is within the spread of the other models (**P3**). However, while variability in the other models increases from the tropics towards the polar region, it remains relatively constant for the EBM. As a result, the polar EBM spectra show lower variability than the more complex simulations. Analyzing skewness and kurtosis reveals that the EBM fails to simulate the full distribution of surface temperatures especially with regards to the tails and can thus not be used in the study of extremes.

Better agreement of the EBM with more complex models on the global scale reflects the fact that while global variability is largely a response to external forcing (Crowley, 2000). However, internal variability arising due to non-linear interactions becomes increasingly relevant at regional and local scale (Goosse et al., 2005), which the EBM lacks (**P1**). Additionally, the increasing similarity in standard deviation towards longer timescales might actually be due to underestimated variability in the more complex models rather than an improvement of the EBM on those timescales.

Overall, the EBM can only capture the linear climate response and has limited the role of many compartments to their average heat capacity, horizontal heat transport and albedo (**P1**). Simulated variability would likely improve if sea ice had an annual cycle or was modeled rather than provided as an annually constant boundary condition. Yet, without considering non-linear interactions and more complex compartments and processes, simulated variability will remain unrealistic. Targeted representations of global variability in an EBM are possible, e.g., by focusing on the scaling with timescale (cf. Lovejoy, 2021; Lovejoy et al., 2021). Instead of aiming for such a realistic representation, a model such as TransEBM can provide insights into the linear response to forcings and give rough estimates as to their relative importance (cf. Sec. 3.1.3, Crowley, 2000).

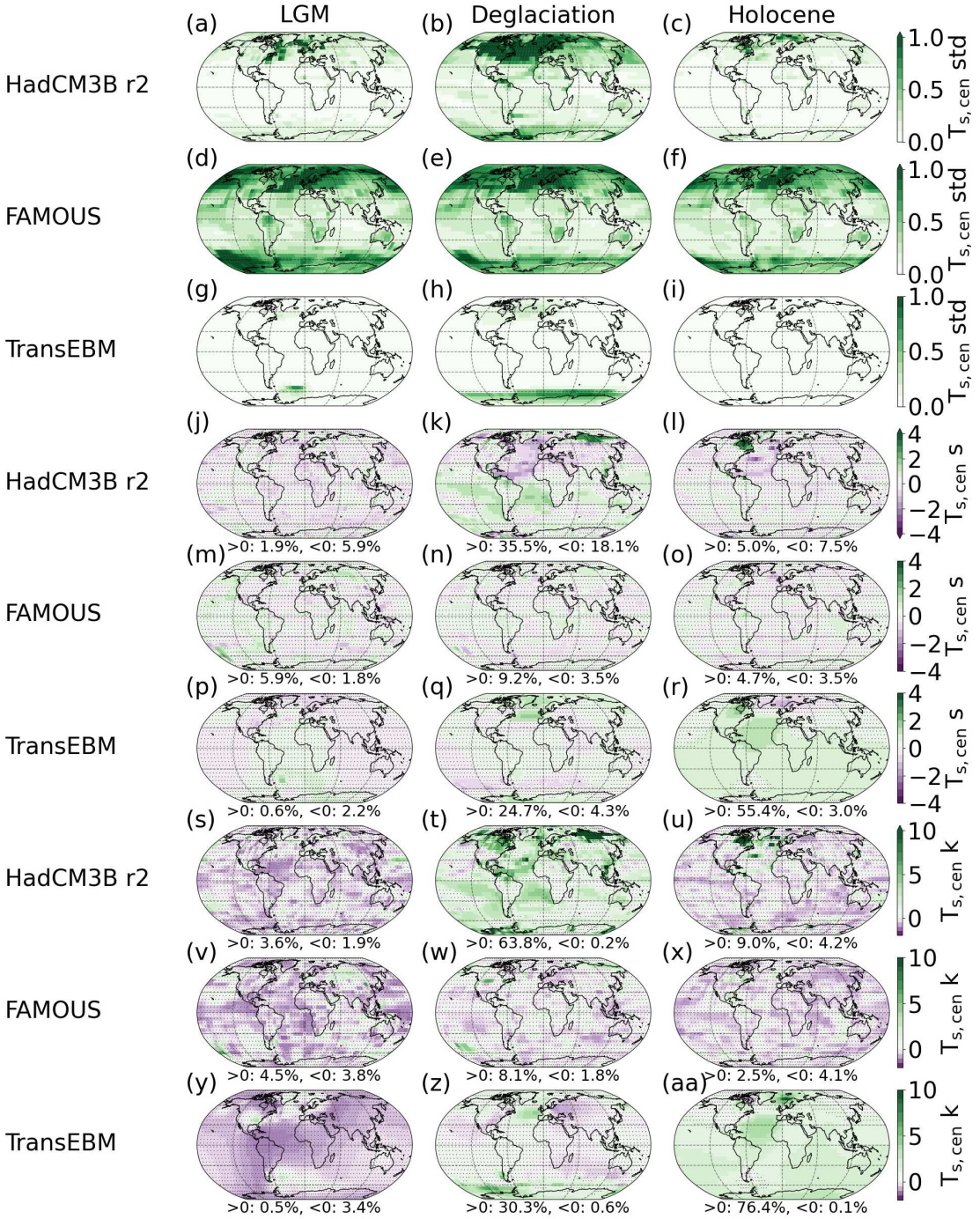


Figure 3.3: Centennial standard deviation (a)–(i), skewness (j)–(r) and kurtosis (s)–(aa) of surface temperature in three simulations: HadCM3B r2, FAMOUS and TransEBM (cf. **P3**). Moments are shown for the LGM (left column), Deglaciation (middle column) and the Holocene (right column). For skewness and kurtosis, the percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are given. Areas of non-significant changes are hatched.

The EMICs included in the Deglacial ensemble display a clear difference in their abilities of simulating variability of surface temperature versus precipitation (**P3**). Comparisons to more complex models and reanalyses (**P3**, Sec. 3.2.2) show that precipitation variability is severely underestimated. Patterns of surface temperature variability, on the other hand, are largely reproduced by the EMIC simulations. Especially in high latitudes, temperature variability is captured well, less so in low latitudes. Still, their simulated temperature variability remains on the lower end of the ensemble and is too small in many places and regions. Discrepancies are larger for skewness and kurtosis, suggesting that the EMICs in the ensemble are not suited for studies of extremes. The EMICs further perform worse in lower latitudes, which affects precipitation in particular. This weakness might be due to a model design focused on the ocean rather than atmospheric circulation and the hydrological cycle in the atmosphere (Goosse et al., 2010b). In this regard, looking at oceanic variability in the EMIC simulations could prove insightful, as could using an EMIC with a more complex atmosphere that improves for example on moisture transport or cloud parameterizations.

Within our ensemble the complexity of the GCMs seems sufficient to capture surface climate variability, at least in comparison to the reconstructions, reanalyses and direct observations available for comparison (**P3**, Sec. 3.2). Those comparison to observations suggest that there is no model in the ensemble that is across the board the best in simulating variability, as the models instead have individual strengths and weaknesses. A lot of the inter-model differences are due to varying forcing protocols further complicating any ranking efforts, as differences between forcing protocols with the same model can be as large as differences between models (cf. Sec. 3.1.3, **P3**). Nevertheless, improved comparisons of the models with observations, especially on multi-decadal and longer timescales and for climate variables outside of temperature, are needed and could inform future model development. Such efforts rely on improved proxy reconstruction and are hindered by a lack of available records, uncertainties for that data relating for example to reconstructions methods, age models or calibrations and the methods for comparing simulations with observations themselves (cf. Sec. 3.2.3). Tackling these uncertainties to provide a comprehensive evaluation of simulated variability was outside the scope of this work. Future work could undertake such a comparison to hone in on model weaknesses regarding variability and thus potentially identify the source of the underestimated regional variability on decadal and longer timescales.

As discussed above, a reliable simulation of variability requires at least the complexity of GCMs (**P3**). However, long-term transient simulations with GCMs and ESMs are restricted by the computational resources they require. Resolution has a major impact on the run-time of a simulation. Within the range of resolutions in our ensemble, we find no impact of model resolution on simulated variability. Confirmation that the effects of increased resolution on variability are at most minor would require an investigation with an ensemble from the individual models at different resolutions. Models included in such an ensemble would need to have parameterizations and implementations of model physics that change as little as possible with resolution. As things stand, resolution does not seem to have a large impact on simulated variability. In high-resolution models, ongoing model development efforts focus on including previously parameterized sub-grid processes, such as convection (e.g., Jucker et al., 2020; Kendon et al., 2021; Liu et al., 2022; Takayabu et al., 2022) or eddies (e.g., Bian et al., 2023; Bryan et al., 2010; Ding et al., 2022).

The physical modeling instead of parameterizations of such processes affects simulated variability (e.g., Ding et al., 2022; Judt and Rios-Berrios, 2021), but their high computational demand currently limits their usage for long simulations.

Another way to reduce computational demand besides limiting spatial resolution, is to employ accelerated forcing. In the Deglaciation ensemble used in our analysis, the FAMOUS simulation uses an acceleration factor of 10 such that forcing changes that would take 100 yrs only take 10 simulation years, shortening the run-time of the simulation. Simulated mean climate changes in FAMOUS are reasonable, but rather small from LGM to Holocene (**P3**). FAMOUS is the only model whose simulated LGM-to-Holocene GMST difference lies outside the range suggested by proxy reconstructions (**P3**, Sec. 3.2.1). Our analysis further shows that variability in FAMOUS matches neither the implicit nor the explicit timescale of the forcing. In fact, simulated variability in FAMOUS stands out as it lacks state-dependency in contrast to other GCM and ESM simulations (**P3**, Fig. 3.4). FAMOUS also produces much higher standard deviation and lacks skewness and kurtosis, suggesting that the simulation of extreme conditions is lacking as well. In fact, with respect to the higher order moments, the FAMOUS simulation is more similar to the EMICs (**P3**). As current changes in GHG conditions are much faster than they were in the past, this difference in variability in response to faster changes in forcing could hint to limitations in using the Deglacial warming to learn about current and future warming. However, variability in CMIP6 simulations is more similar to the other GCM and ESM simulations rather than FAMOUS (cf. Sec. 3.3). Nevertheless, determining the effects of accelerated forcing with certainty would require a comparison of accelerated and non-accelerated runs from the same model.

3.2 Comparison of variability in simulations, paleoclimate reconstructions, reanalyses and instrumental observations

So far, we have provided an inter-model comparison of variability. As models provide only one perspective on surface climate variability, we here supplement the analysis with observations. For long timescales, this implies using paleoclimate reconstructions based on proxies. We compare variability in models to that in reanalyses, reconstructions and observations of the climate from the LGM to present-day. In Sec. 3.2.1, we draw on results from **P2** and **P3** before expanding the analysis from **P3** in the end of the section (Fig. 3.5–3.8). Sec. 3.2.2 then adds a new perspective by comparing the simulations from **P3** to present-day observations and reanalysis. Finally, Sec. 3.2.3 puts the results of **P2**, **P3**, Sec. 3.2.1 and 3.2.2 in a broader context and discusses their implications for future comparisons of simulations with observations.

3.2.1 Reconstructed and simulated perspectives on surface temperature since the Last Glacial Maximum

Both proxy-based reconstructions and simulations of past climates carry significant uncertainty. This pertains to mean climate as well as variability. With respect to mean changes, the sug-

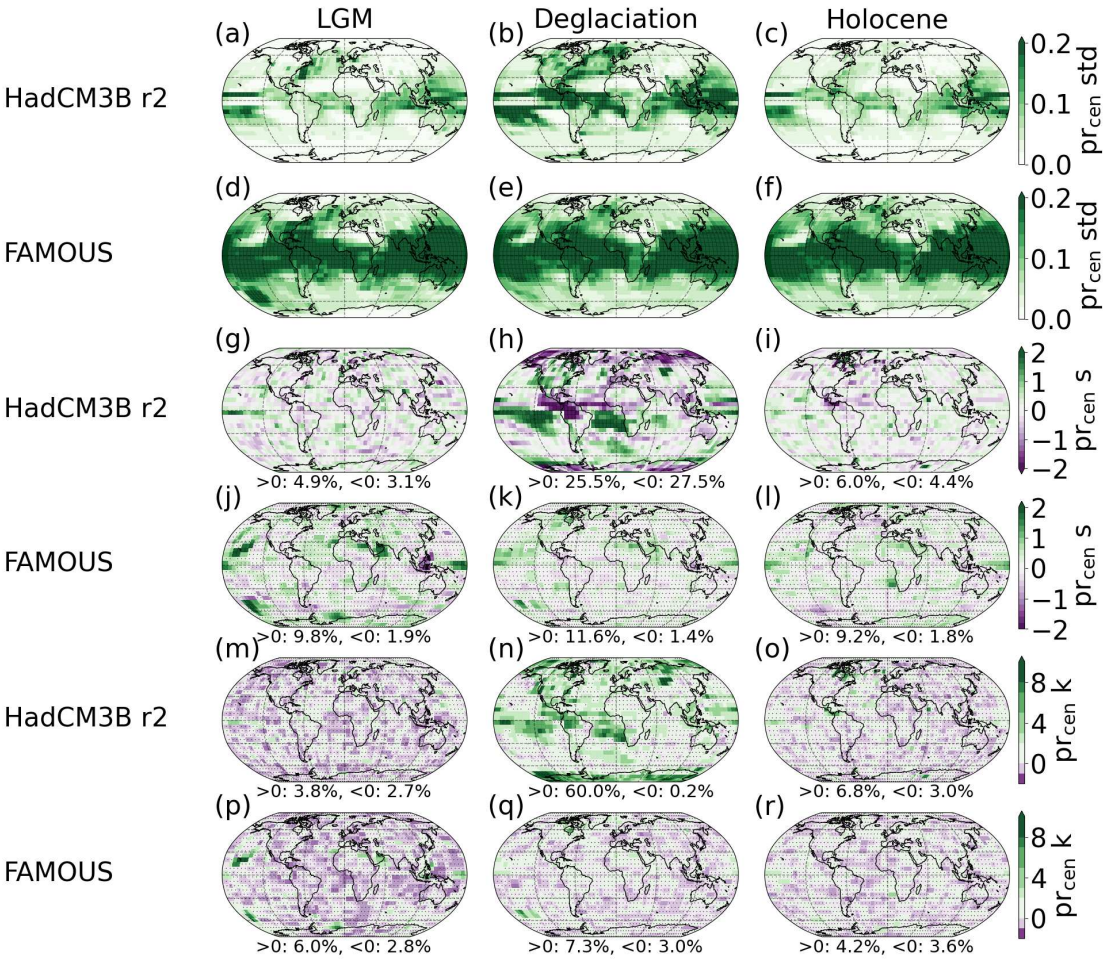


Figure 3.4: Centennial standard deviation (a)–(f), skewness (g)–(l) and kurtosis (m)–(r) of precipitation in two simulations: HadCM3B r2 and FAMOUS (cf. **P3**). Moments are shown for the LGM (left column), Deglaciation (middle column) and the Holocene (right column). For skewness and kurtosis, the percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are given. Areas of non-significant changes are hatched.

gested LGM-to-Holocene temperature difference varies between studies and simulations: data assimilation approaches produce a range of 4.5 to 7.0°C warming (Annan et al., 2022; Osman et al., 2021; Shakun and Carlson, 2010; Tierney et al., 2020), while models simulate a warming between 3.1 and 6.6°C (**P3**). All but one of the simulations in the ensemble used in **P3** (FAMOUS) fall within the range suggested by reconstructions. Patterns of cooling show similarities in simulated and assimilated field reconstructions, with strong LGM cooling in the polar regions of the Northern Hemisphere, particularly northern North America, and western Antarctica (cf. **P2** and Annan et al., 2022; Osman et al., 2021; Tierney et al., 2020).

Annan et al. (2022) demonstrate the impact of the model prior on the overall degree of warming as they created an ensemble based on priors from different LGM simulations from PMIP2 to PMIP4. In the resulting ensemble, the member based on CESM1.2 shows larger LGM cooling than the other members, while a CESM2.1 member was removed as an outlier due to its large equilibrium climate sensitivity, that is to say, due to its large warming in response to increasing levels of atmospheric CO_2 (Zhu et al., 2021). Tierney et al. (2020) and Osman et al. (2021) both use iCESM simulations as a prior, a model version based on CESM1.2, which might explain their relatively large inferred LGM-to-Holocene temperature differences of 6.1°C (5.7, 6.5) and 7.0°C (6.0, 8.0), respectively. Annan et al. (2022) argue that both these and a smaller LGM cooling are plausible.

The temporal evolution of surface temperature reveals notable differences between simulations and LGMR (Osman et al., 2021), particularly during the Deglaciation and abrupt events like the Northern Hemisphere warming at the onset of the Bølling-Allerød (**P3**). Most simulations show an onset of the Deglaciation around 19 kyr BP, with some simulating slight warming before then. The LGMR, on the other hand, cools until almost 17 kyr BP and suggests faster and stronger warming than the models for most of the Deglaciation afterwards. This seems to be driven by the northern polar and mid-latitude regions, a region comparatively well constrained by proxies especially in the North Atlantic. In those regions, the LGMR shows stronger warming than all simulations, resulting in a large decrease in the equator-to-pole temperature gradient.

When it comes to variability, agreement with reconstructions depends on the temporal and spatial scale considered (**P2**). Comparing studies of the Holocene that draw on a variety of proxy-based reconstructions, instrumental records and simulations demonstrates that simulated global variability agrees with observed and reconstructed variability on interannual to millennial timescales (**P2**). On regional scales, however, when several proxy records or grid boxes are aggregated on scales smaller than a hemisphere, results of comparisons of models and observational records vary. While studies are split between reporting agreement or disagreement, depending on region (Ljungqvist et al., 2019; PAGES2k-PMIP3 group, 2015; Zorita et al., 2010), proxy variable and archive (Parsons et al., 2017) or timescale (Cheung et al., 2017; Fredriksen and Rypdal, 2016). Studies that find timescale-dependent differences report agreement on shorter, interannual to decadal, and disagreement on multi-decadal and longer timescales.

Locally, at the level of individual proxy records, instrumental observations or grid boxes in models, this change from agreement to disagreement with timescale exists across studies (**P2**). Usually, the transition happens on multi-decadal timescales, although some studies report a change already on decadal (Ellerhoff and Rehfeld, 2021) and others only on centennial scales

(Bühler et al., 2021; Hébert et al., 2022). Few studies cover periods further back in time than the Holocene, and in fact even the majority of studies considered in **P2** cover at most the last 2000 years. Considering periods beyond the Holocene, Zhu et al. (2019) confirm global agreement of reconstructions with transient simulations starting at 21 kyr BP. Rehfeld et al. (2018), on the other hand, established that the global LGM-to-Holocene variance ratio is significantly larger in proxy records than in simulations and that the same holds at most latitudes. We find that the inclusion of transient volcanic forcing narrows this gap, but fails to close it (**P3**).

These disagreements between simulations and observations have been found for different types of observational records, from reanalysis and instrumental records to proxy records from a variety of climate archives like speleothems, corals, ice cores, sediments, pollen or tree rings (**P2** and references therein). This breadth of archives and recorders strengthens the confidence in the disagreement. Nevertheless, instrumental records only provide limited insights into longer timescales and comparisons to climate variables based on statistical transformations of proxy measurements are uncertain and potentially biased. Locally, lake and marine sediments as well as ice cores often dominate multi-proxy compilations for periods before the Common Era. For these records and comparisons at a local scale, proxy noise from non-climatic processes is the most likely source of bias for variability. For evaluations on larger spatial scales, averaging a collection of records dampens the impact of proxy noise and other sources of uncertainty become more relevant. There are many processes leading to proxy noise, some of which can lead to an underestimation, others to an overestimation of variability in reconstructions. However, even when applying statistical methods or using proxy-system models to remove such errors, model-observation differences persist across reconstructions from different proxy types, which agree with instrumental records for overlapping timescales. Thus, after consideration of various error sources, we conclude that reconstruction uncertainties explain at most part but not the full extent of differences. As observed local variability on decadal and longer timescales tends to be significantly larger than simulated variability, this means that simulations likely lack variability and are too stable, as has been suggested elsewhere (e.g., Valdes, 2011). This could imply that the local and regional variations that exist on annual timescales should persist on longer timescales (**P2**). On annual scales, we find spatial patterns that are less homogeneous with a small spatial extent of temperature anomalies in comparison to decadal and longer timescales. On decadal and longer timescales, spatial correlation of temperature increase and the difference between simulated local and global variability decreases, in contrast to observed variability.

Adding to these comparisons of simulated and observed variability, we compare variability in the Deglacial ensemble and LGMR. Here, we extend the comparison in **P3**, which used the LGMR ensemble mean field by taking the LGMR ensemble members into account. As the only field reconstruction available that covers the period from the LGM to today, the LGMR provides a perspective on the spatial and temporal development of surface temperature during that time, which is informed by proxies. However, the mean reconstruction investigated in **P3** provides a smoothed representation of surface temperature and as such only limited insights into variability. Mean latitudinal profiles and fields are generally similar across ensemble members (Fig. 3.5a and S25). Variability, however, differs considerably across the ensemble and from that found in the ensemble mean field. While the general pattern of enhanced standard deviation towards higher latitudes persists, all investigated ensemble members exhibit larger standard deviation,

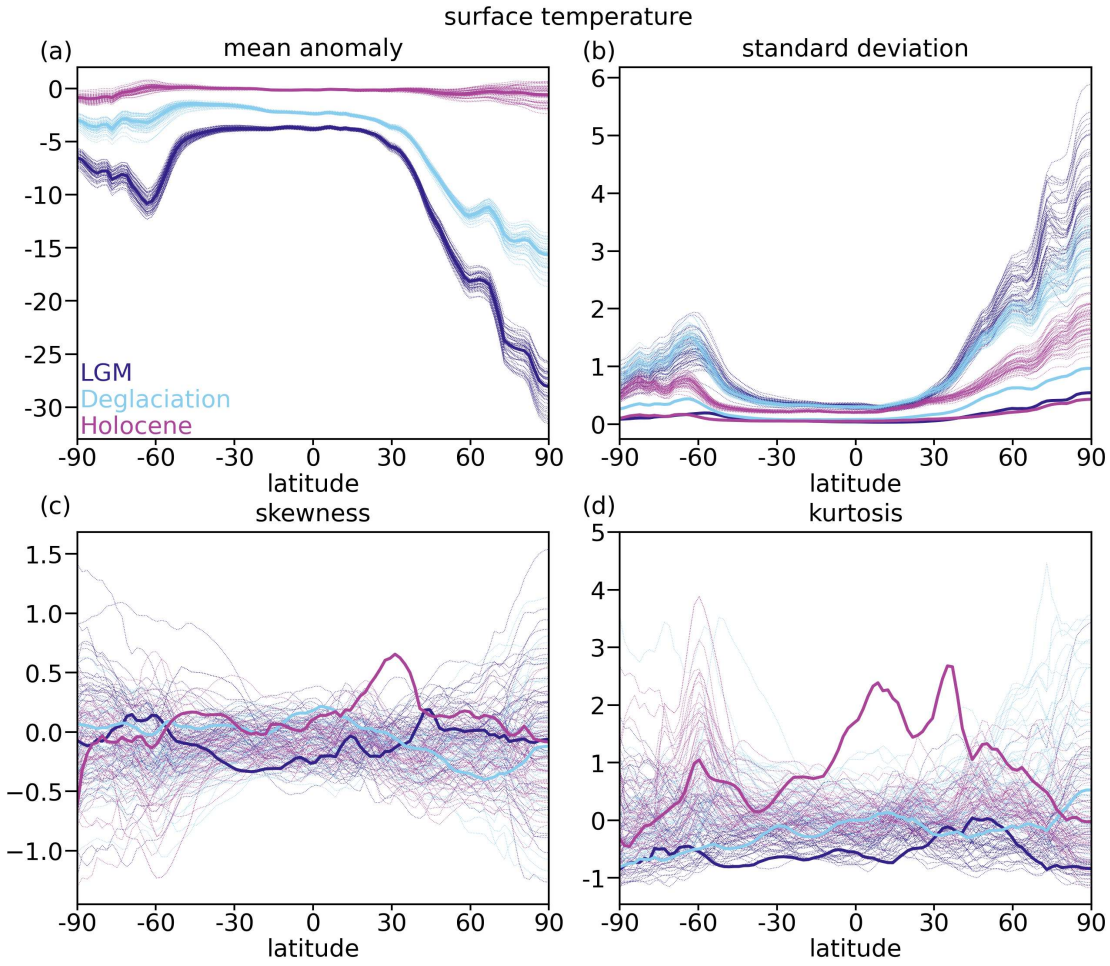


Figure 3.5: Moments across latitudes of the 50 ensemble members (dashed) of the LGMR by Osman et al. (2021), as well as the ensemble mean (solid). Ensemble mean as in **P3** but without regridding. Shown are (a) mean, (b) standard deviation, (c) skewness and (d) kurtosis for LGM (dark blue), Deglaciation (light blue) and Holocene (pink).

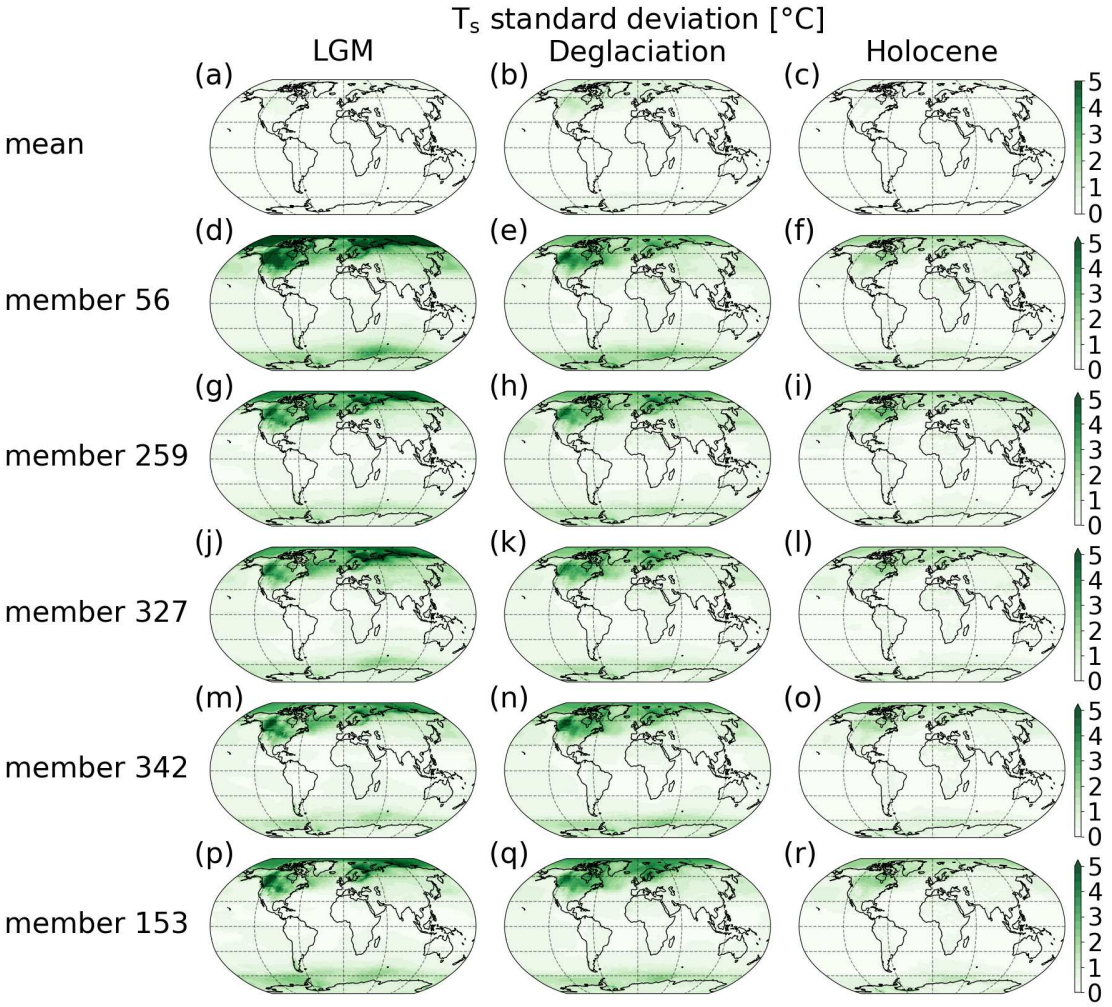


Figure 3.6: Standard deviation of surface temperature with respect to the past 2000 years of the LGMR for the LGM (left column), Deglaciation (middle column) and the Holocene (right column). Shown are in panels (a)–(c) the standard deviation of the ensemble mean (as in **P3** but without regridding) and in panels (d)–(r) that of five randomly selected ensemble members. The latter show considerably larger standard deviation, especially at higher latitudes.

especially at high latitudes, than the ensemble mean and all simulations (Fig. 3.5b and 3.6). In the Northern Hemisphere, standard deviation in the ensemble members keeps increasing with latitude, unlike in models (**P3** Fig. 5) and unlike in the Southern Hemisphere, where it peaks around 60°S. As another notable difference to the ensemble mean and the simulations, the ensemble members mostly show larger (Northern Hemisphere) or similar (Southern Hemisphere) standard deviation during the LGM than the Deglaciation. While the overall amplitudes vary between ensemble members, spatial patterns are generally robust with the largest standard deviation found over northern North America and Northern Europe across periods. That standard deviation during the LGM is higher than during the Deglaciation is likely the result of a lack of proxy constraints during the LGM, especially for the polar regions where records are sparse.

The higher order moments, on the other hand, show no consistent patterns across ensemble members as found for the mean and standard deviation (Fig. 3.5c,d, 3.7 and 3.8). In fact, even the overall direction of change varies between ensemble members across most regions. This might be, because extremes, and thus also the higher moments, are affected by the smoothing and spatial interpolation inherent in field reconstructions (Director and Bornn, 2015; Haylock et al., 2008). In the meridional patterns, there is a tendency towards larger absolute skewness at higher latitudes and kurtosis seems to develop a robust positive peak along 60S in the Southern Pacific during the Holocene (Fig. 3.5c,d). The moments of the ensemble mean also seem to be a poor representation of the ensemble members. Overall, the amplitudes of higher order moments are still smaller than in most simulations and indicate no deviation from normal distributions in most places. This might be a reflection of the ensemble Kalman filter used in the data assimilation process, which assumes that state vectors are normal (Evensen, 1994; Evensen et al., 2022). Despite the assumption of normality, ensemble Kalman filters have been shown to work for many weakly, though not strongly, non-linear systems, but this will depend on the application (Evensen, 1994; Evensen et al., 2022). Due to the sparsity of available proxy records, they provide little constraints on spatial variability patterns, such that the model prior might have a larger influence. As the ensemble members vary within the range of the prior, a lack of variability and spatial variations in the prior could translate to the ensemble members. As variability in the LGMR has not been evaluated previously, it remains unclear whether proxies, model prior or the assimilation algorithm dominate the reconstructed variability at specific timescales and in specific regions.

3.2.2 Variability in models, reanalyses and observations since 1850 CE

For present-day, we compare how the annual moments change over time in two reanalyses (ERA5 and NOAA 20CR v3) and one station product (HadCRUT5) to those Deglacial simulations that extend at least until 1950 CE to ensure sufficient overlap. The reanalyses and HadCRUT5 suggest a similar structure for standard deviation of surface temperature, matched by most simulations except the EBM (Fig. 3.9): Annual standard deviation grows towards higher latitudes and peaks in the polar regions of both hemispheres with a similar amplitude. NOAA 20CR and HadCRUT5 show growing standard deviation in time, although they disagree with respect to the amplitude of that change. Standard deviation in NOAA 20CR grows considerably

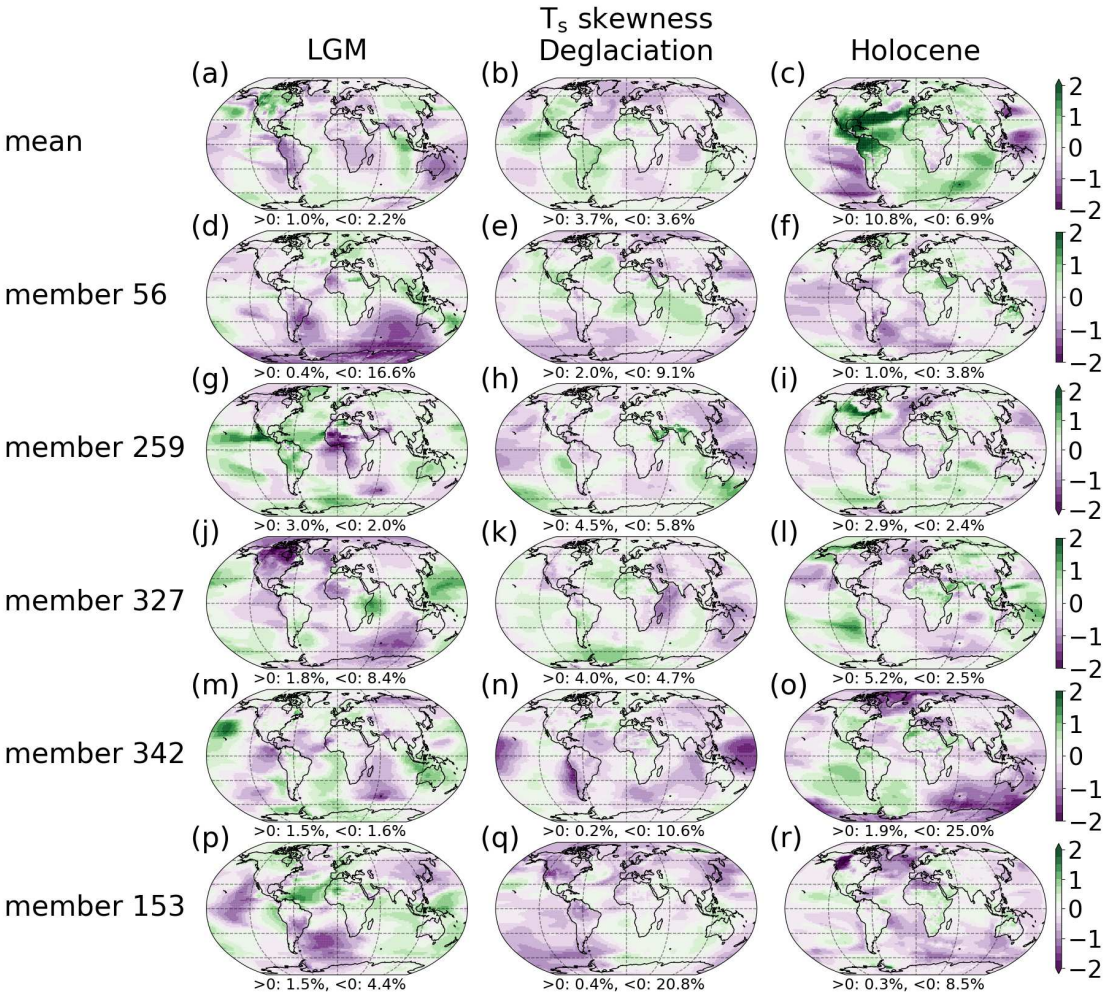


Figure 3.7: Skewness of surface temperature with respect to the past 2000 years of the LGMR for the LGM (left column), Deglaciation (middle column) and the Holocene (right column). Shown are in panels (a)–(c) the skewness of the ensemble mean (as in **P3** but without regridding) and in panels (d)–(r) that of five randomly selected ensemble members. Like the ensemble mean, the individual members rarely show significant deviation from a normal distribution, but patterns between members vary considerably.

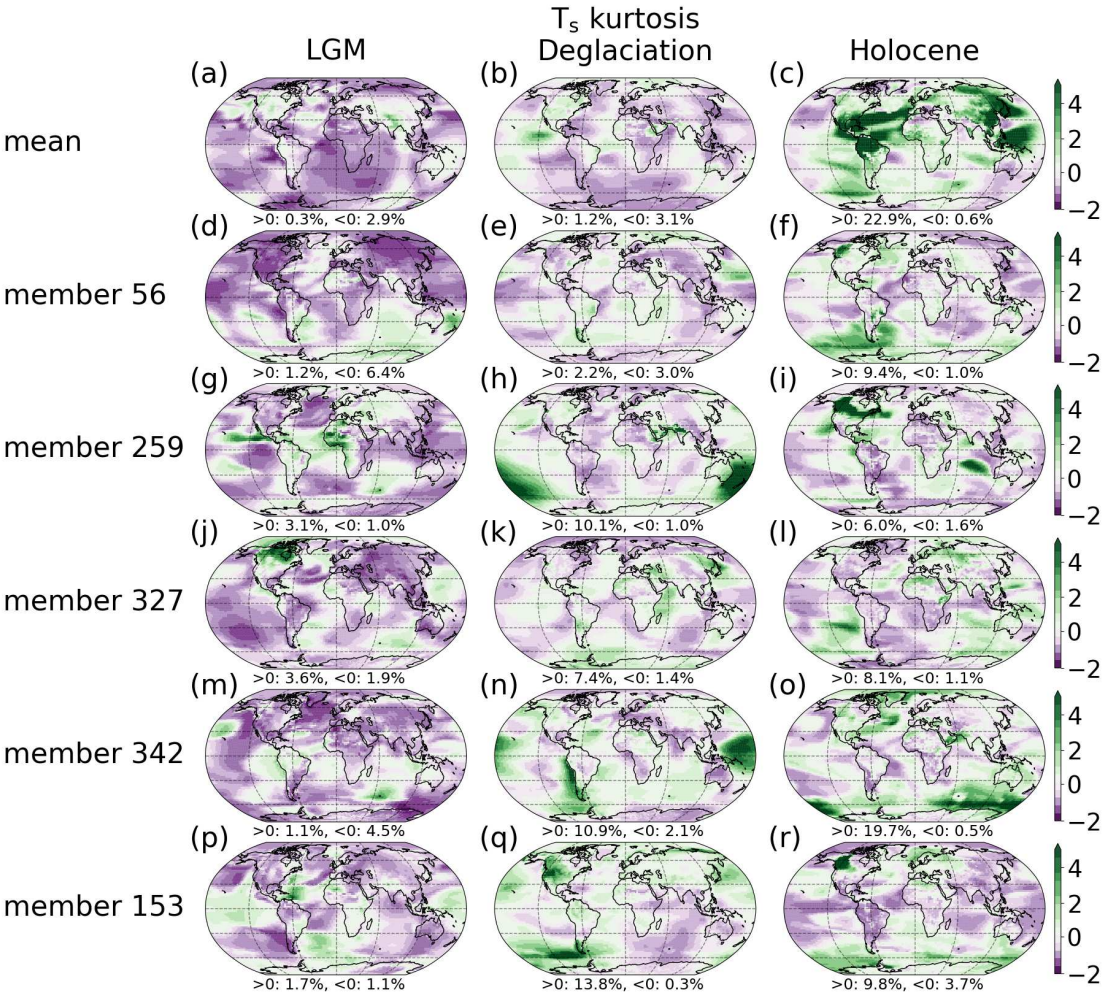


Figure 3.8: Kurtosis of surface temperature with respect to the past 2000 years of the LGMR for the LGM (left column), Deglaciation (middle column) and the Holocene (right column). Shown are in panels (a)–(c) the kurtosis of the ensemble mean (as in **P3** but without regridding) and in panels (d)–(r) that of five randomly selected ensemble members. Like the ensemble mean, the individual members rarely show significant deviation from a normal distribution, but patterns between members vary considerably.

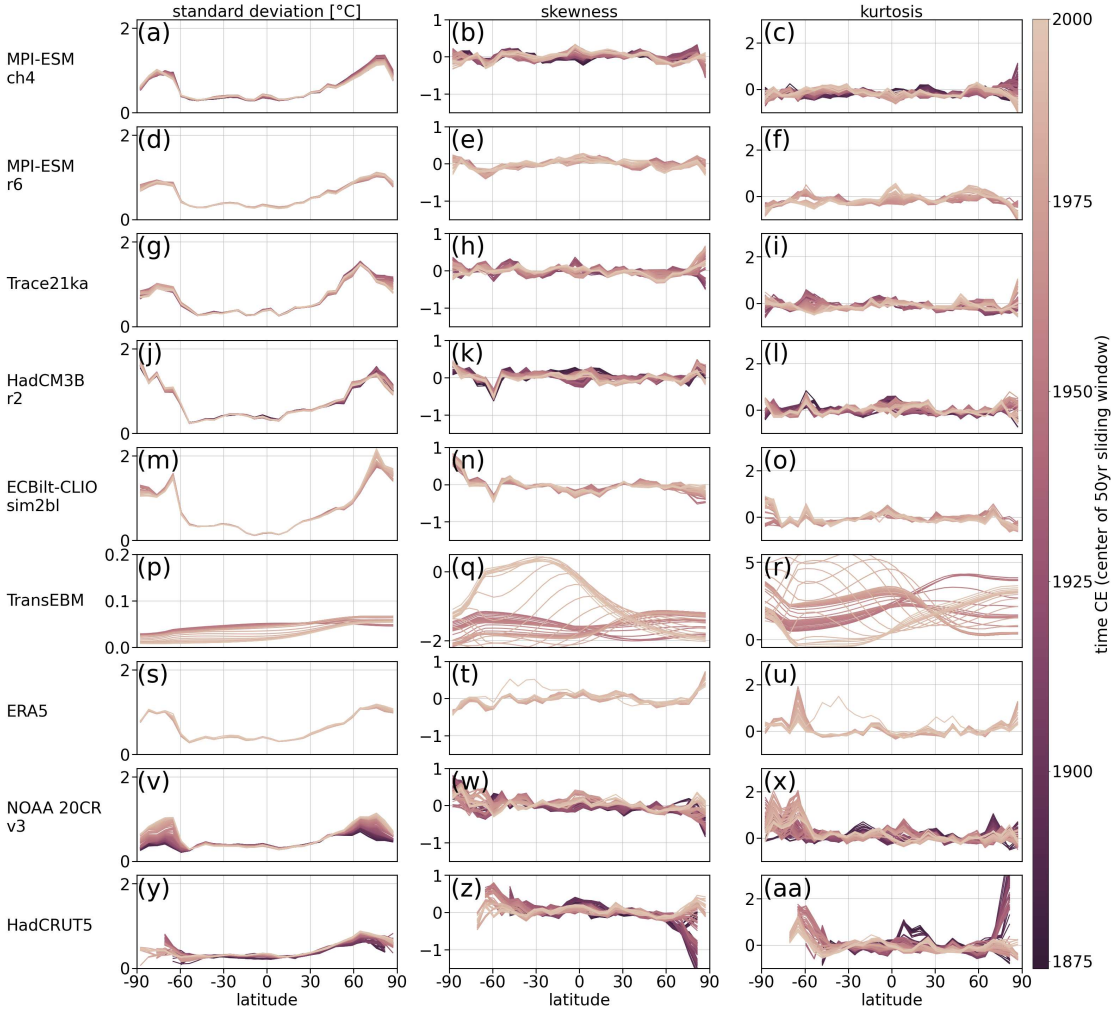


Figure 3.9: Comparison of annual moments of surface temperature in time and over latitude in simulations, reanalyses and an observational compilation. Standard deviation (left column), skewness (middle column) and kurtosis (right column) are computed for 50 yr sliding windows along longitude for the period 1850–2023. Note the different y-axis for the TransEBM run (panels p–r).

with time in the polar regions, likely due to larger sea ice variability as also found by Olonscheck et al. (2021), a trend that the models do not reproduce. ERA5 covers only the end of the period from 1950 onward, when it shows similar amplitudes of standard deviation as NOAA 20CR, whereas standard deviation in HadCRUT5 is smaller. Results for HadCRUT5 are, however, affected by the lack of coverage in the polar regions, especially in the 19th century. As standard deviation tends to be higher in this region (Sec. 3.1.1), HadCRUT5 results have to be interpreted with caution. Furthermore, we use the gridded ensemble mean of HadCRUT5 and do not investigate individual ensemble members. As we have seen for the LGMR (Sec. 3.2.1), variability can be subdued in an ensemble mean in comparison to the ensemble members. Some of the models (HadCM3B, ECBilt-CLIO) simulate larger standard deviation or a slightly different structure, as standard deviation in HadCM3B keeps increasing with latitude in the Southern Hemisphere. With the exception of the EBM, the latitudinal patterns of standard deviation in the models resembles those in the observational products in agreement with expectations from **P2** for annual, regional variability.

For the higher order moments of surface temperature, there are again some differences in the higher latitudes between simulations and observational products (Fig. 3.9). In lower and middle latitudes, on the other hand, skewness and kurtosis remain close to zero, suggesting that distributions are normal there. The EBM remains an exception for the higher order moments as well. This agrees with our findings for the Deglaciation (**P3**). As such, the remainder of the discussion will focus on the other simulations. The observational products suggest larger amplitudes in skewness and kurtosis in the northern high latitudes and larger kurtosis also in the southern high latitudes, none of which is matched by the models. HadCRUT5, in particular, also shows large changes over time in the Northern Hemisphere with skewness switching from negative to positive and kurtosis increasing between the late 19th and late 20th century. While the changes and moments in the Northern Hemisphere also differ between the reanalyses and HadCRUT5, they all agree that there is a peak in kurtosis in the southern polar regions that the models do not simulate. As a consequence, the models might lack in their representation of extreme conditions in the polar regions.

Differences in precipitation distributions over the instrumental era between simulations and reanalyses suggest possible weaknesses in the models in the tropical and polar regions (Fig. 3.10). Among the simulations, only HadCM3B matches the overall amplitude of standard deviation of precipitation in the tropics that the reanalyses consistently show throughout the period. However, instead of a single tropical peak near the equator, the latitudinal pattern in HadCM3B contains a double peak north and south of the equator. The EMIC ECBilt-CLIO misses the peak altogether confirming its weakness in simulating precipitation variability concluded for the Deglaciation (**P3**). For skewness and kurtosis, the reanalyses display two positive peaks north and south of the equator, the amplitudes and position of which vary slightly between them. However, the differences to and between the simulations are considerably larger. The EMIC lacks these regions of increased skewness and kurtosis completely. The GCM and ESM simulations produce other patterns than the reanalyses in at least one of the moments, with the MPI-ESM simulations notably only simulating a Northern Hemisphere peak in both skewness and kurtosis. The changes over time that the NOAA 20CR reanalysis suggests for skewness and kurtosis in the polar regions do not exist in the simulations. In the reanalysis, both skewness and kurtosis

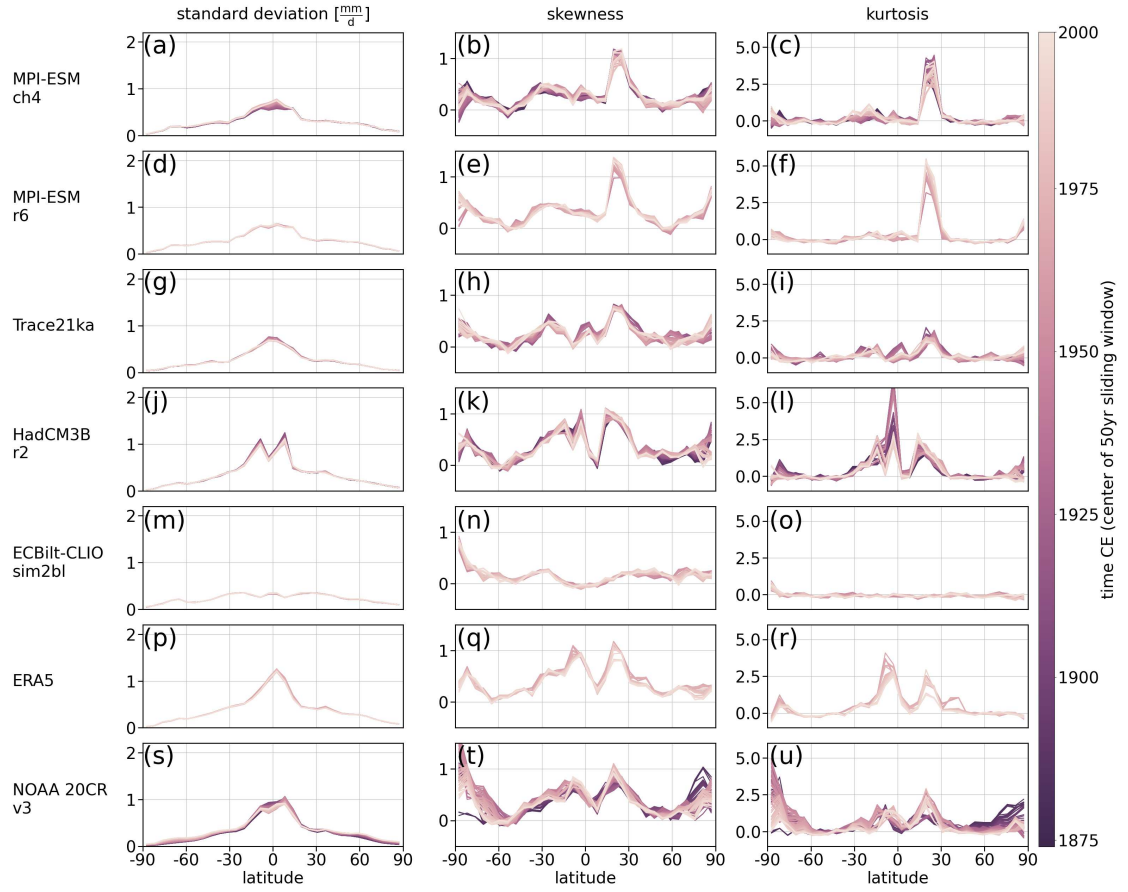


Figure 3.10: Comparison of annual moments of precipitation in time and over latitude in simulations to reanalyses. Standard deviation (left column), skewness (middle column) and kurtosis (right column) are computed for 50 yr sliding windows along longitude for the period 1850–2023.

decrease over time, getting closer to zero, indicating that distributions become more normal. The simulations, on the other hand, only show increased positive skewness, but not kurtosis, in the southern polar regions, but little change over time. Inter-model differences, too, are large, as no two models produce similar latitudinal patterns in either skewness or kurtosis.

Overall, the results for the instrumental era show the large variations among ESMs and GCMs that are also present for the Deglaciation (**P3**). While there are some differences between the reanalyses and HadCRUT5, disagreement with and between the simulations is larger. The simulations tend to compare better for standard deviation than for skewness and kurtosis, indicating potential weaknesses in simulating extremes of surface climate on an annual timescale. These weaknesses appear regionally and are especially strong in the polar regions.

3.2.3 Combining simulated and observed evidence to understand Earth System dynamics

For long-term transient simulations, comparisons of models and observations are challenging, especially with respect to variability. For example, simulations cannot be expected to match the timing of abrupt events in the observational record as internal variability affects their timing (PAGES Hydro2k Consortium, 2017). However, with the increasing amount of transient simulations included in paleoclimate modeling intercomparison projects, tackling the challenges in comparisons of models and observations is crucial. This requires approaches that consider uncertainties on the side of models and observations, which might need to be tailored to the specific applications and circumstances at hand.

As of now, most comparisons of paleoclimate simulations with observations rely on equilibrium simulations (e.g., Braconnot et al., 2012; Kageyama et al., 2021; Rehfeld et al., 2018). When studies consider transient simulations, they mostly focus on comparisons at individual proxy locations (e.g., He et al., 2021; Menviel et al., 2011; PAGES Hydro2k Consortium, 2017; Snoll et al., 2024) or with large-scale mean reconstructions (e.g., Dallmeyer et al., 2022; Mikolajewicz et al., 2025). These studies often consider only mean changes or investigate the most recent past, in particular the common era (**P2**). However, some studies now expand the toolkit for comparing model simulations with observations from various sources with data assimilation approaches or evaluation metrics customized for variability (e.g., Bühler et al., 2021; Choblet et al., 2024; Dolman et al., 2021; Kunz and Laepple, 2021; Weitzel et al., 2024).

Here, we use a gridded field reconstruction for the past 25 kyr created using data assimilation (Osman et al., 2021) as well as historical and modern reanalyses and observations for comparison. The comparison of the simulations to observations suggest some weaknesses in the models which merit further investigation, especially with respect to the representation of extreme conditions and the changes of distributions over time. To this end, a more in-depth comparison to reanalyses and observational records, including potentially individual station records, might prove fruitful to investigate the representation of variability on short timescales. However, these comparisons will remain limited by the short time-span covered by instrumental records. On longer timescales, the comparison to the LGMR highlighted weaknesses in current data assimilation approaches, which do not reflect the variability in the assimilated proxy records. As such,

either improvements in data assimilation methods or comparisons to proxy-based reconstructions will be necessary to evaluate variability in simulations on longer timescales further.

On the model-side, such comparisons are affected by uncertainties in boundary conditions and forcings, as well as in model parameterizations. Computational constraints for long-term transient simulations will further limit the possible resolution. Thus topographic effects and fine-scale processes are parameterized or missing, whereas proxies will record such local effects (PAGES Hydro2k Consortium, 2017). The limitation in resolution can be especially relevant for precipitation, which can be highly localized as is then reflected in observations from stations and proxies, something model grid boxes cannot adequately capture (Osborn and Hulme, 1997; PAGES Hydro2k Consortium, 2017). Focusing on large-scale dynamics or using regional models and downscaling methods can help avoid such issues (PAGES Hydro2k Consortium, 2017). However, downscaling with a regional climate model is not necessarily able to compensate biases in the circulation of the driving global model, meaning that biases in simulated variability can remain (Gómez-Navarro et al., 2015).

Simulations of past climate can further be biased as models are tuned to modern climate (Hourdin et al., 2017; Schmidt et al., 2017). For paleoclimate modeling, this adherence to tuning based solely on modern observations has been called in question (e.g., Gregoire et al., 2011; Hopcroft and Valdes, 2021; Hopcroft et al., 2021). Tuning, at least partially, to metrics of past climate has been suggested as an alternative. Using such an approach, Hopcroft and Valdes (2021) have shown that this can introduce abrupt behavior into a model that is absent if tuning is limited to present-day observations. Due to the limited amount of observational records available for past periods, the paleoclimate metrics and records to include in tuning have to be chosen with care. As long as the evaluation of the resulting model setup includes out-of-sample periods to avoid overtuning, expanding tuning metrics to include paleoclimate reconstructions can compensate for the short record of direct observations.

On the proxy-side, limited spatial and temporal coverage of proxies are a concern. Proxy locations are further biased by availability and accessibility of archives, marine sediments, for example, are biased towards coastal areas (e.g., Jonkers et al., 2020; Paul et al., 2021). These spatial biases will affect the reconstructed spatial structure of variability. Uncertainties regarding individual proxy records or proxy types can be reduced by using data from various proxy types and locations. Non-climatic noise can further affect all stages from the recording of the climate signal in an archive to the signal retrieval, for instance, measurement errors, as well as the calibration and dating will introduce uncertainties into the reconstruction. The effect of noise can further differ by timescale. For example, proxy variables in sedimentary archives integrate variability from short to longer timescales (PAGES Hydro2k Consortium, 2017). As a result, their spectra have more power at longer timescales and less at shorter ones in comparison to the hydroclimatic variables they capture in a so-called reddening of the spectra (Huybers et al., 2016; PAGES Hydro2k Consortium, 2017).

Improving our understanding of variability requires using simulations and observations as complementary resources by tackling these uncertainties. Simulation ensembles using multiple models or multiple forcing protocols can help address issues of uncertain parameterizations and boundary conditions. Here, we have shown, though, that not all models are made equally in

their representation of variability. Taking ensemble means, on the other hand, can be inadequate for analyzing variability, as ensemble members can, for example, have opposing trends (cf. Sec. 3.2.1, **P3**). Additionally, running large ensembles for multiple millennia can be computationally prohibitive. One approach to circumvent this issue is to build statistical models, so-called emulators, that emulate a climate model, but are computationally much more efficient. They have been shown to be successful in reproducing mean changes of surface temperature and precipitation globally (e.g., Castruccio et al., 2014) and regionally (e.g., Barboza et al., 2023). Schillinger et al. (2022) have further demonstrated that a Bayesian approach allows for emulating GMST variability, demonstrating that emulators can capture variability.

Furthermore, machine learning approaches have been implemented in weather forecasting, which are trained to reproduce weather patterns based on reanalysis or observational data and demonstrate skill comparable to traditional weather forecasting models (e.g., Bi et al., 2023; Lam et al., 2023; Rasp et al., 2024). Kochkov et al. (2024) have recently shown that a machine learning approach can also reproduce many features of climate for multiple decades, including tropical cyclones. While variability has not been evaluated, the abilities of such methods look promising, at least on short timescales, and they could offer a computationally efficient alternative to GCMs and ESMs. However, they rely on sufficient training data. For the historical period, these exist in the form of observational datasets, reanalysis and even large collections of climate model simulations and ensembles. Further back in time, data becomes increasingly sparse. For mean changes, ensembles from low and intermediate complexity models can provide a training set, but as we have shown such models are insufficient in their representations of variability (**P3**, Sec. 3.1.4). An evaluation of variability of the climates generated by machine learning models, especially on decadal and longer timescales, could provide clarity on their potential for paleoclimate modeling. If such approaches prove inadequate, machine learning methods might instead be used to improve boundary conditions or parameterizations (O'Gorman and Dwyer, 2018) and then apply those in a GCM or ESM paleoclimate simulation.

Synthesizing individual proxy records can improve comparisons of simulations with reconstructions, for example through data assimilation methods. As data assimilation produces gridded field reconstructions using a climate model constrained by proxy reconstructions, such as the LGMR (Osman et al., 2021), it simplifies the comparison to simulations. However, the computational demand for data assimilation is high. Paleoclimate research thus often uses offline methods or intermediate complexity models (PAGES Hydro2k Consortium, 2017). As shown here (**P3**, Sec. 3.2.1) and in Choblet et al. (2024), spatial smoothing can be an issue for variability in the assimilated product. Non-linear methods that do not require normality of the state vectors could remedy this and are a current focus in data assimilation research (Evensen et al., 2022). For paleoclimate research, Choblet et al. (2024) have demonstrated that data assimilation can be adapted to account for the differing time resolutions of various proxy records and be used on the measured water isotopes directly. Generally, comparing the measured proxy variables directly instead of relying on reconstructed variables like temperature or precipitation can eliminate error sources. Water isotopes, in particular, are measured across many major archives (PAGES Hydro2k Consortium, 2017) and have been integrated in several climate models (e.g., Hoffmann et al., 1998; Risi et al., 2010; Roche, 2013; Tindall et al., 2009; Werner et al., 2011). While running isotope-enabled simulations increases computational demand, removing

the transfer step from measured to climate variable benefits model-observation comparisons, including of variability (Bühler et al., 2022, 2021).

Any comparison of simulations and reconstructions needs to account for sources of non-climatic noise as well as uncertainties in resolution and dating affecting proxy records. To this end, proxy-system models (PSMs) provide a mathematical translation between the measured proxy quantity and the climate variable of interest that mimics the biological, chemical and physical processes affecting the proxy (Evans et al., 2013; PAGES Hydro2k Consortium, 2017; Weitzel et al., 2024). A PSM can be run forward to create pseudo-proxies from simulated climate variables by applying the effects of how the climate variable imprints on a climate archive. It can also be used as an inverse model to remove non-climatic effects and deduce the likely original signal. PSMs are crucial when analyzing variability as noise can alter the signal differently across timescales (Bühler et al., 2021; PAGES Hydro2k Consortium, 2017). They also allow for a better synthesis of records, for example to compute spatial and temporal patterns of variability (Laepple and Huybers, 2014b).

3.3 Surface climate variability under low to high future emission scenarios

The Last Deglaciation is the most recent period of substantial global warming with an increase in GMST by about 4–7°C (Annan et al., 2022; Gulev et al., 2021; Osman et al., 2021; Tierney et al., 2020). Depending on the eventual emissions pathway, expected future warming has the potential to be of a similar magnitude to the warming of the Deglaciation. Considering the colder initial state and the differences in warming speed, how similar changes in the climate system during the Deglaciation are to future projected changes is unclear. To investigate this, we present an analysis of variability changes in future projections not included in **P1–P3**.

The analysis focuses on simulations that are extensions of the transient Deglaciation simulation MPI-ESM ch4 under three SSP scenarios — SSP 1-1.9, 2-4.5 and 5-8.5. Additionally, we analyze all CMIP6 simulations of those three SSP scenarios that run at least until 2300 CE combined with their historical counterparts. The analysis period is thus 1850 – 2300 CE. The required simulation length limits the number of available simulations, as so far most CMIP6 scenario simulations have been run only until 2100 CE. Only MPI-ESM ch4 and GISS E2-1G have a simulation available for each scenario, which is why we concentrate on them here and present results for the other CMIP6 simulations in the appendix (Sec. A2). As in **P3**, we detrend the annual timeseries of surface temperature and precipitation to assure weak stationarity, using a kernel length of 100 yr for the Gaussian filter. We then compute the moments of the distributions over 100 yr sliding windows.

The standard deviation of surface temperature decreases with warming, especially towards higher latitudes, a finding that is consistent for the Deglaciation (**P3**) and the SSPs (Fig. 3.11, S17–S19). This agrees with findings in the literature that report such a decrease with warming from monthly to multi-decadal timescales (e.g., Bathiany et al., 2018; Boer, 2009; Olonscheck et al., 2021; Rehfeld et al., 2020). It has been linked to the reduction of sea ice (Collow et al., 2019;

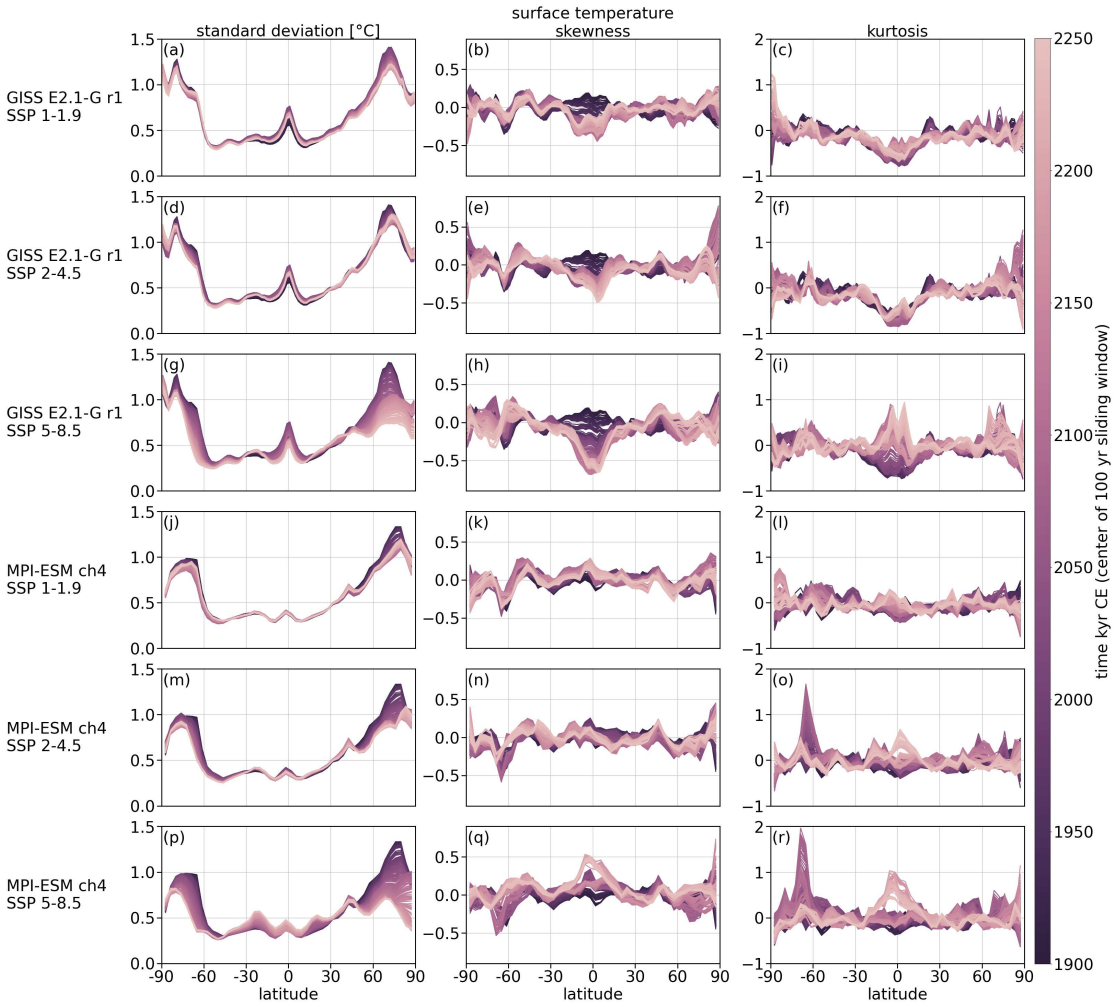


Figure 3.11: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of surface temperature across latitude for combined historical and SSP runs. Shown are the results for (a)–(i) GISS E2-1G r1 from the CMIP6 ensemble and (j)–(r) MPI-ESM ch4. Results for the other simulations from the CMIP6 ensemble can be found in the appendix (Fig. S17–S19). Moments are computed over 100 yr long rolling windows from 1850 – 2300 CE. Time is given as the middle of that rolling window.

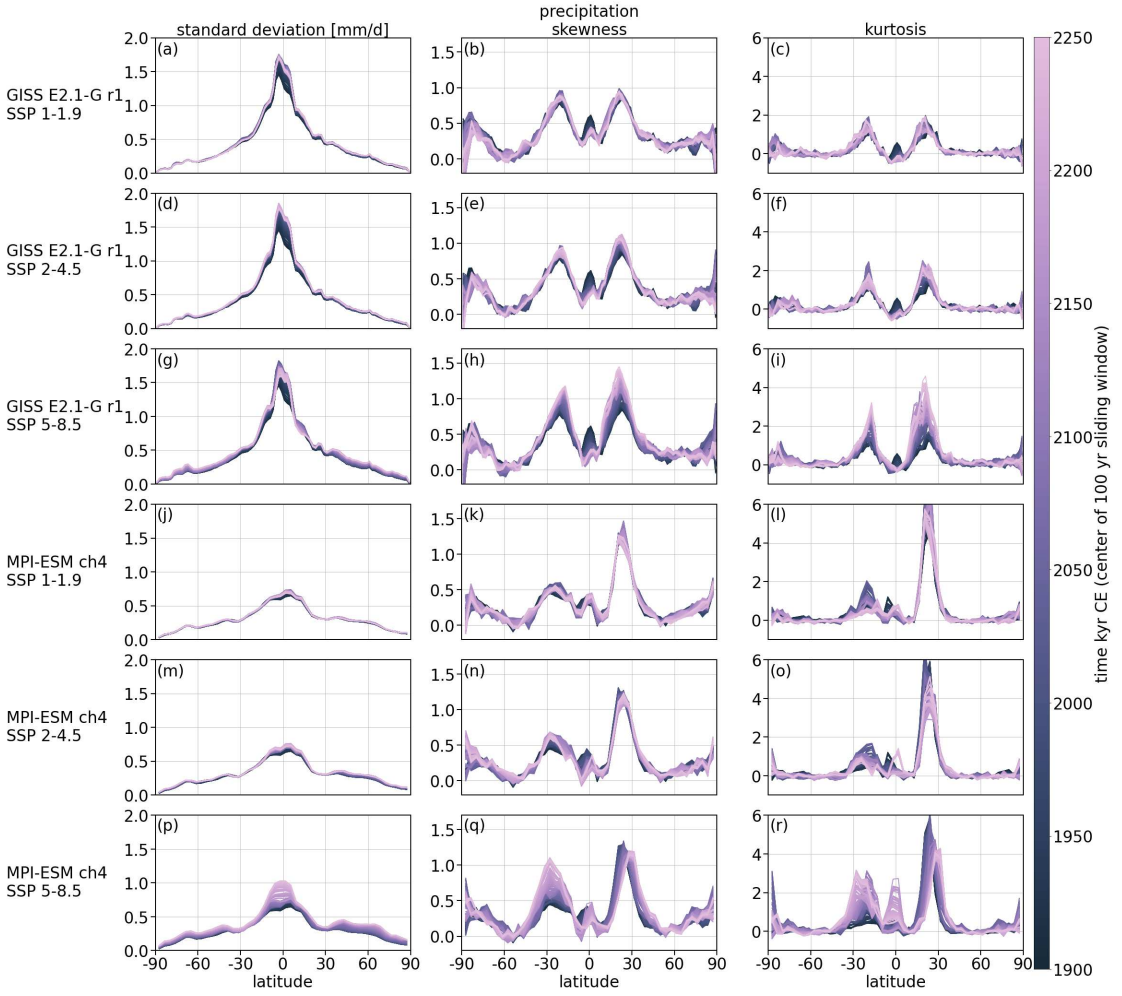


Figure 3.12: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of precipitation across latitude for combined historical and SSP runs. Shown are the results for (a)–(i) GISS E2-1G r1 from the CMIP6 ensemble and (j)–(r) MPI-ESM ch4. Results for the other simulations from the CMIP6 ensemble can be found in the appendix (Fig. S20–S22). Moments are computed over 100 yr long rolling windows from 1850 – 2300 CE. Time is given as the middle of that rolling window.

Dai and Deng, 2021; Huntingford et al., 2013). As in Olonscheck et al. (2021), some simulations show a small initial increase in standard deviation while the sea ice border moves and then a decrease after sea ice has disappeared. For SSP 5-8.5 the decrease is particularly large and can exceed expectations based extrapolation from the lower emission scenarios or the Deglaciation (Fig. 3.13, S23). This is particularly apparent for MPI-ESM ch4, as the relationship between mean and moments in SSP 5-8.5 often contains breaks (Fig. 3.13, 3.14). At these points, the relationship between mean and the other moments suddenly changes and becomes categorically different. For example, while the standard deviation of temperature in MPI-ESM ch4 in the Antarctic regions initially decreases with warming under the SSP scenarios, for large warming it flattens and does not further decrease. In this case, this change is likely related to the disappearance of sea ice at those latitudes.

For precipitation, the Deglaciation and future simulations show opposing trends in standard deviation in the tropics, where overall standard deviation is largest: In the Deglacial simulations, standard deviation of precipitation decreases with warming (Fig. 3.2), whereas it increases in the SSP runs (Fig. 3.12). For some CMIP6 simulations, it starts decreasing towards the end of the SSP 5-8.5 simulations (Fig. S22), but not MPI-ESM ch4 (Fig. 3.2). In the extratropics, on the other hand, there is a trend towards increased standard deviation for both Deglacial and future simulations.

For skewness and kurtosis, we find large inter-model differences (Fig. 3.13, 3.14, S17–S22). For surface temperature, for example, there are large changes in the tropics, which grow larger for higher emission scenarios. However, models disagree on the direction of change, as well as the degree and temporal progression of it, both in the Deglacial and future simulations. For precipitation, SSP 5-8.5 again stands out in comparison to the lower emission scenarios and the Deglaciation. Only some changes in kurtosis in the tropics during the early Deglaciation are similarly large to those in SSP 5-8.5. Otherwise, Deglacial changes in the higher moments of precipitation tend to be more similar to those of SSP 2-4.5.

At times, we find large excursions in skewness and kurtosis with significant non-normality while the radiative forcing is still changing until about 2100 CE (cf. Fig.1.1) and a couple of decades after it stabilizes (Fig. 3.13 and S17–S22). Such disruptions can be found for instance for temperature skewness in the polar regions of the Northern Hemisphere (e.g., Fig. 3.13e) and for kurtosis around 70°S (e.g., Fig. 3.13o, r). For MPI-ESM ch4, we additionally find distinct states in the higher order moments and rather abrupt transitions between them, especially for SSP 5-8.5 (e.g., Fig. 3.13q & r). As the emissions increase and the climate warms, simulated sea ice cover shrinks and more and more areas become ice free, potentially year-round. This will lead to abrupt changes in the local climate of these regions that affect the distributions of surface temperature and precipitation.

The disturbances in the higher order moments appear while emissions are increasing and thus seem to occur through the response of fast-acting components and feedbacks of the climate system. Once the emissions stop, the moments tend to stabilize. This indicates that even in the high emission scenario, stopping the emissions shows effect quickly, lessening the impacts on the tails of the distributions. This potentially implies that magnitude and frequency of extreme events also stabilizes, something that an analysis of extremes could clarify. Additional analysis

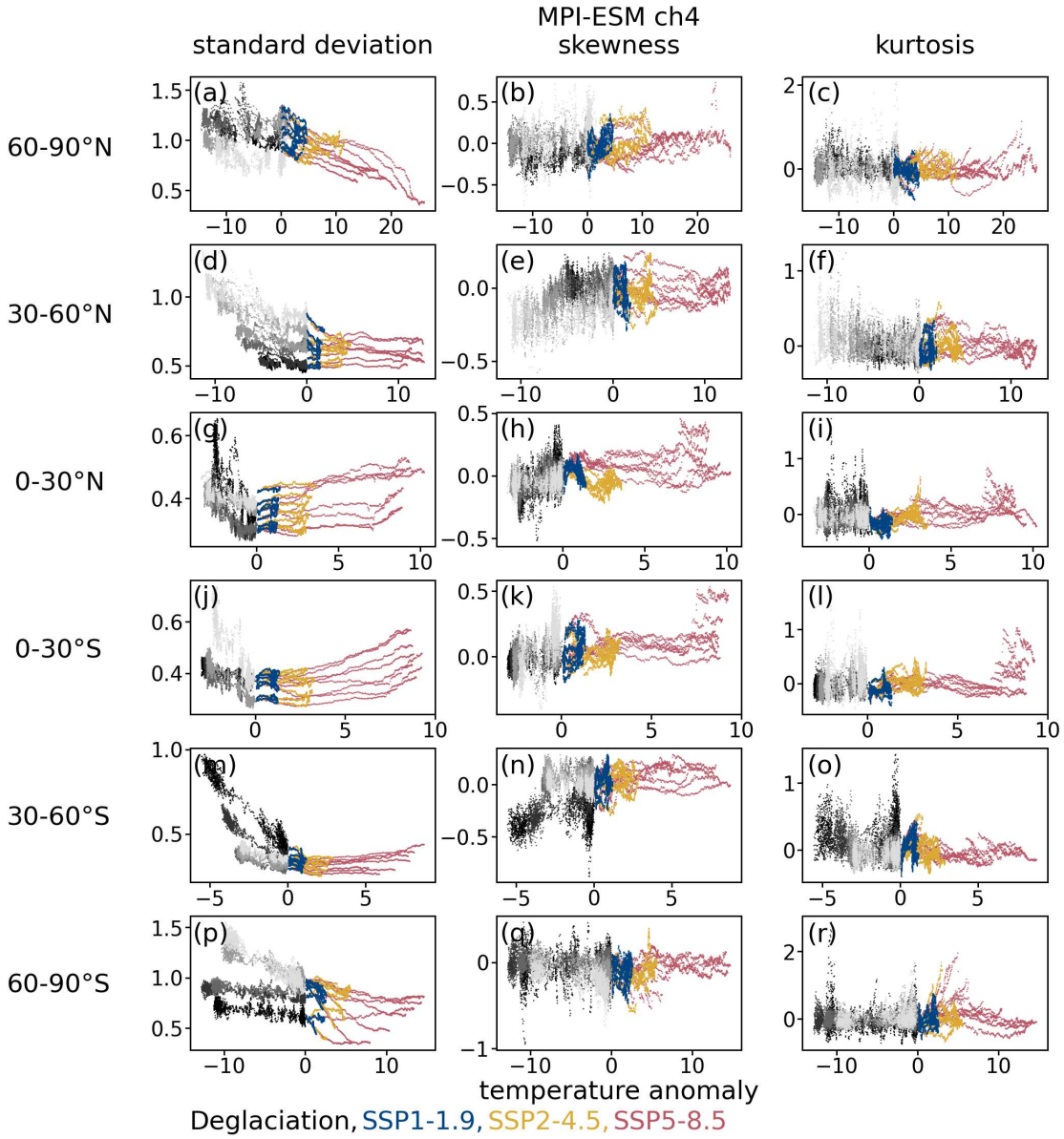


Figure 3.13: Moments (y-axis) of surface temperature across each latitude against the mean anomaly with respect to 1850 − 1899 CE (x-axis) for MPI-ESM ch4. Shown are the Deglaciation, SSP 1-1.9, SSP 2-4.5 and SSP 5-8.5. For the Deglaciation lighter shades of gray are further north within the range of each panel.

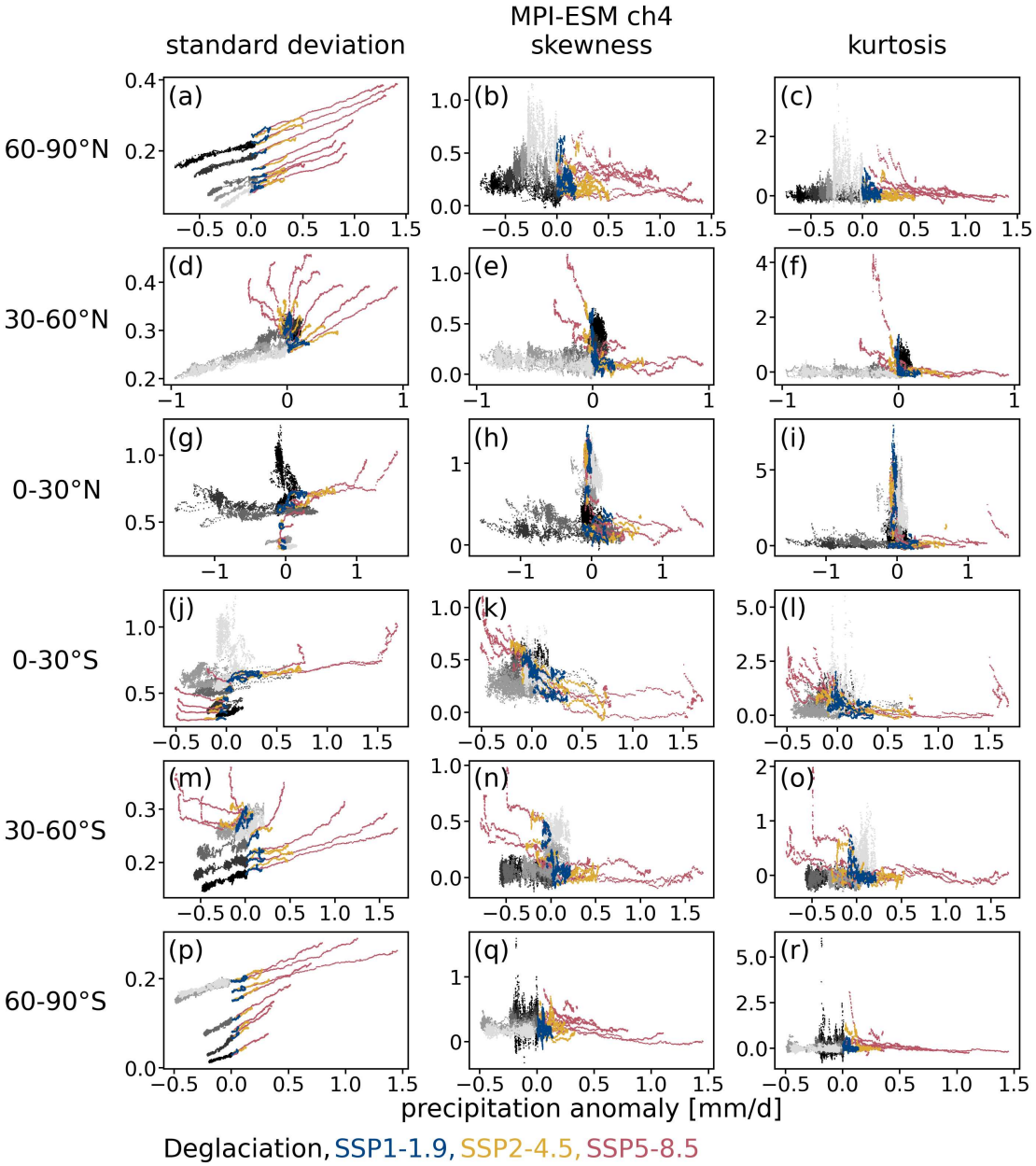


Figure 3.14: Moments (y-axis) of precipitation across each latitude against the mean anomaly with respect to 1850−1899 CE (x-axis) for MPI-ESM ch4. Shown are the Deglaciation, SSP 1-1.9, SSP 2-4.5 and SSP 5-8.5. For the Deglaciation lighter shades of gray are further north within the range of each panel.

should hone in on differences before and after forcing stabilizes, for example by considering idealized forcing scenarios. Such idealized simulations could switch several times between an increasing and constant GHG forcing, while also varying the strength of forcing, the length of its application and the down time in-between. This would allow to investigate whether the excursions in higher order moments are indeed a property of a climate system in short-term or long-term disequilibrium. It could further demonstrate how soon after the onset of increase in emissions, the excursions appear and how soon they disappear once the emissions stop. The relation to the strength of the forcing and the speed of its increase could further be examined.

Our results already show that the speed of forcing matters for the amplitude and frequency of response of the climate system. While there are commonalities between the three scenarios, the moments are also often different at the same warming level (e.g., Fig. 3.13, S23). Sensitivity experiments, which reach the same warming levels but at different speeds, could clarify this link between warming speed and variability. Such experiments could also be used to investigate the relative role of speed and overall magnitude of warming. In our analysis, the low and middle emission scenario mostly show similar trends to the Deglacial simulations. SSP 5-8.5, on the other hand, seemingly exhibits changes in variability without equivalent in the Deglaciation or lower emission scenarios. This suggests that parts of the climate system might enter a different territory at such high emissions, but how much this is driven by the speed versus the magnitude of forcing is, at this point, unclear.

When comparing how moments of the Deglacial and SSP simulations change with the mean, the larger spread in moments at any warming level during the Deglaciation is apparent (Fig. S23, S24). Due to the lower speed of changes, the influence of internal variability is larger during the Deglaciation, as the slower components of the climate system can adapt to the changes. In contrast, the slower components of the climate system lag behind in the future projections and much of their response can be expected to unfold over the centuries and millennia to come, as global mean sea level for example will continue to rise even under SSP 1-1.9 (IPCC, 2021c). Based on the results for the Deglaciation, the moments can likely be expected to display a spread in the future projections once the forcing stabilizes.

We find large inter-model differences for skewness and kurtosis in the SSP scenarios, which warrant further investigation. In particular, larger ensembles for SSP 1-1.9 and SSP 2-4.5 or adding other scenarios could strengthen confidence in these results. Additionally, the continuation of more or new members of the Deglacial ensemble into the future would enable the analysis of the connection between past and future warming. Due to the large inter-model spread, differences between MPI-ESM ch4 and the SSP scenarios cannot, with confidence, be attributed to long-term memory effects present in MPI-ESM ch4, something which additional ensemble members or comparison to MPI-ESM SSP runs not initialized from the Deglacial simulation could remedy in the future.

This analysis provides first insights into future climate variability based on past evidence. However, it is limited by the small number of simulations available across scenarios, both as extensions of the Deglacial ensemble and running until 2300 CE in CMIP6. Specifically, an investigation of memory effects in slow components of the climate system and the influence of the initial state would be of interest. This could be achieved by comparing projections extended

from the transient Deglacial ensemble members to historical and scenario simulations from the same models initialized at 1850 CE. For the historical period, these simulations could then be evaluated against observations to investigate whether there are systematic improvements in simulated variability depending on the initial conditions. Considering high-resolution data, could provide further confidence in these comparisons. This would provide an evaluation of models' skills in simulating variability and could inform future model development. Both are vital to better constrain future climate variability and its changes with magnitude and speed of warming across timescales.

4 Conclusion

Climate variability directly impacts people's everyday lives: It affects the frequency and severity of extremes (**P2**, Ionita et al., 2021; Katz and Brown, 1992; Schär et al., 2004; Simolo and Corti, 2022) and amplifies impacts on societies (Anderson et al., 2019; Hsiang et al., 2013; O'Loughlin et al., 2014) and ecosystems (Seddon et al., 2016; Stenseth et al., 2002). Furthermore, projections (Deser et al., 2012; Hawkins and Sutton, 2009), as well as attribution (Harrington et al., 2021; Hegerl and Zwiers, 2011; Van Oldenborgh et al., 2022) and impact studies (Calel et al., 2020; Schwarzwald and Lenssen, 2022) require the reliable simulation of variability. As such, reported uncertainties in variability in studies of past, present and future climate are concerning (**P3**, Alexander and Perkins, 2013).

Various studies suggest that surface climate variability depends on timescale as well as the background climate state (e.g., Brown et al., 2017; Rehfeld et al., 2018; Shao and Ditlevsen, 2016). However, variability has mainly been investigated in quasi-equilibrium climate states, mostly during the Holocene (**P2**), but rarely during periods of warming on longer than interannual timescales. Here, we addressed these gaps in three publications (**P1**–**P3**) and additional analyses (**AA**) by investigating how spatio-temporal patterns of surface climate variability depend on and change with the mean climate from the LGM to possible future emission scenarios. To this end, we examined the changes of surface climate variability in ensembles of climate model simulations and compared them to observational records. Following our research objectives (**O1**–**O5**), we analyzed how surface climate variability depends on temporal and spatial scales (**O1**) and the background climate state (**O2**) in models and observations. In the simulations, we analyzed the effects of forcings and boundary conditions (**O3**), as well as the complexity of climate models on variability (**O4**). Lastly, we evaluated how surface climate variability changes with past and future warming (**O5**).

The results show

1. **Spatio-temporal patterns of surface climate variability**
   Indicators of surface climate variability depend on the temporal and spatial scale in question (**P2**, **P3**, **AA**). The standard deviation decreases between annual and centennial timescales for both surface temperature and precipitation (**P3**). For temperature, it is the largest at high latitudes and for precipitation in the tropics. For temperature, the skewness and tails of the distribution become more pronounced from annual to centennial scales. For precipitation, on the other hand, both skewness and kurtosis decrease towards longer timescales, but the resulting shape of the distribution depends on the background state.
   Observations and simulations of surface temperature variability agree globally and regionally on annual to interannual timescales (**P2**, **P3**). On decadal and longer timescales, however, regional variability diverges as observed variability is higher than simulated variabi-

lity. Since uncertainties in reconstructions cannot on their own explain these differences, models likely lack regional variability on these timescales (**P2**). Regional climate fluctuations on annual to interannual scales likely persist on longer timescales, which models currently fail to capture.

2. **State-dependency of surface climate variability**
   Surface climate variability depends on the background state and changes with it (**P3**, **AA**). This state-dependency increases from annual to millennial scales for surface temperature and from centennial scales for precipitation. LGM temperature variability is larger than Holocene variability, while the opposite holds for precipitation. However, the simulated temperature variance ratio between LGM and Holocene is smaller than that in reconstructions (**P3**), in agreement with results from equilibrium simulations (Rehfeld et al., 2018; Shi et al., 2022). In contrast to these quasi-equilibrium states, variability tends to be even larger throughout transitionary periods like the Last Deglaciation (**P3**). These patterns depend on the period and timescale. For example, annual temperature standard deviation generally decreases with warming from one century to the next.
   There are large inter-model differences for the higher order moments in all mean states (**P3**, **AA**). As these moments describe the tails of the distributions, disagreement between models casts doubt on their ability to simulate extreme behavior. At times, skewness and kurtosis change abruptly, especially during the early Deglaciation and early Holocene. As such, variability does not change linearly with the background state. Instead, abrupt shifts and changes in scaling are possible as the climate seems to enter new regimes.

3. **Effects of forcings and boundary conditions on simulated surface climate variability**
   The impact of ice sheet reconstructions, meltwater protocol and volcanism is usually stronger on the simulated variability of surface temperature than precipitation (**P3**). On the forcings' characteristic timescales their impact is the largest. However, they also influence indicators of variability on all other timescales investigated here, from annual to millennial. As such, these forcings impact variability on timescales both shorter and longer than their characteristic timescales as they alter slower and faster components of the climate system.
   Among the ice sheet reconstructions, we mostly find higher variability the higher and larger in extent the ice cover is (**P3**). Testing different meltwater protocols demonstrates that river-routing is associated with higher and more realistic patterns of variability than the globally homogeneous distribution of meltwater. The exclusion of meltwater results in a notable lack of variability. The inclusion of volcanic forcing generally increases variability and narrows the gap between simulated and observed regional variability on decadal and longer timescales.

4. **Influence of model complexity on simulated surface climate variability**
   The complexity of GCMs is at least necessary for an adequate representation of variability in models (**P3**). Beyond that, inter-model differences surpass differences in complexity in our Deglacial simulation ensemble. Within the computational limitations imposed on transient simulations that are several ten thousand years long, model resolution also plays no significant role in the simulation of surface climate variability. As such, forcings and

boundary conditions are likely to play a larger role for variability in state-of-the-art models.

5. **Surface climate variability under past and future warming**
   Standard deviation changes are similar in the past and future, although the dependence on the mean can change for high emission scenarios (**AA**). For the higher order moments, on the other hand, there is a large spread between simulations, which thus likely translates into a spread in simulated extremes. As the connection between the moments and extremes is intricate, this merits an explicit investigation comparing changes in temperature and precipitation extremes under past and future warming.
   Our results suggest that variability depends not only on the magnitude but also the speed of warming (**P3**, **AA**). Both contribute to the increased variability we observe during warming episodes in contrast to quasi-equilibrium states, which can appear in abrupt excursions. Once the forcing and the resulting warming stabilize, variability decreases.

This extensive study of surface climate variability extends the understanding of variability to consider the whole distributions of climatic variables. We demonstrate how variability changes from past to present and potential future climate states on annual to millennial timescales. The results highlight the importance of transient, long-term simulations with river-routed meltwater and volcanic forcing for an adequate representation of surface climate variability in models. The impact of meltwater protocol and ice sheet reconstruction further suggests that ice sheet models coupled to AOVGCMs might improve simulated variability and narrow the identified gaps between observed and simulated variability.

Nevertheless, a more in-depth comparison of simulated and observed variability including higher order moments is necessary to further hone in on model weaknesses and improve models. In particular, variables besides surface temperature need more attention to determine whether weaknesses in simulated temperature variability are the exception or the rule. However, the spatial and temporal coverage of proxy records is limited and the calibration of measured proxy properties to climatic variables can be challenging. In this respect, isotope-enabled modeling could prove particularly useful as it allows the simulation of the properties measured in proxies and thus a direct comparison, especially when used in conjunction with proxy system models. For paleoclimate, common data assimilation methods fail to preserve the variability found in the underlying proxy-based reconstructions, possibly due to their sparse coverage and assumptions of normality. Current developments towards non-linear data assimilation approaches could help retain reconstructed variability in assimilated products. Such improved comparisons of simulated and observed variability would help evaluating the relative skill of GCMs and ESMs, for which we found large inter-model differences.

A comprehensive evaluation of simulated variability could then further substantiate the analysis of projected variability presented here, as it would allow identifying models skilled at simulating variability. Currently, an extension to future scenarios is available for only one of the simulations in the Deglacial ensemble, such that our findings could be model specific. As such further simulations are needed to corroborate our results, such as extensions of Deglacial ensemble members into the future and sensitivity experiments that focus on the speed and magnitude of the forcing. Using our extensive analysis of past variability, it would then be possible to verify

to which extent past parallels future variability and to constrain the conditions under which variability enters new territory.

Acknowledgements

I would like to express my sincere gratitude to everyone — colleagues, friends, family. I am deeply grateful that I got to share these past years with you. A special thank you to all of my colleagues and collaborators, whom I had the pleasure to meet, exchange ideas and work with, be it online or in person, at conferences, over dinner or on a mountain in the Allgäu. Thank you for many stimulating and inspiring discussions.

In particular, I would like to thank the STACY/SPACY group, my scientific home of the past years and a few more to come — and what a fantastic home it is. Thank you to all current and past members for all your support and camaraderie and hard work. I will never forget the times we shared, deep in scientific (or not so scientific) discussions, scaling mountains and caves, giving paragliding a try or sharing meals, from the infamous Eblypfanne to what has by now become the stuff of legend, the Schwarzwaldbecher. Of course, I also have to mention our wild goose chases throughout the forests in and around Tübingen and Heidelberg, following sets of weird clues. Or in my case, quite often setting them up. It has been a wild and fun ride with an incredible set of people and I am sure it will continue to be so.

None of this would have been possible without the never-ending support of my supervisor, Kira Rehfeld. You encouraged me when I needed encouraging, challenged me when I needed challenging and have always looked out for me, making me the scientist that I am today. You are a source of constant inspiration to me and I continue to be glad that I was send your way when you started building this group, thank you for having me be a part of it. I am looking forward to working on many more projects with you.

Among the current and former SPACY members, I would like to particularly thank Bea and Janica who accompanied me for most of my PhD journey and who it is wonderful to watch grow in their lives after SPACY. Janica, thank you for being someone I could always turn to, figuratively and, quite often, literally. From the big to the smallest things, our conversations were always enriching and your support was invaluable. Bea, thank you for always finding the right words to cheer me up when necessary. Working with you is always a joy and I am grateful for the time we spend growing alongside each other. Nils, you have been such a core part of my SPACY experience over the last years that it is weird to think that you will be moving on to Bristol soon. None of my work would be the same without your input. Thanks for being a guiding light, in science and in the many times spent roaming Swabia. Markus, to me, you are the beating heart of GUZ. Thank you for your warmth and your unparalleled ability to instill calm and positivity in me. Mo, thanks for always being someone I can confide in and sharing your worries in return. I hope there are only good things waiting in your future and I am sure we will keep expanding our whiteboard of seemingly endless projects and ideas. Valdir, thank you for exploring the Tübingen food scene, the Swabian alps and many a cave with me. All the best for your life as a professor in Brasília, let's meet again soon. Jean-Philippe, thanks for traversing the Palmod universe alongside me and for wrestling with our computing infrastructure with me and the rest of TeamTech and thanks for sharing your passion, weather, with us. Yana, let's visit the ducks again soon or enjoy some figure skating. Muriel, glad to have you with us again, I am looking forward to seeing all the adventures you keep getting up to. Everyone else, you have

made this experience what it was, thank you for that, and to all current and future members, I am looking forward to keep working and spending time with you!

Thanks to everyone who provided feedback on this thesis, Bea, Kira, Markus and Nils, it is better for your input. I would also like to thank my second reviewer, Reinhard Drews, and the remaining members of my examination committee, Monika Fleischer and Paul Bons, for taking the time to evaluate my work. Thank you also to Christine Köhler for keeping me sane.

My deepest gratitude goes to my friends and family, especially, Ricardo, Bruno, Tini, Florian and my parents. Ricardo and Bruno, you have truly been with me through thick and thin. Whatever comes, I know that I can count on you to have my back, just like I will have yours. Tini and Flori, thank you for accompanying me for almost my whole life, for celebrating my highs with me and carrying me through my lows. And to the people who are no longer with us to see this day: I miss you.

Thank you all!

Appendix

A1 Comparison of surface climate variability on annual versus centennial timescales

To investigate the differences with respect to timescale in the moments discussed in Sec. 3.1.1 and whether distributions become more or less normal with increasing timescale, we compute the differences in absolute skewness and kurtosis. For both surface temperature (Fig. S1 & S2) and precipitation (Fig. S3 & S4), the figures show the differences per time period — LGM, Deglaciation and Holocene — for a subset of the Deglacial ensemble. The differences between the annual and centennial moments are often small. With the exception of the erroneous tropical Holocene signal in MPI-ESM (c.f. 3.1.2), large differences mostly occur during the Deglaciation, when the centennial skewness and kurtosis tend to be larger for both surface temperature and precipitation. For surface temperature, the inclusion of volcanic forcing in MPI-ESM r7 leads to larger annual moments, especially in the tropics and for LGM and Holocene (Fig. S1 & S2, panels j & l).

A2 Changes in annual moments of surface climate with time from the Last Glacial Maximum to future scenarios

Temporal changes in moments in the Deglacial ensemble

As part of the analysis of how the moments depend on different timescales in Sec. 3.1.1 and in addition to the analysis from **P3**, we analyze the change in the annual moments over 100 yr rolling windows from the LGM to present-day. Here, we present the results from the main text for surface temperature (Fig. 3.1) and precipitation (Fig. 3.2) split into LGM (Fig. S5 & S8), Deglaciation (Fig. S6 & S9) and Holocene (Fig. S7 & S10) so as to highlight the differences between the periods. As discussed in Sec. 3.1.1, these show that abrupt shifts and large changes in moments mostly happen during the Deglaciation, while changes are more gradual during the LGM and Holocene.

Additionally, we examine state-dependency of the moments by calculating the LGM-to-Holocene ratios of the moments for annual, decadal and centennial timescales (c.f. Sec. 3.1.2). Standard deviation of surface temperature tends to be larger during the LGM across timescales, although there are exceptions for some simulation setups like MPI-ESM r5 and TraCE-21ka (Fig. S11). For precipitation, standard deviation is larger during the Holocene almost everywhere, with similar patterns across simulations (Fig. S14). For both temperature and precipitation, the skewness and kurtosis ratios exhibit seemingly random distributions with almost no clear patterns (Fig. S12, S13, S15 & S16).

Temporal changes of moments in low to high emission future scenarios

In Sec. 3.3, we analyze variability in future projections under three SSP scenarios, 1-1.9, 2-4.5 and 5-8.5. There, we present results for MPI-ESM ch4 and GISS E2.1-G r1. MPI-ESM

ch4 represents an extended member of the Deglacial ensemble, whereas the GISS simulation is a CMIP6 ensemble member, for which all three scenarios are available. Here, we add the other CMIP6 ensemble members. If several runs are available for a scenario for one model configuration, such as the ten SSP 2-4.5 runs available for GISS E2.1-G, we show results for run 1 only. Figures S17, S18 and S19 show the annual moments over 100 yr sliding windows for SSP 1-1.9, 2-4.5 and 5-8.5, respectively, with changes becoming larger and more erratic the higher the emission scenario. Figures S20, S21 and Fig. S22 display the same but for precipitation.

For the comparison between past and future warming, we compare the changes in standard deviation, skewness and kurtosis of surface temperature (Fig. S23) and precipitation (Fig. S24) with their mean in the Deglacial ensemble with the CMIP6 ensemble. Note that for the Deglaciation individual ensemble members are shown, whereas for the CMIP6 ensemble, it is the ensemble mean. However, taking the ensemble mean of the CMIP6 simulations is not responsible for the relative smoothness of the moment to mean relationship exhibited in comparison to the Deglacial simulations. Instead, it is a result of the stronger and faster forcing that suppresses much of the internal variability. For the Deglacial simulations, on the other hand, the mean state changes more slowly. As a result, the simulations sample more of the internal variability of the climate system resulting in a larger range of the moments in comparison to the mean. Figures 3.13 and 3.14 in Sec. 3.3 show the relation between moments and mean but only for MPI-ESM ch4 and similarly exhibit a larger spread in moments during the Deglaciation.

A3 Mean fields of surface temperature in the LGM Reanalysis

Complementing the discussion of the LGMR ensemble in Sec. 3.2.1, Fig. S25 presents the mean fields of the randomly chosen ensemble members. It shows that the mean fields are very similar across ensemble members with only slight differences in the warming patterns.

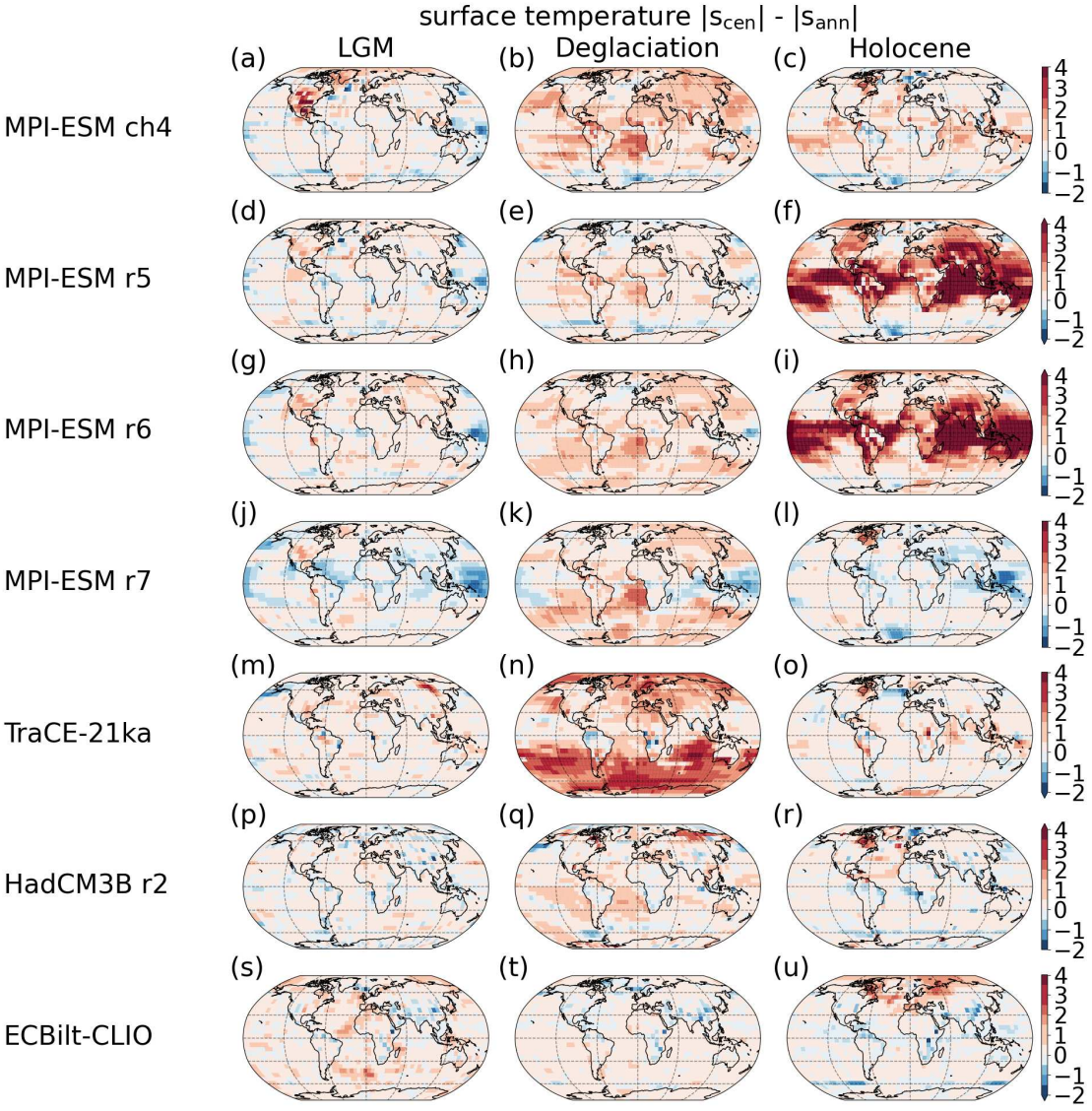


Figure S1: Difference of absolute centennial to annual surface temperature skewness for LGM, Deglaciation and Holocene. Red tones indicate that centennial skewness is larger, blue ones the opposite.

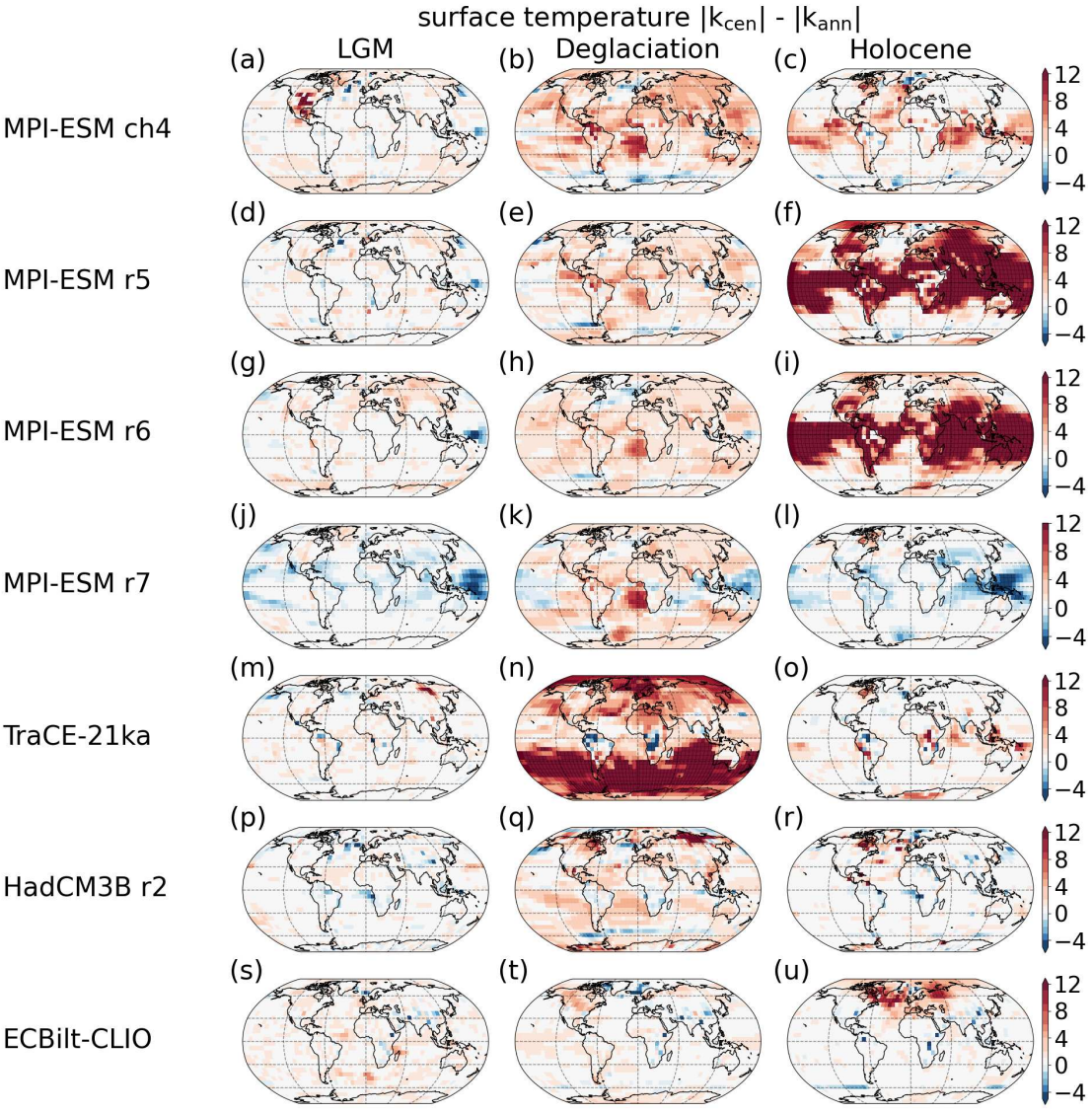


Figure S2: Difference of absolute centennial to annual surface temperature kurtosis for LGM, Deglaciation and Holocene. Red tones indicate that centennial skewness is larger, blue ones the opposite.

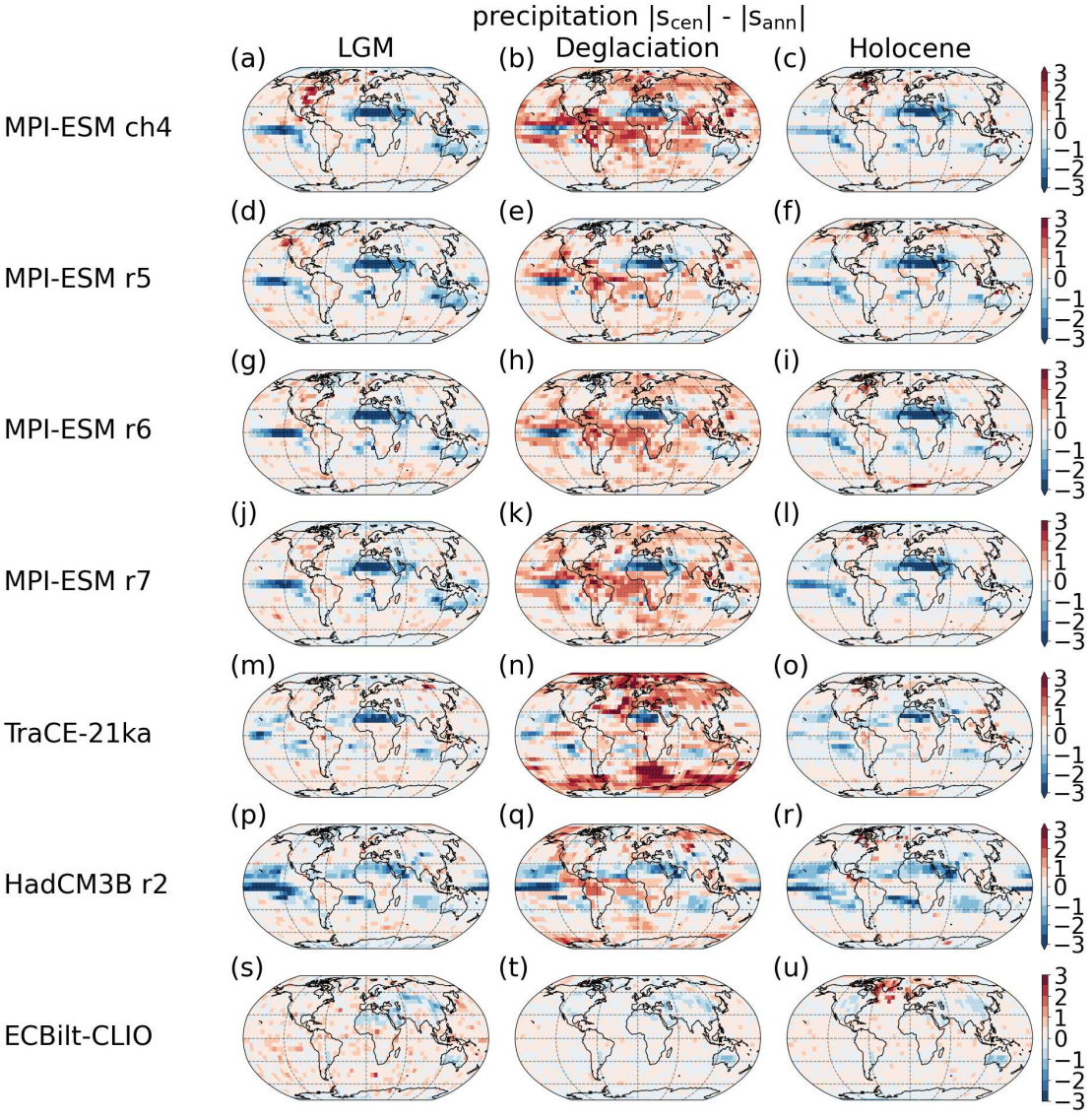


Figure S3: Difference of absolute centennial to annual precipitation skewness for LGM, Deglaciation and Holocene. Red tones indicate that centennial skewness is larger, blue ones the opposite.

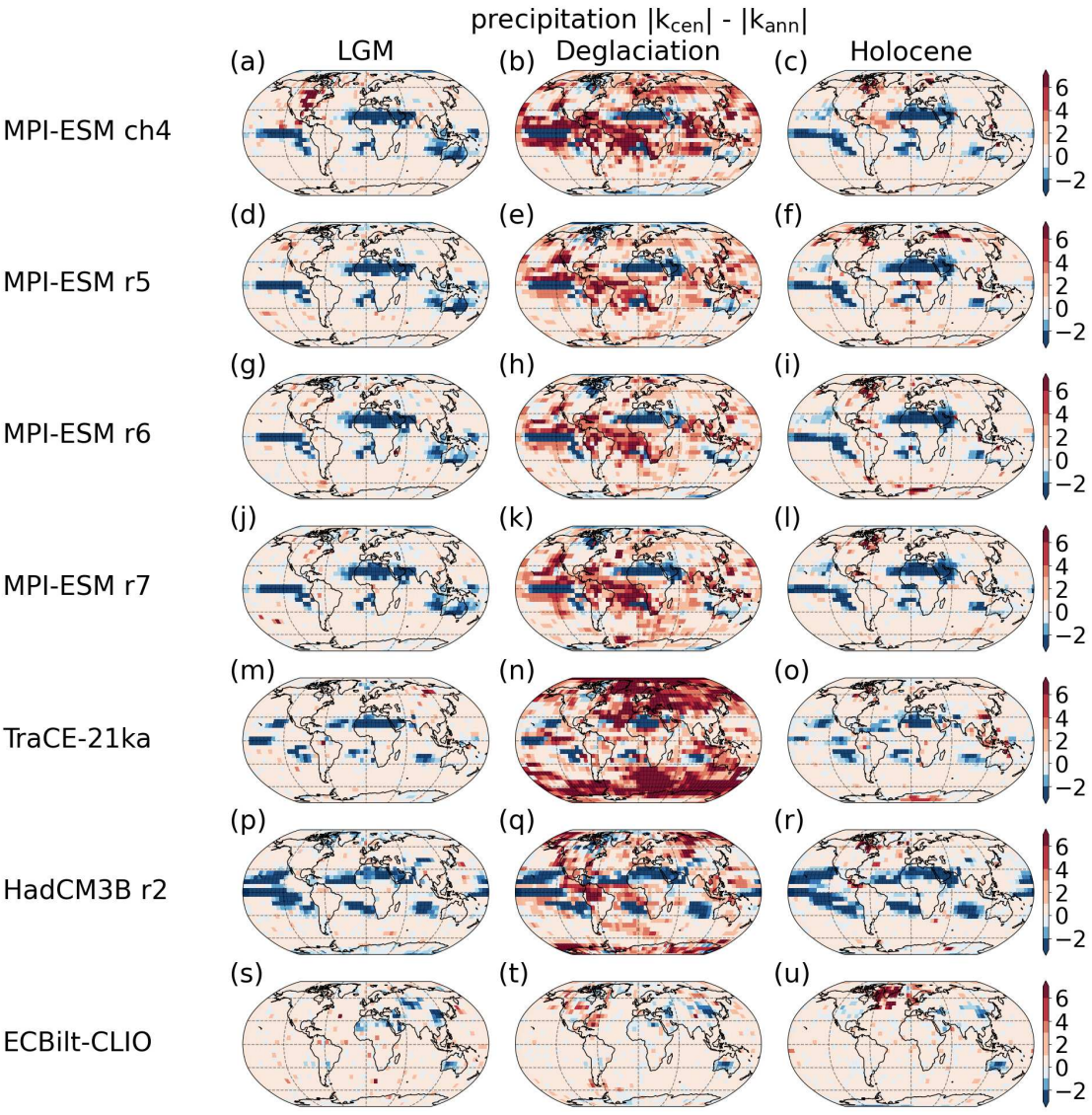


Figure S4: Difference of absolute centennial to annual precipitation kurtosis for LGM, Deglaciation and Holocene. Red tones indicate that centennial skewness is larger, blue ones the opposite.

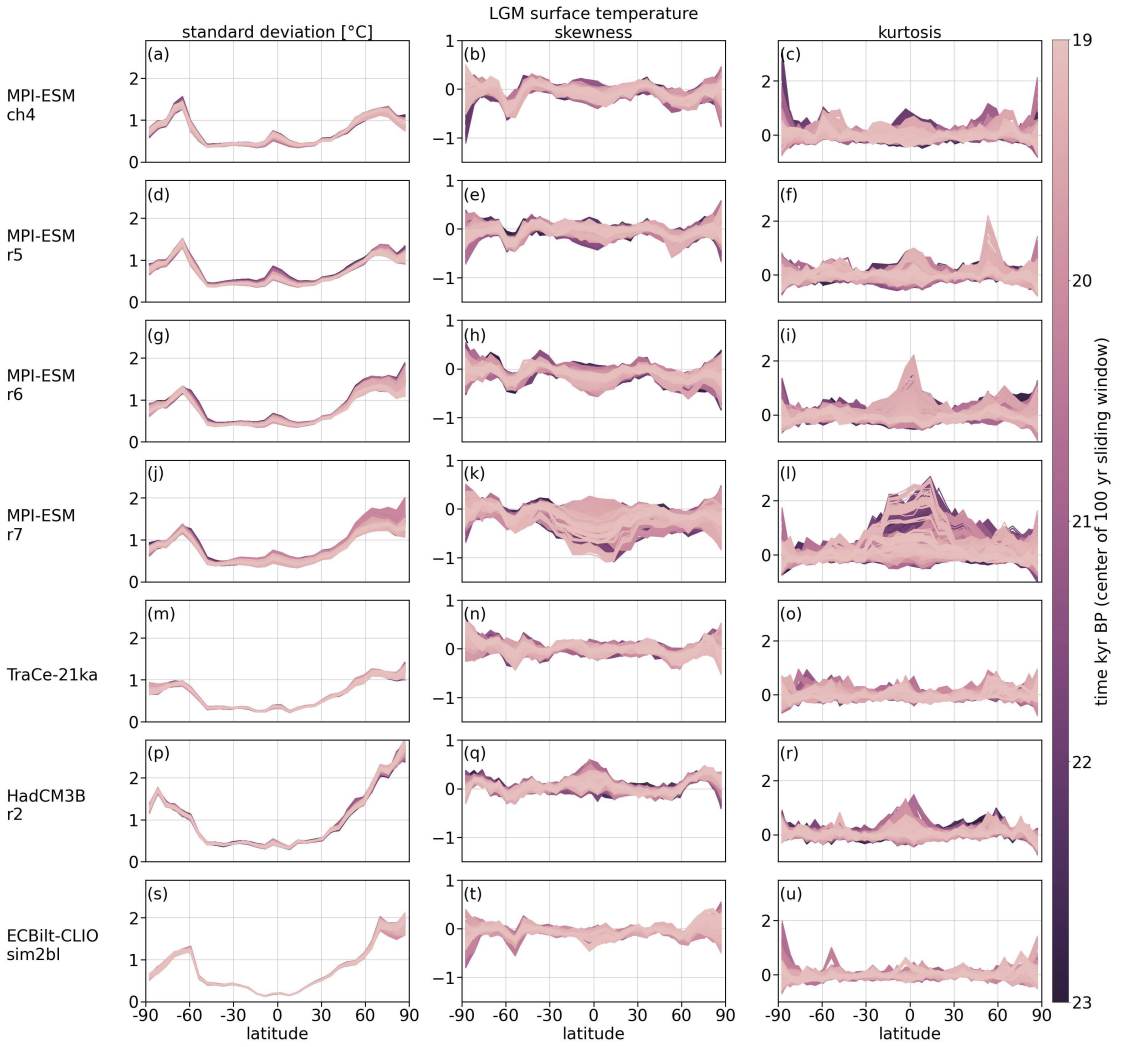


Figure S5: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of surface temperature across latitudes during the LGM. Moments are computed over 100 yr long rolling windows. Time is given as the middle of that rolling window.

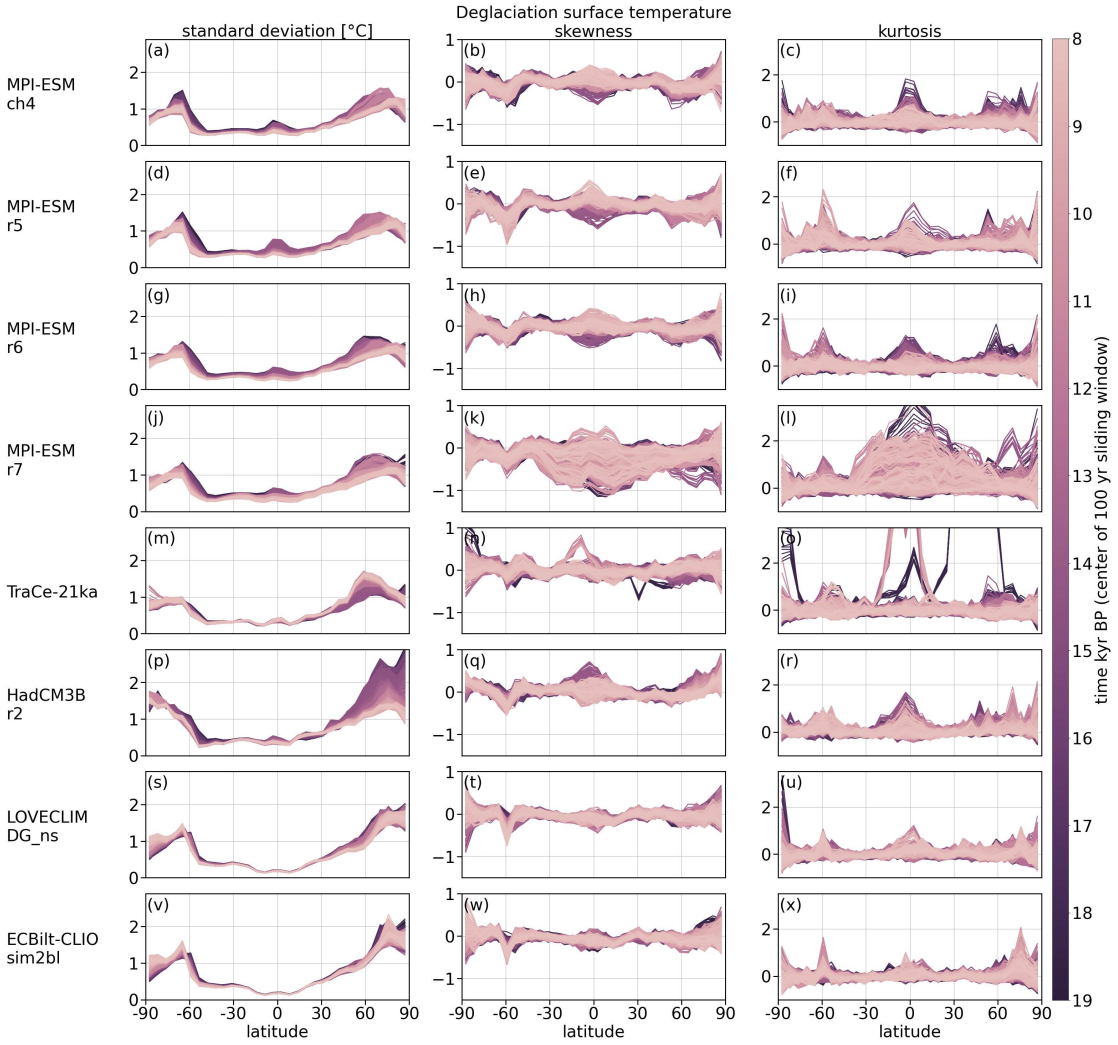


Figure S6: As Fig. S5 but for the Deglaciation.

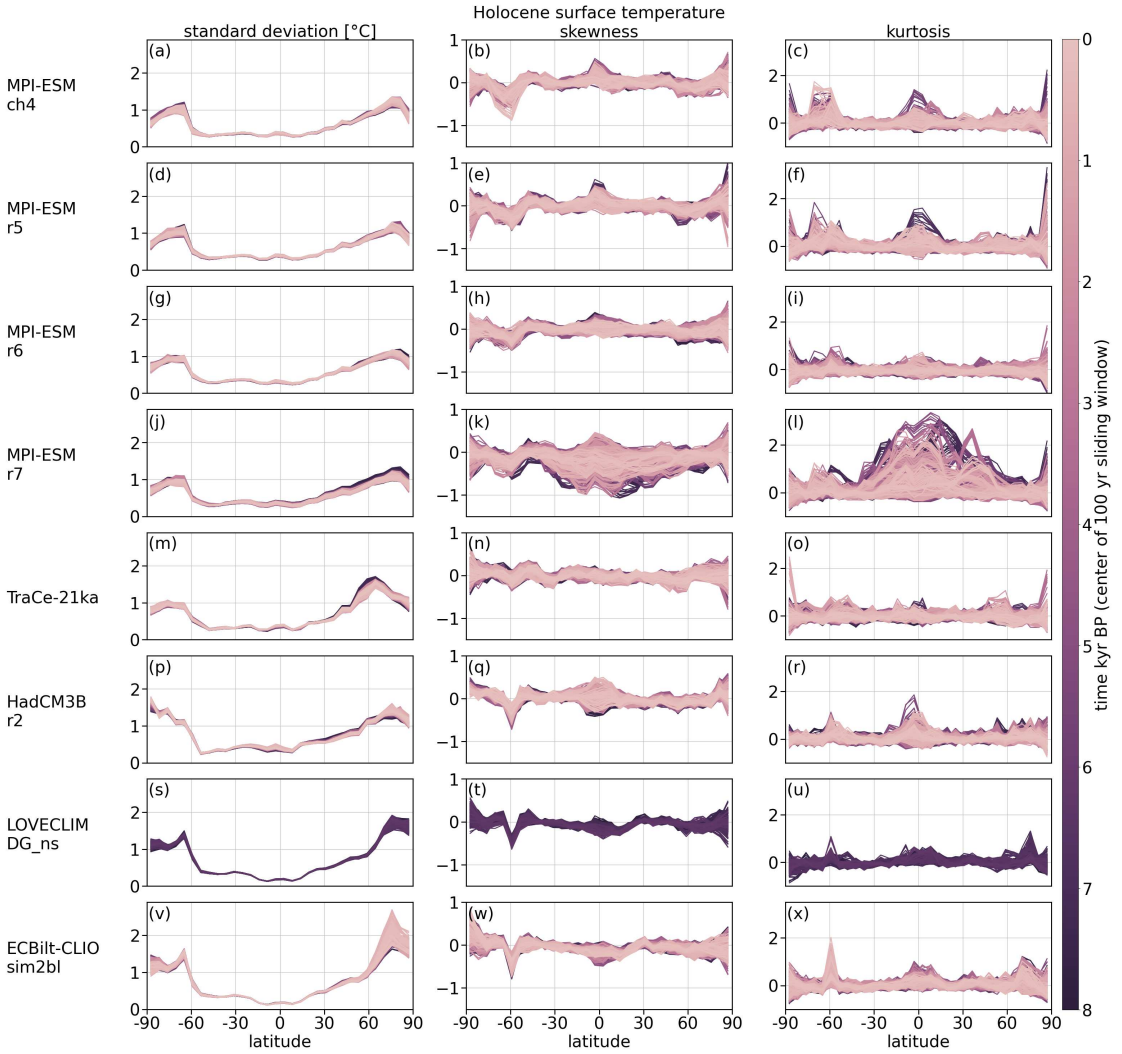


Figure S7: As Fig. S5 but for the Holocene.

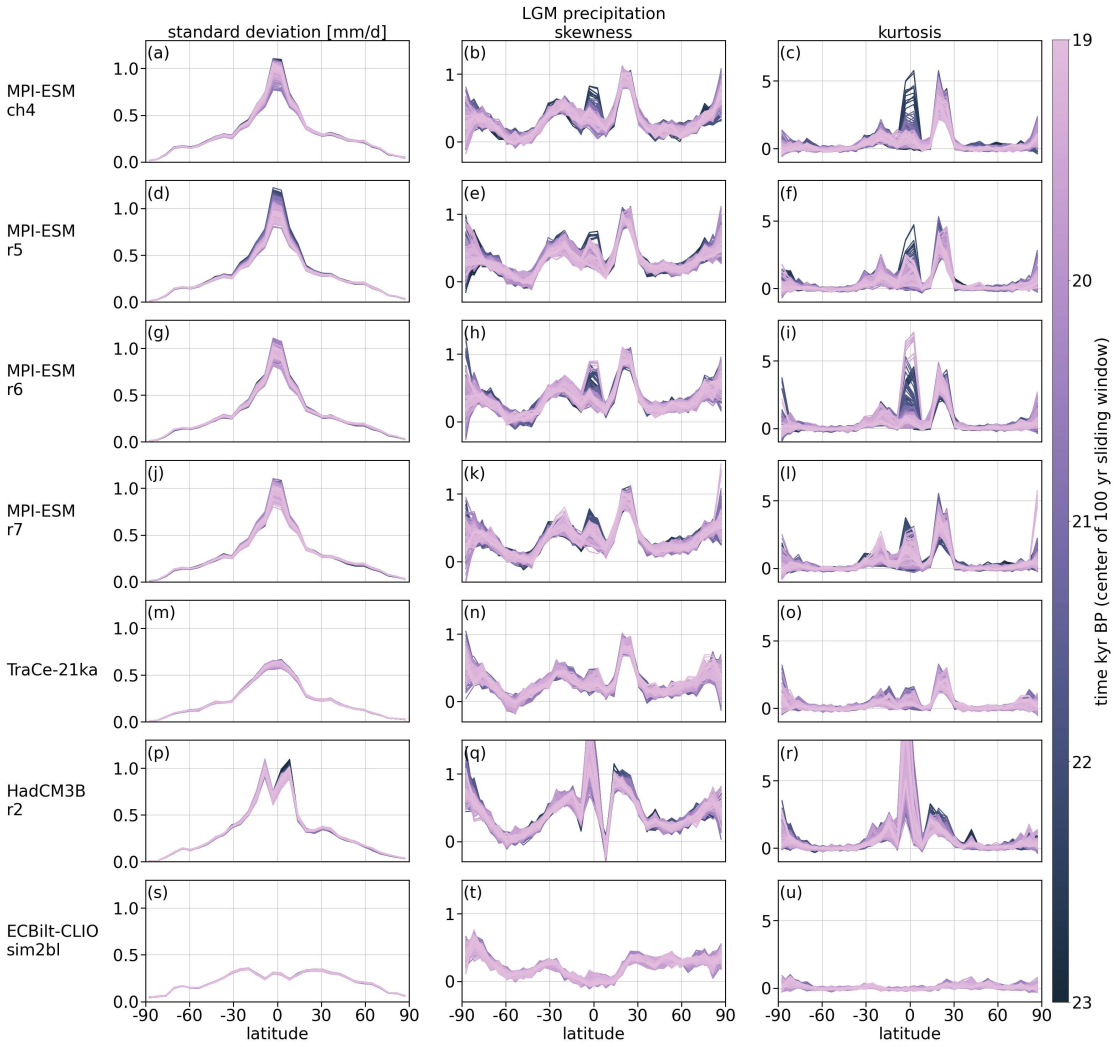


Figure S8: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of precipitation across latitudes during the LGM. Moments are computed over 100 yr long rolling windows. Time is given as the middle of that rolling window.

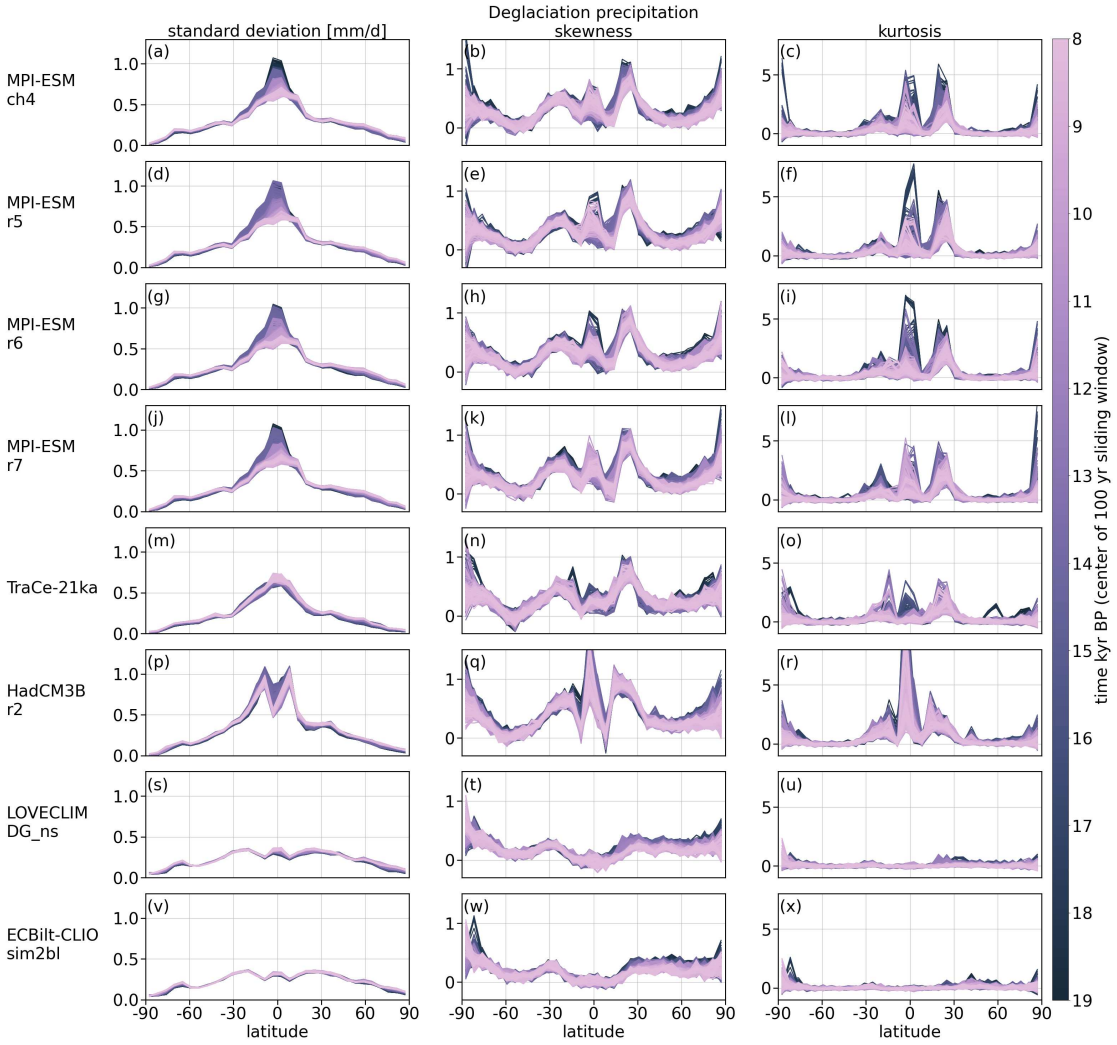


Figure S9: As Fig. S8 but for the Deglaciation.

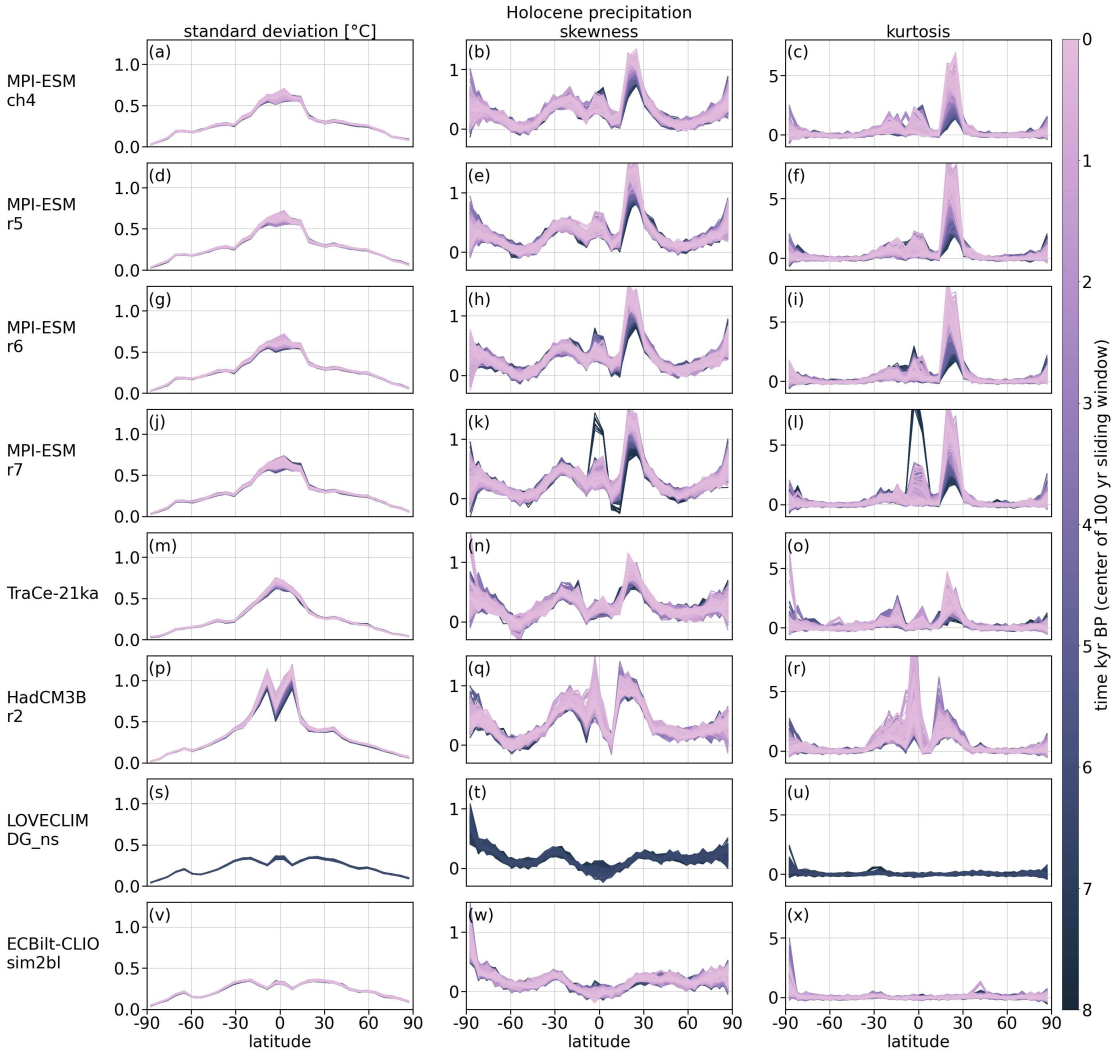


Figure S10: As Fig. S8 but for the Holocene.

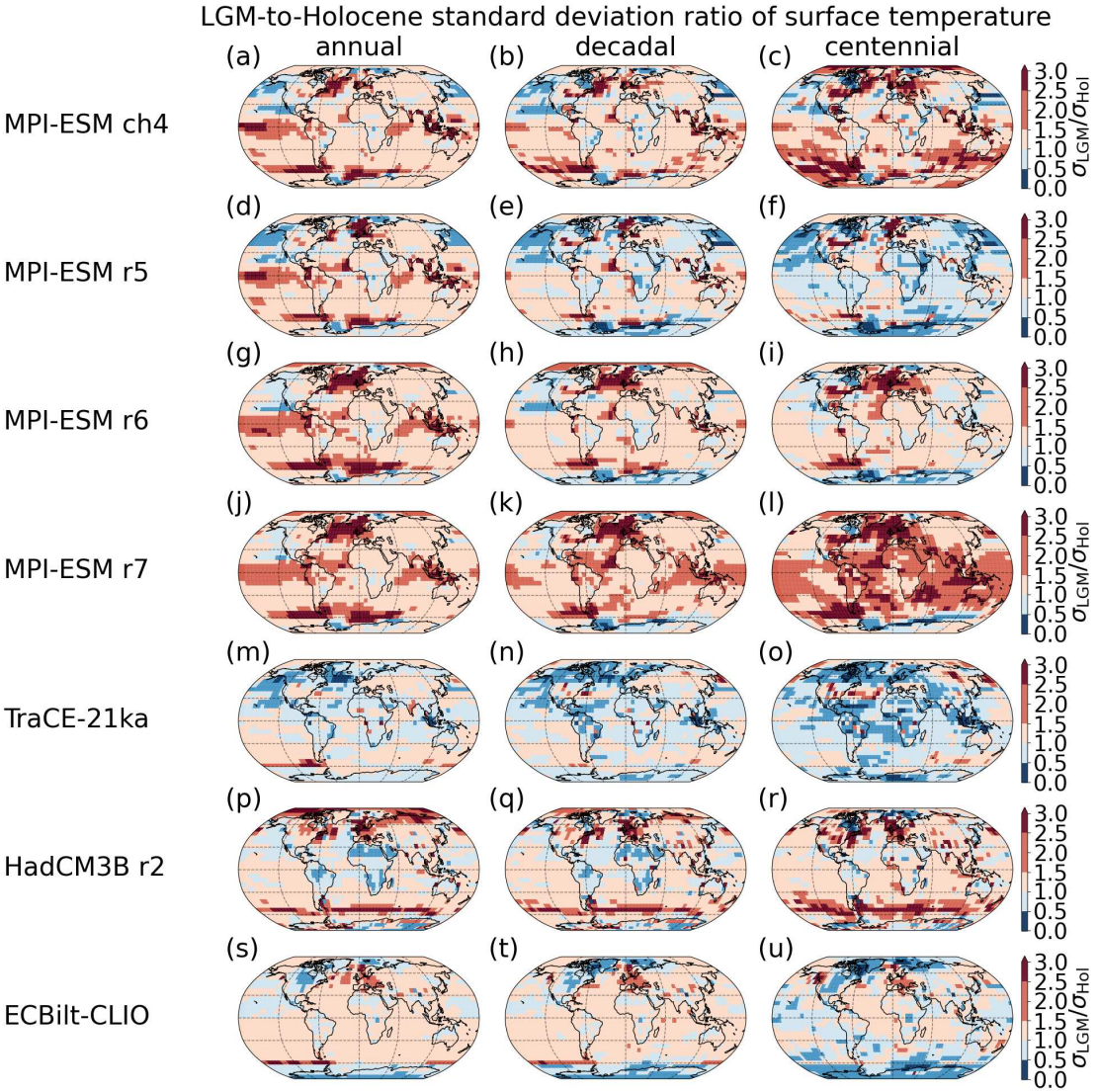


Figure S11: Ratio of standard deviation of surface temperature of the LGM with respect to the Holocene for annual (left column), decadal (middle column) and centennial (right column) timescales.

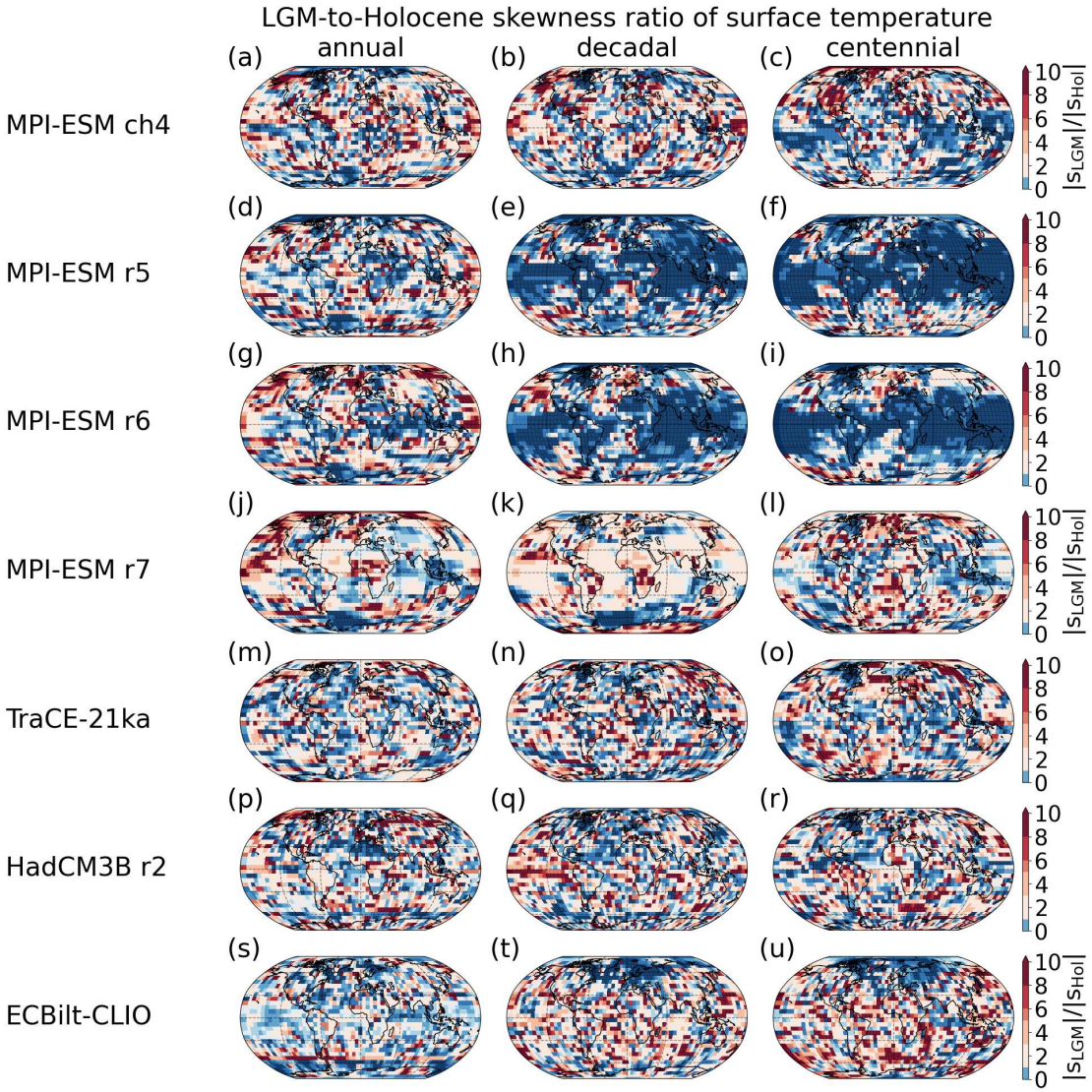


Figure S12: Ratio of skewness of surface temperature of the LGM with respect to the Holocene for annual (left column), decadal (middle column) and centennial (right column) timescales.

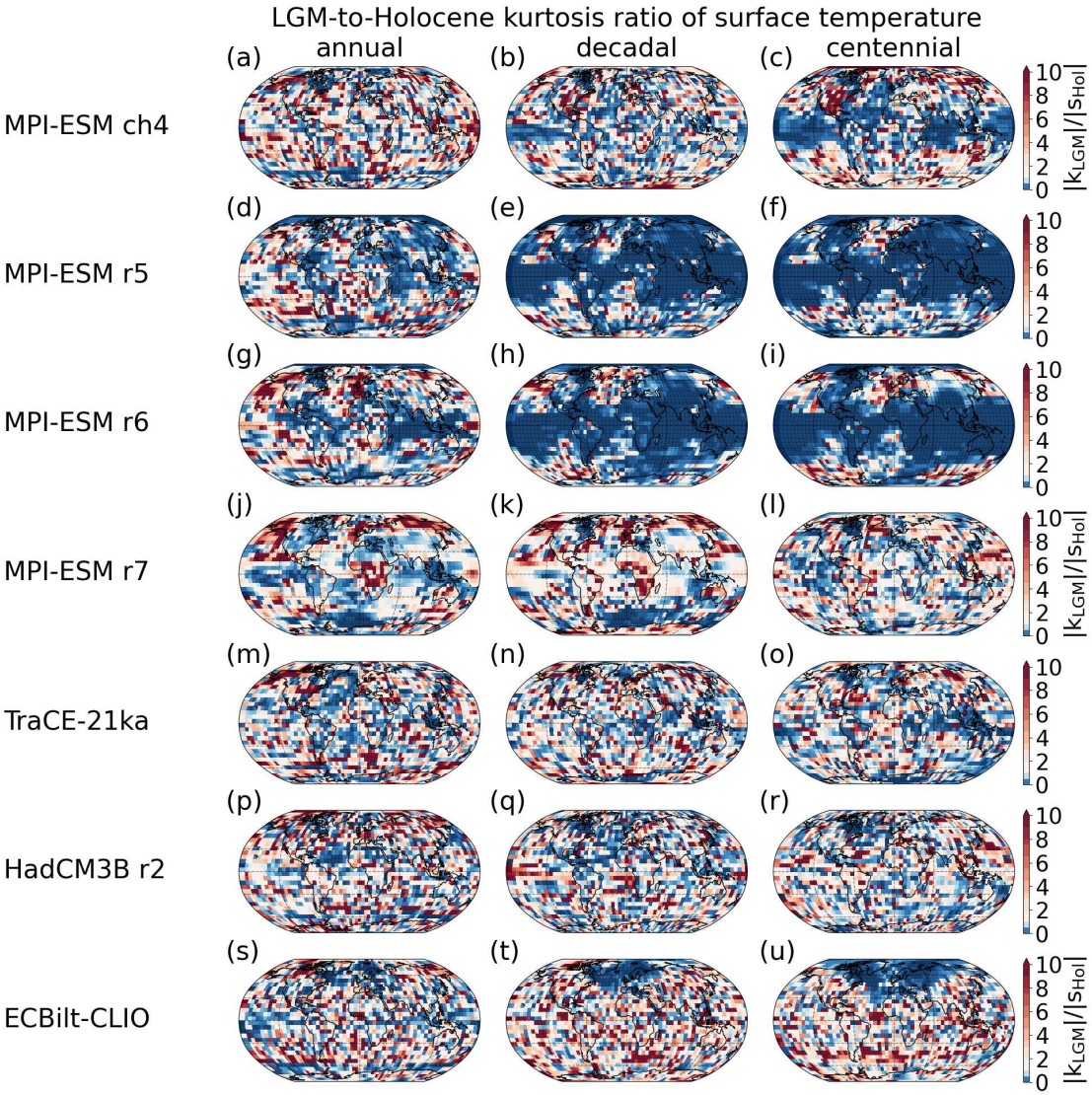


Figure S13: Ratio of kurtosis of surface temperature of the LGM with respect to the Holocene for annual (left column), decadal (middle column) and centennial (right column) timescales.

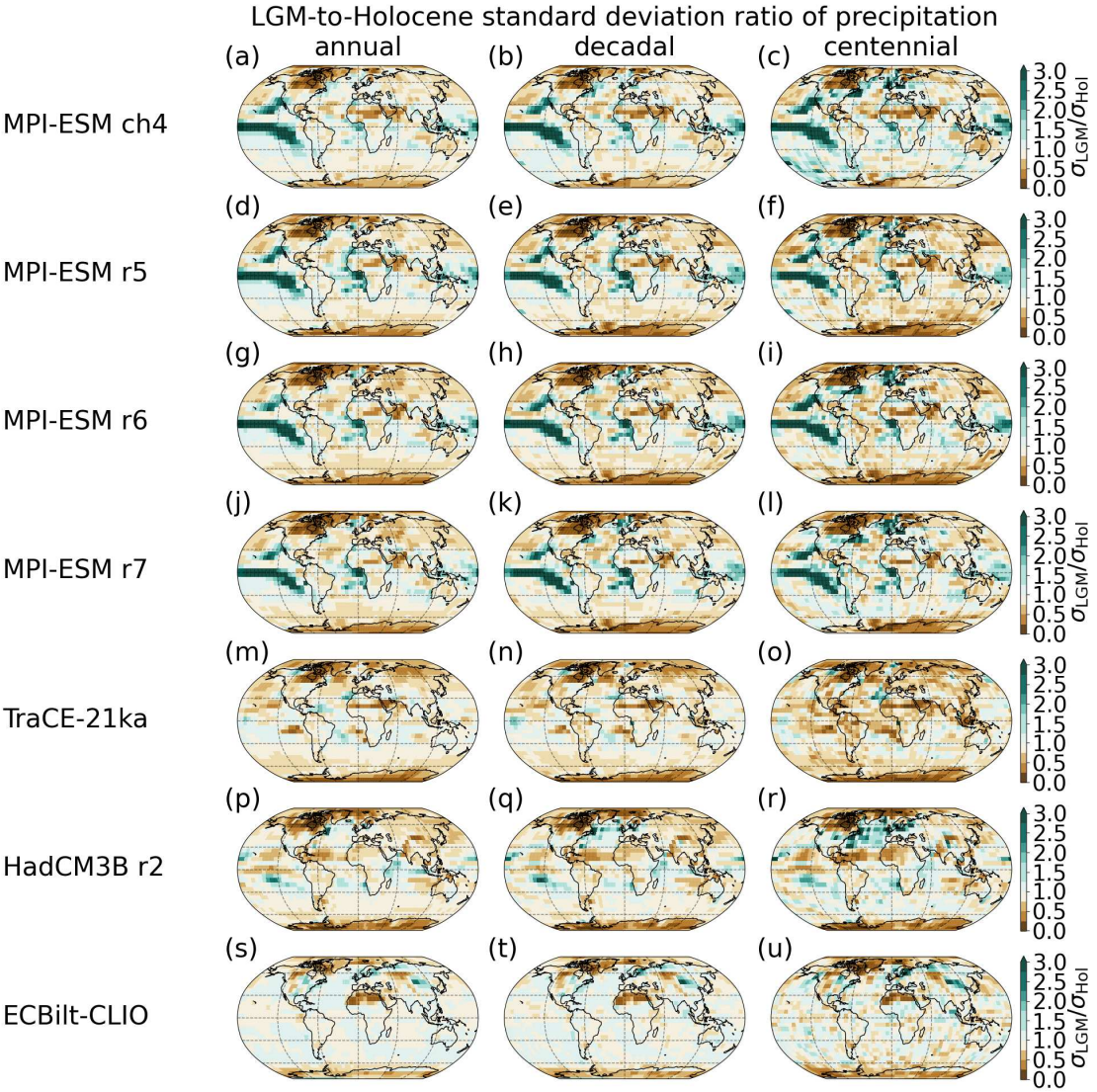


Figure S14: Ratio of standard deviation of precipitation of the LGM with respect to the Holocene for annual (left column), decadal (middle column) and centennial (right column) timescales.

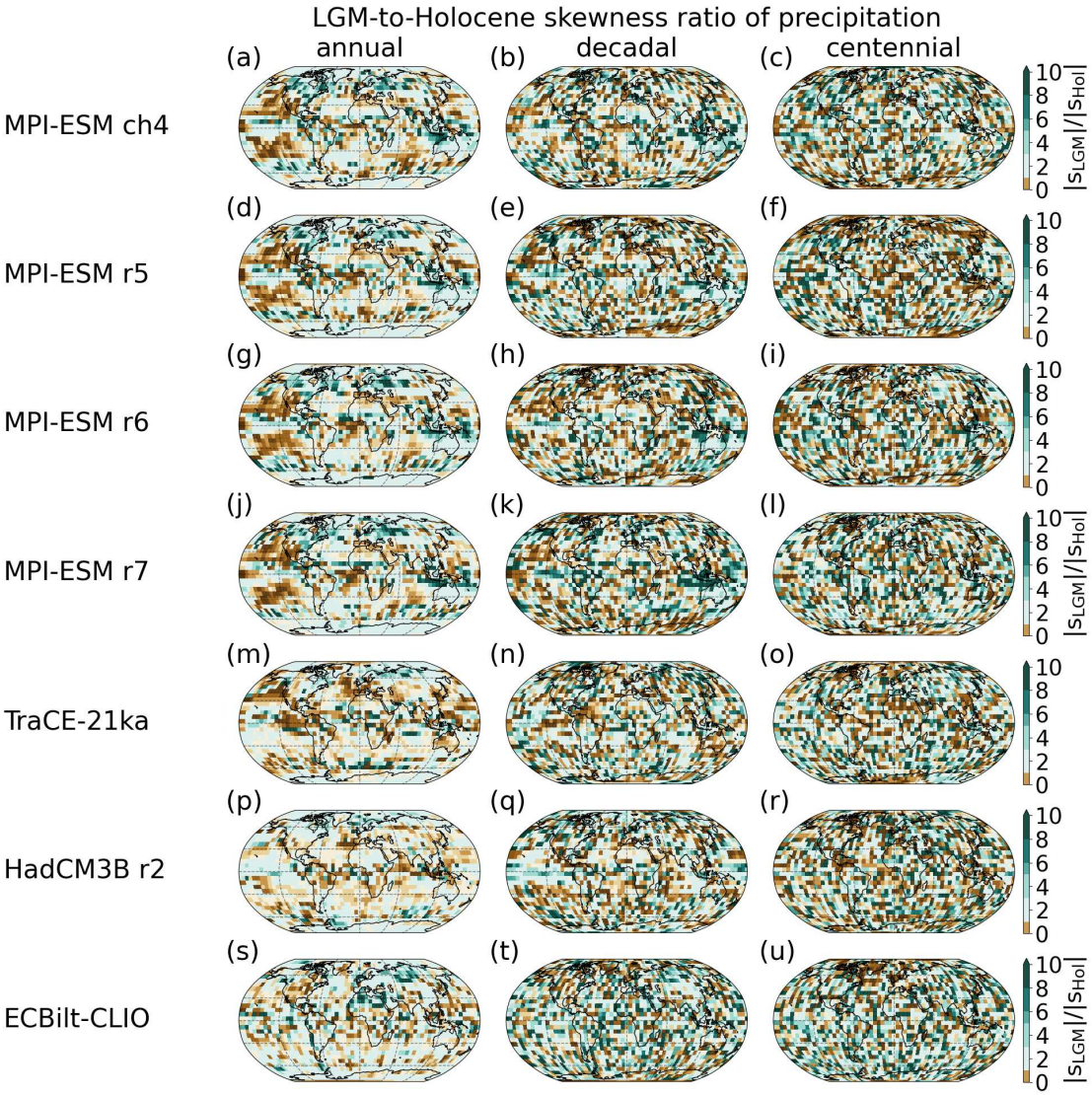


Figure S15: Ratio of skewness of precipitation of the LGM with respect to the Holocene for annual (left column), decadal (middle column) and centennial (right column) timescales.

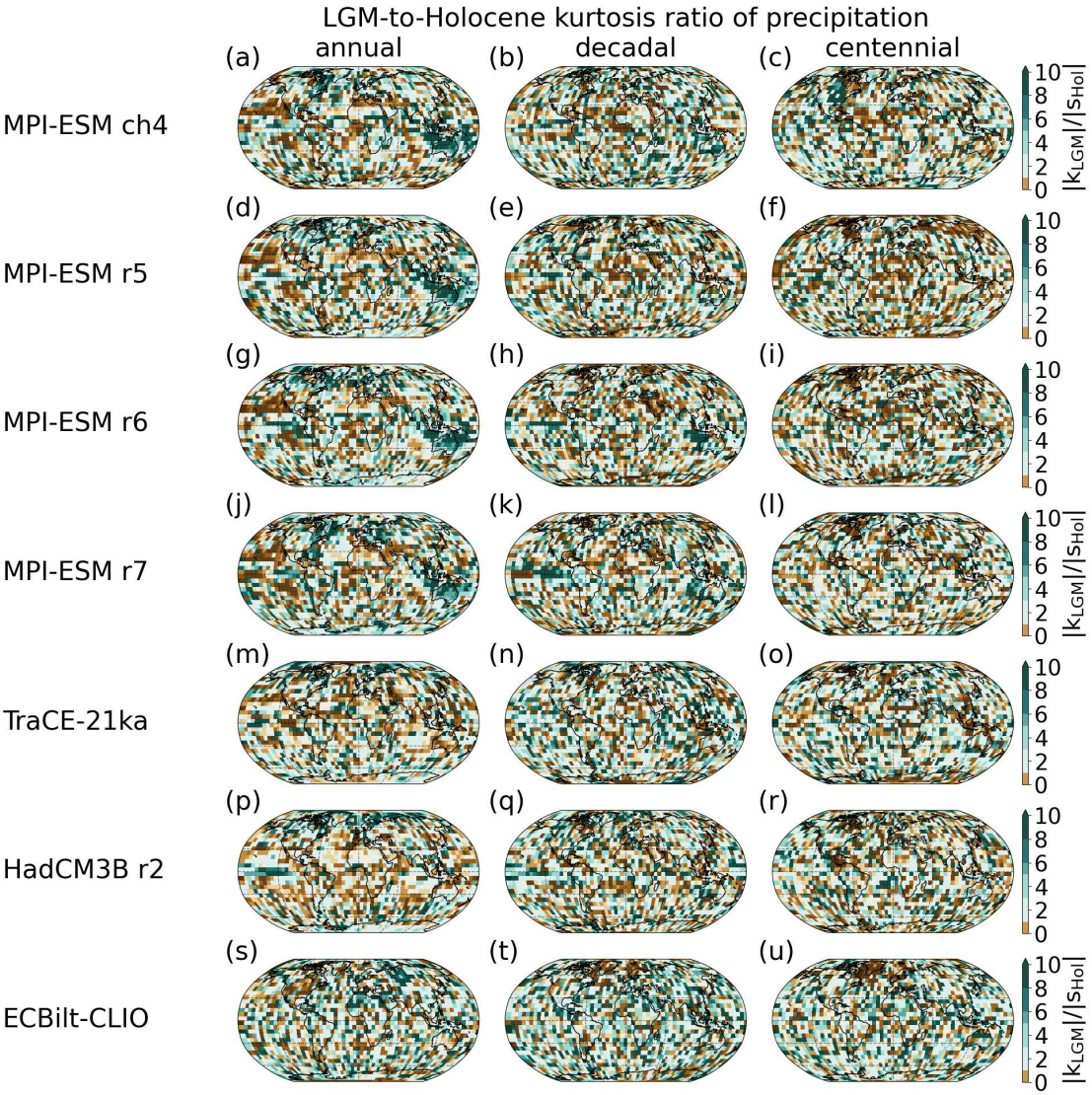


Figure S16: Ratio of kurtosis of precipitation of the LGM with respect to the Holocene for annual (left column), decadal (middle column) and centennial (right column) timescales.

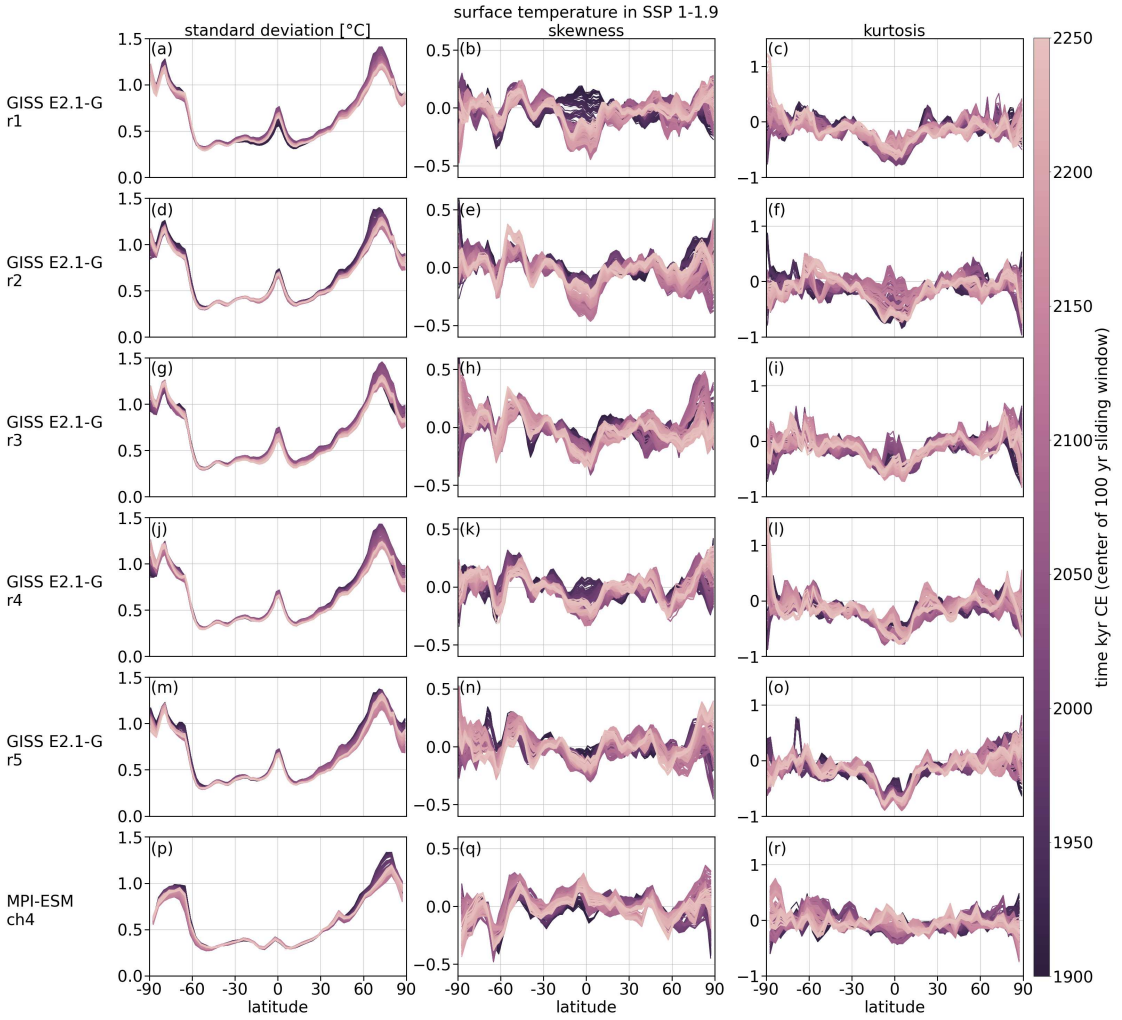


Figure S17: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of surface temperature across latitude for combined historical and SSP 1-1.9 runs from the CMIP6 multi-model ensemble and MPI-ESM ch4. Moments are computed over 100 yr long rolling windows from 1850 – 2300 CE. Time is given as the middle of that rolling window.

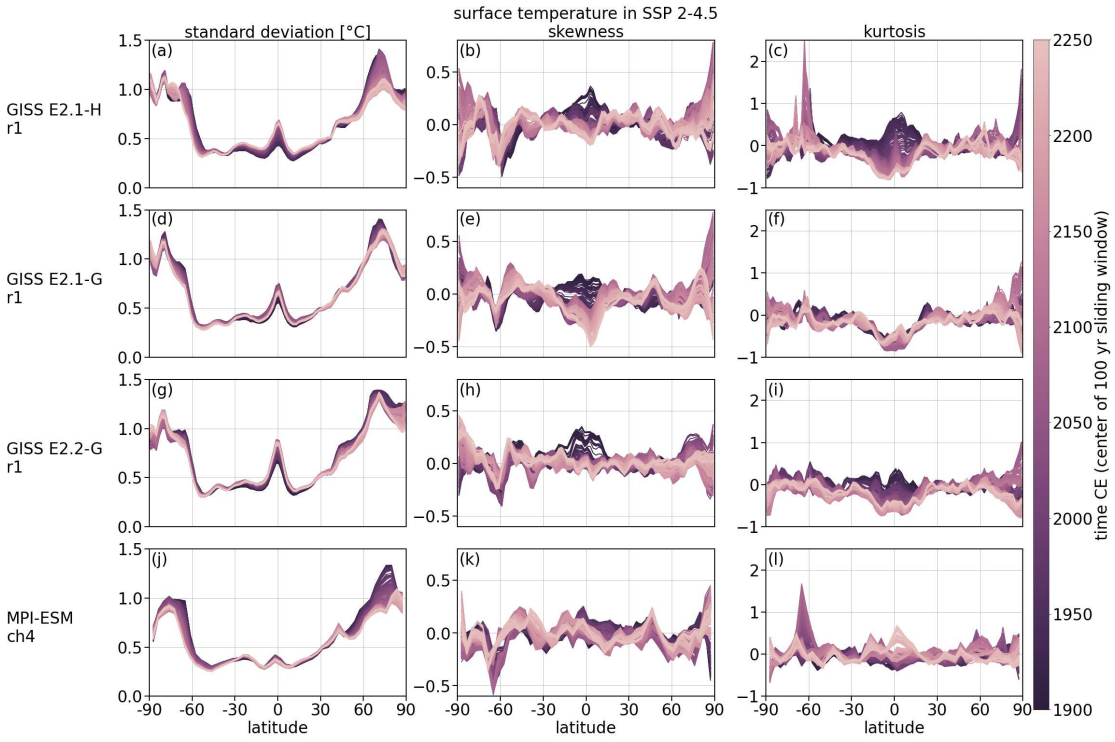


Figure S18: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of surface temperature across latitude for combined historical and SSP 2-4.5 runs from the CMIP6 multi-model ensemble and MPI-ESM ch4. Moments are computed over 100 yr long rolling windows from 1850 – 2300 CE. Time is given as the middle of that rolling window.

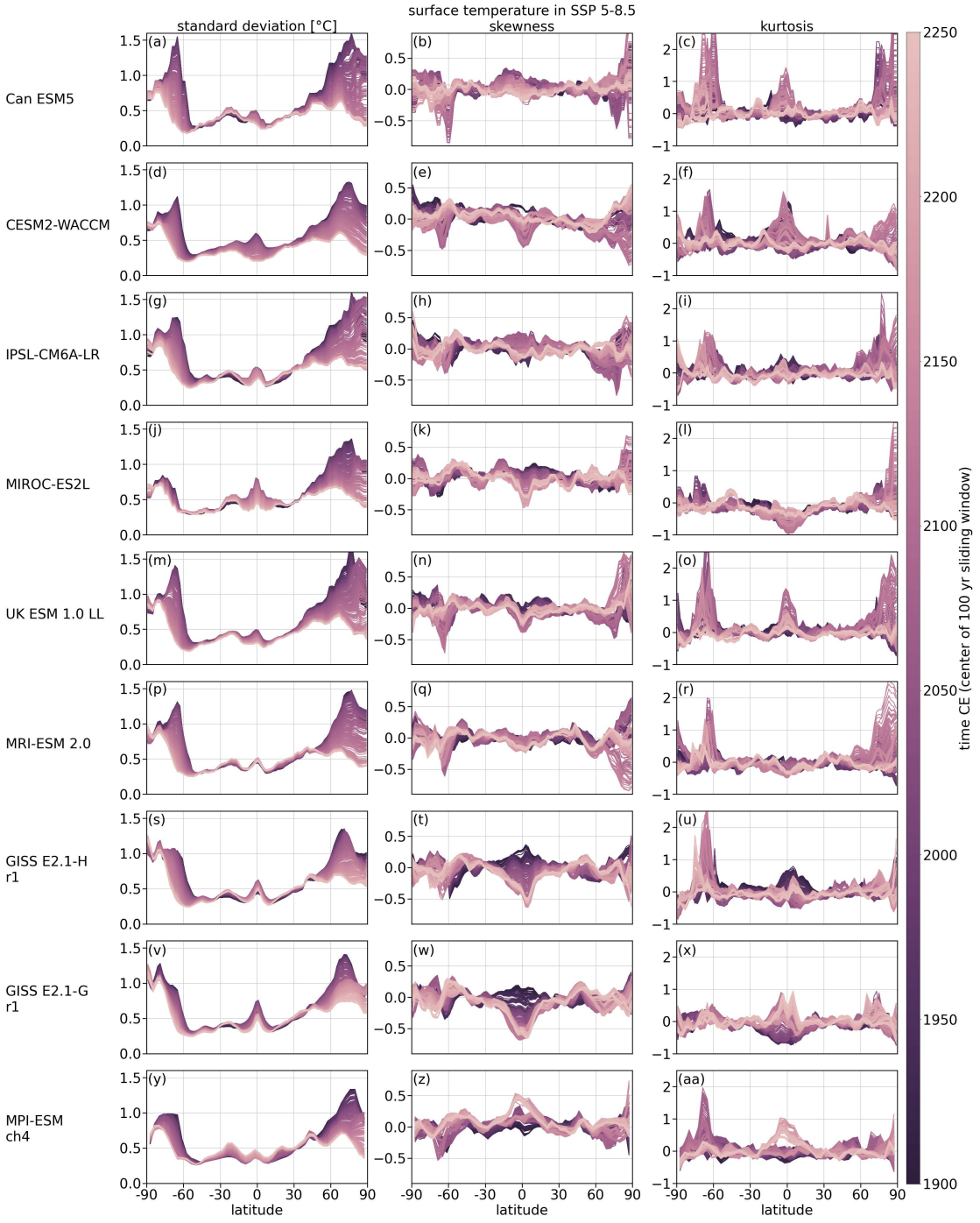


Figure S19: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of surface temperature across latitude for combined historical and SSP 5-8.5 runs from the CMIP6 multi-model ensemble and MPI-ESM ch4. Moments are computed over 100 yr long rolling windows from 1850 – 2300 CE. Time is given as the middle of that rolling window.

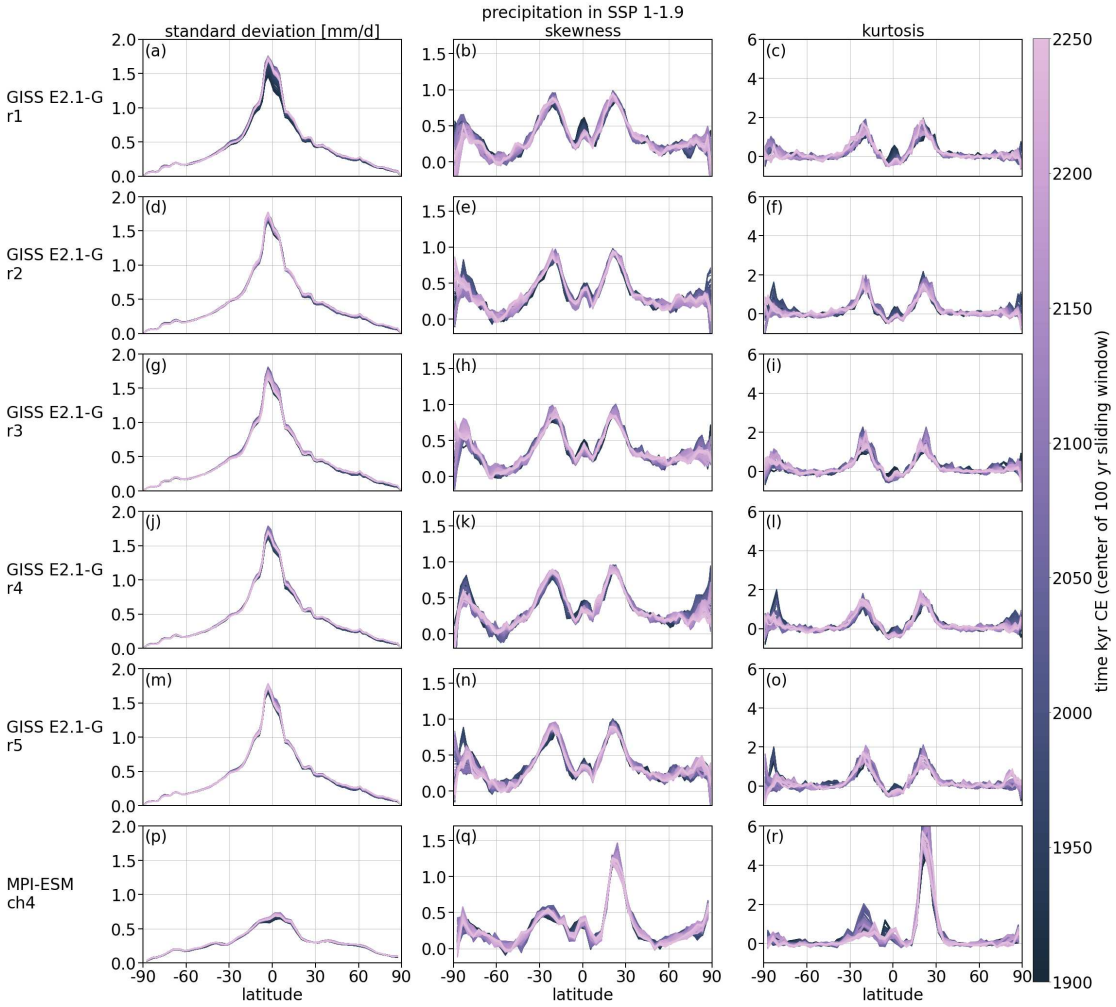


Figure S20: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of precipitation across latitude for combined historical and SSP 1-1.9 runs from the CMIP6 multi-model ensemble and MPI-ESM ch4. Moments are computed over 100 yr long rolling windows from 1850 – 2300 CE. Time is given as the middle of that rolling window.

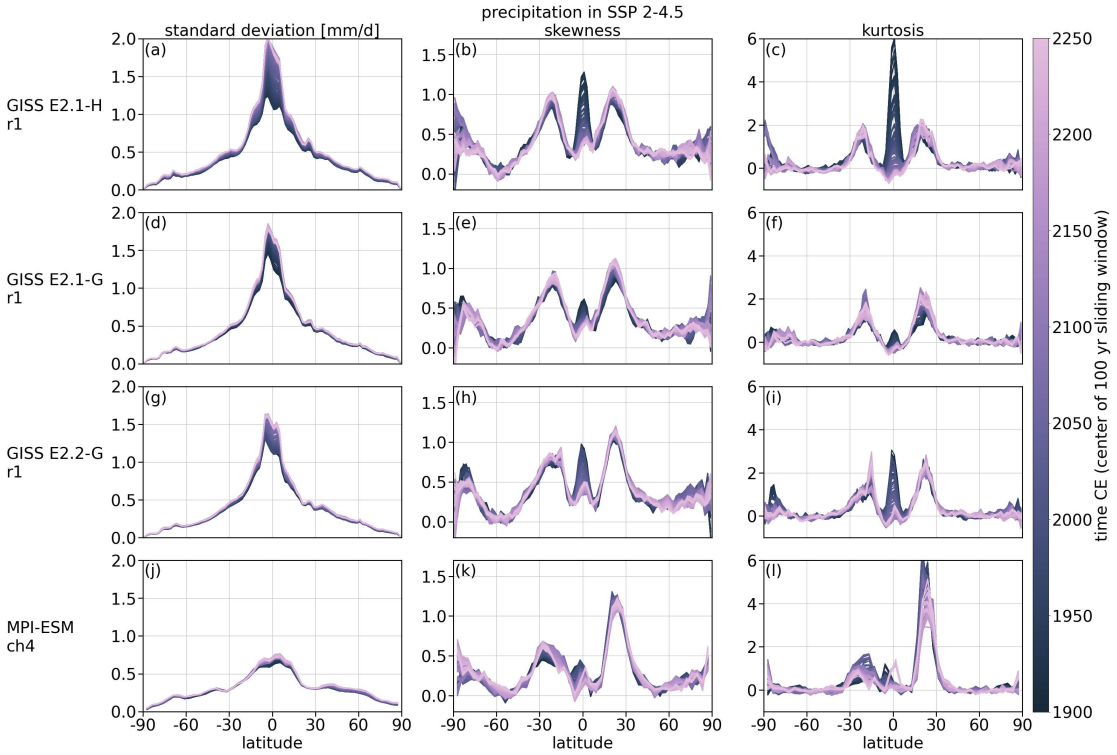


Figure S21: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of precipitation across latitude for combined historical and SSP 2-4.5 runs from the CMIP6 multi-model ensemble and MPI-ESM ch4. Moments are computed over 100 yr long rolling windows from 1850 – 2300 CE. Time is given as the middle of that rolling window.

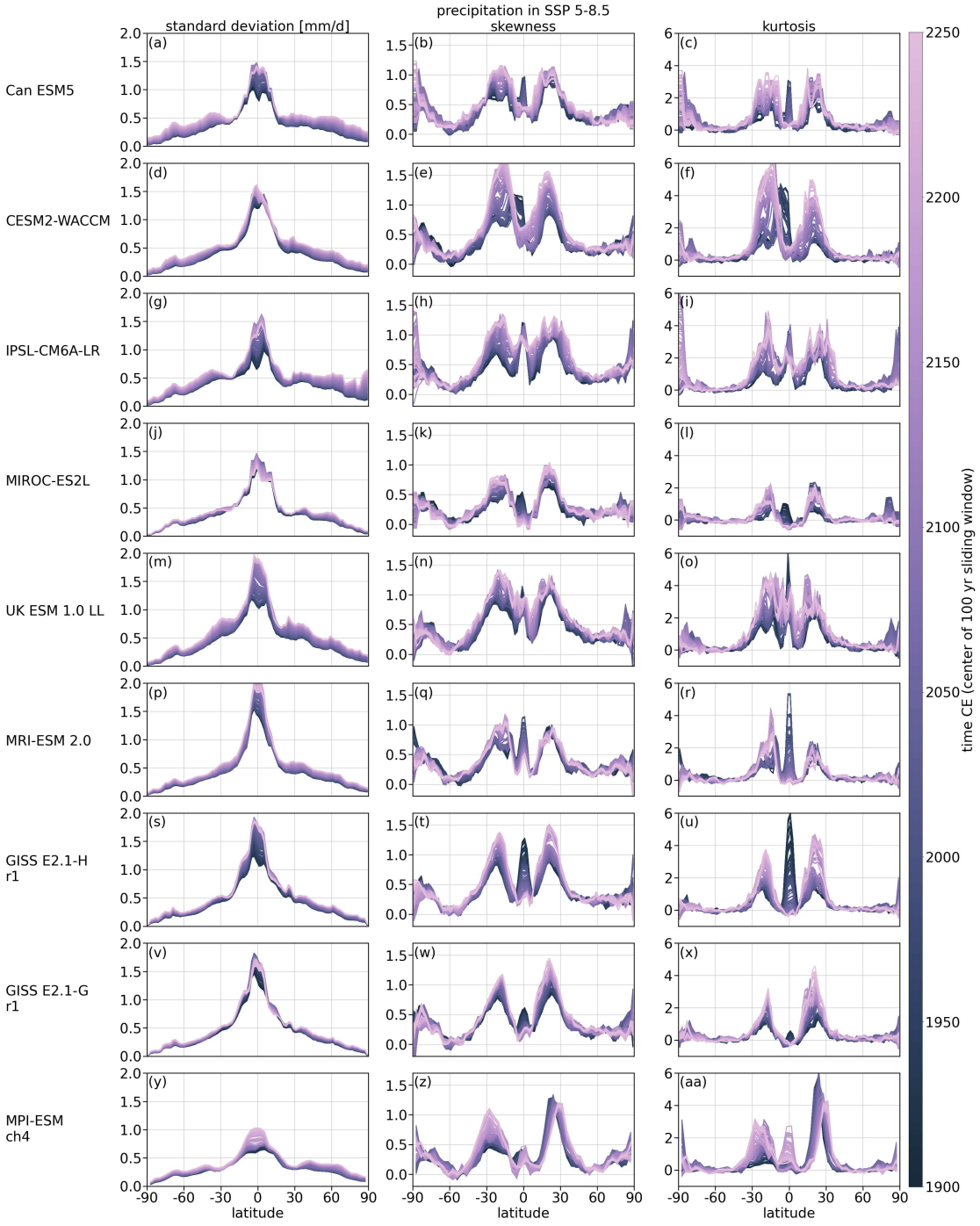


Figure S22: Annual standard deviation (left column), skewness (middle column) and kurtosis (right column) of precipitation across latitude for combined historical and SSP 5-8.5 runs from the CMIP6 multi-model ensemble and MPI-ESM ch4. Moments are computed over 100 yr long rolling windows from 1850 – 2300 CE. Time is given as the middle of that rolling window.

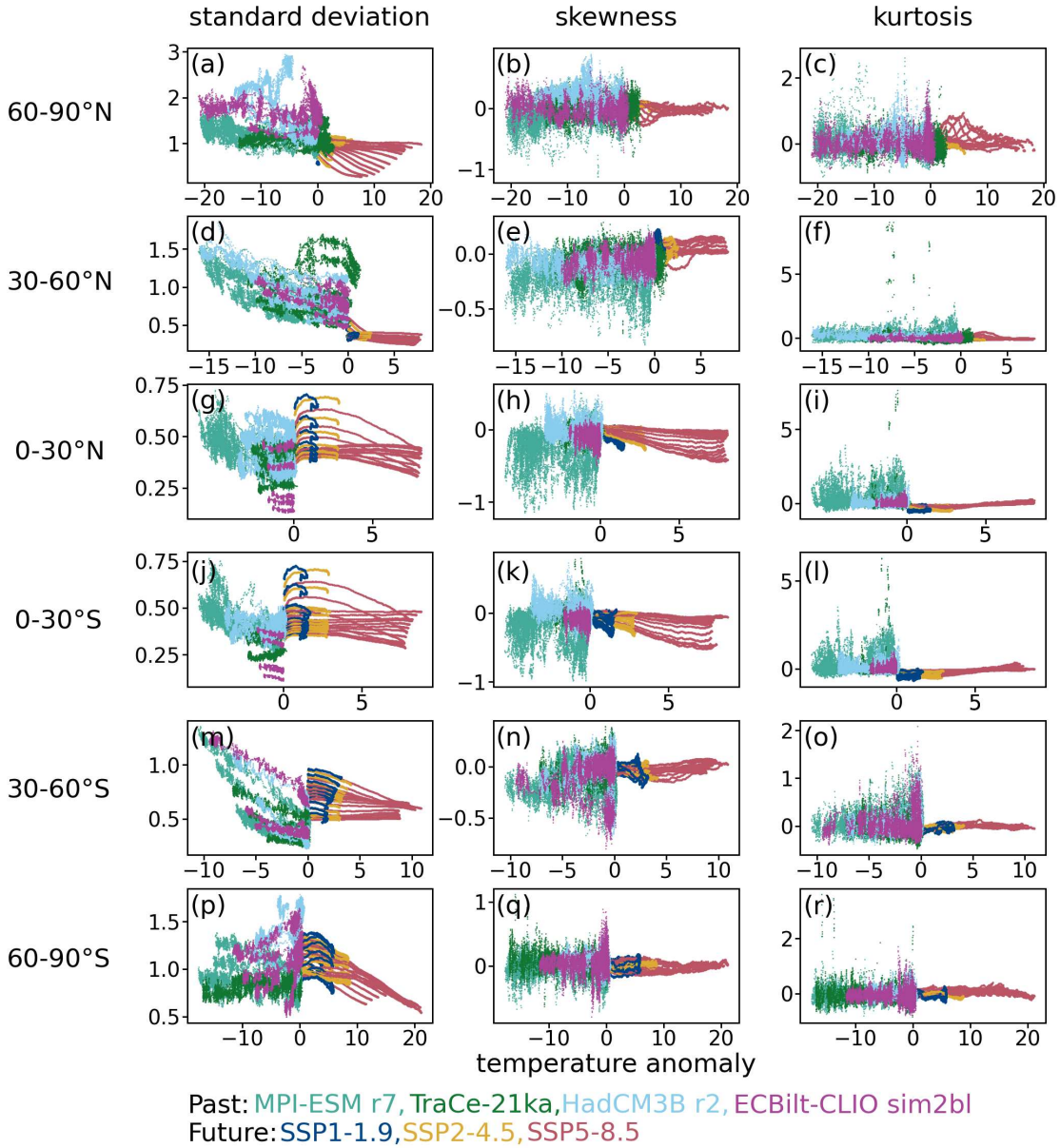


Figure S23: Moments (y-axis) of surface temperature across each latitude against the mean anomaly with respect to 1850 – 1899 CE (x-axis) for four Deglacial simulations as well as the CMIP6 ensemble mean for SSP 1-1.9, SSP 2-4.5 and SSP 5-8.5.

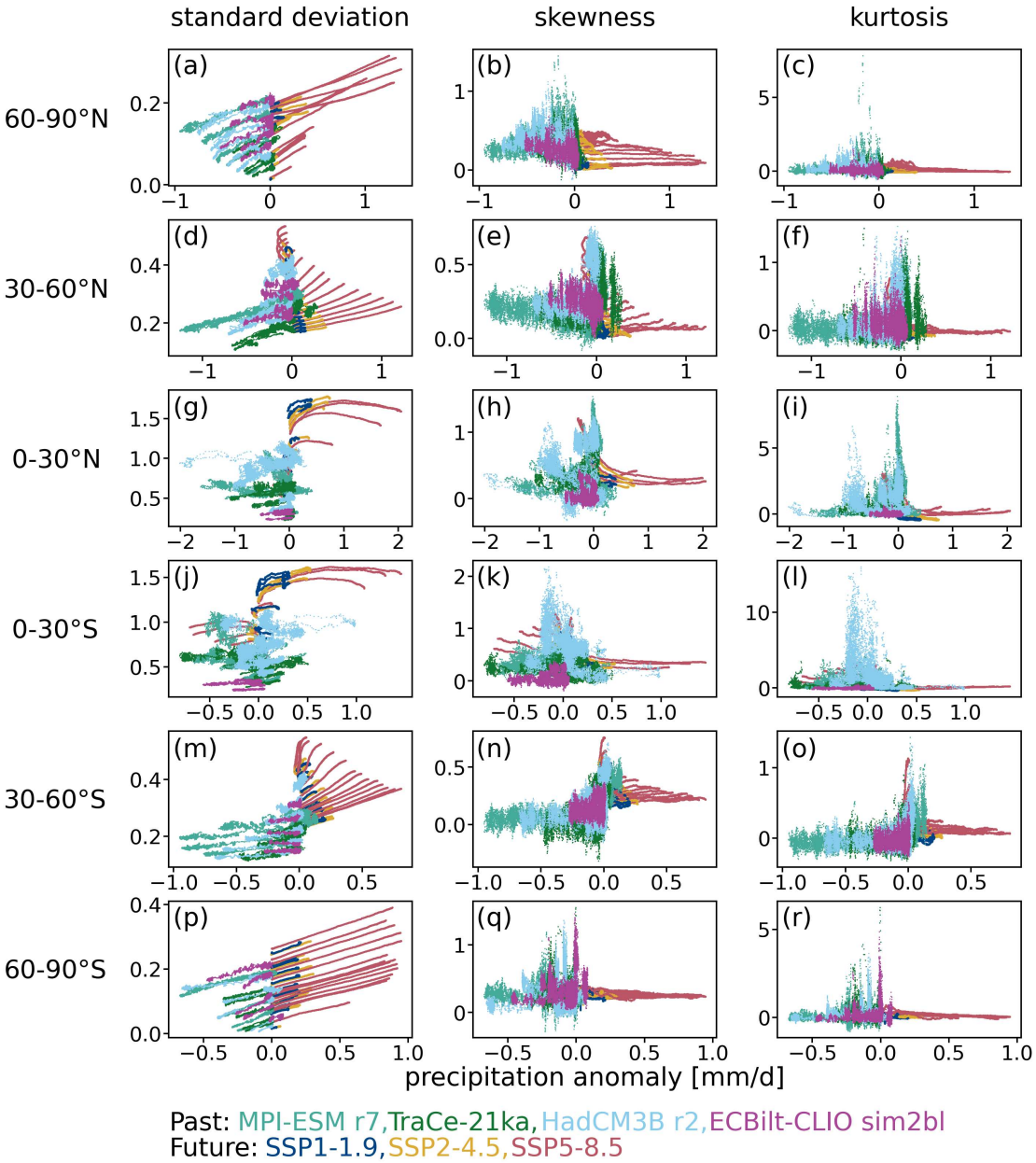


Figure S24: Moments (y-axis) of precipitation across each latitude against the mean anomaly with respect to 1850 – 1899 CE (x-axis) for four Deglacial simulations as well as the CMIP6 ensemble mean for SSP 1-1.9, SSP 2-4.5 and SSP 5-8.5.

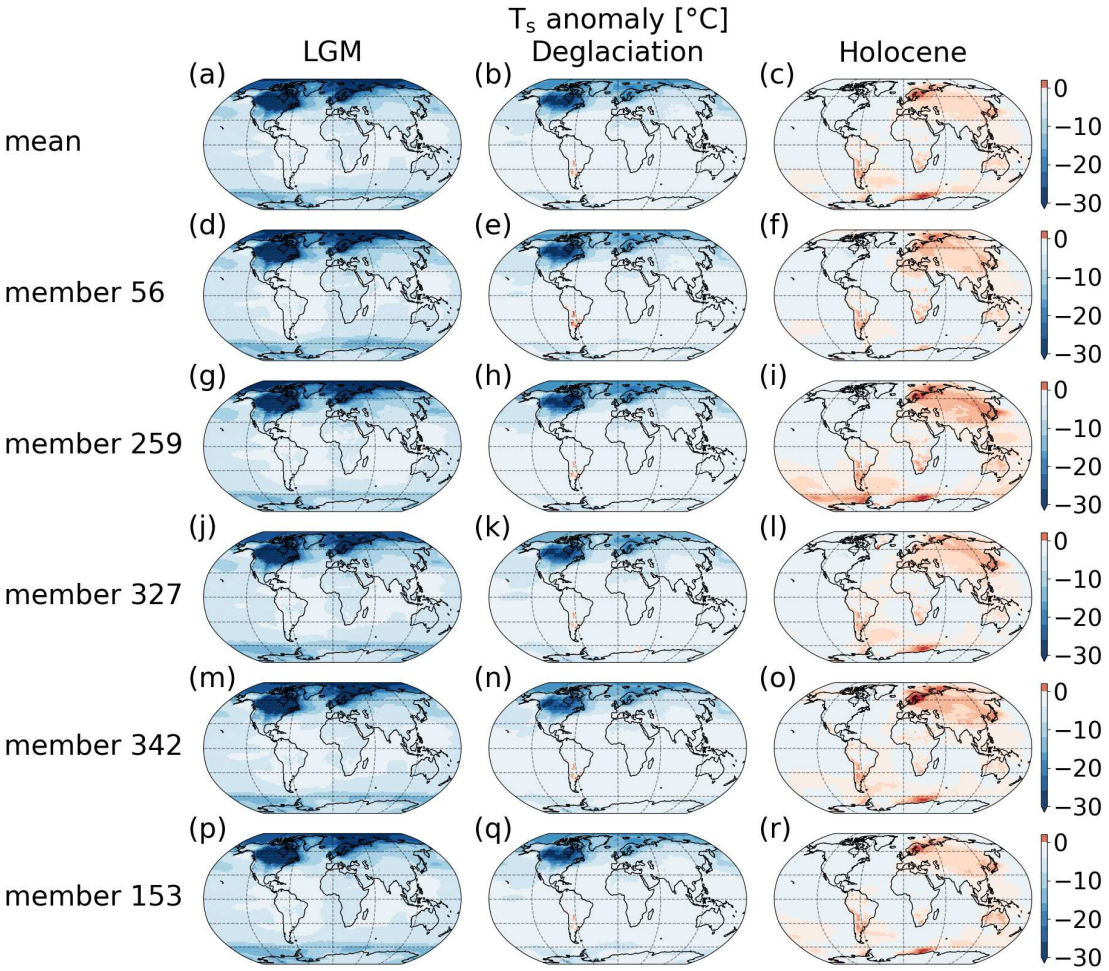


Figure S25: Mean temperature anomaly with respect to the past 2000 years of the LGMR for the LGM (left column), Deglaciation (middle column) and the Holocene (right column). Shown are in panels (a)–(c) the ensemble mean (as in **P3** but without regridding) and in panels (d)–(r) five randomly selected ensemble members. For the mean fields variation between the ensemble members is small.

A4 Publication 1: TransEBM v. 1.0: Description, tuning, and validation of a transient model of the Earth's energy balance in two dimensions

Publication 1 (P1): E. Ziegler & K. Rehfeld (2020): TransEBM v. 1.0: Description, tuning, and validation of a transient model of the Earth's energy balance in two dimensions, Geoscientific Model Development, 14. doi:10.5194/gmd-14-2843-2021

Geosci. Model Dev., 14, 2843–2866, 2021
https://doi.org/10.5194/gmd-14-2843-2021

Geoscientific Model Development
Open Access
EGU

TransEBM v. 1.0: description, tuning, and validation of a transient model of the Earth's energy balance in two dimensions

Elisa Ziegler and Kira Rehfeld

Institute of Environmental Physics, Heidelberg University, Heidelberg, Germany

Correspondence: Elisa Ziegler (eziegler@iup.uni-heidelberg.de) and Kira Rehfeld (krehfeld@iup.uni-heidelberg.de)

Received: 14 July 2020 – Discussion started: 9 October 2020
Revised: 23 February 2021 – Accepted: 3 March 2021 – Published: 20 May 2021

Abstract. Modeling the long-term transient evolution of climate remains a technical and scientific challenge. However, understanding and improving modeling of the long-term behavior of the climate system increases confidence in projected changes in the mid- to long-term future. Energy balance models (EBMs) provide simplified and computationally efficient descriptions of long timescales and allow large ensemble runs by parameterizing energy fluxes. In this way, they can be used to pinpoint periods and phenomena of interest. Here, we present TransEBM, an extended version of the two-dimensional energy balance model by Zhuang et al. (2017a). Transient CO_2, solar insolation, orbital configuration, fixed ice coverage, and land–sea distribution are implemented as effective radiative forcings at the land surface. We show that the model is most sensitive to changes in CO_2 and ice distribution, but the obliquity and land–sea mask have significant influence on modeled temperatures as well. We tune TransEBM to reproduce the 1960–1989 CE global mean temperature and the Equator-to-pole and seasonal temperature gradients of ERA-20CM reanalysis (Hersbach et al., 2015). The resulting latitudinal and seasonal temperature distributions agree well with reanalysis and the general circulation model (GCM) HadCM3 for a simulation of the past millennium (Bühler et al., 2020). TransEBM does not represent the internal variability of the ocean–atmosphere system, but non-deterministic elements and nonlinearity can be introduced through model restarts and randomized forcing. As the model facilitates long transient simulations, we envisage its use in exploratory studies of stochastic forcing and perturbed parameterizations, thus complementing studies with comprehensive GCMs.

1 Introduction

Observational records of climate and proxies are limited in both time and space (Schmidt et al., 2014; Hersbach et al., 2015). As such, climate models are vital for climate research since they can fill in gaps left by observational records. Such gaps exist both during the instrumental era and in paleoclimate, e.g., with respect to transitions and hysteresis in climate (e.g., Rahmstorf et al., 2005; Abe-Ouchi et al., 2013; Ghil, 2015; Fraedrich et al., 2016; Stap et al., 2017) or variability in the climate system (Huybers and Curry, 2006; Laepple and Huybers, 2014; Rehfeld et al., 2020). They allow for the study of questions about past changes in climate like the Holocene temperature conundrum (Liu et al., 2014) or the response to changes in boundary conditions, thus improving the understanding of the past, present, and future states of the climate system (Braconnot et al., 2012; Schmidt et al., 2014). To provide confidence regarding their results, large ensembles are necessary (Collins et al., 2013; Schmidt et al., 2014). However, due to constraints in computational resources, to date state-of-the-art Earth system or general circulation models (GCMs) can only offer transient simulations for limited timescales (Smith and Gregory, 2012; Liu et al., 2009; Stap et al., 2017; Otto-Bliesner et al., 2017) unless they restrict the complexity of the models (e.g., Goosse et al., 2010; Smith and Gregory, 2012; Andres and Tarasov, 2019; Stap et al., 2018). Long timescales in particular have been simulated by taking snapshots of equilibrium states instead (e.g., Singarayer and Valdes, 2010). In this context, low-complexity models such as energy balance models (EBMs) are viable alternatives since their computational efficiency allows the running of fast and repeated simulations.

EBMs compute the temperature field based on the difference between incoming solar and outgoing terrestrial radia-

tion. Unlike GCMs, EBMs do not contain parameterizations of the whole climate system and instead focus on the Earth's energy balance and parameterization of energy fluxes. They were among the first climate models developed and helped build the current understanding of the climate system (North and Kim, 2017). However, many of the early studies cannot be reproduced today, since the models were not preserved (Zhuang et al., 2017a). This motivated Zhuang et al. (2017a) to publish their model code for a two-dimensional EBM, which serves as the basis of this study.

Arrhenius (1896) was the first to estimate the Earth's energy balance in order to calculate its global mean temperature response to the influence of greenhouse gases. In the late 1960s, Budyko (1968) and Sellers (1969) introduced EBMs as one-dimensional models with latitudinal energy transport and investigated their response to changes in solar radiation. Sellers (1969) contributed an early study on the effect of human greenhouse gas emissions. North and Coakley (1979) added a seasonal cycle and later introduced a longitudinal dimension (North et al., 1980, 1983) so that their EBM could simulate a two-dimensional representation of global temperatures.

North et al. (1983) incorporated seasonality in the two-dimensional EBM and investigated the lag of the annual cycle with respect to solar forcing and how seasonality interacts with changes in orbital parameters. They showed how configurations of orbital parameters can lead to colder seasons in the Northern Hemisphere that favor glaciation. Others studied seasonality under varying continental configurations throughout Earth's history (Crowley et al., 1986; Mengel et al., 1988; Crowley et al., 1989; Hyde et al., 1989; Short et al., 1991). For example, examining the effect of changing land–sea distributions on the seasonal cycle, Crowley et al. (1986) and Crowley et al. (1989) found considerable differences in summer temperatures throughout the ages depending on land–sea distribution. Huang and Bowman (1992) conducted a similar study, focusing on the seasonal snowline instability and the response of the ice albedo feedback to variations in land–sea distribution and seasonality.

Many of these studies examined seasonality in conjunction with not just the snowline instability but the whole cryosphere, whose influence on Earth's climate has long been of interest. Sellers (1969) had already examined the idea of an ice-free Arctic in response to continued and increasing greenhouse gas emissions, and numerous studies followed that investigated the small ice cap instability and its relation to the formation of the Antarctic ice sheet (North et al., 1983; North, 1984; Mengel et al., 1988; Crowley and North, 1990; Huang and Bowman, 1992). To examine the small ice cap instability, they simulated whether snow and ice accumulates or melts based on midsummer temperatures. The albedo field was then changed in accordance with the simulated snowline. In this way, these studies were able to simulate the response of the ice cover to changes in solar insolation or greenhouse gas concentration.

Crowley and North (1990) additionally analyzed transitions into glacial states. In fact, several of the studies found abrupt transitions and hysteresis in the creation and disappearance of ice sheets (North et al., 1983; Mengel et al., 1988; Crowley and North, 1990; Zhuang et al., 2014), as well as the possibility of multiple stable states with and without ice co-existing (North et al., 1983; North, 1984; Mengel et al., 1988). These studies built on both Budyko (1968) and Sellers (1969), who already demonstrated that EBMs can be used to study hysteresis in the climate system. They found that changes in solar radiation might lead to transitions between stable states and discussed under which conditions an ice-covered Earth could be possible. All of these studies suggest that there are multiple stable states in the climate system that climate can transition between.

Crowley and North (1988) and Crowley and Hyde (2008) relate such transitions to increasing variability in the climate system. They argue that an increase in variability indicates a predisposition of Earth's climate towards tipping into a different state in response to small changes in boundary conditions. Ghil (2015) investigated climate variability and sensitivity from a system dynamics standpoint to understand variability on all timescales. For this, the study used zero- to two-dimensional EBMs as part of a model hierarchy to show how multiple equilibria arise and how internal variability interacts with forcings. Dortmans et al. (2019) also found evidence for bifurcations and possible transitions between stable states in relation to glaciation. They connected these transitions to particular points in the geological past using a zero-dimensional EBM. At these points, they investigated possible mechanisms that explain sudden glaciations or unusually warm periods, focusing mainly on CO_2, water vapor, and oceanic and atmospheric heat transport. Overall, EBMs are able to simulate hysteresis and tipping points, whereas CMIP5 GCMs struggle to simulate such strong transitions, as shown for Arctic sea ice (Wagner and Eisenman, 2015; Bathiany et al., 2016), as well as the Atlantic meridional overturning circulation and other proposed tipping points (Bathiany et al., 2016).

Further studies of Earth's climatic history include the simulation of a possible climate on the supercontinent Pangaea during the early Late Permian (Crowley et al., 1989) and on Gondwana (Crowley et al., 1986, 1987). Hyde et al. (1989) simulated the past 18 000 years with snapshots taken every 3000 years, and Short et al. (1991) modeled the past 800 000 years in selected regions using snapshot simulations in response to orbital forcing. For such long-term studies in particular, transient simulations provide insights into the build-up of forcings, feedbacks, and their interactions in the climate system. With respect to transient simulations, Harvey and Schneider (1985) presented a pioneering study, which used a hemispherically averaged box-type EBM to examine effects of solar, CO_2, and volcanic forcing. More recently, Stap et al. (2017) created transient simulations of the past 38 million years to study the impact of ice sheets on the rela-

tionship between temperature and CO_2 by coupling an one-dimensional EBM to an one-dimensional ice sheet model.

Moreover, EBMs have been applied to study scaling of the power spectrum of temperature, for example Rypdal et al. (2015) generalized a two-dimensional stochastic–diffusive EBM to analyze scaling features. The temperature response to stochastic forcing has also been examined (e.g., Kim and North, 1991; Rypdal et al., 2015), while Fraedrich et al. (2016) and Meyer et al. (2018) used EBMs to analyze memory effects in the climate system. Dommenget et al. (2019) built a large database of more than 1300 experiments with a two-dimensional EBM. They explored how different feedbacks and components in the climate system, such as the ice–albedo feedback, clouds, or the hydrological cycle respond to changes in forcing. They also produced projections of future climate, as did Hébert et al. (2020).

Many of these studies provide robust results only when using large ensembles of runs that represent many possible realizations of climate or explore the parameter space of the models.

We present TransEBM v. 1.0, a transient extension of the model by Zhuang et al. (2017a), a two-dimensional EBM that simulates the surface temperature field based on the Earth's energy balance. Briefly, the model works as follows: local albedo determines how much of the incoming radiation is reflected back into space by both atmosphere and surface. The atmosphere is averaged with no local differences, and only the surface temperature is computed. The surface temperature changes according to the absorbed energy and the surface's heat capacity. The energy that enters the system is transported so as to balance the temperature gradients caused by latitudinal differences in incoming radiation and the inhomogeneous distribution of land and ice masses. The Earth then emits long-wave radiation depending on the local surface temperature.

The model is presented in detail in Sect. 2, which covers the model description (Sect. 2.1), the parameterization and forcings (Sect. 2.2), the code structure and benchmarking (Sect. 2.3), and the parameter tuning (Sect. 2.4). Section 3 validates the model by starting with the evaluation of the tuning using reanalysis data (Sect. 3.1). Furthermore, it presents a sensitivity test with respect to the forcings (Sect. 3.2) and the results of a past millennium simulation, which is compared to GCM simulations and a multi-proxy reconstruction (Sect. 3.3). We present a short tutorial in Sect. 4, followed by a discussion of the model's strengths and weaknesses, as well as potential applications and extensions, in Sect. 5.

2 TransEBM

2.1 Model description

TransEBM computes a surface temperature field $T(\hat{\boldsymbol{r}}, t)$ that depends on time t and two spatial dimensions in latitude and longitude $\hat{\boldsymbol{r}} = (\theta, \phi)$ according to

$$C(\hat{\boldsymbol{r}})\frac{\partial T}{\partial t} + A + B \cdot T = \nabla \cdot (D(\theta)\nabla T) + S_0 \cdot S_{\mathrm{F}}(\hat{\boldsymbol{r}}, t)\,(1 - a(\hat{\boldsymbol{r}})), \quad (1)$$

with heat capacity $C(\hat{\boldsymbol{r}})$, outgoing long-wave radiation parameters A and B, heat transport coefficient $D(\hat{\boldsymbol{r}})$, solar constant S_0, insolation forcing $S_{\mathrm{F}}(\hat{\boldsymbol{r}}, t)$, and albedo $a(\hat{\boldsymbol{r}})$. As detailed in Zhuang et al. (2017a), this elliptic, partial differential equation is solved for 48 time steps per year using the full multigrid method, which dampens both high- and low-frequency errors (Briggs et al., 2000). As such, four time steps are computed per month for a model year starting in January. Changing the resolution in time is possible but requires adapting the model in several places, such as the seasonal insolation cycle, as well as pre- and post-processing.

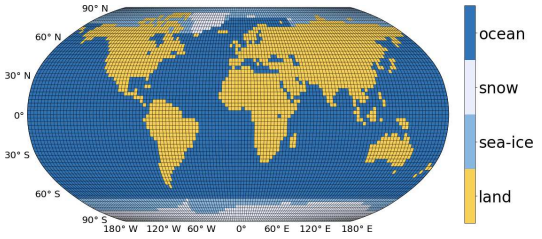


Figure 1. Land, sea, and ice configuration with T42 resolution for 1950 CE as used in modern model runs and given by Zhuang et al. (2017a). Four surface types are differentiated: ocean, land surface, sea ice, and land-based ice or snow.

The model is solved on a T42 Gaussian grid with 128 points in longitude and 64 in latitude. In principle, other resolutions are possible, but since the resolution is defined at different points in the code, any change to the resolution needs to be carefully implemented. Each grid point is categorized as one of four surface types – ocean, land surface, sea-ice, or land-based ice or snow – as shown in Fig. 1 for a modern land–sea configuration. The geography does not contain information about altitude and remains constant throughout a model run. Transient runs with changing land, sea, and especially ice configuration are possible by regularly restarting the model with updated boundary conditions using the climatology from the previous run. Figures 6 and 7 present the structure of the model code, which will be discussed in Sect. 2.3.

2.2 Forcings and parameterization

Four broad categories of parameters exist: heat capacity, albedo, thermal conductivity, and outgoing radiation. Figure 2 shows a conceptualized view of the energy transport of the model with all model parameters and forcings.

The simulated heat capacity changes with the surface type as listed in Table 1 and thus depends on the land, sea, and ice

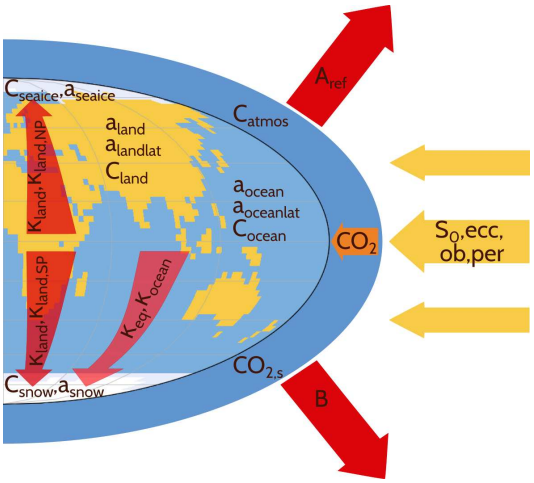


Figure 2. Conceptual view of the forcings, energy transport, and parameterizations in TransEBM. Solar insolation S_0, the orbital parameters, – eccentricity, obliquity, precession of the perihelion – CO_2, land, sea, and ice distribution serve as forcings. The surface types differ by albedo a, which affects the insolation field, and heat capacity C. The energy from incoming solar radiation is transported poleward according to the thermal conductivities κ. The outgoing radiation is calculated based on the parameters A, B, and the response to CO_2 scaled by $CO_{2,s}$.

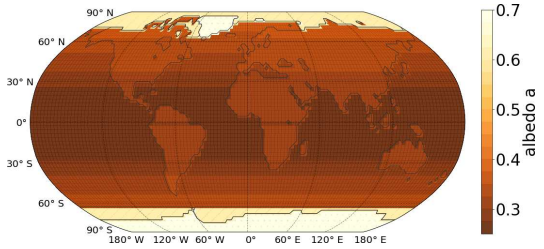


Figure 3. Albedo field for the 1950 CE map shown in Fig. 1. Albedo differs by surface type. Over land and ocean it increases with latitude.

configuration. Due to this dependence, the heat capacity field remains constant during a single model run. For each grid point, the total heat capacity equals the sum of the heat capacity for its surface type and that of the atmospheric column above. For the atmospheric column an average scale height of 7.6 km is assumed. Zhuang et al. (2017a) base the oceanic heat capacity on the assumption of a mixed layer of about 70 m. This can be changed by tuning the oceanic heat capacity as presented in Sect. 2.4. The tuned heat capacity of TransEBM corresponds to a mixed layer of about 96 m.

Local relaxation timescales τ are determined by the total heat capacities $C_{\mathrm{tot}} = C_{\mathrm{atmos}} + C_i$ in the atmospheric column over surface type i and the outgoing radiation coefficient B as $\tau = C \cdot B^{-1}$. Table 2 lists the resulting relaxation times and shows that the longest timescales are simulated for the ocean at about 6 years. Therefore, the contribution of the deep ocean that impacts timescales on the order of 10^2 years and longer (Rohling et al., 2012) is not modeled. For the other surface types, relaxation times are far smaller and range from about 2 weeks to 1 month. Overall, the heat capacity parameters approximate the effect of energy storage in the climate system since TransEBM does not model energy storage explicitly.

It is possible to differentiate surface types beyond those currently used in the model. In fact, additional surface types to distinguish between water reservoirs like the Pacific, Atlantic, and Indian oceans, as well as the Mediterranean Sea and inland seas exist for the parameterization of heat capacity. Using them, the impact of the deep ocean could be modeled more accurately. However, other parts of the parameterization like the albedo are not yet adapted for a larger number of surface types.

The albedo depends on the geography and gets initialized at the beginning of a model run and then remains constant. Table 1 lists the albedo values that are assigned to the different surface types. The albedo fields of land surface and ocean are latitude-dependent. Figure 3 shows the albedo distribution of a present-day continental configuration as seen in Fig. 1. The highest albedo values are in snow-covered areas and the lowest are close to the Equator. Modeled albedo values agree with the ranges found in literature (cf. Ruddiman, 2014).

Heat transport due to sensible and latent heat, as well as oceanic heat transport across isotherms, is described using a heat transport coefficient

$$D = \begin{cases} (\kappa_{\mathrm{equ}} - \kappa_{\mathrm{ocean}})\sin^5(\theta) + \kappa_{\mathrm{ocean}}, \\ \quad \text{oceans} \\ (\kappa_{\mathrm{land}} - \kappa_{\mathrm{land,NP}})\sin^5(\theta) + \kappa_{\mathrm{land,NP}}, \\ \quad \text{Northern Hemisphere} \\ (\kappa_{\mathrm{land}} - \kappa_{\mathrm{land,SP}})\sin^5(\theta) + \kappa_{\mathrm{land,SP}}, \\ \quad \text{Southern Hemisphere}, \end{cases} \tag{2}$$

which is constant throughout one model run. Table 1 lists the values for the thermal conductivities κ, on which D depends. Figure 2 presents a conceptual view of the heat transport in the model. Oceanic heat transport is distinguished from heat transport over land and ice-covered areas. Over land and ice-covered areas, the Northern and Southern hemispheres are treated separately so as to reproduce differences in observed temperature patterns due to the asymmetric distribution and sizes of the land masses. This difference increases towards higher latitudes. Figure 4 shows the resulting latitude-dependent pattern of the heat transport coefficient. Heat transport is large close to the Equator and declines towards higher latitudes, where it evens out.

Table 1. List of all model parameters for the model by Zhuang et al. (2017b) and TransEBM v. 1.0 after the tuning described in Sect. 2.4. Albedo over land and ocean depends on latitude θ via the second Legendre polynomial $P_2(\theta) = 0.5 \cdot (3\sin^2(\theta) - 1)$.

Parameter	Zhuang et al. (2017b)	TransEBM v. 1.0
Heat capacity C [10^6 J K^{-1} m^{-2}]		
C_{land}	1.87	1.87
C_{ocean}	293.02	394.47
C_{seaice}	4.83	4.83
C_{snow}	1.52	1.52
C_{atmos}	4.76	0.79
Albedo a		
$a_{\text{land}} + a_{\text{landlat}}\, P_2$	$0.30 + 0.09\, P_2 \in [0.21, 0.48]$	$0.32 + 0.05\, P_2 \in [0.27, 0.37]$
$a_{\text{ocean}} + a_{\text{oceanlat}}\, P_2$	$0.29 + 0.09\, P_2 \in [0.20, 0.47]$	$0.289 + 0.08\, P_2 \in [0.209, 0.369]$
a_{seaice}	0.60	0.60
a_{snow}	0.70	0.70
Thermal conductivity [W °C^{-1}]		
κ_{equ}	0.65	0.90
κ_{ocean}	0.40	0.40
κ_{land}	0.65	0.65
$\kappa_{\text{land,NP}}$	0.28	0.45
$\kappa_{\text{land,SP}}$	0.20	0.12
Outgoing radiation		
A [W m^{-2}]	210.3	210.2
B [W m^{-2} °C^{-1}]	2.15	2.13
$CO_{2,\text{ref}}$ [ppm]	315	315
$CO_{2,\text{s}}$	5.35	5.35

Table 2. Total heat capacities $C_{\text{tot}} = C_{\text{atmos}} + C_i$ and resulting relaxation times τ for surface type i.

Surface type	C_{tot} [10^6 J K^{-1} m^{-2}]	τ
Land	2.66	14.45 d
Ocean	395.26	5.88 years
Sea ice	5.62	30.54 d
Snow	2.31	12.55 d

The parameterization of outgoing radiation as $A + B \cdot T$ follows Budyko (1968) and uses phenomenological coefficients A and B. Both A and B were estimated from satellite observations in previous studies, but have large uncertainties attached to them (cf. Kyle et al., 1995; North and Kim, 2017). A captures the direct radiative effect of atmospheric CO_2 on radiative forcing. Following Arrhenius (1896), Myhre et al. (1998) established its dependence on CO_2 as

$$A = A_{\text{ref}} - CO_{2,\text{s}} \cdot \log\left(\frac{CO_2}{CO_{2,\text{ref}}}\right), \tag{3}$$

here with reference values $A_{\text{ref}} = 210.2\,\text{W m}^{-2}$ and $CO_{2,\text{ref}} = 315$ ppm and scaling factor $CO_{2,\text{s}} = 5.35$. A_{ref} and $CO_{2,\text{ref}}$ are fixed parameters that are excluded from tuning.

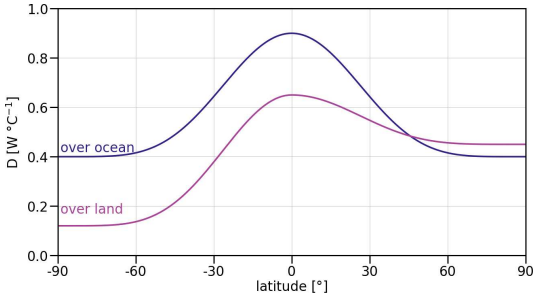


Figure 4. Effective heat transport coefficient D as modeled in TransEBM v. 1.0 and computed from the thermal conductivities κ (cf. Eq. 2). Oceanic heat transport is treated separately from that over land and ice-covered areas. For the latter, the Northern and Southern hemispheres are separated, with larger heat transport coefficients in the north.

Changing them requires the model to be re-tuned due to their influence on the other radiative parameters. Overall, A follows the CO_2 input and can change with each model year.

B describes the radiative damping due to climate feedbacks. It is set to $B = 2.15\,\text{W m}^{-2}\,\text{°C}^{-1}$, well within the

range of satellite measurements (Kyle et al., 1995; North and Kim, 2017). Table 1 summarizes the parameterization of outgoing radiation, which remains constant during a model run.

Insolation is the only component of the parameterization that changes for each time step with the seasonal cycle. It depends on the latitude- and time-dependent insolation $S(\theta, t)$ and the albedo $a(\hat{\boldsymbol{r}})$ via

$$S_{\mathrm{F}} = S(\theta, t)(1 - a(\hat{\boldsymbol{r}})). \tag{4}$$

Berger (1978) relates the insolation to the solar constant at the top of the atmosphere S_0 and the orbital configuration, which are both transient forcings in the model, as

$$S(\theta, t) = \begin{cases} 0, \\ \quad \text{latitudes without sunrise} \\ \frac{S_0}{\rho^2} \sin\theta \sin\delta, \\ \quad \text{latitudes without sunset} \\ \frac{S_0}{\rho^2 \pi}(H_0 \sin\theta \sin\delta + \cos\theta \cos\delta \sin H_0), \\ \quad \text{else.} \end{cases} \tag{5}$$

Here, the orbital configuration is given by declination δ, absolute hour angle of the Sun at sunrise and sunset H_0, and the normalized Earth–Sun distance ρ is as follows:

$$\rho = \left(\frac{1 + e\cos\nu}{1 - e^2} \right)^2, \tag{6}$$

which depends on eccentricity e and the positional angle of Earth on its orbit around the Sun ν as given counterclockwise from perihelion.

Figure 5 shows the relationship between orbital forcing, insolation, and the modeled temperature response. For the three orbital forcings eccentricity, obliquity, and precession, it shows the difference in insolation and temperature during northern summer between the possible extremes of the orbital forcings as listed in Table 3. When examining one orbital parameter, the other ones were set to their listed average values. The simulations were run in equilibrium mode with the tuned model (cf. Sect. 2.4), the land–sea mask as shown in Fig. 1, and constant $CO_2 = 315\,\mathrm{ppm}$ and $S_0 = 1362\,\mathrm{W\,m^{-2}}$.

Precession of the perihelion has the largest effect on the distribution of solar insolation and the resulting changes in temperatures, followed by obliquity. The effect of the changing eccentricity is the smallest. Larger eccentricity results in a global decrease in temperatures between May and July when perihelion is set to occur at vernal equinox in March. This means that the Earth will be farther from the Sun in northern summer with higher eccentricities. The effect is larger in the Northern Hemisphere, particularly in midlatitudes.

A large obliquity between May and July decreases insolation in the Southern Hemisphere and leads to an increase in the Northern Hemisphere. In the Southern Hemisphere, this effect is strongest in the midlatitudes, but disappears over high latitudes. In the Northern Hemispheres, insolation increases up to around 60° N, and then stabilizes in the high latitudes, leading to the largest increase in temperatures at high latitudes and over land in midlatitudes.

For the high precession simulation, summer solstice and perihelion coincide, leading to increased insolation and temperatures between May and July. Overall, the strongest increase in insolation occurs in northern low latitudes and midlatitudes and the highest increase in temperatures over Asia. Temperatures grow large over land and in the northern high latitudes. Appendix B examines the effect of the orbital parameters on insolation and temperatures further by repeating this analysis for November to January and annual averages. For all orbital parameters, the insolation and temperature patterns even out over the ice-covered regions near the poles.

2.3 Model extension and benchmarking

Based on the model code by Zhuang et al. (2017b), the model formulation was extended to enable transient and parallel runs, as well as transient, user-defined forcings for CO_2, solar insolation, and the orbital configuration. The land, sea, and ice configuration can be changed within a simulation via model restarts at prescribed intervals. Additionally, the model configuration was overhauled so as to allow as much customization for experiments as possible. Table 4 summarizes the major expansions of TransEBM v. 1.0 with respect to the model by Zhuang et al. (2017a). A default configuration is provided with the model code. It contains detailed explanations for each variable that expand on the short tutorial provided in Sect. 4.

The model has a preprocessing component visualized in Fig. 6 that imports the experiment's configuration, prepares the geography, and computes the albedo field. Based on this, the main component sets up the specified experiment, key steps of which are presented in Fig. 7. The initial setup is then integrated until either equilibrium or the user-defined number of years is reached. Finally, the output is written to netCDF files.

Only the land, sea, and ice distribution – described in the map section of the configuration – is constant throughout a model run, all other forcings can be transient. For long-term simulations, changing it is essential. This can be achieved by restarting the model based on the restart configuration that is continually updated while running. Such restarts preserve the state of the model and allow for the user to manipulate the run as desired and restart it from any modeled year for which restart information was written. This restart information consists of the temperature and insolation fields in the last time step of the year from which to restart. In a restart, these fields are used as the input for Eq. (1), which is updated and solved using the changed forcings.

The model is written in Fortran 90 and uses bash scripts and CMAKE and as such should run on standard Linux or

Table 3. Orbital parameter extremes used to examine the maximal influence that the parameters have on the modeled distributions of insolation and temperature as shown in Fig. 5. To examine the effect of one parameter, the other two were set to average values and an equilibrium simulation with land, sea, and ice configuration as in Fig. 1 and constant $CO_2 = 315$ ppm and $S_0 = 1362\,\mathrm{W\,m^{-2}}$ was run.

Parameter	Eccentricity	Obliquity	Angle of perihelion*
High	0.06	$24.5° \approx 0.427606$	$\frac{3}{2}\pi \approx 4.712389$
Low	0.005	$22.1° \approx 0.385718$	$\frac{1}{2}\pi \approx 1.570796$
Average	0.0325	0.398808	0

* Angle of perihelion is given from vernal equinox in TransEBM, so that 0°, for example, means perihelion lies at the vernal equinox.

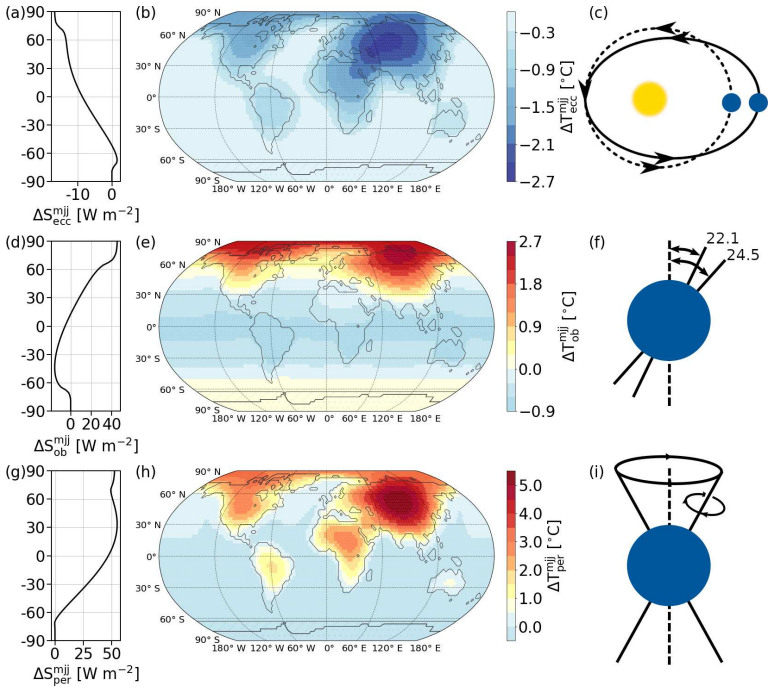


Figure 5. Difference that high and low orbital parameters as listed in Table 3 cause in solar insolation (left) and temperature field (middle) for the averages of the months May, June, and July (MJJ). A schematic illustration of the orbital parameters is added to the right. Equilibrium simulations with land, sea, and ice configuration as in Fig. 1 and constant $CO_2 = 315$ ppm and $S_0 = 1362\,\mathrm{W\,m^{-2}}$. **(a–c)** Eccentricity, **(d–f)** obliquity, and **(g–i)** precession of the position of perihelion. Precession shows the largest impact, followed by obliquity. High eccentricity results in overall lower insolation and temperatures; high precession overall resulted in increased insolation and temperatures. For eccentricity, decreased insolation in the north is found; for precession, on the other hand, insolation increases in the north. High obliquity results in increased insolation in the north and decreased insolation in the south. Precession acts strongest in lower and middle northern latitudes; obliquity and eccentricity, on the other hand, affect the high latitudes the most.

Unix machines. In equilibrium mode, it takes between 35 and 40 years to reach a steady state. For each simulated year, a minimum of 13 MB of output is created. On a 64-bit Ubuntu 18.04.2 LTS machine with a gfortran-7 compiler, simulating one model year takes 1.5 s on a single core of an Intel i5 – 8600K 3.6 GHz processor. This translates into the calculation of 57 600 years in a single day.

2.4 Tuning methodology

The model is tuned to the 1960–1989 climatology from ERA-20CM reanalysis (Hersbach et al., 2015). Target metrics are global mean temperature (GMT), the root-mean-square error of the latitudinal temperature distribution ($\mathrm{RMSE_{lat}}$) and seasonal temperature distribution in the Northern Hemisphere ($\mathrm{RMSE_{seas,nh}}$) and Southern Hemi-

Table 4. Features of the revised model, TransEBM v. 1.0, in comparison to the model by Zhuang et al. (2017a).

	Zhuang et al. (2017a)	TransEBM v. 1.0
Run types	equilibrium only	transient & equilibrium
Restarts	no	yes
Forcings	constant only	user-defined, transient
Map	defined in one file	land–sea mask and ice separate
Configuration	spread out in model code	one external configuration file
Output	T for the last year only	complete time series of T, C, and S possible
Parallelization	no	yes

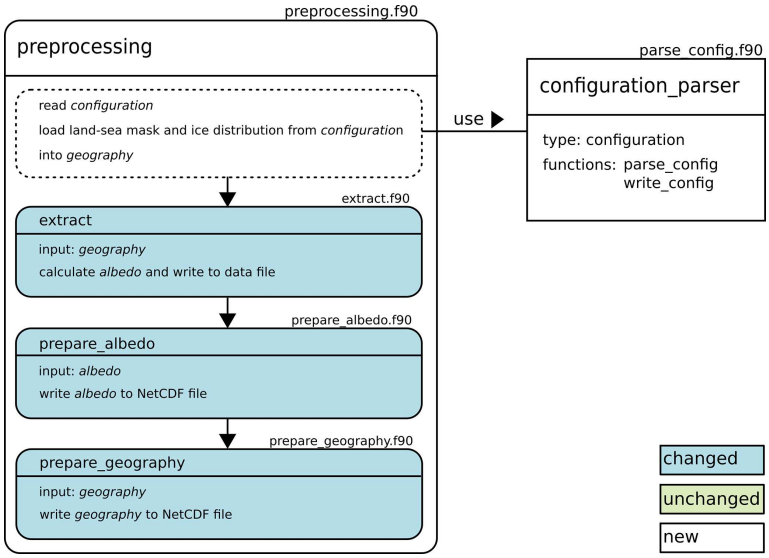


Figure 6. Preprocessing component of the model with new, changed, and unchanged features with respect to the EBM by Zhuang et al. (2017a) marked.

sphere ($RMSE_{seas,sh}$). An ensemble of simulations explores the parameter space with respect to these four metrics as shown in Figs. 8 and 9. These simulations model the same time period with 100 years beforehand to ensure the decay of the transients. The land, sea, and ice configuration was constant and as shown in Fig. 1. Orbital parameters were also constant during the simulations with eccentricity (ecc = 0.016740), obliquity (ob = 0.409253), and longitude of perihelion (per = 1.796257), while CO_2 followed Köhler et al. (2017) and S_0 followed Coddington et al. (2016).

To get a complete picture of how the parameters affect the modeled temperatures, parameter values that disagree with observations or physics were explored as well. During the actual tuning, however, only configurations that agree with observations and physics were considered.

Albedo changes have a strong effect on temperature means and extremes and as such affect all metrics as seen in Fig. 8. They do not influence the shapes of the seasonal temperature distribution though, since the albedo field is constant throughout the model year and as such affects the insolation at all time steps the same way. The thermal conductivities κ describing the heat transport primarily alter the seasonality and Equator-to-pole gradient. Decreasing them sharpens the latitudinal distribution and increases the seasonality and vice versa for increased thermal conductivities. In comparison to albedo, they have a small effect on global mean temperature and produce differences of about 5 °C over the whole explored parameter space. Changing the heat capacities affects the equilibration time and the seasonal cycle, both in regard to the lag of the temperature response to the forcing and the absolute temperature difference between warmest and coldest month. Figure 9 shows that the outgoing radiation parameters A_{ref} and B influence GMT, the latitudinal distribution, and seasonality. By comparison, $CO_{2,s}$ only has a small effect on the seasonal distribution.

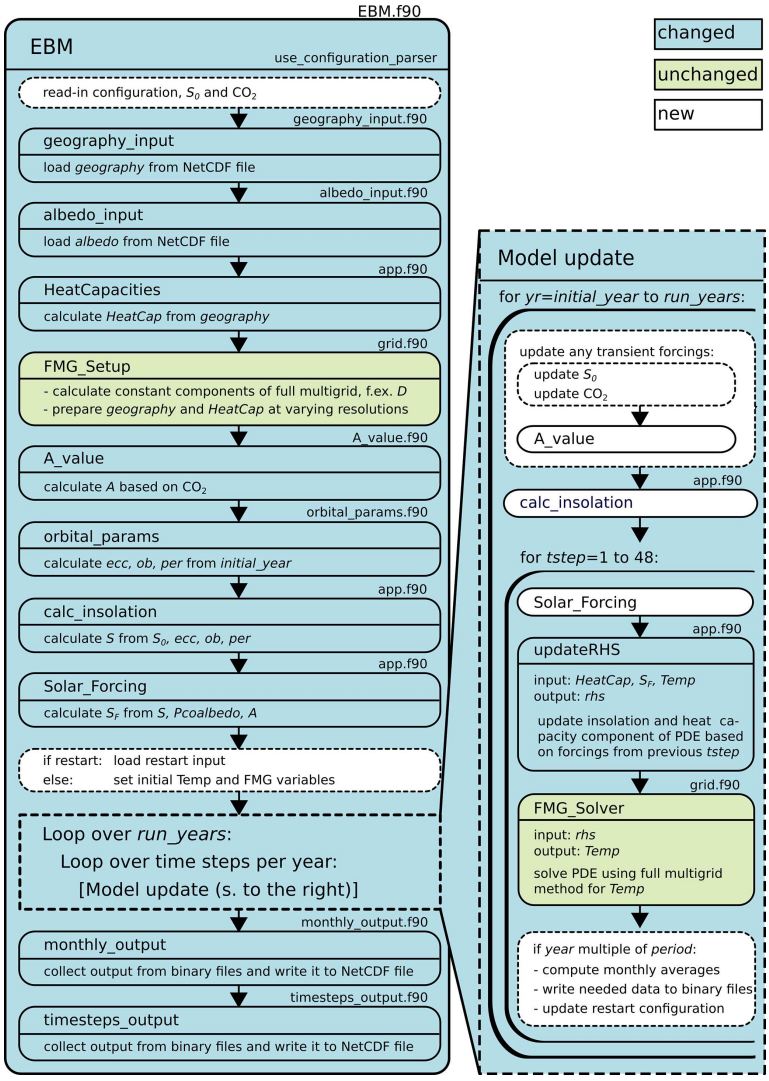


Figure 7. Main model component with new, changed, and unchanged features with respect to the EBM by Zhuang et al. (2017a) marked.

Table 5 summarizes which parameters affect which metric. It shows that only the seasonal distribution can be tuned without affecting the other metrics. Minimizing the root-mean-square errors was prioritized over agreement with the GMT from ERA-20CM $\mathrm{GMT_{ERA}} = 13.27\,^{\circ}\mathrm{C}$. With respect to GMT, any tuning options were only discarded if they produced a change in GMT of 2° or more. The latitudinal and seasonal distributions were prioritized to ensure that the spatially resolved temperature fields show realistic features. Moreover, the latitudinal and seasonal profiles deviated considerably from observations in areas of interest (e.g., the polar regions), whereas the GMT was already very close to that found in the ERA climatology. However, depending on the envisioned applications, the relative weighting of the tuning metrics could change. Since $\mathrm{RMSE_{lat}}$ was the highest, with significant disagreement observed in the polar regions, it was the initial focus of the tuning. Table 6 shows the metrics before and after tuning. The agreement between the latitudinal distributions in ERA-20CM and TransEBM and the seasonal distributions in the Southern Hemisphere improved. $\mathrm{RMSE_{seas,nh}}$ increased slightly, but since this produced a large improvement in the other two RMSEs, this increase

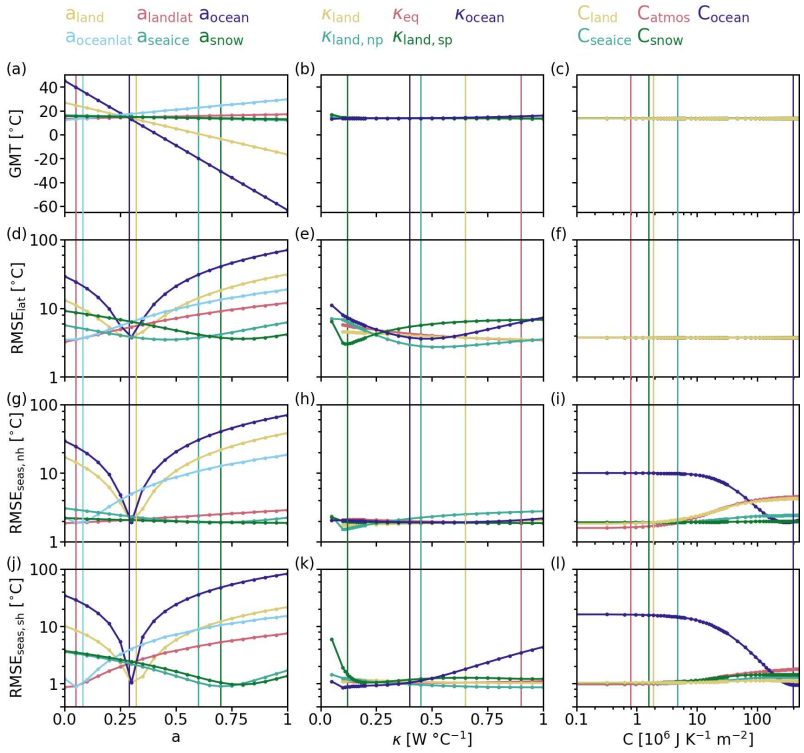


Figure 8. Effect of changes to the parameterization of albedo (first column), heat transport (second column), and heat capacity (third column) on the GMT and the RMSE of the latitudinal ($RMSE_{lat}$) and seasonal temperature profiles in both hemispheres ($RMSE_{seas,nh}$ and $RMSE_{seas,sh}$) with respect to ERA-20CM reanalysis (Hersbach et al., 2015). **(a–c)** GMT, **(d–f)** $RMSE_{lat}$, **(g–i)** $RMSE_{seas,nh}$, and **(j–l)** $RMSE_{seas,sh}$ with all specific parameters listed at the top of each column. Simulations and reanalysis cover 1960–1989 CE. The simulations used the land, sea, and ice configuration shown in Fig. 1 and constant orbital parameters: eccentricity (ecc = 0.016740), obliquity (ob = 0.409253), and longitude of perihelion (per = 1.796257). CO_2 followed Köhler et al. (2017), and S_0 values followed Coddington et al. (2016). Vertical lines indicate the chosen values after tuning. Albedo affects all tuning metrics, while thermal conductivities barely change GMT and heat capacities only manipulate the seasonal temperature distributions.

Table 5. List of the parameter groups and which of the tuning metrics (GMT, $RMSE_{lat}$, and $RMSE_{seas}$) they affect based on the results shown in Fig. 8 and 9. Bracketed markers (x) represent a notably smaller but not negligible influence by that parameter on the tuning metric in comparison to the other parameters.

	Affects		
Parameter	GMT	$RMSE_{lat}$	$RMSE_{seas}$
Heat capacity C			x
Albedo a	x	x	x
Thermal conductivity κ	(x)	x	x
A_{ref}	x	x	x
B	x	x	x
$CO_{2,s}$			(x)

was considered acceptable. Although GMT was barely considered during the tuning, the final GMT is closer to the target.

3 Model validation

3.1 Tuning validation and comparison to EBM by Zhuang et al. (2017a)

The model was tuned using the 1960–1989 climatology of a simulation of the 20th century (1900–1989). As a first validation of that tuning, we compare that whole simulation to ERA-20CM reanalysis data (Hersbach et al., 2015) and a simulation with the EBM based on the parameterization by Zhuang et al. (2017a). Both simulations used the constant modern land, sea, and ice configuration shown in Fig. 1, with

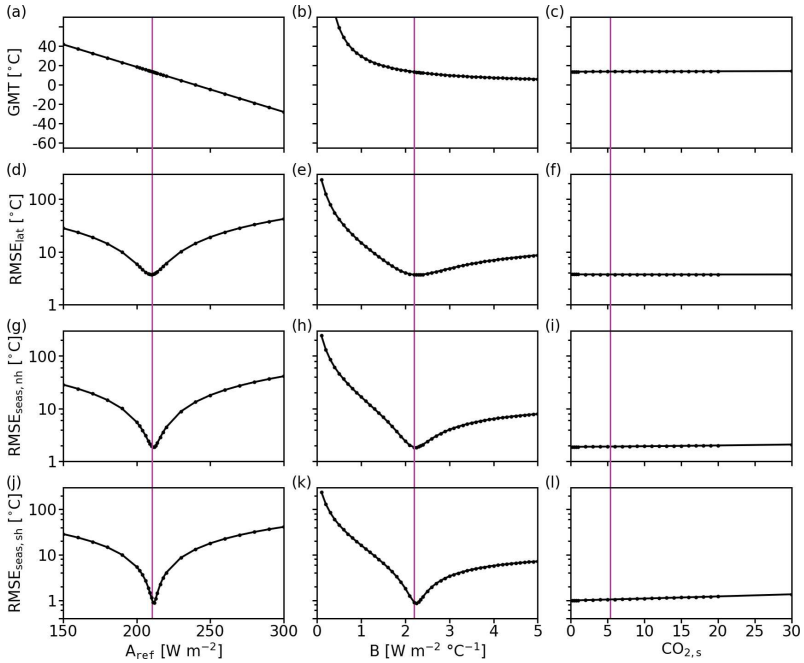


Figure 9. Effect of changes to the parameterization of the outgoing radiation on the GMT and the RMSEs of the latitudinal ($RMSE_{lat}$) and seasonal temperature profiles in both hemispheres ($RMSE_{seas,nh}$ and $RMSE_{seas,sh}$) with respect to ERA-20CM reanalysis (Hersbach et al., 2015). **(a–c)** GMT, **(d–f)** $RMSE_{lat}$, **(g–i)** $RMSE_{seas,nh}$, and **(j–l)** $RMSE_{seas,sh}$. The first column shows results for A_{ref}, the second shows results for B, and the third shows results for $CO_{2,s}$. Simulations and reanalysis cover 1960–1989 CE. The simulations used the land, sea, and ice configuration shown in Fig. 1 and constant orbital parameters: eccentricity (ecc = 0.016740), obliquity (ob = 0.409253), and longitude of perihelion (per = 1.796257). CO_2 followed Köhler et al. (2017), and S_0 values followed Coddington et al. (2016). Vertical lines indicate the chosen values after tuning. GMT reaches up to 257 °C for B; for a better comparison to the other parameters, the full range is not shown in panel **(b)**. Both A and B affect all tuning metrics, whereas changes in $CO_{2,s}$ barely make a difference.

Table 6. Tuning metrics GMT, $RMSE_{lat}$, and $RMSE_{seas}$ for each hemisphere before and after tuning. The metrics were evaluated for the years 1960–1989 CE against ERA-20CM ($GMT_{ERA} = 13.27$ °C). The simulations used the land, sea, and ice configuration shown in Fig. 1 and constant orbital parameters: eccentricity (ecc = 0.016740), obliquity (ob = 0.409253), and longitude of perihelion (per = 1.796257). CO_2 followed Köhler et al. (2017), and S_0 values followed Coddington et al. (2016). All metrics except $RMSE_{seas,nh}$ improved with tuning.

Metric	GMT	$RMSE_{lat}$	$RMSE_{seas,nh}$	$RMSE_{seas,sh}$
Before tuning [°C]	13.73	3.74	1.92	1.03
After tuning [°C]	13.18	1.13	2.04	0.59

transient CO_2 data from Köhler et al. (2017) and S_0 values from Coddington et al. (2016). To follow the evolution of CO_2 and S_0, the tuning used the extension to transient forcings of the model.

Figure 10 summarizes the simulations and compares them to the reanalysis. Figure 10a presents the time series of the annual temperature anomaly and shows that both parameterizations produce very similar anomalies. They follow the overall increasing trend in temperatures that the reanalysis shows. However, the reanalysis shows a slightly larger increasing trend. The trend in the EBM is solely in response to the increasing CO_2 values, so the difference indicates that the EBM is lacking positive feedback mechanisms that enhance the temperature response to increasing CO_2, leading to the EBMs underestimating greenhouse-gas-related warming. Unlike the reanalysis, they also show smoother changes and less inter-annual variability. The 11-year solar cycle, which was superimposed, dominates inter-annual changes in GMT in the simulations.

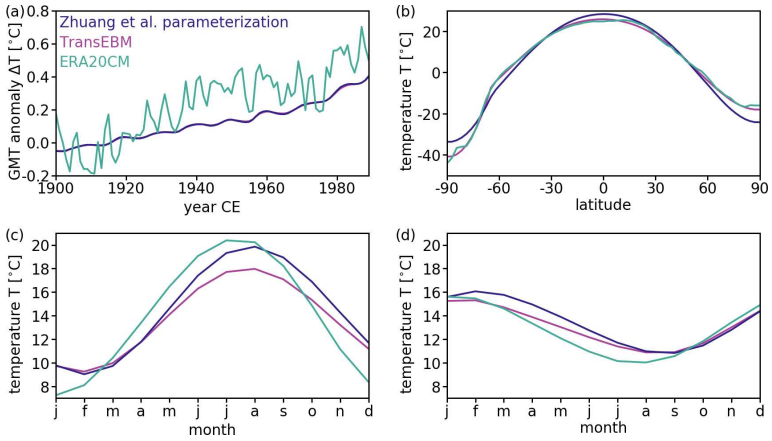


Figure 10. Comparison of a simulation of the 20th century using the parameterization by Zhuang et al. (2017a) with the tuned parameterization introduced in Sect. 2.4 and ERA-20CM reanalysis data (Hersbach et al., 2015). The simulations use the land, sea, and ice configuration from Fig. 1, CO_2 is set according to Köhler et al. (2017), S_0 follows Coddington et al. (2016) and orbital parameters Laskar et al. (2004). **(a)** Time series of the annual temperature anomaly from 1900–1989 CE with respect to the 1900–1929 climatological mean. The different parameterizations show little difference and capture the overall increasing trend, but the reanalysis shows stronger inter-annual variation. **(b)** The average latitudinal profile. Here, TransEBM agrees with reanalysis data better – especially in the polar regions and around the Equator. Seasonal temperature profiles in the Northern Hemisphere **(c)** and Southern Hemisphere **(d)** are both shifted in comparison to the reanalysis. Agreement is better in the Southern Hemisphere, TransEBM underestimates the seasonal cycle in the Northern Hemisphere.

Figure 10b presents the mean latitudinal temperature profiles. TransEBM agrees well with the reanalysis. In particular, it is able to simulate the dip in temperatures around the Equator and the temperatures in the polar regions better after tuning.

Figure 10c and d show the average seasonal temperature profiles in the Northern Hemisphere and Southern Hemisphere, respectively. The simulated average seasonal cycles lag in comparison to the reanalysis. The amplitude of the seasonal cycle is lower than that of the EBM by Zhuang et al. (2017a) as a result of the tuning.

For further analysis the phase lag was computed based on sinusoids fitted to the seasonal cycle for every grid box. The phase lag is computed as the difference between the occurrence of the extrema in the fit to the extrema in the insolation occurring on 21 June and 21 December. Figure 11 shows the resulting phase lags for TransEBM and ERA-20CM. Areas where the seasonal cycle does not follow a sine were excluded from the analysis. Appendix C provides details on the exclusion process. The resulting pattern in phase lags highlights that TransEBM produces larger lags. Over the oceans, the modeled phase lag is significantly higher than that in the reanalysis. Over land and ice, TransEBM simulates longer phase lags as well, but the difference to the reanalysis is smaller than in oceanic areas.

The amplitude of the seasonal cycle is, on average, smaller in the EBM simulations in comparison to the reanalysis, especially in the Northern Hemisphere as seen in Fig. 10 and further examined in Fig. 12. In TransEBM the largest amplitudes of up to 40 d occur in Siberia and Antarctica. In the reanalysis, the largest amplitudes can be found in Siberia as well, although they are shifted in position in comparison to TransEBM. The reanalysis indicates high amplitudes in northern Canada as well. Overall, TransEBM agrees with the patterns in the reanalysis, and simulates rising amplitudes with increasing continental size and towards central continental areas and small amplitudes over the oceans. The patterns are smoother and less variable in comparison to those found in the reanalysis, though.

Following North and Kim (2017, chap. 2), theoretical average values for the amplitude and phase lag of the seasonal cycle in TransEBM can be calculated from the Fourier transform of Eq. (1). This yields the solution

$$\tilde{T}_f = \frac{S_0}{4} \cdot \frac{a_g}{B} \cdot \frac{1}{\sqrt{2\pi i f \tau + 1}}, \tag{7}$$

with global mean albedo $a_g = 0.32$ for the present-day configuration shown in Fig. 3 and relaxation times τ for the specific surfaces as given in Table 2. The amplitude is then

$$|\tilde{T}_f| = \frac{S_0}{4} \cdot \frac{a_g}{B} \cdot \frac{1}{\sqrt{(2\pi f \tau)^2 + 1}}, \tag{8}$$

and the phase lag in radians is

$$\phi_f = \arctan(2\pi f \tau). \tag{9}$$

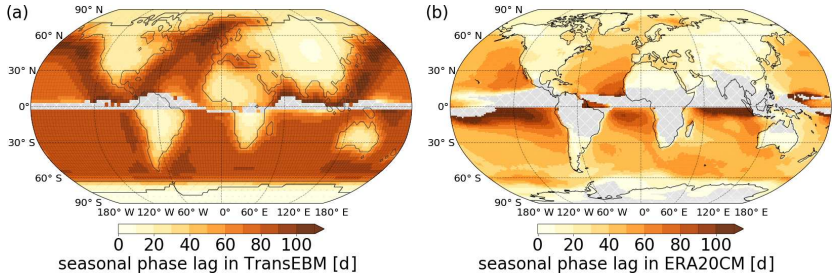


Figure 11. Phase lag of the seasonal cycle in surface temperature in days after solar forcing (solstice) in a 20th century run of TransEBM v. 1.0 and ERA-20CM. The TransEBM simulation is the same as in Fig. 10. The phase lag uses sinusoids fitted to the seasonal cycle. Areas where the seasonal cycle does not follow a clear sine were excluded from the analysis and are shown in grey. TransEBM simulates larger phase lags in comparison to the reanalysis (particularly over the oceans).

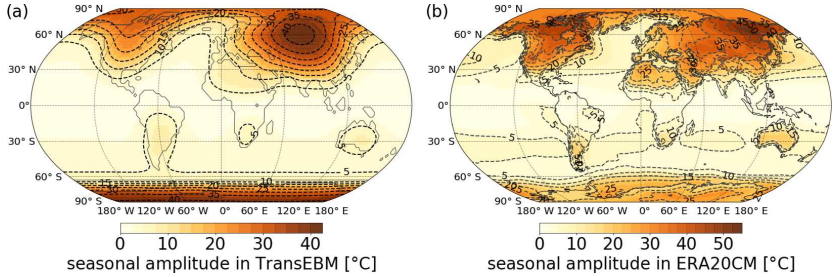


Figure 12. Amplitude of the seasonal cycle in surface temperature during the 20th century in TransEBM v. 1.0 and ERA-20CM based on the extrema in the sinusoids fitted to the seasonal cycles. Simulation as in Fig. 10. TransEBM shows similar patterns to the reanalysis but simulates smaller amplitudes.

Table 7 presents the resulting amplitudes and phase lags for an annual forcing frequency $f = 1/365\,\mathrm{d}$. The values corroborate the low oceanic seasonal amplitudes of only a couple of degrees and the much higher amplitudes over the other surfaces as shown in Fig. 12a. The phase lags also agree well with those found in the simulations (cf. Fig. 11a) with long average phase lags of about 90 d over the ocean and much shorter ones – between 2 weeks and 1 month – over the other surface types.

3.2 Sensitivity test

To test TransEBM's response to different forcings, we designed a three-step sensitivity test utilizing the implemented restarting extension. The model is initialized for conditions corresponding to the Last Glacial Maximum (LGM) with the forcings listed in Table 8. This simulation is run for 100 years with constant forcings. Following this, either one forcing or all of them shift to those found during the mid-Holocene warm period around 6000 years before present (BP). After another 100 years, the simulation switches to conditions for the year 2000 CE. Overall, the simulations thus cover three states in Earth's history with distinct climatic conditions that span a large range of the natural variability in forcings. For the solar constant there would be no significant change, which is why artificial steps were introduced that reveal its effect on the simulated climate. The simulation takes the orbital parameters for the three steps from Laskar et al. (2004) and CO_2 from Köhler et al. (2017). For the land, sea, and ice distribution during the LGM and modern day, the data provided by Zhuang et al. (2017a) were used. They are shown in Fig. 1 and Fig. A1a, respectively. The map for the mid-Holocene warm period was then the result of an interpolation between the two based on the sea level reconstruction by Grant et al. (2012) and using the bathymetry by the National Geophysical Data Center (1993) (cf. Fig. A1b).

Figure 13 shows the temperature time series resulting from the sensitivity test. Among the forcings, the distributions of ice sheets and sea ice and the CO_2 value have the largest effect on GMT. Despite its artificial enhancement, the solar constant has a smaller influence. Changes to the land–sea mask and to obliquity have a significant effect as well, whereas the variations due to eccentricity and precession are negligible. The sum of the temperature changes due to the individual forcing steps is slightly larger than the simultane-

Table 7. Theoretical average phase lags (Eq. 9) and amplitudes (Eq. 8) in TransEBM for the different surface types and annual forcing.

Surface type	Phase lag ϕ_f [rad]	Phase lag $365 \cdot \phi_f (2\pi)^{-1}$ [d]	$\lvert\tilde{T_f}\rvert$ [° C]
Land	0.24	14.2	49.6
Ocean	1.54	89.7	1.38
Sea ice	0.48	28.1	45.2
Snow	0.21	12.4	49.9

Table 8. Step forcings for the sensitivity test. Modeled states are associated with the LGM, the mid-Holocene warm period 6000 BP, and 2000 CE. Orbital parameters follow Laskar et al. (2004), sea level w.r.t. 0 BP is from Grant et al. (2012), and CO_2 is from Köhler et al. (2017). Appendix A shows the land, sea, and ice distributions corresponding to the sea levels of the three modeled periods. The solar constant is chosen such that it, like the other forcings, changes significantly.

Forcing	LGM	6 ka	Year 2000 CE
Obliquity	0.391108	0.422536	0.409093
Precession	0.871267	5.727245	1.796257
Eccentricity	0.017868	0.01918	0.016702
Sea level	−110 m	−6 m	+1 m
Solar constant S_0	1360 W m^{-2}	1365 W m^{-2}	1370 W m^{-2}
CO_2	190.21 ppm	264.36 ppm	368.92 ppm

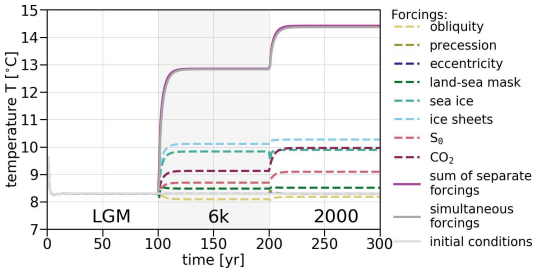


Figure 13. Test of the sensitivity of the model to the different forcings. Conditions for three different states, which cover a large range in forcings, are simulated as described in Table 8. When forcings are applied one at a time, the sum of the temperature change differs slightly from when they act altogether, indicating some nonlinear model behavior. Changes to CO_2 and the distributions of ice sheets and sea ice affect GMT the most.

ous change in all forcings, suggesting that the model contains some nonlinear behavior introduced by the restarts.

3.3 Past millennium climate

To further examine TransEBM's abilities, we simulate the climate of the past millennium. The solar forcing includes the change in radiative forcing due to volcanic eruptions as given by Neukom et al. (2019), who also provided the CO_2 forcing employed in this simulation. The orbital parameters are calculated according to Berger (1978) and the land, sea, and ice distribution is that from Fig. 1. To ensure adequate time for the decay of the transient solution, the simulation is initialized at 800 CE, but evaluated from 850 to 2000 CE. Figure 14 compares the results to a HadCM3 (Gordon et al., 2000; Pope et al., 2000) experiment (ID °xnapa', unpublished data, Kira Rehfeld, Bühler et al., 2020), a past millennium simulation with pre-industrial orography; solar, volcanic, and orbital forcing; and varying greenhouse gases. The temperature fields were regridded to T42 resolution. For the comparison of the GMT time series, the temperature reconstructions by the PAGES 2k Consortium (Neukom et al., 2019), as well as a CESM past millennium ensemble (Otto-Bliesner et al., 2016), were added.

Figure 14a presents the temperature anomalies. As in the 20th century simulations, the overall trends are similar, but TransEBM simulates lower variability than HadCM3 and CESM. The reconstruction also shows larger variability. The response from HadCM3 and CESM to volcanic forcing, seen for example in large negative peaks, is also larger than in TransEBM. As discussed in Sect. 3.1, this is most likely due to a lack of positive feedback mechanisms in TransEBM, which are present in the more detailed descriptions of HadCM3 and CESM. The response to the increase in greenhouse forcing, on the other hand, is similar in TransEBM, CESM, and the reconstruction. The average latitudinal distributions shown in Fig. 14b largely agree. They differ over Antarctica, where HadCM3 simulates lower temperatures, and in the lower latitudes of the Northern Hemisphere, where HadCM3 produces higher temperatures. The seasonal cycles in the Northern Hemisphere and Southern Hemisphere as displayed in Fig. 14, on the other hand, show significant disagreement: in the Northern Hemisphere the amplitude of the cycle is larger by about 5° in HadCM3, due to higher summer temperatures. In both hemispheres, the

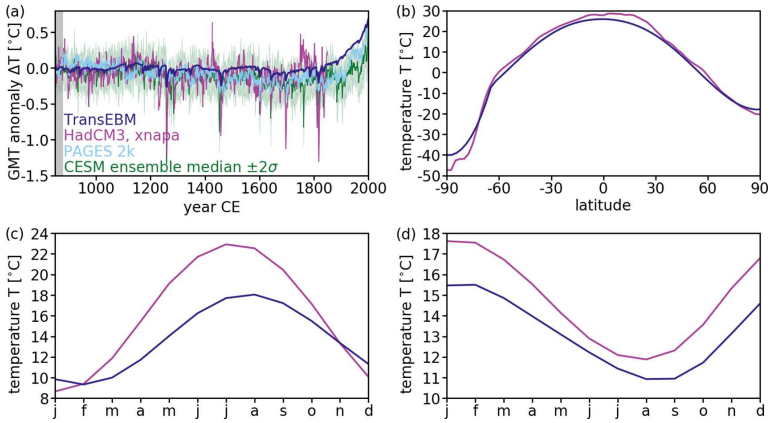


Figure 14. Past millennium simulations and reconstruction. Solar, volcanic, and greenhouse forcings for TransEBM are from Neukom et al. (2019); orbital are parameters set according to Berger (1978); and land, sea, and ice distribution are the same as in Fig. 1. For comparison, the xnapa experiment (ID °xnapa', unpublished data, K. Rehfeld, Bühler et al., 2020) from HadCM3 (Gordon et al., 2000; Pope et al., 2000) is provided until 1850 CE, in addition to GMT in a CESM ensemble (Otto-Bliesner et al., 2016) and a reconstruction for the Common Era (Neukom et al., 2019). Panel **(a)** presents the GMT anomaly with respect to the climatological means from 850 to 879 CE. For CESM, the ensemble median with a spread of 2 standard deviations is shown. The grey shading indicates the reference period. Panel **(b)** shows the average latitudinal temperature profiles; panels **(c)** and **(d)** show the average seasonal temperature profiles for the Northern Hemisphere and Southern Hemisphere, respectively. The overall trends in GMT agree, but HadCM3, CESM and the reconstruction are much more variable. Additionally, HadCM3 and CESM show a stronger response to volcanic forcing. The latitudinal profiles show disagreement only over Antarctica and in the lower latitudes of the Northern Hemisphere. The seasonal cycles disagree, with HadCM3 simulating overall higher temperatures and a larger amplitude in the Northern Hemisphere.

temperature extreme between July and September is shifted, with TransEBM lagging behind HadCM3. In the Southern Hemisphere, HadCM3 generally yields higher temperatures.

4 Short tutorial

For setting up TransEBM, three files need to be potentially changed: *preprocess/preprocess.sh*, *config/config_module.sh*, and *src/Makefile*. The model was tested with Ubuntu 18.04 and the gfortran compiler and should run on Linux and Unix machines. To set up the model, follow the following procedure:

1. Start by installing all necessary packages. Running TransEBM requires Fortran90 or higher and NetCDF[1]. The optional post-processing relies on climate data operators (CDO, Max Planck Institute für Meteorologie).

2. For using a compiler other than gfortan-7, change the compiler name in files *config/config_module.sh*, *preprocess/preprocess.sh*, and *src/Makefile*.

3. Set the paths to the local installation of NetCDF for Fortran in *preprocess/preprocess.sh* and *src/Makefile* via the following steps.

   – In *preprocess/preprocess.sh*, give the include path via *I/usr/include* and the library via the *-lnetcdff* flag. If the installation did not set that flag, provide the path via *L/.../netcdf/lib* instead.

   – In *src/Makefile*, provide the include folder via the variable *NI* and the library as the *-lnetcdff* flag. Give the compiler as *FC=gfortran-7*. Additionally, the configuration folder needs to be provided as *I/home/user/TransEBM/config*. Set this to the path of the configuration folder of TransEBM for the given setup.

4. To switch on the post-processing, un-comment the line *bash postprocess_cdo.sh $EBM_wrk _dir* in *run_ebm.sh* in the top level folder.

An experiment can be set up with a configuration file. The model code includes a default configuration in *config/default_config.conf*, which contains all variables that can be set, as well as a short explanation for all of them. The default configuration is that of an equilibrium experiment starting in 1950 CE with constant $CO_2 = 315$ ppm, solar insolation $S_0 = 1362\,\mathrm{W\,m^{-2}}$, and orbital parameters. It uses

[1]On Ubuntu, the packages *gfortran-7*, *libnetcdf-dev*, and *libnetcdff-dev* can be installed using the apt-get package manager. It is then possible to use *nf-config –all* to determine the details of your installation (e.g., the include directory) that are needed for setting up TransEBM.

the land–sea mask shown in Fig. 1, which is provided in the same folder, and the parameterization described in Sect. 2.4. To design a different experiment, TransEBM requires a configuration file that specifies all desired variables that differ from the default values. Any variables that are not set in that configuration will default to the values given in *config/default_config.conf*. Possible changes are as follows.

1. *Set the land–sea mask and ice distribution*
   Give the path to a text file containing the desired land–sea mask as the parameter *WORLD*. That text file should have 64 rows and 128 columns and contain a number for each grid point that describes the surface type: 1 for land, 2 for sea ice, 3 for land-based ice, and 5 for ocean. It should contain no spaces or empty rows.

2. *Set the CO_2 forcing*
   For a constant forcing, *CO2_VALUE* can be set together with *CO2_FORCED = FALSE*. For a non-constant forcing, a text file with one CO_2 value per year and line can be provided as *CO2_FILE* with *CO2_FORCED = TRUE* and number of rows given as the file length *CO2_FILE_LEN*.

3. *Set the solar constant*
   Setting the solar constant works in exactly the same way as CO_2 using the variable *S0_VALUE* with *S0_FORCED = FALSE* for a constant forcing or *S0_FILE* with *S0_FORCED = TRUE* for a transient solar constant. Remember to set the file length in *S0_FILE_LEN*.

4. *Set the orbital parameters*
   For constant values, set *ECC*, *OB*, and *PER* or have them be calculated following Berger (1978). In the latter case, provide a start year of the experiment as *INITIAL_YEAR* and set *ORB_SRC="code"* and *ORB_FORCED=TRUE*.

5. *Define the length of the experiment*
   Provide the initial year as a Common Era value in *INITIAL_YEAR*, for example as 1950 or −21000 (relevant only if the orbital parameters are computed in the code). Set the number of years for which the simulation should be run as *RUN_YEARS*.

6. *Choose whether it is an equilibrium or transient simulation*
   For *USE_EQUI=true*, the experiment will be stopped once equilibrium or the number of *RUN_YEARS* is reached, whichever happens first. For *USE_EQUI=false*, the simulation will run for as many years as provided in *RUN_YEARS*.

7. *Set the working and the output directory*
   Set them using *WRK_DIR* and *OUT_DIR*.

8. *Set the experiment ID*
   Finally, the experiment should receive an *ID*.

Other possible changes to experiments include the option to change the parameterization and the writing of additional output. Once the configuration is set up, the model can be run by executing *run_ ebm.sh* with the following configuration file as input:

bash run_ebm.sh config/test.conf

If no configuration is specified, the code defaults to *config/default_config.conf*.

5 Discussion

Building on previous modeling efforts, ranging from Budyko (1968) to Zhuang et al. (2017a), we have presented here TransEBM v. 1.0, a transient energy balance model that simulates Earth's temperature field based on CO_2 greenhouse forcing, orbital parameters, and solar constant, as well as land, sea, and ice distribution. The provided tuning is in reference to the 1960–1989 climatology from the ERA-20CM reanalysis (Hersbach et al., 2015). It sampled the parameter space to investigate the sensitivity to the different parameters and, where available, relied on observations to provide boundaries. The tuning period and metrics, with a prioritization of the latitudinal and seasonal temperature profiles, were chosen to test the model's ability to simulate climate in polar regions and with its application for simulations of past climate in mind. The tuning improves the simulated seasonal cycle in the Southern Hemisphere and the latitudinal temperature profile in particular. Thereby, it improves the fit of the temperature profile in polar regions, showing that TransEBM can be used to model temperatures in ice-covered regions and as such provide ensemble runs exploring long timescales.

For the past millennium, there remain small disagreements with reanalysis and GCM simulations over Antarctica, where TransEBM overestimates temperatures towards the pole. This is due to the lack of topography in the model, which is why the elevation of ice sheets cannot be considered. Furthermore, the temperatures near the Equator disagree a little, with lower temperatures simulated by TransEBM for low northern latitudes. After tuning, TransEBM is able to reproduce the dip in temperatures related to the Intertropical Convergence Zone (ITCZ). However, this is an indirect parameterization of atmospheric dynamics that might not translate to climate states other than the tuned one.

TransEBM follows the overall trend in GMT but is unable to capture inter-annual variability due to the lack of dynamical processes modeled. With respect to the seasonal cycle, TransEBM lags the reanalysis data and produces smaller seasonal amplitudes. Moreover, there is less variability found in the spatial distribution of the amplitudes in the model, mirroring the findings of Hyde et al. (1989). The spatial structures of the amplitude and phase lag of the seasonal cycle that they found in their EBM are similar to the ones presented here (cf. Figs. 12 and 11). They also found reduced spatial variability in comparison to observations. The simulated

phase lags are similar as well (cf. Fig. 11 here and Fig. 3a in Hyde et al. (1989)), but the amplitudes differ (cf. Fig. 12 here and Fig. 1a in Hyde et al., 1989). In better agreement with the reanalysis, TransEBM produces amplitudes roughly twice as high as Hyde et al. (1989).

The seasonal cycle in TransEBM is a product of only the seasonality in its insolation fields. As such, it lacks seasonality in the sea ice response, which would influence the parameterization of albedo, heat transport, and heat capacity as simulated by the model. The first trials of including such a seasonal sea ice cycle that could lead to the inclusion of dynamical ice sheets have been unsuccessful since they destabilize the numerical scheme. Such a change to the model therefore requires further research. Furthermore, the model does not currently explicitly consider energy conservation. This is especially relevant for interactive changes to the land, sea, and ice distribution between model restarts.

Overall, TransEBM makes large ensemble runs and long timescales accessible for exploration and can be used to design experiments with models more constrained by computational costs since TransEBM runs quickly with limited computational resources. As an easy-to-use climate model, it is suitable for use in education and for reproducing past studies with energy balance models for which no documented code exists (e.g., Hyde et al., 1989). In order to reproduce studies of hysteresis and abrupt transitions between ice-covered and ice-free states (cf. North et al., 1983; Mengel et al., 1988; Crowley et al., 1989; Short et al., 1991), an extension of the model with interactive ice sheets and sea ice would be necessary. When it comes to forcings, TransEBM underlines the sensitivity to ice configuration and CO_2 forcing of simulated climates in comparison to other forcings, which have smaller impacts on temperatures. Furthermore, it illustrates the large impact that the parameterization of the outgoing radiation and the albedo fields of ocean and land masses have on the model results.

6 Conclusions

We describe a transient energy balance framework suitable for simulating planetary temperatures on long timescales due to its computational efficiency. The model can capture trends in the global mean temperature and simulates latitudinal temperature profiles that agree well with ERA-20CM reanalysis data (Hersbach et al., 2015) and HadCM3 (Gordon et al., 2000; Pope et al., 2000), a general circulation model. There are differences with respect to the seasonal cycle especially in the Northern Hemisphere related to the tuning of the model provided here. TransEBM v. 1.0 has a larger phase lag than that found in reanalysis, and the amplitude of the seasonal cycle in the Northern Hemisphere is smaller. Overall, it simulates less variable temperature fields than reanalysis, multiproxy reconstructions, or more complex models due to its reduced complexity.

Among the modeled forcings, changes to the ice sheets, CO_2 level, and the sea ice configuration are shown to influence the simulated climate the most. Among the parameters, the albedo fields of the oceans and land masses, as well as the parameterization of outgoing radiation, have the largest impact, as even small changes to these parameters can shift the model results significantly. With respect to the seasonality, the heat capacity of the oceans also carries some weight. The model behaves largely linearly, with only a small difference between the separate and simultaneous application of the forcings.

Desirable future improvements include the addition of a seasonal cycle in the sea ice, which currently remains constant during a model run. The inclusion of this seasonality is likely to improve the modeled seasonal cycle. Such a change to the model could encompass all ice masses and have them respond dynamically to the modeled temperature changes. For this to be physically reasonable, consideration of energy conservation within the model is necessary, which is especially relevant for the simulation of climate on long timescales with significant changes to Earth's ice masses. Overall, TransEBM v. 1.0 is accessible for large and long ensemble runs and thus can complement complex models, in particular for the study of the influence of orbital cycles and other radiative forcings on Earth's energy balance.

Appendix A: Land, sea, and ice distributions used for the sensitivity test

Figure A1 shows the land, sea, and ice configuration used for the sensitivity test presented in Sect. 3.2. Figure A1a shows the map for the LGM taken from Zhuang et al. (2017a). Figure A1b shows a map corresponding to the state at 6000 BP. For this, the sea levels from Grant et al. (2012) were taken and applied to the bathymetry by the National Geophysical Data Center (1993) to obtain a land–sea distribution. Based on the difference in sea levels between this time and both the LGM and modern configurations from Fig. A1a and 1, respectively, the difference in ice cover was interpolated. The amount of ice corresponding to the difference in sea level was removed from the LGM map according to a weighting. The weighting preferred the removal of ice closer to the Equator and at the borders of ice cover, that is ice with more ice-free neighbors. Ice was removed randomly if only a fraction of the ice with the same weighting needed to be removed. This interpolation is flawed – it ignores the height of ice sheets, energy conservation, and many more aspects – but for the idealized sensitivity test it suffices.

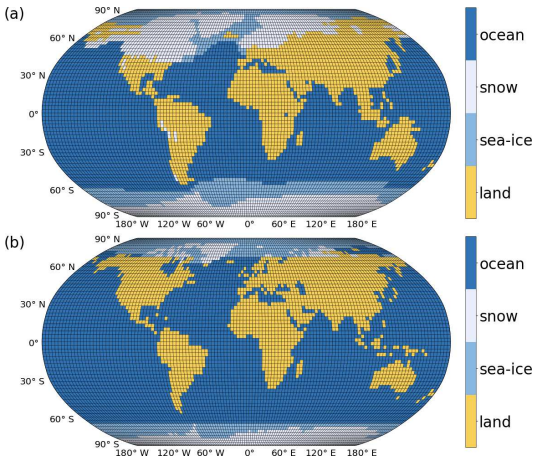


Figure A1. Configuration of land, sea, and ice during **(a)** the LGM and **(b)** the mid–Holocene warm period about 6000 BP as they were used for the sensitivity test discussed in Sect. 3.2. Panel **(a)** is from Zhuang et al. (2017a); panel **(b)** has a land–sea distribution based on the sea level reconstruction from Grant et al. (2012) and the topography by the National Geophysical Data Center (1993). Ice was interpolated according to sea level difference of this state to **(a)** and the modern map from Fig. 1.

Appendix B: Effect of the extremes in orbital forcing on November–December–January and yearly insolation and average temperatures

The relationship between orbital parameters and the modeled insolation and temperature distributions changes throughout the year. Section 2.2 analyzes the effect of extremes in orbital parameters on insolation and temperatures from May to July. Here, Fig. B1 repeats the same analysis for November to January. Overall, precession and obliquity cause similarly large differences in the amplitude of the insolation. For obliquity, this translates into the largest modeled temperature changes between low and high obliquity states among the three orbital parameters. Eccentricity, on the other hand, has the smallest effect both on insolation and temperatures.

All three orbital parameters show a pattern mirrored between the hemispheres from that between May and July shown in Fig. 5. For eccentricity this means that a high eccentricity corresponds to larger insolation and temperatures than are found for a lower eccentricity if obliquity and perihelion are set to the averages noted in Table 3. This net positive effect is larger in the Southern Hemisphere with insolation growing towards middle southern latitudes. The largest increases in temperature are found over Antarctica, followed by a region covering eastern Africa and the Middle East. Altogether, temperatures increase the least over the oceans.

A high obliquity causes a decrease in insolation in the Northern Hemisphere and an increase in the Southern Hemisphere in comparison to the low obliquity simulation. The largest decrease can be found in the middle northern latitudes, and the largest increase can be found in the high southern latitudes. With regard to temperatures, the largest increase occurs in the southern polar regions, but increased temperatures are simulated for the northern polar regions as well. The largest decreases are found over Asia and near the Equator.

With the perihelion at the June solstice, insolation is lower between November and January in comparison to the perihelion occurring at the December solstice. This decrease is largest in the low latitudes of the Southern Hemisphere. It translates to decreased temperatures over land masses, Asia and Africa in particular, and towards the polar regions, which are affected by the lack of energy entering near the Equator, which, if present, would be transported polewards. Precession is the strongest in low southern latitudes, whereas obliquity and eccentricity affect the middle and high latitudes more.

Repeating the analysis for yearly averages yields the results shown in Fig. B2. Most of the differences average out between the simulations with orbital extremes. For the precession of the perihelion in particular, there is virtually no difference between the experiments, as Fig. B2e and f show. Both eccentricity and obliquity show a symmetric pattern in the annual averages. The difference between obliquity extremes is the largest, confirming the results from the sensi-

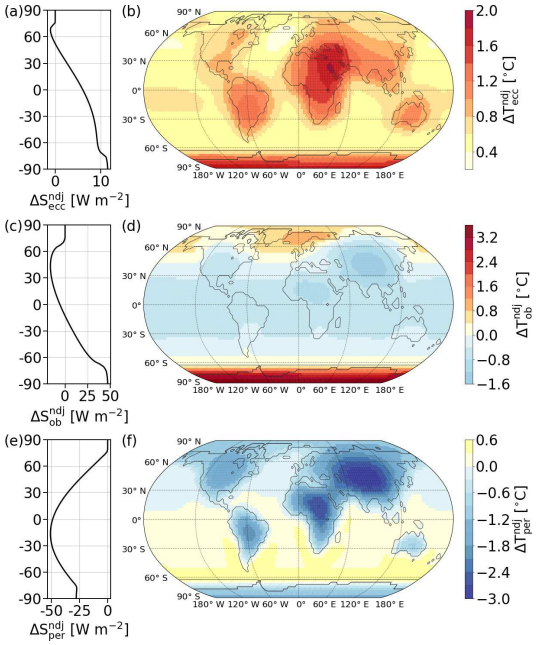


Figure B1. Difference in solar insolation (left) and temperature (right) caused by extremes in the orbital parameters as listed in Table 3 for eccentricity **(a, b)**, obliquity **(c, d)**, and precession of the perihelion **(e, f)** from November to January (NDJ). Model configuration as in Fig. 14. Obliquity has the largest impact on simulated temperatures, closely followed by precession. High eccentricity corresponds to overall increased insolation and temperatures, with the largest increase in insolation towards the middle southern latitudes. Temperatures in the high eccentricity state change the most over Antarctica and continental areas near the Equator with respect to the low eccentricity simulation. A decrease in insolation in the Northern Hemisphere and increase in the Southern Hemisphere characterizes the high obliquity state when comparing it to the low obliquity one. Temperature increases are found in the polar regions, especially over Antarctica. Changes to precession have the strongest impact on low-latitude insolation and on temperatures in continental areas. NDJ insolation and temperatures are larger when perihelion occurs at the December solstice.

tivity study from Sect. 3.2, which also found that obliquity had the most notable effect on the global mean temperature among the orbital parameters. Obliquity affects the high latitudes the most, where a high obliquity increases insolation and temperatures. Close to the Equator, there is a net negative effect on the insolation and temperatures. A high eccentricity, on the other hand, produces the largest effect close to the Equator, with both insolation and temperatures decreasing towards higher latitudes.

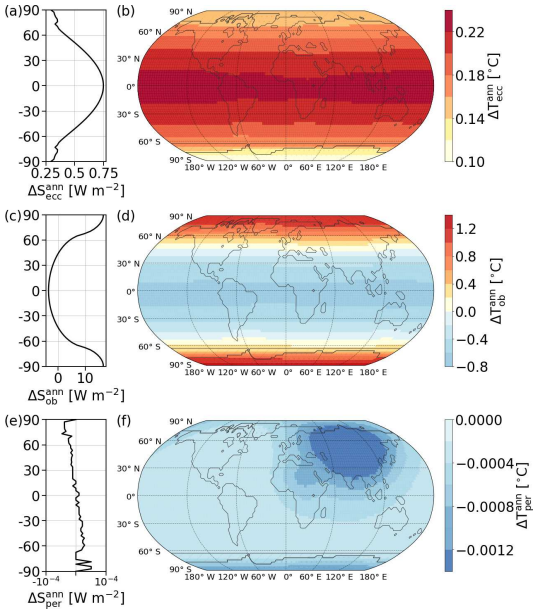


Figure B2. Annually averaged difference in solar insolation (left) and temperature (right) caused by extremes in the orbital parameters as listed in Table 3 for eccentricity **(a, b)**, obliquity **(c, d)**, and precession of the perihelion **(e, f)**. Model configuration is the same as in Fig. 14. Obliquity changes produce the largest difference, whereas the precession of the perihelion averages out throughout the year and has no effect on annual averages. Both eccentricity and obliquity show symmetric patterns. However, eccentricity affects the low latitudes the most, while the difference in the obliquity experiments grows polewards.

Appendix C: Seasonality in the model

The analysis of seasonality discussed in Sect. 3.1 fitted sinusoids to the detrended temperature time series of TransEBM and ERA-20CM for each grid box and then extracted the amplitude and phase lag from these fits. The computation of the phase lags relies on the fits, since TransEBM's low temporal resolution would only allow a coarse estimate. However, there are some grid boxes where such a fit oversimplifies the temperature pattern, especially in the ERA-20CM reanalysis. This is why the phase lag analysis shown in Fig. 11 excluded some grid boxes. Figure C1 delves deeper into these issues by showing the disagreement between the fitted amplitude and the amplitude calculated as the difference between the minimal and maximal temperature. The difference between the min–max and fit amplitudes serves as an indicator for how suitable the fit is for deriving amplitudes. The figure also shows the time series with fits for six exemplary grid boxes.

Overall, the difference between the amplitudes is small in TransEBM. The fitted amplitude is larger by up to 1° over

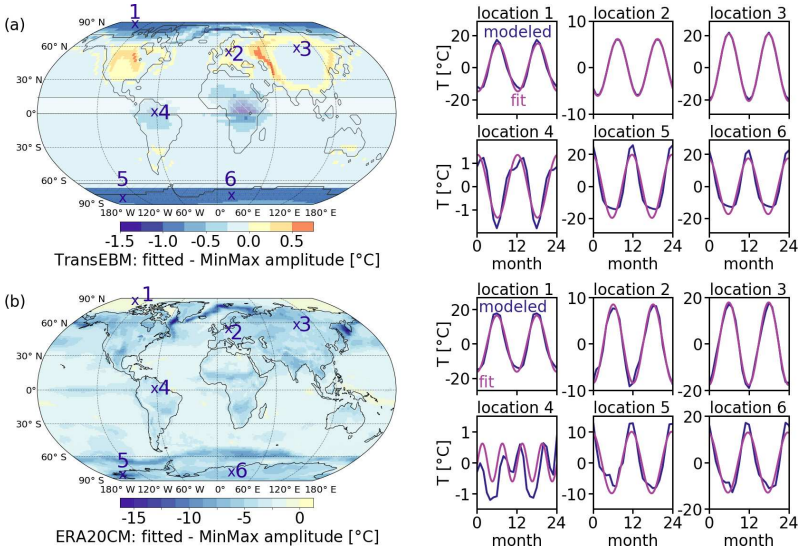


Figure C1. Amplitude of the seasonal cycle extracted from fitting a sine minus the amplitude that results from taking the difference between the maximal and minimal temperature. Model configuration as in Fig. 5. Time series from **(a)** TransEBM and **(b)** ERA-20CM (Hersbach et al., 2015) from 1900 to 1989 CE, both were detrended. Six locations were picked as marked in the maps, for which 2 years of the time series and the fit are shown. For TransEBM, even the maximal differences are relatively small. In the polar regions the min–max amplitude is larger, over Asia and North America it is the other way around. For ERA-20CM the differences are larger, reaching up to around 15° in oceanic polar regions. The fit almost always produces smaller amplitudes. For both, locations 1–3 show places where the fit matches the time series very well. Location 4, on the other hand, lies close to the Equator inside the shaded area where the time series show only small variations throughout the year that are not sinusoidal. Locations 5 and 6 lie in Antarctica, where the fit is still applicable but is not able to fully capture the extrema.

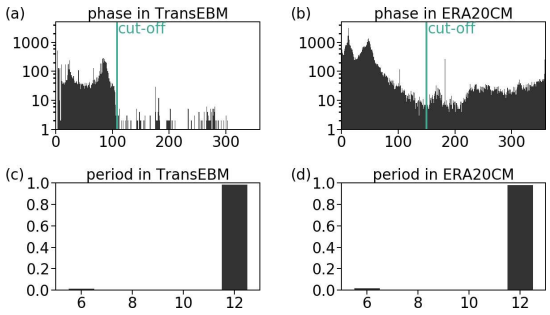


Figure C2. Logarithmic histogram of the phase distribution in **(a)** TransEBM and **(b)** ERA-20CM (Hersbach et al., 2015) with cut-offs at 108 and 150 d, respectively, and histograms of the periods found in **(c)** TransEBM and **(d)** ERA-20CM. Both are based on the sinusoids fitted to the detrended 1900–1989 CE time series. Model configuration as in Fig. 5. Panel **(a)** shows a clear cut-off, after which only few grid boxes have larger phase lags. All of them lie close to the Equator, where the time series is not sinusoidal as shown in Figs. 12a and C1a. **(b)** In the ERA-20CM data a much larger variety of phases occurs and even very large phase lags are found. The cut-off here was chosen such that all locations close to the Equator with non-sinusoidal behavior were excluded. Many of the locations with very large phase lags are in Antarctica, where the temperature maximum sometimes occurs a couple of days before the insolation maximum as shown in Figs. 12b and C1b. Panels **(c)** and **(d)** both show that most grid boxes have a period of around 12 months, with very few exceptions of places where the fitted sine has a period of about 6.

North America and parts of Asia. The min–max amplitude is larger in polar regions, as shown at locations 5 and 6 in Antarctica. In ERA-20CM reanalysis data the differences between the two types of amplitudes are larger, especially in locations with sea ice where they can become more than 15°. It is notable that the fitted amplitude is almost always smaller and only ever larger by a little. This underestimation of the fitted amplitude is visible in locations 5 and 6, where the fit is unable to capture the temperature maxima. Location 4 near the Equator in South America is an example of a grid box that was excluded from the phase lag analysis. It is a location where the variations have a period smaller than 12 months. Since this was taken as a sign that a sine would not fit the underlying time series, such places where excluded from the analysis.

Figure C2 shows the histograms of the phases and periods for both TransEBM and ERA-20CM. A period other than around 12 occurs only in few places, mostly around the Equator where the time series shows no seasonal cycle. The phase histograms were used to define cut-offs to distinguish places where sine fitting was suitable from those, where the time series showed little or no sinusoidal behavior. Placing the cut-off is straightforward in TransEBM, since there is a gap at a phase lag of 108 d, after which there are almost no occurrences. Furthermore, Fig. 12a shows that all of these occurrences are close to the Equator, i.e., in places like location 4 without a 12-month seasonal cycle.

In ERA-20CM, the picture is more complicated, as Fig. C2b shows. Here, the cut-off was set by lowering it until all grid boxes close to the Equator where the time series is not suited to fitting a sine were excluded. This leads to a cut-off at 150 d. In doing so, large parts of Antarctica were excluded as well. As discussed previously in Sect. 3.1, these are places where the temperature maximum occurs slightly before the maximum in insolation, leading to very high computed phase lags, which the analysis omits. While this means that some grid boxes might have been excluded unnecessarily, this procedure allows confidence in all grid boxes that are included.

Code and data availability. The current version of the model is available at https://github.com/paleovar/TransEBM (last access: 30 April 2021) and licensed under the GNU General Public License Version 3. The version of the model used for the analyses presented here is archived on Zenodo (https://doi.org/10.5281/zenodo.3941311, Ziegler and Rehfeld, 2020), together with all input data, configuration files, and scripts to run the model. The model simulations and all other data to produce the plots of this paper, as well as the code that generated the plots, are archived there as well.

Author contributions. EZ and KR conceptualized this study, decided on the methodology, created the analyses, and edited and reviewed the manuscript. EZ implemented the software and computations, performed the visualization of the data, and wrote the manuscript, all supervised by KR.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We are grateful for discussions with Marie-France Loutre regarding our analyses and for the support by the Heidelberg Graduate School for Physics (HGSFP). We thank Beatrice Ellerhoff and all reviewers for their helpful comments on the manuscript.

Financial support. This research has been supported by the Deutsche Forschungsgemeinschaft (grant no. 395588486), the Bundesministerium für Bildung und Forschung (grant no. 01LP1926C), the data storage service SDS@hd, supported by the Ministry of Science, Research and the Arts Baden-Württemberg (MWK) and the German Research Foundation (DFG) (grant no. INST 35/1314-1 FUGG).

Review statement. This paper was edited by Andrea Stenke and reviewed by Colin Zhuang, Shaun Lovejoy, and one anonymous referee.

References

Abe-Ouchi, A., Saito, F., Kawamura, K., Raymo, M. E., Okuno, J., Takahashi, K., and Blatter, H.: Insolation-driven 100,000-year glacial cycles and hysteresis of ice-sheet volume, Nature, 500, 190–193, https://doi.org/10.1038/nature12374, 2013.

Andres, H. J. and Tarasov, L.: Towards understanding potential atmospheric contributions to abrupt climate changes: characterizing changes to the North Atlantic eddy-driven jet over the last deglaciation, Clim. Past, 15, 1621–1646, https://doi.org/10.5194/cp-15-1621-2019, 2019.

Arrhenius, S.: On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground, Philosophical Magazine and Journal of Science, 41, 237–276, 1896.

Bathiany, S., Dijkstra, H., Crucifix, M., Dakos, V., Brovkin, V., Williamson, M. S., Lenton, T. M., and Scheffer, M.: Beyond bifurcation: using complex models to understand and predict abrupt climate change, Dynamics and Statistics of the Climate System, https://doi.org/10.1093/climsys/dzw004, 2016.

Berger, A. L.: Long-Term Variations of Daily Insolation and Quaternary Climatic Changes, J. Atmos. Sci., 35, 2362–2367, https://doi.org/10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2, 1978.

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B., and Zhao, Y.: Evaluation of climate models using palaeoclimatic data, Nat. Clim. Change, 2, 417–424, https://doi.org/10.1038/nclimate1456, 2012.

Briggs, W. L., Hensen, V. E., and McCormick, S. F.: A Multigrid Tutorial, Society for Industrial and Applied Mathematics, 2nd Edn., https://doi.org/10.1137/1.9780898719505, 2000.

Budyko, M. I.: The effect of solar radiation variations on the climate of the Earth, Tellus, 21, 611–619, https://doi.org/10.3402/tellusa.v21i5.10109, 1968.

Bühler, J. C., Roesch, C., Kirschner, M., Sime, L., Holloway, M. D., and Rehfeld, K.: Comparison of the oxygen isotope signatures in speleothem records and iHadCM3 model simulations for the last millennium, Clim. Past Discuss. [preprint], https://doi.org/10.5194/cp-2020-121, in review, 2020.

Coddington, O., Lean, J. L., Pilewskie, P., Snow, M., and Lindholm, D.: A solar irradiance climate data record, B. Am. Meteorol. Soc., 97, 1265–1282, https://doi.org/10.1175/BAMS-D-14-00265.1, 2016.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A., and Wehner, M.: Long-term Climate Change: Projections, Commitments and Irreversibility, book section 12, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1029–1136, https://doi.org/10.1017/CBO9781107415324.024, 2013.

Crowley, T. J. and Hyde, W. T.: Transient nature of late Pleistocene climate variability, Nature, 456, 226–230, https://doi.org/10.1038/nature07365, 2008.

Crowley, T. J. and North, G. R.: Abrupt climate change and Extinction Events in Earth History, Science, 240, 996–1002, 1988.

Crowley, T. J. and North, G. R.: Modeling onset of glaciation, Ann. Glaciol., 14, 39–42, https://doi.org/10.3189/S0260305500008223, 1990.

Crowley, T. J., Short, D. A., Mengel, J. G., and North, G. R.: Role of seasonality in the evolution of climate during the last 100 million years, Science, 231, 579–584, https://doi.org/10.1126/science.231.4738.579, 1986.

Crowley, T. J., Mengel, J. G., and Short, D. A.: Gondwanaland's seasonal cycle, Nature, 329, 803–807, https://doi.org/10.1038/329803a0, 1987.

Crowley, T. J., Hyde, W. T., and Short, D. A.: Seasonal cycle variations on the supercontinent of Pangaea, Geology, 17, 457–460, https://doi.org/10.1130/0091-7613(1989)017<0457:SCVOTS>2.3.CO;2, 1989.

Dommenget, D., Nice, K., Bayr, T., Kasang, D., Stassen, C., and Rezny, M.: The Monash Simple Climate Model experiments (MSCM-DB v1.0): an interactive database of mean climate, cli-

mate change, and scenario simulations, Geosci. Model Dev., 12, 2155–2179, https://doi.org/10.5194/gmd-12-2155-2019, 2019.

Dortmans, B., Langford, W. F., and Willms, A. R.: An energy balance model for paleoclimate transitions, Clim. Past, 15, 493–520, https://doi.org/10.5194/cp-15-493-2019, 2019.

Fraedrich, K., Bordi, I., and Zhu, X.: Climate dynamics on global scale: resilience, hysteresis and attribution of change, CISM International Centre for Mechanical Sciences, Courses and Lectures, 564, 143–159, https://doi.org/10.1007/978-3-7091-1893-1_6, 2016.

Ghil, M.: A Mathematical Theory of Climate Sensitivity or, How to Deal With Both Anthropogenic Forcing and Natural Variability?, in: Climate Change: Multidecadal and Beyond, edited by: Chang, C.-P., Ghil, M., Latif, M., and Wallace, J. M., chap. 2, World Scientific, 31–51, https://doi.org/10.1142/9789814579933_0002, 2015.

Goosse, H., Brovkin, V., Fichefet, T., Haarsma, R., Huybrechts, P., Jongma, J., Mouchet, A., Selten, F., Barriat, P.-Y., Campin, J.-M., Deleersnijder, E., Driesschaert, E., Goelzer, H., Janssens, I., Loutre, M.-F., Morales Maqueda, M. A., Opsteegh, T., Mathieu, P.-P., Munhoven, G., Pettersson, E. J., Renssen, H., Roche, D. M., Schaeffer, M., Tartinville, B., Timmermann, A., and Weber, S. L.: Description of the Earth system model of intermediate complexity LOVECLIM version 1.2, Geosci. Model Dev., 3, 603–633, https://doi.org/10.5194/gmd-3-603-2010, 2010.

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., Mitchell, J. F. B., and Wood, R. A.: The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, Clim. Dynam., 16, 147–168, https://doi.org/10.1007/s003820050010, 2000.

Grant, K. M., Rohling, E. J., Bar-Matthews, M., Ayalon, A., Medina-Elizalde, M., Ramsey, C. B., Satow, C., and Roberts, A. P.: Rapid coupling between ice volume and polar temperature over the past 150,000 years, Nature, 491, 744–747, https://doi.org/10.1038/nature11593, 2012.

Harvey, L. D. and Schneider, S. H.: Transient climate response to external forcing on 10^0-10^4 year time scales, 2. Sensitivity experiments with a seasonal, hemispherically averaged, coupled atmosphere, land, and ocean energy balance model, J. Geophys. Res., 90, 2207–2222, https://doi.org/10.1029/JD090iD01p02207, 1985.

Hébert, R., Lovejoy, S., and Tremblay, B.: An observation-based scaling model for climate sensitivity estimates and global projections to 2100, Clim. Dynam., 56, 1105–1129, https://doi.org/10.1007/s00382-020-05521-x, 2020.

Hersbach, H., Peubey, C., Simmons, A., Berrisford, P., Poli, P., and Dee, D.: ERA-20CM: A twentieth-century atmospheric model ensemble, Q. J. Roy. Meteor. Soc., 141, 2350–2375, https://doi.org/10.1002/qj.2528, 2015.

Huang, J. and Bowman, K. P.: The small ice cap instability in seasonal energy balance models, Clim. Dynam., 7, 205–215, https://doi.org/10.1007/BF00206862, 1992.

Huybers, P. and Curry, W.: Links between annual, Milankovitch and continuum temperature variability, Nature, 441, 329–332, https://doi.org/10.1038/nature04745, 2006.

Hyde, W. T., Crowley, T. J., Kim, K.-Y., and North, G. R.: Comparison of GCM and energy balance model simulations of seasonal temperature changes over the last 18,000 years, J. Climate, 2, 864–887, https://doi.org/10.1175/1520-0442(1989)002<0864:COGAEB>2.0.CO;2, 1989.

Kim, K.-Y. and North, G. R.: Surface temperature fluctuations in a stochastic climate model, J. Geophys. Res., 96, 18573, https://doi.org/10.1029/91JD01959, 1991.

Köhler, P., Nehrbass-Ahles, C., Schmitt, J., Stocker, T. F., and Fischer, H.: A 156 kyr smoothed history of the atmospheric greenhouse gases CO2, CH4, and N2O and their radiative forcing, Earth Syst. Sci. Data, 9, 363–387, https://doi.org/10.5194/essd-9-363-2017, 2017.

Kyle, H. L., Weiss, M., and Ardanuy, P.: Cloud, Surface Temperature, and Outgoing Longwave Radiation for the Period from 1979 to 1990, J. Climate, 8, 2644–2658, https://doi.org/10.1175/1520-0442(1995)008<2644:CSTAOL>2.0.CO;2, 1995.

Laepple, T. and Huybers, P.: Ocean surface temperature variability: Large model–data differences at decadal and longer periods, P. Natl. Acad. Sci. USA, 111, 16682–16687, https://doi.org/10.1073/pnas.1412077111, 2014.

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A. C. M., and Levrard, B.: A long-term numerical solution for the insolation quantities of the Earth, Astron. Astrophys., 428, 261–285, https://doi.org/10.1051/0004-6361:20041335, 2004.

Liu, Z., Otto-Bliesner, B. L., He, F., Brady, E. C., Tomas, R., Clark, P. U., Carlson, A. E., Lynch-Stieglitz, J., Curry, W., Brook, E., Erickson, D., Jacob, R., Kutzbach, J., and Cheng, J.: Transient simulation of last deglaciation with a new mechanism for bolling-allerød warming, Science, 325, 310–314, https://doi.org/10.1126/science.1171041, 2009.

Liu, Z., Zhu, J., Rosenthal, Y., Zhang, X., Otto-Bliesner, B. L., Timmermann, A., Smith, R. S., Lohmann, G., Zheng, W., and Elison Timm, O.: The Holocene temperature conundrum, P. Natl. Acad. Sci. USA, 111, E3501–E3505, https://doi.org/10.1073/pnas.1407229111, 2014.

Max Planck Institute für Meteorologie: Climate Data Operators, available at: https://code.mpimet.mpg.de/projects/cdo, last access: 30 April 2021.

Mengel, J. G., Short, D. A., and North, G. R.: Seasonal snowline instability in an energy balance model, Clim. Dynam., 2, 127–131, https://doi.org/10.1007/BF01053470, 1988.

Meyer, P., Hoell, M., and Kantz, H.: Reproducing Long-Range Correlations in Global Mean Temperatures in Simple Energy Balance Models, J. Geophys. Res.-Atmos., 123, 4413–4422, https://doi.org/10.1002/2017JD028078, 2018.

Myhre, G., Highwood, E. J., Shine, K. P., and Stordal, F.: New estimates of radiative forcing due to well mixed greenhouse gases, Geophys. Res. Lett., 25, 2715–2718, https://doi.org/10.1029/98GL01908, 1998.

National Geophysical Data Center: 5-minute Gridded Global Relief Data (ETOPO5), National Geophysical Data Center, NOAA, https://doi.org/10.7289/V5D798BF, 1993.

Neukom, R., Barboza, L. A., Erb, M. P., Shi, F., Emile-Geay, J., Evans, M. N., Franke, J., Kaufman, D. S., Lücke, L., Rehfeld, K., Schurer, A., Zhu, F., Brönnimann, S., Hakim, G. J., Henley, B. J., Ljungqvist, F. C., McKay, N., Valler, V., and von Gunten, L.: Consistent multidecadal variability in global temperature reconstructions and simulations over the Common Era, Nat. Geosci., 12, 643–649, https://doi.org/10.1038/s41561-019-0400-0, 2019.

North, G. R.: The small ice cap instability in diffusive climate models, J. Atmos. Sci., 41, 3390–3395, https://doi.org/10.1175/1520-0469(1984)041<3390:TSICII>2.0.CO;2, 1984.

North, G. R. and Coakley, J. A.: Differences between seasonal and mean annual energy balance model calculations of climate and climate sensitivity, J. Atmos. Sci., 36, 1189–1204, https://doi.org/10.1175/1520-0469(1979)036<1189:DBSAMA>2.0.CO;2, 1979.

North, G. R. and Kim, K.-Y.: Energy Balance Climate Models, Wiley-VCH, Weinheim, 2017.

North, G. R., Cahalan, R. F., and Coakley, J. A.: Energy-Balance Climate Models, Rev. Geophys. Space Phys., 16, 2000–2001, 1980.

North, G. R., Mengel, J. G., and Short, D. A.: Simple energy balance model resolving the seasons and the continents: application to the astronomical theory of the ice ages, J. Geophys. Res., 88, 6576–6586, https://doi.org/10.1029/JC088iC11p06576, 1983.

Otto-Bliesner, B. L., Brady, E. C., Fasullo, J., Jahn, A., Landrum, L., Stevenson, S., Rosenbloom, N., Mai, A., and Strand, G.: Climate variability and change since 850 CE. An ensemble approach with the community earth system model, B. Am. Meteorol. Soc., 97, 787–801, https://doi.org/10.1175/BAMS-D-14-00233.1, 2016.

Otto-Bliesner, B. L., Braconnot, P., Harrison, S. P., Lunt, D. J., Abe-Ouchi, A., Albani, S., Bartlein, P. J., Capron, E., Carlson, A. E., Dutton, A., Fischer, H., Goelzer, H., Govin, A., Haywood, A., Joos, F., LeGrande, A. N., Lipscomb, W. H., Lohmann, G., Mahowald, N., Nehrbass-Ahles, C., Pausata, F. S. R., Peterschmitt, J.-Y., Phipps, S. J., Renssen, H., and Zhang, Q.: The PMIP4 contribution to CMIP6 – Part 2: Two interglacials, scientific objective and experimental design for Holocene and Last Interglacial simulations, Geosci. Model Dev., 10, 3979–4003, https://doi.org/10.5194/gmd-10-3979-2017, 2017.

Pope, V. D., Gallani, M. L., Rowntree, P. R., and Stratton, R. A.: The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3, Clim. Dynam., 16, 123–146, https://doi.org/10.1007/s003820050009, 2000.

Rahmstorf, S., Crucifix, M., Ganopolski, A., Goosse, H., Kamenkovich, I., Knutti, R., Lohmann, G., Marsh, R., Mysak, L. a., Wang, Z., and Weaver, A. J.: Thermohaline circulation hysteresis: A model intercomparison, Geophys. Res. Lett., 32, 1–5, https://doi.org/10.1029/2005GL023655, 2005.

Rehfeld, K., Hébert, R., Lora, J. M., Lofverstrom, M., and Brierley, C. M.: Variability of surface climate in simulations of past and future, Earth Syst. Dynam., 11, 447–468, https://doi.org/10.5194/esd-11-447-2020, 2020.

Rohling, E. J., Sluijs, A., Dijkstra, H. A., Köhler, P., Van De Wal, R. S., Von Der Heydt, A. S., Beerling, D. J., Berger, A., Bijl, P. K., Crucifix, M., Deconto, R., Drijfhout, S. S., Fedorov, A., Foster, G. L., Ganopolski, A., Hansen, J., Hönisch, B., Hooghiemstra, H., Huber, M., Huybers, P., Knutti, R., Lea, D. W., Lourens, L. J., Lunt, D., Masson-Demotte, V., Medina-Elizalde, M., Otto-Bliesner, B., Pagani, M., Pälike, H., Renssen, H., Royer, D. L., Siddall, M., Valdes, P., Zachos, J. C., and Zeebe, R. E.: Making sense of palaeoclimate sensitivity, Nature, 491, 683–691, https://doi.org/10.1038/nature11574, 2012.

Ruddiman, W. F.: Earth's Climate. Past and Future, W. H. Freeman and Company, 3rd Edn., 2014.

Rypdal, K., Rypdal, M., and Fredriksen, H. B.: Spatiotemporal long-range persistence in earth's temperature field: Analysis of stochastic-diffusive energy balance models, J. Climate, 28, 8379–8395, https://doi.org/10.1175/JCLI-D-15-0183.1, 2015.

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 palaeo-climate comparisons to constrain future projections in CMIP5, Clim. Past, 10, 221–250, https://doi.org/10.5194/cp-10-221-2014, 2014.

Sellers, W. D.: A Global Climatic Model Based on the Energy Balance of the Earth-Atmosphere System, J. Appl. Meteorol., 8, 392–400, https://doi.org/10.1175/1520-0450(1969)008<0392:AGCMBO>2.0.CO;2, 1969.

Short, D. A., Mengel, J. G., Crowley, T. J., Hyde, W. T., and North, G. R.: Filtering of milankovitch cycles by earth's geography, Quaternary Res., 35, 157–173, https://doi.org/10.1016/0033-5894(91)90064-C, 1991.

Singarayer, J. S. and Valdes, P. J.: High-latitude climate sensitivity to ice-sheet forcing over the last 120 kyr, Quaternary Sci. Rev., 29, 43–55, https://doi.org/10.1016/j.quascirev.2009.10.011, 2010.

Smith, R. S. and Gregory, J.: The last glacial cycle: Transient simulations with an AOGCM, Clim. Dynam., 38, 1545–1559, https://doi.org/10.1007/s00382-011-1283-y, 2012.

Stap, L. B., van de Wal, R. S. W., de Boer, B., Bintanja, R., and Lourens, L. J.: The influence of ice sheets on temperature during the past 38 million years inferred from a one-dimensional ice sheet–climate model, Clim. Past, 13, 1243–1257, https://doi.org/10.5194/cp-13-1243-2017, 2017.

Stap, L. B., van de Wal, R. S., de Boer, B., Köhler, P., Hoencamp, J. H., Lohmann, G., Tuenter, E., and Lourens, L. J.: Modeled Influence of Land Ice and CO_2 on Polar Amplification and Paleoclimate Sensitivity During the Past 5 Million Years, Paleoceanogr. Paleoclim., 33, 381–394, https://doi.org/10.1002/2017PA003313, 2018.

Wagner, T. J. W. and Eisenman, I.: How Climate Model Complexity Influences Sea Ice Stability, J.Climate, 28, 3998–4014, https://doi.org/10.1175/jcli-d-14-00654.1, 2015.

Zhuang, K., North, G. R., and Giardino, J. R.: Hysteresis of glaciations in the Permo-Carboniferous, J. Geophys. Res.-(Atmo., 119, 6196–6206, https://doi.org/10.1002/2013JD020524, 2014.

Zhuang, K., North, G. R., and Stevens, M. J.: A NetCDF version of the two-dimensional energy balance model based on the full multigrid algorithm, SoftwareX, 6, 198–202, https://doi.org/10.1016/j.softx.2017.07.003, 2017a.

Zhuang, K., North, G. R., and Stevens, M. J.: Model code: A NetCDF version of two-dimensional energy balance climate model based on the full multigrid method in FORTRAN, available at: https://github.com/ElsevierSoftwareX/SOFTX-D-16-00023 (last access: 6 January 2021), 2017b.

Ziegler, E. and Rehfeld, K.: Data and model from TransEBM v. 1.0: Description, tuning, and validation of a transient model of the Earth's energy balance in two dimensions, Zenodo, https://doi.org/10.5281/zenodo.3941311, 2020.

A5 Publication 2: Regional but not global temperature variability underestimated by climate models at supradecadal timescales

T. Laepple, E. Ziegler, N. Weitzel, R. Hébert, B. Ellerhoff, P. Schoch, B. Martrat, O. Bothe, E. Moreno-Chamarro, M. Chevalier, A. Herbert & K. Rehfeld (2023): Regional but not global temperature variability underestimated by climate models at supradecadal timescales. Nature Geoscience, 16. doi:10.1038/s41561-023-01299-9

nature geoscience

Perspective

https://doi.org/10.1038/s41561-023-01299-9

Regional but not global temperature variability underestimated by climate models at supradecadal timescales

Received: 19 January 2023

Accepted: 21 September 2023

Published online: 06 November 2023

Check for updates

T. Laepple[1,2] ✉, E. Ziegler[3,4], N. Weitzel[3], R. Hébert[1], B. Ellerhoff[3,4,13], P. Schoch[5], B. Martrat[6], O. Bothe[7], E. Moreno-Chamarro[8], M. Chevalier[9,10], A. Herbert[11,12] & K. Rehfeld[3,4]

Knowledge of the characteristics of natural climate variability is vital when assessing the range of plausible future climate trajectories in the next decades to centuries. The reliable detection of climate fluctuations on multidecadal to centennial timescales depends on proxy reconstructions and model simulations, as the instrumental record extends back only a few decades in most parts of the world. Systematic comparisons between model-simulated and proxy-based inferences of natural variability, however, often seem contradictory. Locally, simulated temperature variability is consistently smaller on multidecadal and longer timescales than is indicated by proxy-based reconstructions, implying that climate models or proxy interpretations might have deficiencies. In contrast, at global scales, studies found agreement between simulated and proxy reconstructed temperature variations. Here we review the evidence regarding the scale of natural temperature variability during recent millennia. We identify systematic reconstruction deficiencies that may contribute to differing local and global model–proxy agreement but conclude that they are probably insufficient to resolve such discrepancies. Instead, we argue that regional climate variations persisted for longer timescales than climate models simulating past climate states are able to reproduce. This would imply an underestimation of the regional variability on multidecadal and longer timescales and would bias climate projections and attribution studies. Thus, efforts are needed to improve the simulation of natural variability in climate models accompanied by further refining proxy-based inferences of variability.

Climate variability results from a forced component driven by the planetary energy imbalance and from chaotic variations linked to the internal dynamics of the climate system. While the current long-term increase in global mean temperature is mostly due to climate forcings related to human activities[1], internal variability (Box 1) dominates the regional, short-term changes that are most relevant to societies[2,3]. Knowledge of the climate system's variability is required to anticipate the full range of possible future climate change and to increase the impact of societal mitigation and adaptation efforts. While forced variability is relatively well understood[1], regional internal variability is largely uncertain, and society-relevant changes persisting for decades could be outside the range of current projections[4]. Climate models can provide direct estimates of internal variability at all scales[5]. However, uncertainties are introduced by processes not being included and models mostly being

A full list of affiliations appears at the end of the paper. ✉e-mail: tlaepple@awi.de

https://doi.org/10.1038/s41561-023-01299-9

BOX 1

Glossary of terms

Spatial scales: In this study, local refers to a single proxy record or model grid box and regional to any aggregation of multiple proxy records or grid boxes smaller than a hemisphere.

Temporal scales: Subdecadal denotes interannual to decadal timescales, and supradecadal multidecadal to millennial timescales.

Spectrum: For a given time series, the PSD (here referred to simply as spectrum) provides a measure of variability as a function of timescale. For temperature, this spectrum often follows a power law with a scaling coefficient that describes its temporal persistence. A larger scaling implies stronger correlations across timescales and longer persistence times.

Temporal persistence: The strength of (anti-)correlation of successive values in a time series.

Internal climate variability: Climate variability has a forced component and an unforced (that is, internal) component. Internal variability arises from nonlinear interactions between the various feedbacks and components in the climate system, from stochastic fluctuations and from deterministic chaos. As such, it appears in unforced 'control' model simulations. Forced variability results from external drivers of the climate system, irrespective of whether they are anthropogenic or natural. At first order, internal and forced variability are independent, but forced variability can influence the timing and strength of internal climate variability. Natural variability refers to internal and naturally forced.

Climate model: Unless otherwise stated, we use model as shorthand to denote global atmosphere/ocean general circulation models.

Model–data comparison: In the context of model–data comparison, we use data as shorthand to refer to both instrumental observations and proxy-based reconstructions of surface temperature, but not for results produced by numerical models.

validated against recent observational constraints[6]. Furthermore, computational constraints necessitate parameterizations and only allow small ensemble simulations on long timescales that are needed to constrain longer-than-decadal timescales (here termed supradecadal timescales). While instrumental data can be used to estimate the contribution of internal variability for periods of up to several decades, the limited length of these records and the concurrent changes in anthropogenic forcing imply that fluctuations on supradecadal timescales are poorly constrained. Thus, model simulations remain largely unvalidated for variability on multidecadal timescales and beyond. Approaches that exploit observational records are promising[7], but their robustness and scope need to be further improved using palaeoclimate reconstructions to achieve confidence on those timescales.

Current comparisons of temperature variability in model simulations and reconstructions or instrumental observations (hereafter model–data comparisons) seem contradictory. While some studies have found agreement between the variability simulated by climate models and the variability inferred from instrumental observations[8] or proxy-based reconstructions[9–11], others suggest disagreement for the instrumental period[12,13] and past climates[4,14,15].

Typical examples of global and local temperature reconstructions from proxy data and model simulations reveal that these contradictory findings are a result of the distinct nature of global and local climate variability (Fig. 1). Global mean temperature time series from proxy-based reconstructions (Past Global Changes (PAGES) 2k proxy temperature database[16]) and models (the Coupled Model Intercomparison Project Phase 5 (CMIP5)/Paleoclimate Modelling Intercomparison Project Phase III (PMIP3) and Coupled Model Intercomparison Project Phase 6 (CMIP6) last millennium ensemble[17], referred to here as CMIP5/6; Extended Data Table 1) show similar timings and amplitudes of the anomalies throughout most of the past millennium (Fig. 1a) and the power spectra of model simulations and reconstructions are consistent across most timescales (Fig. 1c). In contrast, local temperature variabilities are largely uncorrelated, as illustrated by the Cariaco sea surface temperature (SST) reconstruction[18] and simulated temperatures for the same location (Fig. 1b). The reconstruction exhibits larger amplitudes than individual simulations and even exceeds the ensemble's spread. On decadal and shorter timescales the proxy and model-based power spectra are more consistent (Fig. 1d), but they diverge for supradecadal timescales. There, model spectra flatten, while the local SST record shows a steep increase in the spectral density towards longer timescales. This behaviour is not limited to this example but can also be found in high-resolution records of the PAGES 2k compilation (ref. 19 and Extended Data Fig. 1), in lower-resolution SST records[14] and in land temperature records[4].

To systematically disentangle the structure of agreement and disagreement between model simulations and reconstructions, we performed a meta-study of the existing literature on surface temperature variability in the Holocene and past millennium, following the respective author's original assessments of model–data consistency (Fig. 2 and Extended Data Fig. 2; Methods). At the global scale, there is a high level of agreement on all timescales[8,10–12,16,19,20]. As the spatial scale decreases from global to hemispheric, the results are more mixed. For the Northern Hemisphere some studies found agreement on decadal to centennial timescales[9,11,21], while others reported weaker simulated variability than that reconstructed on interannual to multicentennial scales[22,23]. Owing to the sparsity of proxy coverage of the Southern Hemisphere, there is a clear bias towards the Northern Hemisphere in global and hemispheric studies. At regional and local scales (Box 1), the findings are strongly timescale-dependent. On interannual scales, most studies found model–data agreement[4,8,12,14,19,21–26] or reported mixed results[13,26–29]. Disagreement emerged on decadal[19] or multidecadal[8,12,14] scales. On supradecadal timescales, the majority of studies showed disagreement between reconstructions and models[4,12–14,19,21,25,27] with reconstructions yielding higher magnitudes of temperature variability than simulations.

This leaves us with the enigma of a model–data match at global scales and a mismatch at local and regional scales for supradecadal timescales. Possible causes for this could be shortcomings in the ability of current climate models to simulate local and regional temperature variability, or systematic deficiencies in global or local temperature reconstructions.

Reconstruction deficiencies alone do not explain the scale enigma

Pre-instrumental temperature estimates are primarily reconstructed from physical and biogeochemical proxies measured in natural archives such as glacier ice, sediments and trees[30]. This proxy information is transformed into estimates of past temperatures using statistical methods. While many sources of uncertainties and biases have been identified in these statistical reconstructions[31–34], we focus here on potential underlying causes of the diagnosed scale enigma.

Studies that have found consistent global variability in reconstructions and simulations typically employed multi-proxy databases comprising predominantly annually resolved records. These have

https://doi.org/10.1038/s41561-023-01299-9

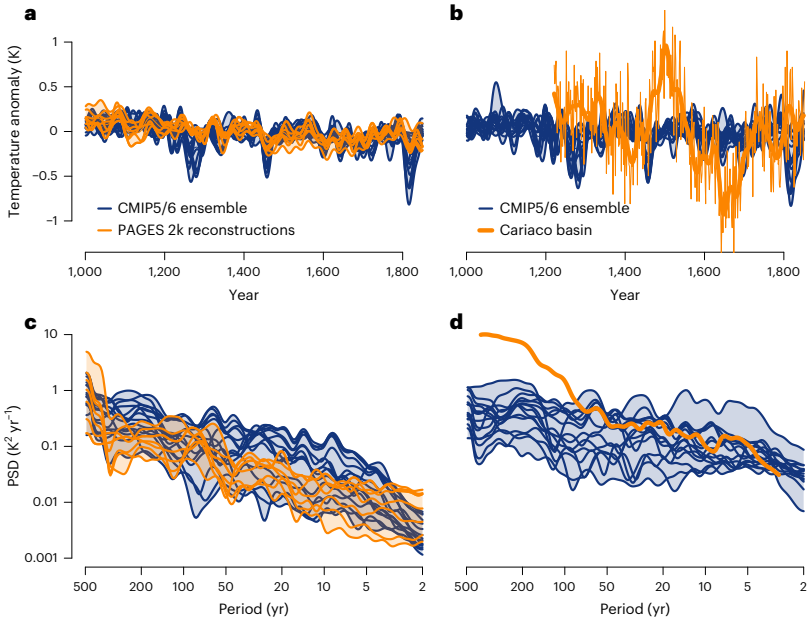


Fig. 1 | Examples of local and global surface temperature variability in models and reconstructions over the 1,000–1,800 CE time span. a, Time series of global mean surface temperature from CMIP5/6 model simulations and all PAGES 2k proxy reconstructions[16]. **b**, Time series of a local SST reconstruction (Cariaco[18]) and of model simulations for the same location (*y*-axis as in (**a**)). The thin lines show annual values, thick lines are low-pass filtered (cutoff frequency 1/30 yr) to focus on the supradecadal variations. **c**,**d**, Power spectral density (PSD) of the global (**c**) and local (**d**) simulated and reconstructed temperature time series on shared *y*-axis. Coloured shading indicates the range of the model and proxy reconstruction ensemble.

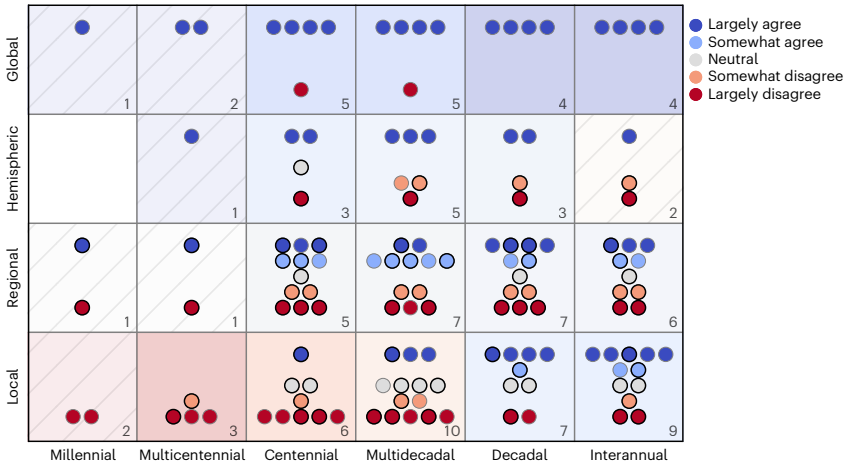


Fig. 2 | Model–data (dis)agreement on Holocene temperature variability in the literature. Agreement as a function of temporal (*x* axis) and spatial (*y* axis) scales. Each circle represents a specific study, colour coded according to the strength of (dis)agreement (Methods). Some studies report different rates of agreement at one spatiotemporal scale; that is, depending on properties such as region or proxy type. Such multiple occurrences in one box are marked by a thick line around the circle. The number of distinct studies is given at the bottom right of each box. Diagonal lines indicate a small number of studies (one or two) for the given spatiotemporal scale. See Extended Data Fig. 2 indicating the specific literature studies (refs. 4,8–14,16,19–2997).

been calibrated using the period that overlaps with the instrumental record (temporal calibration). A number of studies have suggested that global temperature reconstructions might underestimate variability (such as ref. 34 and references therein). The temporal calibration step can underestimate supradecadal variability, particularly when the proxy is used as predictor (direct regression)[35]. Furthermore, current multi-proxy databases are dominated by tree ring records (for example ~60% in the PAGES 2k 2017 database) that can underestimate

supradecadal variability due to the removal of individual growth trends[36]. These arguments have sparked the development of new methods including indirect regression (which avoids the underestimation of low-frequency variability in the calibration step), ensemble methods and Bayesian models[37]. Pseudo-proxy experiments suggest that these new methods do not substantially underestimate supradecadal variability[31]. While the few reconstructions that do not rely on tree ring records exhibit the highest supradecadal variability[38,39] and differing methodologies can produce a spread of up to one order of magnitude, even the global reconstructions with the highest magnitude variability are still consistent with the ensemble of CMIP5/6 simulations (Fig. 1a,c).

Unlike global reconstructions, studies focusing on local scales often rely on proxies from sedimentary archives (lake and marine sediments, glacier ice) that are calibrated using independent data from laboratory experiments[40], physical models[41] or modern spatial relationships[42]. These calibrations are not subject to the underestimation of supradecadal variability found for direct temporal regression[35]. Instead, 'proxy noise' (that is, the variability enhancement from non-climatic processes) has been proposed as the most likely culprit for overestimations of supradecadal variability in these local reconstructions. Examples of such processes are aliasing of the seasonal cycle[43], spatial heterogeneity in the archive[44], recording of site-specific features such as upwelling events that are not representative of mean conditions[45], vital effects or measurement errors. As local variability estimates typically rely on single records, there can be no dampening of proxy noise through averaging. In contrast to these variability-enhancing mechanisms, other non-climatic processes such as bioturbation of sediments[46], isotopic diffusion in ice[47], slow response times of biological systems[48] and homeostasis[49] can lead to timescale-dependent underestimation of local variability. Biases from proxy noise and timescale-dependent temperature sensitivities can be minimized through statistical corrections[50] and the use of proxy system models[51]. These corrections lead to consistent estimates of variability across independent proxy types and with instrumental observations on overlapping timescales[4,14], strengthening the credibility of reconstructed variability. Nonetheless, corrected estimates still show the model–data variability mismatch on supradecadal scales[14,27]. Supporting the evidence in the literature, we found that local variability reconstructed from instrumentally calibrated annually resolved records is consistent with marine and terrestrial proxies and displays a similar model–data variability mismatch on supradecadal timescales (Extended Data Fig. 1).

Chronological uncertainties can complicate the comparison of individual events in proxies and simulations. However, age model errors have little influence on local variability estimates from single records[52]. While averaging across records with age uncertainty leads to an underestimation of variability[4], the age uncertainties for annually resolved records of the Common Era is usually below 10 years (ref. 53). Thus, we expect little to no effect on supradecadal global temperature variability estimates. It is often not possible to unambiguously attribute proxy variations to a specific physical variable (for example, trees can be sensitive to temperature or moisture) or season. However, the mismatch for local supradecadal variability would be even higher if we assumed that proxies represented hydroclimate, instead of temperature variations (Extended Data Fig. 3). In addition, the local model–data disagreement is also found in geochemical proxies that are only sensitive to temperature (for example, Mg/Ca ratios in foraminifera or Sr/Ca ratios in corals). Taken together, it is unlikely that chronological uncertainties or hydroclimate influences on proxies contribute substantially to the scale enigma.

Finally, local proxies may be more variable than model simulations due to differences in spatial scales between local proxy records and a typical grid-box size in climate models (~100 km). However, for surface temperature, the spatial decorrelation length on interannual scales is usually larger than the size of a model grid box, which limits the impact of differences in spatial scale[54]. Moreover, as the spatial coherence of temperature variability increases between interannual and supradecadal timescales[55], discrepancies are expected to diminish towards longer timescales.

The described methodological differences between global and local reconstructions complicate the assessment of systematic biases. Nevertheless, current evidence suggests that neither an overestimation of local variability nor an underestimation of global variability from reconstructions can solely explain the diagnosed scale enigma. Therefore, we argue that model deficiencies substantially contribute to the enigma.

Consequences for the spatial structure of natural climate variability

As a possible explanation for the scale enigma, we hypothesize that while the current generation of climate models can accurately simulate global-scale temperature variability, they are too stable at smaller (local) scales. If correct, this hypothesis would imply that climate models need more spatially independent modes of supradecadal variability.

Global and local variability are not independent, they are connected by the spatial correlation structure of the temperature field. Stronger spatial correlations lead to smaller differences between regional and global temperature variability and vice versa[56]. Comparing the spectra of simulated local and global variability suggests that the ratio between both quantities decreases with longer timescales (Fig. 3a, grey shaded area). This is associated with an increase in the typical spatial extent of temperature anomalies on longer timescales[56] and is supported by instrumental observations on subdecadal scales[12,55]. The relationship between temporal and spatial scales can also be visualized by subtracting two mean temperature fields from successive time steps from one another (for example, from two adjacent years (Fig. 3b), from adjacent decades (Fig. 3c) or from even longer periods (Fig. 3d)). In the CMIP5/6 model simulations, the local and global variability magnitudes converge as timescales increase (grey shading in Fig. 3a), which is associated with a continuously increasing spatial scale of the temperature fluctuations (Fig. 3b–d).

This increase in the spatial extent of temperature anomalies on supradecadal timescales might be an artefact introduced by the climate models, as the reconstructed local temperature variability on these timescales is larger than the simulated variability (Figs. 1 and 2). By adding published estimates of the local variability spectrum from PAGES 2k reconstructions[19] and marine and lake sediment cores[4,14] (Fig. 3a, blue dashed line; Methods) to the model spectrum, we tentatively deduce that the observed supradecadal local variability increases in parallel with the global variability. This would lead to a constant ratio between local and global variability (vertical dashed shading) and imply that, in contrast to the model results, the spatial scale of supradecadal temperature fluctuations might not increase after all. This prompts us to speculate that the true temperature anomalies between centuries might be more similar in structure and amplitude to the simulated temperature anomalies between decades (Fig. 3c), rather than the simulated anomalies between centuries (Fig. 3d).

This implied invariance in the spatial extent of temperature anomalies on supradecadal timescales according to the reconstructions, contrasting the further increase suggested by the models, may result from a different balance between two competing processes: (1) those creating local climate variability and (2) those causing a timescale-dependent suppression of local variability.

It has been proposed that the amplitude of natural forcing or the response to that forcing could be underestimated or be too spatially homogeneous due to shortcomings in the dynamical response in models[57]. Missing components of the climate system such as interactive ice sheets and dynamic vegetation[58–62] might further cause a lack of regional variability, especially when the mechanisms causing the variability are linked to timescale-invariant spatial structures such as the

https://doi.org/10.1038/s41561-023-01299-9

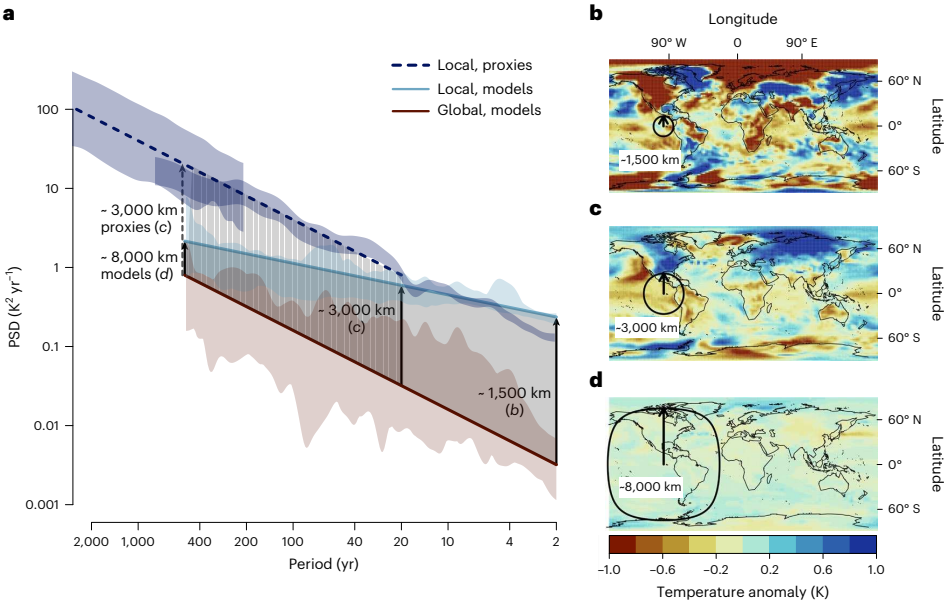


Fig. 3 | Local versus global temperature variability and its implied spatial scale. a, Global (brown) and mean local (light blue) PSD estimates from CMIP5/6 and proxy-based mean local spectra (dark blue) (refs. 4,14,19 and Methods). Lines show slopes PSD ≈ f^{-1} (red and dashed dark blue) and PSD ≈ $f^{-0.4}$ (light blue) for frequencies f; the shading shows the range of all estimates. The ratio between local and global spectra (represented by a distance in this plot as due to the logarithmic y-axis) relates to the typical sizes of temperature anomalies[56] (vertical arrows). For models, the ratio decreases with increasing timescales (grey shading), whereas according to proxies it remains constant on supradecadal timescales (vertical dashed grey shading). **b**–**d**, Typical model temperature anomaly fields for 2 yr (**b**), 20 yr (**c**) and 400 yr (**d**) timescales. Circles show the typical spatial extents of temperature anomalies[55]. We hypothesize that real-world temperature anomalies on supradecadal timescales might be closer to those shown in **c** than those shown in **d**.

topography or land–sea boundaries. On the other hand, horizontal diffusion[63] reduces regional variability at low frequencies. Hence, overly weak or misrepresented ocean eddies[64] and inadequacies in the representation and spatiotemporal propagation of sub-grid-scale processes in current climate models could artificially dampen supradecadal temperature variabilities at regional scales. One notable limitation of the current CMIP protocols for past simulations that could contribute to the local variability deficit of simulations is the lack of knowledge regarding the dynamic initial state of the ocean. The Earth system has a long memory for past forcings[65] and the supradecadal variability has been suggested to echo slowly varying boundary conditions, mainly orbital forcing[10]. Meanwhile, simulations of the past millennium are generally initialized in a state of quasi-equilibrium with the starting boundary conditions.

Our hypothesis of a stronger persistence of regional temperature variability is consistent with other observational and proxy evidence. A lack of temporal persistence resulting in too much fast variation relative to to slow variation of regional temperatures in climate models[66,67] and a local response to forcing that is too weak were proposed as explanations for the so-called 'signal-to-noise paradox'[66]. This denotes the perplexing finding that models predict the observed (regional) temperature variations better than they predict ensemble forecast members. Finally, proxy-based reconstructions of temperature anomalies for specific periods in the past often exhibit more complex spatial patterns than the corresponding fields from simulations[68–70].

Implications for climate projections and attribution efforts

Misrepresented variability would decrease the skill of policy-relevant climate and extreme event projections[71,72], impact assessments[73,74] and attribution studies[75]. The ability of forced climate model simulations to reproduce the magnitude and structure of global temperature variability (Fig. 2) strengthens the confidence in the simulated response to external forcing[9,76] in future climate projections. In contrast, biases in the spatial structure of variability would affect regional climate projections at supradecadal timescales. Likewise, projections might miss abrupt climate transitions as many climate models have been shown to be too stable to accurately simulate such transitions in the past[77]. However, recent progress in Earth system models[60,78] regarding the representation of these (in)stabilities of the climate system is promising.

Long-term variability modulates the occurrence of extreme weather and climate events[72,79], and stronger temporal persistence can increase the frequency of occurrence of such extremes (Fig. 4). As a result, current model-based loss and damage estimates from extreme events and risk projections may be underestimations. This may also affect short-term projections as the return periods of extreme events might be misrepresented (Fig. 4c). Finally, our reliance on instrumental observations can bias the distribution estimates of climate variables (Fig. 4b) as this period might be too short to sample the full range of supradecadal variability. Stronger persistence of climate variations would increase the probability of extreme conditions (Fig. 4a), including very rare and intense events[80]. Compound events may also occur more often, intensifying the vulnerability of ecosystems and societies, and the risk of sustained impacts such as agricultural losses and wildfires[81]. Socio-economic long-term planning for climate change requires a better understanding of this linkage between variability and extreme events[82].

Attribution studies provide vital insights into the impacts of anthropogenic greenhouse gas emissions on current and future climate change, including extreme climate and weather events. The reliability

https://doi.org/10.1038/s41561-023-01299-9

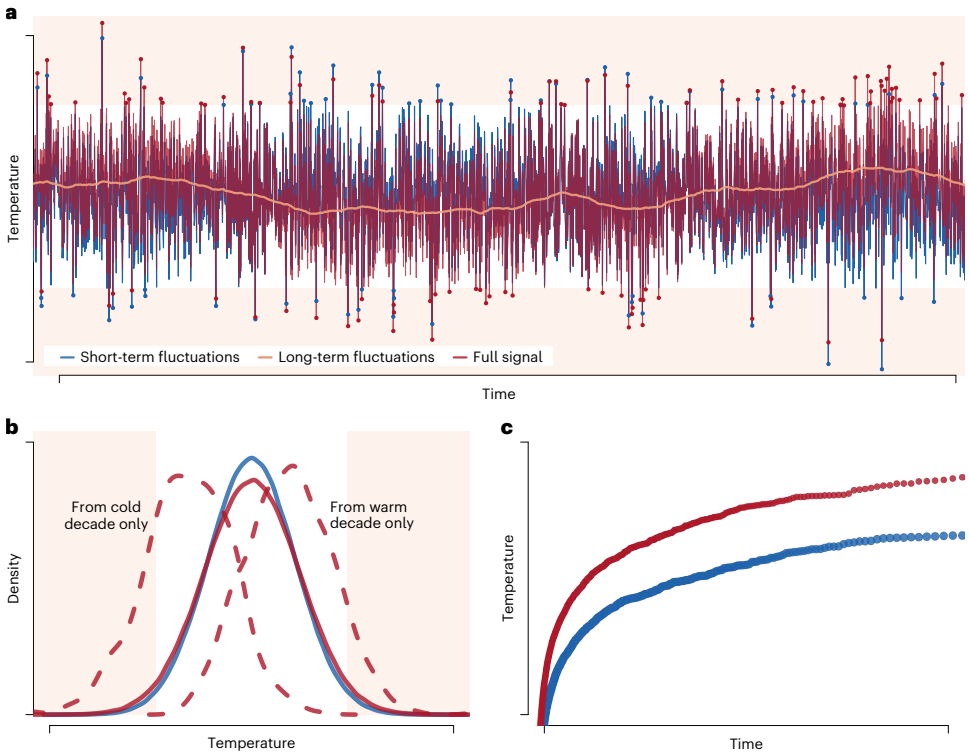


Fig. 4 | Slow variability increases the magnitude of extreme events and decreases their return time in a conceptual time series. a, Conceptual time series of a white noise process representing short-term (that is, weather; light blue) and an autocorrelated process representing long-term fluctuations (that is, decadal-to-centennial variability; orange). The sum of both (red) exemplifies a typical climate time series of 50 years (Methods). Dots indicate extremes more than two standard deviations from the mean (shaded regions). **b**, Probability density function of short-term and full signals (solid lines) and of the full signal estimated from an exceptionally cold and warm decade (dashed lines). **c**, Return levels from the short-term and full signals.

of these assessments depends on accurate representations of the statistical properties of the climate and its decomposition into anthropogenic and natural contributions[83,84], which makes these assessments susceptible to misrepresentations of variability in models[75]. Indeed, extreme event attribution often assumes that decadal and longer internal variability plays only a minor role in the occurrence of extremes[85].

Discrepancies between modelled and observed changes in climatic parameters complicate regional attribution[86] as most attribution studies focus precisely on those local to regional spatial scales at which the supradecadal variability is probably underestimated in models. Thus, if strong and persistent local variability is missing in models, tests for the significance of anthropogenic effects are likely to be biased positive. These arguments highlight the urgent need for an improved understanding of variability and its spatial structure, including the relationship between local and global variability estimates. We could then leverage this new understanding to validate and improve the capabilities of climate models regarding regional projections, for example, through emergent constraints[87], downscaling or bias-correction techniques[88] targeting variability. Incorporating stochastic natural forcing in climate model simulations of future scenarios is another important step in climate modelling for realistic projections of future climate variability[89,90].

Regional projections, risk assessments of extremes and attribution all depend on reliable simulations of local and regional variability. Without an improved representation of variability in climate models, the range of future conditions might not be fully covered. Thus, policies that are based on these estimates will not address the actual required mitigation and adaptation strategies. Overall, underestimation of the full temperature range can lead to a false sense of security (during a decade on the lower end of the range, for example) (Fig. 4), as well as to decades of unexpectedly high temperatures[1] for which policymakers and societies are unprepared.

Closing the existing knowledge gap

To resolve the enigma of global climate models and proxy-based climate reconstructions being in agreement regarding the global average continuum of temperature variability but not at smaller (local to regional) scales, we have argued that regional-scale fluctuations in climate models should be larger and persist for longer. As a result, regional temperatures in current models probably seem too stable on supradecadal timescales and internal variability constitutes only a small fraction of their projection uncertainties, even at regional scales[91]. Additional efforts are required to ensure that natural (forced and internal) variability is simulated accurately to better evaluate its contribution to regional projection uncertainty. This includes investigating long model simulations with transient boundary conditions and simulations with additional coupled Earth system components in terms of their simulated supradecadal variability.

Validation of climate models with respect to palaeoclimate data usually focuses on their ability to simulate the mean state in a given time period such as the Last Glacial Maximum or slow changes, but rarely with respect to variability across timescales; this requires

both a paradigm shift in our approach to validating models and a new data product. In the 1970s, Climate: Long range Investigation, Mapping, and Prediction[92] became a landmark project, providing the first spatially comprehensive estimates of climate change and mapping the climate and ecology of the Last Glacial Maximum. An undertaking of similar magnitude concerning natural variability is now necessary and achievable. It requires the participation of the palaeoclimate, proxy-development and statistics communities. While we already have a large number of palaeoclimate reconstructions that cover several continents and oceanic basins, as well as timescales that range from annual to millennia, this dataset needs to be extended to improve replicability[93] and reduce the number and size of gaps in spatial, temporal and timescale coverage. While this may seem to be a considerable challenge, recent advances in analytical techniques such as mass spectral imaging[94] or laser spectroscopy[95] can be exploited to help achieve this, as these methods allow faster, more efficient and higher-resolution processing of proxy data. We need to combine community-driven synthesis and production of proxy datasets with a process-based understanding of the recording of climate signals in different types of archive[30] based on more site-level monitoring data. This will facilitate the separation of the climate signal from non-climatic noise, the separation of temperature and hydroclimate variability and the identification of the timescales at which the climate signal dominates variability. We also need to further develop and optimize statistical methods to combine the diverse information into reliable and unbiased variability estimates. This effort would facilitate spatially comprehensive estimates of variability as a function of timescale, ideally both for surface temperature and hydroclimatic variables.

Such advances are crucial for benchmarking and improving the representation of variability in climate models and the quantification of future regional climate risks linked to natural variability. Making use of distinct spatial fingerprints of the different mechanisms producing supradecadal variability would help identify the origins of the currently missing local supradecadal variability in models. Furthermore, these advances can provide the basis for characterizing the spatiotemporal covariance structure of climate variability. This in turn will improve the skill of data assimilation[96] and climate field reconstruction methods, permitting us to test our hypothesis for the scale enigma.

Finally, variability is a critical criterion of model development that should be evaluated in validating models and as an emergent property. This would allow us to investigate a broader range of possible model sensitivities and predictions of climate variability at various spatiotemporal scales. The proposed measures are vital for reliable estimates of variability, which in turn advance more robust regional projections, assessments of model-independent constraints and attribution and impact studies. The improved understanding of climate variability would thus contribute to better climate policy choices and mitigation and adaptation efforts.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41561-023-01299-9.

References

1. IPCC *Climate Change 2021: The Physical Science Basis* (eds Masson-Delmotte, V. et al.) (Cambridge Univ. Press, 2021).
2. Degroot, D. et al. Towards a rigorous understanding of societal responses to climate change. *Nature* **591**, 539–550 (2021).
3. Deser, C., Phillips, A., Bourdette, V. & Teng, H. Uncertainty in climate change projections: the role of internal variability. *Clim. Dynam.* **38**, 527–546 (2012).
4. Hébert, R., Herzschuh, U. & Laepple, T. Millennial-scale climate variability over land overprinted by ocean temperature fluctuations. *Nat. Geosci.* **15**, 899–905 (2022).
5. Maher, N., Lehner, F. & Marotzke, J. Quantifying the role of internal variability in the temperature we expect to observe in the coming decades. *Environ. Res. Lett.* **15**, 054014 (2020).
6. Hourdin, F. et al. The art and science of climate model tuning. *Bull. Am. Meteorol. Soc.* **98**, 589–602 (2017).
7. McKinnon, K. A. & Deser, C. Internal variability and regional climate trends in an observational large ensemble. *J. Clim.* **31**, 6783–6802 (2018).
8. Fredriksen, H.-B. & Rypdal, K. Spectral characteristics of instrumental and climate model surface temperatures. *J. Clim.* **29**, 1253–1268 (2016).
9. Crowley, T. J. Causes of climate change over the past 1000 years. *Science* **289**, 270–277 (2000).
10. Zhu, F. et al. Climate models can correctly simulate the continuum of global-average temperature variability. *Proc. Natl Acad. Sci. USA* **116**, 8728–8733 (2019).
11. Fernández-Donado, L. et al. Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium. *Clim. Past* **9**, 393–421 (2013).
12. Laepple, T. & Huybers, P. Global and regional variability in marine surface temperatures. *Geophys. Res. Lett.* **41**, 2528–2534 (2014).
13. Parsons, L. A. et al. Temperature and precipitation variance in CMIP5 simulations and paleoclimate records of the last millennium. *J. Clim.* **30**, 8885–8912 (2017).
14. Laepple, T. & Huybers, P. Ocean surface temperature variability: large model–data differences at decadal and longer periods. *Proc. Natl Acad. Sci. USA* **111**, 16682–16687 (2014).
15. Rehfeld, K., Münch, T., Ho, S. L. & Laepple, T. Global patterns of declining temperature variability from the Last Glacial Maximum to the Holocene. *Nature* **554**, 356–359 (2018).
16. Neukom, R. et al. Consistent multidecadal variability in global temperature reconstructions and simulations over the Common Era. *Nat. Geosci.* **12**, 643–649 (2019).
17. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
18. Black, D. E. et al. An 8-century tropical Atlantic SST record from the Cariaco Basin: baseline variability, twentieth-century warming, and Atlantic hurricane frequency. *Paleoceanogr. Palaeoclimatol.* **22**, PA4204 (2007).
19. Ellerhoff, B. & Rehfeld, K. Probing the timescale dependency of local and global variations in surface air temperature from climate simulations and reconstructions of the last millennia. *Phys. Rev. E* **104**, 064136 (2021).
20. Askjær, T. G. et al. Multi-centennial Holocene climate variability in proxy records and transient model simulations. *Quat. Sci. Rev.* **296**, 107801 (2022).
21. Cheung, A. H. et al. Comparison of low-frequency internal climate variability in CMIP5 models and observations. *J. Clim.* **30**, 4763–4776 (2017).
22. Bothe, O., Jungclaus, J. H. & Zanchettin, D. Consistency of the multi-model CMIP5/PMIP3-past1000 ensemble. *Clim. Past* **9**, 2471–2487 (2013).
23. Collins, M., Osborn, T. J., Tett, S. F. B., Briffa, K. R. & Schweingruber, F. H. A comparison of the variability of a climate model with paleotemperature estimates from a network of tree-ring densities. *J. Clim.* **15**, 1497–1515 (2002).
24. Ault, T. R., Deser, C., Newman, M. & Emile-Geay, J. Characterizing decadal to centennial variability in the equatorial Pacific during the last millennium. *Geophys. Res. Lett.* **40**, 3450–3456 (2013).

25. Bühler, J. C. et al. Comparison of the oxygen isotope signatures in speleothem records and iHadCM3 model simulations for the last millennium. *Clim. Past* **17**, 985–1004 (2021).
26. Zorita, E. et al. European temperature records of the past five centuries based on documentary/instrumental information compared to climate simulations. *Climatic Change* **101**, 143–168 (2010).
27. Dee, S. G. et al. Improved spectral comparisons of paleoclimate models and observations via proxy system modeling: Implications for multi-decadal variability. *Earth Planet. Sci. Lett.* **476**, 34–46 (2017).
28. Franke, J., Frank, D., Raible, C. C., Esper, J. & Brönnimann, S. Spectral biases in tree-ring climate proxies. *Nat. Clim. Change* **3**, 360–364 (2013).
29. PAGES 2k-PMIP3 group. Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium. *Clim. Past* **11**, 1673–1699 (2015).
30. Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M. & Anchukaitis, K. J. Applications of proxy system modeling in high resolution paleoclimatology. *Quat. Sci. Rev.* **76**, 16–28 (2013).
31. Anchukaitis, K. J. & Smerdon, J. E. Progress and uncertainties in global and hemispheric temperature reconstructions of the Common Era. *Quat. Sci. Rev.* **286**, 107537 (2022).
32. Esper, J., Frank, D. C. & Wilson, R. J. S. Climate reconstructions: low-frequency ambition and high-frequency ratification. *Eos* **85**, 113–120 (2004).
33. Kunz, T., Dolman, A. M. & Laepple, T. A spectral approach to estimating the timescale-dependent uncertainty of paleoclimate records – part 1: theoretical concept. *Clim. Past* **16**, 1469–1492 (2020).
34. Christiansen, B. & Ljungqvist, F. C. Challenges and perspectives for large-scale temperature reconstructions of the past two millennia. *Rev. Geophys.* **55**, 40–96 (2017).
35. Osborn, T. J. CLIMATE: the real color of climate change? *Science* **306**, 621–622 (2004).
36. Cook, E. R., Briffa, K. R., Meko, D. M., Graybill, D. A. & Funkhouser, G. The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies. *Holocene* **5**, 229–237 (1995).
37. Tingley, M. P. & Huybers, P. A Bayesian algorithm for reconstructing climate anomalies in space and time. part i: development and applications to paleoclimate reconstruction problems. *J. Clim.* **23**, 2759–2781 (2009).
38. Moberg, A., Mohammad, R. & Mauritsen, T. Analysis of the Moberg et al. (2005) hemispheric temperature reconstruction. *Clim. Dynam.* **31**, 957–971 (2008).
39. Trouet, V. et al. A 1500-year reconstruction of annual mean temperature for temperate North America on decadal-to-multidecadal time scales. *Environ. Res. Lett.* **8**, 024008 (2013).
40. Kim, S.-T. & O'Neil, J. R. Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates. *Geochim. Cosmochim. Acta* **61**, 3461–3475 (1997).
41. Werner, M., Mikolajewicz, U., Heimann, M. & Hoffmann, G. Borehole versus isotope temperatures on Greenland: seasonality does matter. *Geophys. Res. Lett.* **27**, 723–726 (2000).
42. Müller, P. J., Kirst, G., Ruhland, G., Von Storch, I. & Rosell-Melé, A. Calibration of the alkenone paleotemperature index $U_{37}^{K'}$ based on core-tops from the eastern South Atlantic and the global ocean (60°N-60°S). *Geochim. Cosmochim. Acta* **62**, 1757–1772 (1998).
43. Laepple, T. et al. On the similarity and apparent cycles of isotopic variations in East Antarctic snow pits. *Cryosphere* **12**, 169–187 (2018).
44. Zuhr, A. M. et al. Age-heterogeneity in marine sediments revealed by three-dimensional high-resolution radiocarbon measurements. *Front. Earth Sci.* https://doi.org/10.3389/feart.2022.871902 (2022).
45. Peeters, F. J. C., Brummer, G.-J. A. & Ganssen, G. The effect of upwelling on the distribution and stable isotope composition of *Globigerina bulloides* and *Globigerinoides ruber* (planktic foraminifera) in modern surface waters of the NW Arabian Sea. *Glob. Planet. Change* **34**, 269–291 (2002).
46. Berger, W. H. & Heath, G. R. Vertical mixing in pelagic sediments. *J. Mar. Res.* **26**, 134–143 (1968).
47. Johnsen, S. J. in *Isotopes and Impurities in Snow and Ice* Publication No. 118, 210–219 (IAHS-AISH, 1977).
48. Webb, T. Is vegetation in equilibrium with climate? How to interpret late-Quaternary pollen data. *Vegetatio* **67**, 75–91 (1986).
49. Mix, A. in *North America and Adjacent Oceans During the Last Deglaciation* Vol. K-3, 111–135 (Geological Society of America, 1987).
50. Laepple, T. & Huybers, P. Reconciling discrepancies between Uk37 and Mg/Ca reconstructions of Holocene marine temperature variability. *Earth Planet. Sci. Lett.* **375**, 418–429 (2013).
51. Dee, S. et al. PRYSM: an open-source framework for proxy system modeling, with applications to oxygen-isotope systems. *J. Adv. Model. Earth Syst.* **7**, 1220–1247 (2015).
52. Rhines, A. & Huybers, P. Estimation of spectral power laws in time uncertain series of data with application to the Greenland Ice Sheet Project 2 $\delta^{18}O$ record. *J. Geophys. Res. Atmos.* **116**, D01103 (2011).
53. Sigl, M. et al. Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature* **523**, 543–549 (2015).
54. North, G. R., Wang, J. & Genton, M. G. Correlation models for temperature fields. *J. Clim.* **24**, 5850–5862 (2011).
55. Jones, P. D., Osborn, T. J. & Briffa, K. R. Estimating sampling errors in large-scale temperature averages. *J. Clim.* **10**, 2548–2568 (1997).
56. Kunz, T. & Laepple, T. Frequency-dependent estimation of effective spatial degrees of freedom. *J. Clim.* **34**, 7373–7388 (2021).
57. Shindell, D. T., Schmidt, G. A., Mann, M. E., Rind, D. & Waple, A. Solar forcing of regional climate change during the Maunder Minimum. *Science* **294**, 2149–2152 (2001).
58. Bakker, P., Clark, P. U., Golledge, N. R., Schmittner, A. & Weber, M. E. Centennial-scale Holocene climate variations amplified by Antarctic Ice Sheet discharge. *Nature* **541**, 72–76 (2017).
59. Braconnot, P., Zhu, D., Marti, O. & Servonnat, J. Strengths and challenges for transient Mid- to Late Holocene simulations with dynamical vegetation. *Clim. Past* **15**, 997–1024 (2019).
60. Hopcroft, P. O. & Valdes, P. J. Paleoclimate-conditioning reveals a North Africa land–atmosphere tipping point. *Proc. Natl Acad. Sci. USA* **118**, e2108783118 (2021).
61. Bonan, G. B. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* **320**, 1444–1449 (2008).
62. Laguë, M. M., Bonan, G. B. & Swann, A. L. S. Separating the impact of individual land surface properties on the terrestrial surface energy budget in both the coupled and uncoupled land–atmosphere system. *J. Clim.* **32**, 5725–5744 (2019).
63. Rypdal, K., Rypdal, M. & Fredriksen, H.-B. Spatiotemporal long-range persistence in Earth's temperature field: analysis of stochastic–diffusive energy balance models. *J. Clim.* **28**, 8379–8395 (2015).
64. Jüling, A., von der Heydt, A. & Dijkstra, H. A. Effects of strongly eddying oceans on multidecadal climate variability in the Community Earth System Model. *Ocean Sci.* **17**, 1251–1271 (2021).
65. Rypdal, M. & Rypdal, K. Long-memory effects in linear response models of Earth's temperature and implications for future global warming. *J. Clim.* **27**, 5240–5258 (2014).

66. Sevellec, F. & Drijfhout, S. S. The signal-to-noise paradox for interannual surface atmospheric temperature predictions. *Geophys. Res. Lett.* **46**, 9031–9041 (2019).
67. Strommen, K. & Palmer, T. N. Signal and noise in regime systems: a hypothesis on the predictability of the North Atlantic Oscillation. *Q. J. R. Meteorol. Soc.* **145**, 147–163 (2019).
68. Mann, M. E. et al. Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. *Science* **326**, 1256–1260 (2009).
69. Hargreaves, J. C., Annan, J. D., Ohgaito, R., Paul, A. & Abe-Ouchi, A. Skill and reliability of climate model ensembles at the Last Glacial Maximum and mid-Holocene. *Clim. Past* **9**, 811–823 (2013).
70. Weitzel, N., Hense, A. & Ohlwein, C. Combining a pollen and macrofossil synthesis with climate simulations for spatial reconstructions of European climate using Bayesian filtering. *Clim. Past* **15**, 1275–1301 (2019).
71. Blanusa, M. L., López-Zurita, C. J., & Rasp, S. Internal variability plays a dominant role in global climate projections of temperature and precipitation extremes. *Climate Dynamics* **61**, 1931–1945 (2023).
72. Ionita, M., Dima, M., Nagavciuc, V., Scholz, P. & Lohmann, G. Past megadroughts in central Europe were longer, more severe and less warm than modern droughts. *Commun. Earth Environ.* **2**, 61 (2021).
73. Calel, R., Chapman, S. C., Stainforth, D. A. & Watkins, N. W. Temperature variability implies greater economic damages from climate change. *Nat. Commun.* **11**, 5028 (2020).
74. Schwarzwald, K. & Lenssen, N. The importance of internal climate variability in climate impact projections. *Proc. Natl Acad. Sci. USA* **119**, e2208095119 (2022).
75. Harrington, L. J., Schleussner, C.-F. & Otto, F. E. L. Quantifying uncertainty in aggregated climate change risk assessments. *Nat. Commun.* **12**, 7140 (2021).
76. Hausfather, Z., Drake, H. F., Abbott, T. & Schmidt, G. A. Evaluating the performance of past climate model projections. *Geophys. Res. Lett.* **47**, e2019GL085378 (2020).
77. Valdes, P. Built for stability. *Nat. Geosci.* **4**, 414–416 (2011).
78. Klockmann, M., Mikolajewicz, U., Kleppin, H. & Marotzke, J. Coupling of the subpolar gyre and the overturning circulation during abrupt glacial climate transitions. *Geophys. Res. Lett.* **47**, e2020GL090361 (2020).
79. Czymzik, M., Muscheler, R. & Brauer, A. Solar modulation of flood frequency in central Europe during spring and summer on interannual to multi-centennial timescales. *Clim. Past* **12**, 799–805 (2016).
80. Yan, M. & Liu, J. Physical processes of cooling and mega-drought during the 4.2 ka BP event: results from TraCE-21ka simulations. *Clim. Past* **15**, 265–277 (2019).
81. Zscheischler, J. et al. A typology of compound weather and climate events. *Nat. Rev. Earth Environ.* **1**, 333–347 (2020).
82. Deser, C., Knutti, R., Solomon, S. & Phillips, A. S. Communication of the role of natural variability in future North American climate. *Nat. Clim. Change* **2**, 775–779 (2012).
83. Hegerl, G. & Zwiers, F. Use of models in detection and attribution of climate change. *WIREs Clim. Change* **2**, 570–591 (2011).
84. Stott, P. A. et al. Observational constraints on past attributable warming and predictions of future global warming. *J. Clim.* **19**, 3055–3069 (2006).
85. Philip, S. et al. A protocol for probabilistic extreme event attribution analyses. *Adv. Stat. Climatol. Meteorol. Oceanogr.* **6**, 177–203 (2020).
86. van Oldenborgh, G. J. et al. Pathways and pitfalls in extreme event attribution. *Climatic Change* **166**, 13 (2021).
87. Qasmi, S. & Ribes, A. Reducing uncertainty in local temperature projections. *Sci. Adv.* **8**, eabo6872 (2022).
88. Wu, Y. et al. Quantifying the uncertainty sources of future climate projections and narrowing uncertainties with bias correction techniques. *Earth Future* **10**, e2022EF002963 (2022).
89. Bethke, I. et al. Potential volcanic impacts on future climate variability. *Nat. Clim. Change* **7**, 799–805 (2017).
90. Ellerhoff, B. et al. Contrasting state-dependent effects of natural forcing on global and local climate variability. *Geophys. Res. Lett.* **49**, e2022GL098335 (2022).
91. Lehner, F. et al. Partitioning climate projection uncertainty with multiple large ensembles and CMIP5/6. *Earth Syst. Dynam.* **11**, 491–508 (2020).
92. McIntyre, A. et al. *Seasonal Reconstructions of the Earth's Surface at the Last Glacial Maximum* (Geological Society of America, 1981).
93. Comboul, M., Emile-Geay, J., Hakim, G. J. & Evans, M. N. Paleoclimate sampling as a sensor placement problem. *J. Clim.* **28**, 7717–7740 (2015).
94. Wörmer, L. et al. Ultra-high-resolution paleoenvironmental records via direct laser-based analysis of lipid biomarkers in sediment core samples. *Proc. Natl Acad. Sci. USA* **111**, 15669–15674 (2014).
95. Barkan, E. & Luz, B. High precision measurements of $^{17}O/^{16}O$ and $^{18}O/^{16}O$ ratios in H_2O. *Rapid Commun. Mass Spectrom.* **19**, 3737–3742 (2005).
96. Amrhein, D. E., Hakim, G. J. & Parsons, L. A. Quantifying structural uncertainty in paleoclimate data assimilation with an application to the last millennium. *Geophys. Res. Lett.* **47**, e2020GL090485 (2020).
97. Ljungqvist, F. C. et al. Centennial-scale temperature change in last millennium simulations and proxy-based reconstructions. *J. Clim.* **32**, 2441–2482 (2019).

[1]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Potsdam, Germany. [2]MARUM–Center for Marine Environmental Sciences and Faculty of Geosciences, University of Bremen, Bremen, Germany. [3]Department of Geosciences, University of Tübingen, Tübingen, Germany. [4]Department of Physics, University of Tübingen, Tübingen, Germany. [5]Institut für Naturwissenschaften, Geographie und Technik, Pädagogische Hochschule Heidelberg, Heidelberg, Germany. [6]Institute of Environmental Assessment and Water Research (IDAEA-CSIC), Barcelona, Spain. [7]Institute for Coastal Systems—Analysis and Modelling, Helmholtz-Zentrum Hereon, Geesthacht, Germany. [8]Barcelona Supercomputing Center (BSC), Barcelona, Spain. [9]Section Meteorology, Institute of Geosciences, Rheinische Friedrich-Wilhelms-Universität Bonn, Bonn, Germany. [10]Institute of Earth Surface Dynamics, Geopolis, University of Lausanne, Lausanne, Switzerland. [11]ARC Centre of Excellence in Australian Biodiversity and Heritage (CABAH), Canberra, Australian Capital Territory, Australia. [12]School of Culture, History and Language, The Australian National University, Canberra, Australian Capital Territory, Australia. [13]Present address: Department Greenhouse gas emission verification, Deutscher Wetterdienst, Offenbach, Germany. ✉e-mail: tlaepple@awi.de

Methods

Spectral estimates

The PSD estimates were calculated using the multitaper method[98] with three tapers and a time-bandwidth parameter $\omega = 2$. The PSD estimates were smoothed using a Gaussian kernel with a constant width of 0.03 on the (base 10) logarithmic timescale.

Literature review and agreement

We included literature that covered the entire Holocene and applied a ranking system to indicate the level of (dis)agreement between climate reconstructions and models using the following five levels: disagreement, largely disagree, neutral, largely agree, and agreement. Levels were assigned on the basis of the respective study authors' original assessments of their model–data consistency that were based on various methods used for the model–data comparison (for example direct comparison, proxy modelling) and proxy calibration (for example temporal calibration, spatial calibration). We differentiated the results according to temporal and spatial scales wherever possible. Supplementary Table 1 contains the full review on which Fig. 2 is based, together with statements from the original papers that were the basis for assigning the levels of (dis)agreement. Extended Data Fig. 2 explicitly cites the individual studies.

Local and global spectra and relationship to the spatial scale of temperature fluctuations

Global mean spectra (Figs. 1 and 3) were computed as the PSD of the global temperature time series. As an example of a local record, we chose the Cariaco Mg/Ca records as it is one of the highest-resolution marine records that spans most of the past millennium (1200–2000 CE). This record is not affected by bioturbation (due to its laminated sediment) and is based on a classical temperature proxy (Mg/Ca on planktic foraminifera). Mg/Ca is clearly attributed to temperature and we can use an independent calibration (Methods; see ref. 14), circumventing potential issues with temporal calibrations. Local mean spectra (Figs. 1 and 3 and Extended Data Fig. 1) were computed as area-weighted mean spectra of the local (grid box) temperature. For the model spectra, the grid boxes containing the proxy record were selected. The local mean spectrum from the PAGES 2k database (Fig. 3 and Extended Data Fig. 1) corresponds to the estimate in ref. 19. For this estimate, 101 proxy records were considered and selected according to their resolution, number of data points and coverage, as well as their maximum hiatus. The result is robust to the selection criteria[19]. In Fig. 3, the local spectrum from sedimentary records is also shown, covering timescales from 1/200–1/3000 years. The proxy spectra were created from the weighted average of local SST spectra from ref. 14 (arithmetic mean of Mg/Ca and Uk37 proxy-based spectra, where Uk37 is the unsaturation index of C_{37} methyl alkenones[50]) and local terrestrial air temperature spectra from ref. 4 (2/3 × SST + 1/3 × terrestrial temperature). All three spectra show a similar scaling of PSD $\approx f^{-1}$ for frequency f and are therefore parallel to the global mean temperature spectra from CMIP5/6 and the PAGES 2k reconstruction (Fig. 1). This implies that the spatial scale remains nearly constant on supradecadal timescales, independent of the absolute magnitude of local variability. The ratio of local to global variability can be interpreted as effective spatial degrees of freedom that can in turn be translated into a characteristic length scale (an effective correlation radius)[56] (see Fig. 3a and the circles and labelling in Fig. 3b–d). To visualize a typical amplitude and spatial scale of temperature anomalies at different timescales (Fig. 3b–d), we computed a Haar fluctuation at different timescales (that is, the difference between the average of the first and second half of a given time interval) from the MPI-ESM-P past1000 experiment. We show examples for one 2 yr (Fig. 3b), one 20 yr (Fig. 3c) and one 500 yr fluctuation (Fig. 3d); they correspond to, respectively, the average temperature for the year 1050 minus 1051, the years 1050–1059 minus 1060–1069 and the years 1050–1299 minus 1300–1549.

Conceptual relation between slow climate variability and extremes

The time series of short-term variations in Fig. 4 was created as white noise, while the time series of long-term variations was generated as a low-pass filtered (cutoff 1/10 yr) stochastic process with PSD $\approx f^{-1}$. The variance ratio was 5:1. The parameters and time steps were chosen such that the short-term time series corresponds to fluctuations with a characteristic timescale of weeks and the long-term series to decadal fluctuations. In total, we generated a 10,000 yr time series to ensure that the distributions converged. The distribution and return levels were obtained empirically. We used the definition of the return level from the Weibull formula, which relates the inverse rank, i, of the sorted time series vector, $\mathbf{X}$, of length N to the return period $R(x_i)$ via (Box 1):

$$\mathbf{X} = (x_1, \ldots x_N),\ x_i \geq x_{i+1}$$

$$\mathbf{R}(x_i) = 1/(P(x \geq x_i)) = \frac{N+1}{i}$$

Data availability

The PAGES 2k palaeotemperature records (PAGES 2k v.2.0.0) are available at www.ncdc.noaa.gov/paleo/study/21171. The ensemble of global temperature reconstructions based on the PAGES2k16 data are available through the World Data Service (NOAA) Palaeoclimatology at https://www.ncdc.noaa.gov/paleo/study/26872 and via Figshare at https://doi.org/10.6084/m9.figshare.c.4507043. The pollen-based reconstructions are available via PANGEA at https://doi.pangaea.de/10.1594/PANGAEA.930512. The marine proxy data are available via PANGEA at https://doi.org/10.1594/PANGAEA.899489. The CMIP5 millennium simulations are available through the Earth System Grid Federation portal at https://esgf-data.dkrz.de. Source data are provided with this paper.

References

98. Percival, D. B. & Walden, A. T. *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* (Cambridge Univ. Press, 1993).
99. Wu, T. et al. Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century. *J. Geophys. Res. Atmos.* **118**, 4326–4347 (2013).
100. Gent, P. R. et al. The Community Climate System Model Version 4. *J. Clim.* **24**, 4973–4991 (2011).
101. Li, L. et al. The flexible global ocean-atmosphere-land system model, grid-point version 2: FGOALS-g2. *Adv. Atmos. Sci.* **30**, 543–560 (2013).
102. Schmidt, G. A. et al. Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *J. Adv. Model. Earth Syst.* **6**, 141–184 (2014).
103. Dufresne, J.-L. et al. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim. Dynam.* **40**, 2123–2165 (2013).
104. Roeckner, E. et al. Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *J. Clim.* **19**, 3771–3791 (2006).
105. Yukimoto, S. et al. A new global climate model of the Meteorological Research Institute: MRI-CGCM3 —model description and basic performance. *J. Meteorol. Soc. Jpn Ser. II* **90A**, 23–64 (2012).
106. Volodin, E. M. et al. Simulation of the modern climate using the INM-CM48 climate model. *Russ. J. Numer. Anal. Math. Model.* **33**, 367–374 (2018).
107. Hajima, T. et al. Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geosci. Model Dev.* **13**, 2197–2244 (2020).

108. Mauritsen, T. et al. Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and its response to increasing CO_2. *J. Adv. Model. Earth Syst.* **11**, 998–1038 (2019).
109. Yukimoto, S. et al. The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: description and basic evaluation of the physical component. *J. Meteorol. Soc. Jpn Ser. II* **97**, 931–965 (2019).

Acknowledgements

This study was undertaken by members of CVAS and 2k Network, working groups of the Past Global Changes (PAGES) Global Research association. This is a contribution to the SPACE ERC, STACY and PALMOD projects. The SPACE ERC project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 716092). STACY has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, project no. 395588486). This work has also been supported by the German Federal Ministry of Education and Research (BMBF), through the PalMod project (subprojects 01LP1926B (O.B.), 01LP1926D (M.C.) and 01LP1926C (B.E., P.S. and N.W.)) from the Research for Sustainability initiative (FONA). B.E. is supported by the Heinrich Böll Foundation. E.M.-C. was supported by the PARAMOUR project, funded by the Fonds de la Recherche Scientifique–FNRS and the FWO under the Excellence of Science (EOS) programme (grant no. O0100718F, EOS ID no. 30454083). A.H. was supported by a Legacy Grant from the Australian Research Council Centre of Excellence for Australian Biodiversity and Heritage. B.M. was supported by LINKA20102 and the Spanish Ministry of Science and Innovation project CEX2018-000794-S. The work originated from discussions at the CVAS working group of PAGES at a workshop at the Internationales Wissenschaftsforum Heidelberg, which was funded by a Hengstberger Prize. We thank N. Beech, C. Brierley, F. Gonzalez-Rouco and M. MacPartland for comments on earlier drafts of the manuscript. This manuscript uses data provided by the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP and PMIP. We thank the research groups for producing and kindly making their model outputs, measurements and palaeoclimate reconstructions available to us. Editorial assistance, in the form of language editing and correction, was provided by XpertScientific Editing and Consulting Services. We acknowledge support by the Open Access Publication Funds of Alfred-Wegener-Institut Helmholtz Zentrum für Polar- und Meeresforschung.

Author contributions

T.L. led the synthesis and overall analysis; E.Z. coordinated the work. K.R. and T.L. initiated the project through the workshop. R.H. led the writing of the introductory paragraph. E.Z. led the writing on the remainder of the introduction regarding the conflicting evidence in the literature; E.Z., B.M. and P.S. led the literature review; A.H. and all authors contributed to it; B.E. and R.H. produced Fig. 1, P.S. produced Fig. 2. N.W. and T.L. led the writing of the section on reconstruction deficiencies; N.W., E.M.-C. & T.L. led the writing of the section on the consequences for the spatial structure; R.H & T.L. produced Fig. 3. E.Z and B.E. led the writing of the section on the implications for climate projections and attribution efforts; B.E. produced Fig. 4. R.H. led the writing of the concluding section. O.B. provided a critical check of the current literature O.B., N.W., B.E, R.H., E.Z., T.L., K.R., B.M., M.C., reviewed each section in detail. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

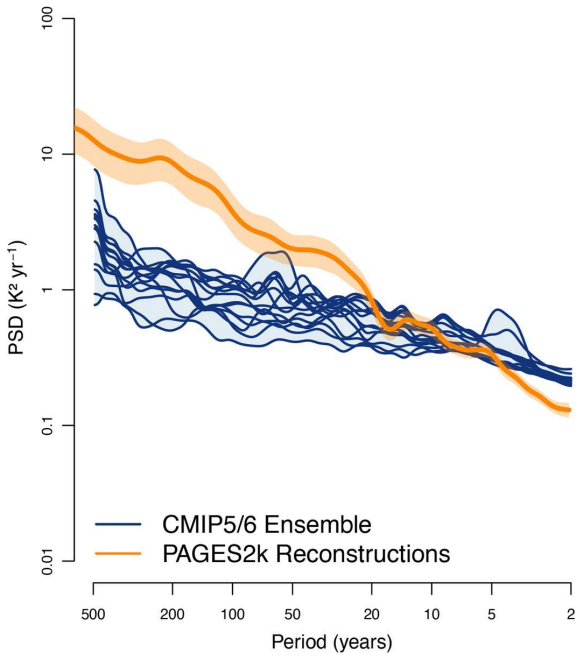
Extended data is available for this paper at https://doi.org/10.1038/s41561-023-01299-9.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41561-023-01299-9.

Correspondence should be addressed to T. Laepple.

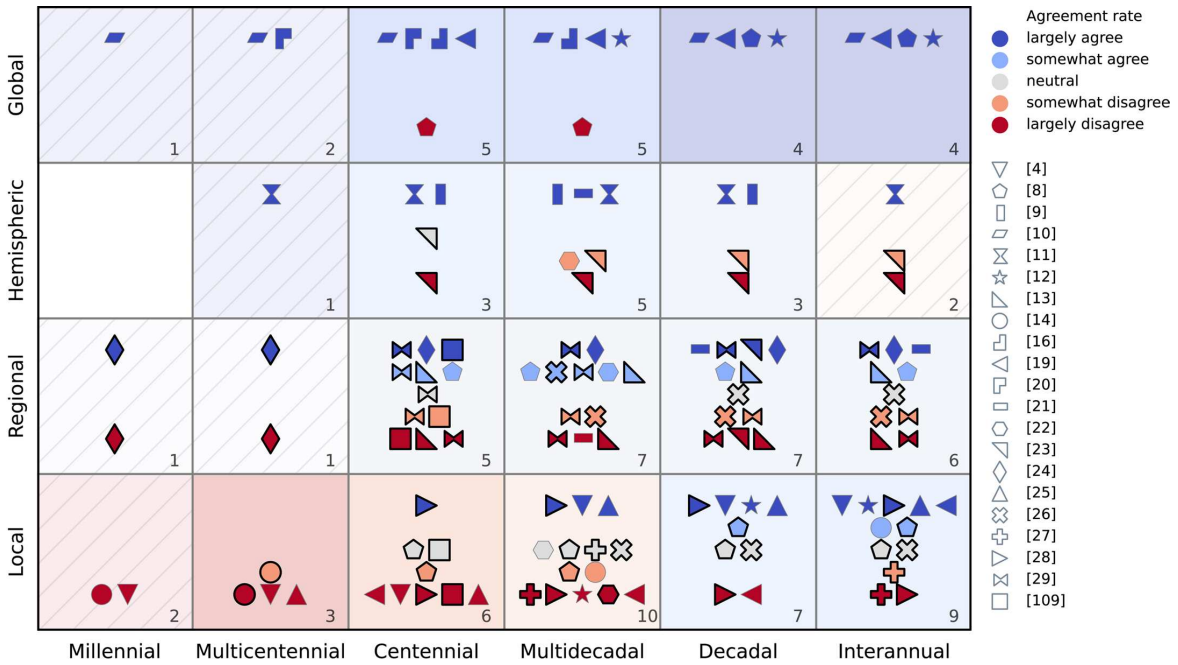
Peer review information *Nature Geoscience* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor(s): James Super, in collaboration with the *Nature Geoscience* team.

https://doi.org/10.1038/s41561-023-01299-9



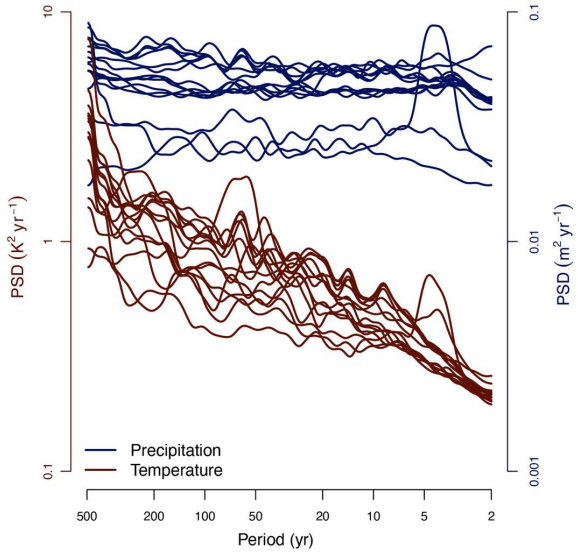
Extended Data Fig. 1 | Spectrum of mean local simulated and reconstructed temperature variability. As in Fig. 1d of the main text but for the mean local temperature spectrum from CMIP5/CMIP6 simulations and from PAGES2k temperature reconstructions (see Methods and Ref. 19). This shows that local variability reconstructed from instrumentally calibrated annually resolved records displays a similar model-data variability mismatch on supra-decadal time-scales as the example reconstruction (Cariaco) shown in Fig. 1d, or other marine or terrestrial records[4,14].

https://doi.org/10.1038/s41561-023-01299-9



Extended Data Fig. 2 | Overview of model-data (dis)agreement in Holocene temperature variability in the literature with explicit references. As in Fig. 2 of the main text, model-data agreement is grouped according to temporal (x-axis) and spatial scale (y-axis). Each symbol represents a specific study (refs. 4,8–14,16,19–2997) and the color-code indicates strength of (dis)agreement. Multiple occurrences in one box can happen when differing results are reported that is depending on reconstruction method or proxy type. Such cases are highlighted with a black border. The number at the bottom right of each box is the number of distinct studies in this box. Dashing of a box indicates only one or two studies for this spatio-temporal scale. Further details can be found in the Methods section.

https://doi.org/10.1038/s41561-023-01299-9



Extended Data Fig. 3 | Local precipitation and local temperature variability show a different temporal scaling. Local mean spectral estimates from CMIP5/6 precipitation (dark blue) and temperature (brown). Across all models, local precipitation variability shows a flatter (more white) scaling than local temperature variability. This implies that the mismatch between simulated and reconstructed local supra-decadal variability would increase, if the proxies would represent a mix of precipitation and temperature (calibrated to temperature units), as the difference in scaling between proxies and simulated precipitation is even larger than between proxies and simulated temperature.

https://doi.org/10.1038/s41561-023-01299-9

Extended Data Table 1 | CMIP5/CMIP6 model experiments. The model experiments were used in Figs. 1 and 3 (refs. 99–109)

Model	Country	Atm. grid spacing (lat x lon)	Experiment	Runs	Generation	Reference
BCC-CSM1.1	China	2.8° x 2.8°	past1000	r1i1p1	CMIP5	Ref[98]
CCSM4	USA	0.9° x 1.25°	past1000	r1i1p1	CMIP5	Ref[99]
FGOALS-gl	China	2.8° x 2.8°	past1000	r1i1p1	CMIP5	Ref[100]
GISS-E2-R	USA	2° x 2.5°	past1000	r1i1p121 r1i1p122 r1i1p123 r1i1p124 r1i1p125 r1i1p126 r1i1p127 r1i1p128 r1i1p1221	CMIP5	Ref[101]
IPSL-CM5A-LR	France	1.9° x 3.8°	past1000	r1i1p1	CMIP5	Ref[102]
MPI-ESM	Germany	1.875° x 1.875°	past1000	r1i1p1	CMIP5	Ref[103]
MRI-CGCM3	Japan	1.125° x 1.125°	past1000	r1i1p1	CMIP5	Ref[104]
INM-CM4-8	Russia	2° x 1.15°	past1000	r1i1p1f1	CMIP6	Ref[105]
MIROC-ES2L	Japan	2.8° x 2.8°	past1000	r1i1p1f2	CMIP6	Ref[106]
MPI-ESM1-2-LR	Germany	1.875° x 1.875°	past2k	r1i1p1f1	CMIP6	Ref[107]
MRI-ESM2-0	Japan	1.125° x 1.125°	past1000	r1i1p1f1	CMIP6	Ref[108]

A6 Publication 3: Patterns of changing surface climate variability from the Last Glacial Maximum to present in transient model simulations

E. Ziegler, N. Weitzel, J.-P. Baudouin, M.-L. Kapsch, U. Mikolajewicz, L. Gregoire, R. Ivanovic, P. J. Valdes, C. Wirths & K. Rehfeld (2025). Patterns of changing surface climate variability from the Last Glacial Maximum to present in transient model simulations. Climate of the Past, 21. doi:10.5194/cp-21-627-2025

Clim. Past, 21, 627–659, 2025
https://doi.org/10.5194/cp-21-627-2025

Climate of the Past
Open Access
EGU

Patterns of changing surface climate variability from the Last Glacial Maximum to present in transient model simulations

Elisa Ziegler[1,2], Nils Weitzel[1], Jean-Philippe Baudouin[1], Marie-Luise Kapsch[3], Uwe Mikolajewicz[3], Lauren Gregoire[4], Ruza Ivanovic[4], Paul J. Valdes[5], Christian Wirths[6], and Kira Rehfeld[1,2]

[1]Department of Geosciences, University of Tübingen, Tübingen, Germany
[2]Department of Physics, University of Tübingen, Tübingen, Germany
[3]Max Planck Institute for Meteorology, Hamburg, Germany
[4]School of Earth and Environment, University of Leeds, Leeds, UK
[5]School of Geographical Sciences, University of Bristol, Bristol, UK
[6]Physics Institute, University of Bern, Bern, Switzerland

Correspondence: Elisa Ziegler (elisa.ziegler@uni-tuebingen.de) and Kira Rehfeld (kira.rehfeld@uni-tuebingen.de)

Received: 10 May 2024 – Discussion started: 30 May 2024
Revised: 24 December 2024 – Accepted: 3 January 2025 – Published: 11 March 2025

Abstract. As of 2023, global mean temperature has risen by about 1.45 ± 0.12 °C with respect to the 1850–1900 pre-industrial (PI) baseline according to the World Meteorological Organization. This rise constitutes the first period of substantial global warming since the Last Deglaciation, when global temperatures rose over several millennia by about 4.0–7.0 °C according to proxy reconstructions. Similar levels of warming could be reached in the coming centuries considering current and possible future emissions. Such warming causes widespread changes in the climate system, of which the mean state provides only an incomplete picture. Instead, fluctuations around the mean and in higher-order statistics need to be considered. Indeed, climate's variability and the distributions of climate variables change with warming, impacting, for example, ecosystems and the frequency and intensity of extremes. However, previous investigations of climate variability focus mostly on measures such as variance, or standard deviation, and on quasi-equilibrium states such as the Holocene or Last Glacial Maximum (LGM). Changes in the tails of distributions of climate variables and transition periods such as the Last Deglaciation remain largely unexplored.

Therefore, we investigate changes of climate variability on annual to millennial timescales in 15 transient climate model simulations of the Last Deglaciation. This ensemble consists of models of varying complexity, from an energy balance model to Earth system models (ESMs), and includes sensitivity experiments, which differ only in terms of their underlying ice sheet reconstruction, meltwater protocol, or consideration of volcanic forcing. The ensemble simulates an increase in global mean temperature of 3.0–6.6 °C between the LGM and Holocene. Against this backdrop, we examine whether common patterns of variability emerge in the ensemble. To this end, we compare the variability in surface climate during the LGM, Deglaciation, and Holocene by estimating and analyzing the distributions and power spectra of surface temperature and precipitation. For analyzing the distribution shapes, we turn to the higher-order moments of variance, skewness, and kurtosis. These show that the distributions cannot be assumed to be normal, a precondition for commonly used statistical methods. During the LGM and Holocene, they further reveal significant differences, as most simulations feature larger temperature variance during the LGM than the Holocene, in line with results from reconstructions.

As a transition period, the Deglaciation stands out as a time of high variance in surface temperature and precipitation, especially on decadal and longer timescales. In general, this dependency on the mean state increases with model complexity, although there is a large spread between models of similar complexity. Some of that spread can be explained by differences in ice sheet, meltwater, and volcanic forcings, revealing the impact of simulation protocols on simulated variability. The forcings affect variability not only on their char-

acteristic timescales. Rather, we find that they impact variability on all timescales from annual to millennial. The different forcing protocols further have a stronger imprint on the distributions of temperature than precipitation. A reanalysis of the LGM exhibits similar global mean variability to most of the ensemble, but spatial patterns vary. However, paleoclimate data assimilation combines model and proxy data information using a Kalman-filter-based algorithm. More research is needed to disentangle their relative impact on reconstructed levels of variability. As such, uncertainty around the models' abilities to capture climate variability likewise remains, affecting simulations of all time periods: past, present, and future. Decreasing this uncertainty warrants a systematic model–data comparison of simulated variability during periods of warming.

1 Introduction

Understanding the response of the climate system during extended periods of global warming is of vital importance given current and projected anthropogenic warming. However, the observational record provides an insufficient data basis due to its short length of only about 150 years and its sparse spatial coverage during the earlier years (e.g., Morice et al., 2012). To extend the record further back in time requires the use of natural climate archives and proxy-based reconstructions. Such reconstructions have many associated uncertainties and limited resolution in time and space. Combining proxy records from different locations into a global field reconstruction introduces additional uncertainties, such as different interpolation and calibration procedures, age models, and proxy biases (Christiansen and Ljungqvist, 2017; Tingley et al., 2012). Climate models, on the other hand, simulate three-dimensional fields of a wide variety of variables that describe the climate. As such, they provide continuous estimates of climate that are limited by model physics and parametrizations but allow detailed investigations of the climate system and its changes on long timescales. Since their simulation length and resolution depend mostly on computational resources, simulation protocols up until the Paleoclimate Modeling Intercomparison Project phase 3 (PMIP3) encompassed mostly equilibrium simulations of past climate states in the form of time slices. Experiments with time-dependent (transient) forcings were limited to short periods like the past millennium or done with accelerated boundary conditions. The latest iteration, PMIP4, added more and longer experiments with transient boundary conditions. This allows more in-depth explorations of past transitions in the climate's mean state, such as the Last Deglaciation, which we examine here using an ensemble of transient simulations.

As the transition from the Last Glacial Maximum (LGM; 23–19 kyr before present[1]) to the current warm period of the Holocene (10.65 kyr BP–present day), the Last Deglaciation was a period of substantial global climate change. Global mean surface temperature (GMST) increased by about 4–7 °C according to proxy-based reconstructions and climate model simulations (Fig. 1a; Gulev et al., 2021; Osman et al., 2021; Annan et al., 2022; Tierney et al., 2020; Shakun and Carlson, 2010). The spread among recent estimates is similar, with some leaning towards the higher end with a warming of 7.0 ± 1.0 °C suggested by Osman et al. (2021) and 6.1 °C (5.7, 6.5) by Tierney et al. (2020) and others towards the lower end such as the estimate of 4.5 ± 0.9 °C by Annan et al. (2022). In simulations, a rise in global mean precipitation (GMP) accompanies this warming (Fig. 1b). During the same period, global mean sea level rose by approximately 120 m as the ice sheets in both hemispheres, but especially the Fennoscandian and Laurentide ice sheets, shrunk (Fig. 1c, Lambeck et al., 2014; Grant et al., 2012). However, there are significant uncertainties associated with ice sheet reconstructions, especially with respect to ice sheet extent and elevation (Stokes et al., 2015; Abe-Ouchi et al., 2015; Ivanovic et al., 2016). In turn, the timing and magnitude of meltwater events, which crucially impact the deglacial climate evolution (Snoll et al., 2024), remain uncertain.

Increasing levels of atmospheric carbon dioxide (CO_2) contributed to and drove this change (Shakun et al., 2012) as they rose from about 193.2 ppm at the onset of the Last Deglaciation to approximately 271.2 ppm at its end (Gulev et al., 2021). During the Holocene, CO_2 levels roughly stabilized until the Industrial Revolution (Fig. 1d). Similarly, atmospheric methane almost doubled from the LGM to the Holocene (Köhler et al., 2017, Fig. 1e). Changes in latitudinal and seasonal insolation distribution favored this rise in atmospheric greenhouse gases (GHGs) and warming (Fig. 1f).

However, considering only the described mean changes is insufficient to capture the full breadth of climate change then and now. Instead, it is necessary to study the climate's variability, too, as reflected in the fluctuations around the mean[2] and in higher-order statistics in space and time (Katz and Brown, 1992). These fluctuations determine the actual climate conditions at any point in time and space and are the focus of our study. They affect the various modes of variability (Rehfeld et al., 2020) and the occurrence and frequency of extremes (Simolo and Corti, 2022; Ionita et al., 2021; Schär et al., 2004; Loikith et al., 2018; Ruff and Neelin, 2012; Laepple et al., 2023). Climate variability further acts across timescales, from intra-annual (i.e., heat waves) and interan-

[1] Here, before present (BP) refers to the year 1950, Common Era (CE). AP denotes the opposite, after present.

[2] We include only changes outside the mean in our use of variability. This is in contrast to the IPCC (2021), which includes any deviation from a given equilibrium state, including the change in the mean with time.

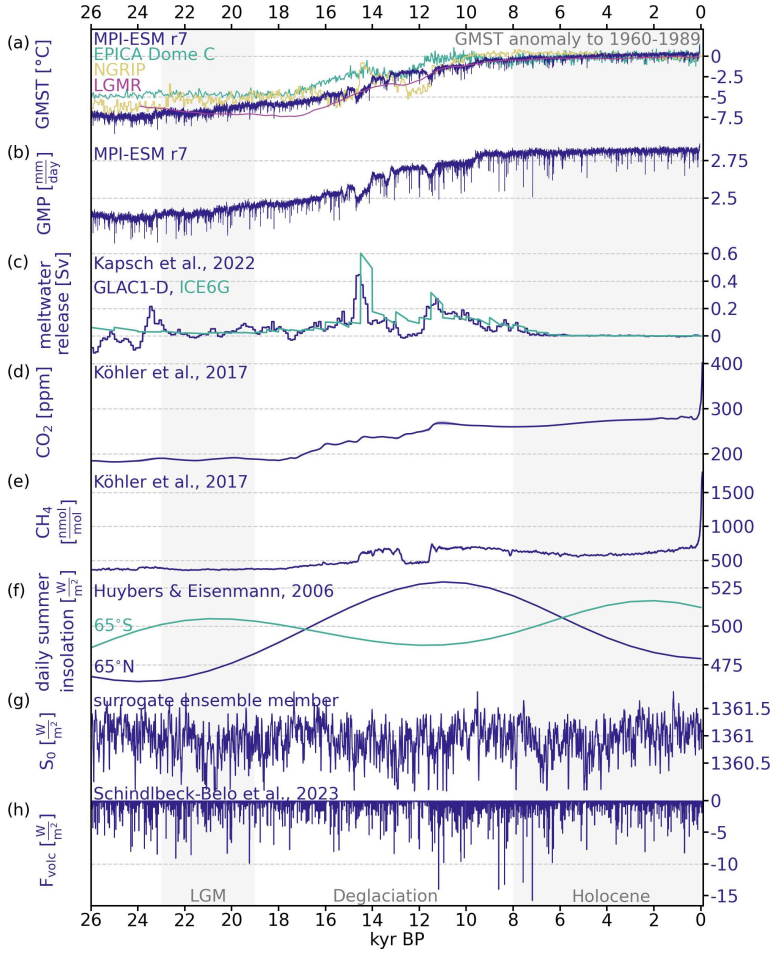


Figure 1. Climate responses and external forcing during the past 26 kyr: **(a)** global mean temperature anomaly (with regard to 1960–1989) as simulated by MPI-ESM, captured in ice cores from Antarctica (EPICA Dome C; Jouzel et al., 2007) and Greenland (NGRIP; Andersen et al., 2004) and reconstructed in the LGM reanalysis (Osman et al., 2021). The proxy records for local temperature derived from EPICA and NGRIP are scaled to GMST using GMST$= 0.5 \times T_{local}$. **(b)** Global mean precipitation as simulated by the Earth system model MPI-ESM, **(c)** meltwater release for ice sheet reconstructions GLAC1-D and ICE6G as used in MPI-ESM r1–r7 (Kapsch et al., 2022), **(d)** atmospheric CO_2 (Köhler et al., 2017) and **(e)** CH_4 levels (Köhler et al., 2017), **(f)** daily insolation at 65°N and 65°S at the summer solstice (Huybers and Eisenman, 2006), **(g)** solar constant from one ensemble member generated as surrogate data based on Steinhilber et al. (2009) following Ellerhoff and Rehfeld (2021) (comparison with Steinhilber et al., 2009 in Fig. S2 in the Supplement), and **(h)** volcanic forcing TephraSynthIce (Schindlbeck-Belo et al., 2024; Sigl et al., 2022).

nual (multi-year droughts, the El Niño–Southern Oscillation (ENSO)) to millennial scales (D–O events) and beyond and across different spatial scales (Franzke et al., 2020; Laepple et al., 2023).

Proxy-based reconstructions suggest that global mean temperature variance was about 4 times higher during the LGM than the Holocene, possibly due to changes in the Equator-to-pole temperature gradient (Rehfeld et al., 2018). This implies a dependence of variability on mean climate. The extent of this state dependency varies regionally; e.g., Rehfeld et al. (2018) find that it is generally larger in the Northern Hemisphere mid- and high latitudes than in the Southern Hemisphere. Models are only partially able to match this LGM-to-Holocene change in temperature variability. Rehfeld et al. (2018) found that interannual to decadal variability is about 30 % higher during the LGM than during

the pre-industrial (PI) in PMIP3 and Coupled Model Intercomparison Project (CMIP) phase 5 simulations. Shi et al. (2022) confirmed this for PMIP3/4 LGM simulations, which have about 20 % larger interannual variability than PI simulations. Models do agree with reconstructions on decreasing global temperature variability (Rehfeld et al., 2020) and mean local variability (Ellerhoff et al., 2022) with warming, especially towards higher latitudes (Ellerhoff et al., 2022), but with some exceptions in the tropics (Rehfeld et al., 2020). Few studies quantitatively considered variability changes over the Last Deglaciation. One, Weitzel et al. (2024), compared millennial and orbital variability in many of the transient simulations considered here with proxy reconstructions. Differences varied in time and space, and no single simulation was identified to best match reconstructions.

Globally, there is mostly agreement between the variance at interannual to centennial timescales in models and reconstructions both during the Holocene (Laepple et al., 2023) and further back in time, including during the Deglaciation (Zhu et al., 2019). On regional and local scales, however, models simulate less variability than reconstructions, especially on multi-decadal timescales and longer (Laepple et al., 2023; Ellerhoff et al., 2022; Rehfeld et al., 2018). Including natural forcing (that is, forcing from solar and volcanic activity) in simulations reduces this difference but cannot close it (Ellerhoff et al., 2022). Opposite to temperature, precipitation variability increases with warming, with some regional exceptions (Rehfeld et al., 2020). This precipitation variability can be linked to mean precipitation changes, as dry regions generally have lower variability and wet regions generally have higher variability (Rehfeld et al., 2020).

The influence of natural forcing demonstrates that significant variability arises in response to external forcings and boundary conditions. Volcanism, in particular, has been identified as a prominent driver of changes in temperature, precipitation, and modes of atmospheric dynamics (Timmreck, 2012; Iles and Hegerl, 2015; Liu et al., 2016; Zanchettin et al., 2015). Its strongest effects manifest on annual timescales (Lovejoy and Varotsos, 2016), as has been found for the past millennium and Common Era (Schurer et al., 2014; Lovejoy and Varotsos, 2016). It further contributed substantially to subdecadal (Le et al., 2016), decadal (Hegerl et al., 2003), and multi-decadal (Schurer et al., 2013) variance. During glacials, strong volcanic eruptions are even suggested as a driver of millennial variability (Baldini et al., 2015). There is conflicting evidence with respect to the dependence of the impacts of volcanic forcing on the background state: in equilibrium simulations of the LGM and the PI period, Ellerhoff et al. (2022) found no state dependency on the global scale for surface temperature variability and only slight differences for precipitation. Bethke et al. (2017), on the other hand, found enhanced variability in future projections on annual to decadal timescales.

Throughout glacial cycles, the cryosphere plays a crucial role for the climate and its variability. This includes sea ice dynamics and changes in ice sheets and associated meltwater releases. Ice sheets and meltwater releases are still commonly simulated as external forcings (Ivanovic et al., 2016). However, reconstructions of ice sheet extent and elevation and associated meltwater pulses entail significant uncertainties (Stokes et al., 2015; Abe-Ouchi et al., 2015; Ivanovic et al., 2016; Izumi et al., 2023). For both the LGM and the Last Deglaciation, simulated climate has been shown to be very sensitive to ice sheet reconstructions (Izumi et al., 2023; Kapsch et al., 2022; Bakker et al., 2020; Ullman et al., 2014). Furthermore, meltwater release as a consequence of melting ice sheets affects ocean circulation and thus deglacial climate as a whole (Kapsch et al., 2022). Consequently, the uncertainties in meltwater scenarios and models' varying sensitivities to freshwater crucially affect the simulation of deglacial climate (Snoll et al., 2024).

For sea ice, a decreasing extent has been shown to reduce the seasonal to interannual standard deviation of temperature, likely due to polar amplification and the sea ice–albedo feedback (Screen, 2014; Huntingford et al., 2013; Screen and Simmonds, 2010; Bathiany et al., 2018). As a response to shrinking sea ice, this feedback reduces the meridional temperature gradient, which has been linked to decreased variability. Collow et al. (2019) demonstrate a decrease in extreme temperatures, both in frequency and magnitude, with decreasing sea ice extent. Loss of sea ice further leads to an increase in scaling (Rehfeld et al., 2020). In addition, Ellerhoff et al. (2022) found that sea ice dynamics are a significant component of local variability on decadal and longer timescales.

Analyses of variability largely focus on variance, especially in paleoclimate studies, as mean and variance suffice to describe a normal (Gaussian) distribution in full, making variance a useful metric in many contexts. For annual to decadal temperature data, assuming normally distributed data is often a good approximation after removing periodic variations like the diurnal or seasonal cycle. However, on shorter timescales, this assumption can break down locally and regionally, where many climate variables are non-normal (Tamarin-Brodsky et al., 2022; Garfinkel and Harnik, 2017; Perron and Sura, 2013; Simolo and Corti, 2022). Such cases necessitate more detailed analyses of the shape of distributions, which higher-order moments allow.

The higher-order moments of skewness and kurtosis facilitate an examination of the asymmetry and heaviness of a distribution's tails, respectively. They have been shown to be pronounced for many atmospheric variables, such as geopotential height, vorticity, wind fields, and specific humidity (Perron and Sura, 2013) on top of temperature (Tamarin-Brodsky et al., 2022; Ruff and Neelin, 2012; Skelton et al., 2020; Volodin and Yurova, 2013) and precipitation (He et al., 2013). All else being equal, an increase in variance already increases the probability of extremes, whereas a decrease would counteract it. However, this can be complicated by additional changes in skewness and kurtosis (McKinnon et al.,

2016), which reveal enhancements or reductions in extremes (Ruff and Neelin, 2012; Simolo and Corti, 2022).

The shape of the tails determines how extremes change with warming, such that, for example, under warming, short tails lead to higher exceedances with respect to fixed hot extreme thresholds than Gaussian tails would (Ruff and Neelin, 2012; Loikith et al., 2018). Additionally, changes in skewness can indicate approaching abrupt shifts. As a system moves towards a tipping point, the weight of the distribution moves towards this point with an increasingly long tail away from it; that is, skewness increases when approaching a point of abrupt change (Guttal and Jayaprakash, 2008). These kinds of early warning signals have been found in weather station and climate model simulation data (Skelton et al., 2020; He et al., 2013), as well as in ecosystems (Guttal and Jayaprakash, 2008).

Overall, changing dynamics in the Earth system will affect the distributions of a climate variable, potentially resulting in changes in skewness or kurtosis. However, the mechanisms linking the climate system to these statistical measures remain unclear (Simolo and Corti, 2022; Perron and Sura, 2013). For surface or near-surface temperature, asymmetry and long tails are found due to horizontal advection along storm tracks (Garfinkel and Harnik, 2017; Ruff and Neelin, 2012). Land–atmosphere interactions, particularly changes in soil moisture, are related to significant changes in skewness for near-surface temperature as well (Berg et al., 2014; Douville et al., 2016). Skewness further reflects a marine versus continental influence (McKinnon et al., 2016). Studies of skewness and kurtosis in the literature use data from the 20th century or future projections, often consider only limited time frames, and mostly focus on daily and seasonal data. To the best of our knowledge, for paleoclimate, no other study has investigated moments higher than standard deviation. As a consequence, the role of higher-order moments on longer timescales, when normality assumptions might break down under a non-stationary climate evolution, and in past climates is unknown. Whether they changed between past climate periods, can indicate past abrupt transitions, or could provide a useful metric for inter-model and model–data comparisons remains similarly unclear.

Here, we evaluate how the variability in surface climate changes from the LGM to the present. The analysis uses an ensemble of transient climate model simulations (Sect. 2.1) that we characterize based on model complexity (Sect. 2.2). As indicators of variability, we focus on changes to the distributions of surface temperature and precipitation, including the higher-order moments (Sect. 3.1) and the power spectrum (Sect. 3.2). We hypothesize (1) that patterns of surface climate variability are state-dependent for the quasi-equilibrium conditions of the LGM and the Holocene, which differ from those during a transitionary state like the Deglaciation; (2) that state- and forcing-induced changes in variability depend on timescale; and (3) that there is a necessary and sufficient level of model complexity for the simulation of variability. To verify these hypotheses, we investigate the dependence of the variability of surface temperature and precipitation on

- background state (Sect. 5.2);
- timescale (Sect. 5.2);
- forcings, particularly ice sheet reconstruction, meltwater forcing protocol, and volcanism (Sect. 5.3); and
- model complexity (Sect. 5.4).

By comparing simulated variability with reconstructions and a reanalysis product, we explore the impact of forcing protocols on model–data agreement (Sect. 4.6.4). Overall, we examine the last global transition in climate to highlight differences between a period of warming in comparison to its preceding and succeeding stable climates.

2 Models and data

We draw on an ensemble of 15 simulations of the Last Deglaciation from climate models of varying complexity (Sect. 2.1). The models range from an energy balance model (EBM) and Earth system models of intermediate complexity (EMICs) to general circulation models (GCMs) and Earth system models (ESMs), which we evaluate regarding their complexity (Sect. 2.2). Furthermore, we compare the simulations to a multi-proxy reconstruction (Sect. 2.3).

2.1 Simulation data

All 15 simulations are transient and cover at least the Last Deglaciation. We separate the simulations into a main set and a sensitivity set. Table 1 provides an overview of the simulations and forcing protocols. The following describes the ensemble in more detail:

- *MPI-ESM ch4* (Kleinen et al., 2023, 2020)
  Model: This main set simulation used a setup of MPI-ESM v.1.2 at a coarse resolution called MPI-ESM-CR (Mauritsen et al., 2019; Mikolajewicz et al., 2018) with a methane cycle (Kleinen et al., 2020). Boundary conditions, including ice sheets, bathymetry, topography (Meccia and Mikolajewicz, 2018) from GLAC1-D (Briggs et al., 2014; Tarasov et al., 2012), and river routing (Riddick et al., 2018) were updated every 10 years. It covers 23 kyr BP until the present day.
  Simulated climate: This run simulates an LGM (23–19 kyr BP) to Holocene (8–0 kyr BP)[3] warming of

[3] Whenever the mean for LGM and Holocene is computed, we follow these definitions, as marked in Fig. 2; that is, for the LGM we consider 23–19 kyr, and for the Holocene we consider 8–0 kyr BP. The mean anomalies in Fig. 2 are plotted with respect to the past 2 kyr.

4.4 °C and wetting of 0.27 mm d^{-1}. [4] At its start, it still cools in the global mean in tandem with an increase in the Equator-to-pole difference in the Southern Hemisphere and in sea ice volume (Fig. 2). It reaches its minimal GMST at around 21 kyr BP. The trend in increasing GMST during the Deglaciation is interrupted by abrupt decreases in GMST thrice: at around 14.5, 13.5, and 11.5 kyr BP. In comparison to other simulations, MPI-ESM ch4 simulates the smallest sea ice cover (Fig. 2c). Its global sea ice fraction is largest between 23 and 17 kyr BP and undergoes cycles of abrupt increase and decrease during the Deglaciation.

– *MPI-ESM r1–r7* (Kapsch et al., 2022, 2021)
Model: These seven simulations were produced using two more setups of MPI-ESM-CR. They also update boundary conditions every 10 years and cover the period from 26 kyr BP to the present day. They use different ice sheet reconstructions – GLAC1-D or ICE-6G_C (in the following ICE6G; Peltier et al., 2015) – and vary by meltwater scenario. Furthermore, a parameter for cloud formation was changed in r5–r7 to remove a cold bias found in r1–r4 (as detailed in the supporting information of Kapsch et al., 2022). The ice sheet reconstructions differ in their original resolution in time, with ICE6G providing updated boundary conditions at 500-year intervals and GLAC1-D at 100-year intervals, and were interpolated here to 10-year resolution. The meltwater scenarios follow the options outlined in the deglacial protocol of PMIP4 (Ivanovic et al., 2016): melt-uniform, melt-routed, and no-melt. These correspond to meltwater being distributed globally, through river-routing or being removed. Simulation r7 also applies a volcanic forcing by Schindlbeck-Belo et al. (2024) that builds on the Holocene reconstruction by Sigl et al. (2022), drawing on tephra records and including synthetic volcanic eruptions to mitigate underestimation of small eruptions. This simulation is part of the main set. Runs 1–6 form part of the sensitivity set. For two simulations, r3 and r4, only centennial means were available; thus they are only considered for the analysis of centennial variability.
Simulated climate: These simulations exhibit the largest warming between the LGM and Holocene of the ensemble with a range from 5.3 °C (for r5) up to 6.6 °C (for r7; Figs. 2a and S1a). The accompanying global mean wetting is also the largest in the ensemble, ranging between 0.30 mm d^{-1} (r5) and 0.39 mm d^{-1} (r1 and r7). Runs 1, 6, and 7 further simulate abrupt cooling periods during the Deglaciation with the same timing as in MPI-ESM ch4. These are the simulations that employ the GLAC1-D ice sheet reconstruction and corresponding meltwater forcing. The remaining runs show either continuous or sometimes abrupt warming during those periods. The sea ice cover in these simulations is generally larger than in MPI-ESM ch4 and shows a stronger decrease towards the Holocene.

– *TraCE-21ka* (He, 2011)
Model: The TraCE-21ka simulation was performed with CCSM3 (Collins et al., 2006) and stretches from 22 kyr BP to 1990 CE. This main set simulation was designed to match proxy data of millennial events such as the Bølling–Allerød and Younger Dryas during the Deglaciation (He, 2011). As such, it applies meltwater forcings in the Northern and Southern hemispheres at various times throughout the Deglaciation to reproduce proxy records (denoted as melt-routed matched). Ice sheets are updated at intervals of 500 years based on a modified version of the ICE5G reconstruction (ICE5G*; He, 2011; Peltier, 2004). As greenhouse gas forcing, TraCE-21ka uses Joos and Spahni (2008) (referred to as J&S in Table 1) with the age model of Monnin et al. (2001).
Simulated climate: Among all simulations, TraCE-21ka tends towards the lower end of GMST and GMP change from the LGM to the Holocene at 4.1 °C and 0.20 mm d^{-1}. It shows abrupt warming around the time of the Bølling–Allerød interstadial (about 14.7–12.9 kyr BP) with subsequent cooling matching the Younger Dryas (circa 12.9–11.7 kyr BP; Fig. 2a). For most of the time period covered, TraCE-21ka produces the largest sea ice cover, with the exception of the EBM (Fig. 2c). This difference becomes particularly large towards the end of the Deglaciation and remains so throughout the Holocene.

– *HadCM3B r1 & r2* (Snoll et al., 2024)
Model: The ensemble contains two simulations from HadCM3B (Valdes et al., 2017) that cover 23 kyr BP to 2 kyr AP. These employ two different meltwater protocols, melt-uniform (r1) and melt-routed (r2), from the PMIP4 protocol to match the ICE6G ice sheet history. The simulations prescribe orbit and greenhouse gases (GHGs) annually, while ICE6G ice sheet, orography, land–sea mask, and bathymetry are updated every 500 years. HadCM3B r1 is part of the sensitivity set, while r2 is included in the main set.
Simulated climate: The GMST difference between the LGM and the Holocene is 4.5 °C for the melt-uniform r1 and 4.8 °C for the melt-routed r2. Similarly, wetting of r1 is weaker at 0.26 mm d^{-1} in comparison to 0.27 mm d^{-1} for r2. The changes in Equator-to-pole gradient are notably small, especially in the Northern Hemisphere (Fig. 2d, e). Sea ice cover shrinks until 14 kyr BP and remains roughly constant thereafter (Fig. 2c).

[4] Throughout the analysis, we use the units Celsius for surface temperature and mm d^{-1} for precipitation.

– *FAMOUS* (Smith and Gregory, 2012)
Model: FAMOUS is a low-resolution, slightly simplified version of HadCM3 (Smith et al., 2008). It is sometimes classified as an EMIC (as in Valdes et al., 2017) or as a low-resolution GCM based on the complexity of its atmosphere model. The simulation used here as part of the main set was run with an acceleration factor of 10 for all forcings, allowing it to cover the last 120 kyr (Smith and Gregory, 2012). The simulation does not consider sea level change; that is, ice sheets are present only where there are no modern ocean grid points. Furthermore, the ICE5G reconstruction and topographic changes (Peltier, 2004) were only applied north of 40° N (ICE5G**). In particular, the Antarctic ice sheet remains unchanged (Smith and Gregory, 2012).
Simulated climate: FAMOUS simulates the smallest global mean change for both surface temperature and precipitation among all simulations at 3.1 °C and 0.15 mm d^{-1}, respectively. Its simulated Equator-to-pole temperature differences are among the largest in the ensemble, but they decrease comparatively little from the LGM to the present day (Fig. 2d, e).

– *LOVECLIM DG_ns* (Menviel et al., 2011)
Model: This version of the EMIC LOVECLIM couples the atmosphere model ECBilt (Opsteegh et al., 1998) to the ocean model CLIO (Campin and Goosse, 1999; Goosse et al., 1999; Goosse and Fichefet, 1999). LOVECLIM DG_ns used ECBilt-CLIO v.3 coupled to a dynamical vegetation model with a terrestrial carbon cycle (Menviel et al., 2011) and is included in the main set. It focuses on the Deglaciation, running from 18–6.2 kyr BP. Employed greenhouse forcing is based on reconstructions from EPICA (Lüthi et al., 2008; Monnin et al., 2001; Spahni et al., 2005) mapped onto the EDC3 age scale (Parrenin et al., 2007). Like TraCE-21ka, it includes meltwater pulses in the North Atlantic and Southern Ocean (melt-routed matched) to reproduce millennial-scale events in the North Atlantic (McManus et al., 2004) and Greenland (Alley, 2000) during this period (Menviel et al., 2011).
Simulated climate: As a result of the employed meltwater pulses, there are a warming and a subsequent cooling event visible in the global mean around the times of the Bølling–Allerød and the Younger Dryas, respectively (Fig. 2a, b). This signal is very strong in the Northern Hemisphere, where LOVECLIM DG_ns exhibits a large reduction in Equator-to-pole temperature gradient alongside these abrupt changes (Fig. 2d). In the Southern Hemisphere, this decrease is more subdued (Fig. 2e). Overall, the simulation shows deglacial warming and wetting comparable to most of the other simulations (Fig. 2a, b).

– *ECBilt-CLIO sim2bl* (Timm and Timmermann, 2007)
Model: The second ECBilt simulation included in the main set uses the same coupled ocean and atmosphere models and covers the period from 21–0 kyr BP (Timm and Timmermann, 2007). It contains no meltwater forcing. For the ice sheets and land–sea mask of the atmosphere model, it applies ICE4G with the East Siberian ice sheet removed (ICE4G*). For the ocean model, the same ice sheet is used but combined with a constant land–sea mask representing present-day conditions (ICE4G* & PD).
Simulated climate: Its simulated mean changes are 3.9 °C and 0.25 mm d^{-1}. Like LOVECLIM DG_ns, it has deglacial warming and wetting comparable to most of the other simulations (Fig. 2a, b). In magnitude, changes in ECBilt-CLIO sim2bl resemble those in LOVECLIM DG_ns (Fig. 2). Their structure is quite different, though, as ECBilt-CLIO sim2bl variables all change in a step-like manner. The simulated sea ice cover is at the upper end of the ensemble at the beginning of the simulation (similarly to TraCE-21ka and MPI-ESM r7) and then, like MPI-ESM r7, reduces drastically towards the Holocene (Fig. 2c).

– *TransEBM* (Sect. S2.1 in the Supplement)
Model: To represent the linear temperature response of the climate system to external forcing, we juxtapose a simulation from an extended version of the 2D energy balance model TransEBM (Ziegler and Rehfeld, 2021) with the other simulations and include it in the main set. Here, it has been extended to include freshwater and zonal volcanic forcing. The simulation covers the surface temperature evolution of the last 26 kyr, with ICE6G boundary conditions updated every 125 or 500 years. Sea ice extent was interpolated between the LGM and present-day states given by Zhuang et al. (2017). Meltwater forcing was assimilated based on the database of sea surface temperature records by Jonkers et al. (2020) (Jonkers assimilated; see Sect. S2.1). The simulation employs the same volcanic forcing as MPI-ESM r7. Sect. S2.1 describes the simulation in more detail.
Simulated climate: TransEBM simulates a GMST difference between the LGM and the Holocene of 4.1 °C, which is at the lower end of the ensemble and comparable to that of TraCE-21ka. Changes in Equator-to-pole difference are similar in magnitude in both hemispheres, unlike most other simulations (Fig. 2d, e). Its sea ice cover is the largest and changes the most during the Deglaciation because EBM models sea ice as a surface type, which covers any given grid cell completely.

To summarize, the main set is made up of MPI-ESM ch4 and r7, TraCE-21ka, HadCM3B r2, FAMOUS, LOVECLIM DG_ns, ECBilt-CLIO sim2bl, and TransEBM. MPI-ESM r1–r6 and HadCM3B r1 form the sensitivity set.

https://doi.org/10.5194/cp-21-627-2025

Table 1. Forcings applied for the transient simulations of the Last Deglaciation. Further description in Sect. 2.1. Labels of main set simulations are bold.

Simulation	Forcing							Reference
	Orbital	Solar constant	GHG	Volcanic	Land–sea mask	Ice sheets	Meltwater	
MPI-ESM ch4	Berger	B&L	Köhler	–	GLAC1-D	GLAC1-D	Melt-routed	Kleinen et al. (2023, 2020)
MPI-ESM r1	Berger	B&L	Köhler	–	GLAC1-D	GLAC1-D	Melt-routed	Kapsch et al. (2022, 2021)
MPI-ESM r2	Berger	B&L	Köhler	–	ICE6G	ICE6G	Melt-routed	Kapsch et al. (2022, 2021)
MPI-ESM r3	Berger	B&L	Köhler	–	ICE6G	ICE6G	No-melt	Kapsch et al. (2022, 2021)
MPI-ESM r4	Berger	B&L	Köhler	–	ICE6G	ICE6G	Melt-uniform	Kapsch et al. (2022, 2021)
MPI-ESM r5	Berger	B&L	Köhler	–	ICE6G	ICE6G	Melt-routed	Kapsch et al. (2022, 2021)
MPI-ESM r6	Berger	B&L	Köhler	–	GLAC1-D	GLAC1-D	Melt-routed	Kapsch et al. (2022, 2021)
MPI-ESM r7	Berger	B&L	Köhler	PalVol V.1	GLAC1-D	GLAC1-D	Melt-routed	Marie Kapsch and Uwe Mikolajewicz, personal communication, 2023
TraCE-21ka	Berger	Constant	CO_2, N_2O, CH_4: J&S	–	ICE5G*	ICE5G	Melt-routed matched	He (2011)
HadCM3B r1	Berger	Constant	PMIP4	–	ICE6G	ICE6G	Melt-uniform	Snoll et al. (2024)
HadCM3B r2	Berger	Constant	PMIP4	–	ICE6G	ICE6G	Melt-routed	Snoll et al. (2024)
FAMOUS	Berger	Constant	CO_2, N_2O, CH_4: EPICA	–	ICE5G**	ICE5G**	–	Smith and Gregory (2012)
LOVECLIM DG_ns	Berger	Berger	CO_2: EPICA, other const.	–	ICE4G	ICE4G	Melt-routed matched	Menviel et al. (2011)
ECBilt-CLIO sim2bl	Berger	Berger	CO_2, N_2O, CH_4: Taylor Dome	–	ICE4G* & PD	ICE4G*	–	Timm and Timmermann (2007)
TransEBM	Berger	Steinhilber surrogate	Köhler	PalVol V.1	ICE6G	ICE6G	Jonkers assimilated	Sect. S2.1

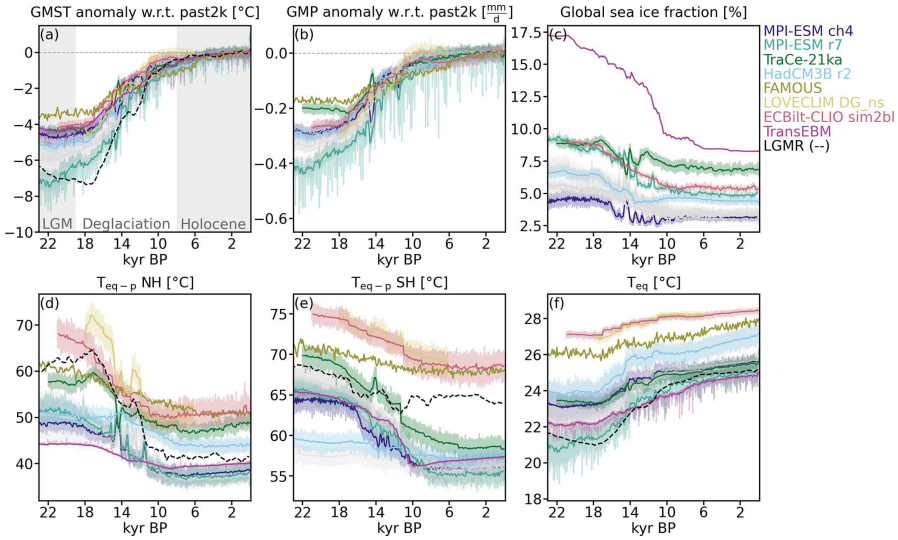


Figure 2. Centennial changes in the main set from the simulation ensemble from the LGM to the Holocene with annual data as shading. **(a)** GMST and **(b)** GMP anomaly with respect to the past 2 kyr. **(c)** Global sea ice fraction. Note that the EBM only allows complete coverage of grid cells by one surface type and therefore has the largest sea ice cover. FAMOUS and LOVECLIM DG_ns are missing, since their sea ice cover was not readily available. **(d, e)** Equator-to-pole temperature difference for the Northern Hemisphere and Southern Hemisphere, respectively, computed as the difference between polar (70–90°) and equatorial (15° S–15° N) temperatures. The latter are shown in panel **(f)**. The sensitivity set is shown in gray and can be found in Fig. S1. LGM (23–19 kyr BP), Deglaciation (19–8 kyr BP), and Holocene (8–0 kyr BP) as used in this study are marked in panel **(a)**.

2.2 The model hierarchy

We construct a hierarchy of the models to summarize the outlined differences and thus understand their effects on all aspects of climate variability. The complexity of the models and simulations differs along several axes: resolution in time and space, complexity of the individual components (e.g., atmosphere, ocean, land surface), their coupling, and their forcing. Constructing a hierarchy of models or simulations helps summarize those differences and thus understand their effect on any given analysis. The relevant axes of comparison might differ between applications. As a consequence, ranking the same models and simulations might produce a different hierarchy depending on the application. Here, we establish a hierarchy focused on features that affect variability and for which the simulations meaningfully differ.

Based on these considerations, we include eight axes of comparison (Fig. 3a). Section S2.2 explores these axes and the classification of the simulations. The resulting hierarchy reveals the various levels of complexity of the different simulations by placing them along each axis. In general, a simulation is considered more complex, the larger the total area it covers. Whenever an axis of the hierarchy does not apply to a simulation, the rank will be at the center of the net; see the lack of dedicated ocean or land hydrology model in TransEBM.

Based on all the factors summarized in Fig. 3a, we separate all simulations into four groups of complexity for parts of our analysis: ESMs (MPI-ESM), GCMs (TraCE-21ka, HadCM3B1, FAMOUS), EMICs (LOVECLIM, ECBilt-CLIO), and the EBM (TransEBM). The categorization follows the overall number of levels reached in the hierarchy. In the end, both applied forcings and complexity of the model components decide the simulation output. Our analysis tries to identify and disentangle the effects of both on simulated variability, with the goal of identifying the complexity both necessary and sufficient for long, transient climate simulations. Since increased complexity implies higher computational demand, a trade-off has to be made between complexity and available resources. Knowing the benefits and limitations of added complexity is thus crucial.

2.3 Global climate reanalysis data

For quantitative comparison, we draw on a spatiotemporally gridded product, the LGM reanalysis LGMR by Osman et al. (2021), which covers the past 24 000 years. LGMR combines model simulations and proxy reconstructions in an offline data assimilation approach for a proxy-constrained estimate of the full field of surface temperature since the LGM. The resulting dataset has a resolution in time of 200 years, allowing a comparison of centennial variability to the results of our

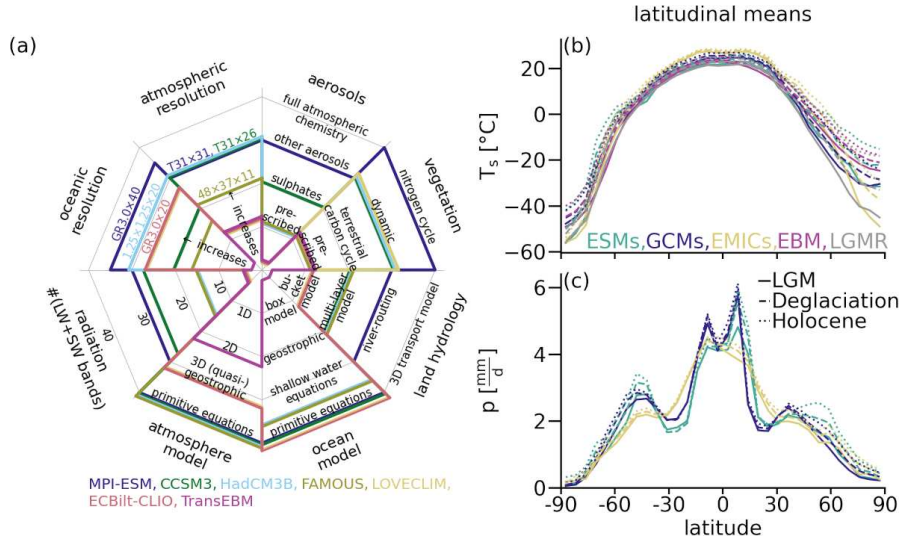


Figure 3. **(a)** Ranking of the models used in this study along eight axes of complexity. The criteria are described in detail in Sect. 2.2 and Table S1. Altogether they establish a hierarchy of the different models: ESMs (MPI-ESM), GCMs (CCSM3, HadCM3B, FAMOUS), EMICs (LOVECLIM, ECBilt-CLIO), and the EBM (TransEBM). Atmospheric and oceanic resolutions of some of the simulations are listed, with the vertical resolution always last. **(b, c)** Latitudinal distributions of mean temperature and precipitation for hierarchy categories based on panel **(a)**. For temperature, the biggest spread between the models and largest overall increases from the LGM to the Holocene can be found in the polar regions. For precipitation, the simulations and periods vary most in the tropics and the mid-latitude bands.

analysis. The reanalysis relies on model priors from 17 time-slice experiments from iCESM1 (Brady et al., 2019) and 539 geochemical proxy records of sea surface temperature. Using a Bayesian forward model, proxy values are estimated for given time steps at every proxy location from the model prior. This produces a forward-modeled proxy value different from the actual proxy value. To take uncertainties and the covariance between proxy location and the climate field into account, this difference is weighted by the Kalman gain for the update of the model prior temperature field. The resulting reanalysis estimates a global warming of 7.0 ± 1.0 °C from the end of the LGM to the PI (with PI defined as 1000–1850 CE), as it is contains an LGM state colder than reconstructed elsewhere (cf. Annan et al., 2022; Tierney et al., 2020; Shakun and Carlson, 2010). However, unlike other reconstructions, LGMR provides a gridded reconstruction of the surface temperature field covering the whole time period of interest here, not just the LGM. Here, we use the ensemble mean as the basis for our calculations of variability.

3 Methods

Climate can be represented by sets of observations in space and time. The field of a climate variable then refers to usually gridded spatial representations of that variable (von Storch and Zwiers, 1999). Conversely, a time series specifies the sequence of observations in time (Chatfield, 2016). As such, climate variables can be treated as random variables with associated probability distributions, and time series represent realizations of a stochastic process. Here, we analyze the statistical properties of the time series of surface temperature and precipitation in space and time by computing their moments and power spectra.

In order to compare the transient simulations, we first regrid them to a common T21 resolution, which is the lowest commonly used resolution in the ensemble. We further compute decadal and centennial means of the annual data to obtain the variability on those timescales. Then, we extract the time periods, LGM (23–19 kyr BP), Deglaciation (19–8 kyr BP), and Holocene (8–0 kyr BP), from all time series (see Fig. 2a). Finally, we remove the trend from the time series using a Gaussian filter with a kernel length equivalent to 4000 years,[5] which is the length of the LGM as the shortest time period we investigate. After detrending, we can assume that the resulting time series are (weakly) stationary, a requirement for the estimation of moments, along with the autocovariance function and thus the spectrum.

[5]Fig. S3 shows the effect of different choices of kernel lengths and compares this method to linear detrending with breakpoints.

3.1 Moments of a probability distribution

The distributions of surface temperature and precipitation cannot be assumed to be normal. Precipitation in particular often has heavier tails than a normal distribution (Franzke et al., 2020). To describe the shape of the distributions of climate variables, we turn to the four moments: mean, standard deviation, skewness, and kurtosis (Fig. 4; von Storch and Zwiers, 1999). We compute them for every grid box and time period. These are then area-averaged globally or zonally when providing the respective means of the moments.

The generalized moments of random variable X of a point A for a sample of size N are defined using the expected value as

$$\mu_n = \mathbb{E}[(X - A)^n] = \frac{1}{N}\sum_{i=1}^{N}(X_i - A)^n, \tag{1}$$

where n designates the nth moment (Papoulis and Pillai, 2002). There exist variations of this definition of the moments depending on normalization or bias correction terms. We introduce them in more detail in Sect. S3.1. Here, we focus on the definitions used in the analysis. To assess the background climate state, we use the arithmetic mean μ for $n = 1$ (Fig. 4a; Papoulis and Pillai, 2002) and the expected values computed about the origin, such that

$$\mu_1 \equiv \mu = \mathbb{E}[X] = \frac{1}{N}\sum_{i=1}^{N} X_i. \tag{2}$$

Considering the moments about the mean μ instead yields

$$m_n = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^n, \tag{3}$$

called nth central moment (Papoulis and Pillai, 2002; von Storch and Zwiers, 1999). For the second moment, variance, we use a symmetric unbiased estimator (Filliben and Heckert, 2024), yielding

$$\sigma^2 = \frac{N}{N-1}m_2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \mu)^2. \tag{4}$$

It describes the spread of the distribution – the larger the variance and its square root standard deviation σ, the larger the spread around mean μ (Fig. 4b).

For the third moment, skewness (s), we use

$$s = \frac{m_3}{m_2^{3/2}}. \tag{5}$$

Skewness describes the (a)symmetry of a distribution (von Storch and Zwiers, 1999). It is zero for a symmetric distribution, e.g., for the normal distribution. For negative skewness, the weight of the distribution is at higher values, with mode and median larger than the mean (Fig. 4c). This implies a stronger tail for lower than higher values: the distribution is "skewed left". For positive skewness, on the other hand, mode and median are smaller than the mean: the distribution is skewed towards higher values. To test whether any skewness found differs significantly from a normal distribution, we test its deviation from normality for significance using a t-test. For this test, the null hypothesis is that the skewness found and that of a corresponding normal distribution are the same and thus 0. We define the threshold for the p-value to be 0.05.

Adapting the skewness definition for $n = 4$, kurtosis, yields a kurtosis of 3 for a normal distribution. To derive an estimator which is 0 for normal distributions, kurtosis is often shifted by -3 to derive excess kurtosis k (Filliben and Heckert, 2024). Based on the fourth and second central moment, we calculate excess kurtosis as

$$k = \frac{m_4}{m_2^2} - 3. \tag{6}$$

We use excess kurtosis throughout the paper; for the sake of brevity, we will refer to it as kurtosis from now on. Kurtosis captures the heaviness of the tails of a distribution (Fig. 4d). If excess kurtosis is negative, the tails are thinner than those of a normal distribution. Conversely, positive excess kurtosis corresponds to heavier tails. Generally, positive kurtosis and skewness co-occur for datasets with more extreme values (Doane and Seward, 2011). As for skewness, we check again for non-normality using the hypothesis test derived by Anscombe and Glynn (1983) with a threshold for the p-value of 0.05. For all computations, we ignore rare not-a-number (nan) values in the temperature or precipitation fields. Changes in moments often occur concurrently and can then both enhance or counteract each other (Fig. 4).

3.2 Spectral analysis

In order to analyze how the variability in surface temperature and precipitation depends on timescale, we further compute the power spectral density (PSD), also called the power spectrum. If a process contains (quasi-)oscillatory components, the spectrum shows a peak at their periodicity with a certain width related to the damping rate of that process. The spectrum's background and scaling reflect the persistence (or memory) of the process (Ditlevsen et al., 2020).

The auto-covariance function for a random variable X_t at times t_1 and t_2 is given by the expectation value of its variance as

$$\gamma(t_1, t_2) = \mathbb{E}[(X(t_1) - \mu(t_1))(X(t_2) - \mu(t_2))], \tag{7}$$

where $\gamma(0) = \mathbb{E}[X^2]$ is the variance.

If the time series samples an ergodic, weakly stationary stochastic process, the auto-covariance and mean are independent of time and thus depend only on lag, $\tau = t_2 - t_1$. Assuming further that the data X_T are an excerpt of a theoretically infinite time series such that they are non-zero only

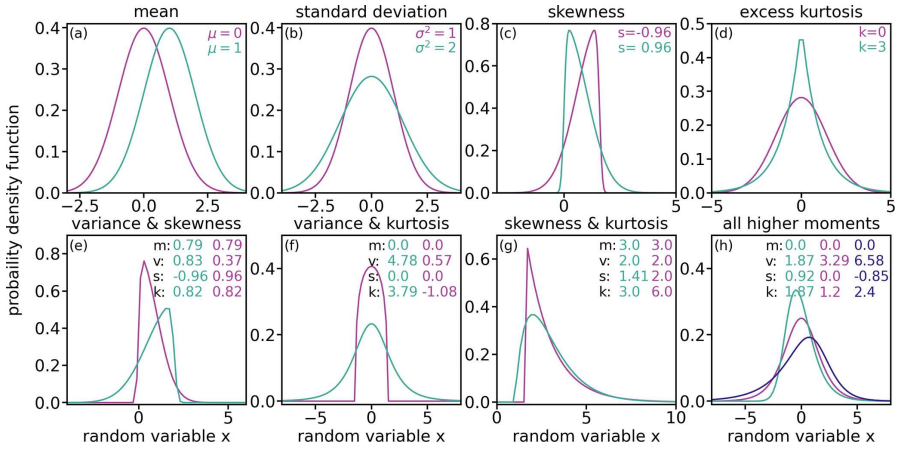


Figure 4. Visualization of changes in the moments of the distribution of a random variable, with individual changes on top and concurrent changes on the bottom. Distributions for a lower (pink) and higher (green) value are shown. Panels **(a)**–**(d)** show changes in just one moment: **(a)** mean, **(b)** variance, **(c)** skewness, and **(d)** excess kurtosis. Panels **(e)**–**(h)** show exemplary combinations of changes in the higher moments with constant mean: **(e)** opposite changes in variance and skewness, **(f)** concurrent change in variance and kurtosis, **(g)** concurrent change in skewness and kurtosis, and **(h)** changes in all higher moments. Figure S4 shows exemplary time series corresponding to the distributions.

for an interval $t \in [-\frac{T}{2}, \frac{T}{2}]$ (Ditlevsen et al., 2020), autocovariance can be written as

$$\gamma(\tau) = \mathbb{E}((X(t) - \mu)(X(t+\tau) - \mu)), \tag{8}$$

$$= \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t)X(t+\tau)\mathrm{d}t. \tag{9}$$

The PSD S for frequency ω is then defined as the Fourier transform $\mathcal{F}$ of the autocovariance

$$S(\omega) = \mathcal{F}(\gamma(\tau))(\omega), \tag{10}$$

$$= \int_{-T/2}^{T/2} \gamma(\tau)\exp^{-i\omega\tau}\,\mathrm{d}\tau. \tag{11}$$

The spectra of climate variables sometimes scale consistently across timescales following a power law with $S(\omega) \propto \omega^{-\beta}$, with β as the so-called scaling coefficient (Fredriksen and Rypdal, 2017; Lovejoy and Varotsos, 2016; Huybers and Curry, 2006; Wunsch, 2003). The scaling coefficient then reflects the persistence of the stochastic process.

To estimate PSDs, we apply the multi-taper method (Thomson, 1982; Percival and Walden, 1993) to the detrended time series. For data of finite length, this method reduces spectral leakage by computing separate spectra for orthogonal windows, so-called tapers, and averages the resulting spectra. Here, we use three tapers and estimate chi-squared distributed confidence intervals. We smooth the resulting spectrum and cut off artifacts at the low- and high-frequency end, such that, for a time series with a time step t_s, a period range of $[2t_s, 1000]$ remains. For comparing the variance in the different time periods across timescales, we further compute the spectral gain following Ellerhoff and Rehfeld (2021) by dividing the spectrum of the LGM and Deglaciation, respectively, by that of the Holocene.

4 Results

We examine changes in variability against a backdrop of a changing mean state, which we examine first (Sect. 4.1). Then, we evaluate temperature moments with respect to their dependence on mean state (LGM, Deglaciation, and Holocene), timescale, and model complexity (Sect. 4.2). Next, we focus on the forcing dependency by analyzing the influence of ice sheet reconstruction (Sect. 4.3.1), meltwater protocol (Sect. 4.3.2), and volcanism (Sect. 4.3.3) on surface temperature variability. Sections 4.4 and 4.5 repeat the analysis for precipitation. Then, we turn to the power spectra of temperature and precipitation, again considering state and forcing dependency, as well as differences related to model complexity (Sect. 4.6). We further compare the temperature spectra to results from the LGM reanalysis.

4.1 Mean state changes from LGM to Holocene across the ensemble

Between the LGM and the Holocene, all simulations show a mean warming and wetting, as evidenced by the increasing trends in GMST and GMP towards the Holocene (Fig. 2). Overall, MPI-ESM r1–r7 exhibit the largest temperature difference between the LGM and the Holocene with an aver-

age increase of 5.6 °C. Among the simulations, the anomaly is largest and the simulated LGM temperature is lowest for the simulations with GLAC1-D as the ice sheet reconstruction. In the whole ensemble, LGM cooling is widespread and especially pronounced in the high latitudes on land, with the exception of a few localized hotspots in a few of the simulations, e.g., an Alaskan warm patch in TraCE-21ka (Fig. S8g, h). Inter-simulation differences are generally larger in the high latitudes, especially in the Northern Hemisphere (Fig. 3b). For precipitation, the picture is more diverse, but, in most places and especially over land, a drier LGM is simulated. Some simulations show a locally wetter LGM in the tropics, a phenomenon mostly confined to the oceans. ESMs and GCMs show similar latitudinal profiles, while the EMICs miss some precipitation in the inner tropics and mid-latitude westerlies (Fig. 3c).

4.2 State and timescale dependency of surface temperature

Analyzing the higher-order moments of surface temperature reveals their dependence on timescale and model complexity (Figs. 5, S5). The standard deviation of surface temperature and its regional differences decrease towards longer timescales (Fig. 5a, d, g). Most of this decrease occurs between annual and decadal timescales. The only exception to this pattern is the EBM, which has low standard deviation across all periods. Differences between the three periods in the ensemble are concentrated in higher latitudes, especially in the northern polar regions. On annual scales, the Holocene standard deviation is smaller there than during the LGM and Deglaciation, which are similar to each other. For decadal and centennial scales, on the other hand, the Deglaciation stands out with higher standard deviation, while the Holocene and the LGM exhibit more similar levels. The LGMR shows similar patterns on centennial scales.

As a measure of asymmetry, skewness is positive (negative) if the weight of the distribution is at lower (higher) values with a high (low) value tail (see Sect. 3.1). Globally and across latitudes, skewness of temperature is usually close to zero, indicating little asymmetry (Figs. 5b, e, h, S5). The EBM is the exception as it shows pronounced negative skew, a signal that shrinks towards longer timescales. On centennial scales, the lack of skewness in the ensemble agrees with the results for the LGMR ensemble mean. In certain latitudinal bands, more significant deviations from zero exist. For example, MPI-ESM r1–6 and TraCE-21ka show positive centennial and, to a lesser degree, decadal skewness in the tropics during the Holocene, in particular over the ocean. This signal disappears with the addition of volcanic forcing in MPI-ESM r7 (Fig. 5h). In the MPI-ESM simulations, it is not reflective of physical processes in the climate system (Ellerhoff and Rehfeld, 2021). Therefore, we exclude it in the discussion of skewness and kurtosis going forward. Furthermore, TraCE-21ka shows a strong bipolar pattern during the Deglaciation on all timescales, with negative skewness in the Southern Hemisphere and positive skewness in the Northern Hemisphere (Figs. 5h and S9n). All other simulations show either no hemispheric pattern or, in the case of some MPI-ESM simulations and, to a lesser degree, HadCM3B, the opposite one, although with smaller magnitudes (Figs. 5h and S9k). For the MPI-ESM simulations, this bipolar pattern mostly shows up in the runs employing the GLAC1-D ice sheet (ch4, r1, r6, r7). The pattern is weaker for the melt-uniform runs (MPI-ESM r4 and HadCM3B r1) and disappears without meltwater forcing (MPI-ESM r3 and HadCM3B r2).

Kurtosis reflects the heaviness of the tails, defined here such that positive (negative) kurtosis corresponds to tails more (less) pronounced than those of the normal distribution (see Sect. 3.1). As for skewness, the kurtosis is mostly small on annual timescales, across periods and simulations in the ensemble and LGMR (Figs. 5c, S5c). Towards longer timescales, some regional differences emerge (Figs. 5f, i, S5f, i). TraCE-21ka again deviates during the Deglaciation, with temperatures that show strong positive kurtosis that is strongest in the high latitudes (Figs. S6o, r, and S9w). The EBM behaves differently on annual and decadal scales, simulating a strong positive kurtosis and thus heavy tails. On centennial scales, the EBM is again close to the more complex models.

4.3 Influence of forcings on the moments of surface temperature distributions

Using the sensitivity set, we investigate the interaction between forcings and moments of temperature distributions, in particular regarding the underlying ice sheet reconstruction (Sect. 4.3.1), meltwater protocol (Sect. 4.3.2), and volcanic forcing (Sect. 4.3.3).

4.3.1 Effect of ice sheet reconstructions on the shape of surface temperature distributions

Changes in standard deviation are regionally limited in response to the prescribed ice sheet reconstruction (Fig. 6a–f). On centennial timescales, ICE6G runs simulate smaller standard deviation in the northern North Atlantic compared to the runs using GLAC1-D (cf. MPI-ESM r1 and r6; Fig. 6a, b, d, e). This coincides with a reduced sea ice cover in these runs and a smaller temperature difference between the LGM and Holocene (Fig. S1a and c). The opposite pattern occurs in areas of Antarctic sea ice, especially the Weddell Sea, where ICE6G runs have higher standard deviation (Fig. 6b, e).

On centennial timescales, more areas in the simulations using GLAC1-D tend to have significant skewness, both positive and negative, than in simulations using ICE6G (Fig. 7). During the Deglaciation, the bipolar pattern of negative skewness in the Northern Hemisphere and positive skew in the Southern Hemisphere that emerges on decadal and centennial timescales is enhanced in the GLAC1-D simu-

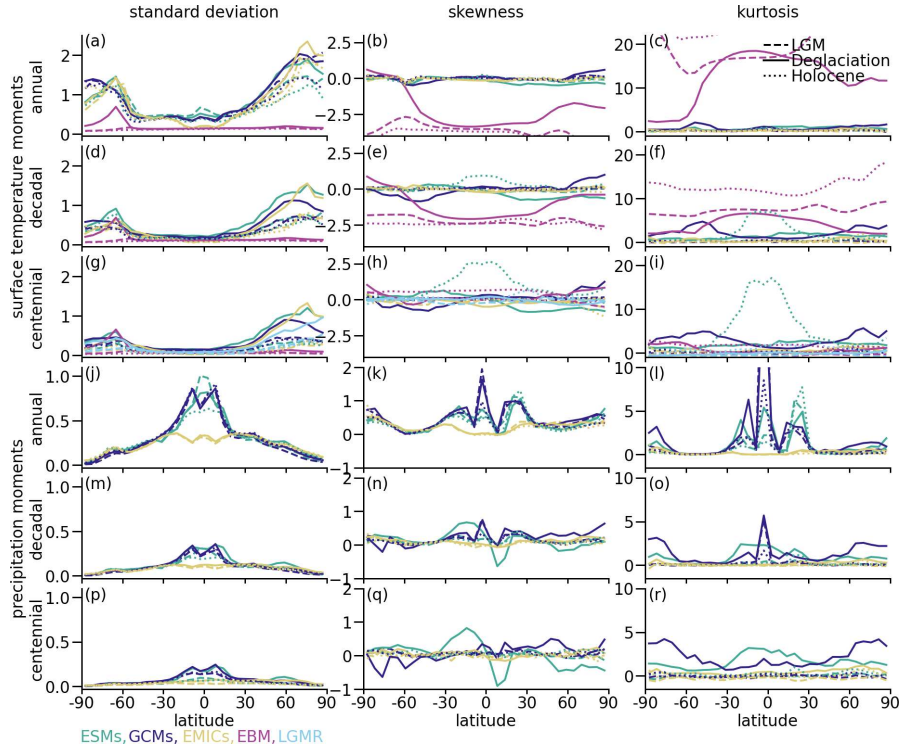


Figure 5. Changes of annual, decadal, and centennial higher-order moments of surface temperature **(a–i)** and precipitation **(j–r)** with latitude. For all simulations, standard deviation (left column, in units of °C for temperature and mm d^{-1} for precipitation), skewness (middle column, dimensionless), and kurtosis (right column, dimensionless) are shown. Results are differentiated according to period (LGM (dashed), Deglaciation (solid), and Holocene (dotted)) and complexity (ESMs (green), GCMs (dark blue), EMICs (yellow), and EBM (pink)). For centennial temperatures, moments of the LGMR ensemble mean are added in light blue. For temperature, the range of skewness and kurtosis of the EBM extend beyond what is shown, as does the kurtosis in HadCM3B for precipitation. Figures S6 and S7 show the individual simulations.

lations (Figs. 7, S13). On decadal and annual scales, the simulations with ICE6G can, at times, show opposite trends in comparison to their respective GLAC1-D simulations (Figs. S13, S14). The chosen ice sheet reconstruction has a limited impact on temperature kurtosis on annual to centennial timescales (Figs. 8, S15, S16).

4.3.2 Effect of meltwater protocols on surface temperature distributions

Meltwater forcing affects the moments particularly during the Deglaciation and in the North Atlantic (Figs. 6, 7, 8). The local melt-routed protocol is associated with the largest moments, and the no-melt scenario is associated with the smallest moments. This holds in particular in the North Atlantic. Furthermore, the melt-routed simulation has the strongest signal in the Southern Ocean.

For skewness, the North Atlantic is associated with a negative signal, again strongest for the melt-routed scenario (Fig. 7). The melt-routed runs further show positive skewness in the southern Atlantic and the southeastern Pacific. On the other hand, the uniform runs do not show consistent patterns: MPI-ESM r4 simulates negative skewness over large parts of the middle and high latitudes in the Northern Hemisphere, whereas HadCM3B r1 only has significant, mostly positive, skewness in a few regions (Fig. 7). The absence of meltwater forcing, as in MPI-ESM r3, results in a notable lack of significant skewness across the globe. Across the ensemble, including meltwater mostly introduces a shift to more positive kurtosis during the Deglaciation, especially in the North Atlantic (Fig. 8). This positive shift is stronger in the melt-routed (HadCM3B r2, MPI-ESM r2) than in the melt-uniform simulations (HadCM3B r1, MPI-ESM r4) on all timescales, although to varying degrees.

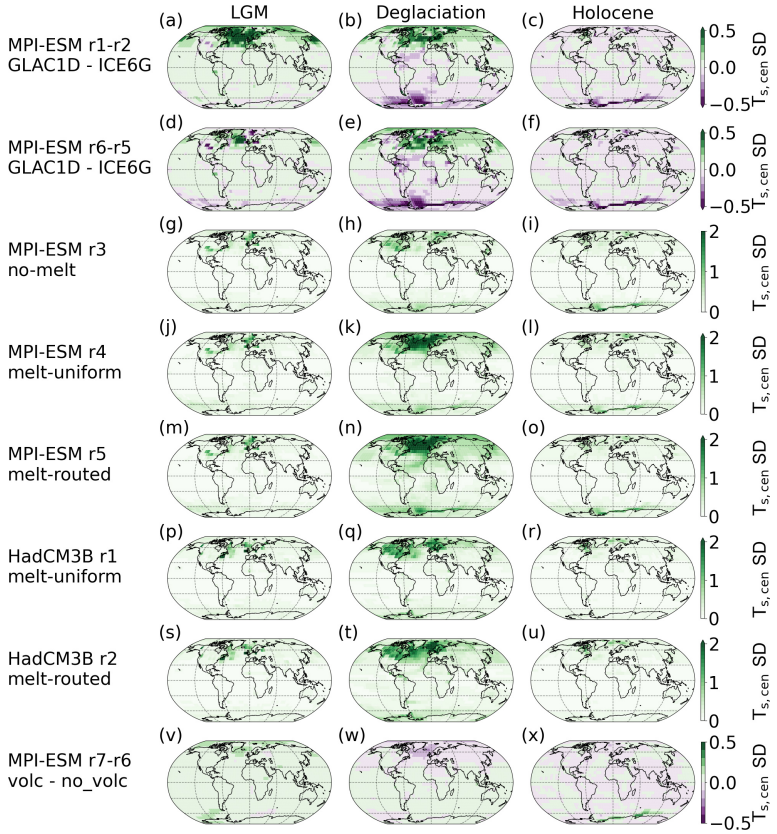


Figure 6. Regional effects of forcings on centennial standard deviation of surface temperature. **(a–f)** Influence of ice sheet forcing as differences between MPI-ESM runs using GLAC1-D and ICE6G. **(g–u)** MPI-ESM **(g–o)** and HadCM3B **(p–u)** simulations following different meltwater protocols. **(v–x)** Difference between MPI-ESM r7 with volcanic forcing and r6 without it.

4.3.3 Effect of volcanism on surface temperature distributions

For standard deviation of temperature, the effects of volcanism are mostly limited to shorter timescales but are overall small (Figs. 6v–x, S12). During the LGM, the run with volcanic forcing has a higher standard deviation than the one without. For the Deglaciation and Holocene, volcanism results in standard deviation that is greater at lower latitudes but smaller at higher latitudes.

Generally, volcanism results in negatively skewed temperature distributions or a reduction in positive skew, since it lowers temperatures after eruptions (Fig. S14j–o). It has the strongest effect on shorter timescales and during the LGM and Holocene. For the LGM, volcanic activity introduces a pronounced negative signal, mostly confined to the tropics. During the Holocene, skewness is decreased as well, turning the unphysical positive signal over the tropics (see Sect. 4.2 and Ellerhoff and Rehfeld, 2021) and most land areas into slightly negative skewness in parts of the tropics and effectively zero elsewhere (Fig. S15). On centennial scales, volcanic activity mainly manifests in skewness during the Holocene, where it again counteracts strong positive skewness in the tropics (Fig. 7r, u).

In contrast to ice sheet and meltwater forcings, volcanic forcing impacts kurtosis on all timescales and for all periods (Figs. 8p–u, S15j–o, S16j–o). For annual temperatures, it shifts the kurtosis to be positive in extended areas, particularly in the tropics and at northern mid-latitudes. This effect persists on decadal and centennial timescales for the LGM and Deglaciation. During the Holocene and on longer timescales, on the other hand, it reduces the low and mid-latitude band of positive kurtosis in the tropics. With volcanism, skewness and kurtosis in MPI-ESM r7 resemble HadCM3B r2 on centennial scales (Figs. 7, 8). On annual scales, however, volcanic forcing introduces skewness and

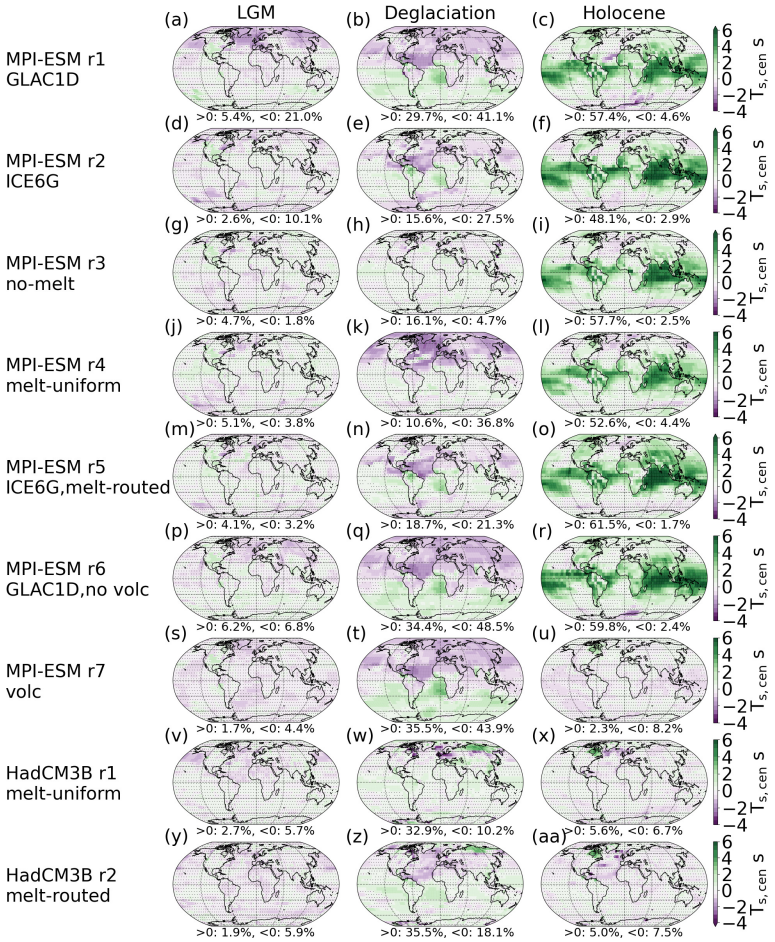


Figure 7. Regional effects of forcings on centennial skewness of surface temperature. Forcings are noted along with the run name for each row. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are given. Areas where changes are non-significant are hatched.

kurtosis patterns unlike any other simulation (Figs. S14, S16).

4.4 State dependency of precipitation at annual to centennial timescales

The standard deviation of precipitation is high in the tropics and decreases towards higher latitudes (Fig. 5). In particular, it is high over the tropical oceans in the region of the intertropical convergence zone (ITCZ; Figs. 5j and S11a–i), where mean precipitation is also highest (Fig. 3c). The tropical band of increased standard deviation exists only to a lesser degree in the EMIC simulations (Fig. S6). State dependency exists on centennial but only very rarely on annual timescales. During the Deglaciation, the tropical pattern of enhanced standard deviation remains on centennial scales but is less pronounced than on annual scales. The spatial patterns of the LGM and Holocene, on the other hand, are more homogeneous, and the tropical standard deviation is similar to that at other latitudes (Fig. 9). The only exception is the FAMOUS simulation, which has enhanced tropical standard deviation for all three periods and the overall largest magnitudes in the ensemble on centennial scales (Fig. S6p).

The higher moments show more diverse patterns for precipitation (Figs. 5, S5). For skewness, the simulations mostly show positive precipitation skewness on annual scales, with some negative skewness in areas near the Equator and little difference between the periods (Fig. S11). The tropics

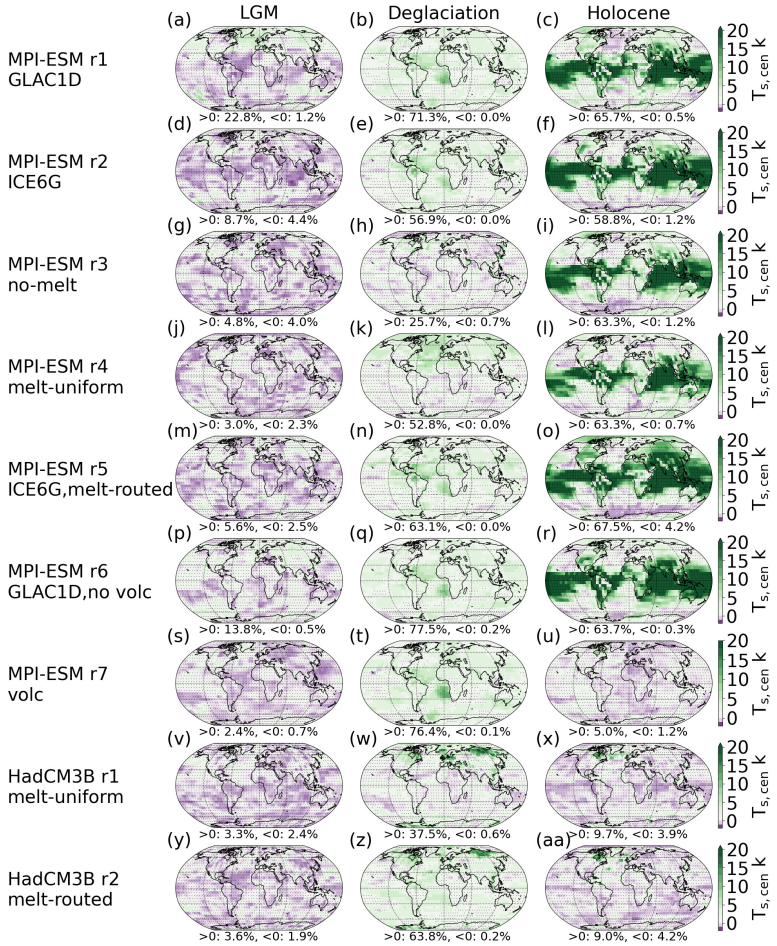


Figure 8. Regional effects of forcings on centennial kurtosis of surface temperature. Forcings are noted along with the run name for each row. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are given. Areas where changes are non-significant are hatched.

also show the largest positive skewness. The EMICs simulate the smallest skewness, although the positive deviation from zero is still significant almost everywhere (Figs. 5, S11). The ESMs and GCMs, on the other hand, simulate very similar patterns on annual scales. Starting on decadal and even more strongly on centennial scales, the patterns diverge between periods and simulations (Figs. 5q, S10j–r). During the LGM and Holocene, centennial skewness is close to zero and thus indicates predominantly symmetric distributions. During the Deglaciation, skewness patterns are far more diverse, with a larger spread and including negative excursions. These center mostly around the Equator but also sometimes in the high northern (for MPI-ESM simulations with a GLAC1-D ice sheet; Fig. 10) or the high southern latitudes (for TraCE-21ka; Fig. S10n). In a bipolar pattern, TraCE-21ka further simulates high positive skewness in the high northern latitudes (Figs. S6n, q and S10n). The EMICs and FAMOUS show almost no significant skew in all periods on decadal and centennial scales (Fig. S6).

Precipitation kurtosis is mostly positive on annual scales across the periods in ESM and GCM simulations, in particular in the tropical regions (Figs. 5, S5, S11). The LGM and Holocene exhibit no significant kurtosis on longer timescales. During the Deglaciation, though, positive kurtosis persists. The EMICs, on the other hand, have some significant kurtosis only during the Deglaciation on annual scales and otherwise show no significant deviation from zero in contrast to the more complex models.

4.5 Changes in precipitation distribution shape in response to forcings

For the moments of precipitation, forcing dependency can mostly be found for meltwater forcing (Figs. 9, 10). Effects of the ice sheet reconstructions are mostly limited to centennial skewness in the tropics, in the North Atlantic zone, and in mid- and high latitudes in Eurasia (Fig. 10). These are the very same areas where temperature skewness changes the most. The skewness patterns in the tropics agree between the reconstructions but are enhanced in the GLAC1-D simulations. An exception is the high northern latitudes during the Deglaciation, where skewness is positive in the ICE6G simulations and negative in the GLAC1-D ones (Fig. 10).

Meltwater forcing affects all Deglacial moments, in particular in the tropics and North Atlantic. Standard deviation is largest in the melt-routed runs (MPI-ESM r2 and HadCM3B r2) and smallest in the no-melt simulation (Fig. 9). Injecting meltwater introduces significant skewness in precipitation distributions during the Deglaciation on centennial timescales (Fig. 10). Both melt-routed simulations (MPI-ESM r5 and HadCM3B r2) have a signal of negative skewness in the eastern equatorial Pacific, with a positive signal to the south of it. Only the positive signal remains somewhat in the melt-uniform runs. These MPI-ESM and HadCM3B runs further exhibit strong negative skewness in high northern latitudes, although in different areas. The influence of meltwater forcing on centennial kurtosis during the Deglaciation shows up predominantly as positive kurtosis (Fig. S7). Melt-routed runs have more positive kurtosis in the tropics, whereas melt-uniform simulations have more in the high northern latitudes. Overall, ice sheet reconstruction, meltwater protocol, and volcanic forcing (see annual moments in Figs. S17, S18) usually affect the moments of temperature more than those of precipitation.

4.6 Spectral analysis of the variability in surface climate

To add to the analysis of variability, we examine the spectra of surface temperature and precipitation during the LGM, Deglaciation, and Holocene.

4.6.1 State and timescale dependency of global and regional surface temperature spectra

Generally, we find temperature spectra that increase towards longer timescales, some of which level off at multi-centennial scales (Fig. 11). This pattern is particularly strong for the Deglaciation, where it can also be found across latitudinal bands (Fig. S21). However, the regional spectra can be flat during the LGM and Holocene, for example, in the tropics or mid-latitudes after a scale break at multi-decadal scales. The spread between the simulations increases towards longer timescales, especially during the Deglaciation and the Holocene. During the LGM, all MPI-ESM simulations show increased variability with a broad peak on interannual scales, such that the spread is comparatively large there. This increased variability originates in the tropical regions (Fig. S21j, m) and relates to the simulated ENSO, which is enhanced in the MPI-ESM simulations during the LGM. None of the other simulations exhibit similarly increased ENSO activity during the LGM, and the MPI-ESM signal has been suggested to be inadvertently amplified (Ellerhoff and Rehfeld, 2021). In the tropics, the spread between simulations remains similar across timescales in all three periods.

The PSD ratios between the LGM and the Holocene depend on simulation and timescale (Fig. 11d, e). In some simulations, the LGM has larger PSD across all timescales (e.g., MPI-ESM r7); for TraCE-21ka, it is the Holocene. For most simulations, it changes with timescale, as many have larger Holocene spectral power on centennial scales but larger LGM power on decadal and millennial scales (Fig. 11d). With very few exceptions, the deglacial spectrum contains the largest power, especially above centennial timescales (Fig. 11e). This pattern mostly holds for regional spectra across latitudinal bands (Figs. S21, S22). This is partially because the increase in power from interannual to millennial timescales is steepest during the Deglaciation, whereas the scaling is smaller for the LGM and the Holocene.

4.6.2 Forcing dependency of the temperature spectrum to ice sheet reconstruction, meltwater forcing, and volcanism

The temperature spectra vary significantly in magnitude and pattern between simulations and in response to external forcing differences (Figs. 11, S21). For GLAC1-D simulations, we find increased variability during the LGM on decadal to centennial timescales, mainly in the Northern Hemisphere mid- and polar latitudes. Meltwater forcing, on the other hand, has the strongest impact during the Deglaciation, also for the northern high latitudes. For the global spectra, there is little difference between runs using the melt-routed (MPI-ESM r2 and HadCM3 r2) versus melt-uniform (MPI-ESM r4 and HadCM3 r1) protocol. However, the run without meltwater forcing has the lowest Deglacial variability among the MPI-ESM simulations. Volcanic forcing strongly impacts the spectrum of simulated surface temperatures. MPI-ESM r7, the run with volcanic forcing, has the largest PSD from interannual up to centennial timescales during all three periods. While the other MPI-ESM runs show a drop in PSD on interannual scales, especially during the LGM, r7 shows a consistent increase in variability until at least centennial scales. On longer timescales, it is also on the upper end of simulated variability.

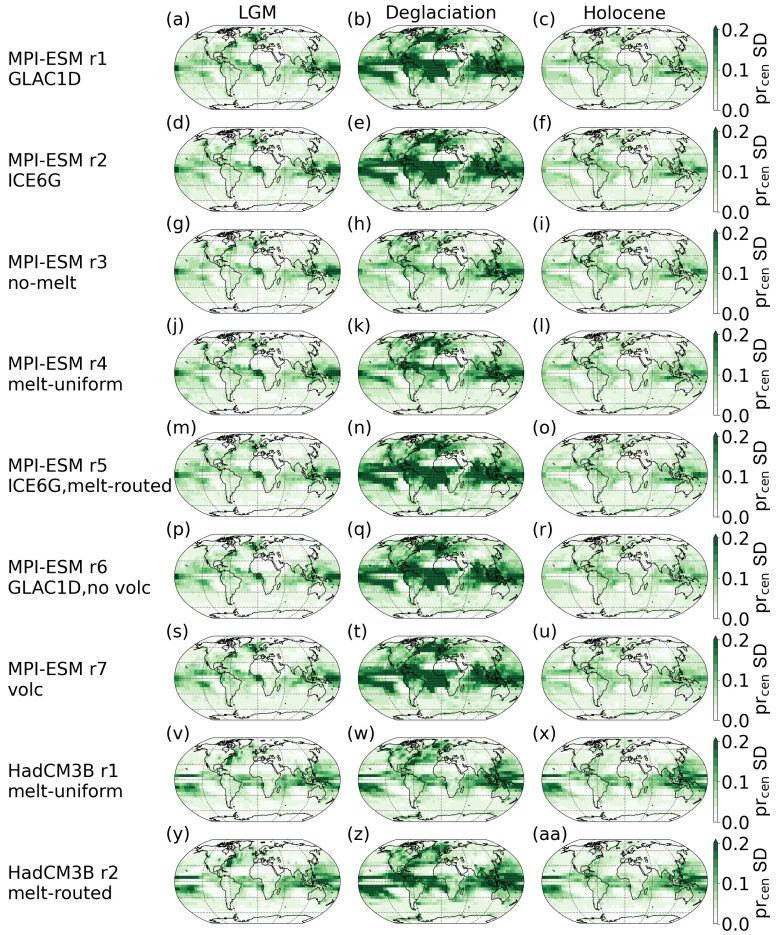


Figure 9. Regional effects of forcings on centennial standard deviation of precipitation. Forcings are noted along with the run name for each row.

4.6.3 Dependence of the spectral power of temperature on model complexity

For the most part, GCMs and EMICs display less spectral power than the ESMs up to multi-centennial scales. MPI-ESM ch4 can be an exception, as it agrees with the other MPI-ESM simulations on interannual scales during the LGM but is then more similar to the non-MPI-ESM simulations on multi-decadal scales. It further exhibits a strong 200–400-year periodicity during the LGM that is absent in all other simulations. This signal originates in the Southern Hemisphere sea ice, grows stronger towards higher latitudes, and extends into the tropics (Figs. S21, S29d, S30, Sect. S7.1). The spectral power of the EBM is at the higher end on decadal to centennial timescales. There, MPI-ESM r7, using the same volcanic forcing reconstruction, is often the only simulation with more power. However, it levels off around centennial scales, with only moderate increases in variability afterwards, such that its variability is among the lowest on millennial scales, indicating a lack of persistence.

4.6.4 Comparison of the simulated surface temperature variability to the LGM reanalysis

The magnitude of the LGMR power spectrum generally falls within the range of the ensemble. During the LGM, it shows similar levels of variability to the GCMs, matching their increase towards millennial scales. For the Deglaciation, it starts at the lower end of variability but again exhibits a strong increase. This increase suggests a larger scaling in comparison to the ensemble within the limited range of timescales covered by the LGMR spectrum. For the

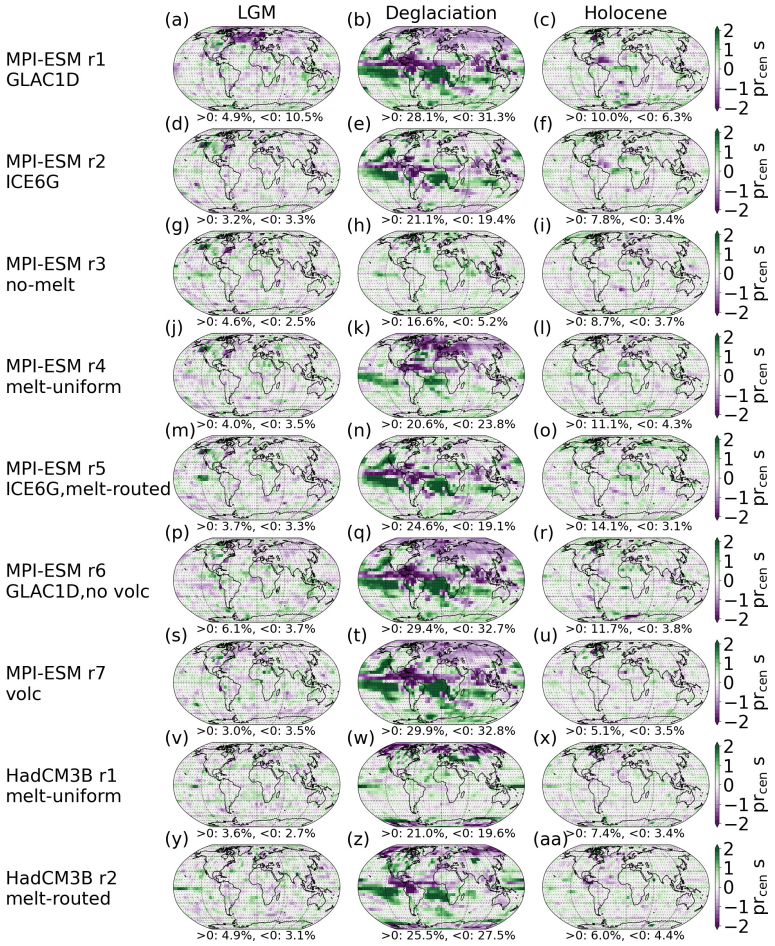


Figure 10. Regional effects of forcings on centennial skewness of precipitation. Forcings are noted along with the run name for each row. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are given. Areas where changes are non-significant are hatched.

Holocene, the LGMR is at the lower end of spectral power and mostly below the range of LGM-to-Holocene spectral ratios found in reconstructions by Rehfeld et al. (2018) (Fig. 11d). Most simulations also fall below that range, especially on decadal to centennial timescales. Above centennial scales, many of the MPI-ESM runs are in agreement with it. Notably, MPI-ESM r7 agrees with the range found by Rehfeld et al. (2018) on most timescales, although it is at the lower end for multi-decadal scales.

4.6.5 Dependency of precipitation spectra on state, timescale, model complexity, and forcing

With the exception of MPI-ESM r7, global spectra are quite flat across short timescales and then feature an increase starting on multi-decadal (MPI-ESM r1–r6) or centennial scales (the remainder of the ensemble). This increase levels off again around millennial scales. For the Deglaciation, the increase is very sharp, whereas it is smaller and often more gradual during the Holocene and LGM. Thus, the largest variability is found during the Deglaciation above centennial timescales, with FAMOUS as the only exception. Regionally, LGM and Holocene spectra can be flat, especially in the tropics, with only some simulations showing an increase in vari-

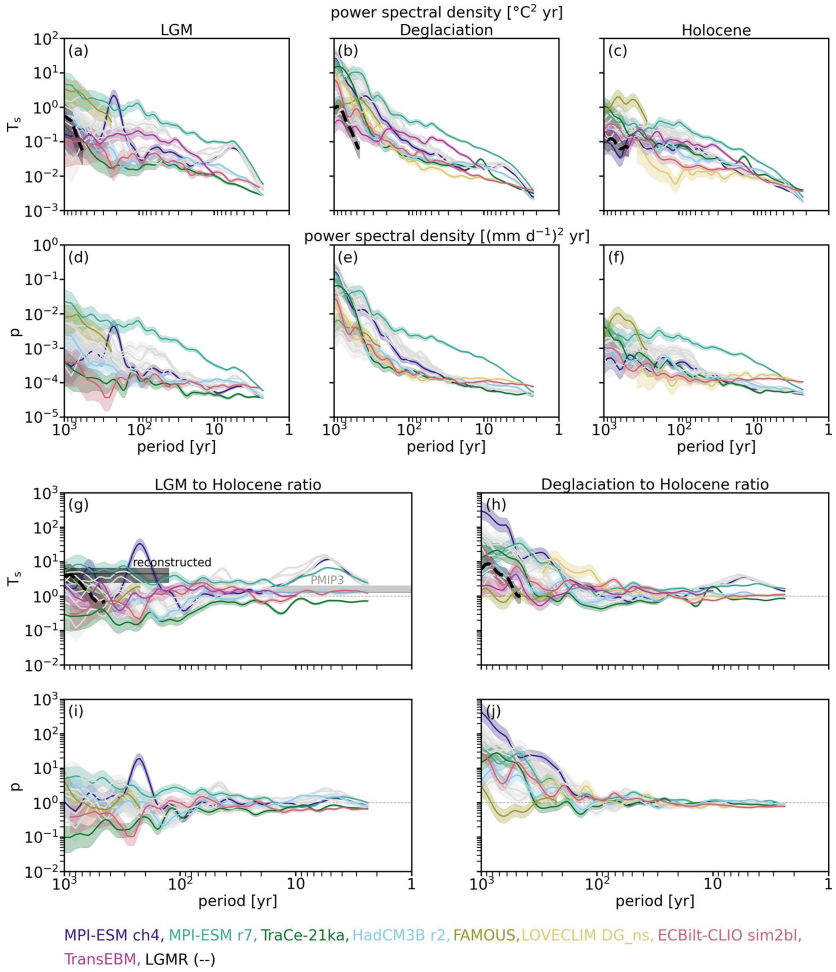


Figure 11. Spectra and spectral ratios of surface temperature (top row) and precipitation (bottom row) variability with chi-squared distributed confidence intervals. Both the x and the y axis are shown on a logarithmic scale. The spectra are separated by time period: **(a, f)** LGM, **(b, g)** Deglaciation, and **(c, h)** Holocene. The spectral ratios highlight the differences between the periods showing the LGM-to-Holocene **(d, i)** and the Deglaciation-to-Holocene **(e, j)** ratios. The sensitivity set (here in gray) is shown in Fig. S20. In panel **(d)**, the estimated ranges of Rehfeld et al. (2018) of the multi-centennial to millennial LGM-to-Holocene variance ratio based on proxy reconstructions (reconstructed) and interannual variability based on the PMIP3 ensemble (PMIP3) are marked for comparison.

ability towards longer timescales for mid- and polar latitudes (Fig. S22). The MPI-ESM simulations generally have larger precipitation variability during the LGM than the Holocene, whereas all other models show the opposite (Fig. 11i). All simulations show larger precipitation variability during the Deglaciation than during the Holocene, with the difference increasing towards longer timescales (Fig. 11j). Precipitation variability is among the highest in the MPI-ESM and lowest in the EMIC simulations, globally and regionally (Fig. S23).

Inspecting the effect of forcings on the spectra of the sensitivity set reveals similar relationships to that for the temperature spectra. Using GLAC1-D leads to larger LGM variability on multi-decadal and longer timescales for mid- and high northern latitudes (Fig. S24). The no-melt protocol shows a distinct lack of Deglacial variability, especially in the Northern Hemisphere. In all periods, MPI-ESM r7 with its volcanic forcing has significantly larger variability than all other simulations on interannual to centennial timescales. This difference is even more pronounced than for the temperature

spectra and reaches up to 1 order of magnitude. Regionally, too, the spectral power of MPI-ESM r7 is always on the upper end, such that it stands out even among the MPI-ESM simulations.

5 Discussion

We investigate variability changes before, during, and after a period of global warming in an ensemble of transient simulations of the Last Deglaciation. Among them, variability differs considerably (see Table 2) depending on the following:

- Timescale. Surface temperature shows a decrease in standard deviation, larger absolute skew, and an increase in kurtosis towards longer timescales (Sect. 4.2). For precipitation, standard deviation decreases with timescale (Sect. 4.4). During the LGM and Holocene, skewness and kurtosis of precipitation tend to decrease, whereas there is usually an increase with timescale during the Deglaciation.

- Background state. Generally, the state dependency of surface temperature increases with timescale for all moments and is largest during the Deglaciation (Sect. 4.2). For precipitation, trends differ between moments and are more complex (Sect. 4.4).

- Forcings. Simulations that differ only by ice sheet reconstruction diverge most on long timescales, although differences can be found even for annual variability (Sect. 4.3.1). For surface temperature, the impacts are largest during the Deglaciation for all moments. For precipitation, the employed ice sheet reconstructions mainly affect skewness (Sect. 4.5).
  The chosen meltwater protocol primarily affects the moments on multi-decadal and longer timescales (Sect. 4.3.2). On these, any kind of meltwater will increase the standard deviation, absolute skewness and kurtosis for both temperature and precipitation, with the largest values for routed meltwater. For temperature, these trends manifest mostly in the North Atlantic, and, for precipitation, they manifest around the equatorial Atlantic and eastern Pacific oceans (Sect. 4.5).
  Volcanic forcing primarily affects the moments of temperature with little effect on those of precipitation (Sect. 4.3.3, Sect. 4.5). Its presence creates a low-temperature tail. For both temperature and precipitation, volcanic forcing increases spectral power on all timescales.

- Model complexity. There are substantial differences in simulated variability between categories of model complexity and models of similar complexity (Sect. 4.2, 4.4). Except for the standard deviation of temperature, EMICs simulate very little change between states and mostly have higher moments close to zero. In this respect, they differ strongly from ESMs and GCMs, which simulate more complex patterns and for which some state dependency exists on all timescales and for all moments.

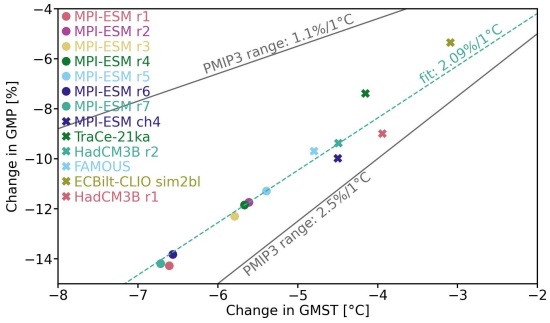


Figure 12. Hydrological sensitivity during the LGM and Holocene: percentage change in LGM-to-Holocene GMP against the change in GMST. The line indicates a 2 % change in precipitation per degree of temperature change. The data are fitted linearly with intercept 0.

Firstly, we discuss our findings on mean state changes, against which we evaluate our hypotheses on state dependency. We then discuss the above findings relating to climate variability.

5.1 Large range in the underlying simulated and reconstructed mean state changes

To understand variability in its context, it is important to assess the simulated mean changes using observational records. The simulated LGM to Holocene changes in GMST range from 3.0 to 6.6 °C (Tables 2 and S1). Proxy-based reconstructions and data assimilation approaches provide similar ranges. Among more recent estimates, Osman et al. (2021) suggest a warming of 7.0 ± 1.0 °C from the Deglaciation onset to the PI, while Tierney et al. (2020) estimate a temperature difference of 6.1 °C (5.7, 6.5). On the other hand, Annan et al. (2022) propose 4.5 ± 0.9 °C and Shakun and Carlson (2010) reconstruct a minimal warming of 4.9 °C for the LGM at 22 kyr BP relative to the Altithermal at around 8 kyr BP. While some of the differences can be explained by different reference periods, uncertainty around the level of warming remains. Agreement is larger with respect to spatial patterns of warming, with larger changes in the Northern Hemisphere than in the Southern Hemisphere, towards higher latitudes in both hemispheres and over land and areas of melting ice sheets. The temporal patterns of GMST change, however, differ a lot between simulations. This includes, but is not limited to, the onset and termination of deglacial warming and the timings of periods of abrupt change.

Precipitation since the LGM is less studied than temperature, and even fewer proxy reconstructions of hydroclimate exist. There is also no data product like the LGMR, which can be used for comparison of spatial patterns. Therefore, we consider simulated hydrological sensitivity to contextualize our results (Fig. 12). The Clausius–Clapeyron relation estimates a 7 % change in saturation water vapor per degree temperature change. This is further constrained based on the surface energy balance and its effect on evaporation and water availability, such that a realistic range for precipitation change per degree temperature change is 1 %–4 % (Li et al., 2013). Here, we find a hydrological sensitivity of about 2.09 % per degree change in GMST. This agrees with the ranges given for CMIP5/PMIP3 equilibrium simulations by Li et al. (2013) of 1.5 %–3 % per degree Kelvin and Rehfeld et al. (2020) of 1.1 %–2.5 % per degree Kelvin.

5.2 Increasing state dependency of variability with timescale

5.2.1 Increased standard deviation and spectral power of surface temperature during the deglacial transition

For all moments of surface temperature, we find that state dependency generally increases with timescale. Simultaneously, the ensemble simulates a reduction in the spread of temperature distributions towards longer timescales with the overall largest values during the Deglaciation. On annual scales, areas of large standard deviation often exhibit large mean changes, too, but this does not hold for longer timescales. At decadal and centennial scales, standard deviation is largest over the high-latitude oceans, particularly in areas with seasonal sea ice cover and for the Deglaciation in the North Atlantic in response to the changes in the Laurentide ice sheet. Because the sea ice cover shrinks with warming and increases during periods of abrupt cooling, standard deviation is larger for the Deglaciation than for the LGM and Holocene across timescales (Fig. S26). This importance of sea ice for local variability is in line with results from Ellerhoff et al. (2022).

The ratio of LGM-to-Holocene variance mostly shows higher LGM variance, with values between 1 and 2 (Tables 2 and S1). These results resemble those of Rehfeld et al. (2018), who found ratios between 1 and 3 on interannual to decadal scales based on CMIP5/PMIP3 equilibrium simulations. The smaller Holocene Equator-to-pole temperature gradient, itself the result of polar amplification, has been suggested as the reason driving the smaller Holocene variability in comparison to the LGM (Rehfeld et al., 2018). Similarly, Shi et al. (2022) conclude that interannual temperature variance in PMIP3/4 LGM simulations is 20 % higher than in PI simulations as a consequence of an increased meridional temperature gradient, in particular at mid-latitudes. Indeed, we find an enhanced meridional temperature gradient during the LGM for some regions and models, which is correlated with the temperature gradient, including at mid-latitudes. However, in other regions and models, we find no such large-scale increase in the gradient nor a correlation to the variance ratio (Fig. S31).

Few model–data comparisons of variance exist that include the LGM. When comparing the LGM and Holocene, proxy reconstructions show that LGM variance is globally about 4 times higher on timescales from 500 to 1750 years (Rehfeld et al., 2018). The differences we find here are smaller (Table 2). For PMIP3 simulations, Rehfeld et al. (2018) similarly found a smaller simulated than reconstructed ratio. Simulated LGM and Holocene spectra resemble each other in their general shape, with increasing scaling variability towards longer timescales. The absence of a characteristic timescale in a stochastic process leads to such scaling behavior with similar statistical properties across scales (Mandelbrot and van Ness, 1968; Ellerhoff and Rehfeld, 2021). Both global and regional spectra (Figs. 11, S21, S22) suggest inter-model differences in scaling and scale breaks. Investigating these in greater detail by computing scaling factors that quantify the relationship between timescale and variability and identifying scale breaks could shed more light on the nature of the scaling behavior. Scaling has been suggested to differ between glacial and interglacial climates (Huybers and Curry, 2006; Nilsen et al., 2016; Lovejoy, 2015; Rypdal et al., 2013; Roe and Steig, 2004), and we also find state-dependent features (Fig. 11). This is in contrast to the lack of state dependency between global spectra of equilibrium LGM and PI simulations found by Ellerhoff et al. (2022). Since the differences are especially apparent on longer timescales, this might point towards the long-term memory effects or transient forcings missing in such equilibrium simulations.

The Deglaciation shows enhanced levels of variance in comparison to the LGM and Holocene on decadal and centennial timescales (Fig. 5d, g) and larger spectral power above centennial scales (Fig. 11e). Northern high latitudes are the largest source of this state dependency, with further significant state dependency in high southern and northern mid-latitudes (Fig. 5d, g). This reflects the dynamic nature of the Deglaciation with the melting of ice sheets, the resulting freshwater flow, and the subsequent reorganization of the climate system. The enhanced variability in the spectra matches the oscillatory behavior that Clark et al. (2012) described, especially on and for millennial scales.

Comparing simulations and reconstructions, Zhu et al. (2019) argue that simulated and reconstructed temperature variability agree on the global scale. That analysis considers long proxy records and reconstructions reaching back to 5 Myr BP and three of the simulations used in this paper: TraCE-21ka, LOVECLIM DG_ns, and ECBilt-CLIO. With respect to agreement on the global scale, our results agree when comparing the simulations to the LGMR. On local and regional scales, however, climate models have repeatedly

been found to underestimate variance on longer timescales (Laepple et al., 2023). Similarly, we find notable differences between regional patterns in the LGMR and the simulations (see Fig. S19).

Considering the longer timescales included in the global spectra, some of the MPI-ESM runs do fall into the range suggested by reconstructions (Fig. 11). The LGMR only falls in the range close to millennial timescales. It shows a steep increase in LGM spectral power towards longer timescales, which translates into an increasing LGM-to-Holocene ratio that is unlike that of most simulations in the ensemble. However, proxy records covering both the LGM and the Holocene are sparse, and both Shi et al. (2022) and Rehfeld et al. (2018) suggest considerable spatial heterogeneity in the variance ratios. The sparse sampling of variance around the globe likely biases the LGMR and thus the comparison. While the LGMR exhibits more similar levels of variability to most of the ensemble, comparison to the other reconstructions suggests a lack of simulated regional variability on multidecadal timescales and beyond. The difference between the LGMR and other reconstructions leaves uncertainty around the models' abilities to capture climate variability and thus a potential lack of variability in future simulations with consequences for projected changes in the frequency and intensity of extremes. To decrease this uncertainty, our findings can provide a basis for more in-depth model–data comparisons of simulated variability during periods of warming. Realizing the full potential of such an analysis requires an ensemble of coordinated experiments using common protocols (as for some of the simulations here) and improved reconstructions of past variability.

5.2.2 Larger absolute surface temperature skewness and kurtosis towards longer timescales

Temperature skewness and kurtosis, describing the asymmetry and heaviness of the tails of a distribution, respectively, deviate more from zero towards longer timescales, indicating more non-Gaussian distributions and changes in extremes (Figs. 5, S5). During the Deglaciation, mid- and high latitudes show enhanced values of skewness and kurtosis in ESMs and GCMs. Changes in skewness can be an early warning signal of abrupt changes (Skelton et al., 2020; He et al., 2013; Guttal and Jayaprakash, 2008). The simulations in our ensemble undergo large-scale changes that can, at times, be abrupt in response to the prescribed forcings and boundary conditions. Regionally, there may be abrupt change due to internal dynamics of atmosphere, sea ice, ocean, or land surface. As our analysis computes variability over the whole LGM, Deglaciation, and Holocene, it is ill-suited to determine whether skewness changes occur locally or regionally ahead of the abrupt changes. This would instead require an investigation into skewness changes over time. Such an analysis would be of particular interest if it included high-resolution proxy data for key variables experiencing abrupt change and for potential tipping elements.

The LGMR contains some enhanced tropical skew and kurtosis, mainly restricted to the Atlantic and continental areas (Fig. S19). During the Deglaciation, the LGMR shows a pattern of negative skew in the North Atlantic consistent with most simulations with any kind of meltwater forcing. It further exhibits state dependency in the spatial patterns of skewness and kurtosis and an increasing global mean kurtosis from the LGM to the Holocene (Fig. S5i). On the other hand, the LGMR shows very few areas of significant deviations from normality and patterns that are generally more smooth than almost all simulations. Here, we use the LGMR ensemble mean as the basis of all computations. Individual ensemble members consistently show larger standard deviation with common spatial patterns (not shown). For skewness and kurtosis, however, patterns are very inconsistent between ensemble members but, as for the ensemble mean, with mostly Gaussian distributions. This is likely related to the data assimilation, as an ensemble Kalman filter assumes that the state vectors are Gaussian (Evensen et al., 2022). While this assumption works for many weakly non-linear systems (Evensen et al., 2022), it is likely insufficient for the tails of the distribution. Furthermore, the higher-order moments are affected by the spatial averaging inherent in field reconstructions from individual proxy sites (Director and Bornn, 2015; McKinnon et al., 2016; Haylock et al., 2008). Due to the sparse coverage in space and time of the proxy records, they only provide weak constraints during the assimilation, especially further back in time. This likely also explains why the variability in the LGMR is largest during the LGM, when proxy availability is smallest. A comparison at individual proxy locations could shed more light on model–data differences and similarities and could ascertain whether the LGMR indeed has too little variability, as indicated by its differences from both simulations and other proxy reconstructions.

5.2.3 Precipitation distributions show trend towards drier and less extreme years in the tropics from the LGM to the Holocene

For precipitation, mean changes are similarly varied as for surface temperature, with a global mean LGM-to-Holocene wetting of 0.15–0.39 mm d^{-1}. The ESMs and GCMs simulate some drying towards the Holocene over the tropical oceans, with wetting almost everywhere else. The moments are generally largest on annual timescales. In the tropics, ESMs and GCMs show some state dependency but not as much as for temperature. Tropical precipitation mostly decreases from the LGM to the Holocene, especially over the ocean. Over the tropical oceans, standard deviation can similarly decrease from the LGM to the Holocene, while it increases almost everywhere else. During the Deglaciation, some high-precipitation years remain, as indicated by the un-

Table 2. Summary of LGM-to-Holocene changes in the moments. For every moment and timescale, the values according to model complexity are listed as ESMs, GCMs, EMICs, and EBM. For the mean, the absolute value of the difference is listed; for the other moments, the ratios are. Ratios are first computed for the individual simulations (see Tables S1 and S2 for surface temperature and precipitation, respectively) and then averaged by category. For EMICs, this only includes ECBilt-CLIO here. Note that, on centennial scales, the ESM and GCM categories include more simulations (MPI-ESM r3 and r4 and FAMOUS) than on the annual and decadal scale, leading among other things to a difference in average mean change. It shows that centennial Holocene skewness of ECBilt-CLIO is very close to zero, which produces a very large EMIC ratio. Very large ratios, as for the skewness of centennial precipitation distributions, are a result of moments very close to zero for the LGM and Holocene.

	$\Delta m_{\mathrm{hol-lgm}}$				$v_{\mathrm{lgm}}/v_{\mathrm{hol}}$				$s_{\mathrm{lgm}}/s_{\mathrm{hol}}$				$k_{\mathrm{lgm}}/k_{\mathrm{hol}}$			
	ESM	GCM	EMIC	EBM	ESM	GCM	EMIC	EBM	ESM	GCM	EMIC	EBM	ESM	GCM	EMIC	EBM
T_{ann}	5.81	4.44	3.95	4.12	1.75	1.27	1.16	1.27	0.07	−2.57	0.87	0.90	0.51	0.88	0.38	0.60
T_{dec}	5.81	4.44	3.95	4.12	1.43	1.25	1.03	1.39	1.76	3.44	0.76	0.98	0.26	0.29	0.09	0.60
$\mathrm{T}_{\mathrm{cen}}$	5.76	4.10	3.95	4.12	1.52	1.50	0.74	1.56	0.25	0.80	0.84	−0.01	−0.26	−0.25	−0.17	−0.29
p_{ann}	0.34	0.24	0.25		1.17	0.82	0.93		1.08	1.09	1.36		1.10	1.63	1.36	
p_{dec}	0.34	0.24	0.25		1.10	0.82	0.93		1.02	1.14	1.33		0.81	1.31	−0.13	
p_{cen}	0.34	0.22	0.25		1.07	0.84	0.88		0.64	1.01	101.84		−0.22	−8.97	−2.44	
LGMR	6.80				1.21				−0.89				−0.47			

changed positive skewness and kurtosis. For the Holocene, however, annual higher-order moments are generally smaller in comparison to the LGM. Towards longer timescales, standard deviation decreases significantly and shows very little state dependency. For skewness and kurtosis, the tropical peaks diminish, but, instead, areas of larger skew or kurtosis emerge in the mid- and high latitudes, especially during the Deglaciation. So, while annual distributions show that there are extreme-precipitation years in the tropics, decadal and centennial distributions demonstrate that such extreme conditions rarely persist for whole decades or centuries. For skewness, dry regions are generally associated with positive skew, since only a high-value tail can exist for low mean precipitation. This is why precipitation mostly has positive skewness.

Some GCM and ESM runs show bipolar skewness patterns between the hemispheres but disagree on whether the negative skewness is in the Northern Hemisphere (MPI-ESM, HadCM3B) or Southern Hemisphere (TraCE-21ka). These bipolar patterns appear in response to meltwater forcing (see Sect. 5.3). The spectra show little state dependency on annual to multi-decadal timescales (Fig. 11i, j). For centennial timescales and longer, the Deglaciation shows a strong increase in variability, strongest in the tropics (Figs. S23 and S24), setting it apart from the LGM and Holocene. The LGM and Holocene differ to a lesser degree but start diverging on centennial timescales, although the simulations disagree whether the LGM or the Holocene has stronger spectral power.

5.3 Dependence of surface climate variability on external forcings

While ice sheet changes, meltwater fluxes and volcanism all have characteristic timescales, differences in these forcings cause variability changes on other timescales as well. As such, forcings, through their non-linear interactions with faster components of the climate system, impact variability on timescales shorter than their characteristic timescales. Conversely, they also impact slower components and thus variability on longer timescales. This implies that, even when investigating variability on short, e.g., interannual, timescales, the initial state of slower components like the oceans can affect simulated variability.

For ice sheet reconstructions, we compare ICE6G with GLAC1-D, which has a more extensive but lower glacial ice cover. The reconstructions have temporal resolutions of 500 and 100 years, respectively. While the reconstructions were interpolated for MPI-ESM r1–7, this difference in the underlying timescale will still affect centennial variability and likely explains some of the increased variability found for simulations using GLAC1-D in comparison to those using ICE6G (see Fig. 6). For the northern latitudes, the comparison indeed reveals a general association of GLAC1-D with larger standard deviation during the LGM and Deglaciation (Fig. 6). On the other hand, GLAC1-D is associated with reduced standard deviation across timescales in parts of the Southern Ocean during the Deglaciation and Holocene related to less variance in sea ice (Fig. S26). Towards longer timescales, the simulations with GLAC1-D are both more positively and negatively skewed, in particular for the Deglaciation. The chosen ice sheet reconstruction thus significantly impacts variability, especially for temperature, and behavior at the tails of the distributions on annual to millennial timescales. However, there is considerable uncertainty in ice sheet reconstructions and corresponding meltwater releases (Stokes et al., 2015; Abe-Ouchi et al., 2015; Ivanovic et al., 2016), an uncertainty that simulated variability thus retains. Many of the differences between simulations

using GLAC1-D and ICE6G can be attributed to their associated meltwater protocols.

A bipolar skewness pattern indicates a meltwater pulse in the area of negative skewness. GLAC1-D introduces meltwater pulses, in particular MWP-1A and MWP-1B, mainly in the Northern Hemisphere. These lead to a freshening of North Atlantic surface water and a slowdown of the Atlantic Meridional Overturning Circulation. As a result, the Northern Hemisphere cools and experiences more cold outliers (negative skewness), while the Southern Hemisphere warms. In ICE6G, on the other hand, MWP-1B is mainly released into the Southern Ocean, counteracting and reducing the dominating bipolar pattern. In TraCE-21ka, the most abrupt freshwater pulse (MWP-1A) is mostly imposed as a freshwater flux into the Southern Ocean and, as such, leads to an opposite pattern in skewness. Kurtosis, similarly, is stronger in GLAC1-D simulations during the LGM and Deglaciation, an effect that grows towards longer timescales. There is a general ranking of simulations by meltwater protocol with variability increasing from no-melt to melt-uniform and finally melt-routed. Since simulations are generally believed to lack variability at least regionally (e.g., Laepple et al., 2023; Weitzel et al., 2024; Rehfeld et al., 2018), this supports the usage of the melt-routed protocol. Since meltwater forcing has such a strong association with variability and the overall deglacial climate evolution (Snoll et al., 2024), variability could help constrain meltwater releases. However, this would require identifying models with high skill regarding simulated variability, currently hindered by large uncertainties in reconstructed variability. Furthermore, tuning meltwater to reproduce reconstructed variability alone is likely too simplified an approach (Weitzel et al., 2024).

Volcanism has the largest impact on annual scales, as it introduces short-term cooling events. However, its impacts are evident on longer timescales as well through non-linear interactions with other components of the climate system, as also noted by, for example, Ellerhoff et al. (2022). This is particularly apparent in the spectra, where volcanism raises the power across and especially on decadal to centennial timescales (Fig. 11). Volcanic forcing mostly increases standard deviation and reduces positive temperature skew across timescales and kurtosis on annual scales as high-temperature outliers become less likely and low-temperature outliers become more likely in response to the negative radiative forcing imposed.

5.4 Dependency of surface climate on model complexity suggests necessary minimal complexity

Our results suggest that there is a required minimal complexity for modeling the variability in surface climate. The EBM only reflects the linear response and does not possess the complexity required to capture changes beyond the mean, while simple models dedicated to variability might (see Lovejoy et al., 2021; Schillinger et al., 2022). The increasing resemblance of the EBM's moments to those of the more complex models towards longer timescales (Fig. S5), on the other hand, suggests that, at the global scale, centennial moments are dominated by the linear response to external forcings. This does not account for spatial patterns, though, which the EBM fails to capture.

The EMIC simulations provide a far better approximation of standard deviation of temperature but fall off for the extreme tails and for precipitation as a whole. The latter suggests that the EMICs lack variability in atmospheric dynamics. The spectral power of the EMICs is almost always on the lower end of the ensemble, further indicating a lack of energy transfer between scales. As such, while they can match variability in proxy reconstructions on the global level (cf. Zhu et al., 2019), they are limited for studies of regional variability. The EMICs included here have reduced atmospheric complexity. This will affect simulated variability and could be different in other EMICs as it is for GCMs.

Among GCMs and ESMs, providing a ranking of simulated variability proves difficult due to the sparse spatial coverage of reconstructions and the lack of literature studying variability during the Deglaciation, although first attempts have been made (Weitzel et al., 2024). Moreover, at the level of complexity of GCMs and ESMs, chosen forcings cause differences between simulations at least as much as the chosen model and its complexity. Due to substantial differences in forcings and boundary conditions inherently arising from different research foci, it can be hard to identify the sources of differences between simulations from different models of similar complexity. While there are some common patterns that emerge, the simulations also disagree in many areas with respect to magnitude and even direction of changes. This variety in the ensemble can be hidden when considering multimodel means. An experimental design geared towards understanding the roles of feedbacks on surface climate variability must take into account external forcing and boundary condition changes, distinguishing interactive effects and prescribed changes in boundary conditions which may, or may not, be physically consistent with the climate evolution. Given the impact of meltwater forcing and its uncertainties, simulations with interactive ice sheets are of particular interest to the study of climate extremes in response to mean changes.

6 Conclusions

The variability in surface climate has undergone considerable changes since the LGM, along with an increase in GMST. Here, we investigate changes in several indicators of variability from the LGM to the Holocene. These include standard deviation and power spectra but also the higher-order moments of skewness and kurtosis, which have, to our knowledge, previously been used only in studies of present-day and future climate. The warming ranges from 3.0–6.6 °C in

the ensemble of 15 transient simulations from models of varying complexity presented here. This result agrees with the estimates found in reconstructions, which also cover a large range. On the whole, we can confirm our hypotheses as (1) the variability between LGM and Holocene changes with the mean background state for both surface temperature and precipitation, and the variability during the Deglaciation is generally larger in comparison. Furthermore, we find that (2) state dependency increases from annual to millennial scales and that the forcings impact variability on all, not just their characteristic timescales, and that (3) there is a minimal complexity needed to simulate adequate levels of variability. In particular, we find that the variability of surface temperature and precipitation depend on the following:

1. Background state. Overall, state dependency of variability increases towards longer timescales. Standard deviation of surface temperature is larger during the LGM than the Holocene, whereas it is the opposite for precipitation. The LGM has little overall temperature skewness and kurtosis, whereas there can be areas of large skewness and kurtosis during the Holocene. For precipitation during the LGM and Holocene, state dependency can be found to some degree on all timescales and for all moments, but no clear patterns emerge. Beginning on decadal scales, the Deglaciation, as the transition between a cold and warmer interglacial climate state, stands out as a period of enhanced variability in comparison to the LGM and Holocene. This is marked by increased variance, more skew (both positive and negative), and larger kurtosis in most simulations.

2. Timescale. Towards longer timescales, the distributions of temperature show a decrease in variance, more absolute skew, and an increase in kurtosis, i.e., enhanced non-Gaussianity. For precipitation, variance decreases, too, while changes to the higher-order moments are mainly limited to the Deglaciation and differ by simulation and region. Volcanism, meltwater forcing, and ice sheet changes affect the distributions of surface temperature and precipitation the most on their characteristic timescales, but their effects can be detected from annual to millennial timescales.

3. Model complexity. Both the EBM and EMICs fail to reproduce patterns of variability found in the (sparse) reconstructions and do not resemble those in the more complex models. In our ensemble, the complexity of a coupled ocean–atmosphere GCM is at least necessary to produce realistic variability patterns. For the GCM and ESM simulations, on the other hand, differences in forcing protocols and boundary conditions play a larger role than model complexity in determining variability in surface temperature and precipitation.

To reach levels of variability comparable to reconstructions, simulations depend upon adequate levels of externally forced variability, including from volcanic forcing, and internal variability. The contribution of internal variability requires at least the minimum complexity of the GCMs in this study. Nevertheless, comparison to some reconstructions of past climate suggests that simulations might lack variability, especially regionally and from multi-decadal timescales onward. While the LGMR provided a first point of comparison, the comparison also raises questions about a potential loss of variability in field reconstruction methods when contrasted to other reconstructions. However, further conclusions necessitate an in-depth model–data comparison, which is limited by the small number of available proxy records, especially for precipitation. Since an improved understanding of the variability in surface climate is crucial, for example, because of its relation to extremes and freshwater availability, a comparison to 20th-century observations could provide clearer evidence on the models' abilities, at least for short timescales. This could complement model evaluation efforts like CMIP and PMIP.

Several factors have emerged that could advance the evaluation of simulated variability and allow a ranking of models: improved benchmarks from reconstructions, more simulations with consistent protocols, ensemble of varying initial conditions, and high-resolution snapshot simulations of past climates. The latter would allow an analysis of the changes in frequencies of extremes against regional mean changes and how extremes on short timescales transfer to longer timescales. As such, it could elucidate the role of higher-order moments in the evaluation of model simulations going forward.

Our results demonstrate that the Deglaciation stands out in comparison to the LGM and Holocene as a period of warming. However, an open question remains to what degree our results hold for future warming. To resolve this question, it is necessary to understand how much the increased variability is related to overall warming versus the initial state with large Northern Hemisphere ice sheets. Interactions between dynamics, forcings, and mean state lead to complex changes in the distributions of surface climate variables. This implies potential changes to extremes on timescales from years to centuries, requiring further investigation.

Code and data availability. The code and data to reproduce the analysis of this paper are available on Zenodo at https://doi.org/10.5281/zenodo.14550167 (Ziegler et al., 2024) under a CC-BY-SA 4.0 license.

Supplement. The supplement related to this article is available online at https://doi.org/10.5194/cp-21-627-2025-supplement.

Author contributions. EZ and KR conceptualized this study and developed the methodology with input from NW. EZ implemented

the analysis and visualized the results using simulations from MK, UM, LG, RI, PJV, and CW. EZ wrote the article with input from KR, NW, JPB, MK, and RI. All authors reviewed and discussed the article.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We thank Thomas Kleinen for discussion of the analysis and his simulation (MPI-ESM ch4). Furthermore, we thank Moritz Adam for his feedback on a previous version of this article. We want to express our gratitude to the two anonymous referees and the editor, Hugues Goosse, for their valuable feedback.

Kira Rehfeld is a member of the Machine Learning Cluster of Excellence, funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) under Germany's Excellence Strategy – EXC no. 2064/1 – grant no. 390727645.

Financial support. This research has been supported by the Deutsche Forschungsgemeinschaft (grant no. 395588486) and the Bundesministerium für Bildung und Forschung through the PalMod project (grant nos. 01LP1926C, 01LP1917B, 01LP2310A, 01LP2311C, and 01LP2302A).

This open-access publication was funded
by the Open Access Publication Fund of the University of Tübingen.

Review statement. This paper was edited by Hugues Goosse and reviewed by two anonymous referees.

References

Abe-Ouchi, A., Saito, F., Kageyama, M., Braconnot, P., Harrison, S. P., Lambeck, K., Otto-Bliesner, B. L., Peltier, W. R., Tarasov, L., Peterschmitt, J.-Y., and Takahashi, K.: Ice-sheet configuration in the CMIP5/PMIP3 Last Glacial Maximum experiments, Geosci. Model Dev., 8, 3621–3637, https://doi.org/10.5194/gmd-8-3621-2015, 2015.

Alley, R. B.: The Younger Dryas Cold Interval as Viewed from Central Greenland, Quaternary Sci. Rev., 19, 213–226, https://doi.org/10.1016/S0277-3791(99)00062-1, 2000.

Andersen, K. K., Azuma, N., Barnola, J. M., Bigler, M., Biscaye, P., Caillon, N., Chappellaz, J., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Flückiger, J., Fritzsche, D., Fujii, Y., Goto-Azuma, K., Grønvold, K., Gundestrup, N. S., Hansson, M., Huber, C., Hvidberg, C. S., Johnsen, S. J., Jonsell, U., Jouzel, J., Kipfstuhl, S., Landais, A., Leuenberger, M., Lorrain, R., Masson-Delmotte, V., Miller, H., Motoyama, H., Narita, H., Popp, T., Rasmussen, S. O., Raynaud, D., Rothlisberger, R., Ruth, U., Samyn, D., Schwander, J., Shoji, H., Siggard-Andersen, M. L., Steffensen, J. P., Stocker, T., Sveinbjörnsdóttir, A. E., Svensson, A., Takata, M., Tison, J. L., Thorsteinsson, T., Watanabe, O., Wilhelms, F., and White, J. W.: High-Resolution Record of Northern Hemisphere Climate Extending into the Last Interglacial Period, Nature, 431, 147–151, https://doi.org/10.1038/nature02805, 2004.

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

Anscombe, F. J. and Glynn, W. J.: Distribution of the Kurtosis Statistic B2 for Normal Samples, Biometrika, 70, 227–234, https://doi.org/10.1093/biomet/70.1.227, 1983.

Bakker, P., Rogozhina, I., Merkel, U., and Prange, M.: Hypersensitivity of glacial summer temperatures in Siberia, Clim. Past, 16, 371–386, https://doi.org/10.5194/cp-16-371-2020, 2020.

Baldini, J. U. L., Brown, R. J., and McElwaine, J. N.: Was Millennial Scale Climate Change during the Last Glacial Triggered by Explosive Volcanism?, Sci. Rep., 5, 17442, https://doi.org/10.1038/srep17442, 2015.

Bathiany, S., Dakos, V., Scheffer, M., and Lenton, T. M.: Climate Models Predict Increasing Temperature Variability in Poor Countries, Sci. Adv., 4, eaar5809, https://doi.org/10.1126/sciadv.aar5809, 2018.

Berg, A., Lintner, B. R., Findell, K. L., Malyshev, S., Loikith, P. C., and Gentine, P.: Impact of Soil Moisture–Atmosphere Interactions on Surface Temperature Distribution, J. Climate, 27, 7976–7993, https://doi.org/10.1175/JCLI-D-13-00591.1, 2014.

Bethke, I., Outten, S., Otterå, O. H., Hawkins, E., Wagner, S., Sigl, M., and Thorne, P.: Potential Volcanic Impacts on Future Climate Variability, Nat. Clim. Change, 7, 799–805, https://doi.org/10.1038/nclimate3394, 2017.

Brady, E., Stevenson, S., Bailey, D., Liu, Z., Noone, D., Nusbaumer, J., Otto-Bliesner, B. L., Tabor, C., Tomas, R., Wong, T., Zhang, J., and Zhu, J.: The Connected Isotopic Water Cycle in the Community Earth System Model Version 1, J. Adv. Model. Earth Sy., 11, 2547–2566, https://doi.org/10.1029/2019MS001663, 2019.

Briggs, R. D., Pollard, D., and Tarasov, L.: A Data-Constrained Large Ensemble Analysis of Antarctic Evolution since the Eemian, Quaternary Sci. Rev., 103, 91–115, https://doi.org/10.1016/j.quascirev.2014.09.003, 2014.

Campin, J.-M. and Goosse, H.: Parameterization of Density-Driven Downsloping Flow for a Coarse-Resolution Ocean Model in z-Coordinate, Tellus A, 51, 412–430, https://doi.org/10.3402/tellusa.v51i3.13468, 1999.

Chatfield, C.: The Analysis of Time Series: An Introduction, Sixth Edition, CRC Press, ISBN 978-0-203-49168-3, 2016.

Christiansen, B. and Ljungqvist, F. C.: Challenges and Perspectives for Large-Scale Temperature Reconstructions of the Past Two Millennia, Rev. Geophys., 55, 40–96, https://doi.org/10.1002/2016RG000521, 2017.

Clark, P. U., Shakun, J. D., Baker, P. A., Bartlein, P. J., Brewer, S., Brook, E., Carlson, A. E., Cheng, H., Kaufman, D. S., Liu, Z., Marchitto, T. M., Mix, A. C., Morrill, C., Otto-

Bliesner, B. L., Pahnke, K., Russell, J. M., Whitlock, C., Adkins, J. F., Blois, J. L., Clark, J., Colman, S. M., Curry, W. B., Flower, B. P., He, F., Johnson, T. C., Lynch-Stieglitz, J., Markgraf, V., McManus, J., Mitrovica, J. X., Moreno, P. I., and Williams, J. W.: Global Climate Evolution during the Last Deglaciation, P. Natl. Acad. Sci. USA, 109, E1134–E1142, https://doi.org/10.1073/pnas.1116619109, 2012.

Collins, W. D., Bitz, C. M., Blackmon, M. L., Bonan, G. B., Bretherton, C. S., Carton, J. A., Chang, P., Doney, S. C., Hack, J. J., Henderson, T. B., Kiehl, J. T., Large, W. G., McKenna, D. S., Santer, B. D., and Smith, R. D.: The Community Climate System Model Version 3 (CCSM3), J. Climate, 19, 2122–2143, https://doi.org/10.1175/JCLI3761.1, 2006.

Collow, T. W., Wang, W., and Kumar, A.: Reduction in Northern Midlatitude 2-m Temperature Variability Due to Arctic Sea Ice Loss, J. Climate, 32, 5021–5035, https://doi.org/10.1175/JCLI-D-18-0692.1, 2019.

Director, H. and Bornn, L.: Connecting Point-Level and Gridded Moments in the Analysis of Climate Data, J. Climate, 28, 3496–3510, https://doi.org/10.1175/JCLI-D-14-00571.1, 2015.

Ditlevsen, P., Mitsui, T., and Crucifix, M.: Crossover and Peaks in the Pleistocene Climate Spectrum; Understanding from Simple Ice Age Models, Clim. Dynam., 54, 1801–1818, https://doi.org/10.1007/s00382-019-05087-3, 2020.

Doane, D. P. and Seward, L. E.: Measuring Skewness: A Forgotten Statistic?, Journal of Statistics Education, 19, 3, https://doi.org/10.1080/10691898.2011.11889611, 2011.

Douville, H., Colin, J., Krug, E., Cattiaux, J., and Thao, S.: Midlatitude Daily Summer Temperatures Reshaped by Soil Moisture under Climate Change, Geophys. Res. Lett., 43, 812–818, https://doi.org/10.1002/2015GL066222, 2016.

Ellerhoff, B. and Rehfeld, K.: Probing the Timescale Dependency of Local and Global Variations in Surface Air Temperature from Climate Simulations and Reconstructions of the Last Millennia, Phys. Rev. E, 104, 1–14, https://doi.org/10.1103/PhysRevE.104.064136, 2021.

Ellerhoff, B., Kirschner, M. J., Ziegler, E., Holloway, M. D., Sime, L., and Rehfeld, K.: Contrasting State-Dependent Effects of Natural Forcing on Global and Local Climate Variability, Geophys. Res. Lett., 49, e2022GL098335, https://doi.org/10.1029/2022GL098335, 2022.

Evensen, G., Vossepoel, F. C., and van Leeuwen, P. J.: Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem, Springer Textbooks in Earth Sciences, Geography and Environment, Springer International Publishing, Cham, ISBN 978-3-030-96708-6 978-3-030-96709-3, https://doi.org/10.1007/978-3-030-96709-3, 2022.

Filliben, J. J. and Heckert, A.: Exploratory Data Analysis, in: NIST/SEMATECH e-Handbook of Statistical Methods NIST/SEMATECH, https://doi.org/10.18434/M32189, 2024.

Franzke, C. L., Barbosa, S., Blender, R., Fredriksen, H. B., Laepple, T., Lambert, F., Nilsen, T., Rypdal, K., Rypdal, M., Scotto, M. G., Vannitsem, S., Watkins, N. W., Yang, L., and Yuan, N.: The Structure of Climate Variability Across Scales, Rev. Geophys., 58, 2, https://doi.org/10.1029/2019RG000657, 2020.

Fredriksen, H. B. and Rypdal, M.: Long-Range Persistence in Global Surface Temperatures Explained by Linear Multibox Energy Balance Models, J. Climate, 30, 7157–7168, https://doi.org/10.1175/JCLI-D-16-0877.1, 2017.

Garfinkel, C. I. and Harnik, N.: The Non-Gaussianity and Spatial Asymmetry of Temperature Extremes Relative to the Storm Track: The Role of Horizontal Advection, J. Climate, 30, 445–464, https://doi.org/10.1175/JCLI-D-15-0806.1, 2017.

Goosse, H. and Fichefet, T.: Importance of Ice-Ocean Interactions for the Global Ocean Circulation: A Model Study, J. Geophys. Res.-Oceans, 104, 23337–23355, https://doi.org/10.1029/1999JC900215, 1999.

Goosse, H., Deleersnijder, E., Fichefet, T., and England, M. H.: Sensitivity of a Global Coupled Ocean-Sea Ice Model to the Parameterization of Vertical Mixing, J. Geophys. Res.-Oceans, 104, 13681–13695, https://doi.org/10.1029/1999JC900099, 1999.

Grant, K. M., Rohling, E. J., Bar-Matthews, M., Ayalon, A., Medina-Elizalde, M., Ramsey, C. B., Satow, C., and Roberts, A. P.: Rapid Coupling between Ice Volume and Polar Temperature over the Past 150,000 Years, Nature, 491, 744–747, https://doi.org/10.1038/nature11593, 2012.

Gulev, S., Thorne, P., Ahn, J., Dentener, F., Domingues, C., Gerland, S., Gong, D., Kaufman, D., Nnamchi, H., Quaas, J., Rivera, J., Sathyendranath, S., Smith, S., Trewin, B., von Schuckmann, K., and Vose, R.: Changing State of the Climate System, 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., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 287–422, https://doi.org/10.1017/9781009157896.004, 2021.

Guttal, V. and Jayaprakash, C.: Changing Skewness: An Early Warning Signal of Regime Shifts in Ecosystems, Ecol. Lett., 11, 450–460, https://doi.org/10.1111/j.1461-0248.2008.01160.x, 2008.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M.: A European Daily High-Resolution Gridded Data Set of Surface Temperature and Precipitation for 1950–2006, J. Geophys. Res.-Atmos., 113, 16, https://doi.org/10.1029/2008JD010201, 2008.

He, F.: Simulating Transient Climate Evolution of the Last Deglaciation with CCSM3, PhD Thesis, University of Wisconsin Madison, Madison, WC, USA, 171 pp., https://trace-21k.nelson.wisc.edu/Publications/He_PhD_dissertation_UW_2011.pdf (last access: 28 February 2025), 2011.

He, W., Wan, S., Jiang, Y., Jin, H., Zhang, W., Wu, Q., and He, T.: Detecting Abrupt Change on the Basis of Skewness: Numerical Tests and Applications, Int. J. Climatol., 33, 2713–2727, https://doi.org/10.1002/joc.3624, 2013.

Hegerl, G. C., Crowley, T. J., Baum, S. K., Kim, K.-Y., and Hyde, W. T.: Detection of Volcanic, Solar and Greenhouse Gas Signals in Paleo-Reconstructions of Northern Hemispheric Temperature, Geophys. Res. Lett., 30, 52, https://doi.org/10.1029/2002GL016635, 2003.

Huntingford, C., Jones, P. D., Livina, V. N., Lenton, T. M., and Cox, P. M.: No Increase in Global Temperature Variability despite Changing Regional Patterns, Nature, 500, 327–330, https://doi.org/10.1038/nature12310, 2013.

Huybers, P. and Curry, W.: Links between Annual, Milankovitch and Continuum Temperature Variability, Nature, 441, 329–332, https://doi.org/10.1038/nature04745, 2006.

Huybers, P. and Eisenman, I.: Integrated Summer Insolation Calculations, IGBP PAGES/World Data Center for Paleoclimatology Data NOAA/NCDC Paleoclimatology Program, Boulder CO, USA, Data Contribution Series no. 2006-079, https://www.ncei.noaa.gov/pub/data/paleo/climate_forcing/orbital_variations/huybers2006insolation/huybers2006b.txt (last access: 28 February 2025), 2006.

Iles, C. E. and Hegerl, G. C.: Systematic Change in Global Patterns of Streamflow Following Volcanic Eruptions, Nat. Geosci., 8, 838–842, https://doi.org/10.1038/ngeo2545, 2015.

Ionita, M., Dima, M., Nagavciuc, V., Scholz, P., and Lohmann, G.: Past Megadroughts in Central Europe Were Longer, More Severe and Less Warm than Modern Droughts, Commun. Earth Environ., 2, 1–9, https://doi.org/10.1038/s43247-021-00130-w, 2021.

IPCC: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/9781009157896, 2021.

Ivanovic, R. F., Gregoire, L. J., Kageyama, M., Roche, D. M., Valdes, P. J., Burke, A., Drummond, R., Peltier, W. R., and Tarasov, L.: Transient climate simulations of the deglaciation 21–9 thousand years before present (version 1) – PMIP4 Core experiment design and boundary conditions, Geosci. Model Dev., 9, 2563–2587, https://doi.org/10.5194/gmd-9-2563-2016, 2016.

Izumi, K., Valdes, P., Ivanovic, R., and Gregoire, L.: Impacts of the PMIP4 Ice Sheets on Northern Hemisphere Climate during the Last Glacial Period, Clim. Dynam., 60, 2481–2499, https://doi.org/10.1007/s00382-022-06456-1, 2023.

Jonkers, L., Cartapanis, O., Langner, M., McKay, N., Mulitza, S., Strack, A., and Kucera, M.: Integrating palaeoclimate time series with rich metadata for uncertainty modelling: strategy and documentation of the PalMod 130k marine palaeoclimate data synthesis, Earth Syst. Sci. Data, 12, 1053–1081, https://doi.org/10.5194/essd-12-1053-2020, 2020.

Joos, F. and Spahni, R.: Rates of Change in Natural and Anthropogenic Radiative Forcing over the Past 20,000 Years, P. Natl. Acad. Sci. USA, 105, 1425–1430, https://doi.org/10.1073/pnas.0707386105, 2008.

Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., Raynaud, D., Schilt, a., Schwander, J., Selmo, E., Souchez, R., Spahni, R., Stauffer, B., Steffensen, J. P., Stenni, B., Stocker, T. F., Tison, J. L., Werner, M., and Wolff, E. W.: Orbital and Millennial Antarctic Climate Variability over the Past 800,000 Years, Science, 317, 793–796, https://doi.org/10.1126/science.1141038, 2007.

Kapsch, M.-L., Mikolajewicz, U., Ziemen, F. A., Rodehacke, C. B., and Schannwell, C.: Analysis of the surface mass balance for deglacial climate simulations, The Cryosphere, 15, 1131–1156, https://doi.org/10.5194/tc-15-1131-2021, 2021.

Kapsch, M. L., Mikolajewicz, U., Ziemen, F., and Schannwell, C.: Ocean Response in Transient Simulations of the Last Deglaciation Dominated by Underlying Ice-Sheet Reconstruction and Method of Meltwater Distribution, Geophys. Res. Lett., 49, 1–11, https://doi.org/10.1029/2021GL096767, 2022.

Katz, R. W. and Brown, B. G.: Extreme Events in Changing Climate Variability Is More Important than Average, Clim. Change, 21, 289–302, https://doi.org/10.1007/BF00139728, 1992.

Kleinen, T., Mikolajewicz, U., and Brovkin, V.: Terrestrial methane emissions from the Last Glacial Maximum to the preindustrial period, Clim. Past, 16, 575–595, https://doi.org/10.5194/cp-16-575-2020, 2020.

Kleinen, T., Gromov, S., Steil, B., and Brovkin, V.: Atmospheric methane since the last glacial maximum was driven by wetland sources, Clim. Past, 19, 1081–1099, https://doi.org/10.5194/cp-19-1081-2023, 2023.

Köhler, P., Nehrbass-Ahles, C., Schmitt, J., Stocker, T. F., and Fischer, H.: A 156 kyr smoothed history of the atmospheric greenhouse gases CO_2, CH_4, and N_2O and their radiative forcing, Earth Syst. Sci. Data, 9, 363–387, https://doi.org/10.5194/essd-9-363-2017, 2017.

Laepple, T., Ziegler, E., Weitzel, N., Hébert, R., Ellerhoff, B., Schoch, P., Martrat, B., Bothe, O., Moreno-Chamarro, E., Chevalier, M., Herbert, A., and Rehfeld, K.: Regional but Not Global Temperature Variability Underestimated by Climate Models at Supradecadal Timescales, Na. Geosci., 16, 958–966, https://doi.org/10.1038/s41561-023-01299-9, 2023.

Lambeck, K., Rouby, H., Purcell, A., Sun, Y., and Sambridge, M.: Sea Level and Global Ice Volumes from the Last Glacial Maximum to the Holocene, P. Natl. Acad. Sci. USA, 111, 15296–15303, https://doi.org/10.1073/pnas.1411762111, 2014.

Le, T., Sjolte, J., and Muscheler, R.: The Influence of External Forcing on Subdecadal Variability of Regional Surface Temperature in CMIP5 Simulations of the Last Millennium, J. Geophys. Res.-Atmos., 121, 1671–1682, https://doi.org/10.1002/2015JD024423, 2016.

Li, G., Harrison, S. P., Bartlein, P. J., Izumi, K., and Colin Prentice, I.: Precipitation Scaling with Temperature in Warm and Cold Climates: An Analysis of CMIP5 Simulations, Geophys. Res. Lett., 40, 4018–4024, https://doi.org/10.1002/grl.50730, 2013.

Liu, F., Chai, J., Wang, B., Liu, J., Zhang, X., and Wang, Z.: Global Monsoon Precipitation Responses to Large Volcanic Eruptions, Sci. Rep., 6, 24331, https://doi.org/10.1038/srep24331, 2016.

Loikith, P. C., Neelin, J. D., Meyerson, J., and Hunter, J. S.: Short Warm-Side Temperature Distribution Tails Drive Hot Spots of Warm Temperature Extreme Increases under Near-Future Warming, J. Climate, 31, 9469–9487, https://doi.org/10.1175/JCLI-D-17-0878.1, 2018.

Lovejoy, S.: A Voyage through Scales, a Missing Quadrillion and Why the Climate Is Not What You Expect, Clim. Dynam., 44, 3187–3210, https://doi.org/10.1007/s00382-014-2324-0, 2015.

Lovejoy, S. and Varotsos, C.: Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forcings, Earth Syst. Dynam., 7, 133–150, https://doi.org/10.5194/esd-7-133-2016, 2016.

Lovejoy, S., Procyk, R., Hébert, R., and Del Rio Amador, L.: The Fractional Energy Balance Equation, Q. J. Roy. Meteor. Soc., 147, 1964–1988, https://doi.org/10.1002/qj.4005, 2021.

Lüthi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fischer, H., Kawamura, K., and Stocker, T. F.: High-Resolution Carbon Dioxide Concen-

tration Record 650,000–800,000 Years before Present, Nature, 453, 379–382, https://doi.org/10.1038/nature06949, 2008.

Mandelbrot, B. B. and van Ness, J. W.: Fractional Brownian Motions, Fractional Noises and Applications, SIAM Rev., 10, 422–437, https://doi.org/10.1137/1010093, 1968.

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., Nam, C. C., Notz, D., Nyawira, S. 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. S., Tian, F., Voigt, A., Vrese, P., Wieners, K. H., 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, J. Adv. Model. Earth Sy., 11, 998–1038, https://doi.org/10.1029/2018MS001400, 2019.

McKinnon, K. A., Rhines, A., Tingley, M. P., and Huybers, P.: The Changing Shape of Northern Hemisphere Summer Temperature Distributions, J. Geophys. Res.-Atmos., 121, 8849–8868, https://doi.org/10.1002/2016JD025292, 2016.

McManus, J. F., Francois, R., Gherardi, J.-M., Keigwin, L. D., and Brown-Leger, S.: Collapse and Rapid Resumption of Atlantic Meridional Circulation Linked to Deglacial Climate Changes, Nature, 428, 834–837, https://doi.org/10.1038/nature02494, 2004.

Meccia, V. L. and Mikolajewicz, U.: Interactive ocean bathymetry and coastlines for simulating the last deglaciation with the Max Planck Institute Earth System Model (MPI-ESM-v1.2), Geosci. Model Dev., 11, 4677–4692, https://doi.org/10.5194/gmd-11-4677-2018, 2018.

Menviel, L., Timmermann, A., Timm, O. E., and Mouchet, A.: Deconstructing the Last Glacial Termination: The Role of Millennial and Orbital-Scale Forcings, Quaternary Sci. Rev., 30, 1155–1172, https://doi.org/10.1016/j.quascirev.2011.02.005, 2011.

Mikolajewicz, U., Ziemen, F., Cioni, G., Claussen, M., Fraedrich, K., Heidkamp, M., Hohenegger, C., Jimenez de la Cuesta, D., Kapsch, M.-L., Lemburg, A., Mauritsen, T., Meraner, K., Röber, N., Schmidt, H., Six, K. D., Stemmler, I., Tamarin-Brodsky, T., Winkler, A., Zhu, X., and Stevens, B.: The climate of a retrograde rotating Earth, Earth Syst. Dynam., 9, 1191–1215, https://doi.org/10.5194/esd-9-1191-2018, 2018.

Monnin, E., Indermühle, A., Dällenbach, A., Flückiger, J., Stauffer, B., Stocker, T. F., Raynaud, D., and Barnola, J.-M.: Atmospheric CO_2 Concentrations over the Last Glacial Termination, Science, 291, 112–114, https://doi.org/10.1126/science.291.5501.112, 2001.

Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones, P. D.: Quantifying Uncertainties in Global and Regional Temperature Change Using an Ensemble of Observational Estimates: The HadCRUT4 Data Set, J. Geophys. Res.-Atmos., 117, 1–22, https://doi.org/10.1029/2011JD017187, 2012.

Nilsen, T., Rypdal, K., and Fredriksen, H.-B.: Are there multiple scaling regimes in Holocene temperature records?, Earth Syst. Dynam., 7, 419–439, https://doi.org/10.5194/esd-7-419-2016, 2016.

Opsteegh, J. D., Haarsma, R. J., Selten, F. M., and Kattenberg, A.: ECBILT: A Dynamic Alternative to Mixed Boundary Conditions in Ocean Models, Tellus A, 50, 348–367, https://doi.org/10.1034/j.1600-0870.1998.t01-1-00007.x, 1998.

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.

Papoulis, A. and Pillai, S. U.: Probability, Random Variables, and Stochastic Processes, McGraw-Hill, ISBN 978-0-07-112256-6, 2002.

Parrenin, F., Barnola, J.-M., Beer, J., Blunier, T., Castellano, E., Chappellaz, J., Dreyfus, G., Fischer, H., Fujita, S., Jouzel, J., Kawamura, K., Lemieux-Dudon, B., Loulergue, L., Masson-Delmotte, V., Narcisi, B., Petit, J.-R., Raisbeck, G., Raynaud, D., Ruth, U., Schwander, J., Severi, M., Spahni, R., Steffensen, J. P., Svensson, A., Udisti, R., Waelbroeck, C., and Wolff, E.: The EDC3 chronology for the EPICA Dome C ice core, Clim. Past, 3, 485–497, https://doi.org/10.5194/cp-3-485-2007, 2007.

Peltier, W.: Global Glacial Isostasy and the Surface of the Ice-Age Earth: The ICE-5G (VM2) Model and GRACE, Annu. Rev. Earth Planet. Sci., 32, 111–149, https://doi.org/10.1146/annurev.earth.32.082503.144359, 2004.

Peltier, W. R., Argus, D. F., and Drummond, R.: Space Geodesy Constrains Ice Age Terminal Deglaciation: The Global ICE-6G_C (VM5a) Model, J. Geophys. Res.-Sol. Ea., 120, 450–487, https://doi.org/10.1002/2014JB011176, 2015.

Percival, D. B. and Walden, A. T.: Spectral Analysis for Physical Applications – Multitaper and Conventional Univariate Techniques, Cambridge University Press, ISBN: 9780511622762, https://doi.org/10.1017/CBO9780511622762, 1993.

Perron, M. and Sura, P.: Climatology of Non-Gaussian Atmospheric Statistics, J. Climate, 26, 1063–1083, https://doi.org/10.1175/JCLI-D-11-00504.1, 2013.

Rehfeld, K., Münch, T., Ho, S. L., and Laepple, T.: Global Patterns of Declining Temperature Variability from the Last Glacial Maximum to the Holocene, Nature, 554, 356–359, https://doi.org/10.1038/nature25454, 2018.

Rehfeld, K., Hébert, R., Lora, J. M., Lofverstrom, M., and Brierley, C. M.: Variability of surface climate in simulations of past and future, Earth Syst. Dynam., 11, 447–468, https://doi.org/10.5194/esd-11-447-2020, 2020.

Riddick, T., Brovkin, V., Hagemann, S., and Mikolajewicz, U.: Dynamic hydrological discharge modelling for coupled climate model simulations of the last glacial cycle: the MPI-DynamicHD model version 3.0, Geosci. Model Dev., 11, 4291–4316, https://doi.org/10.5194/gmd-11-4291-2018, 2018.

Roe, G. H. and Steig, E. J.: Characterization of Millennial-Scale Climate Variability, J. Climate, 17, 1929–1944, https://doi.org/10.1175/1520-0442(2004)017<1929:COMCV>2.0.CO;2, 2004.

Ruff, T. W. and Neelin, J. D.: Long Tails in Regional Surface Temperature Probability Distributions with Implications for Extremes under Global Warming, Geophys. Res. Lett., 39, 13, https://doi.org/10.1029/2011GL050610, 2012.

Rypdal, K., Østvand, L., and Rypdal, M.: Long-Range Memory in Earth's Surface Temperature on Time Scales from Months to Centuries, J. Geophys. Res.-Atmos., 118, 7046–7062, https://doi.org/10.1002/jgrd.50399, 2013.

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., and Appenzeller, C.: The Role of Increasing Temperature Variability in European Summer Heatwaves, Nature, 427, 332–336, https://doi.org/10.1038/nature02300, 2004.

Schillinger, M., Ellerhoff, B., Scheichl, R., and Rehfeld, K.: Separating Internal and Externally Forced Contributions to Global Temperature Variability Using a Bayesian Stochastic Energy Balance Framework, Chaos, 32, 113146, https://doi.org/10.1063/5.0106123, 2022.

Schindlbeck-Belo, J. C., Toohey, M., Jegen, M., Kutterolf, S., and Rehfeld, K.: PalVol v1: a proxy-based semi-stochastic ensemble reconstruction of volcanic stratospheric sulfur injection for the last glacial cycle (140 000–50 BP), Earth Syst. Sci. Data, 16, 1063–1081, https://doi.org/10.5194/essd-16-1063-2024, 2024.

Schurer, A. P., Hegerl, G. C., Mann, M. E., Tett, S. F., and Phipps, S. J.: Separating Forced from Chaotic Climate Variability over the Past Millennium, J. Climate, 26, 6954–6973, https://doi.org/10.1175/JCLI-D-12-00826.1, 2013.

Schurer, A. P., Tett, S. F. B., and Hegerl, G. C.: Small Influence of Solar Variability on Climate over the Past Millennium, Nat. Geosci., 7, 104–108, https://doi.org/10.1038/ngeo2040, 2014.

Screen, J. A.: Arctic Amplification Decreases Temperature Variance in Northern Mid- to High-Latitudes, Nat. Clim. Change, 4, 577–582, https://doi.org/10.1038/nclimate2268, 2014.

Screen, J. A. and Simmonds, I.: The Central Role of Diminishing Sea Ice in Recent Arctic Temperature Amplification, Nature, 464, 1334–1337, https://doi.org/10.1038/nature09051, 2010.

Shakun, J. D. and Carlson, A. E.: A Global Perspective on Last Glacial Maximum to Holocene Climate Change, Quaternary Sci. Rev., 29, 1801–1816, https://doi.org/10.1016/j.quascirev.2010.03.016, 2010.

Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., Otto-Bliesner, B., Schmittner, A., and Bard, E.: Global Warming Preceded by Increasing Carbon Dioxide Concentrations during the Last Deglaciation, Nature, 484, 49–54, https://doi.org/10.1038/nature10915, 2012.

Shi, J., Jiang, D., Tian, Z., and Lang, X.: Enhanced Interannual Variability in Temperature during the Last Glacial Maximum, J. Climate, 35, 5933–5950, https://doi.org/10.1175/JCLI-D-21-0739.1, 2022.

Sigl, M., Toohey, M., McConnell, J. R., Cole-Dai, J., and Severi, M.: Volcanic stratospheric sulfur injections and aerosol optical depth during the Holocene (past 11 500 years) from a bipolar ice-core array, Earth Syst. Sci. Data, 14, 3167–3196, https://doi.org/10.5194/essd-14-3167-2022, 2022.

Simolo, C. and Corti, S.: Quantifying the Role of Variability in Future Intensification of Heat Extremes, Nat. Commun., 13, 7930, https://doi.org/10.1038/s41467-022-35571-0, 2022.

Skelton, A., Kirchner, N., and Kockum, I.: Skewness of Temperature Data Implies an Abrupt Change in the Climate System Between 1985 and 1991, Geophys. Res. Lett., 47, 1–10, https://doi.org/10.1029/2020GL089794, 2020.

Smith, R. S. and Gregory, J.: The Last Glacial Cycle: Transient Simulations with an AOGCM, Clim. Dynam., 38, 1545–1559, https://doi.org/10.1007/s00382-011-1283-y, 2012.

Smith, R. S., Gregory, J. M., and Osprey, A.: A description of the FAMOUS (version XDBUA) climate model and control run, Geosci. Model Dev., 1, 53–68, https://doi.org/10.5194/gmd-1-53-2008, 2008.

Snoll, B., Ivanovic, R., Gregoire, L., Sherriff-Tadano, S., Menviel, L., Obase, T., Abe-Ouchi, A., Bouttes, N., He, C., He, F., Kapsch, M., Mikolajewicz, U., Muglia, J., and Valdes, P.: A multi-model assessment of the early last deglaciation (PMIP4 LDv1): a meltwater perspective, Clim. Past, 20, 789–815, https://doi.org/10.5194/cp-20-789-2024, 2024.

Spahni, R., Chappellaz, J., Stocker, T. F., Loulergue, L., Hausammann, G., Kawamura, K., Flückiger, J., Schwander, J., Raynaud, D., Masson-Delmotte, V., and Jouzel, J.: Atmospheric Methane and Nitrous Oxide of the Late Pleistocene from Antarctic Ice Cores, Science, 310, 1317–1321, https://doi.org/10.1126/science.1120132, 2005.

Steinhilber, F., Beer, J., and Fröhlich, C.: Total Solar Irradiance during the Holocene, Geophys. Res. Lett., 36, 1–5, https://doi.org/10.1029/2009GL040142, 2009.

Stokes, C. R., Tarasov, L., Blomdin, R., Cronin, T. M., Fisher, T. G., Gyllencreutz, R., Hättestrand, C., Heyman, J., Hindmarsh, R. C. A., Hughes, A. L. C., Jakobsson, M., Kirchner, N., Livingstone, S. J., Margold, M., Murton, J. B., Noormets, R., Peltier, W. R., Peteet, D. M., Piper, D. J. W., Preusser, F., Renssen, H., Roberts, D. H., Roche, D. M., Saint-Ange, F., Stroeven, A. P., and Teller, J. T.: On the Reconstruction of Palaeo-Ice Sheets: Recent Advances and Future Challenges, Quaternary Sci. Rev., 125, 15–49, https://doi.org/10.1016/j.quascirev.2015.07.016, 2015.

Tamarin-Brodsky, T., Hodges, K., Hoskins, B. J., and Shepherd, T. G.: A Simple Model for Interpreting Temperature Variability and Its Higher-Order Changes, J. Climate, 35, 387–403, https://doi.org/10.1175/JCLI-D-21-0310.1, 2022.

Tarasov, L., Dyke, A. S., Neal, R. M., and Peltier, W. R.: A Data-Calibrated Distribution of Deglacial Chronologies for the North American Ice Complex from Glaciological Modeling, Earth Planet. Sc. Lett., 315–316, 30–40, https://doi.org/10.1016/j.epsl.2011.09.010, 2012.

Thomson, D. J.: Spectrum Estimation and Harmonic Analysis, P. IEEE, 70, 1055–1096, https://doi.org/10.1109/PROC.1982.12433, 1982.

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.

Timm, O. and Timmermann, A.: Simulation of the Last 21 000 Years Using Accelerated Transient Boundary Conditions, J. Climate, 20, 4377–4401, https://doi.org/10.1175/JCLI4237.1, 2007.

Timmreck, C.: Modeling the Climatic Effects of Large Explosive Volcanic Eruptions, WIREs Climate Change, 3, 545–564, https://doi.org/10.1002/wcc.192, 2012.

Tingley, M. P., Craigmile, P. F., Haran, M., Li, B., Mannshardt, E., and Rajaratnam, B.: Piecing Together the Past: Statistical Insights into Paleoclimatic Reconstructions, Quaternary Sci. Rev., 35, 1–22, https://doi.org/10.1016/j.quascirev.2012.01.012, 2012.

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

Valdes, P. J., Armstrong, E., Badger, M. P. S., Bradshaw, C. D., Bragg, F., Crucifix, M., Davies-Barnard, T., Day, J. J., Farnsworth, A., Gordon, C., Hopcroft, P. O., Kennedy, A. T., Lord, N. S., Lunt, D. J., Marzocchi, A., Parry, L. M., Pope, V., Roberts, W. H. G., Stone, E. J., Tourte, G. J. L., and Williams, J. H. T.: The BRIDGE HadCM3 family of climate models: HadCM3@Bristol v1.0, Geosci. Model Dev., 10, 3715–3743, https://doi.org/10.5194/gmd-10-3715-2017, 2017.

Volodin, E. M. and Yurova, A. Y.: Summer Temperature Standard Deviation, Skewness and Strong Positive Temperature Anomalies in the Present Day Climate and under Global Warming Conditions, Clim. Dynam., 40, 1387–1398, https://doi.org/10.1007/s00382-012-1447-4, 2013.

von Storch, H. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press, ISBN 9780511612336, https://doi.org/10.1017/CBO9780511612336, 1999.

Weitzel, N., Andres, H., Baudouin, J.-P., Kapsch, M.-L., Mikolajewicz, U., Jonkers, L., Bothe, O., Ziegler, E., Kleinen, T., Paul, A., and Rehfeld, K.: Towards spatio-temporal comparison of simulated and reconstructed sea surface temperatures for the last deglaciation, Clim. Past, 20, 865–890, https://doi.org/10.5194/cp-20-865-2024, 2024.

Wunsch, C.: The Spectral Description of Climate Change Including the 100 Ky Energy, Clim. Dynam., 20, 353–363, https://doi.org/10.1007/s00382-002-0279-z, 2003.

Zanchettin, D., Bothe, O., Lehner, F., Ortega, P., Raible, C. C., and Swingedouw, D.: Reconciling reconstructed and simulated features of the winter Pacific/North American pattern in the early 19th century, Clim. Past, 11, 939–958, https://doi.org/10.5194/cp-11-939-2015, 2015.

Zhu, F., Emile-Geay, J., McKay, N. P., Hakim, G. J., Khider, D., Ault, T. R., Steig, E. J., Dee, S., and Kirchner, J. W.: Climate Models Can Correctly Simulate the Continuum of Global-Average Temperature Variability, P. Natl. Acad. Sci. USA, 116, 8728–8733, https://doi.org/10.1073/pnas.1809959116, 2019.

Zhuang, K., North, G. R., and Stevens, M. J.: A NetCDF Version of the Two-Dimensional Energy Balance Model Based on the Full Multigrid Algorithm, SoftwareX, 6, 198–202, https://doi.org/10.1016/j.softx.2017.07.003, 2017.

Ziegler, E. and Rehfeld, K.: TransEBM v. 1.0: description, tuning, and validation of a transient model of the Earth's energy balance in two dimensions, Geosci. Model Dev., 14, 2843–2866, https://doi.org/10.5194/gmd-14-2843-2021, 2021.

Ziegler, E., Weitzel, N., Baudouin, J.-P., Kapsch, M.-L., Mikolajewicz, U., Gregoire, L., Ivanovic, R., Valdes, P., Wirths, C., and Rehfeld, K.: Code and data as a supplement to "Patterns of changing surface climate variability from the Last Glacial Maximum to present in transient model simulations" Zenodo [code, data set], https://doi.org/10.5281/zenodo.14550167, 2024.

Supplement of Clim. Past, 21, 627–659, 2025
https://doi.org/10.5194/cp-21-627-2025-supplement

Climate of the Past Open Access EGU

Supplement of

Patterns of changing surface climate variability from the Last Glacial Maximum to present in transient model simulations

Elisa Ziegler et al.

Correspondence to: Elisa Ziegler (elisa.ziegler@uni-tuebingen.de) and Kira Rehfeld (kira.rehfeld@uni-tuebingen.de)

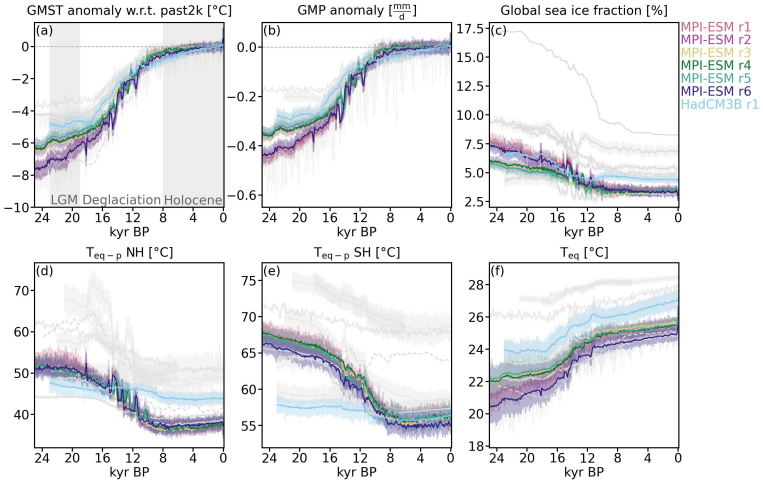


Figure S1. Centennial mean climate variables for the sensitivity set (MPI-ESM r1–r6 and HadCM3B r1) from LGM to Holocene with annual data as shading. The main set is shaded in gray and shown in Fig. 2. (a) GMT and (b) GMP anomaly with respect to the past $2\,\mathrm{kyr}$. (c) Global sea ice fraction. (d, e) Equator-to-pole temperature difference for the Northern and Southern Hemisphere respectively, computed as the difference between the polar ($70°$-$90°$) and equatorial temperatures ($15°$S-$15°$N). The latter are shown in (f).

S1 Main and sensitivity set of simulations

Fig. S1 shows the features of main climate of the sensitivity set analogous to Fig. 2 in the main text.

S2 Supplemental information on the simulations in the Deglacial ensemble

S2.1 The TransEBM Deglacial simulation

Using TransEBM (Ziegler and Rehfeld, 2021), we simulate the linear climate response to the forcings of the last 26k years. The first 25 years of the simulation serve as spin up. The model computes the orbital parameters following Berger (1978). Annual incoming solar radiation is constructed as a surrogate forcing based on a $9300\,\mathrm{yr}$ reconstruction of Holocene total solar irradiance by Steinhilber et al. (2009). From this reconstruction, we construct several artifical timeseries using the Amplitude Adjusted Fourier transform surrogate method with random seeds (Theiler et al., 1992) following Ellerhoff and Rehfeld (2021) to construct a forcing timeseries (Fig. S2). As greenhouse gas forcing, the simulation uses the reconstruction of CO_2 by Köhler et al. (2017). Like MPI-ESM r7, it employs the volcanic forcing PalVol V.1 (Schindlbeck-Belo et al., 2024). TransEBM employs it as an annual averaged volcanic forcing, which we regridded using linear interpolation and then applied to $2.8°$ wide latitudinal bands. Changes to the land sea mask and ice sheets follow ICE6G (Peltier et al., 2015). Since changes in the EBM produce an immediate climate response we interpolate the ice sheets during periods of rapid sea level rise to simulate a more gradual change in ice sheet cover. To this end, we identify areas of land ice, which change between consecutive reconstructions. In these places, ice is gradually removed with sea level change, starting from grid cells closer to equator. At the same latitude ice at places with smaller longitude is removed first. As a result, land ice cover changes in the simulation at intervals between 125 to $500\,\mathrm{yr}$. Sea ice is also modelled as a surface type with specific properties (heat capacity and albedo). As such, it serves as an external forcing in the simulation. We construct it by interpolating between LGM and present-day distributions by (Zhuang

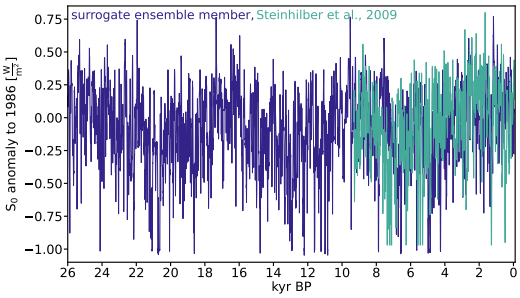


Figure S2. Comparison of the solar forcing reconstruction of Steinhilber et al. (2009) with one ensemble member generated as surrogate data based on it following (Ellerhoff and Rehfeld, 2021).

et al., 2017) with the same method as for land ice. To mimic freshwater induced variability in the North Atlantic heat transport, we modulate the North Atlantic heat diffusivity in the model to match SST reconstructions by (Jonkers et al., 2020).

S2.2 Axes of the model hierarchy

The hierarchy constructed in Sect. 2.1.1 and Fig. 3a includes eight axis: the resolution of atmosphere and ocean, the atmosphere and ocean model, the bands of long- and shortwave radiation, land hydrology, vegetation and aerosols. Resolution does not necessarily equal complexity. Increased resolution can however hold a larger potential for more complex simulated dynamics. For long transient simulations, computational demand limits resolution. Within those limits our analysis contains simulations with horizontal resolutions ranging from 48×37 and T21 to T42 in the atmosphere and from $3.75° \times 2.5°$ to $1.25° \times 1.25°$ in the ocean. Vertical layers in the atmosphere range from none in the case of TransEBM, 3 as the smallest amount of vertical layers in ECBilt to 31 in MPI-ESM. For the ocean, vertical layers are either again none for the EBM, 20 or 40. To summarize these differences in resolution for the hierarchy, atmospheric resolution is ranked by the amount of grid boxes computed as horizontal resolution times vertical resolution. For the ocean model resolution, horizontal and vertical resolution are ranked separately and then the average is taken for the hierarchy. This approach was chosen to be able to account for the variable latitudinal resolution that gets finer towards the equator in TraCE-21ka. Regarding the atmospheric resolution, the EBM and ECBilt make up the lowest tier, FAMOUS is ranked one above and the other GCMs and ESMs constitute the highest tier, mirroring the traditional hierarchy from EBM and EMIC to ESMs. For the oceanic resolution, the picture is more varied as CLIO has a higher resolution and complexity then the atmospheric model ECBilt it is coupled to. As a result, its resolution is larger than that of the ocean models in TraCE-21ka and FAMOUS. The resolution of HadCM3B is larger still and MPI-ESM has the largest oceanic resolution in found in this study. The EBM models the ocean as a surface type and thus ranks as zero.

Overall, the resolutions of all simulations remain limited so as to enable the transient modeling of the Last Deglaciation. However, how much such limited resolutions impacts the simulated climate and thus its study can be unclear. While an increase in resolution allows to better resolve topographic features and can lead to improved representation of regional phenomena and circulation patterns (Liu et al., 2022; Schiemann et al., 2014; Roeckner et al., 2006; Paolini et al., 2022; Kirtman et al., 2012), higher resolution does not always translate to large changes in the simulated climate. However, many of these studies see benefits when hitherto parameterized climatic features can be resolved, e.g. in convection- or eddy-resolving simulations. Otherwise, for many climatic features, the effect of increasing spatial resolution on the ability of simulations to reproduce observations can be smaller than expected. All in all, the resolutions of the simulations considered here thus allow for the study of global climate and its evolution since the LGM.

For the atmosphere model, the hierarchy considers the following options: 0D point model, 1D model (both not represented in this study), 2D model (TransEBM), 3D (quasi-)geostrophic (ECBilt), primitive equations (all others) and could further extend to modeling of non-hydrostatic equations and full Navier-Stokes equations. We further distinguish the radiative models

in the atmospheric model by the total number of long- and shortwave radiative bands, which range from $1+1$ (TransEBM, ECBilt), $6+8$ (FAMOUS, HadCM3B) and $19+8$ (TraCE-21ka) to $14+16$ (MPI-ESM). Regarding the ocean model, many options exist, of which the simulations included here only cover a few: prescribed surface type/ no ocean model (TransEBM), an ocean component with primitive equations with (FAMOUS, HadCM3B) and without (CLIO, TraCE-21ka, MPI-ESM) a rigid lid approximation. The simulations do not include options such as prescribed SSTs, a slab or geostrophic ocean, although they are included in the hierarchy in theory. The land hydrology options range from none (TransEBM), to a single layer with a bucket model (LOVECLIM simulations) and multi-layer models (FAMOUS, HadCM3B, TraCE-21ka), which can include a river-routing model (MPI-ESM). For vegetation, the options are prescribed (TransEBM, LOVECLIM simulations), models with a terrestrial carbon cycle (not included), dynamic vegetation models (FAMOUS, HadCM3B, TraCE-21ka) and a dynamic vegetation model that includes a nitrogen cycle (MPI-ESM). Lastly, the inclusion of aerosols varies from none at all (TransEBM, LOVECLIM simulations), to prescribed (FAMOUS, HadCM3B), modeling of sulfates (TraCE-21ka) and other aerosols in addition to sulfates, e.g. NH_3 (MPI-ESM).

S3 Supplements to the methods section

S3.1 Extended introduction to the moments of a probabililty distribution

The generalized moments of random variable X of a point A for a sample of size N are defined using the expected value as

$$\mu'_n = \mathbb{E}[(X-A)^n] = \frac{1}{N}\sum_{i=1}^{N}(X_i - A)^n, \tag{S1}$$

where n designates the nth moment (Papoulis and Pillai, 2002). There are several points of interest that can be substituted for A. One such choice is the origin, that is $A=0$, such that the nth moment around the origin (Papoulis and Pillai, 2002; von Storch and Zwiers, 1999), sometimes called nth raw moment, is

$$\mu_n = \mathbb{E}[X^n] = \frac{1}{N}\sum_{i=1}^{N}X_i^n. \tag{S2}$$

Thus, the first raw moment of a distribution is the arithmetic mean $\mu_1 \equiv \mu$ (Fig. 4a in the main text, Papoulis and Pillai, 2002).

Setting $A=\mu$ yields the nth central moment (Papoulis and Pillai, 2002; von Storch and Zwiers, 1999)

$$m_n = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^n. \tag{S3}$$

For $n=2$ this corresponds to the variance σ'^2. However, here, we use the symmetric unbiased estimator σ^2 (Filliben and Heckert, 2024) as introduced in the main text.

Moments of order higher than 2 are often normalized using the standard deviation σ' (Kenney and Keeping, 1948). The standardized third central moment, skewness s,

$$s = \frac{m_3}{m_2^{3/2}}\,. \tag{S4}$$

describes the (a)symmetry of a distribution and indicates whether the weight and tails of a distribution are lopsided (von Storch and Zwiers, 1999). This estimator is also called the Fisher-Pearson coefficient (Joanes and Gill, 1998; Filliben and Heckert, 2024). Lastly, we compute the standardized fourth central moment, kurtosis, but shifted by -3 to compute excess kurtosis k (Filliben and Heckert, 2024).

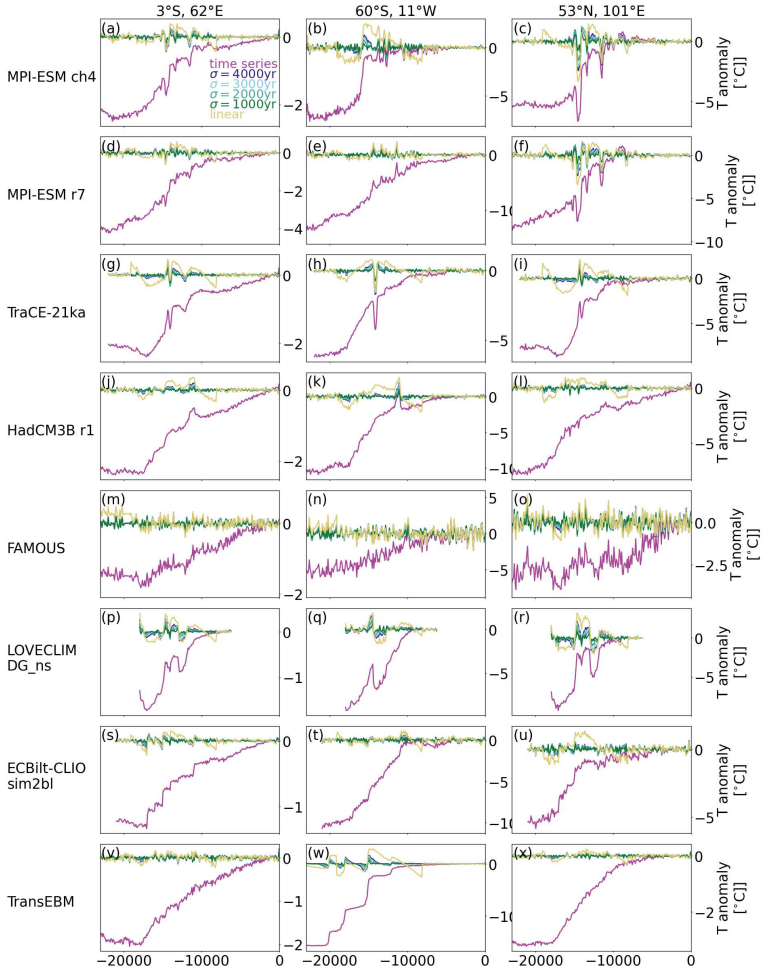


Figure S3. Effects of different versions of detrending on the resulting timeseries used for calculating the moments and spectra in the main text. Shown are different values for the kernel size σ of the Gaussian smoothing, as well as linear detrending with 2 breakpoints.

S3.2 Sensitivity test for different methods for detrending the timeseries

Figure S3 compares different versions of detrending. Among others, it shows the results when using linear detrending with breakpoints at the transitions from LGM to Deglaciation (19kyr BP) and Deglaciation to Holocene (8kyr BP). Different kernel sizes for the Gaussian smoothing are compared as well, from $\sigma = 1000\,\mathrm{yr}$ to the $4000\,\mathrm{yr}$ used for the main manuscript. Linear detrending is the only method that produces notably different results and clear artifacts, e.g. in panels g, k, u and w. While the different kernel sizes produce differences mostly during times of abrupt changes, which will be removed as trend rather than variability as the kernel size gets smaller and thus closer to the timescale of these changes.

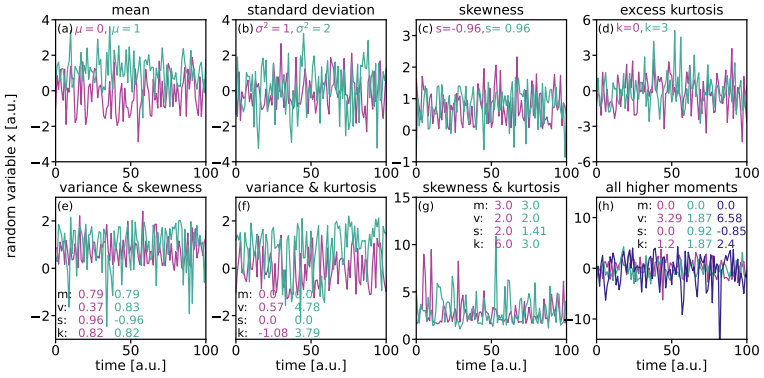


Figure S4. Exemplary timeseries for changes in the moments as sampled from the distributions as shown in Fig. 4. In the upper row only one moment of the corresponding distribution changes, in the bottom row several moments change.

Table S1. LGM to Holocene changes in the moments of surface temperature for all simulations and timescales. For the mean, the absolute value of the difference is listed, for the other moments the ratios.

	$\Delta m_{hol-lgm}$	v_{lgm}/v_{hol}			s_{lgm}/s_{hol}			k_{lgm}/k_{hol}		
	T	T_{ann}	T_{dec}	T_{cen}	T_{ann}	T_{dec}	T_{cen}	T_{ann}	T_{dec}	T_{cen}
MPI-ESM ch4	4.41	1.63	1.49	2.05	4.60	9.26	0.27	0.77	0.14	-0.01
MPI-ESM r1	6.51	1.97	1.69	2.38	-1.43	-0.09	-0.09	0.30	0.02	0.04
MPI-ESM r2	5.52	1.51	1.00	0.92	-1.39	-0.04	-0.04	0.49	0.02	-0.01
MPI-ESM r3	5.69			0.81			0.05			-0.01
MPI-ESM r4	5.57			1.07			0.09			-0.01
MPI-ESM r5	5.30	1.49	0.95	0.81	-1.03	-3e-3	0.03	0.35	0.02	-3e-3
MPI-ESM r6	6.48	1.89	1.47	1.44	-1.59	-0.08	0.01	0.30	0.02	0.03
MPI-ESM r7	6.62	2.03	1.97	2.69	1.27	1.52	1.65	0.84	1.35	-2.08
TraCE-21ka	4.11	0.85	0.74	0.63	1.84	-0.20	0.09	0.92	0.10	-0.16
HadCM3B r1	4.45	1.47	1.52	1.95	-7.69	10.61	-3.10	0.84	0.33	-1.20
HadCM3B r2	4.76	1.47	1.49	1.83	-1.86	-0.08	3.36	0.87	0.43	-0.59
FAMOUS	3.07			1.58			2.84			0.96
ECBilt-CLIO	3.95	1.16	1.03	0.74	0.87	0.76	0.84	0.38	0.09	-0.17
TransEBM	4.12	1.27	1.39	1.56	0.90	0.98	-0.01	0.60	0.60	-0.29

Figure S4 shows example timeseries as drawn from the distributions shown in Fig. 4. The timeseries thus demonstrate the potential effects of differences in one or several moments.

S4 Moments per simulations

S4.1 Global mean moments

Figure S5 shows the global mean moments from LGM to Holocene for surface temperature and precipitation by complexity category. Tables Table S1 and S2 list the LGM-to-Holocene changes in moments per simulations, which are summarized per complexity category in Table 2 of the main text.

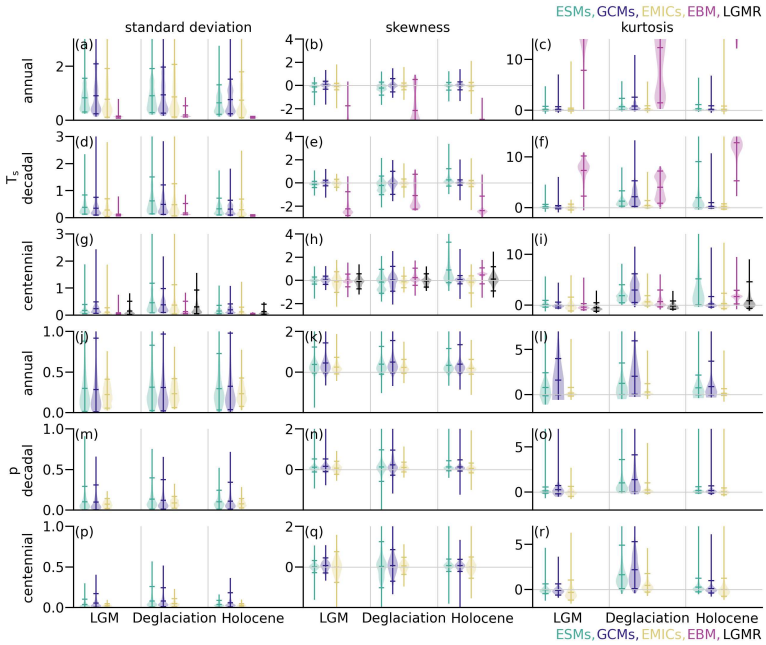


Figure S5. Global annual, decadal and centennial higher order moments of temperature (a-i) and precipitation (j-r). Simulations are summarized by their complexity into ESMs (green), GCMs (blue), EMICs (yellow) and EBM (red) and are shown with the means of the 95th percentiles in space of the individual models. For each moment, standard deviation (left column), skewness (middle column) and kurtosis (right column), results are shown separately for LGM, Deglaciation and Holocene with gray lines separating the periods. For (c) the 95th percentile of the EBM is cut-off for better comparability of the other moments.

S4.2 Zonal moments

Figures S6 and S7 complement Fig. 5 in the main text and show the latitudinal profiles of the moments for the individual simulations.

S4.3 Moments in space

To highlight the differences with timescale, we here show the moments in space according to timescale that are not included in the main text. Figure S8 first shows the mean changes of both surface temperature and precipitation for three examples simulations — MPI-ESM r7, TraCE-21ka and ECBilt-CLIO — and reveals the simulated global warming and wetting from LGM to Holocene across the ensemble. Figure S9 shows the spatial patterns of the higher order moments on centennial scales for the same example simulations for surface temperature. Figures S10 and S11 show the same for precipitation on centennial and annual timescales, respectively.

To complete the analysis of the effects of forcings on the distributions from the main text, Fig. S12 starts by showing the impacts of forcings on annual standard deviation of surface temperature. Further Fig. S13 and S14 display the same for decadal and annual temperature skewness, while Fig. Fig. S15 and S16 present forcing impacts on decadal and annual temperature kurtosis, respectively. For precipitation, we show annual standard deviation (Fig. S17) and skewness. Finally, Fig. S19 presents the spatial changes in mean, standard deviation, skewness and kurtosis for the LGMR.

Table S2. LGM to Holocene changes in the moments of precipitation for all simulations and timescales. For the mean, the absolute value of the difference is listed, for the other moments the ratios.

	$\Delta m_{hol-lgm}$	v_{lgm}/v_{hol}			s_{lgm}/s_{hol}			k_{lgm}/k_{hol}		
	p	p_{ann}	p_{dec}	p_{cen}	p_{ann}	p_{dec}	p_{cen}	p_{ann}	p_{dec}	p_{cen}
MPI-ESM ch4	0.27	1.23	1.18	1.27	1.16	1.21	2.33	1.60	0.98	0.17
MPI-ESM r1	0.39	1.11	1.05	1.19	1.07	0.98	-0.73	1.03	0.67	0.27
MPI-ESM r2	0.31	1.18	1.11	1.07	1.05	1.00	0.18	1.03	0.72	-0.40
MPI-ESM r3	0.34			0.89			1.12			-0.33
MPI-ESM r4	0.32			1.05			0.36			-0.30
MPI-ESM r5	0.30	1.22	1.10	0.95	1.06	1.02	0.21	0.92	0.73	2e-3
MPI-ESM r6	0.37	1.13	1.08	1.05	1.09	0.94	0.54	1.15	0.65	-0.06
MPI-ESM r7	0.39	1.12	1.09	1.13	1.06	0.95	1.09	0.86	1.11	-1.07
TraCE-21ka	0.20	0.82	0.77	0.57	1.00	0.94	0.76	1.00	0.32	-1.29
HadCM3B r1	0.26	0.82	0.84	0.84	1.12	1.19	1.07	1.88	1.53	-2.56
HadCM3B r2	0.27	0.83	0.86	0.91	1.15	1.28	0.77	2.03	2.08	0.47
FAMOUS	0.15			1.03			1.44			-31.55
ECBilt-CLIO	0.25	0.93	0.93	0.88	1.36	1.33	101.84	1.36	-0.13	-2.44

S5 Spectra

Figure S20 shows the global power spectra of the sensitivity set complementing Fig. 11 of the main text, which shows the same for the main set simulations.

S5.1 Regional spectra

Figures S21 (main set) and S22 (sensitivity set) show mean regional temperature spectra for the polar, mid-latitude and tropical regions in the Southern and Northern Hemisphere. Figures S23 and S24 present the same but for precipitation.

S6 Mechanisms

S6.1 Variability of sea ice and its influence on temperature

Moments of centennial sea ice distributions

To investigate the contribution of sea ice changes to changes in variability, Fig. S25, S26, S27 and S28 show the spatial patterns of change in the moments of sea ice for mean, standard deviation, skewness and kurtosis, respectively.

Investigation of the 120–320 yr LGM periodicity in MPI-ESM ch4

The spectra of surface temperature and precipitation show a periodicity of roughly 200 years during the LGM in MPI-ESM ch4 that is largest in the high southern latitudes (Fig. 12, S20–S24). A bandbass analysis of standard deviation for timescales from 120–320 years reveals that this originates in the sea ice of the Southern Ocean: MPI-ESM ch4 simulates particularly high standard deviation in the Southern Ocean, with the exception of the Weddel Sea. The standard deviation during the LGM in ch4 is similar or even higher than during the Deglaciation (Fig. S29) and it is larger than in any of the other simulations during the LGM (Fig. S30).

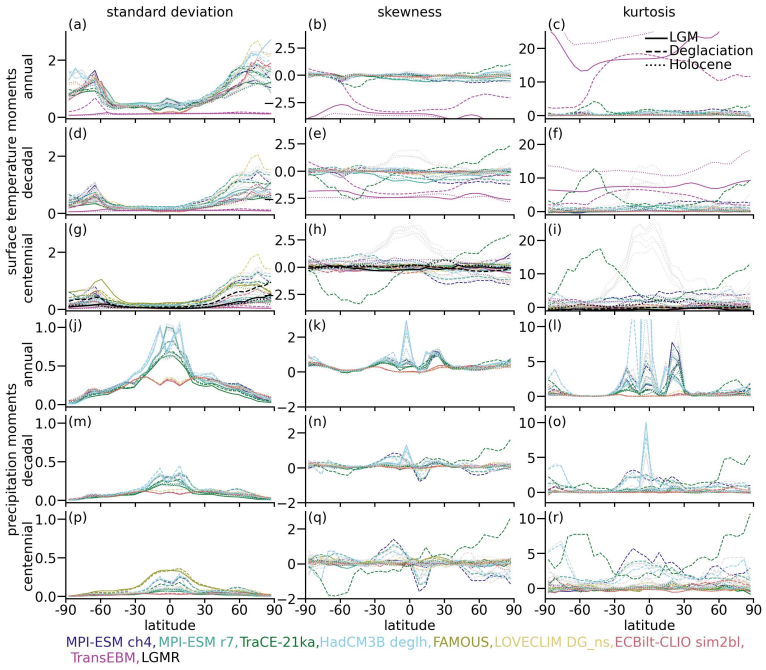


Figure S6. Changes of annual, decadal and centennial higher order moments of surface temperature (a-i) and precipitation (j-r) with latitude as in Fig. 5, but for the individual simulations of the main set. Here, standard deviation (left column), skewness (middle column) and kurtosis (right column) are shown. For each simulation LGM, Deglaciation and Holocene are shown. The sensitivity set of simulations are shown in gray.

S6.2 Change in temperature gradient

Figure S31 shows the simulated relationships between meridional temperature gradients and LGM-to-Holocene variance ratio for all simulations. While an increased meridional temperature gradient during the LGM has been suggested as a cause for increased LGM variance in comparison to the Holocene, our results cannot confirm this hypothesis.

S7 Robustness tests

To test the robustness of our results, we checked how the regridding, various versions of detrending (see Sec. S3) and the definition of the time periods affect the results. The results from a selection of these tests are shown here by showing the changes in latitudinal moments without regridding and different time period definitions in comparison to the version used in the main text.

We find that the effect of the regridding is negligible for all time periods (Fig. S32–S34). For the definitions of the time periods, we find that extending the LGM up until $18\,\mathrm{kyr}$ BP, results in Deglacial changes being included in the LGM for some of the simulations, but has still mostly limited effects, primarily on kurtosis (e.g. Fig. S32r, t, u). Limiting the LGM to run only up until $20\,\mathrm{kyr}$ BP, on the other hand, means that there is far less data for some simulations, e.g. ECBilt-CLIO (Fig. S32p–r). As

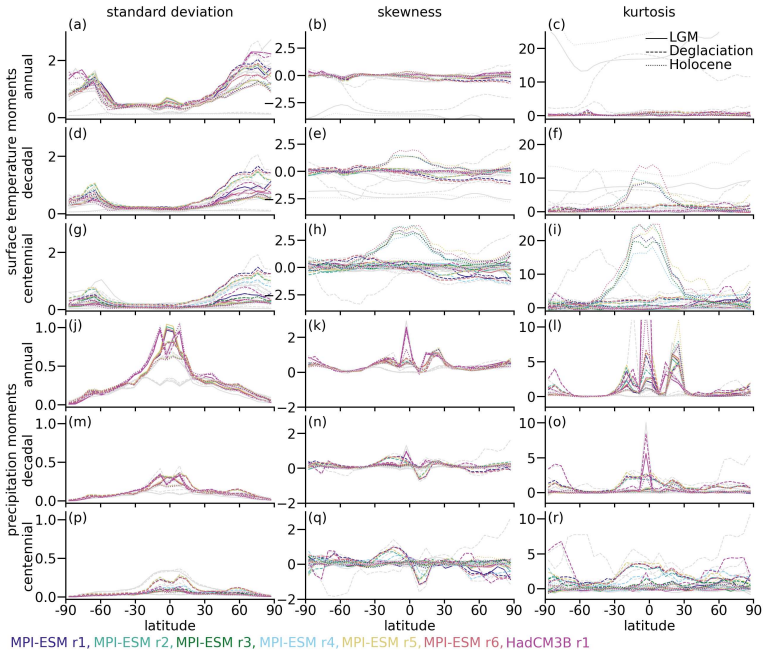


Figure S7. Same as in Fig. S7, but for the sensitivity set.

such $19\,\mathrm{kyr}$ BP as the end point of the LGM seems suitable, since it ensures enough data points for the computations without including a strong warming trend.

Regarding the definitions of onset and close of the Deglaciation, the general patterns hold when these are changed (S33). The largest differences can be found for kurtosis, but the degree of change differs between simulations. There is a similar ranking in overall levels of variability among most of the models: $20\text{–}8 > 20\text{–}9$ & $19\text{–}8 > 20\text{–}10$ & $19\text{–}9 > 19\text{–}10\,\mathrm{kyr}$ BP. An earlier onset and close thus tend to lead to slightly higher kurtosis. However, the differences do not impact any of the conclusions drawn in the main text.

For the Holocene, too, a changing onset mostly has small impacts, primarily on kurtosis (Fig.)S34). For some simulations starting at $10\,\mathrm{kyr}$, and sometimes even $9\,\mathrm{kyr}$, means that the end of Deglaciation is included in the Holocene (e.g. in MPI-ESM r7), impacting moments in high latitudes and especially the Northern Hemisphere. This is most notable for TransEBM (panels v–x). Overall, we find limited effects from both regridding and the definitions of LGM, Deglaciation and Holocene on the results. Where they exist, these are limited to kurtosis and change the overall levels of kurtosis slightly but not the patterns.

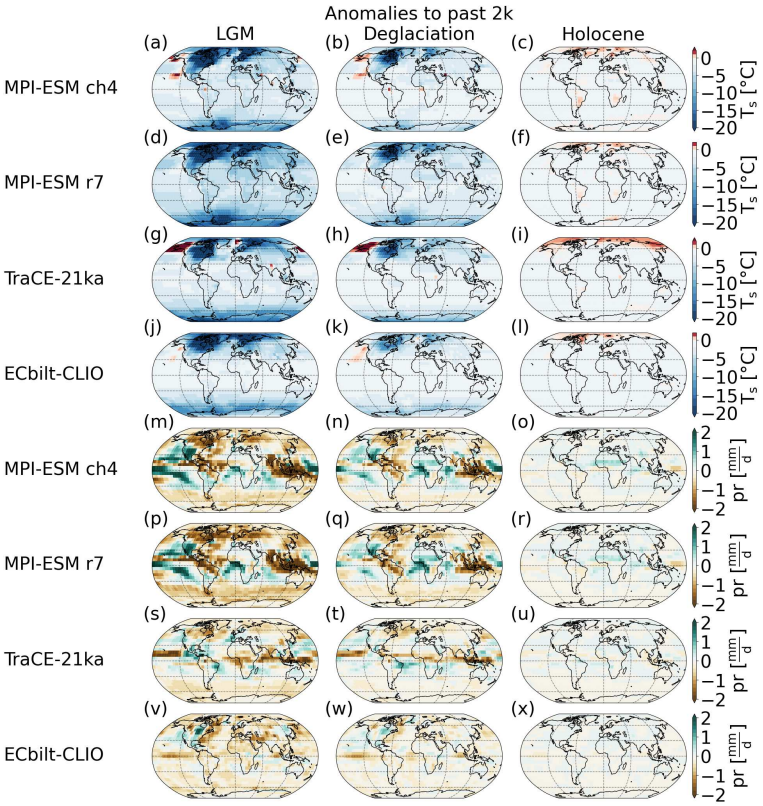


Figure S8. Exemplary mean anomalies of surface temperature (a – l) and precipitation (m – x) for the LGM (left column), Deglaciation (middle column) and Holocene (right column). For each of these, MPI-ESM ch4, MPI-ESM r7, TraCE-21ka and ECBilt-CLIO are shown.

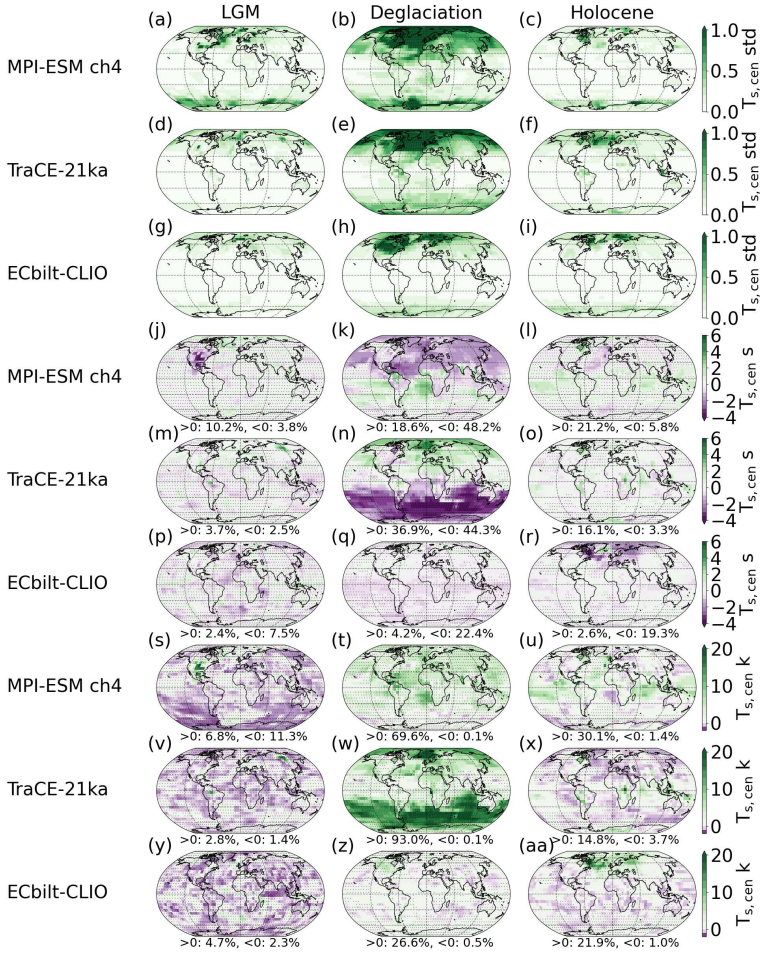


Figure S9. Exemplary patterns of change in centennial standard deviation (a – i), skewness (j – r) and kurtosis (s – aa) of surface temperature for the LGM (left column), Deglaciation (middle column) and Holocene (right column). For each of these, MPI-ESM ch4, TraCE-21ka and ECBilt-CLIO are shown. For skewness and kurtosis, the percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are given and non-significant changes are hatched.

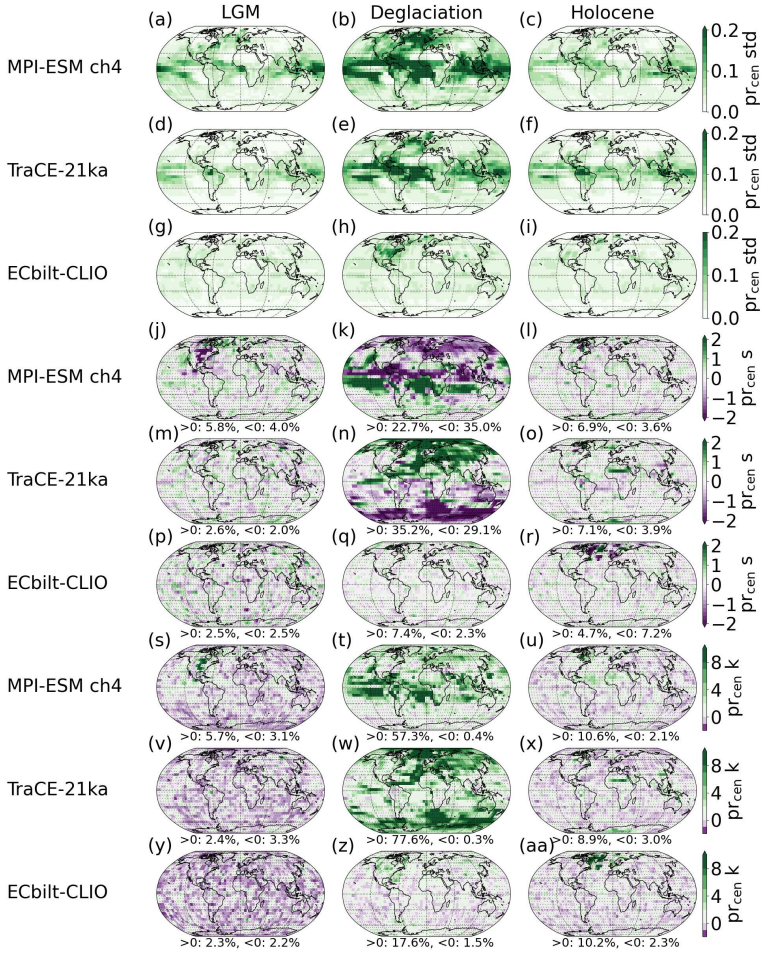


Figure S10. Exemplary patterns of change in centennial standard deviation (a – i), skewness (j – r) and kurtosis (s – aa) of precipitation for the LGM (left column), Deglaciation (middle column) and Holocene (right column). For each of these, MPI-ESM ch4, TraCE-21ka and ECBilt-CLIO are shown. For skewness and kurtosis, the percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are given and non-significant changes are hatched.

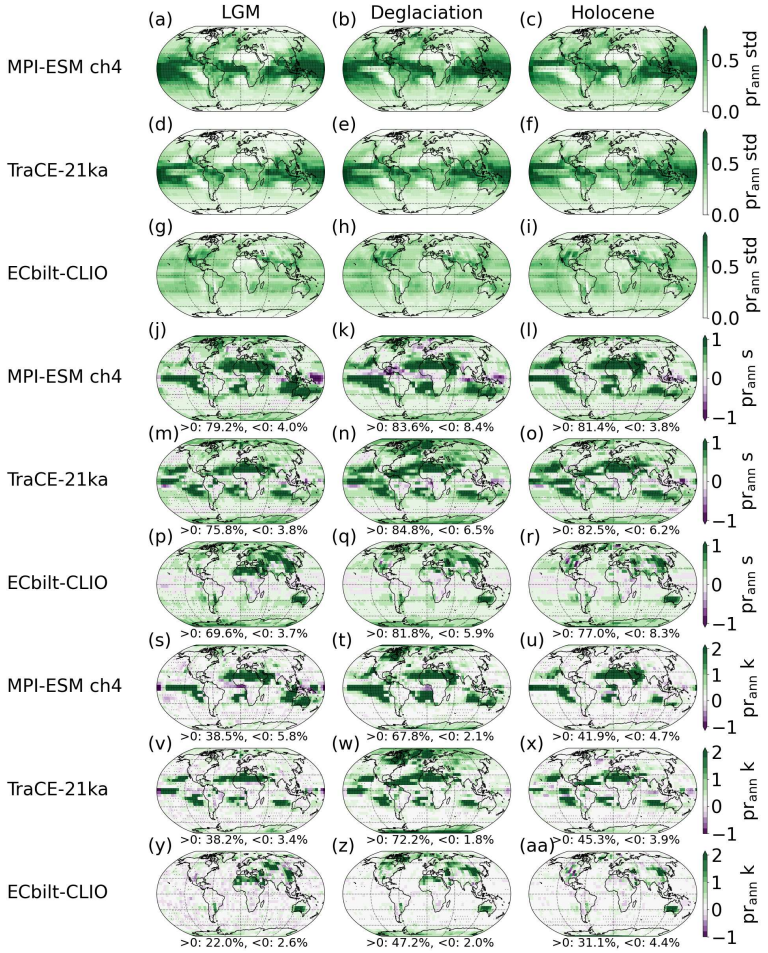


Figure S11. Exemplary patterns of change in annual standard deviation (a – i), skewness (j – r) and kurtosis (s – aa) of precipitation for the LGM (left column), Deglaciation (middle column) and Holocene (right column). For each of these, MPI-ESM ch4, TraCE-21ka and ECBilt-CLIO are shown. For skewness and kurtosis, the percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed and non-significant changes are hatched.

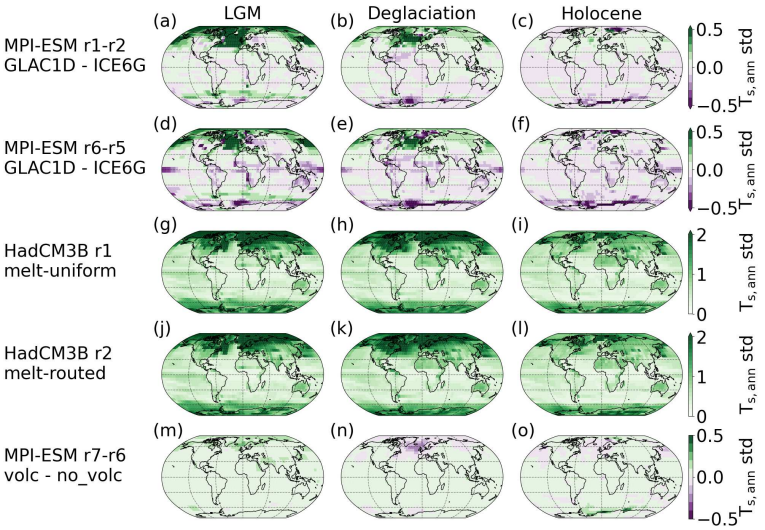


Figure S12. Regional effects of forcings on annual standard deviation of surface temperature. Forcings are noted along with the run name for each row.

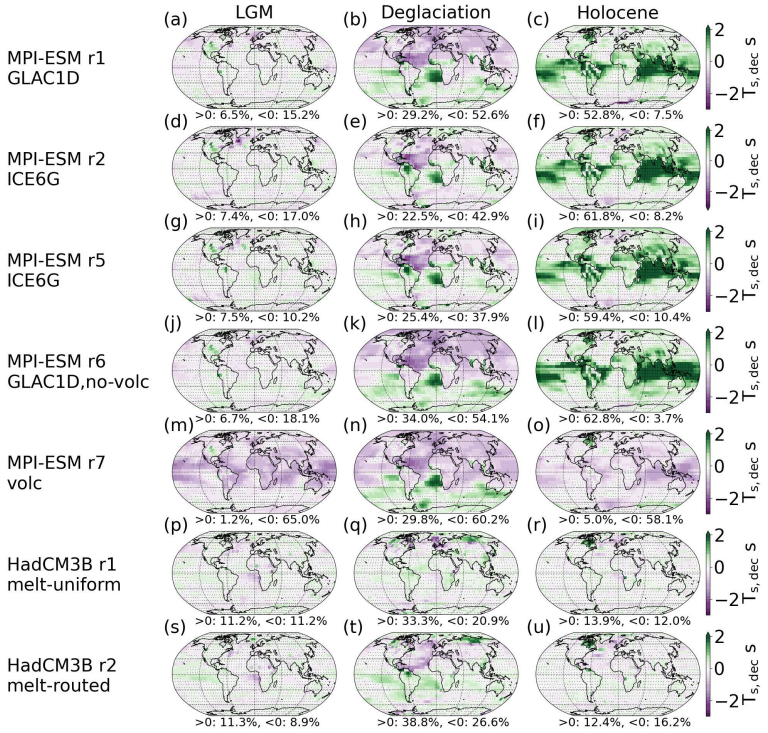


Figure S13. Regional effects of forcings on decadal skewness of surface temperature. Forcings are noted along with the run name for each row. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed. Areas, where changes are non-significant, are hatched.

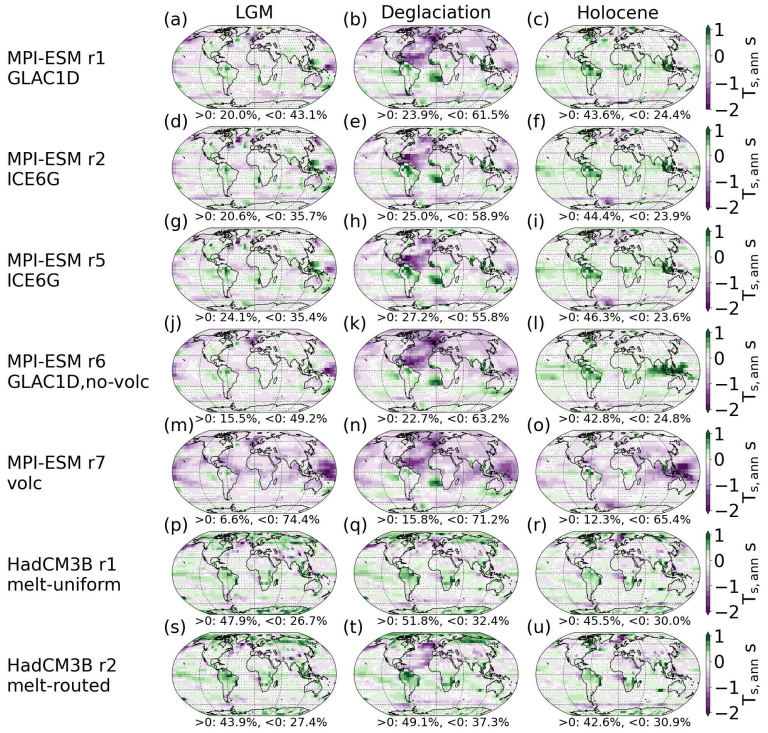


Figure S14. Regional effects of forcings on annual skewness of surface temperature. Forcings are noted along with the run name for each row. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed. Areas, where changes are non-significant, are hatched.

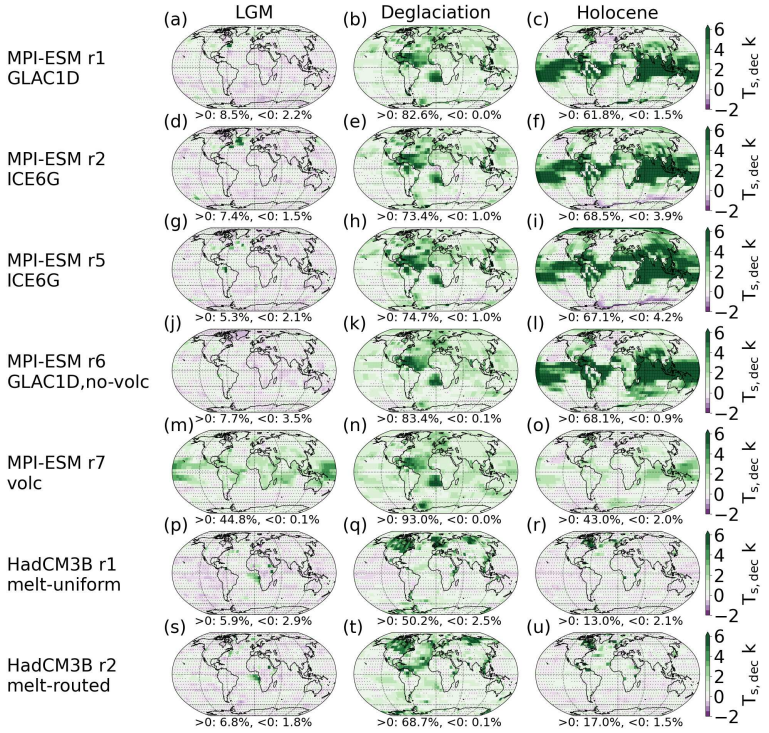


Figure S15. Regional effects of forcings on decadal kurtosis of surface temperature. Forcings are noted along with the run name for each row. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed. Areas, where changes are non-significant, are hatched.

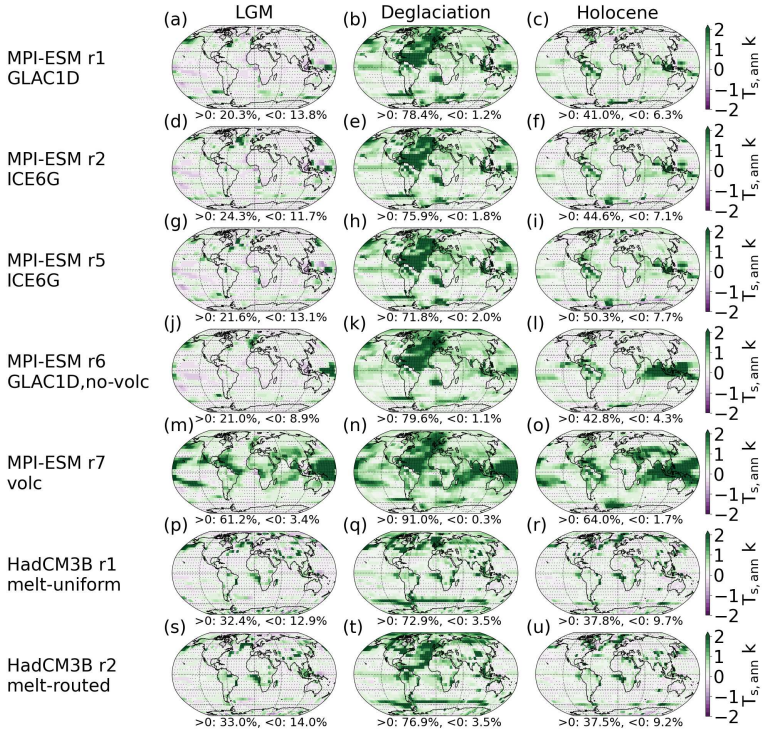


Figure S16. Regional effects of forcings on annual kurtosis of surface temperature. Forcings are noted along with the run name for each row. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed. Areas, where changes are non-significant, are hatched.

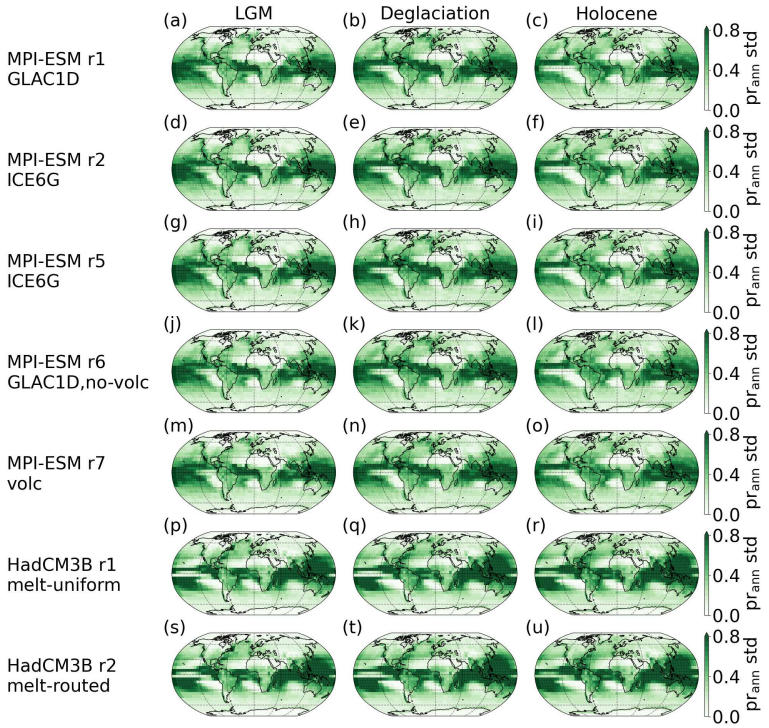


Figure S17. Regional effects of forcings on annual standard deviation of precipitation. Forcings are noted along with the run name for each row.

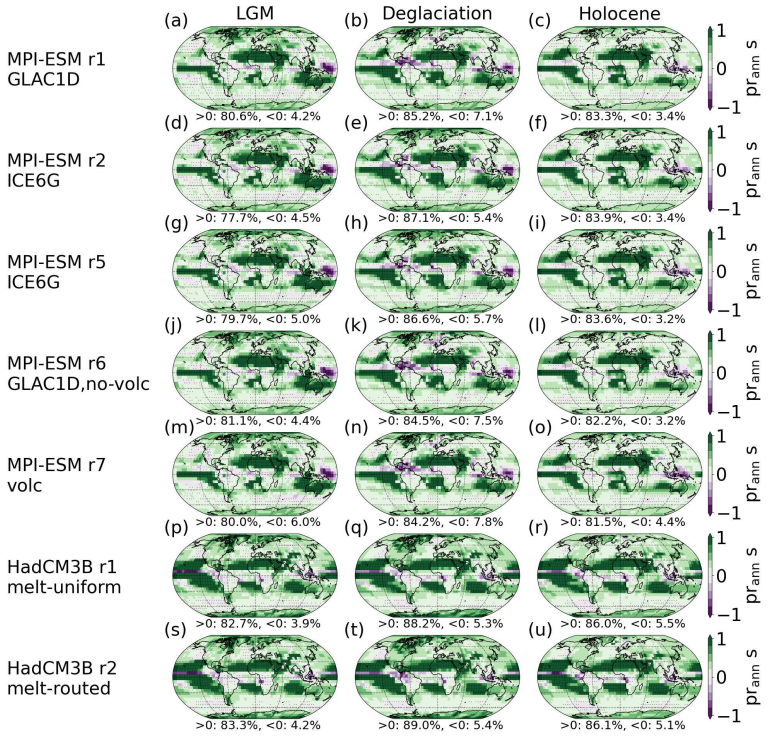


Figure S18. Regional effects of forcings on annual skewness of precipitation. Forcings are noted along with the run name for each row. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed. Areas, where changes are non-significant, are hatched.

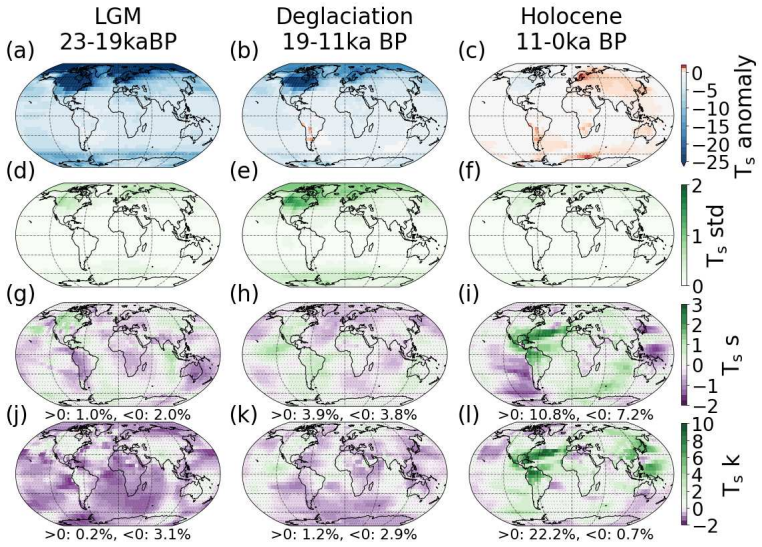


Figure S19. Spatial patterns in the surface temperature moments of LGMR for (a) – (c) mean, (d) – (f) standard deviation, (g) – (i) skewness and (j) – (l) kurtosis. For skewness and kurtosis, percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed and non-significant changes hatched.

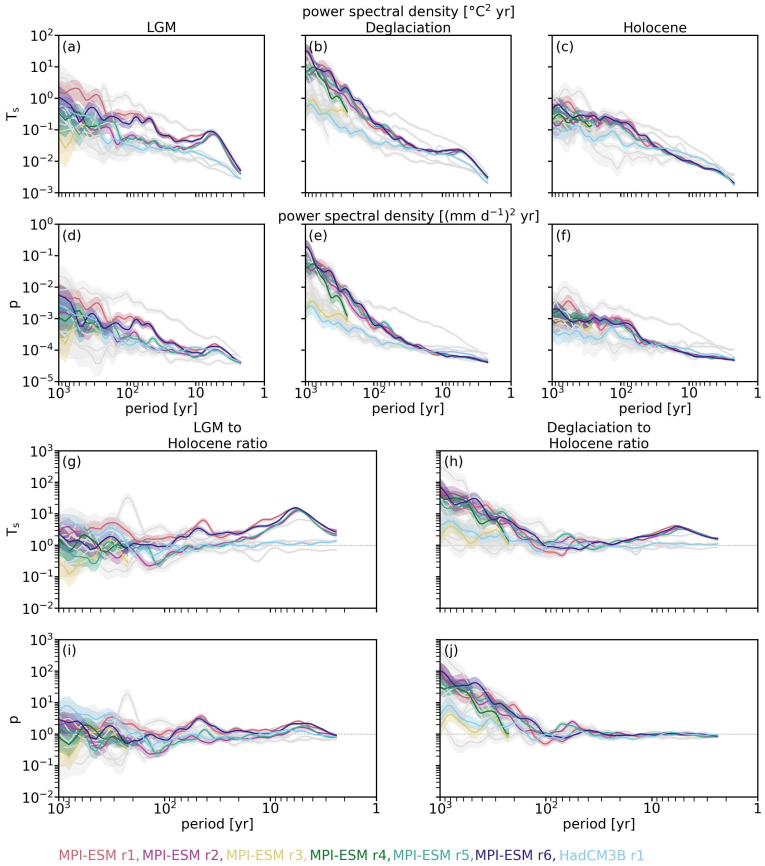


Figure S20. Spectra and spectral ratios of surface temperature (top row) and precipitation (bottom row) variability with chi-squared distributed confidence intervals. The spectra are separated by time period with (a, d) LGM, (b, e) Deglaciation and (c, f) Holocene. The spectral ratios highlight the differences between the periods showing the LGM-to-Holocene (g, i) as well the Deglaciation-to-Holocene ratio (h, j). Shown in gray are the simulations from Fig. 11.

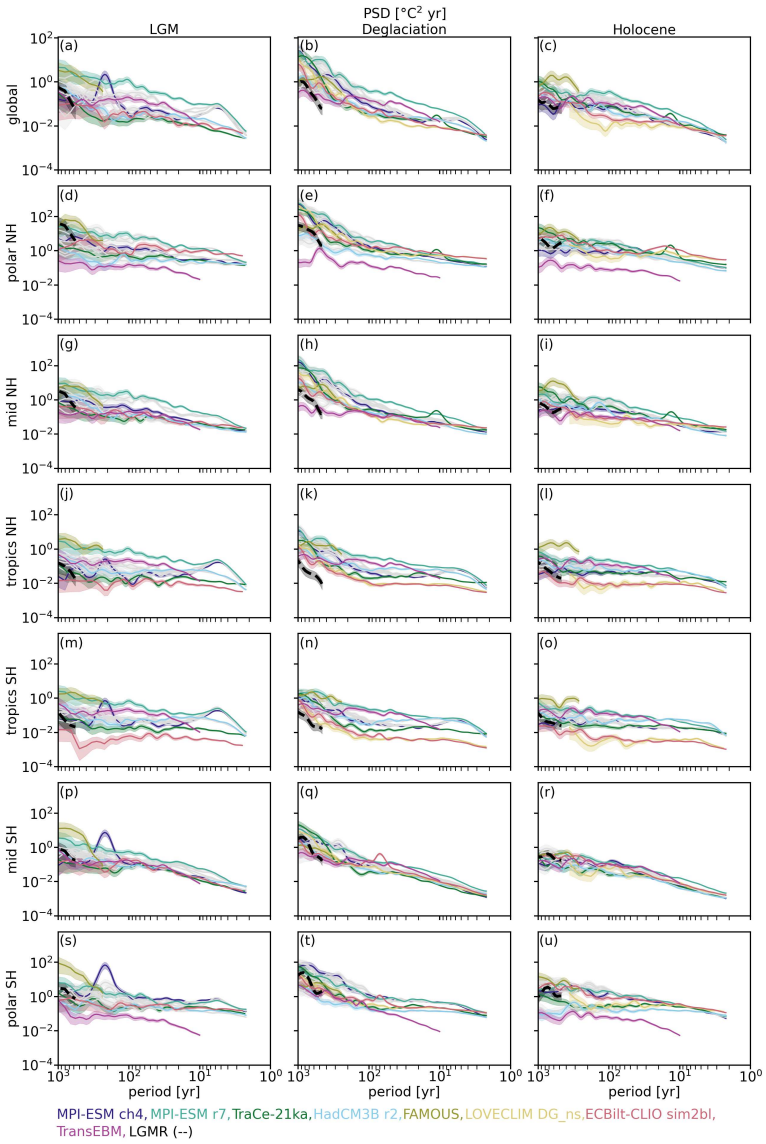


Figure S21. Global (a – c) and regional (d – u) surface temperature spectra for LGM (left column), Deglaciation (middle column) and Holocene (right column) with chi-squared distributed confidence intervals for the main set simulations. The regional spectra correspond to average timeseries of 30 degree wide latitude bands: northern polar region (60–90°N, d–f), northern mid-latitudes (30–60°N, g–i), northern tropics (0–30°N, j–l), southern tropics (0–30°S, m–o), southern mid-latitudes (30–60°S, p–r) and southern polar region (60–90°S, s–u).

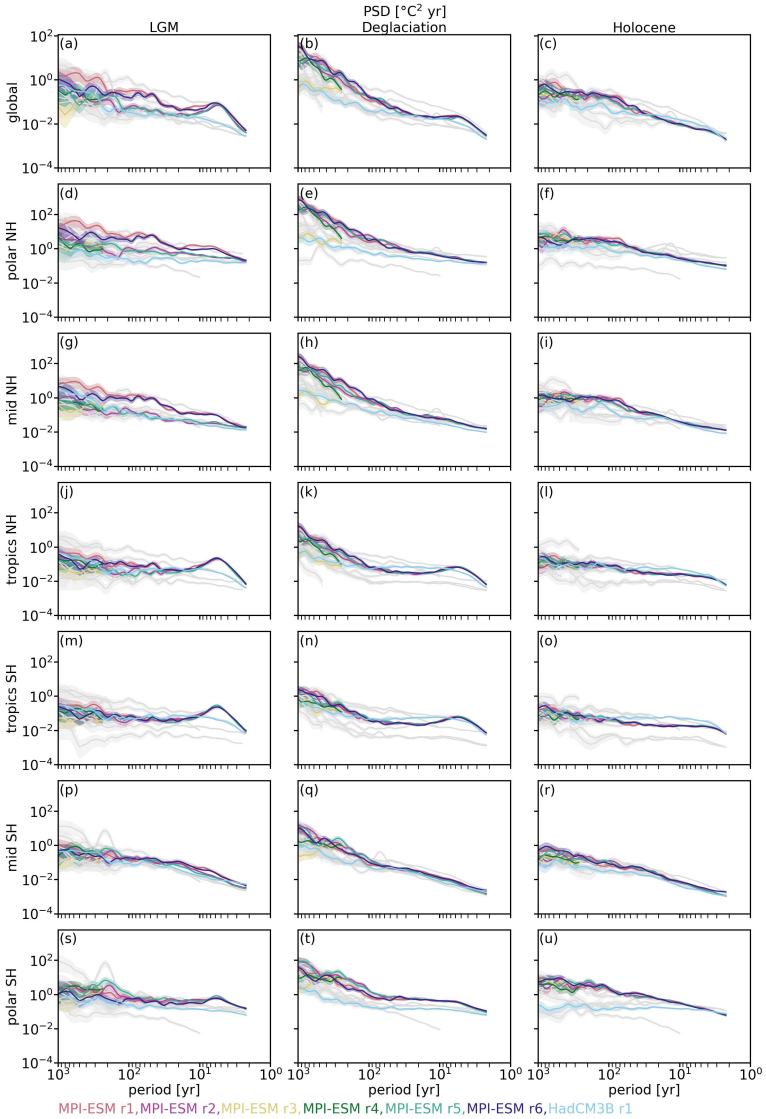


Figure S22. Same as Fig. S21 but for the sensitivity set.

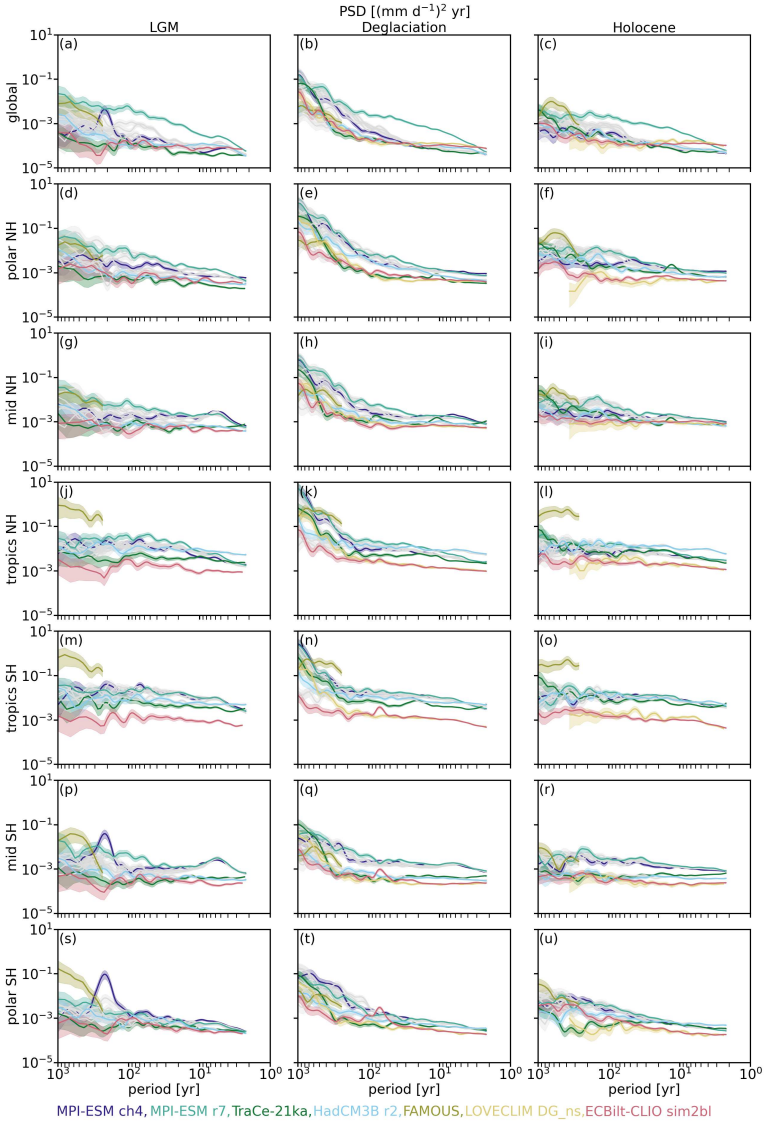


Figure S23. Same as Fig. S21 but for precipitation.

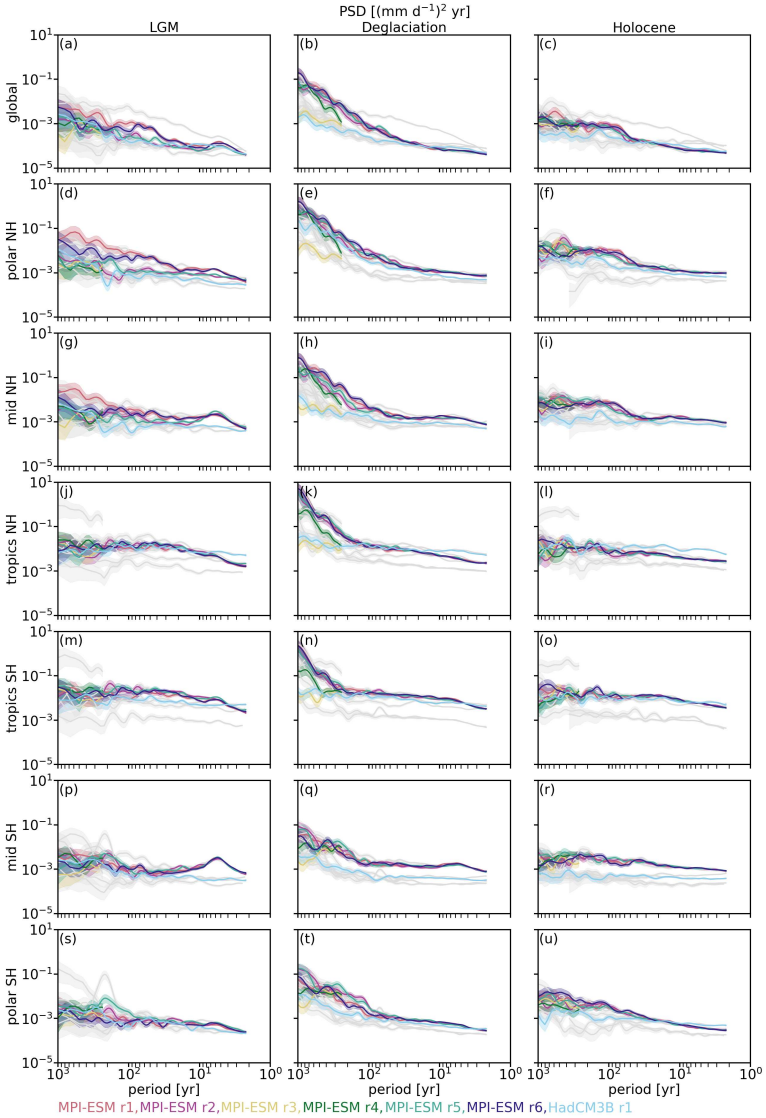


Figure S24. Same as Fig. S23 but for the sensitivity set.

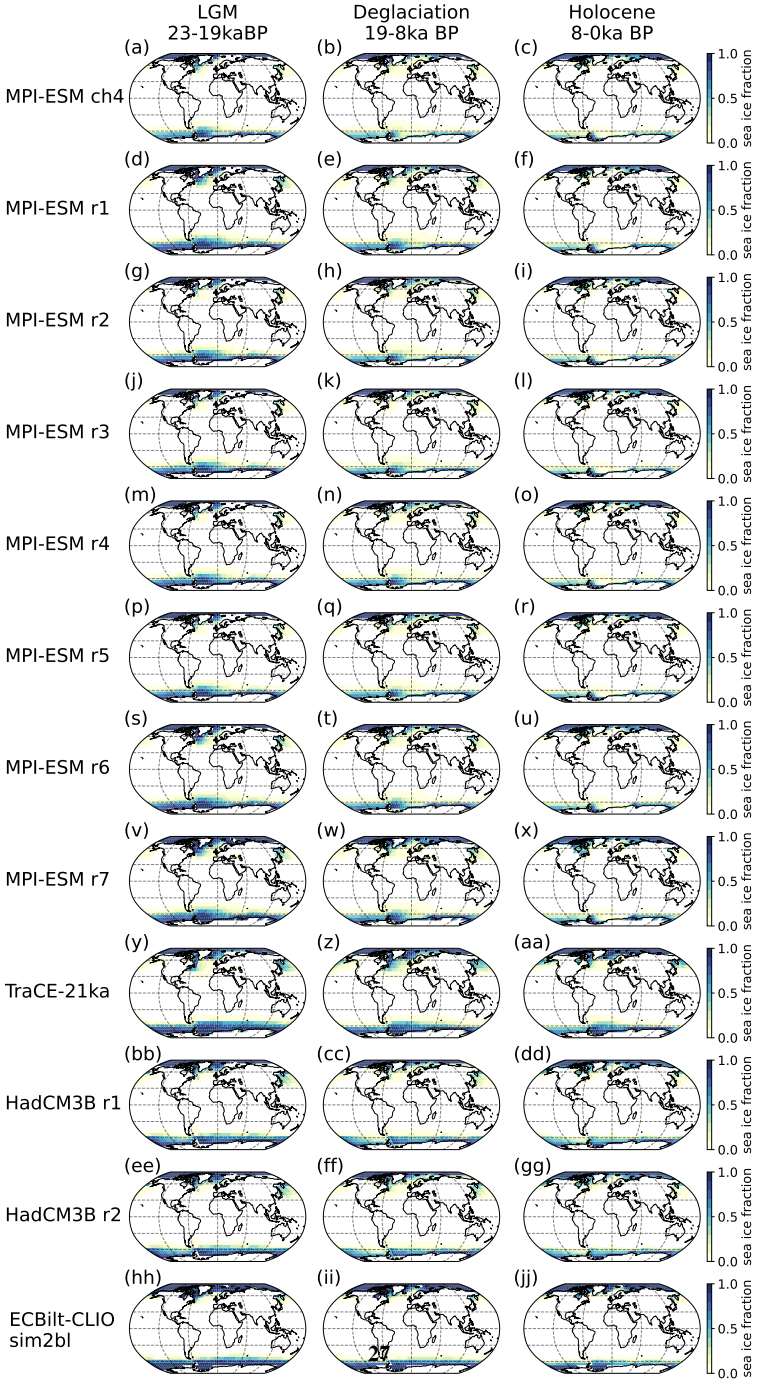


Figure S25. Global mean sea ice fraction from LGM to Holocene for simulations from the main and sensitivity set.

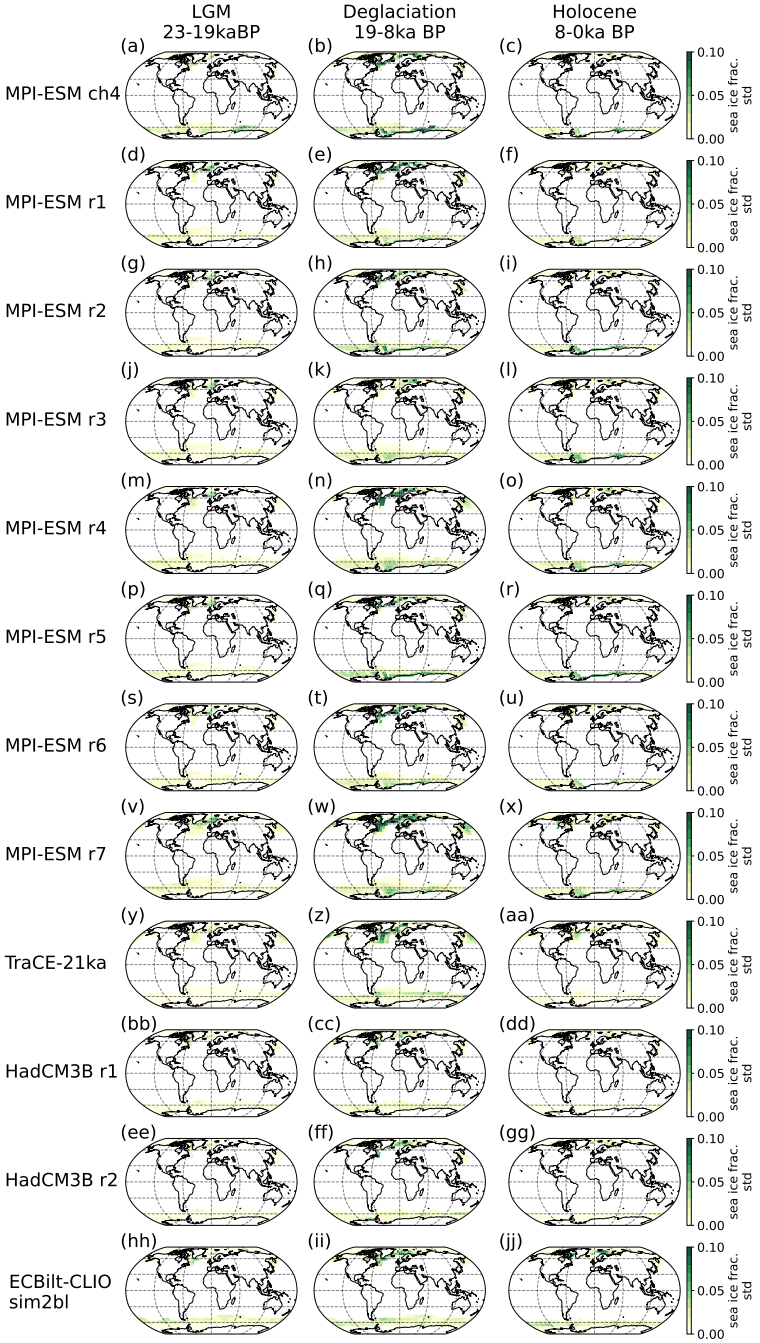


Figure S26. Global standard deviation of sea ice fraction from LGM to Holocene for simulations from the main and sensitivity set.

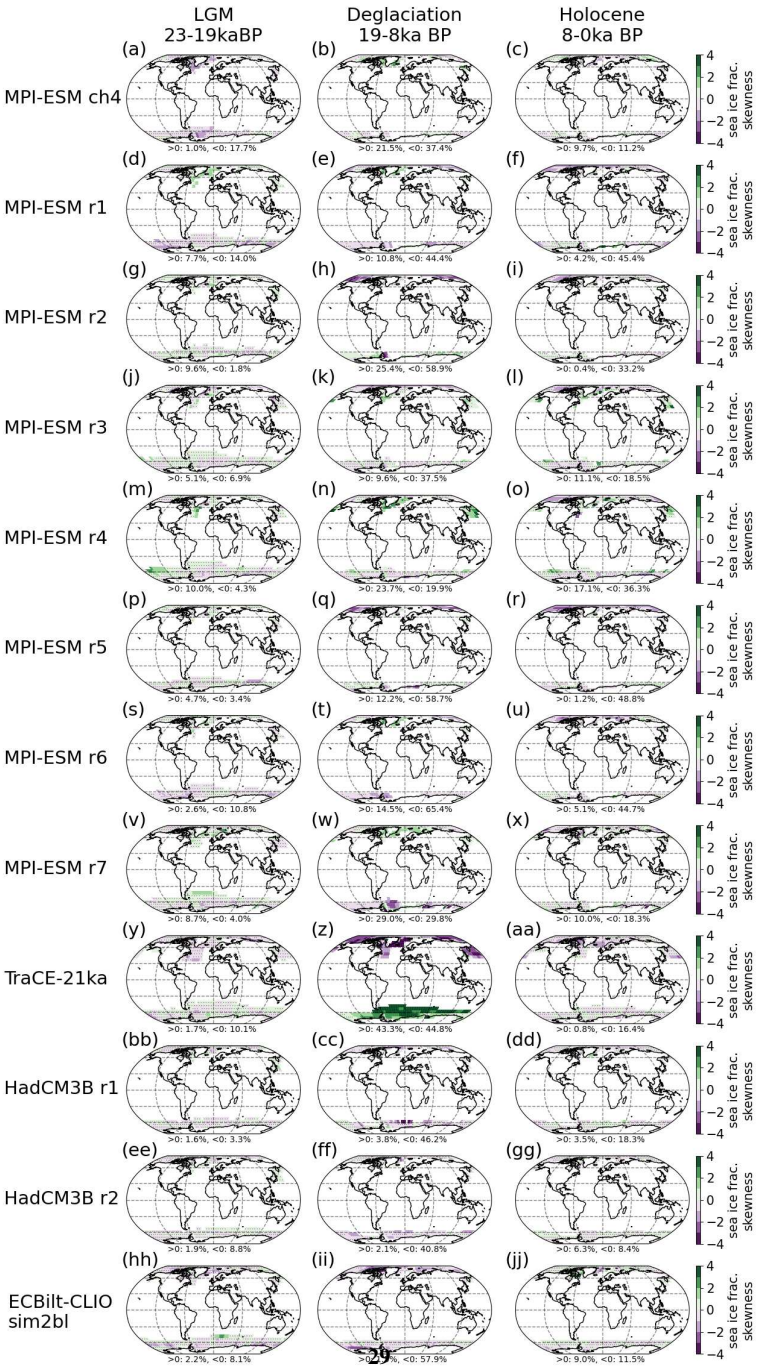


Figure S27. Global skewness of sea ice fraction from LGM to Holocene. Only areas where sea ice is constantly present were included. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed.

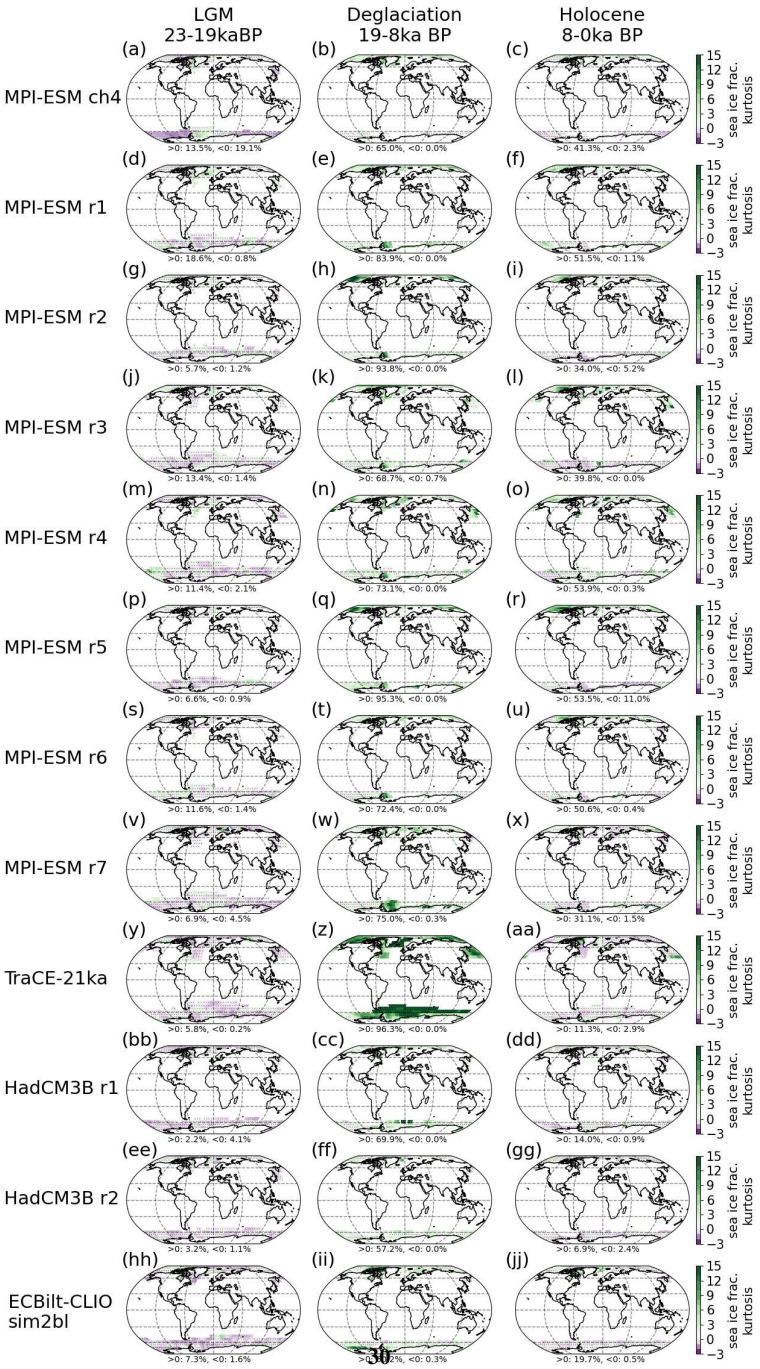


Figure S28. Global kurtosis of sea ice fraction from LGM to Holocene. Only areas where sea ice is constantly present were included. Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed.

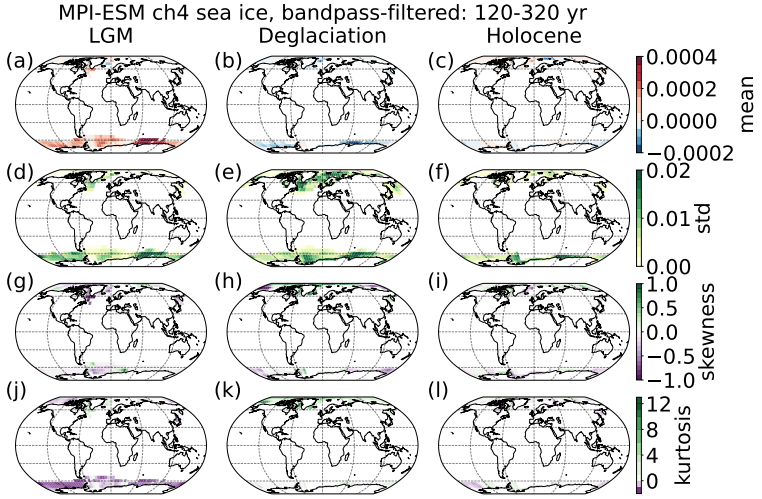


Figure S29. Bandpass-filtered (120–320 yr) moments of MPI-ESM ch4 sea ice from top row to bottom row mean, standard deviation, skewness and kurtosis for LGM (left), Deglaciation (middle) and Holocene (right). Percentages of grid boxes with significant positive and negative deviations from a Gaussian distribution are listed for skewness and kurtosis. Note that areas where sea ice disappears on centennial scale during a period were excluded during the investigation of skewness and kurtosis.

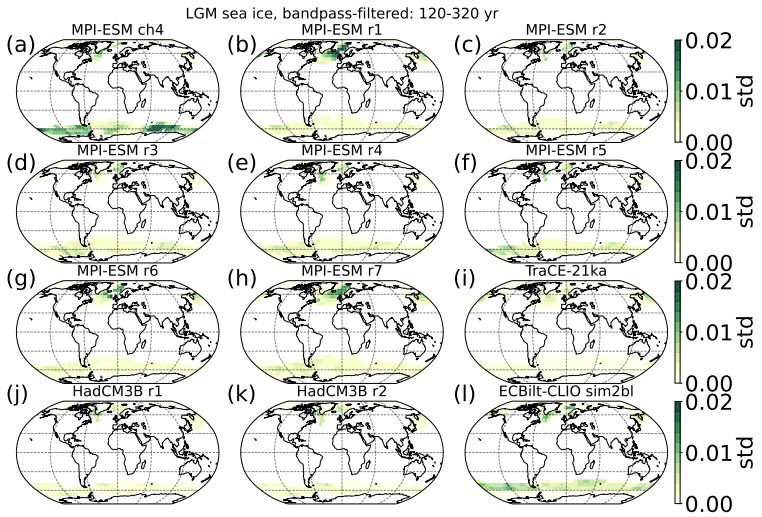


Figure S30. Bandpass-filtered (120–320 yr) standard deviation for simulations from the main and sensitivity set.

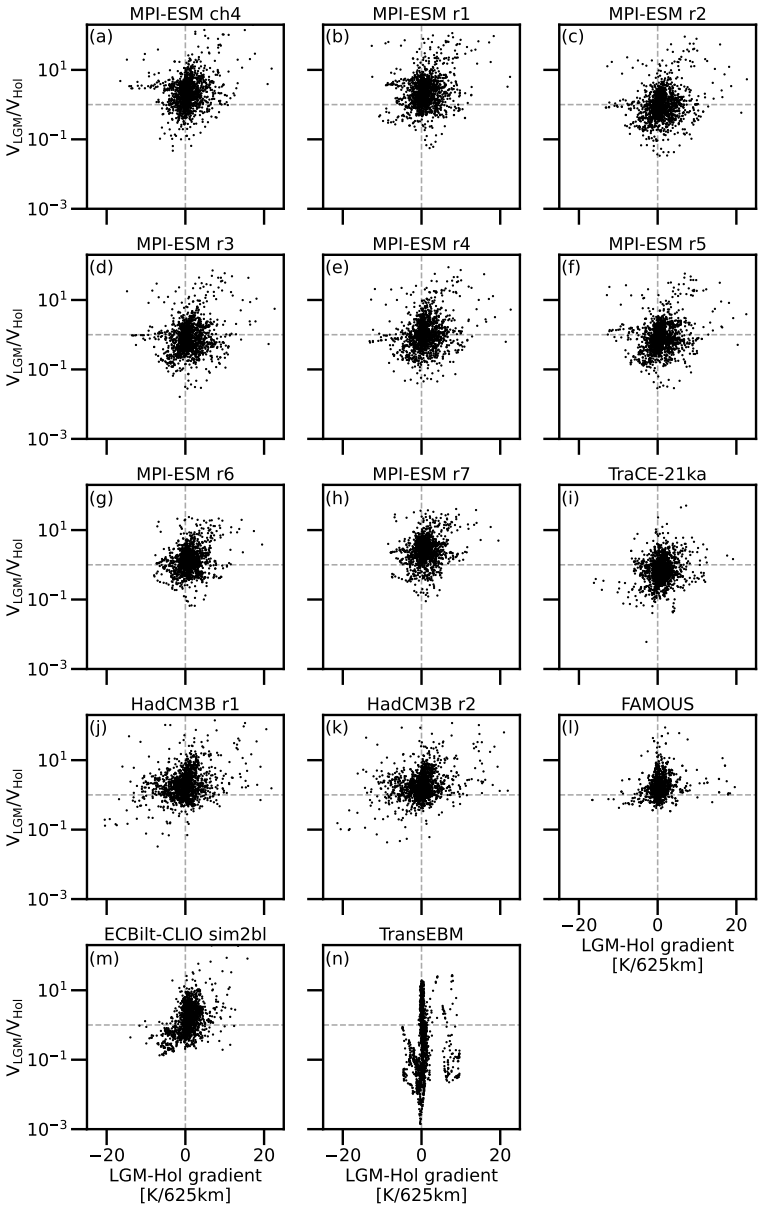


Figure S31. Global LGM-to-Holocene variance ratio against the change in meridional temperature gradient. A variance ratio of 1 and no change in meridional temperature gradient are marked by gray lines.

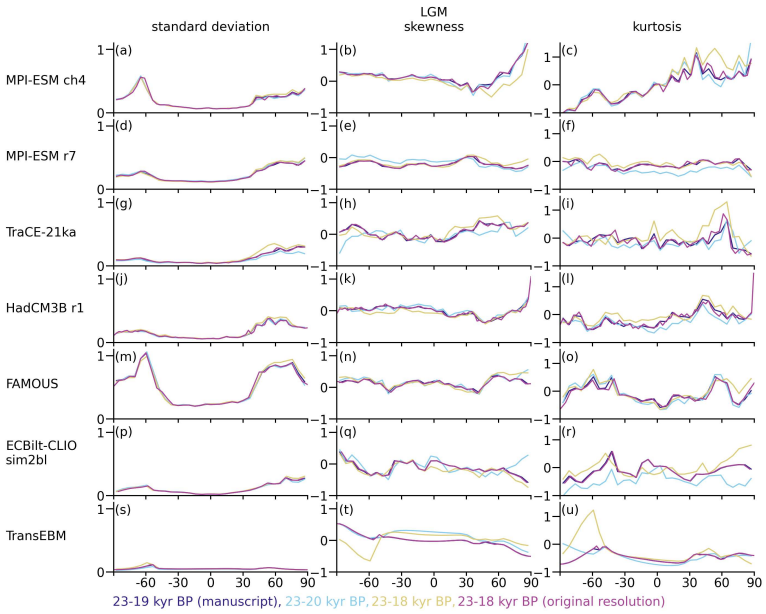


Figure S32. Centennial latitudinal moments during the LGM for different definitions of the LGM and without regridding. Shown are standard deviation (left column), skewness (middle column), kurtosis (right column) for all main set simulations. The version used in the manuscript is shown in dark blue. ECBilt-CLIO was at T21 resolution and thus did not require regridding. Therefore, for this simulation the versions marked "manuscript" and "original resolution" are identical.

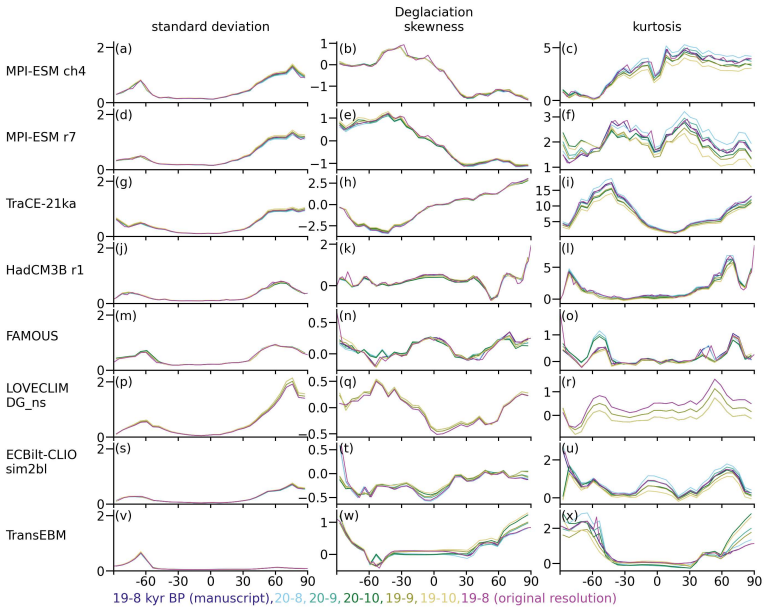


Figure S33. Centennial latitudinal moments during the Deglaciation for different definitions of the Deglaciation and without regridding. Shown are standard deviation (left column), skewness (middle column), kurtosis (right column) for all main set simulations. The version used in the manuscript is shown in dark blue. Note that since LOVECLIM DG_ns only starts at 19kyr, only those versions exist for this simulation. LOVECLIM DG_ns and ECBilt-CLIO were at T21 resolution and thus did not require regridding. Therefore, for them the versions marked "manuscript" and "original resolution" are identical.

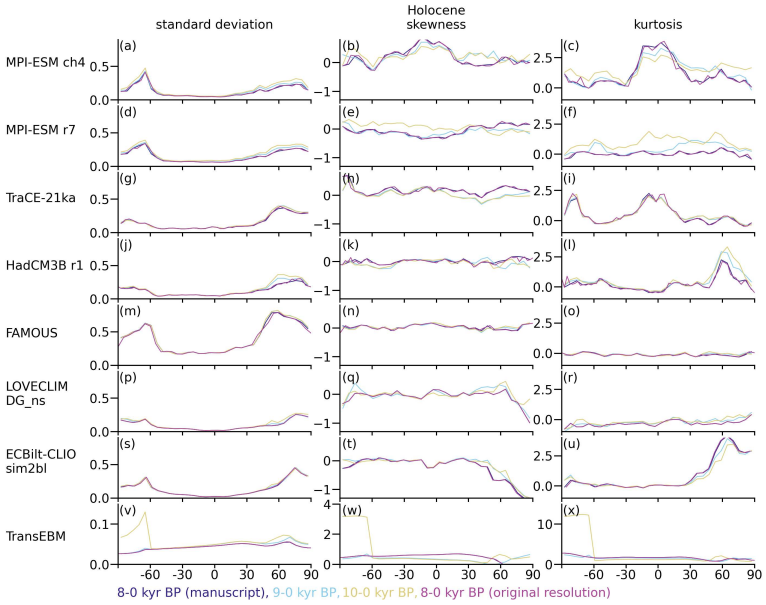


Figure S34. Centennial latitudinal moments during the Holocene for different definitions of the Holocene and without regridding. Shown are standard deviation (left column), skewness (middle column), kurtosis (right column) for all main set simulations. The version used in the manuscript is shown in dark blue. LOVECLIM DG_ns and ECBilt-CLIO were at T21 resolution and thus did not require regridding. Therefore, for them the versions marked "manuscript" and "original resolution" are identical.

References

Berger, A. L.: Long-Term Variations of Daily Insolation and Quaternary Climatic Changes, Journal of the Atmospheric Sciences, 35, 2362–2367, https://doi.org/10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2, 1978.

Ellerhoff, B. and Rehfeld, K.: Probing the Timescale Dependency of Local and Global Variations in Surface Air Temperature from Climate Simulations and Reconstructions of the Last Millennia, Physical Review E, 104, 1–14, https://doi.org/10.1103/PhysRevE.104.064136, 2021.

Filliben, J. J. and Heckert, A.: Exploratory Data Analysis, in: NIST/SEMATECH e-Handbook of Statistical Methods (Last Accessed 29/04/2024), NIST/SEMATECH, 2024.

Joanes, D. N. and Gill, C. A.: Comparing Measures of Sample Skewness and Kurtosis, Journal of the Royal Statistical Society. Series D (The Statistician), 47, 183–189, 1998.

Jonkers, L., Cartapanis, O., Langner, M., McKay, N., Mulitza, S., Strack, A., and Kucera, M.: Integrating Palaeoclimate Time Series with Rich Metadata for Uncertainty Modelling: Strategy and Documentation of the PALMOD 130k Marine Palaeoclimate Data Synthesis, Earth System Science Data Discussions, 908831, 1–53, https://doi.org/10.5194/essd-2019-223, 2020.

Kenney, J. F. and Keeping, E. S.: Mathematics of Statistics, Van Nostrand, 1948.

Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., Kinter, J. L., Loft, R., Rousset, C., Siqueira, L., Stan, C., Tomas, R., and Vertenstein, M.: Impact of Ocean Model Resolution on CCSM Climate Simulations, Climate Dynamics, 39, 1303–1328, https://doi.org/10.1007/s00382-012-1500-3, 2012.

Köhler, P., Nehrbass-Ahles, C., Schmitt, J., Stocker, T. F., and Fischer, H.: A 156 Kyr Smoothed History of the Atmospheric Greenhouse Gases CO2, CH4, and N2O and Their Radiative Forcing, Earth System Science Data, 9, 363–387, https://doi.org/10.5194/essd-9-363-2017, 2017.

Liu, B., Gan, B., Cai, W., Wu, L., Geng, T., Wang, H., Wang, S., Jing, Z., and Jia, F.: Will Increasing Climate Model Resolution Be Beneficial for ENSO Simulation?, Geophysical Research Letters, 49, e2021GL096 932, https://doi.org/10.1029/2021GL096932, 2022.

Paolini, L. F., Athanasiadis, P. J., Ruggieri, P., and Bellucci, A.: The Atmospheric Response to Meridional Shifts of the Gulf Stream SST Front and Its Dependence on Model Resolution, Journal of Climate, 35, 6007–6030, https://doi.org/10.1175/JCLI-D-21-0530.1, 2022.

Papoulis, A. and Pillai, S. U.: Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 2002.

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, 450–487, https://doi.org/10.1002/2014JB011176, 2015.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U.: Sensitivity of Simulated Climate to Horizontal and Vertical Resolution in the ECHAM5 Atmosphere Model, Journal of Climate, 19, 3771–3791, https://doi.org/10.1175/JCLI3824.1, 2006.

Schiemann, R., Demory, M.-E., Mizielinski, M. S., Roberts, M. J., Shaffrey, L. C., Strachan, J., and Vidale, P. L.: The Sensitivity of the Tropical Circulation and Maritime Continent Precipitation to Climate Model Resolution, Climate Dynamics, 42, 2455–2468, https://doi.org/10.1007/s00382-013-1997-0, 2014.

Schindlbeck-Belo, J. C., Toohey, M., Jegen, M., Kutterolf, S., and Rehfeld, K.: PalVol v1: A Proxy-Based Semi-Stochastic Ensemble Reconstruction of Volcanic Stratospheric Sulfur Injection for the Last Glacial Cycle (140–BP), Earth System Science Data, 16, 1063–1081, https://doi.org/10.5194/essd-16-1063-2024, 2024.

Steinhilber, F., Beer, J., and Fröhlich, C.: Total Solar Irradiance during the Holocene, Geophysical Research Letters, 36, 1–5, https://doi.org/10.1029/2009GL040142, 2009.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., and Doyne Farmer, J.: Testing for Nonlinearity in Time Series: The Method of Surrogate Data, Physica D: Nonlinear Phenomena, 58, 77–94, https://doi.org/10.1016/0167-2789(92)90102-S, 1992.

von Storch, H. and Zwiers, F. W.: Statistical Analysis in Climate Research, Cambridge University Press, 1999.

Zhuang, K., North, G. R., and Stevens, M. J.: A NetCDF Version of the Two-Dimensional Energy Balance Model Based on the Full Multigrid Algorithm, SoftwareX, 6, 198–202, https://doi.org/10.1016/j.softx.2017.07.003, 2017.

Ziegler, E. and Rehfeld, K.: TransEBM v. 1.0: Description, Tuning, and Validation of a Transient Model of the Earth's Energy Balance in Two Dimensions, Geoscientific Model Development, 14, 2843–2866, https://doi.org/10.5194/gmd-14-2843-2021, 2021.

List of Figures

List of Tables

Bibliography

Abe-Ouchi, A., Saito, F., Kageyama, M., Braconnot, P., Harrison, S. P., Lambeck, K., Otto-Bliesner, B. L., Peltier, W. R., Tarasov, L., Peterschmitt, J.-Y., and Takahashi, K. (2015). Ice-sheet configuration in the CMIP5/PMIP3 Last Glacial Maximum experiments. *Geoscientific Model Development*, 8(11):3621–3637.

Alexander, L. and Perkins, S. (2013). Debate heating up over changes in climate variability. *Environmental Research Letters*, 8(4):041001.

Anchukaitis, K. J. and Smerdon, J. E. (2022). Progress and uncertainties in global and hemispheric temperature reconstructions of the Common Era. *Quaternary Science Reviews*, 286.

Andersen, K. K., Azuma, N., Barnola, J. M., Bigler, M., Biscaye, P., Caillon, N., Chappellaz, J., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Flückiger, J., Fritzsche, D., Fujii, Y., Goto-Azuma, K., Grønvold, K., Gundestrup, N. S., Hansson, M., Huber, C., Hvidberg, C. S., Johnsen, S. J., Jonsell, U., Jouzel, J., Kipfstuhl, S., Landais, A., Leuenberger, M., Lorrain, R., Masson-Delmotte, V., Miller, H., Motoyama, H., Narita, H., Popp, T., Rasmussen, S. O., Raynaud, D., Rothlisberger, R., Ruth, U., Samyn, D., Schwander, J., Shoji, H., Siggard-Andersen, M. L., Steffensen, J. P., Stocker, T., Sveinbjörnsdóttir, A. E., Svensson, A., Takata, M., Tison, J. L., Thorsteinsson, T., Watanabe, O., Wilhelms, F., and White, J. W. (2004). High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature*, 431(7005):147–151.

Anderson, W. B., Seager, R., Baethgen, W., Cane, M., and You, L. (2019). Synchronous crop failures and climate-forced production variability. *Science Advances*, 5(7):eaaw1976.

Annan, J. D. and Hargreaves, J. C. (2013). A new global reconstruction of temperature changes at the Last Glacial Maximum. *Climate of the Past*, 9(1):367–376.

Annan, J. D., Hargreaves, J. C., and Mauritsen, T. (2022). A new global surface temperature reconstruction for the Last Glacial Maximum. *Climate of the Past*, 18(8):1883–1896.

Askjær, T. G., Zhang, Q., Schenk, F., Ljungqvist, F. C., Lu, Z., Brierley, C. M., Hopcroft, P. O., Jungclaus, J., Shi, X., Lohmann, G., Sun, W., Liu, J., Braconnot, P., Otto-Bliesner, B. L., Wu, Z., Yin, Q., Kang, Y., and Yang, H. (2022). Multi-centennial Holocene climate variability in proxy records and transient model simulations. *Quaternary Science Reviews*, 296:107801.

Babcock, H. W. (1961). The Topology of the Sun's Magnetic Field and the 22-Year Cycle. *The Astrophysical Journal*, 133:572.

Baede, A. P. M., Ahlonsou, E., Ding, Y., Schimel, D., Bolin, B., and Pollonais, S. (2002). The Climate System: An Overview. In *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van Der Linden, X. Dai, K. Maskell, and C.A. Johnson (Eds.)].*, volume 81, page 881. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Bakker, P., Goosse, H., and Roche, D. M. (2022). Internal climate variability and spatial temperature correlations during the past 2000 years. *Climate of the Past*, 18(11):2523–2544.

Baldini, J. U. L., Brown, R. J., and McElwaine, J. N. (2015). Was millennial scale climate change during the Last Glacial triggered by explosive volcanism? *Scientific Reports*, 5(1):17442.

Barboza, L. A., Chou Chen, S. W., Alfaro Córdoba, M., Alfaro, E. J., and Hidalgo, H. G. (2023). Spatio-temporal downscaling emulator for regional climate models. *Environmetrics*, 34(7):e2815.

Barral, A., Gomez, B., Fourel, F., Daviero-Gomez, V., and Lécuyer, C. (2017). CO2 and temperature decoupling at the million-year scale during the Cretaceous Greenhouse. *Scientific Reports*, 7(1):8310.

Bathiany, S., Dakos, V., Scheffer, M., and Lenton, T. M. (2018). Climate models predict increasing temperature variability in poor countries. *Science Advances*, 4(5):eaar5809.

Bazin, L., Landais, A., Lemieux-Dudon, B., Toyé Mahamadou Kele, H., Veres, D., Parrenin, F., Martinerie, P., Ritz, C., Capron, E., Lipenkov, V., Loutre, M.-F., Raynaud, D., Vinther, B., Svensson, A., Rasmussen, S. O., Severi, M., Blunier, T., Leuenberger, M., Fischer, H., Masson-Delmotte, V., Chappellaz, J., and Wolff, E. (2013). An optimized multi-proxy, multi-site Antarctic ice and gas orbital chronology (AICC2012): 120–800 ka. *Climate of the Past*, 9(4):1715–1731.

Bell, E. A., Boehnke, P., Harrison, T. M., and Mao, W. L. (2015). Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. *Proceedings of the National Academy of Sciences*, 112(47):14518–14521.

Berdahl, M. and Robock, A. (2013). Northern Hemispheric cryosphere response to volcanic eruptions in the Paleoclimate Modeling Intercomparison Project 3 last millennium simulations. *Journal of Geophysical Research: Atmospheres*, 118(22):12,359–12,370.

Bereiter, B., Eggleston, S., Schmitt, J., Nehrbass-Ahles, C., Stocker, T. F., Fischer, H., Kipfstuhl, S., and Chappellaz, J. (2015). Revision of the EPICA Dome C CO2 record from 800 to 600 kyr before present. *Geophysical Research Letters*, 42(2):542–549.

Berger, A. L. (1978). Long-Term Variations of Daily Insolation and Quaternary Climatic Changes. *Journal of the Atmospheric Sciences*, 35(12):2362–2367.

Bergström, A., Stringer, C., Hajdinjak, M., Scerri, E. M. L., and Skoglund, P. (2021). Origins of modern human ancestry. *Nature*, 590(7845):229–237.

Bethke, I., Outten, S., Otterå, O. H., Hawkins, E., Wagner, S., Sigl, M., and Thorne, P. (2017). Potential volcanic impacts on future climate variability. *Nature Climate Change*, 7(11):799–805.

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., and Tian, Q. (2023). Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, 619(7970):533–538.

Bian, C., Jing, Z., Wang, H., Wu, L., Chen, Z., Gan, B., and Yang, H. (2023). Oceanic mesoscale eddies as crucial drivers of global marine heatwaves. *Nature Communications*, 14(1):2970.

Blanusa, M. L., López-Zurita, C. J., and Rasp, S. (2023). Internal variability plays a dominant role in global climate projections of temperature and precipitation extremes. *Climate Dynamics*.

Boer, G. J. (2009). Changes in Interannual Variability and Decadal Potential Predictability under Global Warming. *Journal of Climate*, 22(11):3098–3109.

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., de Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J.-L., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M.-A., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J.-Y., Guenet, B., Guez, Lionel, E., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Joussaume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levavasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luyssaert, S., Madec, G., Madeleine, J.-B., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Ottlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N. (2020). Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *Journal of Advances in Modeling Earth Systems*, 12(7):e2019MS002010.

Bowen, G. J., Maibauer, B. J., Kraus, M. J., Röhl, U., Westerhold, T., Steimke, A., Gingerich, P. D., Wing, S. L., and Clyde, W. C. (2015). Two massive, rapid releases of carbon during the onset of the Palaeocene–Eocene thermal maximum. *Nature Geoscience*, 8(1):44–47.

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B., and Zhao, Y. (2012). Evaluation of climate models using palaeoclimatic data. *Nature Climate Change*, 2(6):417–424.

Bradley, R. S. (2015). *Paleoclimatology*. Elsevier, Academic Press.

Brent Dalrymple, G. (2001). The age of the Earth in the twentieth century:a problem (mostly) solved. In Lewis, C. L. E. and Knell, S. J., editors, *The Age of the Earth: From 4004 BC to AD 2002*, page 0. Geological Society of London.

Bretherton, C. S. (2015). Insights into low-latitude cloud feedbacks from high-resolution models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2054):20140415.

Briggs, R. D., Pollard, D., and Tarasov, L. (2014). A data-constrained large ensemble analysis of Antarctic evolution since the Eemian. *Quaternary Science Reviews*, 103:91–115.

Britannica, T. E. o. E. (2024). Encyclopedia Britannica. https://www.britannica.com/science/year.

Brook, E. and Buizert, C. (2018). Antarctic and global climate history viewed from ice cores. *Nature*, 558:200–208.

Brown, F. H., McDougall, I., and Fleagle, J. G. (2012). Correlation of the KHS Tuff of the Kibish Formation to volcanic ash layers at other sites, and the age of early *Homo sapiens* (Omo I and Omo II). *Journal of Human Evolution*, 63(4):577–585.

Brown, P. T., Ming, Y., Li, W., and Hill, S. A. (2017). Change in the magnitude and mechanisms of global temperature variability with warming. *Nature Climate Change*, 7(10):743–748.

Bryan, F. O., Tomas, R., Dennis, J. M., Chelton, D. B., Loeb, N. G., and McClean, J. L. (2010). Frontal Scale Air–Sea Interaction in High-Resolution Coupled Climate Models. *Journal of Climate*, 23(23):6277–6291.

Bühler, J. C., Axelsson, J., Lechleitner, F. A., Fohlmeister, J., LeGrande, A. N., Midhun, M., Sjolte, J., Werner, M., Yoshimura, K., and Rehfeld, K. (2022). Investigating stable oxygen and carbon isotopic variability in speleothem records over the last millennium using multiple isotope-enabled climate models. *Climate of the Past*, 18(7):1625–1654.

Bühler, J. C., Roesch, C., Kirschner, M., Sime, L., Holloway, M. D., and Rehfeld, K. (2021). Comparison of the oxygen isotope signatures in speleothem records and iHadCM3 model simulations for the last millennium. *Climate of the Past*, 17(3):985–1004.

Calel, R., Chapman, S. C., Stainforth, D. A., and Watkins, N. W. (2020). Temperature variability implies greater economic damages from climate change. *Nature Communications*, 11(1):1–5.

Canadell, J., Monteiro, P., Costa, M., Cotrim da Cunha, L., Cox, P., Eliseev, A., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., and Zickfeld, K. (2021). Global carbon and other biogeochemical cycles and feedbacks. In 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., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 673–816. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Castruccio, S., McInerney, D. J., Stein, M. L., Crouch, F. L., Jacob, R. L., and Moyer, E. J. (2014). Statistical emulation of climate model projections based on precomputed GCM runs. *Journal of Climate*, 27(5):1829–1844.

Cenozoic CO2 proxy integration project (CENCO2PIP) consortium (2023). Toward a Cenozoic history of atmospheric CO2. *Science*, 382(6675):eadi5177.

Chatfield, C. (2016). *The Analysis of Time Series: An Introduction, Sixth Edition*. CRC Press.

Chen, D., Rojas, M., Samset, B., Cobb, K., Diongue Niang, A., Edwards, P., Emori, S., Faria, S., Hawkins, E., Hope, P., Huybrechts, P., Meinshausen, M., Mustafa, S., Plattner, G.-K., and Tréguier, A.-M. (2021). Framing, context, and methods. In 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., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 147–286. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cheung, A. H., Mann, M. E., Steinman, B. A., Frankcombe, L. M., England, M. H., and Miller, S. K. (2017). Comparison of Low-Frequency Internal Climate Variability in CMIP5 Models and Observations. *Journal of Climate*, 30(12):4763–4776.

Choblet, M. A., Bühler, J. C., Novello, V. F., Steiger, N. J., and Rehfeld, K. (2024). A continental reconstruction of hydroclimatic variability in South America during the past 2000 years. *EGUsphere*, pages 1–67.

Christiansen, B. and Ljungqvist, F. C. (2017). Challenges and perspectives for large-scale temperature reconstructions of the past two millennia. *Reviews of Geophysics*, 55(1):40–96.

Clark, P. U., Archer, D., Pollard, D., Blum, J. D., Rial, J. A., Brovkin, V., Mix, A. C., Pisias, N. G., and Roy, M. (2006). The middle Pleistocene transition: Characteristics, mechanisms, and implications for long-term changes in atmospheric pCO2. *Quaternary Science Reviews*, 25(23):3150–3184.

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., Mitrovica, J. X., Hostetler, S. W., and McCabe, a. M. (2009). The Last Glacial Maximum. *Science (New York, N.Y.)*, 325(5941):710–714.

Collier, J. C. and Zhang, G. J. (2007). Effects of Increased Horizontal Resolution on Simulation of the North American Monsoon in the NCAR CAM3: An Evaluation Based on Surface, Satellite, and Reanalysis Data. *Journal of Climate*, 20(9):1843–1861.

Collow, T. W., Wang, W., and Kumar, A. (2019). Reduction in Northern Midlatitude 2-m Temperature Variability due to Arctic Sea Ice Loss. *Journal of Climate*, 32(16):5021–5035.

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. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, 137(654):1–28.

Crowley, T. J. (2000). Causes of Climate Change Over the Past 1000 Years. *Science*, 289(5477):270–277.

Crucifix, M. (2012). Oscillators and relaxation phenomena in Pleistocene climate theory. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1962):1140–1165.

Dai, A. and Deng, J. (2021). Arctic Amplification Weakens the Variability of Daily Temperatures over Northern Middle-High Latitudes. *Journal of Climate*, 34(7):2591–2609.

Dallmeyer, A., Kleinen, T., Claussen, M., Weitzel, N., Cao, X., and Herzschuh, U. (2022). The deglacial forest conundrum. *Nature Communications*, 13(1):6035.

Danabasoglu, G., Lamarque, J.-F., 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., van 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. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, 12(2):e2019MS001916.

Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., Hvidberg, C. S., Steffensen, J. P., Sveinbjörnsdottir, A. E., Jouzel, J., and Bond, G. (1993). Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, 364(6434):218–220.

Delworth, T. L., Rosati, A., Anderson, W., Adcroft, A. J., Balaji, V., Benson, R., Dixon, K., Griffies, S. M., Lee, H.-C., Pacanowski, R. C., Vecchi, G. A., Wittenberg, A. T., Zeng, F., and Zhang, R. (2012). Simulated Climate and Climate Change in the GFDL CM2.5 High-Resolution Coupled Climate Model. *Journal of Climate*, 25(8):2755–2781.

Deser, C., Capotondi, A., Saravanan, R., and Phillips, A. S. (2006). Tropical Pacific and Atlantic Climate Variability in CCSM3. *Journal of Climate*, 19(11):2451–2481.

Deser, C., Phillips, A., Bourdette, V., and Teng, H. (2012). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, 38(3-4):527–546.

Ding, M., Liu, H., Lin, P., Meng, Y., Zheng, W., An, B., Luan, Y., Yu, Y., Yu, Z., Li, Y., Ma, J., Chen, J., and Chen, K. (2022). A century-long eddy-resolving simulation of global oceanic large- and mesoscale state. *Scientific Data*, 9(1):691.

Ding, Y., Carton, J. A., Chepurin, G. A., Stenchikov, G., Robock, A., Sentman, L. T., and Krasting, J. P. (2014). Ocean response to volcanic eruptions in Coupled Model Intercomparison Project 5 simulations. *Journal of Geophysical Research: Oceans*, 119(9):5622–5637.

Director, H. and Bornn, L. (2015). Connecting Point-Level and Gridded Moments in the Analysis of Climate Data. *Journal of Climate*, 28(9):3496–3510.

Divine, D. V. and Dick, C. (2006). Historical variability of sea ice edge position in the Nordic Seas. *Journal of Geophysical Research: Oceans*, 111(C1).

Dodd, M. S., Papineau, D., Grenne, T., Slack, J. F., Rittner, M., Pirajno, F., O'Neil, J., and Little, C. T. S. (2017). Evidence for early life in Earth's oldest hydrothermal vent precipitates. *Nature*, 543(7643):60–64.

Dolman, A. M., Kunz, T., Groeneveld, J., and Laepple, T. (2021). A spectral approach to estimating the timescale-dependent uncertainty of paleoclimate records – Part 2: Application and interpretation. *Climate of the Past*, 17(2):825–841.

Dunsworth, H. M. (2010). Origin of the Genus Homo. *Evolution: Education and Outreach*, 3(3):353–366.

Ellerhoff, B., Kirschner, M. J., Ziegler, E., Holloway, M. D., Sime, L., and Rehfeld, K. (2022). Contrasting State-Dependent Effects of Natural Forcing on Global and Local Climate Variability. *Geophysical Research Letters*, 49(10):e2022GL098335.

Ellerhoff, B. and Rehfeld, K. (2021). Probing the timescale dependency of local and global variations in surface air temperature from climate simulations and reconstructions of the last millennia. *Physical Review E*, 104(6):1–14.

Ellerhoff, B. M. (2023). *Structure and Causes of Variability in Earth's Surface Air Temperature at the Local to Global Level*. PhD thesis, Universität Tübingen.

Esper, J., Frank, D. C., and Wilson, R. J. S. (2004). Climate reconstructions: Low-frequency ambition and high-frequency ratification. *Eos, Transactions American Geophysical Union*, 85(12):113–120.

Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M., and Anchukaitis, K. J. (2013). Applications of proxy system modeling in high resolution paleoclimatology. *Quaternary Science Reviews*, 76:16–28.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162.

Evensen, G., Vossepoel, F. C., and van Leeuwen, P. J. (2022). *Data Assimilation Fundamentals: A Unified Formulation of the State and Parameter Estimation Problem*. Springer Textbooks in Earth Sciences, Geography and Environment. Springer International Publishing, Cham.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5):1937–1958.

Fasullo, J. T., Tomas, R., Stevenson, S., Otto-Bliesner, B., Brady, E., and Wahl, E. (2017). The amplifying influence of increased ocean stratification on a future year without a summer. *Nature Communications*, 8(1):1236.

Filliben, J. J. and Heckert, A. (2024). Exploratory Data Analysis. In *NIST/SEMATECH e-Handbook of Statistical Methods (Last Accessed 29/04/2024)*. NIST/SEMATECH.

Fox-Kemper, B., Hewitt, H., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S., Edwards, T., Golledge, N., Hemer, M., Kopp, R., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I., Ruiz, L., Sallée, J.-B., Slangen, A., and Yu, Y. (2021). Ocean, cryosphere and sea level change. In 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., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 1211–1362. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Franke, J., Frank, D., Raible, C. C., Esper, J., and Brönnimann, S. (2013). Spectral biases in tree-ring climate proxies. *Nature Climate Change*, 3(4):360–364.

Franzke, C. L., Barbosa, S., Blender, R., Fredriksen, H. B., Laepple, T., Lambert, F., Nilsen, T., Rypdal, K., Rypdal, M., Scotto, M. G., Vannitsem, S., Watkins, N. W., Yang, L., and Yuan, N. (2020). The Structure of Climate Variability Across Scales. *Reviews of Geophysics*, 58(2).

Fredriksen, H. B. and Rypdal, K. (2016). Spectral characteristics of instrumental and climate model surface temperatures. *Journal of Climate*, 29(4):1253–1268.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E., Hauck, J., Landschützer, P., Le Quéré, C., Luijkx, I. T., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Barbero, L., Bates, N. R., Becker, M., Bellouin, N., Decharme, B., Bopp, L., Brasika, I. B. M., Cadule, P., Chamberlain, M. A., Chandra, N., Chau, T.-T.-T., Chevallier, F., Chini, L. P., Cronin, M., Dou, X., Enyo, K., Evans, W., Falk, S., Feely, R. A., Feng, L., Ford, D. J., Gasser, T., Ghattas, J., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Heinke, J., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jacobson, A. R., Jain, A., Jarníková, T., Jersild, A., Jiang, F., Jin, Z., Joos, F., Kato, E., Keeling, R. F., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Lan, X., Lefèvre, N., Li, H., Liu, J., Liu, Z., Ma, L., Marland, G., Mayot, N., McGuire, P. C., McKinley, G. A., Meyer, G., Morgan, E. J., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K. M., Olsen, A., Omar, A. M., Ono, T., Paulsen, M., Pierrot, D., Pocock, K., Poulter, B., Powis, C. M., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Séférian, R., Smallman, T. L., Smith, S. M., Sospedra-Alfonso, R., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tans, P. P., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., van Ooijen, E., Wanninkhof, R., Watanabe, M., Wimart-Rousseau, C., Yang, D., Yang, X., Yuan, W., Yue, X., Zaehle, S., Zeng, J., and Zheng, B. (2023). Global Carbon Budget 2023. *Earth System Science Data*, 15(12):5301–5369.

Ge, R.-F., Wilde, S. A., Zhu, W.-B., and Wang, X.-L. (2023). Earth's early continental crust formed from wet and oxidizing arc magmas. *Nature*, 623(7986):334–339.

Gent, P. R., Yeager, S. G., Neale, R. B., Levis, S., and Bailey, D. A. (2010). Improvements in a half degree atmosphere/land version of the CCSM. *Climate Dynamics*, 34(6):819–833.

Ghil, M. (2002). Natural climate variability. *Encyclopedia of Global Environmental Change*, 1:544–549.

Gómez-Navarro, J. J., Bothe, O., Wagner, S., Zorita, E., Werner, J. P., Luterbacher, J., Raible, C. C., and Montávez, J. P. (2015). A regional climate palaeosimulation for Europe in the period 1500–1990 – Part 2: Shortcomings and strengths of models and reconstructions. *Climate of the Past*, 11(8):1077–1095.

Goosse, H., Barriat, P. Y., Lefebvre, W., Loutre, M. F., and Zunz, V. (2010a). Introduction to climate dynamics and climate modeling. Online textbook.

Goosse, H., Brovkin, V., Fichefet, T., Haarsma, R., Huybrechts, P., Jongma, J., Mouchet, A., Selten, F., Barriat, P. Y., Campin, J. M., Deleersnijder, E., Driesschaert, E., Goelzer, H., Janssens, I., Loutre, M. F., Morales Maqueda, M. A., Opsteegh, T., Mathieu, P. P., Munhoven, G., Pettersson, E. J., Renssen, H., Roche, D. M., Schaeffer, M., Tartinville, B., Timmermann, A., and Weber, S. L. (2010b). Description of the Earth system model of intermediate complexity LOVECLIM version 1.2. *Geoscientific Model Development*, 3(2):603–633.

Goosse, H., Kay, J. E., Armour, K. C., Bodas-Salcedo, A., Chepfer, H., Docquier, D., Jonko, A., Kushner, P. J., Lecomte, O., Massonnet, F., Park, H. S., Pithan, F., Svensson, G., and Vancoppenolle, M. (2018). Quantifying climate feedbacks in polar regions. *Nature Communications*, 9(1).

Goosse, H., Renssen, H., Timmermann, A., and Bradley, R. S. (2005). Internal and forced climate variability during the last millennium: A model-data comparison using ensemble simulations. *Quaternary Science Reviews*, 24:1345–1360.

Gowan, E. J., Zhang, X., Khosravi, S., Rovere, A., Stocchi, P., Hughes, A. L. C., Gyllencreutz, R., Mangerud, J., Svendsen, J.-I., and Lohmann, G. (2021). A new global ice sheet reconstruction for the past 80 000 years. *Nature Communications*, 12(1):1199.

Gregoire, L. J., Valdes, P. J., Payne, A. J., and Kahana, R. (2011). Optimal tuning of a GCM using modern and glacial constraints. *Climate Dynamics*, 37(3):705–719.

Grootes, P. M., Stuiver, M., White, J. W. C., Johnsen, S., and Jouzel, J. (1993). Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores. *Nature*, 366(6455):552–554.

Gu, G. and Adler, R. F. (2011). Precipitation and Temperature Variations on the Interannual Time Scale: Assessing the Impact of ENSO and Volcanic Eruptions. *Journal of Climate*, 24(9):2258–2270.

Guichard, F. and Couvreux, F. (2017). A short review of numerical cloud-resolving models. *Tellus A: Dynamic Meteorology and Oceanography*, 69(1):1373578.

Guilyardi, E., Gualdi, S., Slingo, J., Navarra, A., Delecluse, P., Cole, J., Madec, G., Roberts, M., Latif, M., and Terray, L. (2004). Representing El Niño in Coupled Ocean–Atmosphere GCMs: The Dominant Role of the Atmospheric Component. *Journal of Climate*, 17(24):4623–4629.

Gulev, S., Thorne, P., Ahn, J., Dentener, F., Domingues, C., Gerland, S., Gong, D., Kaufman, D., Nnamchi, H., Quaas, J., Rivera, J., Sathyendranath, S., Smith, S., Trewin, B., von Schuckmann, K., and Vose, R. (2021). Changing state of the climate system. In 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., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 287–422. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Gumsley, A. P., Chamberlain, K. R., Bleeker, W., Söderlund, U., de Kock, M. O., Larsson, E. R., and Bekker, A. (2017). Timing and tempo of the Great Oxidation Event. *Proceedings of the National Academy of Sciences*, 114(8):1811–1816.

Guttal, V. and Jayaprakash, C. (2008). Changing skewness: An early warning signal of regime shifts in ecosystems. *Ecology Letters*, 11(5):450–460.

Hack, J. J., Caron, J. M., Danabasoglu, G., Oleson, K. W., Bitz, C., and Truesdale, J. E. (2006). CCSM–CAM3 Climate Simulation Sensitivity to Changes in Horizontal Resolution. *Journal of Climate*, 19(11):2267–2289.

Hajima, T., Watanabe, M., Yamamoto, A., Tatebe, H., Noguchi, M. A., Abe, M., Ohgaito, R., Ito, A., Yamazaki, D., Okajima, H., Ito, A., Takata, K., Ogochi, K., Watanabe, S., and Kawamiya, M. (2020). Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geoscientific Model Development*, 13(5):2197–2244.

Hansen, J., Sato, M., Russell, G., and Kharecha, P. (2013). Climate sensitivity, sea level and atmospheric carbon dioxide. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(2001).

Harrington, L. J., Schleussner, C.-F., and Otto, F. E. L. (2021). Quantifying uncertainty in aggregated climate change risk assessments. *Nature Communications*, 12(1):7140.

Hartmann, D. L. (2016). *Global Physical Climatology (Second Edition)*. Elsevier, Boston, 2 edition.

Hasselmann, K. (1976). Stochastic climate models: Part I. Theory. *Tellus A*, 28(6):473–485.

Hausfather, Z., Drake, H. F., Abbott, T., and Schmidt, G. A. (2020). Evaluating the Performance of Past Climate Model Projections. *Geophysical Research Letters*, 47(1):1–10.

Hawkins, E. and Sutton, R. (2009). The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bulletin of the American Meteorological Society*, 90(8):1095–1108.

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research: Atmospheres*, 113(D20).

Hays, J. D., Imbrie, J., and Shackleton, N. J. (1976). Variations in the earth's orbit: Pacemaker of the ice ages. *Science*, 194(4270):1121–1132.

He, C., Liu, Z., Otto-Bliesner, B. L., Brady, E., Zhu, C., Tomas, R., Clark, P., Zhu, J., Jahn, A., Gu, S., Zhang, J., Nusbaumer, J., Noone, D., Cheng, H., Wang, Y., Yan, M., and Bao, Y. (2021). Hydroclimate footprint of pan-Asian monsoon water isotope during the last deglaciation. *Science Advances*, 7(4):eabe2611.

He, W., Wan, S., Jiang, Y., Jin, H., Zhang, W., Wu, Q., and He, T. (2013). Detecting abrupt change on the basis of skewness: Numerical tests and applications. *International Journal of Climatology*, 33(12):2713–2727.

Head, M. J., Gibbard, P., and Salvador, A. (2008). The Quaternary: Its character and definition. *Episodes Journal of International Geoscience*, 31(2):234–238.

Hébert, R., Herzschuh, U., and Laepple, T. (2022). Millennial scale climate variability over land overprinted by ocean temperature fluctuations. *Nature Geoscience*, 15:899–905.

Hegerl, G. and Zwiers, F. (2011). Use of models in detection and attribution of climate change. *WIREs Climate Change*, 2(4):570–591.

Hegerl, G. C., Crowley, T. J., Allen, M., Hyde, W. T., Pollack, H. N., Smerdon, J., and Zorita, E. (2007). Detection of Human Influence on a New, Validated 1500-Year Temperature Reconstruction. *Journal of Climate*, 20(4):650–666.

Heinze, C., Eyring, V., Friedlingstein, P., Jones, C., Balkanski, Y., Collins, W., Fichefet, T., Gao, S., Hall, A., Ivanova, D., Knorr, W., Knutti, R., Löw, A., Ponater, M., Schultz, M. G., Schulz, M., Siebesma, P., Teixeira, J., Tselioudis, G., and Vancoppenolle, M. (2019). ESD Reviews: Climate feedbacks in the Earth system and prospects for their evaluation. *Earth System Dynamics*, 10(3):379–452.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., and Simmons, A. (2023). ERA5 monthly averaged data on single levels from 1940 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*.

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.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049.

Hoffman, P. F., Abbot, D. S., Ashkenazy, Y., Benn, D. I., Brocks, J. J., Cohen, P. A., Cox, G. M., Creveling, J. R., Donnadieu, Y., Erwin, D. H., Fairchild, I. J., Ferreira, D., Goodman, J. C., Halverson, G. P., Jansen, M. F., Le Hir, G., Love, G. D., Macdonald, F. A., Maloof, A. C., Partin, C. A., Ramstein, G., Rose, B. E. J., Rose, C. V., Sadler, P. M., Tziperman,

E., Voigt, A., and Warren, S. G. (2017). Snowball Earth climate dynamics and Cryogenian geology-geobiology. *Science Advances*, 3(11):e1600983.

Hoffmann, G., Werner, M., and Heimann, M. (1998). Water isotope module of the ECHAM atmospheric general circulation model: A study on timescales from days to several years. *Journal of Geophysical Research*, 103:16871–16896.

Hofmann, A. and Bindeman, I. (2023). Earth's first glaciation at 2.9 Ga revealed by triple oxygen isotopes. *Geochemical Perspectives Letters*, 26:20–24.

Holmes, C. R., Woollings, T., Hawkins, E., and de Vries, H. (2016). Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. *Journal of Climate*, 29(6):2221–2236.

Hopcroft, P. O., Kandlbauer, J., Valdes, P. J., and Sparks, R. S. J. (2018). Reduced cooling following future volcanic eruptions. *Climate Dynamics*, 51(4):1449–1463.

Hopcroft, P. O. and Valdes, P. J. (2021). Paleoclimate-conditioning reveals a North Africa land–atmosphere tipping point. *Proceedings of the National Academy of Sciences*, 118(45):e2108783118.

Hopcroft, P. O., Valdes, P. J., and Ingram, W. (2021). Using the Mid-Holocene "Greening" of the Sahara to Narrow Acceptable Ranges on Climate Model Parameters. *Geophysical Research Letters*, 48(6):e2020GL092043.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., Folini, D., Ji, D., Klocke, D., Qian, Y., Rauser, F., Rio, C., Tomassini, L., Watanabe, M., and Williamson, D. (2017). The Art and Science of Climate Model Tuning. *Bulletin of the American Meteorological Society*, 98(3):589–606.

Hsiang, S. M., Burke, M., and Miguel, E. (2013). Quantifying the Influence of Climate on Human Conflict. *Science*, 341(6151):1235367.

Huber, C., Leuenberger, M., Spahni, R., Flückiger, J., Schwander, J., Stocker, T. F., Johnsen, S., Landais, A., and Jouzel, J. (2006). Isotope calibrated Greenland temperature record over Marine Isotope Stage 3 and its relation to CH4. *Earth and Planetary Science Letters*, 243(3):504–519.

Hublin, J.-J., Ben-Ncer, A., Bailey, S. E., Freidline, S. E., Neubauer, S., Skinner, M. M., Bergmann, I., Le Cabec, A., Benazzi, S., Harvati, K., and Gunz, P. (2017). New fossils from Jebel Irhoud, Morocco and the pan-African origin of Homo sapiens. *Nature*, 546(7657):289–292.

Huntingford, C., Jones, P. D., Livina, V. N., Lenton, T. M., and Cox, P. M. (2013). No increase in global temperature variability despite changing regional patterns. *Nature*, 500(7462):327–330.

Huybers, K., Rupper, S., and Roe, G. H. (2016). Response of closed basin lakes to interannual climate variability. *Climate Dynamics*, 46(11):3709–3723.

Huybers, P. and Curry, W. (2006). Links between annual, Milankovitch and continuum temperature variability. *Nature*, 441(7091):329–332.

Iles, C. E. and Hegerl, G. C. (2014). The global precipitation response to volcanic eruptions in the CMIP5 models. *Environmental Research Letters*, 9(10):104012.

Iles, C. E. and Hegerl, G. C. (2015). Systematic change in global patterns of streamflow following volcanic eruptions. *Nature Geoscience*, 8(11):838–842.

Ingram, W. and Bushell, A. C. (2021). Sensitivity of Climate Feedbacks to Vertical Resolution in a General Circulation Model. *Geophysical Research Letters*, 48(12):e2020GL092268.

Ionita, M., Dima, M., Nagavciuc, V., Scholz, P., and Lohmann, G. (2021). Past megadroughts in central Europe were longer, more severe and less warm than modern droughts. *Communications Earth & Environment*, 2(1):1–9.

IPCC (2021a). Annex IV: Modes of variability [Cassou, C., A. Cherchi, Y. Kosaka (eds.)]. In 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., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 2153–2192. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2021b). Annex VII: Glossary [Matthews, J.B.R., V. Möller, R. van diemen, J.S. fuglestvedt, V. Masson-delmotte, C. Méndez, S. Semenov, A. Reisinger (eds.)]. In 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., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 2215–2256. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2021c). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, volume In Press. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC (2021d). Summary for policymakers. In 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., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 3–32. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Ivanovic, R. F., Gregoire, L. J., Kageyama, M., Roche, D. M., Valdes, P. J., Burke, A., Drummond, R., Peltier, W. R., and Tarasov, L. (2016). Transient climate simulations of the deglaciation 21-9 thousand years before present (version 1) - PMIP4 Core experiment design and boundary conditions. *Geoscientific Model Development*, 8:9045–9102.

Johnson, T. E., Kirkland, C. L., Lu, Y., Smithies, R. H., Brown, M., and Hartnady, M. I. H. (2022). Giant impacts and the origin and evolution of continents. *Nature*, 608(7922):330–335.

Jonkers, L., Cartapanis, O., Langner, M., McKay, N., Mulitza, S., Strack, A., and Kucera, M. (2020). Integrating palaeoclimate time series with rich metadata for uncertainty modelling: Strategy and documentation of the PalMod 130k marine palaeoclimate data synthesis. *Earth System Science Data*, 12(2):1053–1081.

Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., Raynaud, D., Schilt, a., Schwander, J., Selmo, E., Souchez, R., Spahni, R., Stauffer, B., Steffensen, J. P., Stenni, B., Stocker, T. F., Tison, J. L., Werner, M., and Wolff, E. W. (2007). Orbital and millennial Antarctic climate variability over the past 800,000 years. *Science (New York, N.Y.)*, 317(5839):793–796.

Jucker, M., Lane, T. P., Vincent, C. L., Webster, S., Wales, S. A., and Louf, V. (2020). Locally forced convection in subkilometre-scale simulations with the Unified Model and WRF. *Quarterly Journal of the Royal Meteorological Society*, 146(732):3450–3465.

Judt, F. and Rios-Berrios, R. (2021). Resolved Convection Improves the Representation of Equatorial Waves and Tropical Rainfall Variability in a Global Nonhydrostatic Model. *Geophysical Research Letters*, 48(14):e2021GL093265.

Kageyama, M., Braconnot, P., Harrison, S. P., Haywood, A. M., Jungclaus, J. H., Otto-Bliesner, B. L., Abe-Ouchi, A., Albani, S., Bartlein, P. J., Brierley, C., Crucifix, M., Dolan, A., Fernandez-Donado, L. L., Fischer, H., Hopcroft, P. O., Ivanovic, R. F., Lambert, F., Lunt, D. J., Mahowald, N. M., Richard Peltier, W., Phipps, S. J., Roche, D. M., Schmidt, G. A., Tarasov, L., Valdes, P. J., Zhang, Q., and Zhou, T. (2018). The PMIP4 contribution to CMIP6 - Part 1: Overview and over-arching analysis plan. *Geoscientific Model Development*, 11(3):1033–1057.

Kageyama, M., Harrison, S. P., Kapsch, M.-L., Lofverstrom, M., Lora, J. M., Mikolajewicz, U., Sherriff-Tadano, S., Vadsaria, T., Abe-Ouchi, A., Bouttes, N., Chandan, D., Gregoire, L. J., Ivanovic, R. F., Izumi, K., LeGrande, A. N., Lhardy, F., Lohmann, G., Morozova, P. A., Ohgaito, R., Paul, A., Peltier, W. R., Poulsen, C. J., Quiquet, A., Roche, D. M., Shi, X., Tierney, J. E., Valdes, P. J., Volodin, E., and Zhu, J. (2021). The PMIP4 Last Glacial Maximum experiments: Preliminary results and comparison with the PMIP3 simulations. *Climate of the Past*, 17(3):1065–1089.

Kapsch, M. L., Mikolajewicz, U., Ziemen, F., and Schannwell, C. (2022). Ocean Response in Transient Simulations of the Last Deglaciation Dominated by Underlying Ice-Sheet Reconstruction and Method of Meltwater Distribution. *Geophysical Research Letters*, 49(3):1–11.

Kapsch, M. L., Mikolajewicz, U., Ziemen, F. A., Rodehacke, C. B., and Schannwell, C. (2021). Analysis of the surface mass balance for deglacial climate simulations. *Cryosphere*, 15(2):1131–1156.

Katz, R. W. and Brown, B. G. (1992). Extreme events in changing climate variability is more important than average. *Climatic Change*, 21(3):289–302.

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., Ackerman, A. S., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Cesana, G., Cheng, Y., Clune, T. L., Cook, B. I., Cruz, C. A., Del Genio, A. D., Elsaesser, G. S., Faluvegi, G., Kiang, N. Y., Kim, D., Lacis, A. A., Leboissetier, A., LeGrande, A. N., Lo, K. K., Marshall, J., Matthews, E. E., McDermid, S., Mezuman, K., Miller, R. L., Murray, L. T., Oinas, V., Orbe, C., García-Pando, C. P., Perlwitz, J. P., Puma, M. J., Rind, D., Romanou, A., Shindell, D. T., Sun, S., Tausnev, N., Tsigaridis, K., Tselioudis, G., Weng, E., Wu, J., and Yao, M.-S. (2020). GISS-E2.1: Configurations and Climatology. *Journal of Advances in Modeling Earth Systems*, 12(8):e2019MS002025.

Kendon, E. J., Prein, A. F., Senior, C. A., and Stirling, A. (2021). Challenges and outlook for convection-permitting climate modelling. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2195):20190547.

Kidder, D. L. and Worsley, T. R. (2010). Phanerozoic Large Igneous Provinces (LIPs), HEATT (Haline Euxinic Acidic Thermal Transgression) episodes, and mass extinctions. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 295(1):162–191.

Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., Kinter, J. L., Loft, R., Rousset, C., Siqueira, L., Stan, C., Tomas, R., and Vertenstein, M. (2012). Impact of ocean model resolution on CCSM climate simulations. *Climate Dynamics*, 39(6):1303–1328.

Klavans, J. M., Poppick, A., Sun, S., and Moyer, E. J. (2017). The influence of model resolution on temperature variability. *Climate Dynamics*, 48(9):3035–3045.

Klein, S. A., Hall, A., Norris, J. R., and Pincus, R. (2017). Low-Cloud Feedbacks from Cloud-Controlling Factors: A Review. *Surveys in Geophysics*, 38(6):1307–1329.

Kleinen, T., Gromov, S., Steil, B., and Brovkin, V. (2021). Atmospheric methane underestimated in future climate projections. *Environmental Research Letters*, 16(9):094006.

Kleinen, T., Gromov, S., Steil, B., and Brovkin, V. (2023). Atmospheric methane since the last glacial maximum was driven by wetland sources. *Climate of the Past*, 19(5):1081–1099.

Kleinen, T., Mikolajewicz, U., and Brovkin, V. (2020). Terrestrial methane emissions from the Last Glacial Maximum to the preindustrial period. *Climate of the Past*, 16(2):575–595.

Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M., and Mann, M. E. (2005). A signature of persistent natural thermohaline circulation cycles in observed climate. *Geophysical Research Letters*, 32(20).

Kobayashi, C. and Sugi, M. (2004). Impact of horizontal resolution on the simulation of the Asian summer monsoon and tropical cyclones in the JMA global model. *Climate Dynamics*, 23(2):165–176.

Kochkov, D., Yuval, J., Langmore, I., Norgaard, P., Smith, J., Mooers, G., Klöwer, M., Lottes, J., Rasp, S., Düben, P., Hatfield, S., Battaglia, P., Sanchez-Gonzalez, A., Willson, M., Brenner, M. P., and Hoyer, S. (2024). Neural general circulation models for weather and climate. *Nature*, 632(8027):1060–1066.

Köhler, P., Nehrbass-Ahles, C., Schmitt, J., Stocker, T. F., and Fischer, H. (2017). A 156 kyr smoothed history of the atmospheric greenhouse gases CO2, CH4, and N2O and their radiative forcing. *Earth System Science Data*, 9(1):363–387.

Kunz, T. and Laepple, T. (2021). Frequency-dependent estimation of effective spatial degrees of freedom. *Journal of Climate*, 34(18):7373–7388.

Laepple, T. and Huybers, P. (2014a). Global and regional variability in marine surface temperatures. *Geophysical Research Letters*, 41(7):2528–2534.

Laepple, T. and Huybers, P. (2014b). Ocean surface temperature variability: Large model–data differences at decadal and longer periods. *Proceedings of the National Academy of Sciences*, 111(47):16682–16687.

Laepple, T. and Lohmann, G. (2009). Seasonal cycle as template for climate variability on astronomical timescales. *Paleoceanography*, 24(4).

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., Ravuri, S., Ewalds, T., Eaton-Rosen, Z., Hu, W., Merose, A., Hoyer, S., Holland, G., Vinyals, O., Stott, J., Pritzel, A., Mohamed, S., and Battaglia, P. (2023). Learning skillful medium-range global weather forecasting. *Science*, 382(6677):1416–1421.

Lambeck, K., Rouby, H., Purcell, A., Sun, Y., and Sambridge, M. (2014). Sea level and global ice volumes from the Last Glacial Maximum to the Holocene. *Proceedings of the National Academy of Sciences of the United States of America*, 111(43):15296–303.

Laskar, J., Joutel, F., and Robutel, P. (1993). Stabilization of the Earth's obliquity by the Moon. *Nature*, 361(6413):615–617.

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A. C. M., and Levrard, B. (2004). A long-term numerical solution for the insolation quantities of the Earth. *Astronomy & Astrophysics*, 428(1):261–285.

Le, T., Sjolte, J., and Muscheler, R. (2016). The influence of external forcing on subdecadal variability of regional surface temperature in CMIP5 simulations of the last millennium. *Journal of Geophysical Research: Atmospheres*, 121(4):1671–1682.

Lear, C. H. and Lunt, D. J. (2016). How Antarctica got its ice. *Science*, 352(6281):34–35.

Lehner, F., Schurer, A. P., Hegerl, G. C., Deser, C., and Frölicher, T. L. (2016). The importance of ENSO phase during volcanic eruptions for detection and attribution. *Geophysical Research Letters*, 43(6):2851–2858.

Leighton, R. B. (1969). A Magneto-Kinematic Model of the Solar Cycle. *The Astrophysical Journal*, 156:1.

Lempert, R. J., Schlesinger, M. E., Bankes, S. C., and Andronova, N. G. (2000). The impacts of climate variability on near-term policy choices and the value of information. *Climatic Change*, 45(1):129–161.

Ligrone, R. (2019). The Great Oxygenation Event. In Ligrone, R., editor, *Biological Innovations That Built the World: A Four-billion-year Journey through Life and Earth History*, pages 129–154. Springer International Publishing, Cham.

Lisiecki, L. E. and Raymo, M. E. (2005). A Pliocene-Pleistocene stack of 57 globally distributed benthic δ 18O records. *Paleoceanography*, 20(1):1–17.

Liu, F., Chai, J., Wang, B., Liu, J., Zhang, X., and Wang, Z. (2016). Global monsoon precipitation responses to large volcanic eruptions. *Scientific Reports*, 6(1):24331.

Liu, S., Zeman, C., Sørland, S. L., and Schär, C. (2022). Systematic Calibration of a Convection-Resolving Model: Application Over Tropical Atlantic. *Journal of Geophysical Research: Atmospheres*, 127(23):e2022JD037303.

Liu, X.-M., Kah, L. C., Knoll, A. H., Cui, H., Wang, C., Bekker, A., and Hazen, R. M. (2021). A persistently low level of atmospheric oxygen in Earth's middle age. *Nature Communications*, 12(1):351.

Ljungqvist, F. C., Zhang, Q., Brattström, G., Krusic, P. J., Seim, A., Li, Q., Zhang, Q., and Moberg, A. (2019). Centennial-scale temperature change in last millennium simulations and proxy-based reconstructions. *Journal of Climate*, 32(9):2441–2482.

Lovejoy, S. (2021). The half-order energy balance equation – Part 2: The inhomogeneous HEBE and 2D energy balance models. *Earth System Dynamics*, 12(2):489–511.

Lovejoy, S., Procyk, R., Hébert, R., and Del Rio Amador, L. (2021). The fractional energy balance equation. *Quarterly Journal of the Royal Meteorological Society*, 147(736):1964–1988.

Lovejoy, S. and Varotsos, C. (2016). Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forcings. *Earth System Dynamics*, 7(1):133–150.

Luthardt, L. and Rößler, R. (2017). Fossil forest reveals sunspot activity in the early Permian. *Geology*, 45(3):279–282.

Lyons, T. W., Reinhard, C. T., and Planavsky, N. J. (2014). The rise of oxygen in Earth's early ocean and atmosphere. *Nature*, 506(7488):307–315.

Maher, N., McGregor, S., England, M. H., and Gupta, A. S. (2015). Effects of volcanism on tropical variability. *Geophysical Research Letters*, 42(14):6024–6033.

Manhes, G., Allègre, C. J., Dupré, B., and Hamelin, B. (1980). Lead isotope study of basic-ultrabasic layered complexes: Speculations about the age of the earth and primitive mantle characteristics. *Earth and Planetary Science Letters*, 47(3):370–382.

Marshall, J. and Plumb, R. A. (2008). *Atmosphere, Ocean, and Climate Dynamics*. Elsevier.

Marti, O., Braconnot, P., Dufresne, J.-L., Bellier, J., Benshila, R., Bony, S., Brockmann, P., Cadule, P., Caubel, A., Codron, F., de Noblet, N., Denvil, S., Fairhead, L., Fichefet, T., Foujols, M.-A., Friedlingstein, P., Goosse, H., Grandpeix, J.-Y., Guilyardi, E., Hourdin, F., Idelkadi, A., Kageyama, M., Krinner, G., Lévy, C., Madec, G., Mignot, J., Musat, I., Swingedouw, D., and Talandier, C. (2010). Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution. *Climate Dynamics*, 34(1):1–26.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S., Péan, C., Berger, S., Chaud, 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. (2021). IPCC, 2021: Summary for Policymakers. 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*, pages 3–32. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

McKinnon, K. A., Rhines, A., Tingley, M. P., and Huybers, P. (2016). The changing shape of Northern Hemisphere summer temperature distributions. *Journal of Geophysical Research: Atmospheres*, 121(15):8849–8868.

McManus, J. F., Oppo, D. W., and Cullen, J. L. (1999). A 0.5-Million-Year Record of Millennial-Scale Climate Variability in the North Atlantic. *Science*, 283(5404):971–975.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R. (2017). Historical greenhouse gas concentrations for climate modelling (CMIP6). *Geoscientific Model Development*, 10(5):2057–2116.

Menviel, L., Timmermann, A., Timm, O. E., and Mouchet, A. (2011). Deconstructing the Last Glacial termination: The role of millennial and orbital-scale forcings. *Quaternary Science Reviews*, 30(9-10):1155–1172.

Mikolajewicz, U., Kapsch, M.-L., Schannwell, C., Six, K. D., Ziemen, F. A., Bagge, M., Baudouin, J.-P., Erokhina, O., Gayler, V., Klemann, V., Meccia, V. L., Mouchet, A., and Riddick, T. (2025). Deglaciation and abrupt events in a coupled comprehensive atmosphere–ocean–ice-sheet–solid-earth model. *Climate of the Past*, 21(3):719–751.

Miller, G. H., Geirsdóttir, Á., Zhong, Y., Larsen, D. J., Otto-Bliesner, B. L., Holland, M. M., Bailey, D. A., Refsnider, K. A., Lehman, S. J., Southon, J. R., Anderson, C., Björnsson, H., and Thordarson, T. (2012). Abrupt onset of the Little Ice Age triggered by volcanism and sustained by sea-ice/ocean feedbacks. *Geophysical Research Letters*, 39(2).

Miller, K. G., Browning, J. V., Schmelz, W. J., Kopp, R. E., Mountain, G. S., and Wright, J. D. (2020). Cenozoic sea-level and cryospheric evolution from deep-sea geochemical and continental margin records. *Science Advances*, 6(20):eaaz1346.

Mitchell, J. M. (1976). An overview of climatic variability and its causal mechanisms. *Quaternary Research*, 6(4):481–493.

Monnin, E., Indermühle, A., Dällenbach, A., Flückiger, J., Stauffer, B., Stocker, T. F., Raynaud, D., and Barnola, J.-M. (2001). Atmospheric CO2 Concentrations over the Last Glacial Termination. *Science*, 291(5501):112–114.

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. (2021). An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set. *Journal of Geophysical Research: Atmospheres*, 126(3):e2019JD032361.

Mounier, A. and Mirazón Lahr, M. (2019). Deciphering African late middle Pleistocene hominin diversity and the origin of our species. *Nature Communications*, 10(1):3406.

Müller, R. D., Sdrolias, M., Gaina, C., Steinberger, B., and Heine, C. (2008). Long-Term Sea-Level Fluctuations Driven by Ocean Basin Dynamics. *Science*, 319(5868):1357–1362.

Muntjewerf, L., Petrini, M., Vizcaino, M., Ernani da Silva, C., Sellevold, R., Scherrenberg, M. D. W., Thayer-Calder, K., Bradley, S. L., Lenaerts, J. T. M., Lipscomb, W. H., and Lofverstrom, M. (2020). Greenland Ice Sheet Contribution to 21st Century Sea Level Rise as Simulated by the Coupled CESM2.1-CISM2.1. *Geophysical Research Letters*, 47(9):e2019GL086836.

Muthers, S., Arfeuille, F., Raible, C. C., and Rozanov, E. (2015). The impacts of volcanic aerosol on stratospheric ozone and the Northern Hemisphere polar vortex: Separating radiative-dynamical changes from direct effects due to enhanced aerosol heterogeneous chemistry. *Atmospheric Chemistry and Physics*, 15(20):11461–11476.

Nakada, M., Kimura, R., Okuno, J., Moriwaki, K., Miura, H., and Maemoku, H. (2000). Late Pleistocene and Holocene melting history of the Antarctic ice sheet derived from sea-level variations. *Marine Geology*, 167(1):85–103.

National Center for Atmospheric Research Staff (Eds) (2017). The Climate Data Guide: Common Spectral Model Grid Resolutions. https://climatedataguide.ucar.edu/climate-tools/common-spectral-model-grid-resolutions.

Neukom, R., Barboza, L. A., Erb, M. P., Shi, F., Emile-Geay, J., Evans, M. N., Franke, J., Kaufman, D. S., Lücke, L., Rehfeld, K., Schurer, A., Zhu, F., Brönnimann, S., Hakim, G. J., Henley, B. J., Ljungqvist, F. C., McKay, N., Valler, V., and von Gunten, L. (2019). Consistent multidecadal variability in global temperature reconstructions and simulations over the Common Era. *Nature Geoscience*, 12(8):643–649.

Nichols, C. I. O., Weiss, B. P., Eyster, A., Martin, C. R., Maloof, A. C., Kelly, N. M., Zawaski, M. J., Mojzsis, S. J., Watson, E. B., and Cherniak, D. J. (2024). Possible Eoarchean Records of the Geomagnetic Field Preserved in the Isua Supracrustal Belt, Southern West Greenland. *Journal of Geophysical Research: Solid Earth*, 129(4):e2023JB027706.

Nilsen, T., Rypdal, K., and Fredriksen, H. B. (2016). Are there multiple scaling regimes in Holocene temperature records? *Earth System Dynamics*, 7(2):419–439.

North, G. R. and Kim, K.-Y. (2017). *Energy Balance Climate Models*. Wiley-VCH, Weinheim.

Ogg, J. G. and Pillans, B. (2008). Establishing Quaternary as a formal international Period/System. *Episodes*, 31(2):230–233.

O'Gorman, P. A. and Dwyer, J. G. (2018). Using Machine Learning to Parameterize Moist Convection: Potential for Modeling of Climate, Climate Change, and Extreme Events. *Journal of Advances in Modeling Earth Systems*, 10(10):2548–2563.

Olonscheck, D., Mauritsen, T., and Notz, D. (2019). Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations. *Nature Geoscience*, 12(6):430–434.

Olonscheck, D. and Notz, D. (2017). Consistently Estimating Internal Climate Variability from Climate Model Simulations. *Journal of Climate*, 30(23):9555–9573.

Olonscheck, D., Schurer, A. P., Lücke, L., and Hegerl, G. C. (2021). Large-scale emergence of regional changes in year-to-year temperature variability by the end of the 21st century. *Nature Communications*, 12(1):2–11.

O'Loughlin, J., Linke, A. M., and Witmer, F. D. W. (2014). Effects of temperature and precipitation variability on the risk of violence in sub-Saharan Africa, 1980-2012. *Proceedings of the National Academy of Sciences of the United States of America*, 111(47):16712–16717.

O'Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., van Ruijven, B. J., van Vuuren, D. P., Birkmann, J., Kok, K., Levy, M., and Solecki, W. (2017). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environmental Change*, 42:169–180.

O'Neill, B. C., Tebaldi, C., van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., Meehl, G. A., Moss, R., Riahi, K., and Sanderson, B. M. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, 9(9):3461–3482.

Orbe, C., Rind, D., Jonas, J., Nazarenko, L., Faluvegi, G., Murray, L. T., Shindell, D. T., Tsigaridis, K., Zhou, T., Kelley, M., and Schmidt, G. A. (2020). GISS Model E2.2: A Climate Model Optimized for the Middle Atmosphere—2. Validation of Large-Scale Transport and Evaluation of Climate Response. *Journal of Geophysical Research: Atmospheres*, 125(24):e2020JD033151.

Osborn, T. J. and Hulme, M. (1997). Development of a Relationship between Station and Grid-Box Rainday Frequencies for Climate Model Evaluation. *Journal of Climate*, 10(8):1885–1908.

Osman, M. B., Tierney, J. E., Zhu, J., Tardif, R., Hakim, G. J., King, J., and Poulsen, C. J. (2021). Globally resolved surface temperatures since the Last Glacial Maximum. *Nature*, 599(7884):239–244.

Otterå, O. H., Bentsen, M., Drange, H., and Suo, L. (2010). External forcing as a metronome for Atlantic multidecadal variability. *Nature Geoscience*, 3(10):688–694.

Otto-Bliesner, B. L., Tomas, R., Brady, E. C., Ammann, C., Kothavala, Z., and Clauzet, G. (2006). Climate Sensitivity of Moderate- and Low-Resolution Versions of CCSM3 to Preindustrial Forcings. *Journal of Climate*, 19(11):2567–2583.

PAGES Hydro2k Consortium, . (2017). Comparing proxy and model estimates of hydroclimate variability and change over the Common Era. *Climate of the Past*, 13(12):1851–1900.

PAGES2k-PMIP3 group (2015). Continental-scale temperature variability in PMIP3 simulations and PAGES 2k regional temperature reconstructions over the past millennium. *Climate of the Past*, 11(12):1673–1699.

Parker, D. E., Legg, T. P., and Folland, C. K. (1992). A new daily central England temperature series, 1772–1991. *International Journal of Climatology*, 12(4):317–342.

Parsons, L. A., Loope, G. R., Overpeck, J. T., Ault, T. R., Stouffer, R., and Cole, J. E. (2017). Temperature and precipitation variance in CMIP5 simulations and paleoclimate records of the last millennium. *Journal of Climate*, 30(22):8885–8912.

Past Interglacials Working Group of PAGES (2016). Interglacials of the last 800,000 years. *Reviews of Geophysics*, 54(1):162–219.

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

Peltier, W. R., Argus, D. F., and Drummond, R. (2015). Space geodesy constrains ice age terminal deglaciation: The global ICE-6G_C (VM5a) model. *Journal of Geophysical Research: Solid Earth*, 120:450–487.

Pendergrass, A. G., Knutti, R., Lehner, F., Deser, C., and Sanderson, B. M. (2017). Precipitation variability increases in a warmer climate. *Scientific Reports*, 7(December):1–9.

Perron, M. and Sura, P. (2013). Climatology of Non-Gaussian Atmospheric Statistics. *Journal of Climate*, 26(3):1063–1083.

Philippon, G., Ramstein, G., Charbit, S., Kageyama, M., Ritz, C., and Dumas, C. (2006). Evolution of the Antarctic ice sheet throughout the last deglaciation: A study with a new coupled climate—north and south hemisphere ice sheet model. *Earth and Planetary Science Letters*, 248(3):750–758.

Pollard, D. and DeConto, R. M. (2009). Modelling West Antarctic ice sheet growth and collapse through the past five million years. *Nature*, 458(7236):329–332.

Poulsen, C. J., Tabor, C., and White, J. D. (2015). Long-term climate forcing by atmospheric oxygen concentrations. *Science*, 348(6240):1238–1241.

Poulton, S. W., Bekker, A., Cumming, V. M., Zerkle, A. L., Canfield, D. E., and Johnston, D. T. (2021). A 200-million-year delay in permanent atmospheric oxygenation. *Nature*, 592(7853):232–236.

Prange, M., Jonkers, L., Merkel, U., Schulz, M., and Bakker, P. (2023). A multicentennial mode of North Atlantic climate variability throughout the Last Glacial Maximum. *Science Advances*, 9(44):eadh1106.

Räisänen, J. (2002). CO2-Induced Changes in Interannual Temperature and Precipitation Variability in 19 CMIP2 Experiments. *Journal of Climate*, 15(17):2395–2411.

Rasp, S., Hoyer, S., Merose, A., Langmore, I., Battaglia, P., Russell, T., Sanchez-Gonzalez, A., Yang, V., Carver, R., Agrawal, S., Chantry, M., Ben Bouallegue, Z., Dueben, P., Bromberg, C., Sisk, J., Barrington, L., Bell, A., and Sha, F. (2024). WeatherBench 2: A Benchmark for the Next Generation of Data-Driven Global Weather Models. *Journal of Advances in Modeling Earth Systems*, 16(6):e2023MS004019.

Rehfeld, K., Hébert, R., Lora, J. M., Lofverstrom, M., and Brierley, C. M. (2020). Variability of surface climate in simulations of past and future. *Earth System Dynamics*, 11(2):447–468.

Rehfeld, K., Münch, T., Ho, S. L., and Laepple, T. (2018). Global patterns of declining temperature variability from the Last Glacial Maximum to the Holocene. *Nature*, 554(7692):356–359.

Risi, C., Bony, S., Vimeux, F., and Jouzel, J. (2010). Water-stable isotopes in the LMDZ4 general circulation model: Model evaluation for present-day and past climates and applications to climatic interpretations of tropical isotopic records. *Journal of Geophysical Research Atmospheres*, 115(12):1–27.

Robock, A. and Mao, J. (1992). Winter warming from large volcanic eruptions. *Geophysical Research Letters*, 19(24):2405–2408.

Roche, D. M. (2013). 18O water isotope in the iLOVECLIM model (version 1.0) - Part 1: Implementation and verification. *Geoscientific Model Development*, 6(5):1481–1491.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Manzini, E., Schlese, U., and Schulzweida, U. (2006). Sensitivity of Simulated Climate to Horizontal and Vertical Resolution in the ECHAM5 Atmosphere Model. *Journal of Climate*, 19(16):3771–3791.

Ruddiman, W. F. (2008). *Earth's Climate Past and Future*. W. H. Freeman and Company, New York, New York, USA.

Ruff, T. W. and Neelin, J. D. (2012). Long tails in regional surface temperature probability distributions with implications for extremes under global warming. *Geophysical Research Letters*, 39(4).

Santer, B. D., Solomon, S., Bonfils, C., Zelinka, M. D., Painter, J. F., Beltran, F., Fyfe, J. C., Johannesson, G., Mears, C., Ridley, D. A., Vernier, J.-P., and Wentz, F. J. (2015). Observed multivariable signals of late 20th and early 21st century volcanic activity. *Geophysical Research Letters*, 42(2):500–509.

Schär, C. and Jendritzky, G. (2004). Hot news from summer 2003. *Nature*, 432(7017):559–560.

Schär, C., Vidale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A., and Appenzeller, C. (2004). The role of increasing temperature variability in European summer heatwaves. *Nature*, 427(6972):332–336.

Schillinger, M., Ellerhoff, B., Scheichl, R., and Rehfeld, K. (2022). Separating internal and externally forced contributions to global temperature variability using a Bayesian stochastic energy balance framework. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(11):113146.

Schlebusch, C. M., Malmström, H., Günther, T., Sjödin, P., Coutinho, A., Edlund, H., Munters, A. R., Vicente, M., Steyn, M., Soodyall, H., Lombard, M., and Jakobsson, M. (2017). Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science*, 358(6363):652–655.

Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J.-C., Hannay, C., Molod, A., Neale, R. B., and Saha, S. (2017). Practice and philosophy of climate model tuning across six US modeling centers. *Geoscientific Model Development*, 10(9):3207–3223.

Schurer, A. P., Tett, S. F. B., and Hegerl, G. C. (2014). Small influence of solar variability on climate over the past millennium. *Nature Geoscience*, 7(2):104–108.

Schwander, J., Jouzel, J., Hammer, C. U., Petit, J.-R., Udisti, R., and Wolff, E. (2001). A tentative chronology for the EPICA Dome Concordia Ice Core. *Geophysical Research Letters*, 28(22):4243–4246.

Schwarzwald, K. and Lenssen, N. (2022). The importance of internal climate variability in climate impact projections. *Proceedings of the National Academy of Sciences*, 119(42):e2208095119.

Seddon, A. W., Macias-Fauria, M., Long, P. R., Benz, D., and Willis, K. J. (2016). Sensitivity of global terrestrial ecosystems to climate variability. *Nature*, 531(7593):229–232.

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. (2019). UKESM1: Description and Evaluation of the U.K. Earth System Model. *Journal of Advances in Modeling Earth Systems*, 11(12):4513–4558.

Sellar, A. A., Walton, J., Jones, C. G., Wood, R., Abraham, N. L., Andrejczuk, M., Andrews, M. B., Andrews, T., Archibald, A. T., de Mora, L., Dyson, H., Elkington, M., Ellis, R., Florek, P., Good, P., Gohar, L., Haddad, S., Hardiman, S. C., Hogan, E., Iwi, A., Jones, C. D., Johnson, B., Kelley, D. I., Kettleborough, J., Knight, J. R., Köhler, M. O., Kuhlbrodt, T., Liddicoat, S., Linova-Pavlova, I., Mizielinski, M. S., Morgenstern, O., Mulcahy, J., Neininger, E., O'Connor, F. M., Petrie, R., Ridley, J., Rioual, J.-C., Roberts, M., Robertson, E., Rumbold, S., Seddon, J., Shepherd, H., Shim, S., Stephens, A., Teixiera, J. C., Tang, Y., Williams, J., Wiltshire, A., and Griffiths, P. T. (2020). Implementation of U.K. Earth System Models for CMIP6. *Journal of Advances in Modeling Earth Systems*, 12(4):e2019MS001946.

Seneviratne, S., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Di Luca, A., Ghosh, S., Iskandar, I., Kossin, J., Lewis, S., Otto, F., Pinto, I., Satoh, M., Vicente-Serrano, S., Wehner, M., and Zhou, B. (2021). Weather and climate extreme events in a changing climate. In 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., editors, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, pages 1513–1766. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Sévellec, F. and Fedorov, A. V. (2015). Unstable AMOC during glacial intervals and millennial variability: The role of mean sea ice extent. *Earth and Planetary Science Letters*, 429(2015):60–68.

Shackleton, N. J. and Opdyke, N. D. (1973). Oxygen isotope and palaeomagnetic stratigraphy of Equatorial Pacific core V28-238: Oxygen isotope temperatures and ice volumes on a 105 year and 106 year scale. *Quaternary Research*, 3(1):39–55.

Shakun, J. D. and Carlson, A. E. (2010). A global perspective on Last Glacial Maximum to Holocene climate change. *Quaternary Science Reviews*, 29(15-16):1801–1816.

Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., Otto-Bliesner, B., Schmittner, A., and Bard, E. (2012). Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature*, 484(7392):49–54.

Shao, Z. G. and Ditlevsen, P. D. (2016). Contrasting scaling properties of interglacial and glacial climates. *Nature Communications*, 7:1–8.

Shi, J., Jiang, D., Tian, Z., and Lang, X. (2022). Enhanced Interannual Variability in Temperature during the Last Glacial Maximum. *Journal of Climate*, 35(18):5933–5950.

Shi, J., Tian, Z., Lang, X., and Jiang, D. (2024). Projected changes in the interannual variability of surface air temperature using CMIP6 simulations. *Climate Dynamics*, 62(1):431–446.

Simolo, C. and Corti, S. (2022). Quantifying the role of variability in future intensification of heat extremes. *Nature Communications*, 13(1):7930.

Skelton, A., Kirchner, N., and Kockum, I. (2020). Skewness of Temperature Data Implies an Abrupt Change in the Climate System Between 1985 and 1991. *Geophysical Research Letters*, 47(23):1–10.

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.-C., 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. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, 145(724):2876–2908.

Smithies, R. H., Lu, Y., Kirkland, C. L., Johnson, T. E., Mole, D. R., Champion, D. C., Martin, L., Jeon, H., Wingate, M. T. D., and Johnson, S. P. (2021). Oxygen isotopes trace the origins of Earth's earliest continental crust. *Nature*, 592(7852):70–75.

Snoll, B., Ivanovic, R., Gregoire, L., Sherriff-Tadano, S., Menviel, L., Obase, T., Abe-Ouchi, A., Bouttes, N., He, C., He, F., Kapsch, M., Mikolajewicz, U., Muglia, J., and Valdes, P. (2024). A multi-model assessment of the early last deglaciation (PMIP4 LDv1): A meltwater perspective. *Climate of the Past*, 20(4):789–815.

Steinman, B. A., Mann, M. E., and Miller, S. K. (2015). Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures. *Science*, 347(6225):988–991.

Stenseth, N. C., Mysterud, A., Ottersen, G., Hurrell, J. W., Chan, K.-S., and Lima, M. (2002). Ecological Effects of Climate Fluctuations. *Science*, 297(5585):1292–1296.

Stevenson, S., Fasullo, J. T., Otto-Bliesner, B. L., Tomas, R. A., and Gao, C. (2017). Role of eruption season in reconciling model and proxy responses to tropical volcanism. *Proceedings of the National Academy of Sciences*, 114(8):1822–1826.

Stocker, T. F. (2011). *Introduction to Climate Modelling*. Springer.

Stokes, C. R., Tarasov, L., Blomdin, R., Cronin, T. M., Fisher, T. G., Gyllencreutz, R., Hättestrand, C., Heyman, J., Hindmarsh, R. C. A., Hughes, A. L. C., Jakobsson, M., Kirchner, N., Livingstone, S. J., Margold, M., Murton, J. B., Noormets, R., Peltier, W. R., Peteet, D. M., Piper, D. J. W., Preusser, F., Renssen, H., Roberts, D. H., Roche, D. M., Saint-Ange, F., Stroeven, A. P., and Teller, J. T. (2015). On the reconstruction of palaeo-ice sheets: Recent advances and future challenges. *Quaternary Science Reviews*, 125:15–49.

Stringer, C. and Galway-Witham, J. (2017). On the origin of our species. *Nature*, 546(7657):212–214.

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.

(2019). The Canadian Earth System Model version 5 (CanESM5.0.3). *Geoscientific Model Development*, 12(11):4823–4873.

Swingedouw, D., Mignot, J., Ortega, P., Khodri, M., Menegoz, M., Cassou, C., and Hanquiez, V. (2017). Impact of explosive volcanic eruptions on the main climate variability modes. *Global and Planetary Change*, 150:24–45.

Swingedouw, D., Ortega, P., Mignot, J., Guilyardi, E., Masson-Delmotte, V., Butler, P. G., Khodri, M., and Séférian, R. (2015). Bidecadal North Atlantic ocean circulation variability controlled by timing of volcanic eruptions. *Nature Communications*, 6(1):6545.

Takayabu, I., Rasmussen, R., Nakakita, E., Prein, A., Kawase, H., Watanabe, S.-I., Adachi, S. A., Takemi, T., Yamaguchi, K., Osakada, Y., and Wu, Y.-H. (2022). Convection-Permitting Models for Climate Research. *Bulletin of the American Meteorological Society*, 103(1):E77–E82.

Tamarin-Brodsky, T., Hodges, K., Hoskins, B. J., and Shepherd, T. G. (2020). Changes in Northern Hemisphere temperature variability shaped by regional warming patterns. *Nature Geoscience*, 13(6):414–421.

Tamarin-Brodsky, T., Hodges, K., Hoskins, B. J., and Shepherd, T. G. (2022). A Simple Model for Interpreting Temperature Variability and Its Higher-Order Changes. *Journal of Climate*, 35(1):387–403.

Tarasov, L., Dyke, A. S., Neal, R. M., and Peltier, W. R. (2012). A data-calibrated distribution of deglacial chronologies for the North American ice complex from glaciological modeling. *Earth and Planetary Science Letters*, 315–316:30–40.

Tierney, J. E., Zhu, J., King, J., Malevich, S. B., Hakim, G. J., and Poulsen, C. J. (2020). Glacial cooling and climate sensitivity revisited. *Nature*, 584(7822):569–573.

Timmreck, C. (2012). Modeling the climatic effects of large explosive volcanic eruptions. *WIREs Climate Change*, 3(6):545–564.

Tindall, J. C., Valdes, P. J., and Sime, L. C. (2009). Stable water isotopes in HadCM3: Isotopic signature of El Niño-Southern Oscillation and the tropical amount effect. *Journal of Geophysical Research Atmospheres*, 114(4):1–12.

Toohey, M., Krüger, K., Niemeier, U., and Timmreck, C. (2011). The influence of eruption season on the global aerosol evolution and radiative impact of tropical volcanic eruptions. *Atmospheric Chemistry and Physics*, 11(23):12351–12367.

Trenberth, K. E. and Dai, A. (2007). Effects of Mount Pinatubo volcanic eruption on the hydrological cycle as an analog of geoengineering. *Geophysical Research Letters*, 34(15).

Valdes, P. (2011). Built for stability. *Nature Geoscience*, 4(7):414–416.

Van Oldenborgh, G. J., Wehner, M. F., Vautard, R., Otto, F. E. L., Seneviratne, S. I., Stott, P. A., Hegerl, G. C., Philip, S. Y., and Kew, S. F. (2022). Attributing and Projecting Heatwaves Is Hard: We Can Do Better. *Earth's Future*, 10(6):e2021EF002271.

Villmoare, B., Kimbel, W. H., Seyoum, C., Campisano, C. J., DiMaggio, E. N., Rowan, J., Braun, D. R., Arrowsmith, J. R., and Reed, K. E. (2015). Early Homo at 2.8 Ma from Ledi-Geraru, Afar, Ethiopia. *Science*, 347(6228):1352–1355.

Volodin, E. M. and Yurova, A. Y. (2013). Summer temperature standard deviation, skewness and strong positive temperature anomalies in the present day climate and under global warming conditions. *Climate Dynamics*, 40(5):1387–1398.

von Storch, H. and Zwiers, F. W. (1999). *Statistical Analysis in Climate Research*. Cambridge University Press.

Wade, D. C., Abraham, N. L., Farnsworth, A., Valdes, P. J., Bragg, F., and Archibald, A. T. (2019). Simulating the climate response to atmospheric oxygen variability in the Phanerozoic: A focus on the Holocene, Cretaceous and Permian. *Climate of the Past*, 15(4):1463–1483.

Walker, M., Johnsen, S., Rasmussen, S. O., Popp, T., Steffensen, J.-P., Gibbard, P., Hoek, W., Lowe, J., Andrews, J., Björck, S., Cwynar, L. C., Hughen, K., Kershaw, P., Kromer, B., Litt, T., Lowe, D. J., Nakagawa, T., Newnham, R., and Schwander, J. (2009). Formal definition and dating of the GSSP (Global Stratotype Section and Point) for the base of the Holocene using the Greenland NGRIP ice core, and selected auxiliary records. *Journal of Quaternary Science*, 24(1):3–17.

Wang, T., Miao, J., Wang, H., and Sun, J. (2021). Influence of Strong Tropical Volcanic Eruptions on Daily Temperature and Precipitation Extremes Across the Globe. *Journal of Meteorological Research*, 35(3):428–443.

Weitzel, N., Andres, H., Baudouin, J.-P., Kapsch, M.-L., Mikolajewicz, U., Jonkers, L., Bothe, O., Ziegler, E., Kleinen, T., Paul, A., and Rehfeld, K. (2024). Towards spatio-temporal comparison of simulated and reconstructed sea surface temperatures for the last deglaciation. *Climate of the Past*, 20(4):865–890.

Werner, M., Langebroek, P. M., Carlsen, T., Herold, M., and Lohmann, G. (2011). Stable water isotopes in the ECHAM5 general circulation model: Toward high-resolution isotope modeling on a global scale. *Journal of Geophysical Research Atmospheres*, 116(15):1–14.

Westerhold, T., Marwan, N., Drury, A. J., Liebrand, D., Agnini, C., Anagnostou, E., Barnet, J. S. K., Bohaty, S. M., De Vleeschouwer, D., Florindo, F., Frederichs, T., Hodell, D. A., Holbourn, A. E., Kroon, D., Lauretano, V., Littler, K., Lourens, L. J., Lyle, M., Pälike, H., Röhl, U., Tian, J., Wilkens, R. H., Wilson, P. A., and Zachos, J. C. (2020). An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science*, 369(6509):1383–1387.

Wetherald, R. T. (2009). Changes of Variability in Response to Increasing Greenhouse Gases. Part II: Hydrology. *Journal of Climate*, 22(22):6089–6103.

Whitehouse, P. L., Bentley, M. J., and Le Brocq, A. M. (2012). A deglacial model for Antarctica: Geological constraints and glaciological modelling as a basis for a new model of Antarctic glacial isostatic adjustment. *Quaternary Science Reviews*, 32:1–24.

Wild, M., Folini, D., Schär, C., Loeb, N., Dutton, E. G., and König-Langlo, G. (2013). The global energy balance from a surface perspective. *Climate Dynamics*, 40(11-12):3107–3134.

Wilde, S. A., Valley, J. W., Peck, W. H., and Graham, C. M. (2001). Evidence from detrital zircons for the existence of continental crust and oceans on the Earth 4.4 Gyr ago. *Nature*, 409(6817):175–178.

Willeit, M., Ilyina, T., Liu, B., Heinze, C., Perrette, M., Heinemann, M., Dalmonech, D., Brovkin, V., Munhoven, G., Börker, J., Hartmann, J., Romero-Mujalli, G., and Ganopolski, A. (2023). The Earth system model CLIMBER-X v1.0 – Part 2: The global carbon cycle. *Geoscientific Model Development*, 16(12):3501–3534.

WMO (2024). *State of the Global Climate 2023*. World Meteorological Organization,.

Yokoyama, Y., Esat, T. M., Thompson, W. G., Thomas, A. L., Webster, J. M., Miyairi, Y., Sawada, C., Aze, T., Matsuzaki, H., Okuno, J., Fallon, S., Braga, J.-C., Humblet, M., Iryu, Y., Potts, D. C., Fujita, K., Suzuki, A., and Kan, H. (2018). Rapid glaciation and a two-step sea level plunge into the Last Glacial Maximum. *Nature*, 559(7715):603–607.

Young, G. M., Brunn, V. v., Gold, D. J. C., and Minter, W. E. L. (1998). Earth's Oldest Reported Glaciation: Physical and Chemical Evidence from the Archean Mozaan Group (~2.9 Ga) of South Africa. *The Journal of Geology*, 106(5):523–538.

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., Tsujino, H., Deushi, M., Tanaka, T., Hosaka, M., Yabu, S., Yoshimura, H., Shindo, E., Mizuta, R., Obata, A., Adachi, Y., and Ishii, M. (2019). The Meteorological Research Institute Earth System Model Version 2.0, MRI-ESM2.0: Description and Basic Evaluation of the Physical Component. *Journal of the Meteorological Society of Japan. Ser. II*, 97(5):931–965.

Zachos, J. C., Dickens, G. R., and Zeebe, R. E. (2008). An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. *Nature*, 451(7176):279–283.

Zanchettin, D., Bothe, O., Graf, H. F., Lorenz, S. J., Luterbacher, J., Timmreck, C., and Jungclaus, J. H. (2013). Background conditions influence the decadal climate response to strong volcanic eruptions. *Journal of Geophysical Research Atmospheres*, 118(10):4090–4106.

Zanchettin, D., Bothe, O., Lehner, F., Ortega, P., Raible, C. C., and Swingedouw, D. (2015). Reconciling reconstructed and simulated features of the winter Pacific/North American pattern in the early 19th century. *Climate of the Past*, 11(6):939–958.

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., Hegerl, G., Robock, A., Pausata, F. S. R., Ball, W. T., Bauer, S. E., Bekki, S., Dhomse, S. S., LeGrande, A. N., Mann, G. W., Marshall, L., Mills, M., Marchand, M., Niemeier, U., Poulain, V., Rozanov, E., Rubino, A., Stenke, A., Tsigaridis, K., and Tummon, F. (2016). The Model Intercomparison Project on the climatic response toVolcanic forcing (VolMIP): Experimental design and forcing input data for CMIP6. *Geoscientific Model Development*, 9(8):2701–2719.

Zeebe, R. E., Ridgwell, A., and Zachos, J. C. (2016). Anthropogenic carbon release rate unprecedented during the past 66 million years. *Nature Geoscience*, 9(4):325–329.

Zhu, F., Emile-Geay, J., McKay, N. P., Hakim, G. J., Khider, D., Ault, T. R., Steig, E. J., Dee, S., and Kirchner, J. W. (2019). Climate models can correctly simulate the continuum of global-average temperature variability. *Proceedings of the National Academy of Sciences of the United States of America*, 116(18):8728–8733.

Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Poulsen, C. J., Tierney, J. E., Lofverstrom, M., and DiNezio, P. (2021). Assessment of Equilibrium Climate Sensitivity of the Community Earth System Model Version 2 Through Simulation of the Last Glacial Maximum. *Geophysical Research Letters*, 48(3):e2020GL091220.

Zhuang, K., North, G. R., and Stevens, M. J. (2017). A NetCDF version of the two-dimensional energy balance model based on the full multigrid algorithm. *SoftwareX*, 6(2017):198–202.

Zorita, E., Moberg, A., Leijonhufvud, L., Wilson, R., Brázdil, R., Dobrovolný, P., Luterbacher, J., Böhm, R., Pfister, C., Riemann, D., Glaser, R., Söderberg, J., and González-Rouco, F. (2010). European temperature records of the past five centuries based on documentary/instrumental information compared to climate simulations. *Climatic Change*, 101(1):143–168.