Probing the timescale dependency of local and global variations in surface air temperature from climate simulations and reconstructions of the last millennia

Beatrice Ellerhoff [1,*] and Kira Rehfeld [1,2,†]

[1] *Institute of Environmental Physics, Ruprecht-Karls-Universität Heidelberg, INF 229, 69120 Heidelberg, Germany*

[2] *Geo- und Umweltforschungszentrum (GUZ), Universität Tübingen, Schnarrenbergstr. 94-96, 72076 Tübingen, Germany*

(Received 2 February 2021; revised 15 June 2021; accepted 9 December 2021; published 27 December 2021)

Earth's climate can be understood as a dynamical system that changes due to external forcing and internal couplings. Essential climate variables, such as surface air temperature, describe this dynamics. Our current interglacial, the Holocene (11 700 yr ago to today), has been characterized by small variations in global mean temperature prior to anthropogenic warming. However, the mechanisms and spatiotemporal patterns of fluctuations around this mean, called temperature variability, are poorly understood despite their socioeconomic relevance for climate change mitigation and adaptation. Here we examine discrepancies between temperature variability from model simulations and paleoclimate reconstructions by categorizing the scaling behavior of local and global surface air temperature on the timescale of years to centuries. To this end, we contrast power spectral densities (PSD) and their power-law scaling using simulated and observation-based temperature series of the last 6000 yr. We further introduce the spectral gain to disentangle the externally forced and internally generated variability as a function of timescale. It is based on our estimate of the joint PSD of radiative forcing, which exhibits a scale break around the period of 7 yr. We find that local temperature series from paleoclimate reconstructions show a different scaling behavior than simulated ones, with a tendency towards stronger persistence (i.e., correlation between successive values within a time series) on periods of 10 to 200 yr. Conversely, the PSD and spectral gain of global mean temperature are consistent across data sets. Our results point to the limitation of climate models to fully represent local temperature statistics over decades to centuries. By highlighting the key characteristics of temperature variability, we pave a way to better constrain possible changes in temperature variability with global warming and assess future climate risks.

DOI: 10.1103/PhysRevE.104.064136

I. INTRODUCTION

The variability of surface air temperature is present on all spatial and temporal scales, from synoptic and seasonal changes to long-term variations on periods of years to multimillennia. On the one hand, it arises from internal processes, such as the El Niño-Southern Oscillation (ENSO) [1]. On the other hand, the temperature varies due to external forcing, such as the greenhouse effect [2,3]. Understanding the internally generated and externally forced variability has been suggested to be at least as necessary for evaluating climate risks for society and ecosystems as projecting the global mean temperature [4]. Available instrumental observations are limited to a small time span, leading to challenges in quantifying temperature variability. Paleoclimate reconstructions extend the characterization of temperature variability and can be compared to global circulation models (GCMs) [5–7]. However, discrepancies between model and paleoclimate data remain to be resolved, especially on the local level and on periods between years and centuries [8–12].

Characterizing local temperature variability is crucial for predicting extremes [6], not only to minimize short-term damage but also to design long-term strategies, including urban planning and food cultivation [13]. Variability of global temperature on periods above years is relevant to the understanding of long-term changes [14] as well as climate sensitivity [15]. Assessing the temporal correlation structure of temperature series by means of scaling behavior and persistence is particularly important for distinguishing externally forced trends from natural changes [16]. It could affect the confidence in future projections and attribution studies [17,18]. Therefore, one of the main topics to be investigated here is the characteristics of local and global temperature variability on periods of years to centuries from model simulations and observation-based data of the last millennia.

To determine how the variability of a temperature series is distributed with timescales τ, we make use of the power spectral density (PSD) $S(\tau)$, known as spectrum. It can be obtained from the Fourier transform of the autocorrelation function (see Appendix A) [19,20]. The spectrum was shown to often follow a power law

$$S(\tau) \sim \tau^{\beta}, \tag{1}$$

with spectral exponent β and period τ [21,23–27], especially on decadal-to-centennial scales [22,28,29]. We refer to this behavior (1) as temporal scaling since the temperature signal has no preferred timescale and is statistically similar across periods τ. The exact determination of the start and end points of a scaling interval is not part of this study.

*beatrice.ellerhoff@iup.uni-heidelberg.de

†Now at Geo- und Umweltforschungszentrum (GUZ), Universität Tübingen, Schnarrenbergstr. 94-96, 72076 Tübingen, Germany.

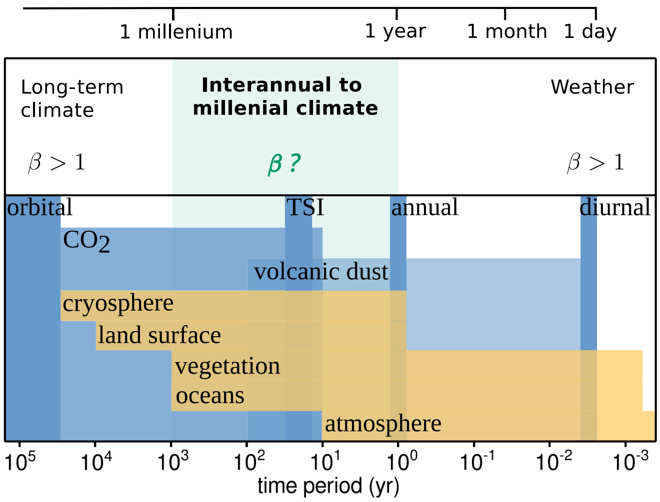

FIG. 1. Characteristic timescales relevant to surface air temperature variability of climatic drivers (dark blue) and climate subsystems (yellow) [32–34]. The weather and long-term climate is characterized by $\beta > 1$ for local and global mean temperature. On interannual to millennial timescales the statistical properties of temperature fluctuations remain to be determined, especially at the local scale. The TSI bar highlights the dominant variations in recent total solar irradiance observations.

Long-range memory stochastic processes are suitable to describe temperature signals with temporal scaling [25,26,30]. Among those, fractional Gaussian noise (fGn) is a stationary process and exhibits a spectral exponent $\beta \in (-1, 1)$ on sufficiently long periods (see Appendix B). Fractional Brownian motion (fBm) is a nonstationary process that shows $\beta \in (1, 3)$. The scaling exponent β relates to the decay of the autocovariance function (Appendix B) and indicates how strongly the values within a time series are correlated (or anticorrelated). It is therefore regarded as a measure of the strength of temporal persistence (or antipersistence) [26,31].

Particular scaling behavior with $\beta \approx 2$ [23,25] is typical for the weather regime (hours to weeks) and can be explained by atmospheric turbulence [35,36]. In the long-term climate, regional and global mean temperatures show similar behavior ($\beta > 1$) [24,25,29] due to the presence of nonlinear processes, such as the temperature-albedo feedback [37]. On timescales between years to millennia, the temperature is constantly influenced by the interaction of all climate subsystems and by volcanic, solar, as well as CO_2 forcing (Fig. 1). Estimates of the spatially dependent scaling behavior of local temperature on these timescales differ [22,25]. On the global scale, many studies find $\beta \approx 1$ [28,29]. However, Lovejoy *et al.* has identified a change from the so-called macroweather regime ($\beta \approx 0.8$ on periods of 10 days to 40 yr) to the climate regime ($\beta \approx 1.8$ on periods from 40 yr to 80 000 yr) [25].

In this manner, previous works find ambiguity in the interpretation of local and global temperature scaling, and it remains to be determined whether simulations and reconstructions qualitatively agree in scaling behavior $\beta < 1$ or $\beta > 1$. The so-called "$1/f$ noise" ($\beta = 1$) corresponds to a process with power spectral density proportional to the period. For $\beta > 1$, the relative contribution

$$\frac{\int_{f'/2}^{f'} S(f)\,df}{\int_{f'}^{2f'} S(f)\,df} = \frac{1 - 2^{\beta-1}}{2^{1-\beta} - 1} = 2^{\beta-1} \tag{2}$$

to the variance is larger from slow timescales compared to faster ones for all frequency intervals $f'/2$ to $2f'$ within a scaling interval [38]. With increasing $\beta > 1$, the fBM is said to exhibit "nonlinear pseudotrends" [30] (Appendix B). Thus, for understanding climate variability and for modeling purposes, the systematic estimate of the scaling exponent β allows to assess the behavior of fluctuation levels across timescales [26]. Moreover, the differentiation between forced and unforced changes poses a challenge to understanding temperature variability [39,40]. Beyond the analysis of Haar fluctuations of a few forcing reconstructions [41–43], spectral analysis of climatic drivers and their frequency-dependent linkage to the temperature response remains incomplete.

We investigate the timescale dependency of local and global surface air temperature variability by analyzing power spectral densities from a few hours to a thousand years, thereby extending and improving on earlier work [24,25,29,44]. We use model simulations and observation-based data, which we introduce in Sec. II. To estimate the PSD and determine its power-law scaling on periods of 10 to 200 yr, we use state-of-the-art methods described in Sec. III. This allows us to contrast regional and global spectra (Sec. IV A), spatial patterns (Sec. IV B), and the agreement of simulated and observation-based estimates (Sec. III C). Along with that, we discuss the joint PSD from various radiative forcings, which allows us to calculate the spectral gain and study the externally forced variability in Sec. III D. Based on our reconstruction of the PSD of surface air temperature for the last millennia, we evaluate the consistency of spectral characteristics across the data sets considered. In Sec. V we elaborate on the stronger persistence of temperature on local than global level as well as remaining discrepancies. Finally, we discuss how our findings could help improve climate model simulations and understand Earth's climate dynamics.

II. DATA

We investigate the timescale-dependent distribution of surface air temperature variability using model simulations, observation-based data, and radiative forcing reconstructions. The model simulations include ten transient runs from GCM experiments [45]. The observation-based data consists of reanalysis data, instrumental measurements, and the paleoclimate reconstructions from the Past Global Changes 2k (PAGES2k) network [46]. We use 12 reconstructions of climatic drivers, including solar, volcanic, orbital, and CO_2 forcing. All temperature and radiative forcing signals are specific to the Mid- and Late-Holocene (the last 6000 yr), with a focus on the Common Era (0 to 2000 CE). The supplemental tables S1–S3 [47] summarize their key specifications.

A. Model simulations

Each of the ten GCM runs considered features a transient, albeit different forcing and a comparable spatiotemporal resolution. The CESM-LME 1 [48] and MPI-M LM [49] experiments serve as representative runs of the last millennium. We analyze them at two temporal resolutions (one month, six hours) to capture both the high- and low-frequency variability within our available computing capacities (see Fig. S7 [47]). CESM 1 past 2k [50] is included as a slightly newer run for the Common Era. To cover the Mid-Holocene, we use simulations from the IPSL [51] (denoted IPSL-p6k) and ECHAM5/MPI-OM [52] (denoted ECH5/MPIOM-p6k) of the last 6000 yr. From the TraCE-21k [53] simulation, we also consider only the last 6000 yr to retain comparability and to avoid potential biases due to significant shifts in the mean state of climate. The Mid-Holocene runs were temporally averaged to a bimonthly resolution to reduce computational costs. To test for the influence of human-induced climate change on our results, we include the HadCM3 LM1 simulation [12], covering the period from 850 to 1850 CE. Furthermore, we compare our results to the pre-industrial (PI) control runs from CESM-LME 1 and MPI-M LM, as well as the TraCE-21k-ORB run, which is solely forced by orbital changes.

B. Observation-based data

In addition to the simulations, we analyze the monthly resolved HadCRUT4 (Hadley Centre/Climatic Research Unit Temperature) instrumental records, ranging from 1850 to 2019 [54]. However, most of the grid-box time series are not available as continuous measurements as required for spectral analysis. Therefore, we retain only those 104 grid boxes with coverage greater than 150 yr after interpolating gaps of up to two months. While the Northern Hemisphere is comparatively well covered up to 72.5°N, only nine grid boxes remain for the Southern Hemisphere. Therefore, this selection comes at the expense of spatial resolution but offers a higher spectral resolution on longer timescales. To further explore the potential effect of these spatiotemporal constraints, we include the ERA5 (European Centre for Medium-Range Weather Forecasts Reanalysis 5th generation) temperature reanalysis for the years 1979 to 2019 [55]. Along with CESM-LME 1 and MPI-M LM, we analyze the ERA5 data at both six-hourly and monthly resolution for the same reasons as mentioned earlier.

In addition to direct temperature observations and reanalysis, we analyze paleoclimate data. Paleoclimate records hold preserved biological, chemical, and physical tracers ("proxies") of past climate. The number of temperature records from paleoclimate data with subcentennial resolution is limited. Recent progress has been made by improved calibration and pseudoproxy methods within the PAGES2k network [56]. Therefore, we base our analysis on their newest global multiproxy database for temperature reconstructions of the Common Era [46]. It gathers 692 records from trees, ice, sediment, corals, speleothems, and documentary evidence with a resolution between weeks and centuries. The records are spread over 648 locations, including all continental regions and major ocean basins.

For investigating the variability of global mean surface temperature, we use the seven spatially weighted statistical reconstructions for the last 2000 yr provided by PAGES2k [46]. To estimate the mean of local spectra, we choose records from the PAGES2k database according to their resolution ($\leqslant$80 yr), their number of data points ($\geqslant$20), their coverage ($\geqslant$20 yr), as well as their maximum hiatuses ($\leqslant$160 yr). To reliably deduce the scaling of the PSD from individual records, we select the records according to our scales of interest (Table I), similar to [26,57]. Ice core records were excluded from our analysis since they require additional consideration of signal-to-noise ratios at the subcentennial timescales [58,59].

TABLE I. Requirements on irregularly sampled time series $x(t)$ for analyzing power-law scaling on timescales $\tau \in [\tau_1, \tau_2]$. We apply this scheme for $\tau_1 = 10$ and $\tau_2 = 200$ yr in Secs. IV B and IV C.

Parameter	Value
Number of data points (N)	$\geqslant 50$
Mean temporal resolution [$\langle t_{i+1} - t_i \rangle$]	$\leqslant \tau_1$
Coverage ($t_N - t_1$)	$\geqslant 3\tau_2$
Length of hiatuses [$\max(t_{i+1} - t_i)$]	$\leqslant 5\tau_1$

C. Radiative forcing

External forcing contributes significantly to temperature variability and is an essential part of reliable climate projections [40,60,61]. We study its spectral properties using forcing reconstructions, widely implemented in GCM experiments and coordinated within the Palaeoclimate Model Intercomparison Project (PMIP3/PMIP4) [62,63]. This includes five solar [64–68], one CO_2 [63], and two volcanic [60,69] forcing reconstructions as well as Berger's numerical solution for orbital forcing [70]. Furthermore, we calculate diurnal insolation changes from the hour angle of the sun [71]. We also use a more recently published volcanic [72] and high-resolution solar forcing [73] reconstruction as well as CO_2 measurements [74]. We neglect land-use forcing [75] which is much lower in amplitude and variability than the other forcings considered here.

All forcing reconstructions are rescaled to radiative forcing equivalents, which express their respective change in the Earth's radiation balance in Watts per square meter ($\mathrm{Wm^{-2}}$). We apply the widely used formula $5.35 \ln([CO_2]/278\mathrm{ppm})\,\mathrm{Wm^{-2}}$ to rescale CO_2 concentrations $[CO_2]$, given in parts per million (ppm) [76]. The stratospheric aerosol optical depth (AOD) from volcanic eruptions is rescaled by $(-20)^{-1}\ \mathrm{Wm^{-2}}/\mathrm{AOD}$ [77]; however, the optimal conversion factor is still a matter of debate [78]. Additional uncertainties arise from the wide spread of reconstructions for volcanic and solar forcing. To account for this and the choice of conversion factor, we simulate the joint PSD of radiative forcing by a Monte Carlo approach described in Appendix E. Here "joint" indicates that the PSD of radiative forcing is calculated by linear summation of the mean PSD from different types of climatic drivers, rescaled to their radiative forcing equivalents.

III. METHODS

Spectral analysis is the primary tool used here for studying the timescale-dependent variability and scaling of temperature series. To minimize uncertainties in the spectral analysis of proxy records, we use state-of-the-art approaches for irregularly sampled time series [79]. Statistical estimators further test for the agreement between simulations and paleoclimate data. We apply linear response theory to derive the spectral gain and investigate the forced temperature response.

A. Spectral analysis

Power spectral analysis requires the assumption that the underlying time series can be described as a weakly stationary, stochastic process with time-independent mean and autocovariance [80]. We therefore linearly detrend all time series as it is standard for temperature analysis [9,26,81,82]. The agreement of the PSD from disjoint time intervals in Fig. S13 [47] provides evidence that stationarity is sufficiently fulfilled. We use the multitaper method with three windows [83,84] and chi-square distributed uncertainties to compute the PSD. The two lowest frequencies were omitted to reduce biases of the multitaper method [24]. For visual purposes, we apply a logarithmic Gaussian smoothing filter of constant width (0.005 decibels) [85]. Mean spectra were calculated by interpolation to the lowest resolution, binning into equally spaced log-frequency intervals, and taking the average with equal weights [24]. This requires the statistical independence of the averaged values [43]. The spectral exponent β is calculated by linear regression to the logarithm of (1) on periods between $\tau_1 = 10$ and $\tau_2 = 200$ yr after binning the PSD into equally spaced log-frequency intervals to more uniformly weight the estimate and avoid low-frequency biases [24,26,29,86]. In the case of seven proxy records with an insufficient resolution, the scaling is estimated on their corresponding spectral resolution, but always at least between 20 and 200 yr (Fig. S4 [47]). The uncertainty of the spectral exponent, $\Delta\beta$, is given by the standard error of the linear regression model $\Delta\beta_{lm}$, except for irregularly temperature series.

B. Uncertainties for irregular temperature series

Spectral analysis of proxy records, which are typically not sampled in regular time steps, is more prone to errors than that of regular time series. We aim to minimize biases by accounting for the number of data points, temporal resolution, total coverage, and hiatuses' length when selecting the records (Table I). We find that the mean temporal resolution of a proxy record approximates well the optimal interpolation time step. Nevertheless, the interpolation introduces uncertainties which are not captured by $\Delta\beta_{lm}$. Similar to Laepple *et al.* [79], we quantify this additional uncertainty $\Delta\beta_{\mathrm{int}}$ in four steps: (1) For each record with spectral exponent β, we simulate $N = 100$ surrogate time series with annual resolution and a power-law scaling $\beta_n \approx \beta$ and $n \in [1, N]$. (2) We form the surrogate's block average over the proxy record's irregular time steps and obtain N surrogate time series at record resolution. (3) We interpolate the surrogate time series, calculate the multitaper spectrum, and extract the scaling exponent $\beta_{n,lm}$ from linear regression in the same way as for the proxy record (Fig. S8 [47]). (4) We calculate the mean deviation $\Delta\beta_{\mathrm{int}} = \frac{1}{N}\sum_{n=1}^{N} |\beta_{n,lm} - \beta_n|$ of the ensemble. The uncertainty of the individual fits $\Delta\beta_{n,lm}$ is negligible compared to the mean deviation $\Delta\beta_{\mathrm{int}}$. We obtain the uncertainty of the record's spectral exponent from both, the uncertainty of the initial fit $\Delta\beta_{lm}$ and due to interpolation $\Delta\beta_{\mathrm{int}}$ via quadratic summation: $\Delta\beta = \sqrt{(\Delta\beta_{lm})^2 + (\Delta\beta_{\mathrm{int}})^2}$.

C. Statistical analysis of spectral exponents

We quantify the agreement of simulated and reconstructed β-values using percent agreement, categorical agreement, and Kappa statistics. Beforehand, we extract the simulated temperature at the proxy record location by bilinear interpolation of neighboring grid boxes to achieve the best possible comparability between record and simulation. Percent agreement p_0 gives the percentage of locations at which the confidence range $\beta \pm \Delta\beta$ from simulation and reconstruction overlap. The agreement by category, here referred to as categorical agreement p_c, is calculated with the help of $\nu = 0.32$, the mean uncertainty of β from all proxy records considered. We then assign the three categories *low* ($\beta < 1 - \nu$), *high* ($1 + \nu \leqslant \beta$), and *intermediate* ($1 - \nu \leqslant \beta < 1 + \nu$) to the spectral exponent β. The *intermediate* regime prevents incorrect assignment. To verify the reliability of categorical agreement, we calculate the kappa statistics

$$\kappa = (p_c - p_e)/(1 - p_e) \tag{3}$$

with expected percent agreement p_e by category [87]. The latter can be obtained from $p_e = \frac{1}{N^2}\sum_{c=1}^{3} n_{c,m} n_{c,p}$ where c is the category, N the number of locations and n the number of times that models (m) and proxy records (p) have predicted category c. The κ-coefficient quantifies the reliability from no agreement beyond chance ($\kappa = 0$) to full agreement ($\kappa = 1$). Negative κ indicates agreement that is beyond change, for example, due to systematic biases.

D. Spectral gain

We investigate how climatic drivers influence the global mean temperature at period τ by calculating the spectral gain

$$G^2(\tau) = \frac{S_T(\tau)}{S_F(\tau)}. \tag{4}$$

Here $S_T(\tau)$ is the PSD of the global mean temperature and $S_F(\tau)$ the PSD of radiative forcing (see also Appendix C). The gain requires the assumption that the global mean temperature can be well approximated as a linear function of the forcing [27,88,89] and that different types of radiative forcing add linearly [90–93]. To this end, we focus on timescales between years and centuries when additivity is a valid assumption and nonlinearities in the global mean temperature are sufficiently small [42,43]. The main practical problem that confronts us is that the gain might be subject to a sampling bias due to our data sets choice. Therefore, we perform a Monte Carlo simulation of the PSD of radiative forcing and the global mean temperature, as well as the spectral gain as described in Appendix E.

IV. RESULTS AND DISCUSSION

A. Global mean and mean of local spectra

In order to study the timescale dependency of global mean temperature, we present its power spectral density in Fig. 2(b). It shows the characteristic background continuum, spectral

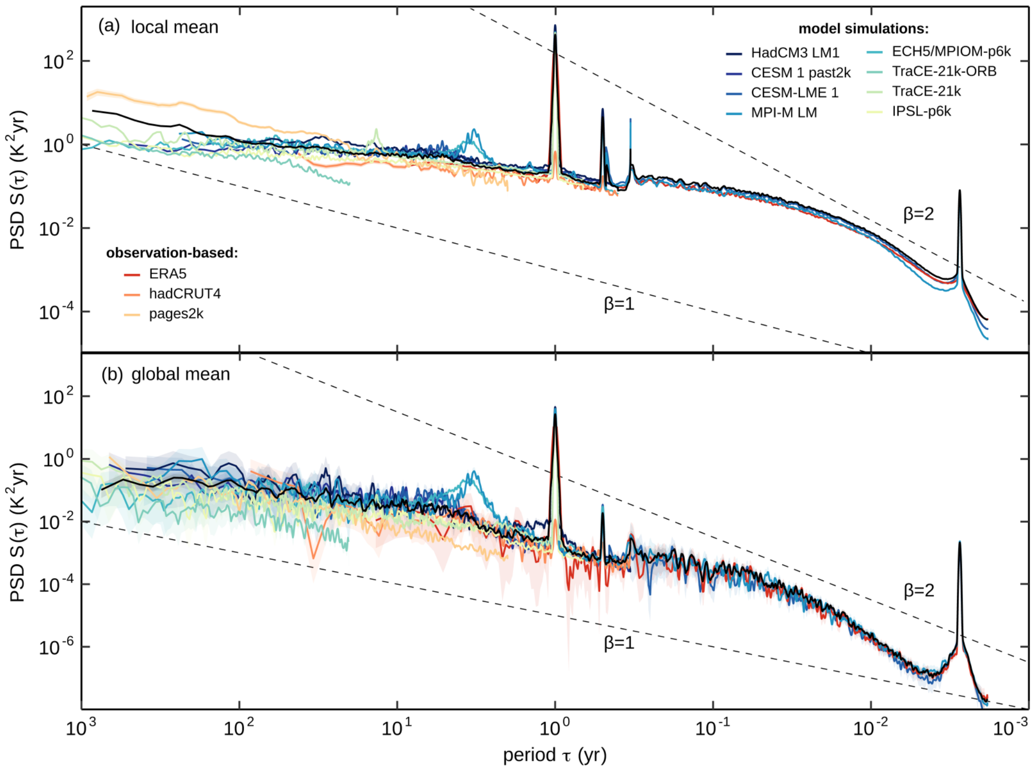

FIG. 2. (a) Mean power spectral densities (PSD) of local temperature from model simulations and observation-based data on periods from hours to 1000 yr for the Holocene. (b) PSD of global mean temperature. The dashed lines with slope β and arbitrary y-intercept in the log-log graph indicate the scaling behavior for visual comparison. The ensemble means (black solid lines) were formed using equal weights across the model group M_0 (see Table S1 [47]).

peaks, and higher harmonics associated with the diurnal and annual cycle. Overall, the PSDs tend to agree between the data sets, albeit with some differences on the interannual scale and when compared to the Trace21k ORB run. The Trace21k-ORB run is solely forced by orbital changes and therefore shows less variability than the ensemble mean. The broad spectral peak on interannual periods reveals an artificially amplified ENSO in the shared MPI-M LM and ECHAM5/MPI-OM ocean component [94]. For a better visibility, PI control runs are separately shown in the supplementary Fig. S6 [47]. Overall, the PSD largely agrees among different data sets, especially towards shorter timescales.

We find a power-law scaling of $\beta \approx 1$ on timescales longer than 10 yr in line with previous results [25,26,28]. The PSD decreases more strongly towards shorter periods, which is characteristic of the weather regime [25,36]. Similar to Nilsen *et al.* [26], we find no evidence for significant changes in scaling behavior around the centennial scale. One limitation of previous work that found scale breaks is that the spectra were estimated across nonstationary shifts in climate, such as the deglaciation [29], and with a change in proxies and archives [24].

We present the area-weighted mean spectra of the local (grid box) temperature in Fig. 2(a). Compared to the global mean in Fig. 2(b), the power increases and the spectral slope decreases, in line with [81]. The spectra agree on periods below 10 yr, except for the artificially amplified ENSO signal mentioned earlier. Moreover, we find a narrow peak at 13 yr, associated with an unrealistic variability in the northern North Atlantic of the TraCE-21k run, similar to [95,96]. Remarkably, the decadal-to-centennial variability of the reconstructed temperature is increased by one to two orders of magnitude compared to the simulations. The spectral exponent is smaller for models ($\beta < 1$) compared to paleoclimate data ($\beta \approx 1$).

This finding verifies that models show less regional temperature variability and that the mismatch increases towards longer timescales. The results are robust to sampling from the PAGES2k database and the influence of anthropogenic climate change (Fig. S10 [47]). One shortcoming of forming the area-weighted mean PSD is that the uncertainty quantification requires the assumption of independent spatial degrees of freedom of the temperature field. Due to the presence of spatial correlations, an estimate of the effective spatial degrees of freedom and their dependence on the underlying timescale would be needed to resolve this limitation [97].

B. Spatial patterns of persistence

To further investigate the mismatch on local scaling properties, we compare the spatial dependence of temperature persistence from simulations and paleoclimate data in Fig. 3. The simulations largely exhibit small-magnitude scaling exponents ($-1 < \beta < 1$), whereas proxy records were found to also show $\beta > 1$. In this manner, the magnitude of local temperature fluctuations from model simulations often shows no dependence on the decadal-to-centennial timescale. However, approximately half of the proxy records show a variance that grows on increasingly long periods (see also Fig. S11 [47]).

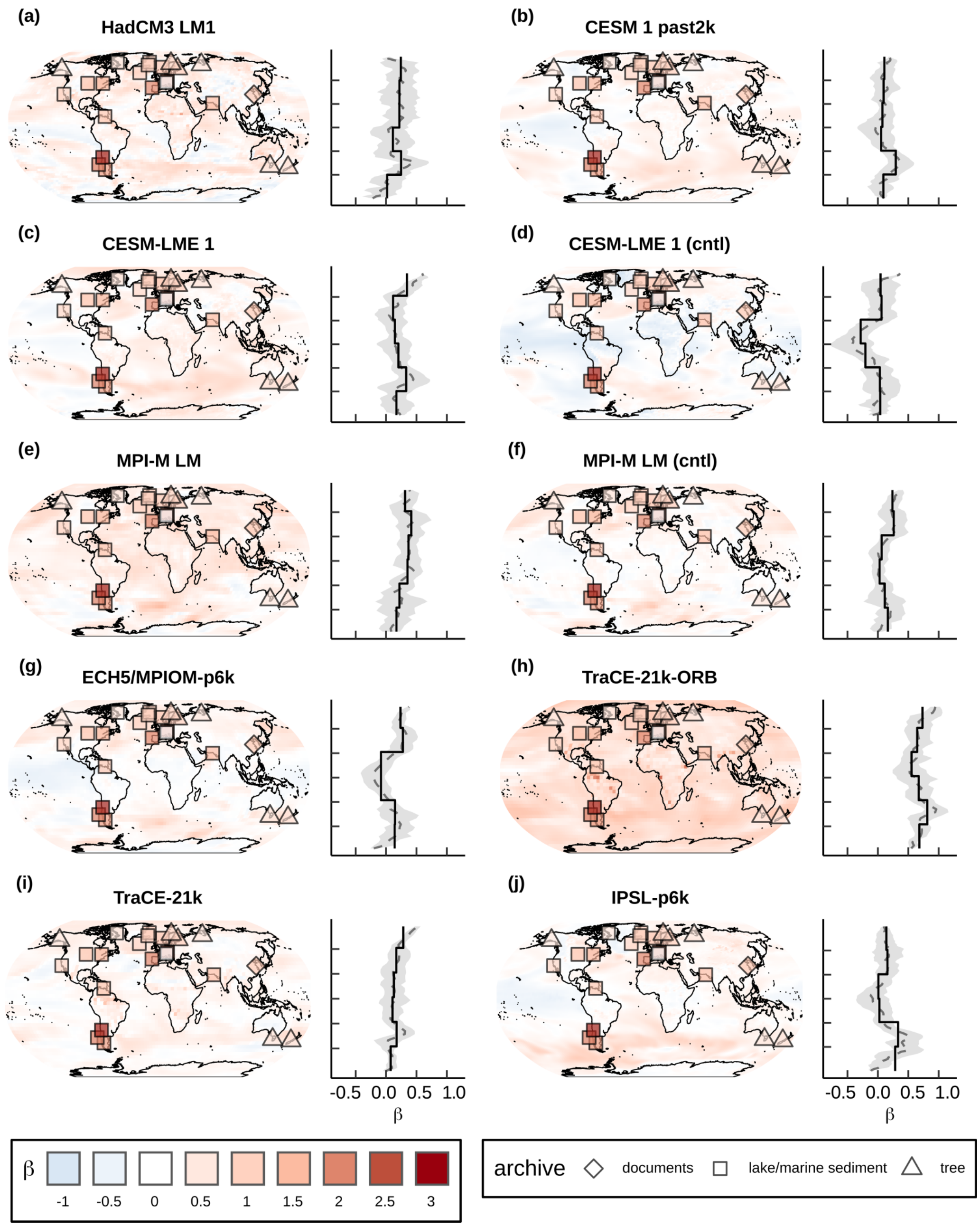

FIG. 3. Local temperature persistence on timescales from $\tau_1 = 10$ to $\tau_2 = 200$ yr across multiple climate simulations and selected proxy records from the PAGES2k database. Colors from blue to red indicate the scaling behavior ranging from $\beta = -1$ to $\beta = 3$. Symbols indicate the scaling of proxy records from different natural archives. The background of each panel shows the β-values fitted to the PSD of the local grid box temperature from simulations. Zonal mean values (dashed curves) are given next to the map, with means (solid curves) over latitude intervals (with breaks at −60, −30, 0, 30, and 60° N) and gray shaded confidence intervals. The spatial coverage of proxy records is not sufficient for robust mean estimates, which is why only simulation data are shown here.

From both simulations and paleoclimate data, we can strengthen the argument by Fredriksen *et al.* [81] that there is no latitudinal dependence of β (Fig. 3), in contrast to previous studies, suggesting a possible linkage to the strength of the seasonal cycle [24]. Inspecting the simulations' β-values (background of Fig. 3), we find a small land-sea contrast. Strongest scaling occurs in the Southern Oceans in line with previous findings [81]. Ocean-sea ice interactions with characteristic timescales of the order of centuries and a generally increased internal variability over the oceans might explain these results.

We find generally lower values for the slope β in the ENSO and Indo-Pacific region. This could be attributed to the fact that (quasi-)oscillatory signals, such as active modes of internal variability, are reflected in the PSD as broad peaks and hence cannot be described by a scaling law. On the other hand, this finding is stronger in PI control runs compared to fully forced runs [Figs. 3(c)–3(f)]. Thus, residual effects of the re-

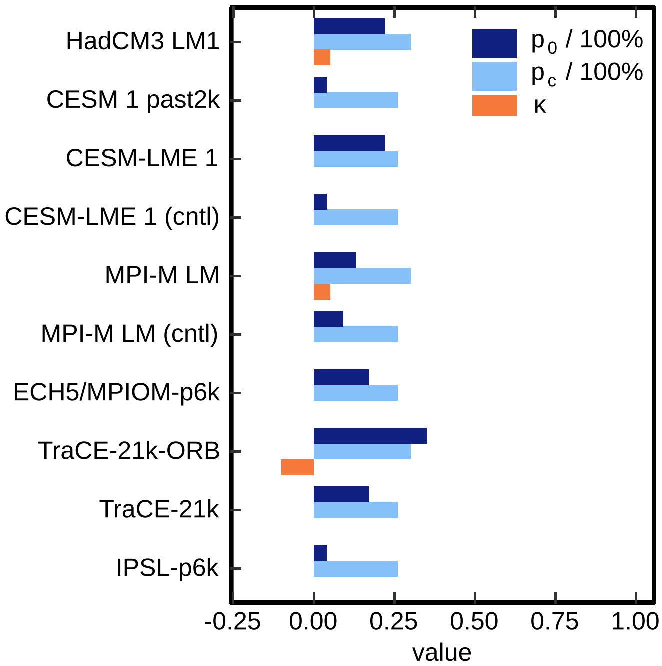

FIG. 4. Percentage agreement p_0, categorical agreement p_c and interrater reliability κ of local temperature persistence from simulations and paleoclimate data. The measures were calculated from a set of bilinearly interpolated simulation records and the proxy record at 23 different locations. Missing orange bars indicate no agreement beyond chance and, therefore, zero interrater reliability ($\kappa = 0$).

cent global warming trend might play an additional role [98]. A systematic bias becomes clear from the spatially almost uniform β-values of Trace21k-ORB [Fig. 3(h)]. In line with Fig. 2, we explain this by the lack of forcing mechanisms on interannual to multidecadal timescales in the aforementioned simulation.

Marine and lake sediments, as well as the archived documents, follow the general trend of increased β-values compared to simulations. Tree ring records agree well with most simulations in North America and Siberia, but not necessarily at the coast of Australia and northern Europe. Discrepancies such as those in southern South America could reflect the proxies' strength in representing local conditions, for example, topography. However, noise sources in the climate signal recording and preservation, such as bioturbation, can influence proxy records. Further separating the signal content from noise sources in paleoclimate reconstructions can help refine our findings [99,100].

C. Statistical agreement of temperature persistence

We further investigate the question of temperature scaling by a statistical analysis of β-values from simulations and paleoclimate reconstructions. It is based on the detailed uncertainty quantification outlined in Sec. III B. Our results show that reconstructions and simulations agree in less than 30% of locations within the scope of uncertainties (Fig. 4). To single out the scaling behavior of temperature signals, we study the agreement by category. We find approximately 25% of agreement within the categories $\beta < 1 - \nu$ (*low*) and $\beta > 1 + \nu$ (*high*). Although widely accepted [101], categorical and percentage agreement suffer from the limitation to ignore any agreement by chance. Therefore, we investigate the κ-statistics (orange bar in Fig. 4) and verify that there is no agreement beyond chance ($\kappa = 0$) for almost all models. Only MPI-M LM and HadCM3 LM1 show any, if poor agreement ($\kappa \approx 0.1$), whereas Trace21k-ORB shows even lower agreement than expected by chance ($\kappa < 0$) due to its systematic bias.

The disagreement could be attributed to both paleoclimate data and simulations. A systematic bias could arise, for example, through the recent, nonstationary global warming trend. Therefore, we repeat our analysis with all time series cut at 1850. In particular, anthropogenic warming slightly increases long-term temperature variability and thus scaling behavior, but not significantly (Figs. S6, S9, and S10 [47]). Further uncertainties could arise from our choice of statistical estimator for the scaling exponent β. Maximum likelihood estimation (MLE) should generally be preferred over linear regression (LR) because of its mathematical soundness and skillfulness [102]. We find that MLE is indeed more accurate for regular time series with $\beta > 0$ (Fig. S14 [47]). However, LR allows for estimation of $\beta < 1$, unlike MLE which assumes $\beta > 1$ [102]. In addition, for the characteristics of our empirical data, the differences between the two methods are not significant for $\beta > 0$ (Fig. S15 [47]). Therefore, linear regression represents the preferred estimator for our analysis. Regardless of the chosen estimator, we observe a slight tendency towards increased scaling exponents for irregularly sampled data (Fig. S15 [47]), similar to Lucke *et al.* [100]. Our uncertainty quantification carefully accounts for these potential errors due to irregular sampling and interpolation by simulating their influence using surrogates (Fig. S8 [47]).

We do not expect other systematic biases for the paleoclimate data since we base our results on multiple archives and proxies, and no systematic spatial pattern is discernible (Fig. S11 [47]). In particular, the cross-correlations between the 23 proxy data sets are weakly positive (0.02 on average with 95% quantiles of -0.17 to 0.21). The assumption of spatial independence necessary for robust statistical analysis (Fig. S16 [47]) therefore appears fully satisfied. The models' resolutions are another possible element of uncertainty that impacts variability over a wide range of timescales [103–105]. We here facilitate intermodel comparison by using state-of-the-art GCMs with comparable spatial and temporal resolutions, but computational costs precluded higher resolutions. The latter might be necessary to improve the representation of decadal variability and response to external forcing. In particular, the increased scaling exponents ($\beta > 1$) from paleoclimate data could indicate that nonlinear processes from an interactive carbon cycle and dynamical ice sheets might not be sufficiently represented in models.

D. The forced temperature response

Climatic drivers are not constant in time and thus affect the surface air temperature on multiple timescales. To investigate the forced temperature response, we present spectra for the main climatic drivers in Fig. 5. The PSD of orbital forcing consists of the diurnal and annual cycle as well as a back-

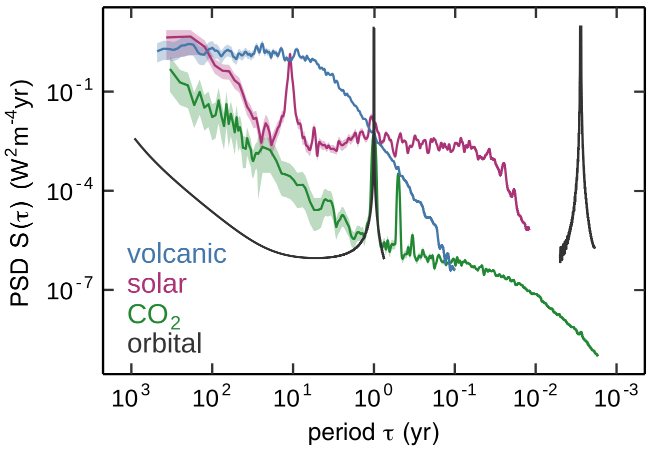

FIG. 5. Power spectral densities from radiative forcings. Details on the reconstructions considered here are summarized in Table S2 and Fig. S5 [47].

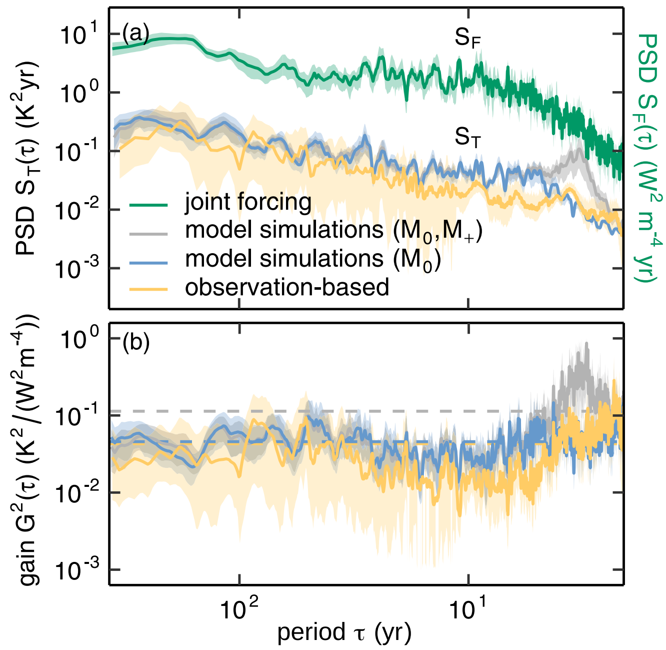

FIG. 6. Monte Carlo simulation of PSD (a) and spectral gain (b) using temperature and forcing reconstructions as well as model simulations. Shaded confidence intervals lie between the 5% and 95% quantiles. We consider only models from the groups M_0 and M_+ (Table S1 [47]) to exclude model artifacts and to represent the historical temperature response in the best possible way. Notably, M_+ contains those simulations with amplified ENSO [94]. (b) Dashed lines indicate the mean variance ratio $\langle S_T\rangle/\langle S_F\rangle$.

ground continuum on longer timescales. Higher harmonics on monthly timescales were omitted. We calculate the mean volcanic, solar, and CO_2 spectra using an equally weighted average of spectra from multiple data sets (Fig. S5 [47]). The CO_2 forcing follows the orbital forcing. The PSD of solar forcing again contains more power and has a pronounced peak around the 11 yr solar cycle. Multiple theories and paleoclimate reconstructions suggest the increased variability on centennial to millennial periods due to the long-term behavior of solar activity [106].

Volcanic forcing dominates interannual to centennial scales and undergoes a scale break around the period of 7 yr, estimated using the goodness of fit [102]. Above decadal scales, it follows a white noise spectrum with constant variance. However, the intermittency of volcanic eruptions might have led to biases in the spectral characteristics [42]. We verify our results using an analytical approach described in Appendix D. Remarkably, the derived PSD of an ideal, intermittent time series with Poisson distributed return times explains our findings. We further demonstrate the scale break by a Monte Carlo simulation of the joint PSD of radiative forcing in Fig. 6(a). This finding raises the question of how the spectrum with a scale break translates into the continuous spectrum in Fig. 2(b).

We address this question by calculating the spectral gain (4) on periods between years and centuries in Fig. 6(b). Here observation-based data include HadCRUT4, ERA5, and PAGES2k again. To account for the model artifacts explained above, we calculate the gain from the model simulation group M_0 and together with group M_+ (see Table S1 [47]). We find that the spectral gain is similar from observation-based data and the model simulation group M_0, which is the one without artificially amplified ENSO. This suggests that both follow a similar distribution of timescale-dependent variability, as already indicated by Fig. 2(b). Large parts of the gain show constant behavior, which is most pronounced in M_0. In a simplified way, the gain might be approximated by an ideal linear amplifier or damper of the forcing with comparable internal variability on all timescales. However, we also find a dip around decadal scales, which is strongest in the gain from measurements. Inspecting Fig. 6(a), this can be explained by forming the ratio between a spectrum with a scale break ($\beta > 1 \rightarrow \beta \approx 0$) and one with moderate scaling ($\beta \approx 1$).

From this standpoint, internal variability slightly grows on periods from years to centuries when slow processes in the oceans, vegetation, land surface, and cryosphere become increasingly active (Fig. 1). While the model simulations follow this general pattern, they may not represent its amplitude correctly, for example, due to the lack of feedback mechanisms. In addition, a too high model diffusivity could cause the suppression of low-frequency variability in model simulations due to a faster energy dissipation over temporal scales [8]. The PAGES2k multiproxy reconstruction, stemming from palaeoclimate data, possibly underestimates internal variability on interannual scales. However, the mean variance ratios in Fig. 5(b) of the model estimates agree with those from observations in the global mean. This leaves us with a conundrum: the global mean temperature based on model simulations and observations is mostly consistent in its variability, scaling, and response to forcing. Notwithstanding, locally, the models show a much lower variance on longer timescales and different scaling behavior than reconstructions. Thus, it appears that the statistics of local fluctuations need to be optimized in models but without significantly altering global properties. To this end, the study of unforced ("spontaneous") oscillations [107] and abrupt transitions [108,109] in the climate system is one promising approach to improve the representation of local variation. Furthermore, higher-resolved ocean and atmosphere models with additional mechanisms such as ice sheet dynamics and an interactive carbon cycle might increase long-range dependence and persistence of local temperature in the future.

V. CONCLUSION

In summary, we have investigated the question of temperature variability on the timescale of years to centuries. To this end, we have presented power spectral densities for both local and global surface air temperature from simulation and observation-based data of the last millennia. On this basis, we concluded that locally there is a stronger scaling and increased variance in reconstructions as compared to simulations. Using statistical analysis, we found that local temperature series extracted from simulations and paleoclimate reconstructions show different scaling behavior, with proxy records hinting at a stronger persistence. Furthermore, we have largely extended the spectral analysis of climatic drivers by estimating the joint PSD from CO_2, solar, volcanic, and orbital forcing using Monte Carlo simulation. Hereby, we discovered a scale break at the period of approximately 7 yr. Moreover, we have presented the spectral gain, describing the timescale-dependent forced temperature response. We found that it is mostly consistent across data sets and indicates an increasing internal variability on timescales of decades to centuries.

Our analysis of the spectral gain was limited to global average values and those timescales where linearity can be reasonably assumed [42,43,110]. Nonlinearities are inherent to the climate system, for example, due to the temperature-albedo feedback. Thus, it will be necessary to examine their possible effects on multiple spatiotemporal scales to further extend this work. Studying nonlinearities could also shine new light on the mechanisms of scaling in Earth's climate, which are not yet fully understood and might be linked to nonlinearities as well [6]. Furthermore, we have focused on the current interglacial, the Holocene. This is because climate variability has been demonstrated to depend on the mean climate state [82]. Furthermore, major shifts in climate could potentially violate the basic assumption of weak stationarity for spectral analysis. Thus, the conclusions laid out here cannot be readily applied to other climate states, such as glacial periods, which is an issue for future studies. Clearly, understanding the dependence of temperature variability on global warming demands additional work.

Ideally, our findings should be replicated by employing models with increased internal variability on longer timescales and paleoclimate data that provides improved spatiotemporal resolution. In particular, investigating the relationship between spatial and temporal disagreement is a key task for future analyses. Optimized analysis of noise sources and spectral analysis of (pseudo-)proxy records could help to expand the data basis of proxy records with decadal resolution [59,111,112]. Regarding climate models, an improved representation of processes that increase Earth's long-term memory, such as an interactive carbon cycle and dynamical ice sheets, might strengthen the long-range dependence and persistence of surface air temperature. A better understanding of unforced low-frequency oscillations as well as abrupt changes will be necessary to improve the representation of local fluctuations and could further help to understand nonlinear feedback and possible bifurcations in the climate system. Future studies could also continue to explore how internally generated and externally forced variability compares on different spatial scales. Research on the interrelation between internal and forced changes, as well as local, regional, and global variability, might prove important and could be conducted using single-forcing experiments from ensembles of model simulations.

Managing climate risks requires a detailed understanding of temperature variability. Locally and on timescales between years and centuries, there is an urgency to address discrepancies to make further progress in climate modeling. In this study, we have singled out the key characteristics of temperature variability and showed that the timescale dependency of local temperature variations from observation-based data and model simulations differs. Our results have demonstrated that the scaling behavior and spectral gain are easy-to-use yet effective and promising tools for investigating variability in Earth's dynamic climate.

Code to reproduce all figures is available at [113].

ACKNOWLEDGMENTS

This manuscript is based upon 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 listed in Tables S1 and S2 for producing and making available their data from model outputs, measurements, paleoclimate, and forcing reconstructions. This study benefited from discussions within the CVAS working group, a working group of the Past Global Changes (PAGES) project. We thank T. Gasenzer, T. Kunz, and N. Weitzel for discussions and J. Bühler, M. Casado, M. Schillinger, and E. Ziegler for helpful comments on the manuscript. We are grateful to Aimé Fournier and one anonymous referee for their constructive and valuable review. This research has been funded by the Heidelberg Graduate School for Physics, by the PalMod project (subProject no. 01LP1926C), and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Project No. 395588486.

APPENDIX A: RELATION BETWEEN POWER SPECTRAL DENSITY AND VARIANCE

The power spectral density of a weakly stationary, stochastic process is given by the Fourier transform of the autocorrelation $S(f) = \mathcal{F}\{R(h)\}$ with frequency f and lag $h = t_2 - t_1$ between two points in time [19,20]. For zero lag and zero mean, the integral of the PSD corresponds to the variance of the signal [80]. Instead of frequency, we use the period $\tau = 1/f$ to express the PSD and spectral gain. The integration of expression (1) is divergent for $\beta < 1$ and $f \to \infty$ which requires a high-frequency cutoff, such as described by Lovejoy *et al.* [38]. In case of temperature time series considered here, this is naturally defined by the temporal resolution, setting the maximum frequency.

APPENDIX B: AUTOCOVARIANCE OF LONG-RANGE MEMORY PROCESSES

Fractional Brownian motion (fBm) and fractional Gaussian noise (fGn) are fully described by their correlation properties [30,114], summarized below. The autocovariance function of fBm $B(t)$ reads

$$\gamma(t',t) = \langle B(t')B(t)\rangle = \frac{V_\beta}{2}(|t|^{\beta-1} + |t'|^{\beta-1} - |t'-t|^{\beta-1}) \propto 1 + \left|\frac{t'}{t}\right|^{\beta-1} - \left|1-\frac{t'}{t}\right|^{\beta-1} \tag{B1}$$

for $1 < \beta < 3$. V_β is a positive constant factor related to $\langle (B(t') - B(t))^2\rangle = V_\beta\, |t'-t|^{\beta-1}$. By definition, fGn is the series of stationary increments $B(t') - B(t)$ and shows spectral exponent $-1 < \beta' = \beta - 2 < 1$ for $f \ll 1/\pi\Delta t$ with $\Delta t = t' - t$. Its autocovariance

$$\gamma(h) = \langle [B(t+1+h) - B(t+h)][B(t+1) - B(t)]\rangle = \frac{V_\beta}{2}|h-1|^{\beta'+1} - 2|h|^{\beta'+1} + |h+1|^{\beta'+1} \tag{B2}$$

depends only on the lag $h \in \mathbb{Z}$, where we set $\Delta t = 1$ without loss of generality. The fGn has a power spectrum of the form [115]

$$S(f) \propto \frac{\sin^2(\pi\Delta t f)}{|2\pi\Delta t f|^{\beta'+2}}, \tag{B3}$$

with the slowly varying factor

$$\sin^2(\pi\Delta t f) \xrightarrow[f/f_{\max}\to 0]{} (\pi\Delta t)^2 f^2, \qquad f_{\max} = 1/\pi\Delta t. \tag{B4}$$

Considering positive frequencies $f > 0$, the spectrum (B3) can be approximated by the power law $S(f) \sim 1/f^{\beta'}$ if $f \ll f_{\max}$. For $f \gtrsim f_{\max}$, however, the fGn has a similar spectral shape to fBm [114]. We account for this by considering sufficiently long periods. To give an example, $10^{0.58}\ \mathrm{yr}^{-1} \lesssim f_{\max} \lesssim 10^{2.7}\ \mathrm{yr}^{-1}$ corresponds to 6 h $\lesssim \Delta t \lesssim$ 1 mo.

For all $|t'/t| \gg 1$, the covariances (B1) keep growing for $\beta > 2$ (persistence) and stay bounded for $\beta < 2$ (antipersistence). As a result, Eq. (B1) involves "nonlinear pseudo-trends" [30] for $B(t')$ conditioned on $B(t)$, which diverge for $\beta > 2$ and converge for $\beta < 2$. According to Eq. (B2), fGn is persistent for $\beta' > 0$ and antipersistent for $\beta' < 0$. Ordinary Brownian motion corresponds to $\beta = 2$ and white noise to $\beta' = 0$. The sequence of partial sums of the autocovariance function diverges for fGn with $\beta' > 0$ and fBm with $\beta > 2$. The process is nonsummable and said to possess long-range memory.

APPENDIX C: SPECTRAL GAIN FOR LINEAR SYSTEMS

In a time-invariant linear system, the output

$$y(t) = \int_{-\infty}^{\infty} h(u)x(t-u)\,du \tag{C1}$$

is given by the input time series $x(t)$ and the impulse response function $h(u)$ [80]. The Fourier transform $H(f) = \mathcal{F}\{h(u)\} = G(f)e^{i\phi(f)}$ gives the frequency response function, also called the transfer function. $G(f)$ and $\phi(f)$ are the gain and phase, respectively. The integral (C1) corresponds to a product in frequency space $\mathcal{F}\{y(t)\} = H(f)\mathcal{F}\{x(t)\}$. This relates the PSD of the output $S_y(f)$ to the one of the input $S_x(f)$ via

$$S_y(f) = |H(f)|^2 S_x(f) = G^2(f)S_x(f). \tag{C2}$$

APPENDIX D: ANALYTICAL SOLUTION TO THE PSD OF INTERMITTENT VOLCANIC FORCING

We investigate the power spectral density of intermittent volcanic forcing by approximating the eruption time series in a simplified way as a stochastic signal $X(t) = \delta(t - t_i)$. This function is zero at all times except t_i, when an event of unique amplitude occurs. We denote $T_i = t_i - t_{i-1}$ the time intervals between two events. We use the fact that the PSD cannot be calculated only from the covariance, but also from the Laplace transform $S(X, f) = 2\lim_{\epsilon\to 0}\langle |\mathcal{L}(X(t), \frac{\epsilon}{2} - 2\pi i f)|^2\rangle$ [116]. Based on this approach, the power spectral density

$$S(f) = \mu_T\frac{1 - |\rho(f)|^2}{|1-\rho(f)|^2}, \quad f > 0 \tag{D1}$$

becomes a function of the Fourier transform of the probability density function $\rho(f) = \mathcal{F}\{\rho(T)\}$ and the inverse mean interval between two events $\mu_T = \langle T\rangle^{-1}$ [116,117]. An exponentially decaying probability distribution $\rho(T) = \mu_T \exp(-\mu_T T)\Theta(T)$ for volcanic forcing is suggested [118], and we have checked this for the data sets considered. The Fourier transform reads $\rho(f) = \mu_T(\mu_T + 2\pi i f)^{-1}$ such that $1 - |\rho(f)|^2 = |1 - \rho(f)|^2$. As a consequence, the PSD (D1) takes a constant value. We can observe this white noise behavior in Figs. 5 and 6(a) on timescales longer than a few years, which is on the order of characteristic return times for eruptions. Below these timescales, the variability considerably drops. This analytical result provides an independent verification of the PSD for volcanic forcing and its scale break.

APPENDIX E: MONTE CARLO SAMPLING OF THE SPECTRAL GAIN

We simulate the spectral gain (4), as well as the PSD of global mean temperature and the joint PSD of radiative forcing using a Monte Carlo approach with $N = 1000$ realizations to account for sampling biases. The PSD of global mean temperature is sampled for three groups: the observation-based data, the model simulations from group M_0, and those from M_0 together with M_+ (Table S1 [47]). Here only models from the groups M_0 and M_+ are considered to exclude model artifacts and to represent the historical temperature response in the best possible way.

We sample the simulation-based PSD from the average PSD of the simulations using uniformly distributed random weights. To obtain the observation-based PSD, we use the global mean temperature from HadCRUT4, ERA5, and a 7000-member reconstruction ensemble provided by PAGES2k [46]. This ensemble allows us to sample the PSD by randomly selecting one ensemble member and form the mean of its spectrum with that of the ERA5 and HadCRUT4 temperature. The joint PSD of radiative forcing is calculated from all forcing reconstructions considered in this work except the Fröhlich *et al.* solar forcing, which has too low temporal resolution above interannual scales (Table S3 and Fig. S5 [47]). We assume the PSD of CO_2 and orbital forcing as fixed since its spectral power is comparatively low on multidecadal scales. We sample the PSD of solar forcing by using uniformly distributed weights when forming the average PSD of all solar reconstructions. Similarly, the PSD of volcanic forcing is obtained. In addition, we randomly vary the conversion factor between $(-18)^{-1}$ and $(-25)^{-1}$ $\mathrm{Wm}^{-2}/\mathrm{AOD}$ [78]. The joint PSD of radiative forcing is calculated by linear summation of the PSD from CO_2, orbital, solar, and volcanic forcing.

Using this sampling scheme, our Monte Carlo produces two outcomes: First, we compute the PSD of global mean temperature and the joint PSD of radiative forcing by simulating an ensemble of N realizations for both forcing and response. Second, we sample the spectral gain directly from the quotient (4) in each of the N realizations. In both cases, the average of the generated N-member ensemble and its 5% and 95% quantiles constitute the result of our Monte Carlo simulation.

[1] J. Bjerknes, A possible response of the atmospheric Hadley circulation to equatorial anomalies of ocean temperature, Tellus **18**, 820 (1966).

[2] S. Arrhenius, XXXI. On the influence of carbonic acid in the air upon the temperature of the ground, London Edinburgh Dublin Philos. Mag. J. Sci. **41**, 237 (1896).

[3] J. B. J. Fourier, Remarques générales sur les températures du globe terrestre et des espaces planétaires, Ann. Chim. Phys. **27**, 136 (1824).

[4] R. W. Katz and B. G. Brown, Extreme events in a changing climate: Variability is more important than averages, Clim. Chan. **21**, 289 (1992).

[5] M. Ghil and V. Lucarini, The physics of climate variability and climate change, Rev. Mod. Phys. **92**, 035002 (2020).

[6] C. L. E. Franzke, S. Barbosa, R. Blender, H.-B. Fredriksen, T. Laepple, F. Lambert, T. Nilsen, K. Rypdal, M. Rypdal, M. G. Scotto *et al.*, The structure of climate variability across scales, Rev. Geophys. **58**, e2019RG000657 (2020).

[7] J. E. Tierney *et al.*, Past climates inform our future, Science **370**, eaay3701 (2020).

[8] T. Laepple and P. Huybers, Global and regional variability in marine surface temperatures, Geophys. Res. Lett. **41**, 2528 (2014).

[9] T. Laepple and P. Huybers, Ocean surface temperature variability: Large model-data differences at decadal and longer periods, Proc. Nat. Acad. Sci. U. S. A. **111**, 16682 (2014).

[10] L. A. Parsons, G. R. Loope, J. T. Overpeck, T. R. Ault, R. Stouffer, and J. E. Cole, Temperature and precipitation variance in CMIP5 simulations and paleoclimate records of the last millennium, J. Clim. **30**, 8885 (2017).

[11] F. C. Ljungqvist, Q. Zhang, G. Brattström, P. J. Krusic, A. Seim, Q. Li, Q. Zhang, and A. Moberg, Centennial-scale temperature change in last millennium simulations and proxy-based reconstructions, J. Clim. **32**, 2441 (2019).

[12] J. C. Bühler, C. Roesch, M. Kirschner, L. Sime, M. D. Holloway, and K. Rehfeld, Comparison of the oxygen isotope signatures in speleothem records and iHadCM3 model simulations for the last millennium, Clim. Past **17**, 985 (2021).

[13] W. B. Anderson, R. Seager, W. Baethgen, M. Cane, and L. You, Synchronous crop failures and climate-forced production variability, Sci. Adv. **5**, eaaw1976 (2019).

[14] M. Crucifix, A. D. Vernal, and C. Franzke, Centennial to millennial climate variability, Past Global Changes Mag. **25**, 131 (2017).

[15] M. Rypdal, H. B. Fredriksen, E. Myrvoll-Nilsen, K. Rypdal, and S. H. Sørbye, Emergent scale invariance and climate sensitivity, Climate **6**, 93 (2018).

[16] C. L. E. Franzke and T. J. O'Kane (eds.), *Nonlinear and Stochastic Climate Dynamics* (Cambridge University Press, Cambridge, 2017).

[17] T. P. Barnett *et al.*, Detection and attribution of recent climate change: A status report, Bull. Am. Meteorol. Soc. **80**, 2631 (1999).

[18] N. Bindoff *et al.*, Detection and attribution of climate change: From global to regional, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Intergovernmental Panel on Climate Change (Cambridge University Press, Cambridge, 2013), pp. 867–952.

[19] N. Wiener, Generalized harmonic analysis, Acta Math. **55**, 117 (1930).

[20] A. Khintchine, Korrelationstheorie der stationären stochastischen Prozesse, Math. Ann. **109**, 604 (1934).

[21] C. Wunsch, The spectral description of climate change including the 100 ky energy, Clim. Dyn. **20**, 353 (2003).

[22] K. Fraedrich, U. Luksch, and R. Blender, $1/f$ model for long-time memory of the ocean surface temperature, Phys. Rev. E **70**, 037301 (2004).

[23] C. Franzke, Nonlinear trends, long-range dependence, and climate noise properties of surface temperature, J. Clim. **25**, 4172 (2012).

[24] P. Huybers and W. Curry, Links between annual, Milankovitch and continuum temperature variability, Nature (London) **441**, 329 (2006).

[25] S. Lovejoy, A voyage through scales, a missing quadrillion and why the climate is not what you expect, Clim. Dyn. **44**, 3187 (2015).

[26] T. Nilsen, K. Rypdal, and H. B. Fredriksen, Are there multiple scaling regimes in Holocene temperature records? Syst. Dyn. Rev. **7**, 419 (2016).

[27] H. B. Fredriksen and M. Rypdal, Long-range persistence in global surface temperatures explained by linear multibox energy balance models, J. Clim. **30**, 7157 (2017).

[28] K. Rypdal, L. Østvand, and M. Rypdal, Long-range memory in Earth's surface temperature on time scales from months to centuries, J. Geophys. Res.: Atmosph. **118**, 7046 (2013).

[29] F. Zhu, J. Emile-Geay, N. P. McKay, G. J. Hakim, D. Khider, T. R. Ault, E. J. Steig, S. Dee, and J. W. Kirchner, Climate models can correctly simulate the continuum of global-

average temperature variability, Proc. Nat. Acad. Sci. U. S. A. **116**, 8728 (2019).
[30] B. B. Mandelbrot and J. W. Van Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev. **10**, 422 (1968).
[31] B. D. Malamud and D. L. Turcotte, Self-affine time series: Measures of weak and strong persistence, J. Stat. Plan. Infer. **80**, 173 (1999).
[32] J. P. Peixto and A. H. Oort, Physics of climate, Rev. Mod. Phys. **56**, 365 (1984).
[33] E. J. Rohling *et al.*, Making sense of palaeoclimate sensitivity, Nature (London) **491**, 683 (2012).
[34] E. J. Rohling, G. Marino, G. L. Foster, P. A. Goodwin, A. S. von der Heydt, and P. Köhler, Comparing climate sensitivity, past and present, Ann. Rev. Marine Sci. **10**, 261 (2018).
[35] F. Zhang, Y. Qiang Sun, L. Magnusson, R. Buizza, S. J. Lin, J. H. Chen, and K. Emanuel, What is the predictability limit of midlatitude weather? J. Atmos. Sci. **76**, 1077 (2019).
[36] J. D. Pelletier, Natural variability of atmospheric temperatures and geomagnetic intensity over a wide range of time scales, Proc. Nat. Acad. Sci. U. S. A. **99**, 2546 (2002).
[37] J. Lohmann and P. D. Ditlevsen, Random and externally controlled occurrences of Dansgaard–Oeschger events, Clim. Past. **14**, 609 (2018).
[38] S. Lovejoy and F. Lambert, Spiky fluctuations and scaling in high-resolution EPICA ice core dust fluxes, Clim. Past **15**, 1999 (2019).
[39] K. Marvel, G. A. Schmidt, R. L. Miller, and L. S. Nazarenko, Implications for climate sensitivity from the response to individual forcings, Nat. Clim. Change **6**, 386 (2016).
[40] A. P. Schurer, G. C. Hegerl, M. E. Mann, S. F. B. Tett, and S. J. Phipps, Separating forced from chaotic climate variability over the past millennium, J. Clim. **26**, 6954 (2013).
[41] S. Lovejoy and D. Schertzer, Stochastic and scaling climate sensitivities: Solar, volcanic and orbital forcings, Geophys. Res. Lett. **39**, L11702 (2012).
[42] S. Lovejoy and C. Varotsos, Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forcings, Syst. Dyn. Rev. **7**, 133 (2016).
[43] K. Rypdal and M. Rypdal, Comment on "Scaling regimes and linear/nonlinear responses of last millennium climate to volcanic and solar forcing" by S. Lovejoy and C. Varotsos (2016), Syst. Dyn. Rev. **7**, 597 (2016).
[44] J. Mitchell, An overview of climatic variability and its causal mechanisms, Quat. Res. **6**, 481 (1976).
[45] A. Henderson-Sellers and K. McGuffie, *A Climate Modelling Primer* (Wiley, Chichester, 1987).
[46] PAGES 2k Consortium, Consistent multidecadal variability in global temperature reconstructions and simulations over the Common Era, Nat. Geosci. **12**, 643 (2019).
[47] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.104.064136 for further information on data sets and benchmarks.
[48] B. L. Otto-Bliesner, E. C. Brady, J. Fasullo, A. Jahn, L. Landrum, S. Stevenson, N. Rosenbloom, A. Mai, and G. Strand, Climate variability and change since 850 CE an ensemble approach with the Community Earth System Model, Bull. Am. Meteorol. Soc. **97**, 735 (2016).
[49] J. H. Jungclaus *et al.*, Climate and carbon-cycle variability over the last millennium, Clim. Past **6**, 723 (2010).
[50] Y. Zhong, A. Jahn, G. H. Miller, and A. Geirsdottir, Asymmetric cooling of the atlantic and pacific arctic during the past two millennia: A dual observation-modeling study, Geophys. Res. Lett. **45**, 12,497 (2018).
[51] P. Braconnot, D. Zhu, O. Marti, and J. Servonnat, Strengths and challenges for transient Mid- to Late Holocene simulations with dynamical vegetation, Clim. Past **15**, 997 (2019).
[52] N. Fischer and J. H. Jungclaus, Evolution of the seasonal temperature cycle in a transient Holocene simulation: Orbital forcing and sea-ice, Clim. Past **7**, 1139 (2011).
[53] Z. Liu, Transient simulation of last deglaciation with a new mechanism for Bølling-Allerød warming, Science **325**, 310 (2009).
[54] C. P. Morice, J. J. Kennedy, N. A. Rayner, and P. D. Jones, Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set, J. Geophys. Res.: Atmos. **117**, D08101 (2012).
[55] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara *et al.*, The ERA5 global reanalysis, Q. J. R. Meteorol. Soc. **146**, 1999 (2020).
[56] PAGES2k Consortium., A global multiproxy database for temperature reconstructions of the Common Era, Sci. Data **4**, 170088 (2017).
[57] J. W. Kantelhardt, Fractal and multifractal time series, in *Mathematics of Complexity and Dynamical Systems*, edited by R. A. Meyers (Springer, New York, 2011), pp. 463–487.
[58] T. Laepple, T. Münch, M. Casado, M. Hoerhold, A. Landais, and S. Kipfstuhl, On the similarity and apparent cycles of isotopic variations in East Antarctic snow pits, Cryosphere **12**, 169 (2018).
[59] M. Casado, T. Münch, and T. Laepple, Climatic information archived in ice cores: Impact of intermittency and diffusion on the recorded isotopic signal in Antarctica, Clim. Past **16**, 1581 (2020).
[60] T. J. Crowley, Causes of climate change over the past 1000 years, Science **289**, 270 (2000).
[61] G. C. Hegerl, T. J. Crowley, M. Allen, W. T. Hyde, H. N. Pollack, J. Smerdon, and E. Zorita, Detection of human influence on a new, validated 1500-year temperature reconstruction, J. Clim. **20**, 650 (2007).
[62] P. Braconnot, S. P. Harrison, M. Kageyama, P. J. Bartlein, V. Masson-Delmotte, A. Abe-Ouchi, B. Otto-Bliesner, and Y. Zhao, Evaluation of climate models using palaeoclimatic data, Nat. Clim. Change **2**, 417 (2012).
[63] G. A. Schmidt *et al.*, Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1), Geosci. Model Dev. **5**, 185 (2012).
[64] G. Delaygue and E. Bard, An Antarctic view of Beryllium-10 and solar activity for the past millennium, Clim. Dyn. **36**, 2201 (2011).
[65] F. Steinhilber, J. Beer, and C. Fröhlich, Total solar irradiance during the Holocene, Geophys. Res. Lett. **36**, L19704 (2009).
[66] Y. Wang, J. L. Lean, and N. R. Sheeley, Jr., Modeling the Sun's magnetic field and irradiance since 1713, Astrophys. J. **625**, 522 (2005).
[67] R. Muscheler, F. Joos, J. Beer, S. A. Müller, M. Vonmoos, and I. Snowball, Solar activity during the last 1000 yr inferred from radionuclide records, Quat. Sci. Rev. **26**, 82 (2007).

[68] L. E. Vieira and S. K. Solanki, Evolution of the solar magnetic flux on time scales of years to millenia, Astron. Astrophys. **509**, A100 (2010).

[69] C. Gao, A. Robock, and C. Ammann, Volcanic forcing of climate over the past 1500 years: An improved ice core-based index for climate models, J. Geophys. Res.: Atmos. **113**, D23111 (2008).

[70] A. L. Berger, Long-term variations of daily insolation and Quaternary climatic changes, J. Atmos. Sci. **35**, 2362 (1978).

[71] M. Crucifix, Palinsol–Package (R) (2016), https://CRAN.R-project.org/package=palinsol.

[72] M. Toohey and M. Sigl, Reconstructed volcanic stratospheric sulfur injections and aerosol optical depth, 500 BCE to 1900 CE, version 2, World Data Center for Climate (WDCC) at DKRZ, https://cera-www.dkrz.de/WDCC/ui/cerasearch/entry?acronym=eVolv2k_v2 (2017).

[73] C. Fröhlich, Solar irradiance variability since 1978: Revision of the PMOD composite during solar cycle 21, Space Sci. Rev. **125**, 53 (2006).

[74] C. D. Keeling, R. B. Bacastow, A. E. Bainbridge, C. A. Ekdahl Jr., P. R. Guenther, L. S. Waterman, and John F. S. Chin, Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii, Tellus **28**, 538 (1976).

[75] J. Pongratz, C. Reick, T. Raddatz, and M. Claussen, A reconstruction of global agricultural areas and land cover for the last millennium, Global Biogeochem. Cycles **22**, GB3018 (2008).

[76] G. Myhre, E. J. Highwood, K. P. Shine, and F. Stordal, New estimates of radiative forcing due to well mixed greenhouse gases, Geophys. Res. Lett. **25**, 2715 (1998).

[77] G. A. Schmidt *et al.*, Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0), Geosci. Model Dev. **4**, 33 (2011).

[78] G. Myhre *et al.*, Anthropogenic and natural radiative forcing supplementary naterial, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by Intergovernmental Panel on Climate Change (Cambridge University Press, Cambridge, 2014), pp. 1–30.

[79] T. Laepple and P. Huybers, Reconciling discrepancies between Uk37 and Mg/Ca reconstructions of Holocene marine temperature variability, Earth Planet. Sci. Lett. **375**, 418 (2013).

[80] C. Chatfield and H. Xing, *The Analysis of Time Series: An Introduction with R*, 7th ed., Chapman & Hall/CRC Texts in Statistical Science (CRC Press, Boca Raton, FL, 2019).

[81] H. B. Fredriksen and K. Rypdal, Spectral characteristics of instrumental and climate model surface temperatures, J. Clim. **29**, 1253 (2016).

[82] K. Rehfeld, T. Münch, S. L. Ho, and T. Laepple, Global patterns of declining temperature variability from the Last Glacial Maximum to the Holocene, Nature (London) **554**, 356 (2018).

[83] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques* (Cambridge University Press, Cambridge, 1993).

[84] P. Yiou, E. Baert, and M. F. Loutre, Spectral analysis of climate data, Surv. Geophys. **17**, 619 (1996).

[85] J. W. Kirchner, Aliasing in $1/f^{a}$ noise spectra: Origins, consequences, and remedies, Phys. Rev. E **71**, 066110 (2005).

[86] L. Østvand, T. Nilsen, K. Rypdal, D. Divine, and M. Rypdal, Long-range memory in internal and forced dynamics of millennium-long climate model simulations, Syst. Dyn. Rev. **5**, 295 (2014).

[87] J. L. Fleiss and J. Cohen, The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability, Educ. Psychol. Meas. **33**, 613 (1973).

[88] O. Geoffroy, D. Saint-Martin, D. J. Olivié, A. Voldoire, G. Bellon, and S. Tytéca, Transient climate response in a two-layer energy-balance model. Part I: Analytical solution and parameter calibration using CMIP5 AOGCM experiments, J. Clim. **26**, 1841 (2013).

[89] D. G. MacMynowski, H.-J. Shin, and K. Caldeira, The frequency response of temperature and precipitation in a climate model, Geophys. Res. Lett. **38**, L16711 (2011).

[90] A. Kirkevåg, T. Iversen, J. E. Kristjánsson, Ø. Seland, and J. B. Debernard, On the additivity of climate response to anthropogenic aerosols and CO_2, and the enhancement of future global warming by carbonaceous aerosols, Tellus, Ser. A: Dyn. Meteorol. Oceanogr. **60**, 513 (2008).

[91] H. Shiogama, D. A. Stone, T. Nagashima, T. Nozawa, and S. Emori, On the linear additivity of climate forcing-response relationships at global and continental scales, Int. J. Climatol. **33**, 2542 (2013).

[92] G. A. Meehl, W. M. Washington, C. M. Ammann, J. M. Arblaster, T. M. Wigley, and C. Tebaldi, Combinations of natural and anthropogenic forcings in twentieth-century climate, J. Clim. **17**, 3721 (2004).

[93] V. Ramaswamy and C. T. Chen, Linear additivity of climate response for combined albedo and greenhouse perturbations, Geophys. Res. Lett. **24**, 567 (1997).

[94] J. Jungclaus (private communication) (2020).

[95] G. Danabasoglu, On multidecadal variability of the Atlantic meridional overturning circulation in the community climate system model version 3, J. Clim. **21**, 5524 (2008).

[96] T. Kunz and T. Laepple, Frequency-dependent estimation of effective spatial degrees of freedom, J. Clim. **34**, 7373 (2021).

[97] T. Kunz and T. Laepple, Time-scale dependent estimation of spatial degrees of freedom, in *EGU General Assembly Conference Abstracts* (EGU General Assembly, Vienna, Austria, 2018), p. 17995.

[98] S.-W. Yeh, J.-S. Kug, B. Dewitte, M.-H. Kwon, B. P. Kirtman, and F.-F. Jin, El Niño in a changing climate, Nature (London) **461**, 511 (2009).

[99] M. Reschke, K. Rehfeld, and T. Laepple, Empirical estimate of the signal content of Holocene temperature proxy records, Clim. Past **15**, 521 (2019).

[100] L. J. Lücke, G. C. Hegerl, A. P. Schurer, and R. Wilson, Effects of memory biases on variability of temperature reconstructions, J. Clim. **32**, 8713 (2019).

[101] J. L. Fleiss, B. Levin, and M. C. Paik, *Statistical Methods for Rates and Proportions*, 3rd ed., Wiley Series in Probability and Statistics (Wiley, New York, 2013).

[102] A. Clauset, C. R. Shalizi, and M. E. J. Newman, Power-law distributions in empirical data, SIAM Rev. **51**, 661 (2009).

[103] B. P. Kirtman *et al.*, Impact of ocean model resolution on CCSM climate simulations, Clim. Dyn. **39**, 1303 (2012).

[104] J. M. Klavans, A. Poppick, S. Sun, and E. J. Moyer, The influence of model resolution on temperature variability, Clim. Dyn. **48**, 3035 (2017).

[105] D. L. Hodson and R. T. Sutton, The impact of resolution on the adjustment and decadal variability of the Atlantic meridional overturning circulation in a coupled climate model, Clim. Dyn. **39**, 3057 (2012).

[106] L. J. Gray *et al.*, Solar influences on climate, Rev. Geophys. **48**, (2010).

[107] E. Hertwig, F. Lunkeit, and K. Fraedrich, Low-frequency climate variability of an aquaplanet, Theor. Appl. Climatol. **121**, 459 (2015).

[108] W. R. Peltier and G. Vettoretti, Dansgaard-Oeschger oscillations predicted in a comprehensive model of glacial climate: A "kicked" salt oscillator in the Atlantic, Geophys. Res. Lett. **41**, 7306 (2014).

[109] M. Klockmann, U. Mikolajewicz, H. Kleppin, and J. Marotzke, Coupling of the subpolar gyre and the overturning circulation during abrupt glacial climate transitions, Geophys. Res. Lett. **47**, e2020GL090361 (2020).

[110] L. Fernández-Donado *et al.*, Large-scale temperature response to external forcing in simulations and reconstructions of the last millennium, Clim. Past **9**, 393 (2013).

[111] T. Kunz, A. M. Dolman, and T. Laepple, A spectral approach to estimating the timescale-dependent uncertainty of paleoclimate records—Part 1: Theoretical concept, Clim. Past **16**, 1469 (2020).

[112] A. M. Dolman, T. Kunz, J. Groeneveld, and T. Laepple, A spectral approach to estimating the timescale-dependent uncertainty of paleoclimate records–Part 2: Application and interpretation, Clim. Past. **17**, 825 (2021).

[113] https://github.com/paleovar/TimescaleDependency.

[114] A. Fournier (private communication) (2021).

[115] S. Mallat, Wavelet zoom, in *A Wavelet Tour of Signal Processing*, 3rd ed. (Academic Press, Boston, 2009), pp. 205–261.

[116] R. L. Stratonovich, *Topics in the Theory of Random Noise* (Gordon and Breach, New York, 1967).

[117] B. Lindner, Superposition of many independent spike trains is generally not a Poisson process, Phys. Rev. E **73**, 022901 (2006).

[118] P. Papale, Global time-size distribution of volcanic eruptions on Earth, Sci. Rep. **8**, 6838 (2018).