Climate models can correctly simulate the continuum of global-average temperature variability

Feng Zhu[a], Julien Emile-Geay[a,1], Nicholas P. McKay[b], Gregory J. Hakim[c], Deborah Khider[a,d], Toby R. Ault[e], Eric J. Steig[f], Sylvia Dee[g], and James W. Kirchner[h,i,j]

[a]Department of Earth Sciences, University of Southern California, Los Angeles, CA 90089; [b]School of Earth and Sustainability, Northern Arizona University, Flagstaff, AZ 86011; [c]Department of Atmospheric Sciences, University of Washington, Seattle, WA 98195; [d]Information Sciences Institute, University of Southern California, Marina del Rey, CA 90292; [e]Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY 14853; [f]Department of Earth and Space Sciences, University of Washington, Seattle, WA 98195; [g]Department of Earth, Environmental, and Planetary Sciences, Rice University, Houston, TX 77005; [h]Department of Environmental System Sciences, Swiss Federal Institute of Technology (ETH) Zürich, CH-8092 Zürich, Switzerland; [i]Mountain Hydrology Research Unit, Swiss Federal Research Institute for Forest, Snow and Landscape Research (WSL), CH-8903 Birmensdorf, Switzerland; and [j]Department of Earth and Planetary Science, University of California, Berkeley, CA 94720

Edited by Mark H. Thiemens, University of California, San Diego, La Jolla, CA, and approved March 12, 2019 (received for review June 25, 2018)

Climate records exhibit scaling behavior with large exponents, resulting in larger fluctuations at longer timescales. It is unclear whether climate models are capable of simulating these fluctuations, which draws into question their ability to simulate such variability in the coming decades and centuries. Using the latest simulations and data syntheses, we find agreement for spectra derived from observations and models on timescales ranging from interannual to multimillennial. Our results confirm the existence of a scaling break between orbital and annual peaks, occurring around millennial periodicities. That both simple and comprehensive ocean–atmosphere models can reproduce these features suggests that long-range persistence is a consequence of the oceanic integration of both gradual and abrupt climate forcings. This result implies that Holocene low-frequency variability is partly a consequence of the climate system's integrated memory of orbital forcing. We conclude that climate models appear to contain the essential physics to correctly simulate the spectral continuum of global-mean temperature; however, regional discrepancies remain unresolved. A critical element of successfully simulating suborbital climate variability involves, we hypothesize, initial conditions of the deep ocean state that are consistent with observations of the recent past.

climate variability | spectral analysis | scaling laws | model evaluation

A grand challenge for climate science is to accurately simulate low-frequency variability (changes occurring on scales longer than a few years). Of particular interest is the temporal spectrum of surface temperature, whose peaks indicate dominant oscillations and whose continuum describes energy transfers between scales (1). This continuum is often characterized by its scaling exponent β, where the power spectral density (PSD) S and the frequency f satisfy the power law relationship:

$$S(f) \propto f^{-\beta}. \quad [1]$$

The larger the exponent is, the longer the memory of past events. A spectral depiction of climate change dates back to ref. 2, which investigated Holocene climate variability in the North Atlantic sector using various records and connected the observed spectral pattern to the thermal inertia of the ocean and cryosphere. Two years later, Mitchell (3) gave an early qualitative overview of the spectrum of climate variability based on causal mechanisms. Later, Pelletier (4) estimated the PSD of regional atmospheric temperature from synoptic to multimillennial and longer scales, using instrumental and ice-core data, and explained the observed scaling exponents with a vertical turbulent transport model. In a landmark paper, Huybers and Curry (5) added many more data sources, extended the analysis to much longer timescales, and proposed that "annual, Milankovitch and continuum temperature variability together represent the response to deterministic insolation forcing" (ref. 5, p. 329). They identified two distinct scaling regimes, with a break at centennial scales, but did not provide an explanation for this break.

Recent studies have looked for similar behavior in temperature fields simulated by climate models and show that the scaling exponents that describe the simulated temperature variability are too small compared with those from instrumental (6, 7) and paleoclimate obervations (8–10). There are at least four reasons to reserve caution in this comparison. First, climate proxies are known to filter climate inputs (11, 12), so simulated temperatures and proxy measurements are not directly comparable (13, 14). Second, the comparisons done to date have included a limited number of proxies with subcentennial resolution ($<$20); it is therefore critical to update this picture with more complete data syntheses, including annually-resolved observations. Third, the model evaluations mentioned above have focused on simulations of the past millennium (850–1850 CE); no systematic comparison has been carried out with longer transient simulations. Finally, a lack of global coverage beyond the past millennium restricted previous studies to focus on regional temperature variability, yet global temperature is more informative of changes in Earth's energy budget. Thus, we focus here on the global signal.

Here we address these challenges and find a variety of climate models to be consistent with scaling behavior observed across a range of paleoclimate archives. The robustness of this

Significance

Climate models are foundational to formulations of climate policy and must successfully reproduce key features of the climate system. The temporal spectrum of observed global surface temperature is one such critical benchmark. This spectrum is known to obey scaling laws connecting astronomical forcings, from orbital to annual scales. We provide evidence that the current hierarchy of climate models is capable of reproducing the increase in variance in global-mean temperature at low frequencies. We suggest that successful climate predictions at decadal-to-centennial horizons hinge critically on the accuracy of initial and boundary conditions, particularly for the deep ocean state.

Author contributions: F.Z., J.E.-G., N.P.M., T.R.A., and S.D. designed research; F.Z. and J.E.-G. performed research; F.Z., J.E.-G., D.K., and J.W.K. contributed analytic tools; F.Z., J.E.-G., N.P.M., G.J.H., D.K., S.D., and J.W.K. analyzed data; and F.Z., J.E.-G., N.P.M., G.J.H., D.K., T.R.A., E.J.S., S.D., and J.W.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: julieneg@usc.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1809959116/-/DCSupplemental.

Published online April 15, 2019.

result has important implications for climate predictability. We also provide an explanation for the transition between scaling regimes.

Completing the Continuum

We first estimate the spectrum of global-average temperature variability, leveraging new measurements and data syntheses (refs. 15 and 18–21 and Table 1) as well as improved spectral methods (ref. 26 and *Methods*). Notably, the latest Past Global Changes, 2k Network Working Group (PAGES2k) compilation (18), which gathers observations from coral, glacial ice core, marine and lake sedimentary, speleothem, tree-ring, and documentary archives, allows us to fill the aforementioned spectral gap in the centennial to millennial band. Individual spectra exhibit scaling behavior (lack of obvious scaling breaks; *SI Appendix*, Fig. S7 and section 5) for all archive types except for glacier ice, where high-frequency signals are known to be damped by a range of processes, including diffusion and compaction (12, 27). The median scaling exponents are close to 0.45, with no clear spatial trend (*SI Appendix*, Fig. S8). Our analysis confirms the existence of scaling behavior in the decadal-to-centennial range, robustly across archive types. Each archive is affected by different processes and timescales (11, 12), distorting spectra in various ways. However, none of these processes can create scaling on their own (14), suggesting that the observed scaling behavior is a property of the climate and not the proxy archives. This picture may change when more complete proxy system models are considered.

Next, we use this dataset to estimate global mean surface temperature with a state-of-the-art paleoclimate state estimation methodology, the Last Millennium Reanalysis (LMR) (16, 17). LMR uses an ensemble Kalman filter to optimally combine information from transient climate model simulations and annually-resolved paleoclimate observations (18) (*SI Appendix*, section 2). Importantly, this approach uses climate models only to provide physically plausible covariances within and among climate fields; all of the temporal variability, and thus the spectral behavior, originates in the paleoclimate observations. Fig. 1 shows scaling behavior in the PAGES2k-based LMR estimate, with decadal-to-centennial scaling exponents around unity, in good agreement with global instrumental temperature (HadCRUT4, ref. 15).

Progressing toward lower frequencies, we consider the ice-core–based reconstruction from EPICA Dome C (21), which nearly doubles the coverage of the Vostok data used by Huybers and Curry (5) (800 vs. 420 ky), as well as two recent estimates based on marine sediments: the global average surface temperature reconstruction of ref. 19 (S16 GAST), based primarily on sea-surface temperature proxies (alkenones, Mg/Ca, and faunal assemblages) and the latest benthic stack based on foraminiferal δ18O (ProbStack, ref. 20). All three of these datasets show consistently steep centennial-to-orbital scaling exponents around 2.5 (Fig. 1). As before, the robustness across different paleoclimate archive, sensor, and observation types is a cogent indication that they are features of the climate, not proxy-related artifacts. The EPICA Dome C (EDC) spectrum flattens considerably at scales shorter than millennial, which can be traced to its local nature (*SI Appendix*, Fig. S6) and possible aliasing of the annual cycle (28).

Overall, Fig. 1 highlights a scaling break between the decadal-to-centennial band ($\beta \approx 1$) and the centennial-to-Milankovitch band ($\beta \approx 2.5$). This confirms the existence of the two scaling regimes pointed out by Huybers and Curry (5), who placed this transition at centennial scales. In contrast, our analysis shifts this scaling break to the vicinity of the millennial scale. As shown in *SI Appendix*, section 4, the location of the break in frequency space is quite variable from one record to the next, and only with global syntheses does it emerge at millennial scales. One possibility is that the frequency of the scaling break depends on weighting proxies that record more of the global-mean, rather than a local signal (e.g., figure 7 in ref. 29). Another possibility is that the spectral break is peculiar to the time of analysis, i.e., not a property of the stationary statistics of the climate system. In any case, do climate model simulations even contain a spectral transition?

Simulating the Continuum

We now consider long transient integrations of general circulation models, including the Palaeoclimate Modeling Intercomparison Project Phase 3 (PMIP3) last millennium (PMIP3 past1000) simulations (25) appended with "historical" Coupled Model Intercomparison Project Phase 5 (CMIP5) simulations and the TraCE-21ka suite of experiments (22). We also include simulations from two Earth System Models of Intermediate Complexity (EMICs) covering the last deglaciation, DG_{ns} (23) and SIM2bl (24) (*SI Appendix*, section 2). Over the decadal-to-centennial band, Fig. 2 (*Upper*) indicates that the PMIP3 simulations share similar scaling exponents around unity, consistent with observed spectra (Fig. 1). Specifically, Fig. 2, *Upper Inset* shows the distribution of the scaling exponents of the PMIP3 simulations against those obtained from the PAGES2k-based LMR, in which we find good agreement for the central quantiles and large inter-model spread. This agreement stands in contrast to previous work showing differences between spectra in climate models and data (8–10, 14, 30).

One important distinction between our study and previous ones is that our comparisons focus on global, rather than regional, variability since the Last Glacial Maximum. Given a model's finite resolution, accurately modeling local and regional variability is more difficult than modeling global variability (*SI Appendix*, section 4). Model biases diminish from local to continental scales, and local variability at small spatial scales, reflecting short temporal scales, is smoothed. As shown by refs. 29 and 31, this smoothing steepens global spectra relative to local spectra.

Table 1. The overview information of the instrumental observations, reanalysis, reconstructions based on proxy records, and model simulations that are used in this paper

Dataset	Type	Time span	Exponent estimation scales, y
Observations/reconstructions			
HadCRUT4 (15)	Instrumental	1850–2017 AD	1/6–50
PAGES2k/LMR GAST (16, 17)	Reanalysis	1–2000 AD	2–1,000
S16 GAST (19)	Reconstruction	2 MyBP–950 AD	2,000–100,000
ProbStack (20)	Proxy	5 MyBP–1950 AD	10,000–100,000
EDC (21)	Reconstruction	800 MyBP–1911 AD	1,500–50,000
Model simulations			
TraCE-21ka (22)	Deglaciation	22 kyBP–1979 AD	400–2,000, 20–400
DG_{ns} (23)	Deglaciation	18 kyBP–3 kyBP	400–2000, 20–400
SIM2bl (24)	Deglaciation	21 kyBP–1949 AD	400–2,000, 20–400
PMIP3 (25)	Last millennium	850–1850 AD, 850–2011 AD	2–500

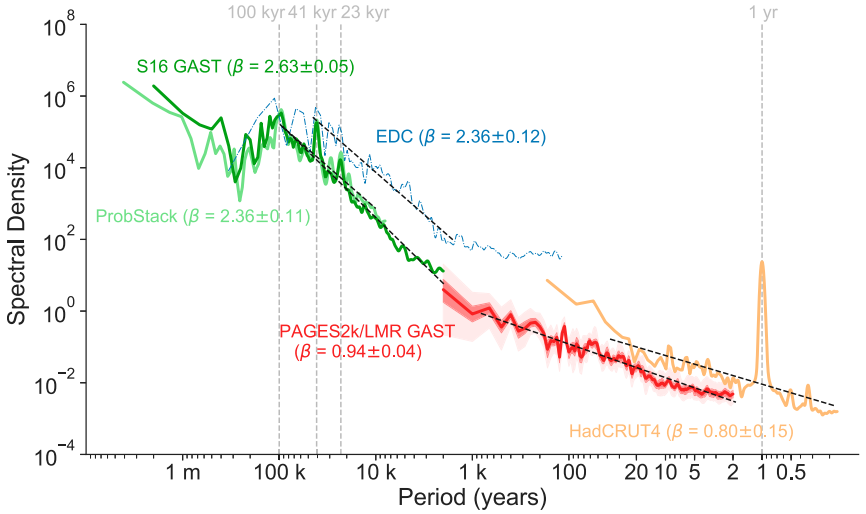

Fig. 1. A spectral estimate of the global-average surface temperature variability using instrumental and paleoclimate datasets (scaled degrees Kelvin), as well as proxy-based reconstructions of surface temperature variability. HadCRUT4: The Met Office Hadley Center gridded dataset of global historical surface temperature anomalies (15). PAGERS2k/LMR GAST: The Last Millennium Reanalysis framework (16, 17) applied to the PAGES2k dataset (18). The thick red curve denotes the median power spectral density (PSD), the dark red shaded area denotes the interquartile range, and the light red shaded area denotes the central 95% range, from 2.5% to 97.5%. S16 GAST: The reconstruction of global average surface temperature (19). ProbStack: A probabilistic Pliocene–Pleistocene stack of benthic δ^{18}O (20). βs denote the estimated scaling exponents over each appropriate frequency band. Details of their estimation are presented in *Methods* and *SI Appendix*, section 3. The regional dataset EPICA Dome C (EDC) Ice Core 800-ky Deuterium Data and Temperature Estimates (21) is included as a point of comparison.

Fig. 2, *Lower* shows a spectral analysis of the TraCE-21ka, DG_{ns}, and SIM2bl transient simulations, which cover the last 10–20 ky (*SI Appendix*, section 2). All three simulations show a similar scaling break around timescales of 300–1,000 y. Henceforth, we define β_{CM} as the centennial-to-millennial scaling exponent (estimated over 400–2,000 y) and β_{DC} as the decadal-to-centennial scaling exponent (estimated over 20–400 y). All three simulations display $\beta_{CM} \approx 2.5$ and $\beta_{DC} \approx 1$ consistent with the observed spectra (Fig. 1). These results are robust to definitions of the scaling ranges (*SI Appendix*, section 9). However, these simulated β_{DC}s arise for different reasons than in the PMIP3 past1000 simulations: (*i*) None of these deglacial simulations are subject to volcanic aerosol forcing, the largest source of low-frequency variance for PMIP3 past1000 simulations (32); (*ii*) DG_{ns} does not include the industrial warming period (Table 1), yet it shows similar β_{DC} compared with TraCE-21ka and SIM2bl, as well as modern and paleoclimate observations, implying that the industrial warming period is not the only explanation for $\beta_{DC} \approx 1$ (*Climate Implications*); (*iii*) the TraCE-21ka experiment was designed in part to capture climate variability inferred from Greenland ice-core records and forced, e.g., through freshwater fluxes, to capture that variability. This presents the possibility of circular logic to conclusions based on TraCE-21ka spectra, although the simulation does remarkably well in reproducing the phase and magnitude of millennial-scale variability in Southern Hemisphere records, for which it was not tuned (33).

A Tale of Two Regimes

What physical mechanisms underlie the scaling break? Nilsen et al. (34) suggest that Holocene temperature reconstructions are consistent with a single scaling regime and that the scaling break likely originates from the large-amplitude Dansgaard–Oeschger (D-O) events of the past glacial period. As long as the analyzed records contain such abrupt events, they argue that one should expect a scaling break in the frequency domain. This is supported by our analysis of 253 Holocene records (*SI Appendix*, Fig. S12), showing a single scaling regime throughout the interval.

To further test this idea, we investigate the spectral density of the output of the TraCE-21ka full simulation before and after 8 kyBP (*SI Appendix*, Fig. S13 and section 7). This choice avoids the 8.2-ky event, thus delineating a period of rapid transition before this point (the deglaciation) and a stable climate afterward. The result indicates that the time series before 8 kyBP shows a PSD similar to that of the full series, while the time series after 8 kyBP lacks a scaling break. This suggests that the scaling break originates in the early part of the time series. A scalogram of the TraCE-21ka full simulation (*SI Appendix*, Fig. S14) reveals two underlying factors for the scaling break: (*i*) abrupt, large-amplitude events as suggested above and (*ii*) the gradual transition from glacial to interglacial states. The first factor is identified by the energetic area in the scalogram around 12.5 kyBP between periods of 500 y and 2,000 y, coincident with the simulated Bölling–Alleröd/Younger Dryas couplet. This mechanism is reproducible using simple models (*SI Appendix*, section 7). Such impulses create bumps in the PSD (*SI Appendix*, Figs. S15 and S17, *Right*), which cascade down to smaller scales, disappearing at periods near 300 y, where the scaling break occurs. The second factor is visible around periods of 5 ky, between 7 kyBP and 20 kyBP, and reflects orbitally driven changes in the Earth system (*SI Appendix*, section 8). They mirror the pattern seen in the CO_2 time series obtained from the EPICA Dome C ice core (*SI Appendix*, Fig. S22) as well as the best estimates of the ice-volume equivalent sea-level function (*SI Appendix*, Fig. S23).

To further disentangle the influence of the various forcings, we leverage the TraCE-21ka single-forcing experiments (22). Fig. 3 confirms that orbital forcing (ORB) acts as the second factor, driving the slow transition from glacial to interglacial states; this is sufficient to generate a scaling break at millennial scales (orange curve). Forcing from greenhouse gases (GHG) and transient ice sheets (ICE) acts to amplify this transition. On the other hand, transient Northern Hemisphere meltwater fluxes (MWF) act as the first factor: They generate a bump in the PSD that shifts the scaling break to periods near 300 y. Because these signals do not propagate instantly around the globe, the break would be expressed in

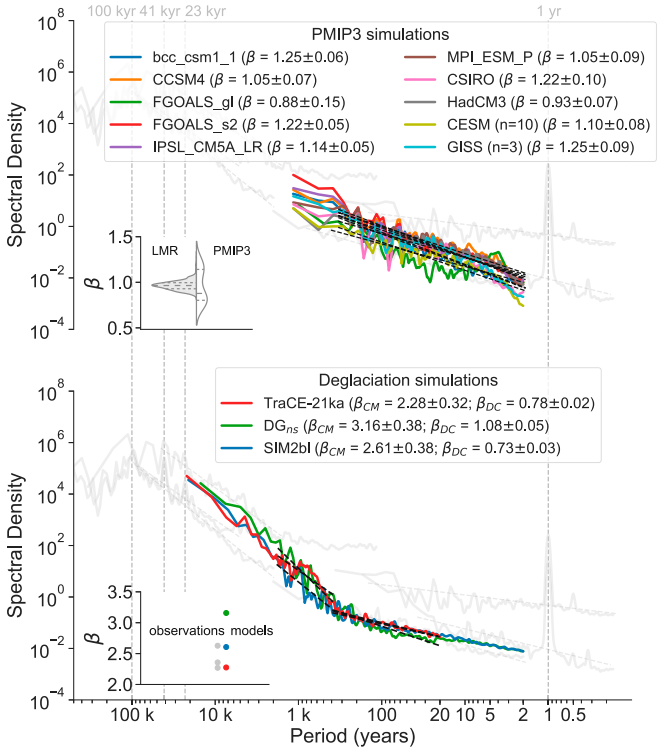

Fig. 2. The power spectral density (PSD) of transient model simulations. (*Upper*) β is estimated over 2–500 y. *Inset* compares distributions of the scaling exponents (estimated over 2–500 y) of GAST in PAGES2k-based LMR vs. the PMIP3 simulations. (*Lower*) β_{CM} is the centennial-to-millennial–scale exponent estimated over 400–2,000 y, while β_{DC} is the decadal-to-centennial–scale exponent estimated over 20–400 y. *Inset* compares the β values of the model simulations (red, green, and blue circles) and those of the observations (gray circles). The gray curves are identical to those in Fig. 1. Note that curves labeled "CESM" and "GISS" are the ensemble average of the PSDs of 10 and 3 members for each model, respectively.

different records at different scales, ranging from centennial to millennial.

Climate Implications

By incorporating a wide range of proxy data, models, and data assimilation approaches to climate variability, we find two scaling regimes linking orbital to annual scales, as Huybers and Curry (5) found from regional records. These regimes are robust across multiple observation types. Increasing the density of records in the centennial band by two orders of magnitude, we find that the regime transition for global-average temperature variability occurs at millennial scales.

At scales shorter than millennial, we find good agreement between modeled and observationally derived scaling exponents. Yet the spectra are qualitatively different: The same CMIP5 models have been shown to oversimulate interannual variance and undersimulate decadal variance (7, 10, 35). This results in steeper scaling at high frequencies and flatter scaling at decadal and longer timescales, in the past 1,000 simulations.

At scales up to 10^4 y, we find that models of varying complexity closely reproduce the observed scaling laws over the past deglaciation, including the scaling break around 10^3 y. In the TraCE-21k simulations, this transition is primarily driven by orbital forcing and modulated by freshwater fluxes. This raises the question of what level of complexity is required for models to correctly reproduce the observed continuum. All models considered here lack interactive ice sheets, and most lack an interactive carbon cycle. Such models therefore require information about these systems to be supplied via boundary conditions to reproduce observed climate trajectories (and therefore, spectra). In reality, of course, insolation is the only true forcing on these timescales; ice sheet topography, greenhouse gas levels, and freshwater fluxes all are Earth system responses to this forcing. The response of the climate system to insolation forcing is state dependent, which introduces a stochastic (unpredictable) component to the response. A surprising finding is that even simplified models like ECBilt-CLIO (used in SIM2bl) can produce a realistic global-average temperature continuum at submillennial scales when supplied with information about gradual climate forcings over the past deglaciation (Fig. 2, *Lower*).

In contrast, most of the decadal-to-centennial variability in PMIP3 past1000 simulations originates from volcanic forcing (32). This dominance is partly the result of such simulations being overly sensitive to stratospheric aerosol loading (36), due to incomplete representations of stratospheric aerosol chemistry (37). However, our analysis suggests that a decadal-to-centennial climate continuum could be inherited from boundary conditions that far predate the last millennium. This is supported by the linear analysis of ref. 38, which expressed temperature at time t as a convolution between the transient forcing and the impulse response function to that forcing at all past instants. Comparing simulated and observed temperature, one evaluates not only the model (which approximates the impulse response), but also the forcing. It is a distinct possibility that such evaluations improperly place the blame on the models, while it should lie in the forcing.

Put another way, a plausible explanation for our results is that, on a global scale, the past millennium still contains echoes of the deglaciation. The systems' adjustment to smoothly varying insolation generates substantial decadal-to-centennial variability at the surface, despite the forcings containing little energy at these scales (*SI Appendix*, Figs. S21–S23). Thus, our results affirm and extend Huybers and Curry's (5) conclusion that continuum temperature variability is an integral part of the response to insolation forcing; decadal-to-centennial variability in global-average temperature is partly a consequence of changes in Earth's orbital parameters. This is also consistent with the theoretical results of ref. 39.

We surmise that the lack of low-frequency variability in PMIP3 past1000 simulations before 1850 AD is related to these simulations being initialized from a quasi-steady state in equilibrium with boundary conditions characteristic of 850 AD. Were the

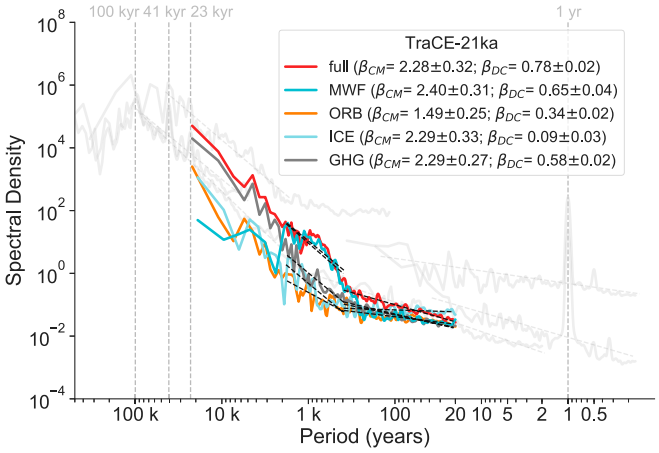

Fig. 3. Effect of forcings on scaling behavior in the TraCE-21ka simulations. The full simulation is forced by transient Northern Hemisphere meltwater fluxes (MWF), orbital forcing (ORB), changing continental ice sheets (ICE), and transient greenhouse gas forcing (GHG). Conventions are identical to those in Fig. 2.

same PMIP3 models to be initialized from an ocean state that bore the imprint of the last deglaciation, we predict that they would exhibit more vigorous internal variability at decadal-to-centennial scales and that the fraction of surface temperature variance imputable to volcanic forcing would be relatively lower. This "echoes" hypothesis may be tested in dedicated experiments with a hierarchy of climate models.

That the ocean state should integrate forcings over a long period is not a novel idea (39–41); the surprising implication is that this property could directly bear on the amplitude of climate variability at scales far shorter than orbital and therefore on the perceived reliability of model-based predictions on societally relevant horizons. Were the echoes hypothesis to be confirmed by subsequent studies, it would bear favorably on the prospects for prediction, at least for the global mean. This suggests two pathways to sensibly evaluate model behavior. The first one is to use comprehensive Earth system models (including dynamic ice sheets) to simulate the climate continuum with sole knowledge of orbital forcing. This would require integrating such models over a full glacial cycle (ideally, multiple ones), which is presently in reach of only the largest available supercomputers (42). The second way would be to initialize shorter simulations (e.g., past1000) from a state informed by paleo-observations of the deep ocean state and diagnose their temperature continuum as done here. Advances in paleoclimate state estimation (43) may soon make this possible.

Scaling behavior is nearly universal, appearing in contexts as diverse as fluid dynamics (44), hydrochemistry (26), metabolism (45, 46), economic growth (47), and city size (48). The existence of scaling behavior in climate time series is therefore unsurprising, although there is currently no consensus on its cause (1, 38, 41). For this reason, it is not obvious whether the climate models investigated here reproduce this behavior for the right reasons. Indeed, the notable consistency between scaling behavior in such a wide range of models and that in observations suggests that it is a relatively crude constraint. A stricter metric would be to aim for consistency at local or regional scales, which are of greatest interest for adaptation and planning and where discrepancies between models and observations remain (7, 13). Enhanced data availability may also enable the evaluation of higher-order spectral moments, which could help reveal other differences. Based on present evidence, however, we conclude that it would be premature to dismiss the capabilities of Earth system models to predict global trends on societally relevant timescales (10–100 y). We suggest that the key to simulating the climate continuum over such scales lies in properly initializing the low-frequency coupled state of the climate system, especially the ocean; how to do so remains an open problem.

Methods

Spectral Estimation. Because paleoclimate data are often unevenly sampled in the time domain, a common strategy for their analysis is to first perform interpolation so that traditional spectral analysis methods, such as the periodogram (49) or multitaper method (MTM) (50), can be applied. However, interpolation can bias spectral estimation as well. To sidestep interpolation, the Lomb–Scargle Fourier transform (51, 52) is often used, yet it is known to overestimate the amplitudes at the high-frequency end (53). Moreover, Fourier transform-based methods assume stationary processes, an oft-violated assumption in geophysical time series. The resulting edge effects are typically mitigated by detrending (54), an imperfect fix because of the intrinsic difficulty of identifying the trend without compromising the signal.

We address these challenges via the weighted wavelet Z-transform (WWZ) (55), which suppresses the energy leakage caused by the data gaps. It is wavelet based and therefore does not rely on interpolation or detrending. In particular, we use its variant (26), in which basis rotations mitigate the numerical instability that occurs in pathological cases with the original algorithm.

The WWZ method has one adjustable parameter, a decay constant that balances the time resolution and frequency resolution of the wavelet analysis. The smaller this constant is, the sharper the peaks. We choose the value 0.001 to obtain smooth spectra that lend themselves to better scaling exponent estimation, while still capturing the main periodicities. For the purpose of showing the scalogram, we use the larger value $(8\pi^2)^{-1}$, justified elsewhere (55, 56). The method is implemented via the Pyleoclim Python package (57). Details are provided in *SI Appendix*, section 3.

Estimation of Scaling Exponents. Taking the log on both sides of Eq. **1** yields $\log S \propto (-\beta) \log f$. Therefore, β is estimated via linear regression in log space. To mitigate biases arising from nonuniform spacing in log coordinates (more points are located in the high-frequency side than in the low-frequency side), we apply Huybers and Curry's (5) frequency binning procedure. When estimating the scaling exponents of the HadCRUT4 dataset, the annual cycle is removed to avoid biasing the estimate. We also estimate the scaling exponents over frequency ranges where the power law is well followed, which leads to different frequency intervals for different series. Conclusions do not depend sensitively on these definitions, as similar results are obtained with overlapping intervals.

ACKNOWLEDGMENTS. We thank Raphi Neukom for compiling global means of these simulations. We acknowledge the World Climate Research Program's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the PMIP3 modeling groups for producing and making available their model output. The US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support for CMIP. PAGES2k Consortium members are acknowledged for providing input proxy data. F.Z. was supported by the University of Southern California. J.E.-G., D.K., and N.P.M acknowledge support from the US National Science Foundation (Awards ICER-1541029 and EAR-1347213). J.E.-G., G.J.H., and E.J.S. acknowledge support from the National Oceanographic and Atmospheric Administration (Award NA14OAR4310176).

1. Lovejoy S, Schertzer D (2013) *The Weather and Climate: Emergent Laws and Multifractal Cascades* (Cambridge Univ Press, Cambridge, UK), 1st Ed.
2. Kutzbach JE, Bryson RA (1974) Variance spectrum of Holocene climatic fluctuations in the North Atlantic sector. *J Atmos Sci* 31:1958–1963.
3. Mitchell JM (1976) An overview of climatic variability and its causal mechanisms. *Quat Res* 6:481–493.
4. Pelletier JD (1998) The power spectral density of atmospheric temperature from time scales of 10^{-2} to 10^{6} yr. *Earth Planet Sci Lett* 158:157–164.
5. Huybers P, Curry W (2006) Links between annual, Milankovitch and continuum temperature variability. *Nature* 441:329–332.
6. Govindan RB, et al. (2002) Global climate models violate scaling of the observed atmospheric variability. *Phys Rev Lett* 89:028501.
7. Laepple T, Huybers P (2014) Global and regional variability in marine surface temperatures. *Geophys Res Lett* 41:2528–2534.
8. Laepple T, Huybers P (2014) Ocean surface temperature variability: Large model–data differences at decadal and longer periods. *Proc Natl Acad Sci USA* 111: 16682–16687.
9. Ault TR, Cole JE, Overpeck JT, Pederson GT, Meko DM (2014) Assessing the risk of persistent drought using climate model simulations and paleoclimate data. *J Clim* 27:7529–7549.
10. Parsons LA, et al. (2017) Temperature and precipitation variance in CMIP5 simulations and paleoclimate records of the last millennium. *J Clim* 30:8885–8912.
11. Evans MN, Tolwinski-Ward SE, Thompson DM, Anchukaitis KJ (2013) Applications of proxy system modeling in high resolution paleoclimatology. *Quat Sci Rev* 76:16–28.
12. Dee S, et al. (2015) PRYSM: An open-source framework for PRoxY system modeling, with applications to oxygen-isotope systems. *J Adv Model Earth Syst* 7: 1220–1247.
13. Laepple T, Huybers P (2013) Reconciling discrepancies between Uk37 and Mg/Ca reconstructions of Holocene marine temperature variability. *Earth Planet Sci Lett* 375:418–429.
14. Dee SG, et al. (2017) Improved spectral comparisons of paleoclimate models and observations via proxy system modeling: Implications for multi-decadal variability. *Earth Planet Sci Lett* 476:34–46.
15. Morice CP, Kennedy JJ, Rayner NA, Jones PD (2012) Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *J Geophys Res* 117:D08101.
16. Hakim GJ, et al. (2016) The last millennium climate reanalysis project: Framework and first results. *J Geophys Res Atmos* 121:6745–6764.
17. Tardif R, et al. (2018) Last millennium reanalysis with an expanded proxy database and seasonal proxy modeling. *Clim Past Discuss* 2018:1–37.
18. PAGES2k Consortium (2017) A global multiproxy database for temperature reconstructions of the Common Era. *Sci Data* 4:170088.
19. Snyder CW (2016) Evolution of global temperature over the past two million years. *Nature* 538:226–228.

20. Ahn S, Khider D, Lisiecki LE, Lawrence CE (2017) A probabilistic pliocene–pleistocene stack of benthic δ18O using a profile hidden Markov model. *Dyn Stat Clim Syst* 2: 1–16.
21. Jouzel J, et al. (2007) Orbital and Millennial Antarctic climate variability over the past 800,000 years. *Science* 317:793–796.
22. Liu Z, et al. (2009) Transient simulation of last deglaciation with a new mechanism for Bølling-Allerød warming. *Science* 325:310–314.
23. Menviel L, Timmermann A, Timm OE, Mouchet A (2011) Deconstructing the last glacial termination: The role of millennial and orbital-scale forcings. *Quat Sci Rev* 30:1155–1172.
24. Timm O, Timmermann A (2007) Simulation of the last 21 000 years using accelerated transient boundary conditions. *J Clim* 20:4377–4401.
25. Braconnot P, et al. (2012) Evaluation of climate models using palaeoclimatic data. *Nat Clim Change* 2:417–424.
26. Kirchner JW, Neal C (2013) Universal fractal scaling in stream chemistry and its implications for solute transport and water quality trend detection. *Proc Natl Acad Sci USA* 110:12213–12218.
27. Cuffey KM, Steig EJ (1998) Isotopic diffusion in polar firn: Implications for interpretation of seasonal climate parameters in ice-core records, with emphasis on central Greenland. *J Glaciol* 44:273–284.
28. Laepple T, et al. (2018) On the similarity and apparent cycles of isotopic variations in East Antarctic snow pits. *Cryosphere* 12:169–187.
29. Rypdal K, Rypdal M, Fredriksen HB (2015) Spatiotemporal long-range persistence in Earth's temperature field: Analysis of stochastic–diffusive Energy balance models. *J Clim* 28:8379–8395.
30. Lovejoy S, Schertzer D, Varon D (2013) Do GCMs predict the climate . . .or macroweather? *Earth Syst Dynam* 4:439–454.
31. Fredriksen HB, Rypdal K (2015) Spectral characteristics of instrumental and climate model surface temperatures. *J Clim* 29:1253–1268.
32. Schurer AP, Tett SFB, Hegerl GC (2013) Small influence of solar variability on climate over the past millennium. *Nat Geosci* 7:104–108.
33. Pedro JB, et al. (2016) The spatial extent and dynamics of the Antarctic Cold Reversal. *Nat Geosci* 9:51–55.
34. Nilsen T, Rypdal K, Fredriksen HB (2016) Are there multiple scaling regimes in Holocene temperature records? *Earth Syst Dynam* 7:419–439.
35. Ault TR, Deser C, Newman M, Emile-Geay J (2013) Characterizing decadal to centennial variability in the equatorial pacific during the last millennium. *Geophys Res Lett* 40:3450–3456.
36. Stoffel M, et al. (2015) Estimates of volcanic-induced cooling in the northern hemisphere over the past 1,500 years. *Nat Geosci* 8:784–788.
37. LeGrande AN, Tsigaridis K, Bauer SE (2016) Role of atmospheric chemistry in the climate impacts of stratospheric volcanic injections. *Nat Geosci* 9:652–655.
38. Fredriksen HB, Rypdal M (2017) Long-range persistence in global surface temperatures explained by linear multibox energy balance models. *J Clim* 30: 7157–7168.
39. Rypdal M, Rypdal K (2014) Long-memory effects in linear response models of Earth's temperature and implications for future global warming. *J Clim* 27:5240–5258.
40. Hasselman K (1976) Stochastic climate models. Part I. Theory. *Tellus* 28:473–485.
41. Fraedrich K, Luksch U, Blender R (2004) $1/f$ model for long-time memory of the ocean surface temperature. *Phys Rev E* 70:037301.
42. Abe-Ouchi A, et al. (2013) Insolation-driven 100,000-year glacial cycles and hysteresis of ice-sheet volume. *Nature* 500:190–193.
43. Emile-Geay J, Erb MP, Hakim GJ, Steig EJ, Noone DC (2017) Climate dynamics with the last millennium reanalysis. *PAGES Mag* 25:162.
44. Lovejoy S, Schertzer D, Lilley M, Strawbridge KB, Radkevich A (2008) Scaling turbulent atmospheric stratification. I: Turbulence and waves. *Q J R Meteorol Soc* 134:277–300.
45. West GB, Brown JH (2005) The origin of allometric scaling laws in biology from genomes to ecosystems: Towards a quantitative unifying theory of biological structure and organization. *J Exp Biol* 208:1575–1592.
46. Marquet PA, et al. (2014) On theory in ecology. *BioScience* 64:701–710.
47. Stanley HE, Mantegna RN (1995) Scaling behaviour in the dynamics of an economic index. *Nature* 376:46–49.
48. Lobo J, Bettencourt LMA, Strumsky D, West GB (2013) Urban scaling and the production function for cities. *PLoS One* 8:e58407.
49. Schuster A (1898) On the investigation of hidden periodicities with application to a supposed 26 day period of meteorological phenomena. *J Geophys Res* 3:13–41.
50. Thomson DJ (1982) Spectrum estimation and harmonic analysis. *Proc IEEE* 70:1055–1096.
51. Lomb NR (1976) Least-squares frequency analysis of unequally spaced data. *Astrophys Space Sci* 39:447–462.
52. Scargle JD (1982) Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys J* 263:835–853.
53. Schulz M, Mudelsee M (2002) REDFIT: Estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. *Comput Geosci* 28:421–426.
54. Wu Z, Huang NE, Long SR, Peng CK (2007) On the trend, detrending, and variability of nonlinear and nonstationary time series. *Proc Natl Acad Sci USA* 104: 14889–14894.
55. Foster G (1996) Wavelets for period analysis of unevenly sampled time series. *Astron J* 112:1709.
56. Witt A, Schumann AY (2005) Holocene climate variability on millennial scales recorded in Greenland ice cores. *Nonlinear Process Geophys* 12:345–352.
57. Khider D, Zhu F, Hu J, Emile-Geay J (2018) *LinkedEarth/Pyleoclim_util: Pyleoclim release v0.4.8*. Available at https://doi.org/10.5281/zenodo.1340732. Accessed August 6, 2018.