Improved spectral comparisons of paleoclimate models and observations via proxy system modeling: implications for multi-decadal variability

S.G. Dee^{a,*}, L. Parsons^b, G. Loope^b, J. T. Overpeck^b, T.R. Ault^c, J. Emile-Geay^d

^a*Department of Earth, Environmental, and Planetary Sciences, Brown University, Providence, RI 02912*

^b*Department of Geosciences, University of Arizona, Tucson, AZ 85721 USA*

^c*Department of Earth and Atmospheric Sciences, Cornell University, Ithaca, NY, USA*

^d*Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA*

Abstract

The spectral characteristics of paleoclimate observations spanning the last millennium suggest the presence of significant low-frequency (multi-decadal to centennial) variability in the climate system. Since this low-frequency climate variability is critical for climate predictions on societally-relevant scales, it is essential to establish whether General Circulation models (GCMs) are able to simulate it faithfully. Recent studies find large discrepancies between models and paleoclimate data at low frequencies, prompting concerns surrounding the ability of GCMs to predict long term, high-magnitude variability under greenhouse forcing (??). However, efforts to ground climate model simulations directly in paleoclimate observations are impeded by fundamental differences between models and the proxy data: proxy systems often record a multivariate and/or nonlinear response to climate, precluding a direct comparison to GCM output. In this paper we bridge this gap via a forward modeling approach, coupled to an isotope-enabled GCM. This allows us to disentangle the various contributions to signals embedded in ice cores, speleothem calcite, corals, tree-ring width, and tree-ring cellulose. The paper addresses the following questions: (1) do forward modeled “pseudoproxies” exhibit variability comparable to proxy data? (2) if not, which processes alter the shape of the spectrum of simulated climate variability, and are these processes broadly distinguishable from climate? We apply our method to representative case studies, and parlay these insights into an analysis of the PAGES2k database (?). We find that current proxy system models (PSMs) can help resolve model-data discrepancies on interannual to decadal timescales, but cannot account for the mismatch in variance on multi-decadal to centennial timescales. We conclude that, specific to this set of PSMs and isotope-enabled model, the paleoclimate record may exhibit larger low-frequency variability than GCMs currently simulate, indicative of

*Corresponding author
Preprint submitted to Elsevier, April 18, 2019
Email addresses: sylvia_11.dee@brown.edu (S.G. Dee), lukeaparsons@email.arizona.edu (L. Parsons), G.Loope@arizona.edu (G. Loope), jto@email.arizona.edu (J. T. Overpeck), toby.ault@cornell.edu (T.R. Ault), julieneg@usc.edu (J. Emile-Geay)

incomplete physics and/or forcings.

9 *Keywords:* climate variability, general circulation models, data-model comparison, paleoclimatology

10 **1. Introduction**

11 Our understanding of the complex dynamics of climate response to anthropogenic forcing rests jointly
12 upon observations over the instrumental period, general circulation models (GCMs), and paleoclimate data.
13 GCMs provide a basis for exploring the mechanisms driving forced and stochastic climate variability; how-
14 ever, improved predictions of decadal- to centennial-scale hydroclimatic variability from GCMs may de-
15 pend crucially on constraints from high-resolution paleoclimate observations (e.g. ??). Such data provide
16 much-needed statistics for climate variability and augment the relatively short instrumental record. Thus,
17 combining data from both models and high-resolution paleoclimate records yields meaningful advances for
18 understanding future climate.

19 Constraining climate models with paleoclimate data requires a robust method for comparing the two. Re-
20 cently, a number of studies have compared GCM simulations and paleoclimate data in the frequency domain,
21 applying spectral analysis to both the simulated and observed climate record. For temperature, precipitation,
22 or any other key indicator in a paleoclimate archive, comparing the power spectral densities (PSDs) across
23 models and data allows one to assess the dominant modes of variability in both signals (???). Recently, ??
24 showed that commonly employed proxies for Holocene sea surface temperature (SST) exhibit a spectrum
25 of SST variability inconsistent with GCM simulations on millennial timescales. Similarly, ? found that
26 last-millennium terrestrial records from western North American exhibit larger low-frequency variability
27 (and larger spectral slopes) when compared to the suite of CMIP5 Last-Millennium GCM simulations (??).
28 While the absolute variability simulated in climate models is different from the shape of the power spec-
29 trum (which measures variability as a function of timescale), the two are closely related (we evaluate both
30 via Supplementary Information, SI hereafter); the spectrum observed in these paleoclimate records implies
31 scaling behavior originating from the climate system, and high variability on longer timescales. Scaling
32 behavior can also imply longer climate-system memory of extreme events, such as megadrought (??). Thus,
33 the mismatch in the shape of the spectrum simulated by GCMs vs. that observed in paleoclimate data has
34 been invoked as deficiencies in the ability of GCMs to simulate climate with a level of realism required

35 for predicting decadal to centennial variability (??). Such findings harbor important implications about risk
36 prediction using climate models (e.g. future drought in the southwest U.S.(?)).

37 The direct comparison of climate model output with paleoclimate observations involves three main chal-
38 lenges (e.g. ?): (1) Internal variability in models is not directly comparable to paleoclimate data in time;
39 (2) biases in climate models limit their ability to correctly simulate extremes in hydroclimate; (3) proxy
40 archives naturally filter and distort the original climate signal, confounding direct comparisons of paleocli-
41 mate data to climate model variables. To address the first two of these issues, comparing PSDs removes
42 model biases while comparing time-scale dependent variances, and ignores phase relationships (which are
43 not expected not match because of natural climate variability, *inter alia*), allowing a more robust analysis of
44 the partitioning of variance across different timescales in models vs. data (?).

45 In this study, we take additional measures to address the third challenge, which relates to the filtering of
46 the initial climate signal by proxy systems. A conversion step is needed to translate between model output
47 and the proxy signal. Accomplishing a major part of this conversion, recent advances in climate modeling
48 have allowed for the explicit incorporation of stable water isotope tracers in both the atmosphere and the
49 ocean (see Table S7, SI). For water isotope-based proxy systems, stable water isotopes translate between
50 the dynamical climate model variables (e.g. temperature and precipitation) and the geochemical signal that
51 the proxy data encode (e.g. $\delta^{18}O$ of precipitation). Adding water isotope physics to GCMs provides crucial
52 insight, helping to determine the drivers of isotopic variations observed in proxy data and associated climate
53 patterns (?). Embedded water-isotope physics bring us closer to a direct comparison between models and
54 data, but do not account for physical processes by which proxy systems alter and subsequently record the
55 original climate signal. In an effort to avoid assumptions inherent to inverse approaches (e.g. inverse-method
56 or calibration-based reconstructions in paleoclimate), we turn to proxy system modeling (for a review, see
57 ??), and employ a new approach using both water isotope physics and proxy system models (PSMs) as
58 tools for simulating each individual process that alters the original climate signal (be it biological, physical,
59 or geochemical). Dynamical and isotope variables are *translated to proxy units* for a direct comparison
60 between GCM output and observations (a forward approach).

61 Our study builds upon the analysis of ? and ??? by employing this forward approach for data-model
62 comparison in the frequency domain. In general, there are two methods that allow for a meaningful com-

63 parison of proxy and model spectra. One is the inverse-method correction of the proxy spectra accounting
 64 for the distortion applied by the recording processes (e.g. ?, which essentially applies an inverted forward
 65 model of the proxy), and one is the forward modeling employed in this manuscript, which in many cases
 66 efforts increased flexibility. In this study, we use forward modeling to disentangle the multivariate influences
 67 on proxy data using state-of-the-art PSMs for ice cores, corals, tree-ring cellulose, speleothem calcite (?)
 68 and tree-ring width (?). Within this novel framework, we address the following questions: (1) are there
 69 proxy system processes that alter the spectrum of simulated (hydro)climatic variability, and are the impacts
 70 of these processes distinguishable from climate in spectral space? (2) accounting for these processes, do
 71 GCM+PSM-driven pseudoproxies exhibit spectral characteristics comparable to proxy observations?

72 Section 2 outlines our experimental design, and Section 3 gives results showing case studies for the
 73 piece-wise transformation of the climate signal down to proxy units. We extend this analysis to a global
 74 scale using the PAGES2k Phase 1 Network in Section 4. Finally, we discuss the limitations and caveats of
 75 our approach, and suggestions for future research, in Section 5.

76 2. Methods

77 2.1. GCM & PSM-Generated Pseudoproxies

78 To provide climate model estimates of hydroclimate variability over the last millennium, as well as
 79 climate fields for the PSM-generated network, we use the water isotope enabled GCM SPEEDY-IER (?) (see
 80 SI Section S8 for details). We forced a transient simulation of SPEEDY-IER with sea surface temperatures
 81 from the last millennium simulation (?) of the CCSM4 coupled model (?), spanning 850-2005 (1000-2005
 82 considered for this study). We generate synthetic proxy time series using ‘proxy system models’ (PSMs
 83 ??). PSMs convert the simulated climate (e.g. temperature, precipitation) into a proxy time series. A given
 84 PSM includes three sub-models, each of which mimics a separate modification of the original input signal
 85 as it would occur in nature: (1) a *sensor* model, which describes any physical, geochemical or biological
 86 processes altering the climate signal; (2) an *archive model*, which accounts for any processes that affect
 87 the emplacement of the signal in the proxy medium, and (3) an *observation* model, which accounts for
 88 dating uncertainties and analytical errors in the final measurement made on the paleoclimate data (?). The

sub-model framework of PSMs helps to quantify changes that occur at each stage of the climate signal's evolution through the proxy system.

Each proxy type employs its own unique PSM. We used VS-Lite (?) to generate tree ring width records for all of the tree proxy locations using temperature and precipitation fields from the isotope-enabled model SPEEDY-IER. We model ice core, coral, speleothem, and tree cellulose records using fields from CCSM4/SPEEDY-IER coupled with a synthesis of previously published models for water isotopes in high-resolution proxy data (PRYSM v.1.0, ?). We apply these models to the individual case study locations listed below in Section 3 and to the larger PAGES2k Phase 1 network (Section 4 (?). The complicated nature of proxy data (e.g. chronological uncertainties and impacts on phasing) precludes point-to-point comparisons of time series, and thus there is a strong case for comparing simulated proxy to the observations in the frequency domain.

2.2. Data-Model Comparison in the Frequency Domain

To assess both models and paleoclimate data in frequency space, a power law scaling parameter β , defined as $S(f) \propto f^{-\beta}$ (where f is frequency and $S(f)$ is the power spectral density (??)) characterizes the distribution of variance in a system over a given timescale, and gives some indication of the spectrum's shape. In general, a high, positive value of β implies a red spectrum, with more variance on longer timescales. A negative value of β implies a 'blue' spectrum, with more variance on shorter timescales, and a 'white' spectrum ($\beta \sim 0$) implies uniform variance distribution across all timescales (?). We use the spectral slope (β) as the key metric in this study for comparing climate model output to paleoclimate observations. We estimate power spectra of the climate model data and annually-resolved paleoclimate data using Thomson's multi-taper method (?) and estimate power spectra for unevenly-spaced paleoclimate data (speleothems only) using the Lomb-Scargle method (??). While both methods contain known biases (???) the two methods yield similar results for evenly spaced data (see SI Sec. S6, S7), and are necessary because the cave paleoclimate time series are not annually resolved and unevenly spaced in time. We apply the same methodology to all records uniformly, reducing the impacts of the methodological biases on our results. β is calculated as the least-squares regression of the log-log transformed spectral density and frequency. Following the methodology of ? and ?, we first bin spectral densities (at equal $0.2 \cdot \log_{10}$ intervals) to

116 avoid overemphasis of high-frequency variance in regression calculations. We calculate β according to
 117 timescale of variability: β_D is the decadal to centennial variability parameter (referenced in the text as *low-*
 118 *frequency variability*), and β_I is the interannual variability parameter (referenced in the text as *high-frequency*
 119 *variability*). We calculated β_D and β_I with slopes restricted to periods of [20 to 200] and [2.5 to 8] years,
 120 respectively. Characteristic spectral slopes for large-scale temperature fields and in GCM simulations have
 121 been quantified in a number of studies(?????), and are summarized in the SI (Section S4). Throughout the
 122 results that follow, we compare the piece-wise slopes of the PSDs across the two major (interannual, decadal
 123 to centennial) timescales.

124 3. Case Studies

125 We employ high-resolution records for five major proxy types, including corals, tree-ring width, ice core
 126 $\delta^{18}O$, tree ring cellulose $\delta^{18}O$, coral $\delta^{18}O$, and speleothem calcite $\delta^{18}O$ as test cases. We opted to employ
 127 annually to near-annually resolved records to maximize the potential for frequency range comparison with
 128 GCM data (i.e. frequencies from 1/2 years to 1/2*length of record) and to avoid the complicating effects of
 129 age offsets (?), which can blur the precision of the spectrum (e.g., see ?). Illustrative case studies spanning
 130 distinct climatic zones demonstrate the viability of our approach. We collected published data for each site
 131 and then compared the spectra of each time series to a PSM-generated record for the same location. Proxy
 132 data types, locations, time spans, mean resolution of data, and citations are given in Table 1.

133 A preliminary model experiment using white noise climate input and five case studies demonstrate how
 134 proxy system processes alter the input climate signal. For each proxy type we computed the GCM-derived
 135 spectrum for each record in a piece-wise fashion: power spectral density of GCM-simulated precipitation
 136 and temperature, GCM-simulated water isotope variables, and finally the full PSM output to compare with
 137 the observed spectrum. The ‘perfect model’ design tracking climate to proxy space includes assumptions
 138 at each step, and identifying these uncertainties explicitly is often challenging. Here we attempt to identify
 139 where discrepancies arise when comparing paleoclimate observations to climate model data.

140 In general, using a single proxy location to constrain climate models is erroneous because of disparities
 141 of scale. Various approaches including downscaling or bias correction can help to minimize such problems,
 142 or paleoclimate data can be aggregated to match GCM grid cell size. Here, we use single or multiple point-

Site	Lat	Lon	Observation	Citation	Dates (CE)	Mean Reso- lution (yrs)
Palmyra Island	5.89	-162.1	Coral $\delta^{18}O_{\text{ARAGONITE}}$	(??)	1146-1998*	1
Palestina Cave	-5.92	-77.35	Cave $\delta^{18}O_{\text{CALCITE}}$	(?)	421-1928	4
Cascayunga Cave	-6.07	-77.18	Cave $\delta^{18}O_{\text{CALCITE}}$	(?)	1088-1907	2
Huagapo Cave	-11.27	-75.79	Cave $\delta^{18}O_{\text{CALCITE}}$	(?)	559-2000	5
Diva de Maura Cave	-12.37	-41.57	Cave $\delta^{18}O_{\text{CALCITE}}$	(?)	-815-1911	5
Lhamcoka, Tibet	31.817	88.1	Tree $\delta^{18}O_{\text{CELLULOSE}}$	(?)	1193-1996	1
Boibar, Pakistan	36.37	74.59	Tree $\delta^{18}O_{\text{CELLULOSE}}$	(?)	1000-1998	1
Austfonna, Norway	79.8	24.02	Ice Core $\delta^{18}O$	(?)	1400-1953	1
Lomonosovfonna, Svalbard	78.87	17.4	Ice Core $\delta^{18}O$	(?)	1400-1997	1
NGRIP, Greenland	75.1	-42.32	Ice Core $\delta^{18}O$	(?)	0-1995	1
Quelccaya, Peru	-13.93	-70.83	Ice Core $\delta^{18}O$	(?)	226-2009	1
Malpais, New Mexico	34.97	-108.18	Tree Ring Width	(?)	-130-1992	1
Wild Rose, Colorado	39.02	-108.23	Tree Ring Width	(?)	1000-2002	1
Upper Wright Lakes, California	36.62	-118.37	Tree Ring Width	(?)	-216-1992	1
Beef Basin, Utah	37.93	-109.8	Tree Ring Width	(?)	350-2005	1

Table 1: List of sites forward modeled and evaluated in this study. By building a workflow connecting paleoclimate proxy data, a climate model, and proxy system models, we evaluate spectral scaling characteristics for five proxy types in four regional case studies. Corals, Ice Cores, Speleothems, Tree-Ring Width, Tree-Ring Cellulose. * Note that the Palmyra coral record is continuous in the interval 1146-1464: we use the previously published, continuous part of the record in this study. This segment has been updated since (?).

based observations alongside a single model grid-cell simulated pseudo proxy to identify proxy processes which influence the shape of the power spectrum, but we acknowledge that robust data-model comparisons require the use of proxy data aggregated from a broad region (e.g. Ocean2k ?). For each proxy type, we attempt to answer whether the mismatch arises from a lack of low-frequency variability simulated by the GCM SPEEDY-IER, or from a data-model comparison strategy problem. Our results are presented as a function of timescale: interannual (β_I) vs. decadal to centennial (β_D). For completeness, we report absolute variance for all case studies and the PAGES2k data in SI Section S3.

3.1. Spectral Fingerprinting of Proxy Systems

As a first pass, we forced each PSM with white noise climate inputs to assess the impact of proxy system processes alone on the shape of the spectra. In contrast to studies which impose the effects of autocorrelation mathematically (or smooth the data using a Gaussian filter, for example) (e.g. ?), for each proxy type, the PSMs resolve the spectrum that results from proxy system processes alone, as well as the resulting β values; we can then quantify the ‘reddening’ that occurs independently from climate due to

autocorrelation processes. Fig. 1 shows the median and 95% confidence intervals of the PSDs for 1000 simulations generated with randomly generated Gaussian white noise input climate (PSM parameters for the white noise exp. given in SI Table S1 and Section S2). As we will demonstrate in the case studies that follow, the effects of processes such as diffusion, karst residence time, and soil-moisture memory give rise to large positive β values over interannual timescales (Fig. 1). For ice cores, speleothems, and tree ring widths, the white noise + PSM simulations demonstrate significant autocorrelation with $\beta > 0.8$ on interannual timescales (2-8 years). Thus, *even without variable climate inputs, the proxies themselves impose a characteristic spectral shape at interannual timescales*. However, the cellulose oxygen isotope PSM does not impose any spectral scaling outside of the input, and returns a white spectrum. For all proxy types, the spectra revert to the shape of the white input climate signal on decadal and longer timescales. Under different PSM formulations these spectra could change significantly, and this non-unicity proves a large source of uncertainty. We find that while proxy system processes modeled here may account for significant reddening at the interannual frequency band, they do not appear to impart reddening at low frequencies. We acknowledge that this lack of low frequency reddening could be alternatively regarded as a deficiency of the PSMs at lower frequencies, or poorly represented (or unknown) proxy system processes; however, the lack of low-frequency amplification is potentially an expression of reality.

3.2. Proxy Results

3.2.1. Corals

The Palmyra fossil coral record resolves ENSO variability over the last millennium at annual resolution, and captures both local and large-scale tropical Pacific SST variability (??). Coral $\delta^{18}O$ also reflects the oxygen isotope composition of seawater, which tends to closely track local changes in salinity. Our first case study concerns parametric uncertainties in coral proxy systems and their influence on spectral shape. We employ a proxy system model for oxygen isotopes in coral aragonite to convert ocean model (CCSM4) output (SST, SSS) to $\delta^{18}O$ (?). Table 2 gives the β values for simulated vs. observed $\delta^{18}O_{\text{CORAL}}$. The coral PSM, as described in ?, is a simple bivariate linear model:

$$\delta^{18}O_{\text{pseudocoral}} = \alpha_1 \cdot SST + \alpha_2 \cdot SSS \quad (1)$$

Coral	β_I (Obs)	β_I (Sim)	Different?	p-value	β_D (Obs)	β_D (Sim)	Different?	p-value
Palmyra	1.95	1.11			1.26	1.25		
PAGES2k Corals	1.36	0.88	no	0.16	1.02	0.03	yes	0.0002
Ice Core	β_I (Obs)	β_I (Sim)	Different?	p-value	β_D (Obs)	β_D (Sim)	Different?	p-value
Austfonna	1.66	0.88			0.86	0.83		
Lomonosovfonna	0.95	1.17			0.26	0.74		
NGRIP	0.99	0.73			-0.48	-0.09		
Quelccaya	0.82	1.08			-0.36	0.16		
PAGES2kIce Cores	1.87	1.08	yes	0.01	0.16	0.18	no	0.76
Cellulose	β_I (Obs)	β_I (Sim)	Different?	p-value	β_D (Obs)	β_D (Sim)	Different?	p-value
Lhamcoka	0.27	0.42			0.91	-0.54		
Boibar	0.88	0.41			-0.02	-0.55		
Speleothem	β_I (Obs)	β_I (Sim)	Different?	p-value	β_D (Obs)	β_D (Sim)	Different?	p-value
Palestina Cave	n/a	n/a			0.67	0.56		
Cascayunga Cave	n/a	n/a			1.41	0.60		
Huagapo Cave	n/a	n/a			-0.19	0.31		
Diva de Maura Cave	n/a	n/a			1.96	0.08		
Tree Ring Width	β_I (Obs)	β_I (Sim)	Different?	p-value	β_D (Obs)	β_D (Sim)	Different?	p-value
CA_mod_negex	-0.01	0.43			0.30	0.16		
CA_VS-Lite		0.87				0.19		
CO_mod_negex	-0.01	0.28			0.79	0.09		
CO_VS-Lite		0.85				0.52		
NM_mod_negex	0.12	-0.44			-0.18	0.68		
NM_VS-Lite		1.01				0.42		
UT_mod_negex	0.60	0.65			0.08	0.56		
UT_VS-Lite		0.71				-0.54		
PAGES2k Tree Rings	0.59	-0.04	yes	<0.0001	0.49	0.13	yes	<0.0001

Table 2: β_I and β_D values for coral, ice core, speleothem, and tree ring cellulose sites, simulated vs. observed for all case studies and the PAGES2k Database. β_I is calculated as the mean slope between 2 and 8 years (interannual), and β_D is the mean slope between 20 and 200 years (decadal to centennial). Mean PAGES2k vs. PSM simulated overall β values for Coral, Ice Core, and Tree Ring PAGES2k sites with test statistics for difference-of-mean Mann-Whitney/Wilcoxon Rank-Sum tests are given alongside case studies. For TRW, β values were calculated for the four tree ring width sites listed in Table 1, simulated vs. observed, using the Modified Negative Exponential (negex) detrending method and simply calculating the slope after simulating the TRW using the forward model VS-Lite (no detrending). We report difference of means test statistics for the PAGES2k data only, as we have a full distribution of values for these data as opposed to the single-point case studies, where a difference of means test is not appropriate. See SI Table S6 for a comparison of six different detrending methods including RCS, Modified Negative Exponential (negex), Linear, Spline, Hegershoff.

174 where coefficient $\alpha_1 = -0.22\text{‰}/^\circ\text{C}$ is the relationship between oxygen isotopic equilibrium and formation
175 temperature of carbonates, and α_2 is the empirical estimate for the regional $\text{SSS-}\delta^{18}\text{O}_{\text{SW}}$ slope reported by

176 ? To investigate the impact of parametric uncertainties on the measured signal, we changed (α_2) to alter
 177 the degree of the coral $\delta^{18}O$ sensitivity to local salinity. With $\alpha_2 = 0$, the corals become simple linear
 178 responders to temperature. As a second experiment, measured α_2 slopes were increased by 200% to mimic
 179 the case where the corals exhibit heightened salinity sensitivity. The results are shown in Fig. 2.

180 Uncertainty in α_2 results in only minimal obfuscation of the original climate signal (SST). Regions with
 181 large variability in salinity will exhibit heightened sensitivity to uncertainty in the slope of the regional
 182 relationship between SSS and $\delta^{18}O_{SW}$, as the contribution of salinity anomalies to the total simulated coral
 183 signal is amplified (?). However, the narrow window due to salinity variations around the power spectrum
 184 shows that the corals are generally strong SST proxies (or, possibly, that the GCM completely underplays
 185 salinity variability). Testing the effects of parametric uncertainty for the corals provides an example of how
 186 PSMs can be used to inform data-model comparison. More interestingly, discrepancies exist between the
 187 simulated and observed power spectrum on decadal to centennial timescales.

188 *Interannual Scaling.* Both the simulated and observed coral time series show high variance on interannual
 189 timescales, with β_I (observed) = 1.95, β_I (simulated) = 1.11. The 2-8 year interannual band cutoff captures
 190 variance on ENSO timescales. However, observed β_I does significantly exceed modeled β_I values (see
 191 pink curve in Fig. 2), and may suggest either (1) undersensitivity of the coral PSM to variable SSTs, (2)
 192 biological or geochemical effects in the real corals that are not captured by the PSM, or (3) an erroneous
 193 ENSO representation in the GCM.

194 *Decadal to Centennial Scaling.* Despite agreement in β_I , at the 20-200 year band, Fig. 2 shows a pronounced
 195 difference between the observed (dotted black line) and simulated (pink line) coral $\delta^{18}O$ values. The differ-
 196 ence in the the simulated vs. observed β_D for the coral case study is 1.25 vs. 1.26, indicating that simulated
 197 values on decadal timescales are roughly equivalent. However, in examining the larger pool of PAGES2k
 198 coral data given in Table 2, the PSM seems to vastly underestimate decadal variability. Further, if we in-
 199 stead evaluate both in terms of absolute variance, the Palmyra record exhibits larger σ at the decadal band
 200 as compared to the PSM-simulated data (SI Section S3). While the PSM-generated pseudo-coral captures
 201 interannual SST variability similar to observations, the PSM seems not to account for the larger variance in
 202 the observations on longer timescales, and this discrepancy remains even when uncertainties in the coral's

203 sensitivity to salinity and $\delta^{18}O_{SW}$ are taken into account.

204 3.2.2. Ice Cores

205 We selected ice core data from four different sites representing a wide geographic range (see Tab. 1). The
206 Quelccaya ice cores record a combination of precipitation and temperature changes driven by tropical Pacific
207 SST variability (?). The Greenland (NGRIP) and Svalbard ice cores (Austfonna and Lomonosovfonna) also
208 record a combination of temperature and precipitation variability (??).

209 Observed and simulated β values for ice cores are given in Table 2, and Fig. 3 shows the power spectra
210 for four well-known ice core sites with annually-resolved data spanning the last millennium. The simulated
211 ice cores illustrate the dominance of temperature vs. precipitation variability on the $\delta^{18}O_{ICE}$ signal; $\delta^{18}O_P$
212 variance closely tracks the temperature spectrum at each site. Diffusion emerges as a dominant control on
213 the spectra for ice cores at high frequencies, but as with the coral data, simulated and observed ice core data
214 tend to diverge at lower frequencies.

215 *Interannual Scaling.* Fig. 3 illustrates the comparatively larger variance loss on interannual timescales in ice
216 cores due to diffusion and compaction processes. This occurs at all sites for the simulated ice cores as well
217 as observed. The positive β_I values (see Table 2) are a result of processes that give rise to autocorrelation
218 in the proxy data. Autocorrelation due to diffusion in the firn emerges at high frequencies (especially if
219 simulated snow accumulation rates at the ice core site are too low). The temperature and precipitation data,
220 both simulated and instrumental, do not agree with the proxy data on these timescales: resultant β_I values
221 are highly influenced by proxy processes.

222 Indeed, when we discount the effects of diffusion, there is more agreement in β between the ice core
223 records and the simulated temperature, precipitation, and water isotope ratios. For example, Fig. 4 shows
224 the simulated vs. observed spectra for NGRIP varying the diffusion length (σ) in the ice core PSM in
225 1000 random simulations from $\frac{1}{2}$ to $2 \cdot \sigma$, and we find that variance loss at high frequencies increases
226 with increasing diffusion length. Conversely, when we remove the diffusion and compaction model, the
227 simulated $\delta^{18}O_{PRECIP}$ (dark blue line in Fig. 4) shows a much flatter spectral slope; the simulated $\delta^{18}O_{ICE}$
228 shows agreement with the observed spectrum for NGRIP when the diffusion and compaction processes
229 are applied (purple line). These results suggest that the diffusion model component of the ice core PSM

230 correctly estimates the effects of down-core diffusion at high-frequencies, reddening the power spectrum at
231 interannual but not at decadal timescales.

232 *Decadal to Centennial Scaling.* On decadal to centennial timescales, differences in the observed vs. sim-
233 ulated spectral slopes are more modest than for interannual, but three of the records tend to increasingly
234 diverge at low frequencies (see Fig. 3). Examining Fig. 3, the spectral characteristics of the simulated
235 $\delta^{18}O_{\text{PRECIP}}$ vs. the observed ice core values exhibit some agreement on multi-decadal frequencies, but the
236 model does not simulate comparable variance in the observations on longer (>centennial) timescales (see
237 Fig. 3). This suggests that neither the GCM, the water isotope physics in the GCM, nor the PSM can account
238 for observed low frequency variability.

239 3.2.3. Speleothems

240 We selected four tropical South American Speleothem Records which are sub-decadally resolved with
241 a mean resolution 2-5 years (refs: Tab. 1). These records have been interpreted as an archive of variability
242 in local precipitation amount (i.e. Diva de Maura, ?), intensity of convection, and the overall strength of
243 the South American Monsoon (???). Hydroclimatic variability in this region closely follows both tropical
244 Atlantic SSTs as well as tropical Pacific SST variability (?). We simulate cave dripwater using a conceptual
245 model which takes temperature, precipitation amount, and the $\delta^{18}O$ of weighted precipitation into account.
246 The PSM also simulates the groundwater storage (i.e. karst transit time, τ) (??)

247 *Interannual Scaling.* Simulated β values for $\delta^{18}O$ of speleothem calcite differ substantially from observa-
248 tions (see values in Table 2). We elected to exclude the β_I calculation from our analysis of the observations
249 because the data were not annually resolved; after binning the spectral data, there were not enough points
250 to generate meaningful values for the interannual slope (see SI Section S7). Nevertheless, to investigate the
251 impacts of the proxy system on β_I , we experiment with a range of values for karst transit time to isolate the
252 effects of parametric uncertainties in our representation of the karst system for one site (Cascayunga). For
253 the initial simulation, all site PSM-generated data assume the karst transit time (τ) is one year. However,
254 spectral characteristics of the simulated cave dripwater signal are largely dependent on groundwater storage
255 time: β_I increases sharply with longer groundwater storage time (Fig. 5). The spectrum of simulated drip

256 water values for $\tau = 2$ years most closely resembles the observed time series over interannual to decadal
257 timescales.

258 The parameter space simulation shown in Fig. 5 allows to quantify how much of the change in β is due
259 to karst processes alone (and see Fig. 1b). As with the ice core data, autocorrelation exists in karst systems
260 on interannual timescales. Mixing in the karst and soil moisture includes the isotopic memory of water
261 that enters the system from previous years, generating steeper β values. Characterizing the true spectrum of
262 climate variability for these caves is thus complicated if karst parameters are poorly constrained or unknown.

263 *Decadal to Centennial Scaling.* At Cascayunga cave, a steeper spectral slope emerges in the paleoclimate
264 observations ($\beta_D = 1.41$) compared to the GCM ($\beta_D = 0.60$) at low frequencies. Referring back to Fig. 1b.,
265 the ‘PSM+white noise’ imposed β_D for speleothems is 0.12, which is smaller than the observed data-model
266 discrepancy in β_D for all cave sites (Table 2). For speleothem data from Cascayunga and Diva de Maura,
267 observed low frequency variability greatly exceeds that simulated by the climate model, proxy system effects
268 aside (see SI Figure S3).

269 The speleothem PSM highlights the fact that on interannual to decadal timescales, we can essentially
270 obtain a β value in agreement with observations simply as a function of the karst parameters. On longer
271 timescales, the simulated spectra tend to flatten while the observed spectra continue to show increased low-
272 frequency variance, potentially indicative of climate processes resulting in a spectrum similar to what we
273 would expect from a power law system (see Fig. 5).

274 3.2.4. Tree Ring Cellulose

275 We employ two published records of Asian Tree-Ring Cellulose $\delta^{18}O$ (Table 1), both which demonstrate
276 the tree ring cellulose isotope ratios are sensitive to regional precipitation and humidity. Our proxy system
277 model for oxygen isotopes in wood cellulose converts water-isotope enabled model output to tree ring cel-
278 lulose $\delta^{18}O$ (?). Table 2 gives the β values for simulated vs. observed $\delta^{18}O_{\text{CELLULOSE}}$ and power spectra for
279 both simulated and observed oxygen isotopes in tree cellulose are shown in Fig. 6.

280 *Interannual Scaling.* On interannual time scales, the simulated spectral slopes vary compared to those of
281 observed Asian tree ring series. The discrepancy is particularly apparent for the Boibar site (panel 6a.),

282 where $\beta_I = 0.88$ (vs. 0.41 simulated). We hypothesized that autocorrelation due to soil water storage and
283 isotopic mixing prior to use of the water by the tree would lead to a steepening of the spectra of the GCM +
284 PSM simulated cellulose records. However, the β_I values for simulated oxygen isotopes in tree cellulose are
285 variable compared to observations.

286 In this case, it is possible that the proxy system processes that redden the signal on interannual timescales
287 are poorly represented, or not represented at all by the GCM or PSM. Potential confounding factors include
288 improper simulation of soil moisture storage and isotope ratios by the GCM, or a failure of the PSM to
289 capture secondary tree growth effects, for example.

290 *Decadal to Centennial Scaling.* The GCM+PSM simulated cellulose time series do not capture the steeper β
291 values observed in the measured data at Lhamcoka, and in fact are negative; none of the processes described
292 in the PSM contribute to scaling behavior in the simulated signal. As discussed, this could reflect gaps in
293 our understanding of the complexity of how tree cellulose oxygen isotopes operate, or indicative of a lack of
294 variability simulated by the GCM. The discrepancies worsen on longer timescales (Fig. 6), and the absolute
295 variances for Lhamcoka and Boibar are 0.46 and 0.32 (observed), respectively, as compared to 0.02 and
296 0.009 (simulated) (see SI Section S3).

297 3.2.5. Tree Ring Width

298 Previous work by ? suggested that biological proxies such as trees tend to add in autocorrelation, which
299 makes proxy records ‘redder’ than the background climate they are recording. Hydroclimate tree-ring prox-
300 ies often reflect soil moisture rather than temperature or precipitation, which exhibits higher PSD at decadal
301 frequencies. Since we rely on these records for reconstructions, these reconstructions may tend toward more
302 low-frequency variability compared to the input climate signal. Proxy system modeling addresses this spec-
303 tral bias: we model the growth response of trees with climate for Western North America using the VS-Lite
304 forward model (?). The four sites we consider (Tab. 1 for references) record over 1000 years of climate
305 and is sensitive to a combination of precipitation and temperature. However, tree growth is influenced by a
306 complex relationship with age which must be removed by detrending the data before a chronology can be
307 used as a climate proxy. Invariably, detrending either partially removes the low-frequency climate signal
308 from the record, or over-emphasizes some of the age-growth relationship masquerading as climate (?). In

309 orsder to fully mimic real tree-ring chronologies, we use VS-Lite to model each individual tree at the site
 310 using simulated temperature and precipitation then add in the age-growth curves from the actual trees cal-
 311 culated using the regional curve standardization method (RCS) (??). We also incorporate a small error term
 312 to mimic the natural variability within a forest, such that 60% of each trees variance is controlled by climate
 313 and 40% is noise. As each tree is modeled individually, most of the noise is removed when we combine
 314 multiple trees into a single chronology. Finally, we pre-whiten and build a chronology with the pseudoproxy
 315 trees using ARSTAN (?) and six different detrending methods (RCS, negative exponential, modified neg-
 316 ative exponential, linear, spline, and Hugesshoff) using the same methodology employed when creating a
 317 real-world chronology.

318 To illustrate the effects of detrending methodology on retrieved climate spectrum, Fig. 7 shows the
 319 temperature, precipitation, modified negative-exponential detrended proxy vs. pseudoproxy chronology, and
 320 the associated β values are reported in Table 2, (bottom panel, and see SI Tab. S6). In addition, for each site,
 321 we plot the VS-Lite-only PSM simulation without added age-growth curves or detrending to demonstrate
 322 the importance of the age-growth relationship and detrending method on the power spectrum. Finally, to
 323 estimate reasonable errors for our VS-Lite generated pseudoproxy data, we performed 100 Monte-Carlo
 324 simulations resampling within a range of reported growth parameter errors.

325 *Interannual Scaling.* The substantial difference in β for the TRW PSM (see Fig. 1c., $\beta_I = 1.53$, $\beta_D = 0.06$)
 326 stems from the seasonal growth parameterization of VS-Lite: autocorrelation arises and reddens β_I because
 327 the forward model includes part of the previous year's growth in the current season's growth. However,
 328 in the final spectrum given for the TRW time series in Fig. 7, the large positive β_I values are not readily
 329 apparent. The detrending applied to each of the TRW chronologies tends toward a blue spectrum, removing
 330 variance on both interannual and decadal timescales.

331 *Decadal to Centennial Scaling.* For each site, the spectrum of observed tree ring width proxy agrees well (in
 332 terms of similarity of power spectra) with the pseudoproxy detrended with the same method. To illustrate
 333 this fully, Figures S7 and S8 in the SI show the full spectra for five frequently used detrending options
 334 used on each of our four sites. The power spectra for each of the detrending methods diverge at periods
 335 greater than 200 years. We choose to cut off the calculations for β at the 200 year period, so the divergence

of detrending methods is relatively modest (Table 2). We find that when comparing climate models to TRW data, the detrending method has a large impact on the low-frequency spectral characteristics of both: aggressive detrending methods tend to remove low frequency variability (demonstrated by Table 2). Table 2 also illustrates the RCS method is most conservative in maintaining low-frequency TRW variability. In general, using the same detrending method for both proxy and pseudoproxy is essential.

4. Spectral characteristics of the PAGES2k Network

We extend our analysis to a larger number and a wider geographic range of sites using the PAGES2k Phase 1 Network (?). The PAGES2k data serves as an expanded test of our results in Section 3.2, and allows us to assess our interpretations of how proxy systems affect the simulated spectra on a broad geographic scale. Both sets of results are summarized in Table 2, and our experimental treatment of the PAGES2k data is described in detail in SI Section S9.

First, we compare GCM+PSM simulated proxy data (SPEEDY-IER, CCSM4) data to observed proxy data from the PAGES2k network. Our analysis includes ice core $\delta^{18}O$ [$N_{OBS}=22$, $N_{PSM}=22$], coral $\delta^{18}O$ [$N_{OBS}=10$, $N_{PSM}=10$], and tree ring width [$N_{OBS}=407$, $N_{PSM}=116$]. Differences in site numbers reflect the fact that multiple proxy sites are often collected in a single model grid cell. Raw GCM β value distributions are compared to the GCM-PSM output in Fig. 8 to examine the impact of translating GCM data to proxy units (i.e. what is the contribution of the PSM alone?) The figure shows climate fields (e.g. temperature, precipitation, SST, SSS) from the GCM plotted alongside PSM-simulated proxy data. We find that the ice core PSM β value distribution is significantly higher than both the precipitation and temperature β distributions due to diffusion in the PSM. The tree ring PSM (VS-lite) β distribution is slightly lower than that of temperature, and quite similar to precipitation. Finally, the coral PSM β value distribution appears to be a combination of SSS and SST (Fig. 8c). On decadal-centennial scales, PSM β_D value distributions appear to incorporate elements of their associated climate input variables (especially for ice cores, which overlap with temperature and precipitation distributions). Although interannual tree ring β_I value distributions overlap with both temperature and precipitation, coral β_I value distributions are lower than SSS and higher than SST (SI Fig. S5 and S6).

There is limited agreement in the distribution of β values in observed vs. PSM-generated ice core and

tree ring data, illustrated in Fig. 8. However, according to a Wilcoxon/Mann-Whitney (Rank-Sum) test, the observed and simulated β values are significantly different for all three types of paleoclimate archives, depending on timescale (Table 2). Simulated ice core β_I values are too large, whereas the simulated tree ring width β_I values are too small compared to the PAGES2k data (Figure 8a, b). As discussed in Section 3.2.2, differences between the simulated and observed differences in β are timescale-dependent. The PAGES2k and PSM β_D values are not significantly different for ice cores, and β_I for both simulated and observed corals are similar on interannual timescales. But, the mismatch remains in coral and tree ring width data at decadal scales, and in ice core and tree ring data at interannual scales (SI Fig. S5 and S6).

Complementary to this analysis, we calculated the absolute variance in the modeled vs. observed data for the PAGES2k sites (SI Figure S2) and find that for all proxy types, the range of absolute variance in the PAGES2k observations exceeds the range of variances simulated by the GCM+PSM pseudo proxies at decadal timescales. The enhanced low-frequency variability in the PAGES2k corals and TRW at decadal timescales suggests agreement with the results of our case studies (see Table 2): PSMs help explain differences in observed vs. simulated variance at interannual timescales, but as we increase the timescale of variability to decadal and centennial periods, high-resolution archives like tree-rings and corals tend toward larger variance than the GCM simulates, even with PSMs (Fig. 8e, f).

5. Discussion

We reevaluated observed disagreement between archives of past climate variability and a water isotope-enabled GCM simulation by including conceptual forward models describing proxy system processes. In doing so, we provided a proof of concept demonstrating the usefulness of proxy system modeling in data-model comparison: without a bridge translating between GCMs and proxy data, one would be uncertain of the true differences between simulated and observed decadal to centennial scale climate variability. With consideration of the complex ways in which proxy systems may alter the input climate signal, PSMs allow us to quantify the effects of these processes explicitly. While many previous studies have evaluated models and data in the frequency domain, we extended this analysis by incorporating PSMs prior to calculating the spectra of GCM data. We find that translating the GCM simulation to proxy units matters, as demonstrated through several case studies. Physical processes and measurement biases exist in proxy systems that act to

obfuscate the original climate signal and alter the spectrum of variability on interannual timescales. However, these confounding factors are not directly related to the continuum of climate variability, and thus the source of the discrepancy between climate model and proxy data may sometimes have origins other than a shortcoming of the GCM simulations.

To address this, PSMs allow us to explicitly resolve and capture processes that generate emergent autocorrelation at high frequencies for each proxy type (rather than accounting for it after the fact). We identified processes consistent with an auto-regressive (AR1) behavior (mixing in the karst, diffusion, and seasonal growth memory) that have a large impact on β_I of the simulated proxy data, but which do not reflect a true climate signal. As demonstrated in Section 3.1, the magnitude of the change in β_I due to autocorrelation-generating processes alone can be quantified explicitly using this framework. In Section 3.2, we showed that constraining proxy system parameters may prove essential for a robust understanding of model/data discrepancies in the frequency domain. One might conclude in error, for example, that the GCMs are ‘wrong’ if they cannot replicate low frequency climate variability observed in archives for precipitation amount (e.g. speleothems) when, in fact, the karst system mixing controls decadal-centennial scaling behavior in caves. Caution is needed when comparing models and data when such confounding proxy system effects are poorly constrained; the GCM+PSM framework narrows the gap between raw GCM data and paleoclimate data, affording heightened awareness of confounding proxy system processes.

Importantly, we find that modeling proxy system processes helps resolve model-data discrepancies on interannual to decadal timescales, but does not account for the mismatch in variance on longer (multi-decadal to centennial) timescales. The paleoclimate archives contain more variance on longer timescales, independent of known proxy processes, than Earth system models currently simulate for surface temperature and precipitation. In agreement with studies such as ? and ??), our analysis suggests that for many proxy types, our GCM simulation falls short of replicating the spectral characteristics and absolute variance (SI Sec. S3, Figure S2) of the paleoclimate archives at decadal to centennial timescales, even with the additional bridge provided by the water isotope physics and PSMs. However, the simplified nature of both the climate and proxy system models limits what can be learned from them. Some important long-term climate processes may lacking from this last millennium simulation, performed with an intermediate complexity AGCM. Additionally, proxy archives may harbor additional reddening or scaling processes not included in our simplified

418 PSMs. Nonetheless, the broader PAGES2k-PSM comparison conducted in Section 4 generally suggest that
419 greater care is needed in contrasting raw GCM data directly with paleoclimate data, due to the reddening
420 effect of many proxies on the input climate signal.

421 We acknowledge a number of limitations of this work. In practice, it is difficult to diagnose how much
422 of the ‘mismatch’ between models and data is due to uncertainties in the paleoclimate data itself, PSM error,
423 GCM error, or a combination of all three. While our framework attempts to aid in these diagnostics, confir-
424 mation of error sources requires further attention. Additionally, we were restricted to a single water-isotope
425 enabled model with a long transient simulation, and as a result our ability to test the reconciliatory power
426 of PSMs with multiple models is limited. At present, a repository of publicly archived water-isotope simu-
427 lations spanning the last millennium is unavailable. In future work, we hope to strengthen our conclusions
428 using ensemble runs with multiple isotope-enabled GCMs, or by repeating our analysis using a suite of dif-
429 ferent spectral methods (??). Finally, variability observed in paleoclimate data is not fully characterized by
430 its spectral properties alone. Novel methods for times series analysis have been explored in recent decades
431 (e.g., techniques in nonlinear dynamics), and future work using these methods may help improve our com-
432 parative data-model analyses and lend further insight toward the underlying causes of the discrepancies.

433 Translating model variables to proxy space using our best understanding of the proxy system physics
434 and chemistry is useful for robust proxy-GCM data comparison. However, this conceptual bridge is far from
435 complete. While the intermediate complexity PSMs used in this study help close the gap between models
436 and data, much work remains to reduce biases in GCMs and to improve the physical representation of proxy
437 systems by PSMs, which are, at present, extremely rudimentary. This analysis includes just a few PSMs,
438 and would be enhanced through the development of additional PSMs. Further, it is unclear whether more
439 advanced PSMs would generate scaling behavior as a result of processes within the archive alone; while
440 our results suggest the opposite, this is potentially conditional on the type and complexity of PSM. With
441 forthcoming advances in PSMs, much insight stands to be gained in the realm of model validation.

442 In closing, we note that the mismatch between climate variability in GCMs and proxy archives harbors
443 implications for the predictability of extreme climate events: if GCMs fail to simulate scaling behavior in
444 the climate system, we may be underestimating statistics surrounding future temperature and precipitation
445 changes, as well as the weather phenomena that accompany these changes (e.g. ?). In particular, we leave

446 ourselves vulnerable to impacts of unpredicted and unexpected low-frequency climate variability. Moreover,
447 recent studies have highlighted the dependence of climate change rate estimates on measurement timespan,
448 and suggest that the real pace of abrupt climate variability in the past may be dramatically underestimated
449 (?). Given the growing importance of decadal climate prediction, it is important that both the modeling
450 and paleoclimate communities work together to resolve a *best estimate* of both absolute low-frequency
451 climate variability σ_D and β_D using robust methodology. We hope that the work presented in this paper lays
452 the groundwork for more advanced data-model comparison strategies and will lend further insight to this
453 important open question.

454 **Acknowledgements**

455 This work was funded by Brown Universitys Institute for Environment and Society via the Voss Postdoc-
456 toral Fellowship, as well as NOAA Climate Change Data and Detection grant NA10OAR4310115 and the
457 National Science Foundation EaSM2 Grant (AGS1243125). The authors would like to thank James Russell
458 for helpful thoughts and feedback on this project. In addition, the authors thank Dr. Thomas Laepple for his
459 help in reviewing this manuscript and helping to improve the models and methodology, and Thomas Münch
460 for his help improving the ice core model. The authors additionally thank three anonymous reviewers and
461 editor Dr. Heather Stoll for their helpful insights which improved this manuscript.

462 **References**

Figure 1: **Spectra and β values by proxy type and as a function of timescale.** Estimated Power Spectral Density for a purely white noise input climate signal, leaving the effects of the proxy system only. Each PSM was forced with white noise climate input 10,000 times. The median spectrum is shown as the solid line, and shading represents the 95% confidence interval. a. Ice Core $\delta^{18}O$, b. Carbonates (Coral and Speleothem $\delta^{18}O$) c. TRW, Tree Ring Cellulose $\delta^{18}O$.

Figure 2: **Coral Aragonite Model-Data Comparison:** simulated (pink) and measured (black) spectrum for Palmyra coral data, as well as the effects of altering α_2 at Palmyra island alongside the climate inputs to the model (SST, SSS). Straight orange lines are the calculated mean slopes across both decadal to centennial (20-200, β_D) and interannual (2-8, β_I) timescales. The black-dotted envelope represents the full range of outcomes given the perturbed salinity parameter space. Low frequency variance observed in the coral data (black dashed line) greatly exceeds that of the simulated pseudo coral data (pink), and the discrepancy grows with increasing timescale. Also shown: simulated sensitivity to salinity changes and potential parameter uncertainties in the PSM (α_2), effects on β , (faint black dots), which does not help account for the model-data discrepancies. We removed the mean from all fields prior to computed the PSD.

Figure 3: **Ice Core Model-Data Comparison.** Estimated Power Spectral Density for Simulated vs. Observed $\delta^{18}O$ of ice at four sites: (a) NGRIP, (b) Quelccaya, (c) Austfonna, (d) Lomonosovfonna. The PSD for each PSM-generated record is shown (red) alongside climate inputs of temperature (dark blue), precipitation (light blue), and $\delta^{18}O_P$ (purple). In all cases, the observed ice core record is plotted in black (dashed). Straight orange lines are the calculated mean slopes across both decadal to centennial (20-200, β_D) and interannual (2-8, β_I) timescales. Diffusion in the ice core PSM causes higher-than-observed spectral slopes at high frequencies, and may suggest that the simulation of diffusion and compaction is overestimated. PSM-simulated ice cores do not capture the observed higher variance at low-frequencies. We removed the mean from all fields prior to computed the PSD.

./F4_NGRIP_Diffusion_3_13.pdf

Figure 4: **Ice Core Model-Data Comparison at NGRIP: effect of diffusion length (σ) on β .** Estimated power spectral density for simulated (purple) vs. observed (black) $\delta^{18}O$ of ice and the effect of varying diffusion length parameters. The grey shaded region spans experiments resampling the data 1000 times varying the diffusion length from $\frac{1}{2} \cdot \sigma$ to $2 \cdot \sigma$. In the case where $\sigma = 0$ (the royal blue line, $\delta^{18}O_{\text{PRECIP}}$), GCM-simulated data is generally in agreement with the observations.

Figure 5: **Speleothem Calcite Model-Data Comparison: the effect of Karst Transit Time (τ) on β .** Simulated vs. Observed $\delta^{18}O$ of speleothem calcite at Cascayunga Cave. PSD is plotted, simulated using an ensemble of plausible values (6 months to 5 years) for the karst transit time parameter, τ . The spectrum reddens as the transit time increases. The orange solid line indicates the mean β slope for the observations. Figure illustrates broadening disagreement between simulated and observed speleothem $\delta^{18}O$ approaching decadal-centennial timescales.

Figure 6: **Tree-Ring Cellulose Model-Data Comparison.** Estimated Power Spectral Density for Simulated (green) vs. Observed (black) $\delta^{18}O$ of tree ring cellulose at (a) Boibar, Pakistan and (b) Lhamcoka, Tibet. We removed the mean from all fields prior to computed the PSD.

Figure 7: **Tree-Ring Width Model-Data Comparison:** the effect of detrending method on β . For each site, the original measured TRW is plotted in black, along with the PSM (VS-Lite) inputs of temperature (red), precipitation (blue), and the simple VS-Lite generated pseudoproxy data (dark green dashed), which does not include detrending or growth-curve. The negative-exponential detrending method was applied to both the real proxy data and the VS-Lite generated pseudoproxy data (bright green). To estimate reasonable errors for simulated TRW, we took reported errors for the estimated growth parameters in 100 randomized possibilities (grey).

Figure 8: **Distribution of proxy β_I, β_D values, GCM+PSM vs. Observed, for the PAGES2k Phase 1 dataset:** a. ice cores (blue, left), tree-ring width (green, center), and coral data (orange, right) from PAGES2k (lighter colors) and GCM/PSM (darker colors). (a, b., d., e.): temperature (red), precipitation (dark blue) and PSM ice cores (light blue, right); SPEEDY-IER temperature (red), precipitation (dark blue), and PSM tree-ring width (green, center); and (c., f.): CCSM4 sea-surface temperature (pink), sea-surface salinity (purple), and PSM coral data (orange, right).