Accepted Manuscript

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

S.G. Dee, L. Parsons, G. Loope, J.T. Overpeck, T.R. Ault, J. Emile-Geay

PII: S0012-821X(17)30420-X
DOI: http://dx.doi.org/10.1016/j.epsl.2017.07.036
Reference: EPSL 14578

To appear in: *Earth and Planetary Science Letters*

Received date: 27 September 2016
Revised date: 16 June 2017
Accepted date: 23 July 2017

Please cite this article in press as: Dee, S.G., et al. Improved spectral comparisons of paleoclimate models and observations via proxy system modeling: implications for multi-decadal variability. *Earth Planet. Sci. Lett.* (2017), http://dx.doi.org/10.1016/j.epsl.2017.07.036

Highlights

- We compare climate variability observed in paleoclimate data to GCM simulations.
- Proxy system modeling is used to enhance this data-model comparison in the frequency domain.
- Paleoclimate records exhibit larger low-frequency variability than GCMs currently simulate.

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 (Laepple and Huybers, 2014a,b). 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 (PAGES2K Consortium, 2013). 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

---

*Corresponding author
Email addresses: sylvia_11_dee@brown.edu (S.G. Dee), lukeaparsons@email.arizona.edu (L. Parsons), (G. Loope), jto@email.arizona.edu (J. T. Overpeck), toby.ault@cornell.edu (T.R. Ault), julieneg@usc.edu (J. Emile-Geay)

GCMs currently simulate, indicative of incomplete physics and/or forcings.

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

1. Introduction

Our understanding of the complex dynamics of climate response to anthropogenic forcing rests jointly upon observations over the instrumental period, general circulation models (GCMs), and paleoclimate data. GCMs provide a basis for exploring the mechanisms driving forced and stochastic climate variability; however, improved predictions of decadal- to centennial-scale hydroclimatic variability from GCMs may depend crucially on constraints from high-resolution paleoclimate observations (e.g. Mann et al., 2009; PAGES2K Consortium, 2013). Such data provide much-needed statistics for climate variability and augment the relatively short instrumental record. Thus, combining data from both models and high-resolution paleoclimate records yields meaningful advances for understanding future climate.

Constraining climate models with paleoclimate data requires a robust method for comparing the two. Recently, a number of studies have compared GCM simulations and paleoclimate data in the frequency domain, applying spectral analysis to both the simulated and observed climate record. For temperature, precipitation, or any other key indicator in a paleoclimate archive, comparing the power spectral densities (PSDs) across models and data allows one to assess the dominant modes of variability in both signals (Kutzbach, 1976; Hays et al., 1976; Huybers and Curry, 2006). Recently, Laepple and Huybers (2014a,b) showed that commonly employed proxies for Holocene sea surface temperature (SST) exhibit a spectrum of SST variability inconsistent with GCM simulations on millennial timescales. Similarly, Ault et al. (2013) found that last-millennium terrestrial records from western North American exhibit larger low-frequency variability (and larger spectral slopes) when compared to the suite of CMIP5 Last-Millennium GCM simulations (Taylor et al., 2012; Landrum et al., 2013). While the absolute variability simulated in climate models is different from the shape of the power spectrum (which measures variability as a function of timescale), the two are closely related (we evaluate both via Supplementary Information, SI hereafter); the spectrum observed in these paleoclimate records implies scaling behavior originating from the climate system, and high variability on longer timescales. Scaling behavior can also imply longer climate-system memory of extreme events, such as megadrought (Ault et al., 2013, 2014). Thus, the mismatch in the shape of the spectrum simulated by GCMs vs. that observed in paleoclimate data has been invoked as deficiencies in the ability of GCMs to simulate climate with a level of realism required for predicting decadal to centennial variability (Laepple and Huybers, 2014a,b). Such findings harbor important implications about risk prediction using climate models (e.g. future drought in the southwest U.S.(Ault et al., 2014)).

The direct comparison of climate model output with paleoclimate observations involves three main challenges (e.g. Ault et al., 2013): (1) Internal variability in models is not directly comparable to paleoclimate data in time; (2) biases in climate models limit their ability to correctly simulate extremes in hydroclimate; (3) proxy archives naturally filter and distort the original climate signal, confounding direct comparisons of paleoclimate data to climate model variables. To address the first two of these issues, comparing PSDs removes model biases while comparing time-scale dependent variances, and ignores phase relationships (which are not expected not match because of natural climate variability, *inter alia*), allowing a more robust analysis of the partitioning of variance across different timescales in models vs. data (Ault et al., 2013).

In this study, we take additional measures to address the third challenge, which relates to the filtering of the initial climate signal by proxy systems. A conversion step is needed to translate between model output and the proxy signal. Accomplishing a major part of this conversion, recent advances in climate modeling have allowed for the explicit incorporation of stable water isotope tracers in both the atmosphere and the ocean (see Table S7, SI). For water isotope-based proxy systems, stable water isotopes translate between the dynamical climate model variables (e.g. temperature and precipitation) and the geochemical signal that the proxy data encode (e.g. $\delta^{18}O$ of precipitation). Adding water isotope physics to GCMs provides crucial insight, helping to determine the drivers of isotopic variations observed in proxy data and associated climate patterns (Sturm et al., 2010). Embedded water-isotope physics bring us closer to a direct comparison between models and data, but do not account for physical processes by which proxy systems alter and subsequently record the original climate signal. In an effort to avoid assumptions inherent to inverse approaches (e.g. inverse-method or calibration-based reconstructions in paleoclimate), we turn to proxy system modeling (for a review, see Evans et al., 2013; Dee et al., 2015a), and employ a new approach using both water isotope physics and proxy system models (PSMs) as tools for simulating each individual process that alters the original climate signal (be it biological, physical, or geochemical). Dynamical and isotope variables are *translated to proxy units* for a direct comparison between GCM output and observations (a forward approach).

Our study builds upon the analysis of Ault et al. (2013) and Laepple and Huybers (2013, 2014a,b) by employing this forward approach for data-model comparison in the frequency domain. In general, there are two methods that allow for a meaningful comparison of proxy and model spectra. One is the inverse-method correction of the proxy spectra accounting for the distortion applied by the recording processes (e.g. Laepple and Huybers, 2013, which essentially applies an inverted forward model of the proxy), and one is the forward modeling employed in this manuscript, which in many cases efforts increased flexibility. In this study, we use forward modeling to disentangle the multivariate influences on proxy data using state-of-the-art PSMs for ice cores, corals, tree-ring cellulose, speleothem calcite (Dee et al., 2015a) and tree-ring width (Tolwinski-Ward et al., 2010). Within this novel framework, we address the following questions: (1) are there proxy system processes that alter the spectrum of simulated (hydro)climatic variability, and are the impacts of these processes distinguishable from climate in spectral space? (2) accounting for these processes, do GCM+PSM-driven pseudoproxies exhibit spectral characteristics comparable to proxy observations?

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

2. Methods

2.1. GCM & PSM-Generated Pseudoproxies

To provide climate model estimates of hydroclimate variability over the last millennium, as well as climate fields for the PSM-generated network, we use the water isotope enabled GCM `SPEEDY-IER` (Dee et al., 2015b) (see SI Section S8 for details). We forced a transient simulation of `SPEEDY-IER` with sea surface temperatures from the last millennium simulation (Landrum et al., 2013) of the CCSM4 coupled model (Gent et al., 2011), spanning 850-2005 (1000-2005 considered for this study). We generate synthetic proxy time series using ‘proxy system models’ (PSMs Evans et al., 2013; Dee et al., 2015a). PSMs convert the simulated climate (e.g. temperature, precipitation) into a proxy time series. A given PSM includes three sub-models, each of which mimics a separate modification of the original input signal as it would occur in nature: (1) a *sensor* model, which describes any physical, geochemical or biological processes altering the climate signal; (2) an *archive model*, which accounts for any processes that affect the emplacement of the signal in the proxy medium, and (3) an *observation* model, which accounts for dating uncertainties and analytical errors in the final measurement made on the paleoclimate data (Dee et al., 2015a). 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 (Tolwinski-Ward et al., 2010) 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,` Dee et al., 2015a). 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 (PAGES2K Consortium, 2013). 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 (Pelletier and Turcotte, 1997; Huybers and Curry, 2006)) 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 (Ault et al., 2013). 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 (Thomson, 1982) and estimate power spectra for unevenly-spaced paleoclimate data (speleothems only) using the Lomb-Scargle method (Lomb, 1975; Scargle, 1982). While both methods contain known biases (Schulz and

Mudelsee, 2002; Vyushin and Kushner, 2009; Rehfeld et al., 2011) 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 Huybers and Curry (2006) and Ault et al. (2013), we first bin spectral densities (at equal $0.2 \cdot \log 10$ intervals) to avoid overemphasis of high-frequency variance in regression calculations. We calculate β according to timescale of variability: β_D is the decadal to centennial variability parameter (referenced in the text as *low-frequency variability*), and β_I is the interannual variability parameter (referenced in the text as *high-frequency variability*). We calculated β_D and β_I with slopes restricted to periods of [20 to 200] and [2.5 to 8] years, respectively. Characteristic spectral slopes for large-scale temperature fields and in GCM simulations have been quantified in a number of studies(Fraedrich and Blender, 2003; Huybers and Curry, 2006; Vyushin and Kushner, 2009; Vyushin et al., 2009; Henriksson et al., 2015; Fredriksen and Rypdal, 2015), and are summarized in the SI (Section S4). Throughout the results that follow, we compare the piece-wise slopes of the PSDs across the two major (interannual, decadal to centennial) timescales.

3. Case Studies

We employ high-resolution records for five major proxy types, including corals, tree-ring width, ice core $\delta^{18}O$, tree ring cellulose $\delta^{18}O$, coral $\delta^{18}O$, and speleothem calcite $\delta^{18}O$ as test cases. We opted to employ annually to near-annually resolved records to maximize the potential for frequency range comparison with GCM data (i.e. frequencies from 1/2 years to 1/2*length of record) and to avoid the complicating effects of age offsets (Comboul et al., 2014), which can 'blur' the precision of the spectrum (e.g., see Dee et al., 2015a). Illustrative case studies spanning distinct climatic zones demonstrate the viability of our approach. We collected published data for each site and then compared the spectra of each time series to a PSM-generated record for the same location. Proxy data types, locations, time spans, mean resolution of data, and citations are given in Table 1.

A preliminary model experiment using white noise climate input and five case studies demonstrate how proxy system processes alter the input climate signal. For each proxy type we computed the GCM-derived

Site	Lat	Lon	Observation	Citation	Dates (CE)	Mean Resolution (yrs)
Palmyra Island	5.89	-162.1	Coral $\delta^{18}O_{\mathrm{ARAGONITE}}$	(Cobb et al., 2003; Emile-Geay et al., 2013a)	1146-1998*	1
Palestina Cave	-5.92	-77.35	Cave $\delta^{18}O_{\mathrm{CALCITE}}$	(Apaéstegui et al., 2014)	421-1928	4
Cascayunga Cave	-6.07	-77.18	Cave $\delta^{18}O_{\mathrm{CALCITE}}$	(Reuter et al., 2009)	1088-1907	2
Huagapo Cave	-11.27	-75.79	Cave $\delta^{18}O_{\mathrm{CALCITE}}$	(Kanner et al., 2013)	559-2000	5
Diva de Maura Cave	-12.37	-41.57	Cave $\delta^{18}O_{\mathrm{CALCITE}}$	(Novello et al., 2012)	-815-1911	5
Lhamcoka, Tibet	31.817	88.1	Tree $\delta^{18}O_{\mathrm{CELLUOSE}}$	(Wernicke et al., 2015)	1193-1996	1
Boibar, Pakistan	36.37	74.59	Tree $\delta^{18}O_{\mathrm{CELLUOSE}}$	(Treydte et al., 2009)	1000-1998	1
Austfonna, Norway	79.8	24.02	Ice Core $\delta^{18}O$	(Isaksson et al., 2005)	1400-1953	1
Lomonosovfonna, Svalbard	78.87	17.4	Ice Core $\delta^{18}O$	(Isaksson et al., 2005)	1400-1997	1
NGRIP, Greenland	75.1	-42.32	Ice Core $\delta^{18}O$	(Vinther et al., 2010)	0-1995	1
Quelccaya, Peru	-13.93	-70.83	Ice Core $\delta^{18}O$	(Thompson et al., 2013b)	226-2009	1
Malpais, New Mexico	34.97	-108.18	Tree Ring Width	(Grissino-Mayer, 1995)	-130-1992	1
Wild Rose, Colorado	39.02	-108.23	Tree Ring Width	(Woodhouse et al., 2006)	1000-2002	1
Upper Wright Lakes, California	36.62	-118.37	Tree Ring Width	(Bunn et al., 2005)	-216-1992	1
Beef Basin, Utah	37.93	-109.8	Tree Ring Width	(Pederson et al., 2011)	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 (Cobb et al., 2003).

spectrum for each record in a piece-wise fashion: power spectral density of GCM-simulated precipitation and temperature, GCM-simulated water isotope variables, and finally the full PSM output to compare with the observed spectrum. The 'perfect model' design tracking climate to proxy space includes assumptions at each step, and identifying these uncertainties explicitly is often challenging. Here we attempt to identify where discrepancies arise when comparing paleoclimate observations to climate model data.

In general, using a single proxy location to constrain climate models is erroneous because of disparities of scale. Various approaches including downscaling or bias correction can help to minimize such problems, or paleoclimate data can be aggregated to match GCM grid cell size. Here, we use single or multiple point-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 Tierney et al., 2015). 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.

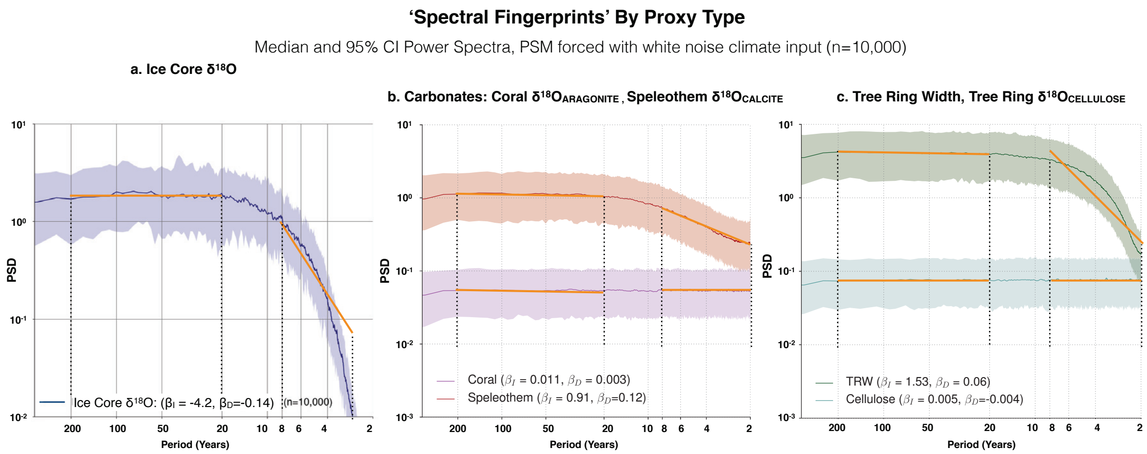

Figure 1: **Spectra and β values by proxy type and as a function of timescale.** Estimated Power Spectral Density for a purely white noise input climate signal, leaving the effects of the proxy system only. Each PSM was forced with white noise climate input 10,000 times. The median spectrum is shown 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 $\delta^{18}O$.

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. Cook et al., 2004), 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 (Cobb et al., 2003; Emile-Geay et al., 2013a). 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 $\delta^{18}O$ (Thompson et al., 2011). Table 2 gives the β values for simulated vs. observed $\delta^{18}O_{\text{CORAL}}$. The coral PSM, as described in Thompson et al. (2011), is a simple bivariate linear model:

$$\delta^{18}O_{pseudocoral} = \alpha_1 \cdot SST + \alpha_2 \cdot SSS \tag{1}$$

where coefficient $\alpha_1 = -0.22‰/°C$ is the relationship between oxygen isotopic equilibrium and formation temperature of carbonates, and α_2 is the empirical estimate for the regional SSS-$\delta^{18}O_{SW}$ slope reported by LeGrande and Schmidt (2006). To investigate the impact of parametric uncertainties on the measured signal, we changed (α_2) to alter the degree of the coral $\delta^{18}O$ sensitivity to local salinity. With $\alpha_2 = 0$, the corals become simple linear responders to temperature. As a second experiment, measured α_2 slopes were increased by 200% to mimic the case where the corals exhibit heightened salinity sensitivity. The results are shown in Fig. 2.

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, Hugershoff.

Uncertainty in α_2 results in only minimal obfuscation of the original climate signal (SST). Regions with large variability in salinity will exhibit heightened sensitivity to uncertainty in the slope of the regional relationship between SSS and $\delta^{18}O_{SW}$, as the contribution of salinity anomalies to the total simulated coral signal is amplified (Thompson et al., 2013a). However, the narrow window due to salinity variations around the power spectrum shows that the corals are generally strong SST proxies (or, possibly, that the GCM completely underplays salinity variability). Testing the effects of parametric uncertainty for the corals provides an example of how PSMs can be used to inform data-model comparison. More interestingly, discrepancies exist between the simulated and observed power spectrum on decadal to centennial timescales.

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

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

3.2.2. Ice Cores

We selected ice core data from four different sites representing a wide geographic range (see Tab. 1). The Quelccaya ice cores record a combination of precipitation and temperature changes driven by tropical Pacific SST variability (Thompson et al., 2013b). The Greenland (NGRIP) and Svalbard ice cores (Austfonna and

Lomonosovfonna) also record a combination of temperature and precipitation variability (Vinther et al., 2010; Isaksson et al., 2005).

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

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

Indeed, when we discount the effects of diffusion, there is more agreement in β between the ice core records and the simulated temperature, precipitation, and water isotope ratios. For example, Fig. 4 shows the simulated vs. observed spectra for NGRIP varying the diffusion length (σ) in the ice core PSM in 1000 random simulations from $\frac{1}{2}$ to $2 \cdot \sigma$, and we find that variance loss at high frequencies increases with increasing diffusion length. Conversely, when we remove the diffusion and compaction model, the simulated $\delta^{18}O_{\mathrm{PRECIP}}$ (dark blue line in Fig. 4) shows a much flatter spectral slope; the simulated $\delta^{18}O_{\mathrm{ICE}}$ shows agreement with the observed spectrum for NGRIP when the diffusion and compaction processes are applied (purple line). These results suggest that the diffusion model component of the ice core PSM correctly estimates the effects of down-core diffusion at high-frequencies, reddening the power spectrum at interannual but not at decadal timescales.

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

3.2.3. Speleothems

We selected four tropical South American Speleothem Records which are sub-decadally resolved with a mean resolution 2-5 years (refs: Tab. 1). These records have been interpreted as an archive of variability in local precipitation amount (i.e. Diva de Maura, Novello et al. (2012)), intensity of convection, and the overall strength of the South American Monsoon (Kanner et al., 2013; Reuter et al., 2009; Apaéstegui et al., 2014). Hydroclimatic variability in this region closely follows both tropical Atlantic SSTs as well as tropical Pacific SST variability (Yoon and Zeng, 2010; Nobre et al., 1991). We simulate cave dripwater using a conceptual model which takes temperature, precipitation amount, and the $\delta^{18}O$ of weighted precipitation into account. The PSM also simulates the groundwater storage (i.e. karst transit time, τ) (Partin et al., 2013; Dee et al., 2015a)

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

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

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

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

3.2.4. Tree Ring Cellulose

We employ two published records of Asian Tree-Ring Cellulose $\delta^{18}O$ (Table 1), both which demonstrate the tree ring cellulose isotope ratios are sensitive to regional precipitation and humidity. Our proxy system model for oxygen isotopes in wood cellulose converts water-isotope enabled model output to tree ring cellulose $\delta^{18}O$ (Evans, 2007). Table 2 gives the β values for simulated vs. observed $\delta^{18}O_{\text{CELLULOSE}}$ and power spectra for both simulated and observed oxygen isotopes in tree cellulose are shown in Fig. 6.

Interannual Scaling. On interannual time scales, the simulated spectral slopes vary compared to those of observed Asian tree ring series. The discrepancy is particularly apparent for the Boibar site (panel 6a.), where β_I = 0.88 (vs. 0.41 simulated). We hypothesized that autocorrelation due to soil water storage and isotopic mixing prior to use of the water by the tree would lead to a steepening of the spectra of the GCM + PSM simulated cellulose records. However, the β_I values for simulated oxygen isotopes in tree cellulose are variable compared to observations.

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

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

3.2.5. Tree Ring Width

Previous work by Franke et al. (2013) suggested that biological proxies such as trees tend to add in autocorrelation, which makes proxy records ‘redder’ than the background climate they are recording. Hydroclimate tree-ring proxies often reflect soil moisture rather than temperature or precipitation, which exhibits higher PSD at decadal frequencies. Since we rely on these records for reconstructions, these reconstructions may tend toward more low-frequency variability compared to the input climate signal. Proxy system modeling addresses this spectral bias: we model the growth response of trees with climate for Western North America using the VS-Lite forward model (Tolwinski-Ward et al., 2010). The four sites we consider (Tab. 1 for references) record over 1000 years of climate and is sensitive to a combination of precipitation and temperature. However, tree growth is influenced by a complex relationship with age which must be removed by detrending the data before a chronology can be used as a climate proxy. Invariably, detrending either partially removes the low-frequency climate signal from the record, or over-emphasizes some of the age-growth relationship masquerading as climate (Cook et al., 1995). In orsder to fully mimic real tree-ring chronologies, we use VS-Lite to model each individual tree at the site using simulated temperature and precipitation then add in the age-growth curves from the actual trees calculated using the regional curve standardization method (RCS) (Briffa et al., 1992; Dupouey et al., 1992). We also incorporate a small error term to mimic the natural variability within a forest, such that 60% of each tree's variance is controlled by climate and 40% is noise. As each tree is modeled individually, most of the noise is removed when we combine multiple trees into a single chronology. Finally, we pre-whiten and build a chronology with the pseudoproxy trees using ARSTAN (Cook, 1985) and six different detrending methods (RCS, negative exponential, modified negative exponential, linear, spline, and Hugershoff) using the same methodology employed when creating a real-world chronology.

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

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

Decadal to Centennial Scaling. For each site, the spectrum of observed tree ring width proxy agrees well (in terms of similarity of power spectra) with the pseudoproxy detrended with the same method. To illustrate this fully, Figures S7 and S8 in the SI show the full spectra for five frequently used detrending options used on each of our four sites. The power spectra for each of the detrending methods diverge at periods 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 (PAGES2K Consortium, 2013). 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 Ault et al. (2013) and Laepple and Huybers (2014a,b)), 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 PSMs. Nonetheless, the broader PAGES2k-PSM comparison conducted in Section 4 generally suggest that greater care is needed in contrasting raw GCM data directly with paleoclimate data, due to the reddening effect of many proxies on the input climate signal.

We acknowledge a number of limitations of this work. In practice, it is difficult to diagnose how much of the ‘mismatch’ between models and data is due to uncertainties in the paleoclimate data itself, PSM error, GCM error, or a combination of all three. While our framework attempts to aid in these diagnostics, confirmation of error sources requires further attention. Additionally, we were restricted to a single water-isotope enabled model with a long transient simulation, and as a result our ability to test the reconciliatory power of PSMs with multiple models is limited. At present, a repository of publicly archived water-isotope simulations spanning the last millennium is unavailable. In future work, we hope to strengthen our conclusions using ensemble runs with multiple isotope-enabled GCMs, or by repeating our analysis using a suite of different spectral methods (Vyushin and Kushner, 2009; Rehfeld et al., 2011). Finally, variability observed in paleoclimate data is not fully characterized by its spectral properties alone. Novel methods for times series analysis have been explored in recent decades (e.g., techniques in nonlinear dynamics), and future work using these methods may help improve our comparative data-model analyses and lend further insight toward the underlying causes of the discrepancies.

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

In closing, we note that the mismatch between climate variability in GCMs and proxy archives harbors implications for the predictability of extreme climate events: if GCMs fail to simulate scaling behavior in the climate system, we may be underestimating statistics surrounding future temperature and precipitation changes, as well as the weather phenomena that accompany these changes (e.g. Ault et al., 2014). In particular, we leave ourselves vulnerable to impacts of unpredicted and unexpected low-frequency climate viability.

Moreover, recent studies have highlighted the dependence of climate change rate estimates on measurement timespan, and suggest that the real pace of abrupt climate variability in the past may be dramatically underestimated (Kemp et al., 2015). Given the growing importance of decadal climate prediction, it is important that both the modeling and paleoclimate communities work together to resolve a *best estimate* of both absolute low-frequency climate variability σ_D and β_D using robust methodology. We hope that the work presented in this paper lays the groundwork for more advanced data-model comparison strategies and will lend further insight to this important open question.

Acknowledgements

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

References

Apaéstegui, J., Cruz, F., Sifeddine, A., Espinoza, J., Guyot, J., Khodri, M., Strikis, N., Santos, R., Cheng, H., Edwards, L., et al., 2014. Hydroclimate variability of the South American Monsoon System during the last 1600 yr inferred from speleothem isotope records of the north-eastern Andes foothills in Peru. Climate of the Past Discussions 10 (1), 533–561.

Ault, T. R., Cole, J. E., Overpeck, J. T., Pederson, G. T., Meko, D. M., 2014. Assessing the risk of persistent drought using climate model simulations and paleoclimate data. Journal of Climate 27 (20), 7529–7549.

Ault, T. R., Cole, J. E., Overpeck, J. T., Pederson, G. T., St. George, S., Otto-Bliesner, B., Woodhouse, C. A., Deser, C., 2013. The continuum of hydroclimate variability in western North America during the last millennium. Journal of Climate 26 (16), 5863–5878.

Briffa, K. R., Jones, P. D., Bartholin, T. S., Eckstein, D., Schweingruber, F. H., Karlen, W., Zetterberg, P., Eronen, M., 1992. Fennoscandian summers from A.D. 500: temperature changes on short and long timescales. Climate dynamics 7 (3), 111–119.

Bunn, A. G., Graumlich, L. J., Urban, D. L., 2005. Trends in twentieth-century tree growth at high elevations in the Sierra Nevada and White Mountains, USA. The Holocene 15 (4), 481–488.

Cobb, K. M., Charles, C. D., Cheng, H., Edwards, R. L., 2003. El Niño/Southern Oscillation and tropical Pacific climate during the last millennium. Nature 424 (6946), 271–276.

Comboul, M., Emile-Geay, J., Evans, M., Mirnateghi, N., Cobb, K. M., Thompson, D. M., 2014. A probabilistic model of chronological errors in layer-counted climate proxies: applications to annually banded coral archives. Climate of the Past 10 (2), 825–841.

Cook, E. R., 1985. A time series analysis approach to tree ring standardization (dendrochronology, forestry, dendroclimatology, autoregressive process), Ph.D. Dissertation, University of Arizona.

Cook, E. R., Briffa, K. R., Meko, D. M., Graybill, D. A., Funkhouser, G., 1995. The *segment length curse* in long tree-ring chronology development for palaeoclimatic studies. The Holocene 5 (2), 229–237.
URL http://hol.sagepub.com/content/5/2/229.abstract

Cook, E. R., Woodhouse, C. A., Eakin, C. M., Meko, D. M., Stahle, D. W., 2004. Long-term aridity changes in the western United States. Science 306 (5698), 1015–1018.

Dee, S., Emile-Geay, J., Evans, M., Allam, A., Steig, E., Thompson, D., 2015a. PRYSM: An open-source framework for PRoxY System Modeling, with applications to oxygen-isotope systems. Journal of Advances in Modeling Earth Systems 7 (3), 1220–1247.

Dee, S., Noone, D., Buenning, N., Emile-Geay, J., Zhou, Y., 2015b. SPEEDY-IER: A fast atmospheric GCM with water isotope physics. Journal of Geophysical Research: Atmospheres 120 (1), 73–91.

Dupouey, J., Denis, J., Becker, M., 1992. A new method of standardization for examining long term trends in tree-ring chronologies. Lundqua Report (Sweden).

Emile-Geay, J., Cobb, K. M., Mann, M. E., Wittenberg, A. T., 2013/03/14 2013a. Estimating Central Equatorial Pacific SST variability over the Past Millennium. Part 1: Methodology and Validation. Journal of Climate.

Evans, M., Tolwinski-Ward, S., Thompson, D., Anchukaitis, K., 2013. Applications of proxy system modeling in high resolution paleoclimatology. Quaternary Science Reviews 76 (0), 16 – 28.
URL http://www.sciencedirect.com/science/article/pii/S0277379113002011

Evans, M. N., Jul. 2007. Toward forward modeling for paleoclimatic proxy signal calibration: A case study with oxygen isotopic composition of tropical woods. Geochemistry Geophysics Geosystems 8 (7).
URL http://www.agu.org/pubs/crossref/2007/2006GC001406.shtml

Fraedrich, K., Blender, R., 2003. Scaling of atmosphere and ocean temperature correlations in observations and climate models. Physical Review Letters 90 (10), 108501.

Franke, J., Frank, D., Raible, C. C., Esper, J., Brönnimann, S., 2013. Spectral biases in tree-ring climate proxies. Nature Climate Change 3 (4), 360–364.

Fredriksen, H.-B., Rypdal, K., 2015. Spectral characteristics of instrumental and climate model surface temperatures. Journal of Climate (2015).

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., et al., 2011. The Community Climate System Model version 4. Journal of Climate 24 (19), 4973–4991.

Grissino-Mayer, H. D., 1995. Tree-ring reconstructions of climate and fire history et El Malpais National Monument, New Mexico.

Hays, J. D., Imbrie, J., Shackleton, N. J., et al., 1976. Variations in the Earth's orbit: pacemaker of the ice ages. American Association for the Advancement of Science.

Henriksson, S., Räisänen, P., Silen, J., Järvinen, H., Laaksonen, A., 2015. Improved power-law estimates from multiple samples provided by millennium climate simulations. Theoretical and Applied Climatology 119 (3-4), 667–677.

Huybers, P., Curry, W., 05 2006. Links between annual, Milankovitch and continuum temperature variability. Nature 441 (7091), 329–332.
URL http://dx.doi.org/10.1038/nature04745

Isaksson, E., Divine, D., Kohler, J., Martma, T., Pohjola, V., Motoyama, H., Watanabe, O., 2005. Climate oscillations as recorded in Svalbard ice core $\delta^{18}O$ records between AD 1200 and 1997. Geografiska Annaler: Series A, Physical Geography 87 (1), 203–214.

Kanner, L. C., Burns, S. J., Cheng, H., Edwards, R. L., Vuille, M., 2013. High-resolution variability of the South American summer monsoon over the last seven millennia: insights from a speleothem record from the central Peruvian Andes. Quaternary Science Reviews 75, 1–10.

Kemp, D. B., Eichenseer, K., Kiessling, W., 2015. Maximum rates of climate change are systematically underestimated in the geological record. Nature Communications 6.

Kutzbach, J. E., 1976. The nature of climate and climatic variations. Quaternary Research 6 (4), 471–480.

Laepple, T., Huybers, P., 2013. Reconciling discrepancies between Uk37 and Mg/Ca reconstructions of Holocene marine temperature variability. Earth and Planetary Science Letters 375, 418–429.

Laepple, T., Huybers, P., 2014a. Global and regional variability in marine surface temperatures. Geophysical Research Letters 41 (7), 2528–2534.
URL http://dx.doi.org/10.1002/2014GL059345

Laepple, T., Huybers, P., 2014b. Ocean surface temperature variability: Large model–data differences at decadal and longer periods. Proceedings of the National Academy of Sciences 111 (47), 16682–16687.

Landrum, L., Otto-Bliesner, B. L., Wahl, E. R., Conley, A., Lawrence, P. J., Rosenbloom, N., Teng, H., 2013. Last millennium climate and its variability in CCSM4. Journal of Climate 26 (4), 1085–1111.

LeGrande, A. N., Schmidt, G. a., 2006. Global gridded data set of the oxygen isotopic composition in seawater. Geophysical Research Letters 33 (12), 1–5.
URL http://www.agu.org/pubs/crossref/2006/2006GL026011.shtml

Lomb, N., 1975. A spectrographic study of beta Centauri. Monthly Notices of the Royal Astronomical Society 172 (3), 639–647.

Mann, M. E., Zhang, Z., Rutherford, S., Bradley, R. S., Hughes, M. K., Shindell, D., Ammann, C., Faluvegi, G., Ni, F., 11 2009. Global Signatures and Dynamical Origins of the Little Ice Age and Medieval Climate Anomaly. Science 326 (5957), 1256–1260.

Nobre, C. A., Sellers, P. J., Shukla, J., 1991. Amazonian deforestation and regional climate change. Journal of Climate 4 (10), 957–988.

Novello, V. F., Cruz, F. W., Karmann, I., Burns, S. J., Stríkis, N. M., Vuille, M., Cheng, H., Lawrence Edwards, R., Santos, R. V., Frigo, E., et al., 2012. Multidecadal climate variability in Brazil's Nordeste during the last 3000 years based on speleothem isotope records. Geophysical Research Letters 39 (23).

PAGES2K Consortium, 05 2013. Continental-scale temperature variability during the past two millennia. Nature Geosci 6 (5), 339–346.
URL `http://dx.doi.org/10.1038/ngeo1797`

Partin, J., Quinn, T., Shen, C.-C., Emile-Geay, J., Taylor, F., Maupin, C., Lin, K., Jackson, C., Banner, J., Sinclair, D., Huh, C.-A., 2013. Multidecadal rainfall variability in South Pacific Convergence Zone as revealed by stalagmite geochemistry. Geology.
URL `http://geology.gsapubs.org/content/early/2013/09/06/G34718.1.abstract`

Pederson, G. T., Gray, S. T., Woodhouse, C. A., Betancourt, J. L., Fagre, D. B., Littell, J. S., Watson, E., Luckman, B. H., Graumlich, L. J., 2011. The unusual nature of recent snowpack declines in the North American Cordillera. Science 333 (6040), 332–335.

Pelletier, J. D., Turcotte, D. L., 1997. Long-range persistence in climatological and hydrological time series: analysis, modeling and application to drought hazard assessment. Journal of Hydrology 203 (1), 198–208.

Rehfeld, K., Marwan, N., Heitzig, J., Kurths, J., 2011. Comparison of correlation analysis techniques for irregularly sampled time series. Nonlinear Processes in Geophysics 18 (3), 389–404.

Reuter, J., Stott, L., Khider, D., Sinha, A., Cheng, H., Edwards, R. L., 2009. A new perspective on the hydroclimate variability in northern South America during the Little Ice Age. Geophysical Research Letters 36 (21).

Scargle, J. D., 1982. Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data. The Astrophysical Journal 263, 835–853.

Schulz, M., Mudelsee, M., 2002. Redfit: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. Computers & Geosciences 28 (3), 421–426.

Sturm, C., Zhang, Q., Noone, D., Feb. 2010. An introduction to stable water isotopes in climate models: benefits of forward proxy modelling for paleoclimatology. Climate of the Past 6 (1), 115–129. URL `http://www.clim-past.net/6/115/2010/`

Taylor, K. E., Stouffer, R. J., Meehl, G. A., 2012. An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society 93 (4), 485–498.

Thompson, D., Ault, T., Evans, M., Cole, J., Emile-Geay, J., 2011. Comparison of observed and simulated tropical climate trends using a forward model of coral $\delta^{18}O$. Geophysical Research Letters 38, L14706.

Thompson, D. M., Ault, T., Evans, M., Cole, J., Emile-Geay, J., LeGrande, A., 2013a. Coral-model comparison highlighting the role of salinity in long-term trends. El Niño-Southern Oscillation.

Thompson, L., Mosley-Thompson, E., Davis, M., Zagorodnov, V., Howat, I., Mikhalenko, V., Lin, P.-N., 2013b. Annually resolved ice core records of tropical climate variability over the past ~1800 years. Science 340 (6135), 945–950.

Thomson, D. J., 1982. Spectrum estimation and harmonic analysis. Proc. IEEE 70(9), 1055–1096.

Tierney, J. E., Abram, N. J., Anchukaitis, K. J., Evans, M. N., Giry, C., Kilbourne, K. H., Saenger, C. P., Wu, H. C., Zinke, J., 2015. Tropical sea surface temperatures for the past four centuries reconstructed from coral archives. Paleoceanography 30 (3), 226–252.

Tolwinski-Ward, S. E., Evans, M. N., Hughes, M. K., Anchukaitis, K. J., Nov. 2010. An efficient forward model of the climate controls on interannual variation in tree-ring width. Climate Dynamics 36 (11-12),

2419–2439.
URL http://www.springerlink.com/index/10.1007/s00382-010-0945-5

Treydte, K. S., Frank, D. C., Saurer, M., Helle, G., Schleser, G. H., Esper, J., 2009. Impact of climate and co_2 on a millennium-long tree-ring carbon isotope record. Geochimica et Cosmochimica Acta 73 (16), 4635–4647.

Vinther, B. M., Jones, P. D., Briffa, K. R., Clausen, H. B., Andersen, K. K., Dahl-Jensen, D., Johnsen, S. J., 2 2010. Climatic signals in multiple highly resolved stable isotope records from Greenland. Quaternary Science Reviews 29 (3–4), 522–538.
URL http://www.sciencedirect.com/science/article/pii/S0277379109003655

Vyushin, D. I., Kushner, P. J., 2009. Power-law and long-memory characteristics of the atmospheric general circulation. Journal of Climate 22 (11), 2890–2904.

Vyushin, D. I., Kushner, P. J., Mayer, J., 2009. On the origins of temporal power-law behavior in the global atmospheric circulation. Geophysical Research Letters 36 (14).

Wernicke, J., Grießinger, J., Hochreuther, P., Bräuning, A., 2015. Variability of summer humidity during the past 800 years on the eastern Tibetan Plateau inferred from $\delta^{18}O$ of tree-ring cellulose. Climate of the Past 11 (2), 327–337.

Woodhouse, C. A., Gray, S. T., Meko, D. M., 2006. Updated streamflow reconstructions for the Upper Colorado River basin. Water Resources Research 42 (5).

Yoon, J.-H., Zeng, N., 2010. An Atlantic influence on Amazon rainfall. Climate dynamics 34 (2-3), 249–264.

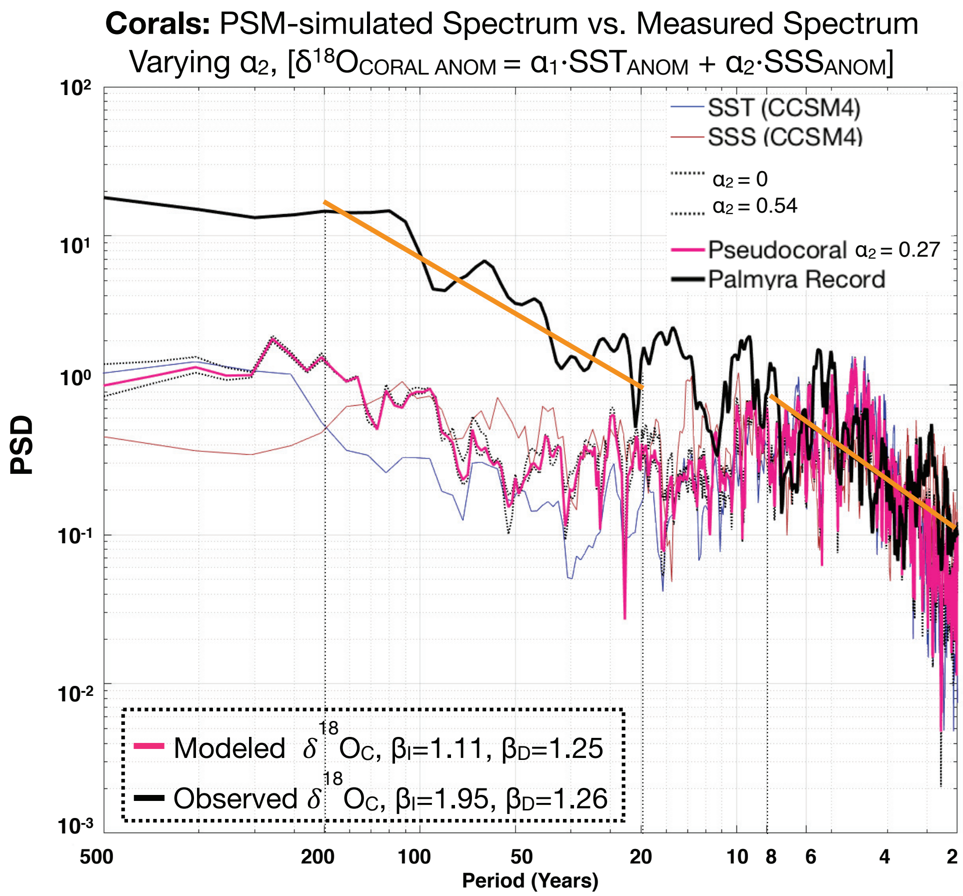

Figure 2: **Coral Aragonite Model-Data Comparison:** simulated (pink) and measured (black) spectrum for Palmyra coral data, as well as the effects of altering α_2 at Palmyra island alongside the climate inputs to the model (SST, SSS). Straight orange lines are the calculated mean slopes across both decadal 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.

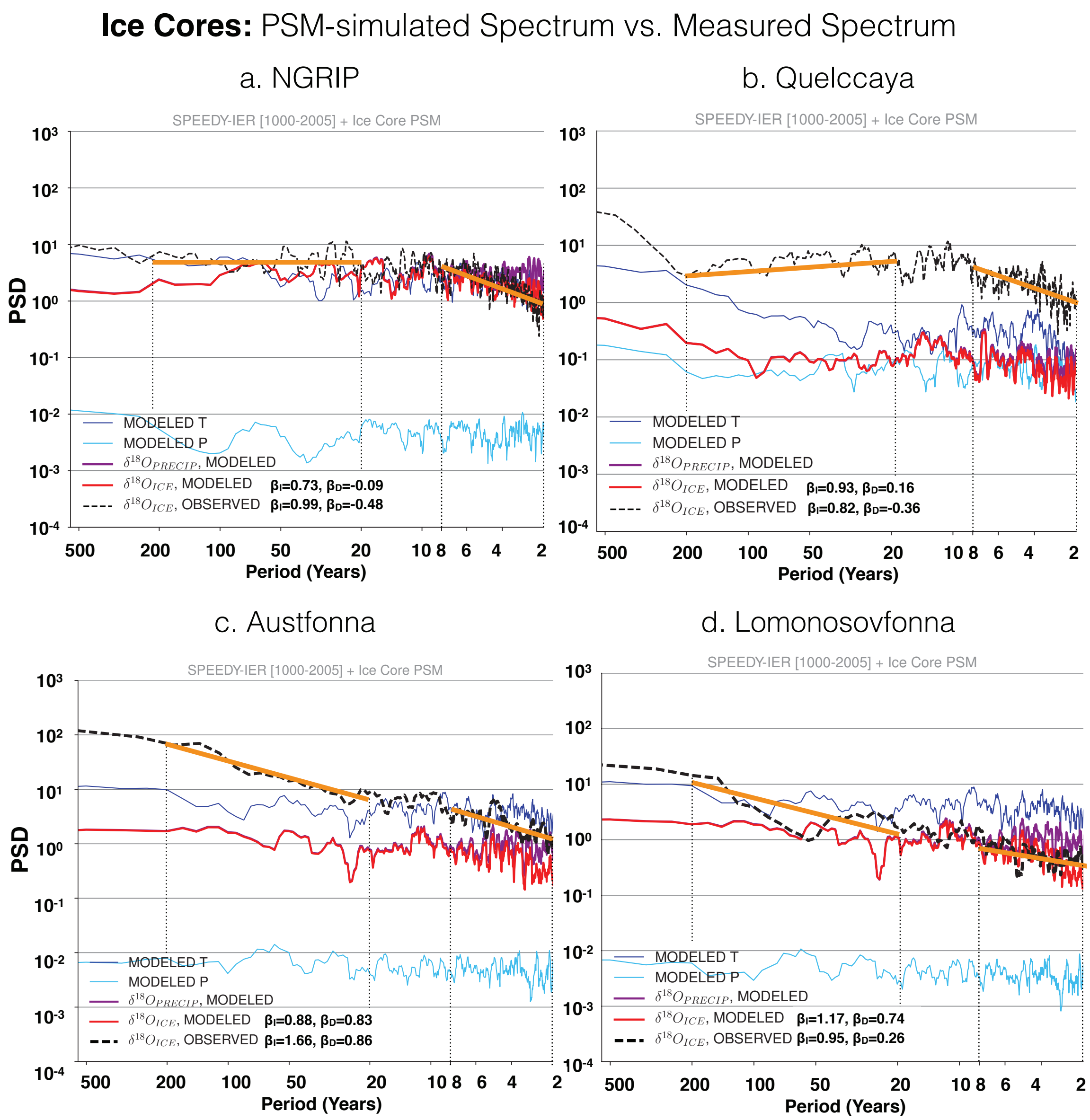

Figure 3: **Ice Core Model-Data Comparison.** Estimated Power Spectral Density for Simulated vs. Observed $\delta^{18}O$ of ice at four sites: (a) NGRIP, (b) Quelccaya, (c) Austfonna, (d) Lomonosovfonna. The PSD for each PSM-generated record is shown (red) alongside climate inputs of temperature (dark 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.

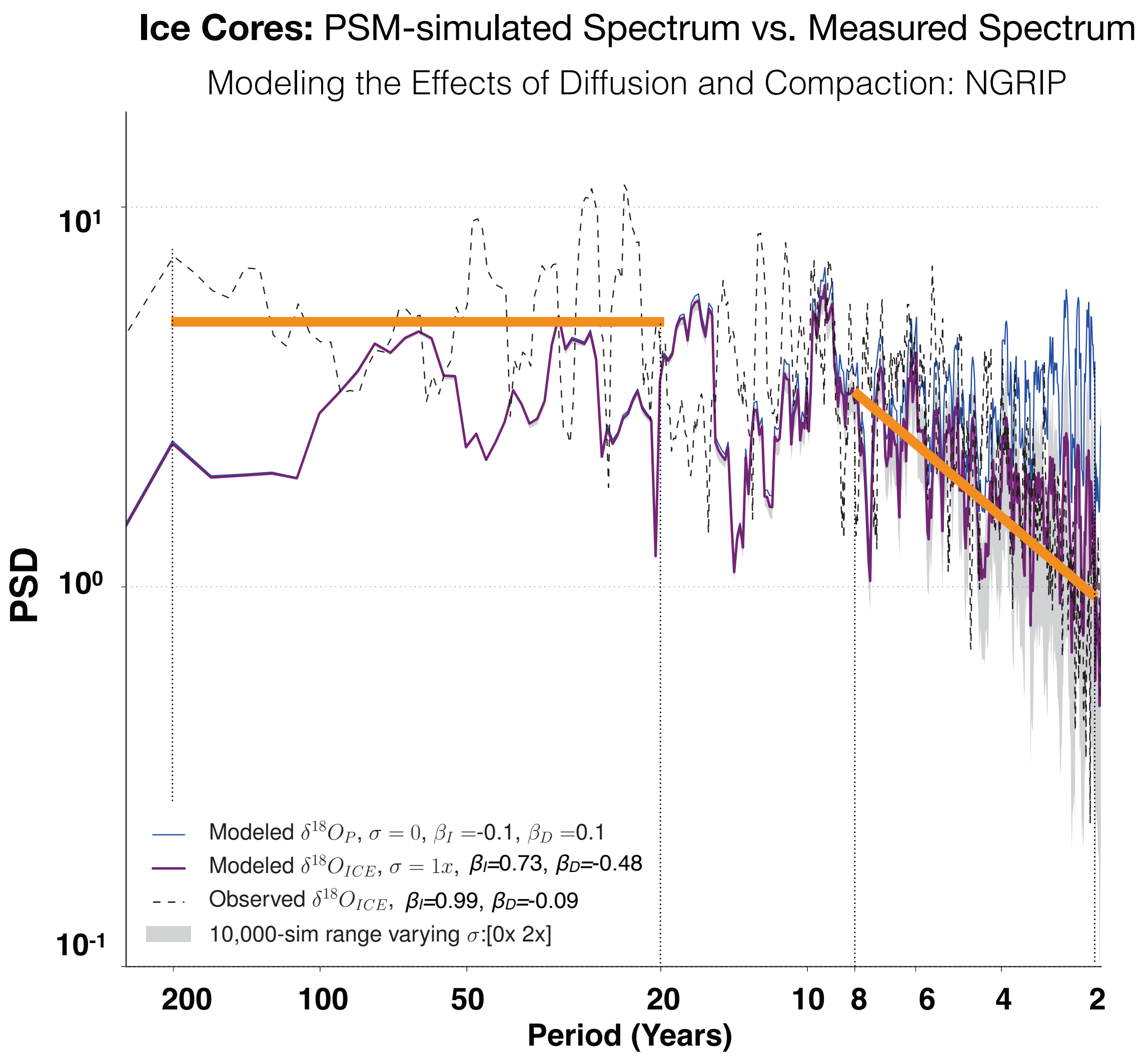

Figure 4: **Ice Core Model-Data Comparison at NGRIP: effect of diffusion length (σ) on β.** Estimated power spectral density for simulated (purple) vs. observed (black) $\delta^{18}O$ of ice and the effect of varying diffusion length parameters. The grey shaded region spans experiments resampling the 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.

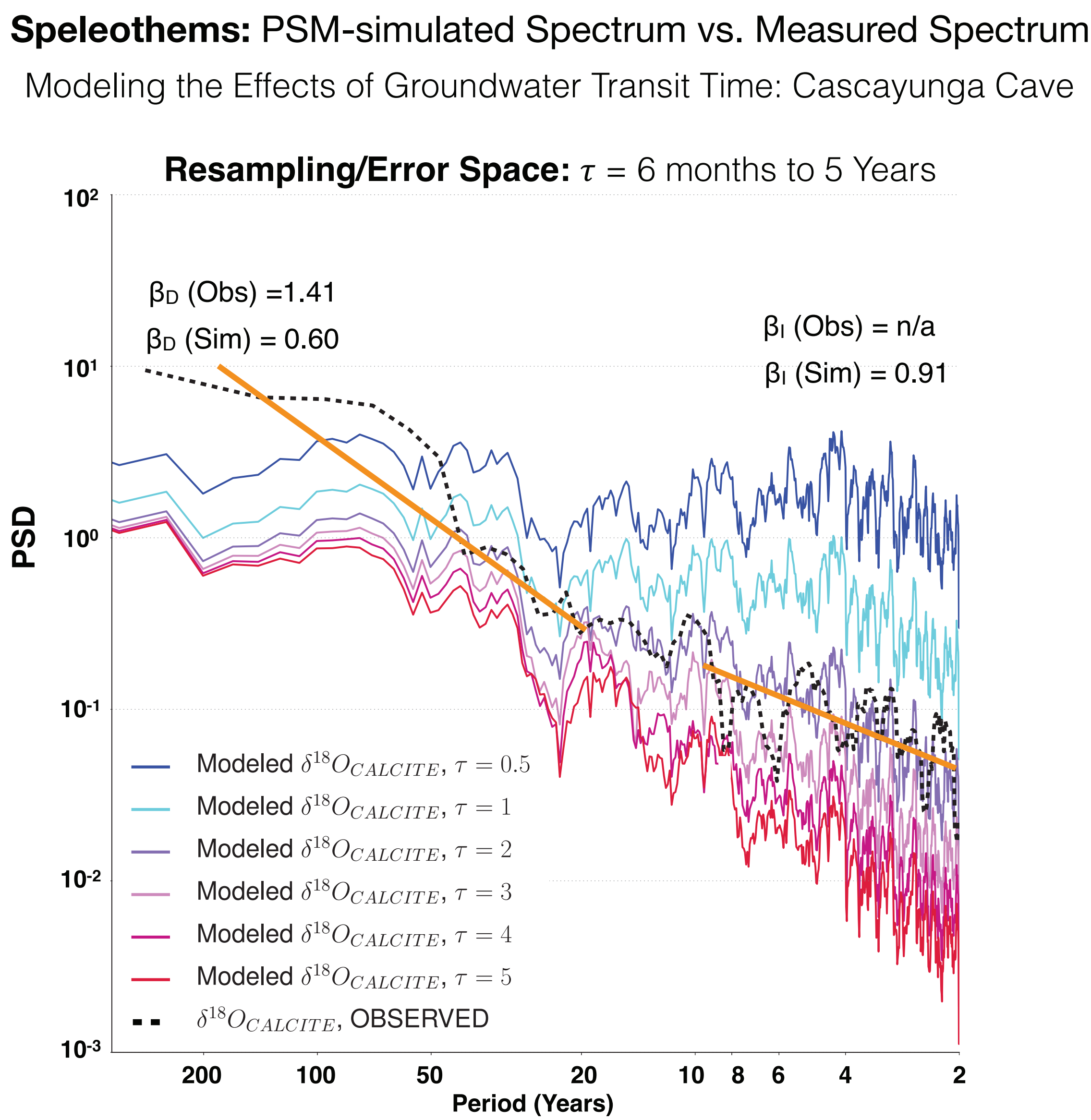

Figure 5: **Speleothem Calcite Model-Data Comparison: the effect of Karst Transit Time (τ) on β.** Simulated vs. Observed $\delta^{18}O$ of speleothem calcite at Cascayunga Cave. PSD is plotted, simulated using an ensemble of plausible values (6 months to 5 years) for the karst transit time parameter, 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.

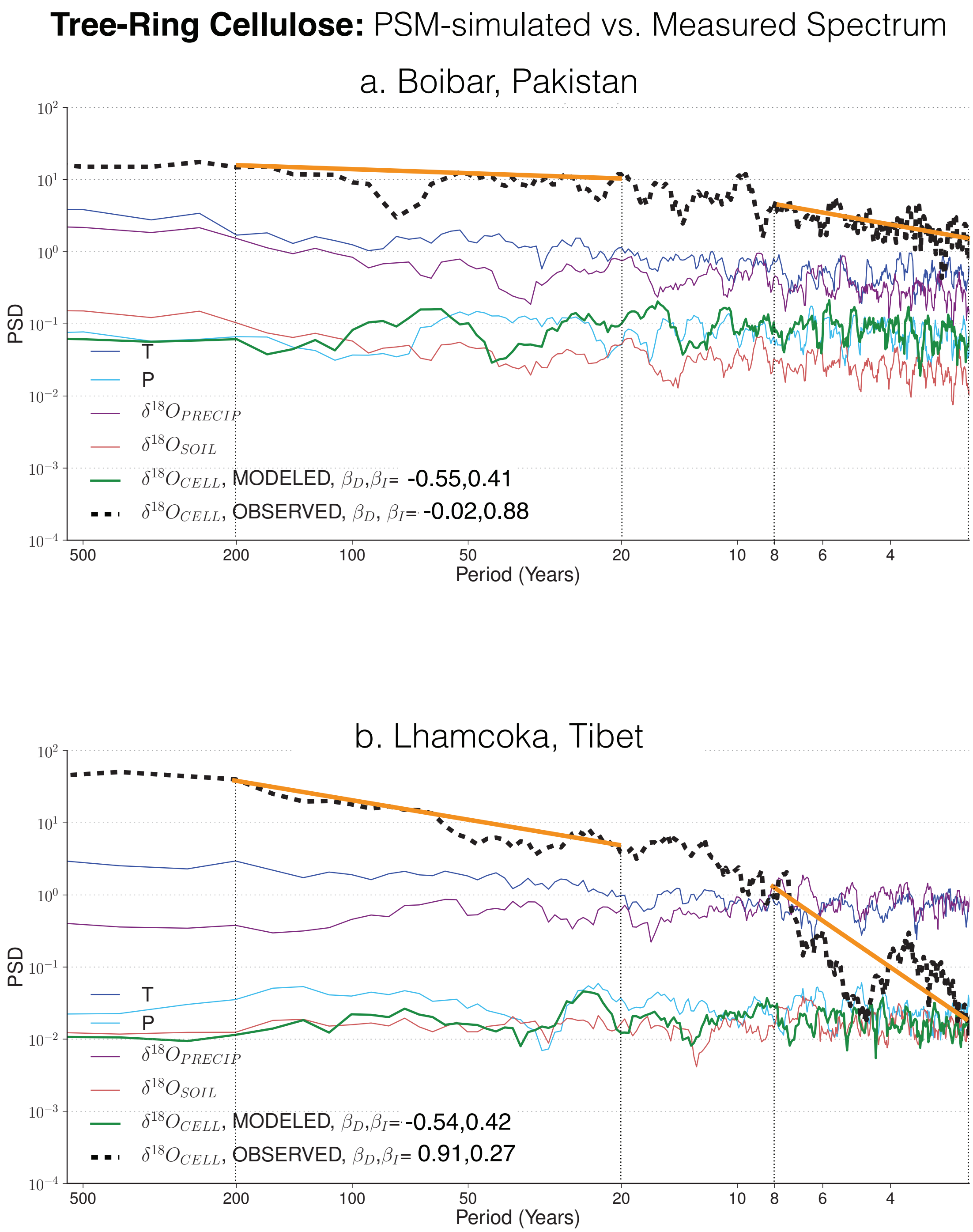

Figure 6: **Tree-Ring Cellulose Model-Data Comparison.** Estimated Power Spectral Density for Simulated (green) vs. Observed (black) $\delta^{18}O$ of tree ring cellulose at (a) Boibar, Pakistan and (b) Lhamcoka, Tibet. We removed the mean from all fields prior to computed the PSD.

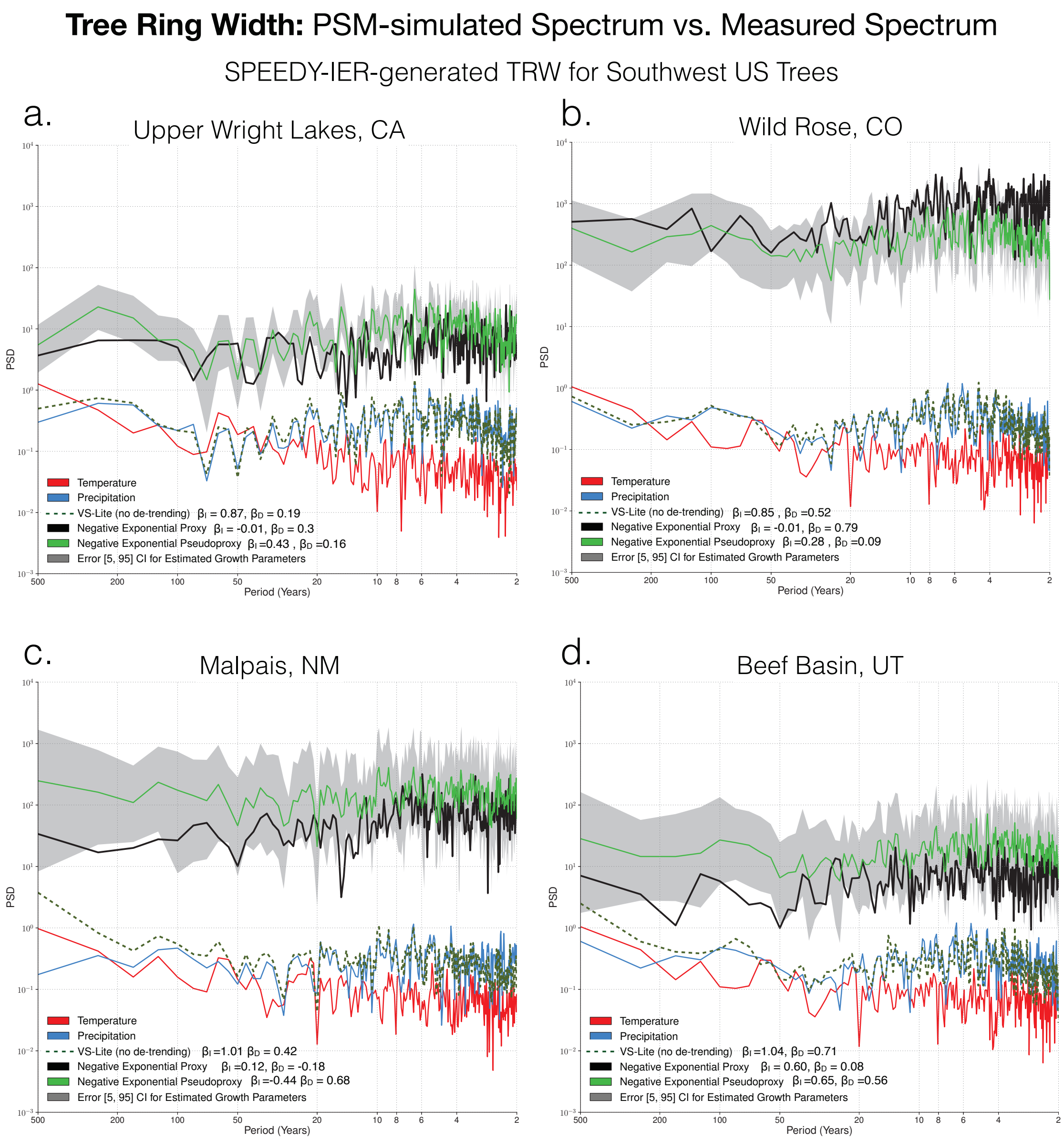

Figure 7: **Tree-Ring Width Model-Data Comparison:** the effect of detrending method on β. For each site, the original measured TRW is plotted in black, along with the PSM (VS-Lite) inputs of temperature (red), precipitation (blue), and the simple VS-Lite generated pseudoproxy data (dark green dashed), 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).

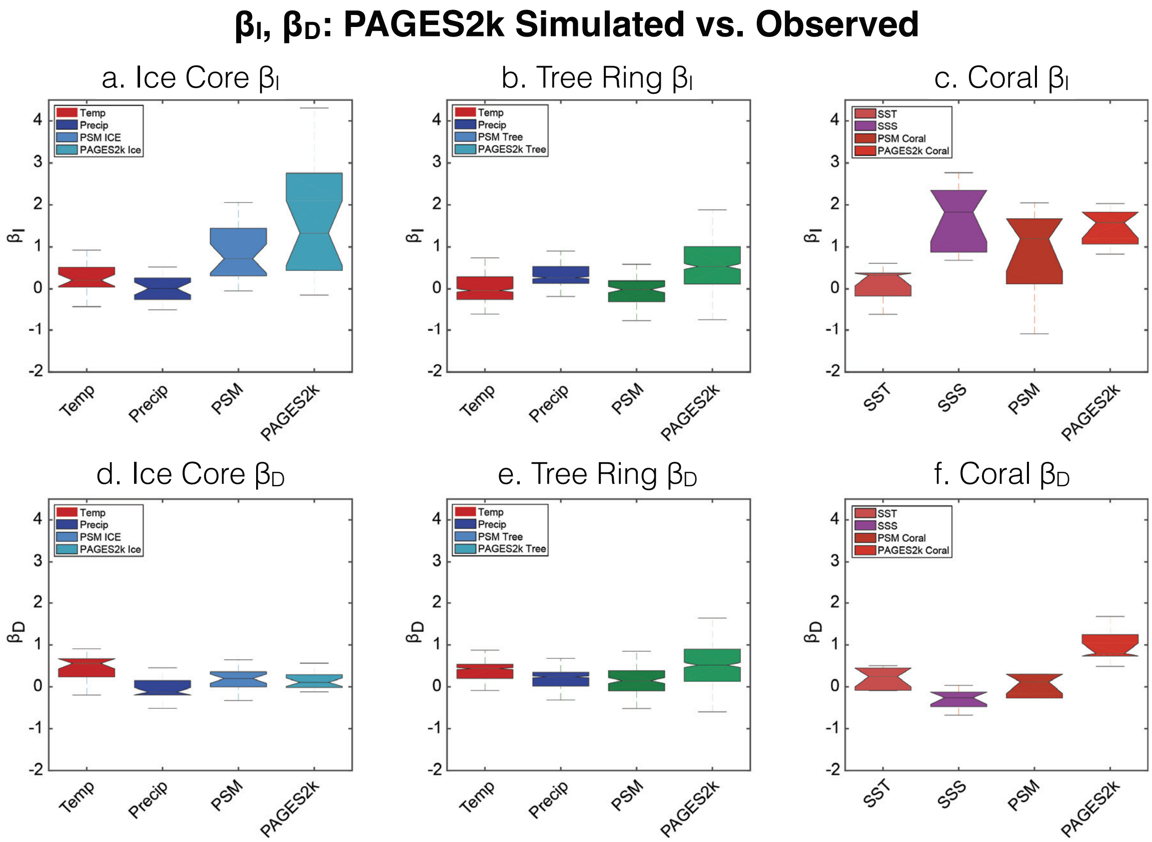

Figure 8: **Distribution of proxy β_I, β_D values, GCM+PSM vs. Observed, for the PAGES2k Phase 1 dataset**: a. ice cores (blue, left), tree-ring width (green, center), and coral data (orange, right) from PAGES2k (lighter colors) and GCM/PSM (darker colors). (a, b., d., e.): temperature (red), precipitation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