



Robust periodic signals in proxy records with chronological uncertainty and variable temporal resolution

István Gábor Hatvani ^{a,*}, Péter Tanos ^b, Manfred Mudelsee ^{c,d,e}, Zoltán Kern ^a

^a Institute for Geological and Geochemical Research, Research Centre for Astronomy and Earth Sciences, Eötvös Loránd Research Network, Budaörsi út 45, H-1112, Budapest, Hungary

^b Hungarian University of Agriculture and Life Sciences, Technical Institute, Páter Károly utca 1, H-2100, Gödöllő, Hungary

^c Climate Risk Analysis, Kreuzstrasse 27, Heckenbeck, 37581, Bad Gandersheim, Germany

^d Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bussestrasse 24, 27570, Bremerhaven, Germany

^e Advanced Climate Risk Education, Kreuzstrasse 27, Heckenbeck, 37581, Bad Gandersheim, Germany

ARTICLE INFO

Article history:

Received 20 April 2021

Received in revised form

6 November 2021

Accepted 22 November 2021

Available online 8 December 2021

Handling Editor: A. Voelker

Keywords:

Sedimentary records

Spectral analysis

Speleothems

Deep sea sediments

Uneven spacing

Chronological uncertainty

Holocene

Pleistocene

CUSP model

ABSTRACT

The detection of periodic fluctuations is important in the quest for a deeper understanding of the drivers of past climates in the interests of being better able to understand the climate changes which are likely in the decades to come. Paleoclimatological information derived from natural archives is typically accompanied by chronological uncertainty and variable temporal resolution, both of which complicate the analysis of their time series and both are often ignored. This has, however, changed in recent years with the development of statistical tools supporting, e.g. spectral analysis tasks, which aim to take these problems into account. In cases where the original data is no longer available, it may not be possible to assess the reliability of published reports of periods detected. In this study we aim to test/model whether or not a signal for a given period can be robustly detected from a sedimentary proxy record considering its mean sampling resolution and degree of chronological uncertainty. To achieve this aim, annually sampled time series free of gaps and timescale error were modeled with white or red noise, resampled in a controlled way to simulate different time resolutions with timescale uncertainty, ultimately mimicking a real-life sedimentary record. In fact, an ensemble of potential timescales was retrieved, and their spectral characteristics explored. It was found that: (i) although sampling frequency (i.e. temporal spacing) is limiting from the side of the smallest-period, (ii) at higher mean sampling resolutions, it can ameliorate the detectability of periodic signals even in the presence of timescale error; furthermore, (iii) the increase in mean sampling resolution is less influential in an autocorrelated time series, since more information is retained due to the phenomenon of persistence. An online tool called CUSP was also developed, which gives a suggestion whether a given period can be considered to be present in a robust way utilizing our test results.

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Natural archives (e.g. speleothems, lake- and ocean sediments) are globally distributed and preserve information about the environmental conditions that prevailed during their formation, thus providing vital information on past climate(s) (Bradley, 1999). As

such, these archives are valuable for intra- and intercontinental assessment of past global changes and allow the investigation of paleoclimatic conditions and their variability (Sorooshian and Martinson, 1995; Wefer et al., 1999). One of the key tools in exploring the climate signal is spectral analysis (Trauth, 2021; Yiu et al., 1996). This can improve the prediction of climatological changes and provide information crucial to the understanding of the relationship between driving forces and the changes they induce, for example, via cross power spectral density analysis (Kunz et al., 2020). Paleoclimatological information (both pristine proxies and inferred reconstructions) derived from sedimentary records is, however, typically accompanied by variable temporal resolution and timescale error complicating time series analysis (Bradley,

Abbreviations: CUSP, Chronological Uncertainty and Spectral Peaks.

* Corresponding author.

E-mail addresses: hatvaniig@gmail.com (I.G. Hatvani), tanospeter@gmail.com (P. Tanos), mudelsee@climate-risk-analysis.com (M. Mudelsee), zoltan.kern@gmail.com (Z. Kern).

1999; Mudelsee, 2014).

Quaternary paleoclimatic archives can be accurately dated using radiometric methods and/or lamina counting (Walker, 2005). The age-depth relationship can be determined for the sequence and described with the use of age-depth models constructed with different statistical methods (Trachsel and Telford, 2016). Special efforts are made (e.g. Blaauw and Christen, 2011; Breitenbach et al., 2012; Hercman and Pawlak, 2012) to improve 'uncertainty aware' age-depth modeling.

The random and systematic error of the dating method is inevitably propagated onto the age-depth model. While a systematic error in the form of a temporal shift is a critical issue in (i) the evaluation of correlations with other (independently dated) records (Franke and Donner, 2019) or (ii) the estimation of a lead-lag relationship (Hatvani et al., 2018; Mudelsee, 2001), the analysis of spectra is obviously not influenced by the presence of a time shift. Hence, we assume that systematic dating errors do not interfere with the spectral assessment techniques studied here to such an extent that would render the comparison of the spectra meaningless. In other words, the present study focuses on the effects of random uncertainty.

This is why the development of techniques capable of evaluating the relationship between series of data taking into account uncertainties represents an important task for the paleoclimate community if full advantage is to be taken of paleoclimate records (Haam and Huybers, 2010; Mudelsee et al., 2009). A task supplementary to this is the development of tools to determine how much the results are biased if the previously mentioned circumstances are unaccounted for.

The uncertainties in the dating of proxy archives and variable temporal resolution originating from continuous, but not linear sedimentation may strongly bias the observed spectral characteristics (Meyers et al., 2008; Zeeden et al., 2018). The development of statistical tools accounting for chronological uncertainty is an active topic (Amrhein, 2020; Franke and Donner, 2019; Zeeden et al., 2018). In numerous cases, however, these phenomena are ignored in the course of spectral analysis (Amrhein, 2020). Such tools can usually be used for spectral analysis, as they are already equipped to account for these complex circumstances, but their use requires raw data sets and a knowledge of the size(s) of chronological uncertainties. These complicating circumstances have, however, often been ignored, and if the original data are not available, it is not possible to assess the reliability of the periods detected as reported in publications, even with the use of tools capable of the correct handling of such tasks, for example, GeoChronR (McKay et al., 2021) in R, *Acycle* (Li et al., 2019) in MATLAB. Nevertheless, it is of vital importance that information on whether the documented period can be considered robustly present in the studied record under the given circumstances (chronological uncertainty, mean sample resolution).

Originally, "robust" is a technical term from statistical literature (Box, 1953), attesting to the fact that a data-analytical method gives reliable results even in the presence of violations of the assumptions made in the course of the analysis. Hence, one may call a method, or a result obtained using a particular approach, robust against timescale uncertainties of a certain size.

The present study aims to fill this niche by testing whether a given period can be considered to be present in a robust way in a time series if the mean chronological uncertainty and mean sampling resolution (i.e. temporal spacing) are taken into consideration. This is done by modeling the spectral bias caused by timescale error on simulated time series resembling the characteristics of real-life sedimentary proxy records. The model forms the basis of an online tool called CUSP, short for Chronological Uncertainty and Spectral Peaks, which gives a suggestion whether a

given period can be considered to be present in a robust way. It is expected that the detectability of a periodic signal is (i) degraded with the increase of chronological uncertainty and (ii) increases with higher sampling resolution.

Section 2 describes the derivation of the modeled periodic time series with chronological uncertainty and different sampling resolution and the methodology proposed for the investigation of the spectral bias caused by these. Section 3 presents the results and shows the practical applicability to real-life examples and an online tool derived from the model. Section 4 provides an outlook describing the possible future perspectives and possible developments of the model. Section 5 summarizes the novelty of the model presented in the study and concludes the new findings. Finally, Section 6 gives the availability of the software developed within the framework of the research.

2. Materials and methods

2.1. Modeled periodic time series with chronological uncertainty and different mean sampling resolution

The main steps in deriving the time series mimicking real-life sedimentary records were to derive single-period time series, both with white and red noise (Part 1), at the same time derive a simulated 'time-step'-depth scale with timescale uncertainty (Part 2) and match them together based on their 'time-step' number (Part 3). Lastly, the time series are randomly rarefied in a controlled way (Part 4) (Fig. 1); this procedure is then repeated 1000 times to generate an ensemble.

The procedure of deriving the time series mimicking a sedimentary record can be formalized as follows:

Part 1: An undisturbed time scale was taken, defined as: $T(i)$, $i = 1, \dots, 7500$, so $T(i)$ are evenly spaced with a spacing equal to 'time-step'.

Part 1.2: Sinusoidal signals were generated as $X_{\sin}^j(i) = A \times \sin\left(2\pi \frac{1}{P^j} T(i)\right)$ where $A = 40$ and P^j is an element of $\{25, 50, 75, 100, 125, 150, 175, 200, 250, 300\}$ and j is the index.

Part 1.3: White and red noise series were generated for each sinusoidal signal (Fig. 2).

Part 1.3.1: White noise $\phi_{\text{white}}^j(i)$ was generated as a set of normal distributed $N(0, 1)$ random variates (Eq. (1))

$$\phi_{\text{white}}^j(i) = \varepsilon_{N(0,1)}(i), \quad i = 1, \dots, 7500. \quad (1)$$

Part 1.3.2: Red noise $\phi_{\text{red}}^j(i)$ was generated using a random standard normal variable, and a temporally autocorrelated variable (Eq. (2)):

$$\phi_{\text{red}}^j(1) = \varepsilon_{N(0,1)}(1), \quad i = 2, \dots, 7500. \quad (2)$$

$$\phi_{\text{red}}^j(i) = a \times \phi_{\text{red}}^j(i-1) + \varepsilon_{N(0,1-a^2)}(i), \quad \dots, i = 2, \dots, 7, 500$$

For red noise, the lag-1 coefficient $a = 0.5$, which is considered a conservative value in the characterization and reflection of the persistence characteristics of natural records (Mudelsee, 2014; von Storch et al., 2009); regarding the random innovation, one uses $1 - a^2$, since then the process is stationary (i.e., constant mean and constant variance); for further details see Mudelsee (2014): Section 2.1.1 therein.

Part 1.4: Composites of noisy time series were generated for the white- and red-noise models (Eqs. (3) and (4)) scaling the noise

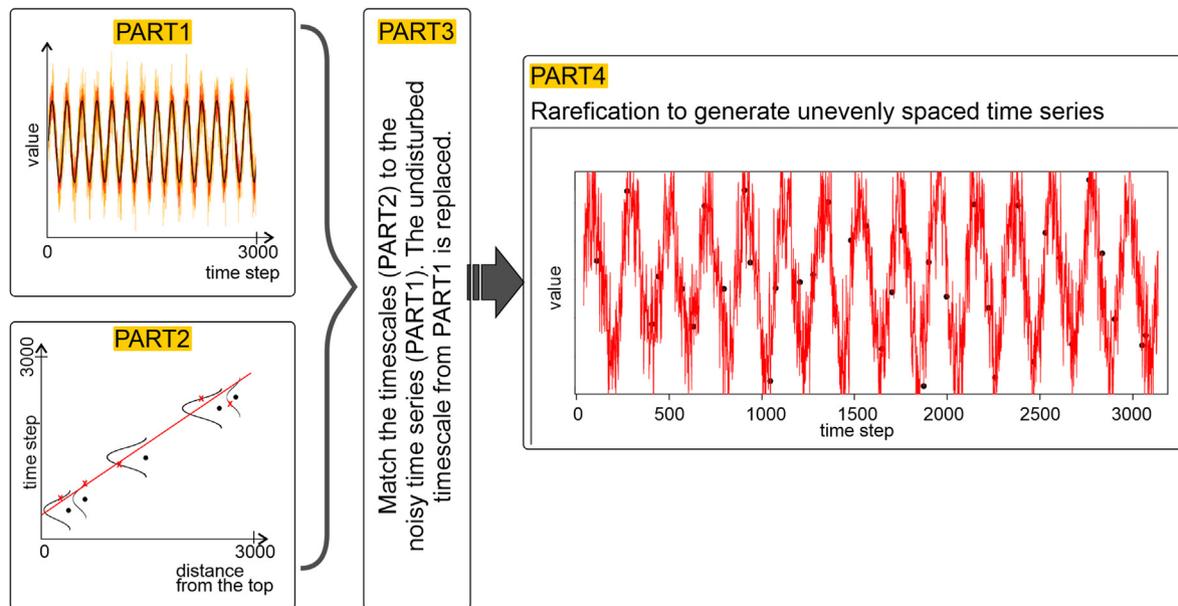


Fig. 1. Flowchart illustrating the construction of periodic signals. Part 1: time series with an implemented periodic signal (black line), white noise (shown, for visibility, as the orange line) and red noise (red line). Part 2: Creating different time scales. Black dots represent randomly selected five depths from a perfect time scale and the red dots the chosen value retrieved randomly from the normal distributions' bell curves, while the red line indicates the best fit linear model. Part 3: Matching the timescales to the noisy periodic series, Part 4: Rarefying the data to obtain the unevenly spaced jittered time series. Red line represents the time series, and the black dots are the actually used values. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

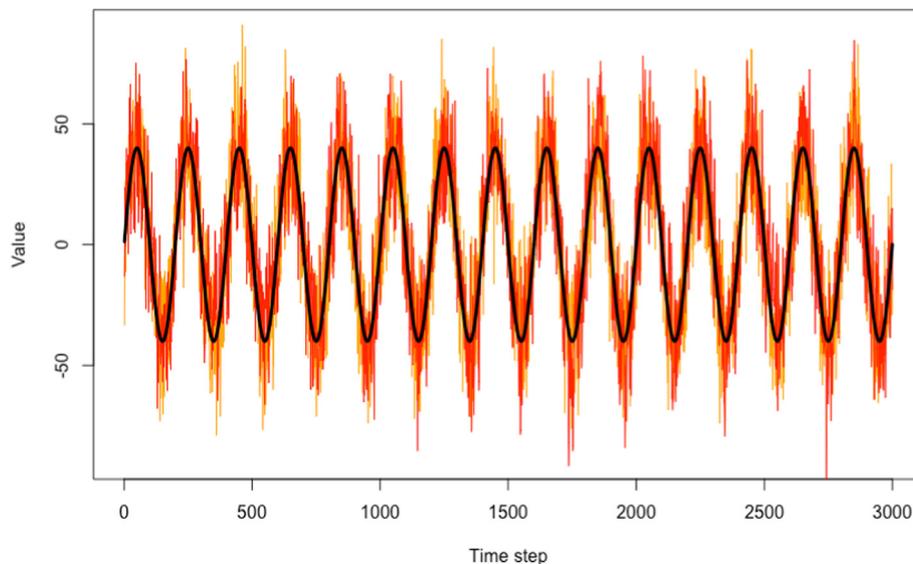


Fig. 2. Periodic time series showing 3000 steps (with an undisturbed timescale) with an implemented periodic signal ($f = 0.005 \text{ time step}^{-1}$) (black line) and white noise (shown, for visibility, as the orange line) and red noise (red line). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

series (shown in Part 3) with $0.333 \times \text{VAR}(X_{\text{sin}}^j(i))$:

$$X_{\text{white}}^j(i) = X_{\text{sin}}^j(i) + 0.333 \times \text{VAR}(X_{\text{sin}}^j(i)) \times \phi_{\text{white}}^j(i) \quad (3)$$

$$X_{\text{red}}^j(i) = X_{\text{sin}}^j(i) + 0.333 \times \text{VAR}(X_{\text{sin}}^j(i)) \times \phi_{\text{red}}^j(i). \quad (4)$$

In the following calculations, only the first n elements will be used from the periodic time series where:

$$n = \begin{cases} 3000, & P < 150 \\ P \times 25, & P \geq 150 \end{cases}.$$

Part 2: A series of steps was generated, $Z(i), i = 1, \dots, n$; the steps were considered to be distances from the top (step 1) to bottom (step n) of a sedimentary sequence. Out of these five depths were selected (i.e. top, bottom, and three additional randomly chosen depth in between) to serve as dating locations, as in the process of tie-point modeling, where the aim is to estimate ages, along with their uncertainties, at multiple depths downcore (McKay et al., 2021). Then normal distributions were derived for all five dating

locations with the mean of the distribution being equal to the numeric value of the depth itself. The standard deviation of the normal distributions was set to half of the predefined degree of chronological uncertainty (Table 1) for the sake of reproducibility.

A value was retrieved randomly from the five normal distributions to serve as the age at the chosen depth. Only ages older than any previous were accepted. Model dates for $z(i)$ were retrieved from the best fit linear model with the aim of mimicking a distorted timescale:

$$\tilde{T}(i) = i = 1, \dots, n. \quad (5)$$

The procedure was repeated 1000 times for each combination of j and noise:

$$\tilde{T}_{j,\text{white}}^s(i) \text{ and } \tilde{T}_{j,\text{red}}^s(i) \quad \text{where } s = 1, \dots, 1000, i = 1, \dots, n \quad (6)$$

these ensembles served to mimic the chronological uncertainty (Unc).

Part 3: Timescales were matched to the noisy periodic series based on ‘time-step’ number:

$$\left\{ \tilde{T}_{j,\text{white}}^s(i), X_{\text{white}}^j(i) \right\} \text{ and } \left\{ \tilde{T}_{j,\text{red}}^s(i), X_{\text{red}}^j(i) \right\}. \quad (7)$$

Altogether 20 periodic time series were assessed, each with 1000 different simulated time scales for both red- and white noise models.

Part 4: All of the time series were rarefied/resampled with the difference in temporal spacing between the selected data points - mean sampling resolution (MSR) - spanning from five steps to a maximum relative to the implanted periods (Table 1); for an example see Fig. 3. In this way, MSR represents the number of time steps which could be yr or kyrs for example.

Technically, MRS and Unc were relativized to the implanted period (P) in each set, in which the minimum and the maximum MSR is 5 and $P/2.5$, respectively, while Unc varies between $P/5$ and $P \times 4$, resulting in 1786 combinations (Table 1). The increment was 5 and 10 steps for MSR and Unc respectively up to $P = 125$. For larger values, according to preliminary assessment, the increment was increased to save computational resources without losing information (Table 1).

2.2. Spectral analysis and evaluation

The spectral bias caused by the chronological uncertainty and uneven spacing was assessed by determining the period with the highest power (peak period) from the original unmodified record with no chronological uncertainty and even sampling (MSR = 1 time step per sample), and the modified ones (MSR >1) using REDFIT (Schulz and Mudelsee, 2002). The shift in the location of the

Table 1

Spectral characteristics of the time series created (implanted period) and the resampling criteria: Mean sampling resolution (MSR) and age uncertainty (Unc).

Period (P)	MSR (from, to, increment)	Unc (from, to, increment)
25	5, 10 ($P/2.5$), 5	5 ($P/5$), 100 ($P \times 4$), 10
50	5, 20, 5	10, 200, 10
75	5, 30, 5	15, 300, 10
100	5, 40, 5	20, 400, 10
125	5, 50, 5	25, 500, 10
150	5, 60, 5	30, 600, 30
175	5, 70, 13	35, 700, 35
200	5, 80, 15	40, 800, 40
250	5, 100, 19	50, 1000, 50
300	5, 120, 23	60, 1200, 60

peak period in the modified periodic time series resembling a sedimentary record was compared to the known location of the peak period in the unmodified time series. In the case of any given Unc-and-MSR combination, the spectral characteristics of all 1000 records from the REDFIT ensemble spectra were evaluated and compared to that of the unmodified one. The number of cases was counted when the shift in the peak period was <5% of the implanted period out of the 1000 cases (e.g. Fig. 4), in other words it was within the chosen tolerance interval ranging from $0.95 \times P$ to $1.05 \times P$. The percentage of well detected significant cycles within the tolerance interval are referred to as the success rate (sr), while $1-sr$ accounts for the false positives and the false negatives together.

Because single-period datasets are used, aliasing is avoided, thus there is no source of additional false positives to occur than the ones described above. In addition, MSR was chosen in such a way that the upper bound of the ‘Nyquist frequency’ (Blackman and Tukey, 1958) would not go above the chosen period (Table 1).

Specifically, REDFIT determined if the powers of the time series’ Lomb–Scargle–Fourier transform periodograms were significant against red noise background from a first-order autoregressive process. A Welch I window with two overlapping (50%) segments was chosen, the oversampling parameter was set as $ofac = 24$, the highest frequency factor was set as $hifac = 1$, while the number of Monte Carlo simulations ($nsim$) to obtain the significance levels was chosen to be 1000, similar to real life applications with sedimentary proxy records (Hatvani et al., 2018; Holzkämper et al., 2004; Neff et al., 2001). The red-noise boundary was estimated to be equal to the bias-corrected 90% chi-squared limit of a fitted AR(1) process. It could have also been a viable approach to use a power-law model to describe background climate variability as in e.g. Huybers and Curry (2006), Pelletier (1998), Vyushin and Kushner (2009). However, in practice it does not make a great change (over the major part of the frequency interval) whether AR(1) or power-law are applied, see Mudelsee (2014): Section 2.5.2 therein. Moreover, in spectrum estimation (and many more other test situations) the preference is for significance, not power (Mudelsee, 2014; Priestley, 1981).

As a final step, using weighted multiple regression surfaces of different - up to the third - degrees were fitted on the combinations of the Unc - MSR - Period values passing through the origin centering on $sr = 95\%$. This was achieved by subjecting the data to inverse distance weighting (w) (Eq. (8))

$$w_i = \begin{cases} 1, & sr = 95 \\ 1/|95 - sr|, & sr \neq 95 \end{cases}, \quad i = 1, 2, \dots, 1786 \text{ (see Table 1)} \quad (8)$$

The goodness of the different order fits was evaluated using Bayesian information criterion (BIC; Priestley (1981), adjusted R^2 ($adj. R^2$) and root mean square error (RMSE).

The intention was to provide an empirical approximation at a confidence level of 95%. In practical terms, any detected period that is above the value of the plane for a given Unc - MSR pair can be considered to be robustly detectable at this level of confidence (see later, Fig. 5).

All computations were done in R (R Core Team, 2019), using the `redfit()` function of the `dp1R` package (Bunn, 2008), and with the scripts written by the authors.

3. Results and discussion

The assessment of spectral ensembles indicated that with an increasing MSR (i.e. increasingly rarefied time series), frequency uncertainty also increases (Fig. 4), a result in accordance with the observations obtained in an analysis consisting of the application of

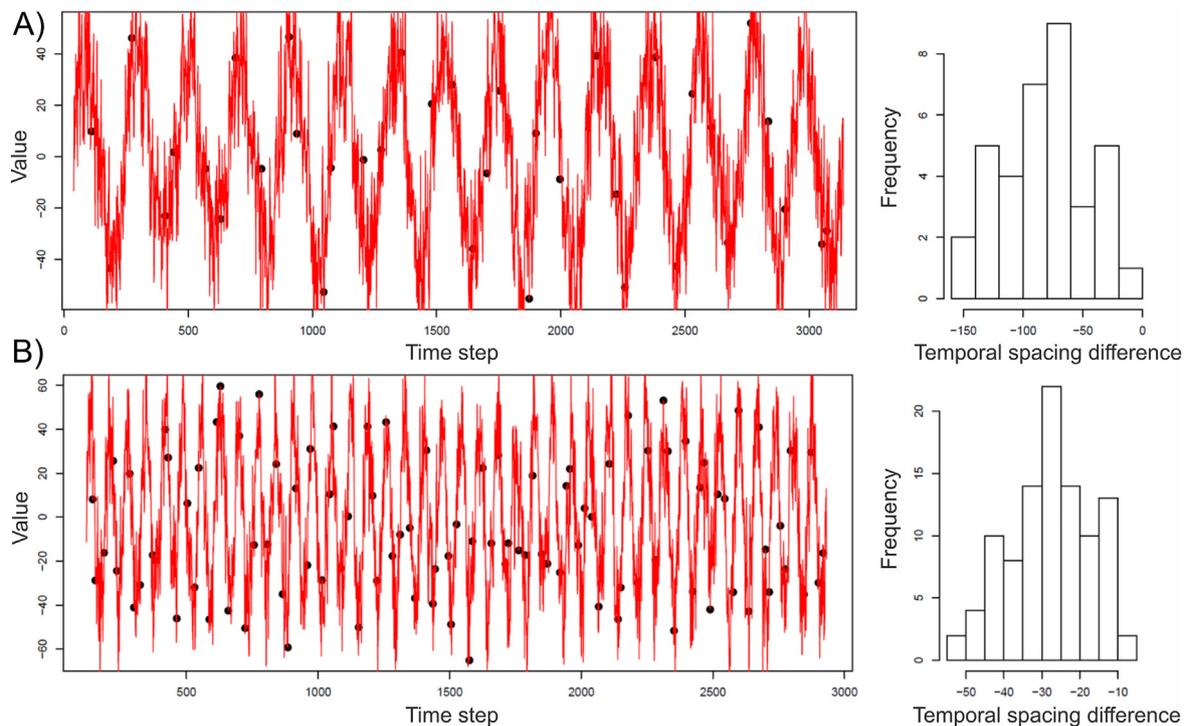


Fig. 3. Illustration of the unevenly spaced characteristic of synthetic time series. Continuous- (red hairline) and selected data points (black dots) for the white noise A) and red noise B) time series with the corresponding histograms (right panels) showing the difference in temporal spacing between the selected data points (see Part 4). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Monte Carlo simulations with predefined spectral properties and timescale error models (Mudelsee et al., 2009). Additionally, the smallest robustly detectable period increases with an increasing MSR (Fig. 4), a phenomenon analogous to the Nyquist frequency under equidistant sampling (Blackman and Tukey, 1958).

To obtain a more in-depth picture on how the robustly detectable periods vary with Unc and MSR, surfaces of different order were fitted to the sr values, and their fit statistics evaluated (Table A1) to see which approximates the 95% sr the best (see Section 2.2). BIC and $adj. R^2$ increased towards higher orders with a practically negligible magnitude between the different orders. In the meanwhile, the increase in the proportion of explained variance was continuous, with $adj. R^2 > 0.91$ for the linear model already (Table A1). Therefore, to avoid unnecessary overfitting - see A1 and Hawkins (2004) - the linear model was chosen to indicate the $sr = 95\%$ threshold (Table A1).

In the white-noise model the robustly detectable period depends ever-more steeply on MSR as Unc increases, in the red-noise model this dependence is far less pronounced (Fig. 5). One plausible cause of this is the substantial difference in the persistence of the model data. In an autocorrelated time series, rarefaction is less influential, due to the presence of persistence (Mudelsee, 2014), and this is reflected in the less steep slope of the red-noise model (Fig. 5b).

In terms of the primary research question(s), it was generally observed for both models that the implanted periods become less robustly detectable with increasing Unc (Mudelsee et al., 2009) and increasing MSR. Moreover, the same periodic peak could be detected robustly even at relatively higher Unc , when MSR was relatively low, that is, the given sedimentary proxy record was more frequently sampled. For example, in the case of $Unc = 600$ time step and $MSR = 70$ time step per sample, periods higher than 421 time steps and 430 time steps could be robustly detected in the

case of white- and red-noise models, respectively; at $Unc = 600$ step and a much better MSR of 10 time step per sample, periods could be robustly detected already from 350 steps and 375 time steps upwards, respectively (Fig. 5).

3.1. Applications

We show the actual applicability of the CUSP model (specifically the red noise one) employing environmental parameters recorded in natural proxies with various spectral- and chronological characteristics. Speleothems (Table 2) and marine sediments (Table 3) were chosen from literature on the basis of having been the subject of spectral analysis. Because sedimentary proxy records usually display strong autocorrelation, in the following the red-noise model will be used to evaluate the robustness of their pre-determined spectral characteristics.

3.1.1. Sub-Milankovitch periodicities of Holocene records

Periods on the decadal to centennial scale are frequently reported in paleoclimate studies concerning geochemical parameters, for example, those of speleothems. The key chronological characteristics of the speleothems relevant to the present work were collected from their original chronologies (Unc and MSR) and also on account of the fact that their alternative chronologies had recently been recalculated by the SISAL working group (Comas-Bru et al., 2020b). Of the recalculated chronologies, only variants using Bayesian modeling - Bacon (Blaauw and Christen, 2011) and, BChron (Haslett and Parnell, 2008) - were considered, because these mainly improve the assessment of uncertainties of age-depth models (Trachsel and Telford, 2016) and outperform classical ones by providing more realistic precision estimates, including at low to average dating densities and are much more robust against dating scatter and outliers (Blaauw et al., 2018).

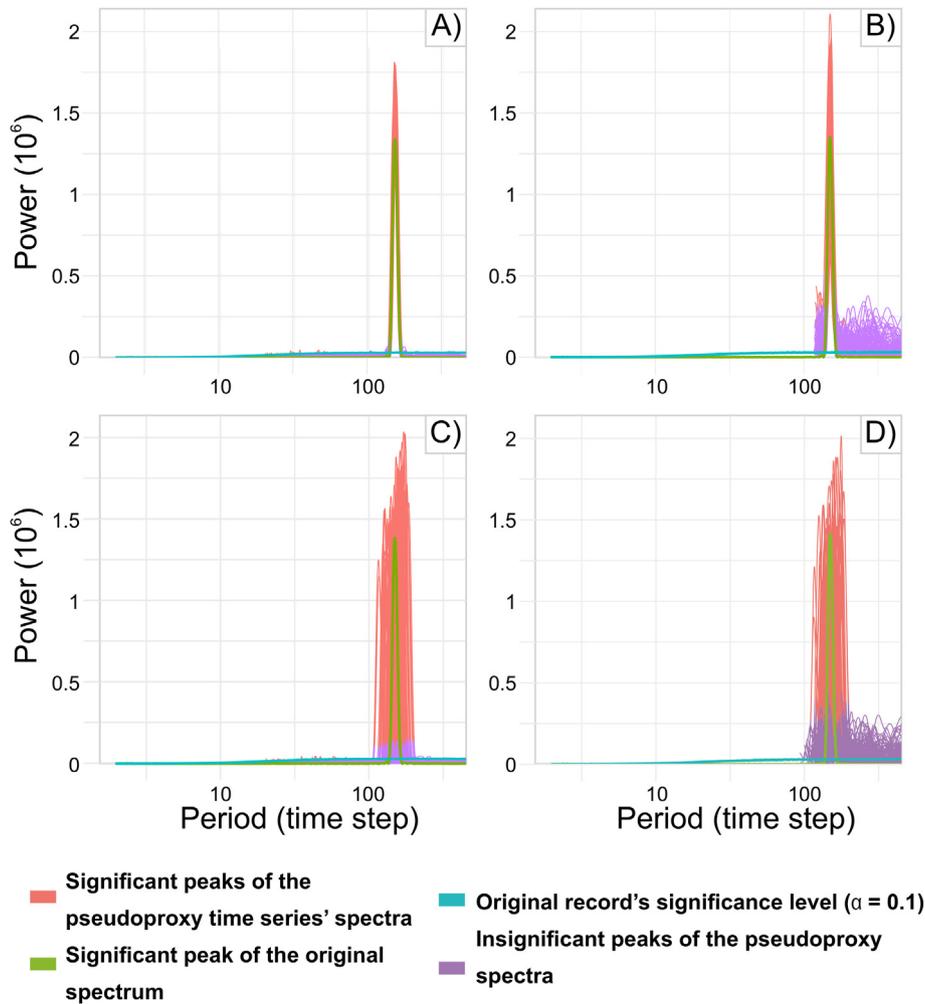


Fig. 4. Illustration of spectral bias resulting from chronological uncertainty and decreased sampling resolution. REDFIT ensemble-spectra for the time series with a 150-step implanted period and $Unc = 50$ time step, $MSR = 10$ time step per sample, 100% of the peak periods are within the 5% tolerance interval A), $Unc = 50$, $MSR = 60$, 100% within the tolerance interval B), $Unc = 500$ time step, $MSR = 10$ time step per sample, 31% in the tolerance interval C) and $Unc = 500$ time step, $MSR = 60$ time step per sample and 30% within the 5% tolerance interval D).

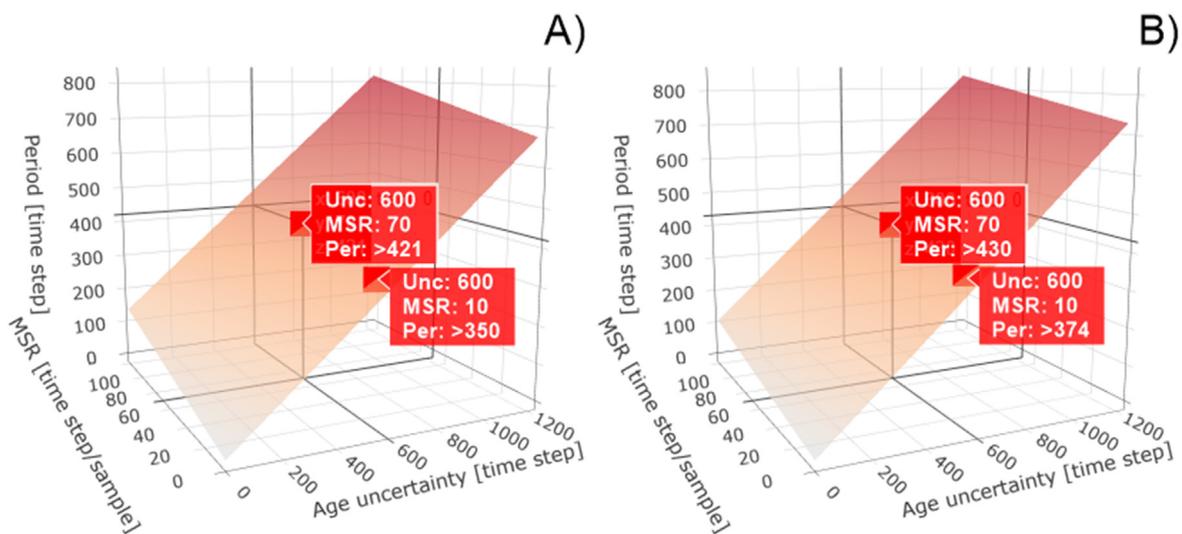


Fig. 5. Fitted linear surface indicating the critical level of robustly detectable periodic signals in the presence of chronological uncertainty and discontinuous sampling in the case of an underlying white- A) and red-noise model B). The fitted surface approximates the 95% success rate (sr). The periods above the fitted surface can be considered as robustly detectable at the given Unc and MSR values. The boxes on the surfaces indicate two pairs of examples on how the detectability increases with an increasing sampling resolution, that is, smaller MSR . (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 2

The chronological characteristics of environmental parameters recorded in speleothem $\delta^{18}\text{O}$ records. The chronological characteristics of the records were taken from the original studies and the SISAL_v2 database. The last column indicates the periods documented in the original studies. Values shown in italics indicate that the periods found can be taken as robustly present under the current chronological circumstances, while bold indicates that they cannot according to the 'red noise' CUSP model. In the case of the DAN-D record there was a difference in the degree of robustness depending on the chronological model assessed; this is indicated in the first line and in the text.

Speleothem name	Literature		Interp age		Bacon		BChron		Documented period(s) (yr)
	MSR (yr sample ⁻¹)	Unc (yrs)	MSR	Unc	MSR	Unc	MSR	Unc	
DAN-D (Berkelhammer et al., 2010)	0.5 (based on supplement)	~12	0.5	22.3	0.5	42	NA	NA	~90
Pink Panther (Asmerom et al., 2007)	17	NA	17.8	NA	17.7	379	18	368	~700, 128, 104, 86, 78, 65–68
A1 (Cosford et al., 2008)	~8	NA	8.1	Not provided	NA	NA	8.1	225	2656, 49, 30, 24, 21, 20
MGY (Yadava and Ramesh, 2007)	1.1	NA	1.1	7	NA (lamina counted)				132, 21, 2.6 (unbiased power spectrum)

Table 3

The chronological characteristics of isotopic records from marine sediments. The penultimate column shows the periods documented in the original studies and the value in bold indicates that the given period cannot be taken as robustly present under the current chronological circumstances according to the 'red noise' CUSP model.

Entity name	MSR	Unc	Documented period(s)	Timestep measured in
$\delta^{18}\text{O}$ PATCH record (RC11-120 & E49-18) (Hays et al., 1976)	3	~9	106, 43, 24, 19	kyrs
$\delta^{15}\text{N}$ Geob 7139-2 (De Pol-Holz et al., 2007)	83	1154	20000; ~220^a	yr

^a The period was documented to surpass the critical 99.89% 'false alarm' level.

In the case of the Dandak $\delta^{18}\text{O}$ record (DAN-D), a single statistically significant ($\alpha = 0.9$) peak was documented around the 90 yr cycle (Berkelhammer et al., 2010). The MSR = 0.5 yr sample⁻¹ calculated from the SISAL_v2 chronologies (Comas-Bru et al., 2020a) of the DAN-D record is in perfect agreement with that in the original study. The UNC archived along with the original chronology is higher, and the Bacon-remodeled chronology even higher (Table 2). The record has a strikingly good mean sampling resolution (MSR = 0.5 yr sample⁻¹) and a remarkably low degree of chronological uncertainty, thus the ~90 yr period was proven to be robustly detectable (Table 2). However, considering the Bacon chronology, due to the higher modeled degree of chronological uncertainty, the robustness of the same period is not confirmed (Table 2). The observation that the robustness of detectability depends on the chronological uncertainties, which vary from model to model, indicates a need for caution: a requirement not to take into account only one, but multiple realistic age-depth models and consider those spectral peaks which are not specific to the unique model approach.

In the case of the speleothem $\delta^{18}\text{O}$ record from the Pink Panther Cave, the MSR documented in the SISAL_v2 database was as good as in the original study (17 yr sample⁻¹); no Unc was, however, documented alongside this (Asmerom et al., 2007). Thus, the SISAL_v2 chronologies had to be taken into account (Table 2) which varied according to the methods applied. In the study of Asmerom et al. (2007), numerous significant ($\alpha = 0.9$) spectral peaks were recorded above the red noise background in the Pink Panther $\delta^{18}\text{O}$ record. Although the MSR of the record was good, its degree of chronological uncertainty was considerably higher (Unc >365 time step) in all age-depth models. Thus, regardless of the method applied, the periodic signals in the record below ~170 yrs cannot be taken as being robustly present in the current chronological circumstances (Table 2). In the A1 speleothem record's $\delta^{18}\text{O}$ record (Cosford et al., 2008), the MSR calculated from the SISAL_v2 database was again similar to that of the original study (~8 yr sample⁻¹); once again, however, no Unc was documented alongside this. Using the SISAL_v2 chronologies in a way similar to the Pink Panther $\delta^{18}\text{O}$ record, only the relatively long period was found to be robustly present in the A1 record (Table 2).

The MGY speleothem record's $\delta^{18}\text{O}$ record from the Akalagavi

Cave was lamina-counted and had no documented chronological uncertainty in its original study, and it was available in the SISAL_v2 database. Significant ($\alpha = 0.95$) spectral peaks were documented in the original study (Yadava and Ramesh, 2007), of which the two largest (132, and 21 yrs) proved to be robustly present, while the smallest was not (Table 2). A general conclusion can be drawn that with the exception of the lamina-counted MGY record and the notably well dated DAN-D record $\delta^{18}\text{O}$ spectrum, in the case of the other records only the relatively longer periods can be considered as robustly present in the data (Table 2).

3.1.2. Milankovitch cycles of Pleistocene records

Multimillennial periodicities are frequently reported in paleoclimate studies concerning the geochemical parameters of, for example, marine sediments. As in the case of natural proxy records, the key chronological characteristics are Unc and MSR, and it was these which were collected from the original studies in order to provide examples of the applicability of the proposed method on the orbital scale.

For instance, the periodic signals in the $\delta^{18}\text{O}$ time series of the combined PATCH record of two subantarctic deep-sea cores (RC11-120 & E49-18) (Hays et al., 1976) were all found to be robustly present since MSR and Unc were sufficiently small considering the documented periods given in kyrs. In the meanwhile, the high-resolution bulk sedimentary $\delta^{15}\text{N}$ data from the southern edge of the present-day oxygen minimum zone of the eastern South Pacific -core GeoB 7139-2 (De Pol-Holz et al., 2007)- indicated a periodic incursion of $\delta^{15}\text{N}$ in the orbital band (ca. 20 kyr), although this was insignificant at the critical 99.89% 'false alarm' level. In addition, a significant ~220 yr period was also indicated in the spectrum (Table 3).

Interestingly, with regard to the former period, the authors could not "rule out that for longer time series, the ~20 ka cycle indeed becomes significant", thus they cautiously consider its results; yet the CUSP model found it robustly present in the signal. The latter, ~220 yr period proved to be statistically significant (Fig. 6 in De Pol-Holz et al. (2007)), while it was not found to be robustly present in the chronological circumstances in the record. This supports the decision of De Pol-Holz et al. (2007) not to discuss this ~220 yr signal, which is possibly uncertain given the presence of

this timescale error.

It should be noted that the “negative” examples presented in Section 3.1 do not mean that these documented spectral peaks are not statistically present in the data at a given confidence interval; neither do they mean that such spectral characteristics could not have been recorded by the original records. It only implies that at a given sampling resolution (considered “insufficient”) and degree of chronological uncertainty, the chance of spectral peak(s) flagged as not robust appearing in the spectrum being spurious is 5%.

4. Outlook

In the future version(s) of CUSP the following developments are planned that could further widen its applicability primarily by varying key parameters such as: signal to noise ratio, AR(1) noise, AR(1) parameter value, power-law exponents, age-depth models, length of a series, etc.; and by applying multi-period signals. Thus, CUSP should be then useful for many paleoclimatic communities and those researchers working on instrumental and future (modeled) data.

Further testing possibilities: establishing a proxy time series from measured proxies and dating points using adequate age-depth modeling software. Determining the memory and strength of the white/red noise of this proxy time series (Mudelsee, 2002) and use the derived parameters to construct sets of artificial time series with different periods. This would provide estimates about the robustness of periods using the same temporal characteristics as the proxy time series under investigation.

5. Conclusions

In the present work an online tool has been introduced together with an underlying model that helps researchers in the rapid assessment of the robustness of spectral peaks against timescale errors on the basis of documented mean sampling resolution and degree of chronological uncertainty in the records.

Via modeling and real-world application, chronological uncertainty was found to be the most limiting factor, and that therefore higher sampling resolution can ameliorate the detectability of periodic signals even in the presence of timescale error. An increase in mean sampling resolution was shown to be less influential in an autocorrelated time series, since more information is retained, due to the phenomenon of persistence. As the study demonstrates, if underlying raw data are not provided, but the chronological characteristics – such as mean sampling resolution and chronological uncertainty – of a given sedimentary record are available, the robustness of a documented periodic signal can be determined with the use of the tool provided (see Sect. 6). Unfortunately, even today, these characteristics are not always documented in the respective studies.

6. Software availability short description

The CUSP online tool for determining the robust presence of a spectral peak present at a given degree of chronological uncertainty (Unc) and mean sampling resolution (MSR) is available here: <http://geochem.hu/CUSP/>. The tool has been tested to work in all major browsers.

In employing CUSP, a chronological uncertainty and mean sampling resolution pair should be provided by the user, along with a period that is planned to be evaluated. CUSP in turn, will yield (i) the value of the threshold period which can still be robustly detected and (ii) a suggestion as to whether the investigated period can robustly be detected under the given chronological circumstances or not. Users can choose between the white and red noise

models, with red noise as the default, to better reflect the persistence characteristics of natural records (Mudelsee, 2014; von Storch et al., 2009).

As stated in Section 2.1, the present model results are universally applicable to any temporal measurement unit, i.e. the measurement unit of the time steps can be chosen to be for example, yr, or kyr, as long as it is used uniformly in the case of all three parameters. If “time step” is taken as “year” the presented model is applicable to the exploration of the robustness of periodic signals on decadal to centennial scales. However, if “time step” is taken as “kyrs” then millennial (e.g. orbital) cycles can be investigated.

If the period sought for is $< MSR \times 2$, then due to the Nyquist cutoff frequency it cannot be evaluated (Press et al., 1996), thus the following output is given by CUSP: “The given sampling rate is probably insufficient to carry the period sought for”.

Author statement

ZK conceived the study idea, all authors took part in developing the methodology used, IGH formulated the model code with help from PT. IGH performed the simulations with input from all authors. The first draft of the manuscript was written by IGH and ZK with input from MM and PT. All authors took part in the revision of the manuscript. The FLAE approach was employed in the sequence of authors; see 10.1371/journal.pbio.0050018.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

On behalf of project 'Limitations of spectral analysis of sedimentary proxy records, tested on simulated time series with timescale error and variable temporal resolution' we thank for the usage of ELKH Cloud (<https://science-cloud.hu/>) that significantly helped us achieving the results published in this paper. The work of PT was supported by the New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund, Hungary under grant No. ÚNKP-20-4-II-SZIE-27, while IGH was supported by the Bolyai János Scholarship of the Hungarian Academy of Sciences and the Ministry of Human Capacities under grant No. NTP-NFTÖ-21-B-0049. This is No. 74 of the 2ka Palaeoclimatology Research Group. The English version was proofread by Paul Thatcher, which authors are thankful for. Authors are also grateful for the detailed, very constructive and encouraging comments of three independent Reviewers.

Appendix A

Table A1

Statistics of the surfaces fitted to the $sr = 95\%$ values in the Unc-MSR-Period space. Chosen model indicated with italics. *N*: number of points, *BIC*: Bayesian information criterion, *adj. R²*: adjusted R², *RMSE*: residual mean square error.

	Red noise			White noise		
Degree	1	2	3	1	2	3
<i>N</i>	1786					
<i>BIC</i>	22659	22373	22211	22599	22423	22238
<i>adj. R²</i>	0.935	0.945	0.95	0.911	0.92	0.929
<i>RMSE</i>	133.4	127.8	123.2	124.2	124.2	124.8

When choosing the linear model over the higher degree ones, it was taken into account that:

- (i) a linear model seems already sufficiently describing the explored process, $adj. R^2 > 0.9$ and the difference between the BIC values for the different degrees is ~1%. In the error there was practically no difference in the case of the white noise model, while in the case of red noise it decreased with less than 5% (Table A1)
- (ii) there is no evidence that to characterize the explored relationship, the complexity of a higher degree surface would be required, on the contrary it may introduce nonexistent complexity
- (iii) if a higher degree surface is used than necessary, it would violate parsimony, and overfitting has to be avoided in general (Hawkins, 2004).

Table A2

Summary table of coefficients of the linear multiple regression. S.E.: standard error. Since both white- and red noise models were forced to pass through the origin, their intercept equals zero.

	Coefficient	Estimate	S.E.	t val.	p
White noise	polym(Unc, MSR)1.0	0.564	0.010	56.64	$< 2 \times 10^{-16}$
	polym(Unc, MSR)0.1	1.182	0.049	24.2	
Red noise	polym(Unc, MSR)1.0	0.609	0.009	69.33	
	polym(Unc, MSR)0.1	0.924	0.044	20.89	

References

Amrhein, D.E., 2020. How large are temporal representativeness errors in paleoclimatology? *Clim. Past* 16, 325–340.

Asmerom, Y., Polyak, V., Burns, S., Rasmussen, J., 2007. Solar forcing of Holocene climate: new insights from a speleothem record, southwestern United States. *Geology* 35, 1–4.

Berkelhammer, M., Sinha, A., Mudelsee, M., Cheng, H., Edwards, R.L., Cannariato, K., 2010. Persistent multidecadal power of the Indian summer monsoon. *Earth Planet Sci. Lett.* 290, 166–172.

Blaauw, M., Christen, J.A., 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Anal.* 6, 457–474.

Blaauw, M., Christen, J.A., Bennett, K.D., Reimer, P.J., 2018. Double the dates and go for Bayes — impacts of model choice, dating density and quality on chronologies. *Quat. Sci. Rev.* 188, 58–66.

Blackman, R.B., Tukey, J.W., 1958. The measurement of power spectra from the point of view of communications engineering — Part I. *Bell Syst. Tech. J.* 37, 185–282.

Box, G.E.P., 1953. Non-normality and tests on variances. *Biometrika* 40, 318–335.

Bradley, R.S., 1999. *Paleoclimatology: Reconstructing Climates of the Quaternary*. Elsevier.

Breitenbach, S.F.M., Rehfeld, K., Goswami, B., Baldini, J.U.L., Ridley, H.E., Kennett, D.J., Pruffer, K.M., Aquino, V.V., Asmerom, Y., Polyak, V.J., Cheng, H., Kurths, J., Marwan, N., 2012. Constructing proxy records from age models (COPRA). *Clim. Past* 8, 1765–1779.

Bunn, A.G., 2008. A dendrochronology program library in R (dplR). *Dendrochronologia* 26, 115–124.

Comas-Bru, L., Atsawaranunt, K., Harrison, S., Members, S.W.G., 2020a. SISAL (Speleothem Isotopes Synthesis and Analysis Working Group) Database Version 2.0, 2.0. University of Reading.

Comas-Bru, L., Rehfeld, K., Roesch, C., Amirnezhad-Mozhdehi, S., Harrison, S.P., Atsawaranunt, K., Ahmad, S.M., Ait Brahimi, Y., Baker, A., Bosomworth, M., Breitenbach, S.F.M., Burstyn, Y., Columbu, A., Deininger, M., Demény, A., Dixon, B., Fohlmeister, J., Hatvani, I.G., Hu, J., Kaushal, N., Kern, Z., Labuhn, I., Lechleitner, F.A., Lorrey, A., Martrat, B., Novello, V.F., Oster, J., Pérez-Mejías, C., Scholz, D., Scroxton, N., Sinha, N., Ward, B.M., Warken, S., Zhang, H., 2020b. SISALv2: a comprehensive speleothem isotope database with multiple age-depth models. *Earth Syst. Sci. Data* 12, 2579–2606.

Cosford, J., Qing, H., Eglinton, B., Matthey, D., Yuan, D., Zhang, M., Cheng, H., 2008. East Asian monsoon variability since the Mid-Holocene recorded in a high-resolution, absolute-dated aragonite speleothem from eastern China. *Earth*

Planet Sci. Lett. 275, 296–307.

De Pol-Holz, R., Ulloa, O., Lamy, F., Desileau, L., Sabatier, P., Hebbeln, D., 2007. Late Quaternary variability of sedimentary nitrogen isotopes in the eastern South Pacific Ocean. *Paleoceanography* 22, PA2207.

Franke, J.G., Donner, R.V., 2019. Correlating paleoclimate time series: sources of uncertainty and potential pitfalls. *Quat. Sci. Rev.* 212, 69–79.

Haam, E., Huybers, P., 2010. A test for the presence of covariance between time-uncertain series of data with application to the Dongge Cave speleothem and atmospheric radiocarbon records. *Paleoceanography* 25, PA2209.

Haslett, J., Parnell, A., 2008. A simple monotone process with application to radiocarbon-dated depth chronologies. *J. Roy. Stat. Soc.: Ser. C Appl. Stat.* 57, 399–418.

Hatvani, I.G., Kern, Z., Leél-Össy, S., Demény, A., 2018. Speleothem stable isotope records for east-central Europe: resampling sedimentary proxy records to obtain evenly spaced time series with spectral guidance. *Earth Syst. Sci. Data* 10, 139–149.

Hawkins, D.M., 2004. The problem of overfitting. *J. Chem. Inf. Comput. Sci.* 44, 1–12.

Hays, J.D., Imbrie, J., Shackleton, N.J., 1976. Variations in the earth's orbit: pacemaker of the ice ages. *Science* 194, 1121–1132.

Hercman, H., Pawlak, J., 2012. MOD-AGE: an age-depth model construction algorithm. *Quat. Geochronol.* 12, 1–10.

Holzschläger, S., Mangini, A., Spötl, C., Mudelsee, M., 2004. Timing and progression of the Last Interglacial derived from a high alpine stalagmite. *Geophys. Res. Lett.* 31, L07201.

Huybers, P., Curry, W., 2006. Links between annual, Milankovitch and continuum temperature variability. *Nature* 441, 329–332.

Kunz, T., Dolman, A.M., Laepple, T., 2020. A spectral approach to estimating the timescale-dependent uncertainty of paleoclimate records — Part 1: theoretical concept. *Clim. Past* 16, 1469–1492.

Li, M., Hinnov, L., Kump, L., 2019. Acycle: time-series analysis software for paleoclimate research and education. *Comput. Geosci.* 127, 12–22.

McKay, N.P., Emile-Geay, J., Khider, D., 2021. geoChronR — an R package to model, analyze, and visualize age-uncertain data. *Geochronology* 2021, 149–169.

Meyers, S.R., Sageman, B.B., Pagani, M., 2008. Resolving Milankovitch: consideration of signal and noise. *Am. J. Sci.* 308, 770–786.

Mudelsee, M., 2001. The phase relations among atmospheric CO₂ content, temperature and global ice volume over the past 420 ka. *Quat. Sci. Rev.* 20, 583–589.

Mudelsee, M., 2002. TAUEST: a computer program for estimating persistence in unevenly spaced weather/climate time series. *Comput. Geosci.* 28, 69–72.

Mudelsee, M., 2014. *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*, 2nd ed. Springer, Cham, Switzerland.

Mudelsee, M., Scholz, D., Röthlisberger, R., Fleitmann, D., Mangini, A., Wolff, E.W., 2009. Climate spectrum estimation in the presence of timescale errors. *Nonlinear Process Geophys.* 16, 43–56.

Neff, U., Burns, S.J., Mangini, A., Mudelsee, M., Fleitmann, D., Matter, A., 2001. Strong coherence between solar variability and the monsoon in Oman between 9 and 6 kyr ago. *Nature* 411, 290–293.

Pelletier, J.D., 1998. The power spectral density of atmospheric temperature from time scales of 10–2 to 106 yr. *Earth Planet Sci. Lett.* 158, 157–164.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1996. *Numerical Recipes in C*. Cambridge university press, Cambridge.

Priestley, M.B., 1981. *Spectral Analysis and Time Series: Probability and Mathematical Statistics*.

R Core Team, 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schulz, M., Mudelsee, M., 2002. REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. *Comput. Geosci.* 28, 421–426.

Sorooshian, S., Martinson, D.G., NRC, Committee, C.R., 1995. *Proxy indicators of climate*. In: *Natural Climate Variability on Decade-To-Century Time Scales*. National Academy Press, Washington, D.C., USA.

Trachsel, M., Telford, R.J., 2016. All age–depth models are wrong, but are getting better. *Holocene* 27, 860–869.

Trauth, M.H., 2021. Spectral analysis in Quaternary sciences. *Quat. Sci. Rev.* 270, 107157.

von Storch, H., Zorita, E., González-Rouco, F., 2009. Assessment of three temperature reconstruction methods in the virtual reality of a climate simulation. *Int. J. Earth Syst. Sci.* 98, 67–82.

Vyushin, D.I., Kushner, P.J., 2009. Power-law and long-memory characteristics of the atmospheric general circulation. *J. Clim.* 22, 2890–2904.

Walker, M., 2005. *Quaternary Dating Methods*. John Wiley and Sons.

Wefer, G., Berger, W.H., Bijma, J., Fischer, G., 1999. Clues to ocean history: a brief overview of proxies. In: Fischer, G., Wefer, G. (Eds.), *Use of Proxies in Paleoceanography: Examples from the South Atlantic*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–68.

Yadava, M.G., Ramesh, R., 2007. Significant longer-term periodicities in the proxy record of the Indian monsoon rainfall. *N. Astron.* 12, 544–555.

Yiou, P., Baert, E., Loutre, M.F., 1996. Spectral analysis of climate data. *Surv. Geophys.* 17, 619–663.

Zeeden, C., Kaboth, S., Hilgen, F.J., Laskar, J., 2018. Taner filter settings and automatic correlation optimisation for cyclostratigraphic studies. *Comput. Geosci.* 119, 18–28.