Nonlin. Processes Geophys., 25, 175–200, 2018
https://doi.org/10.5194/npg-25-175-2018

Nonlinear Processes in Geophysics
Open Access
EGU

A general theory on frequency and time–frequency analysis of irregularly sampled time series based on projection methods – Part 2: Extension to time–frequency analysis

Guillaume Lenoir[1] **and Michel Crucifix**[1,2]

[1]Georges Lemaître Centre for Earth and Climate Research, Earth and Life Institute,
Université catholique de Louvain, 1348, Louvain-la-Neuve, Belgium
[2]Belgian National Fund of Scientific Research, rue d'Egmont, 5, 1000 Brussels, Belgium

Correspondence: Guillaume Lenoir (guillaume.lenoir@hotmail.com)

Received: 16 June 2017 – Discussion started: 4 July 2017
Revised: 4 December 2017 – Accepted: 6 December 2017 – Published: 5 March 2018

Abstract. Geophysical time series are sometimes sampled irregularly along the time axis. The situation is particularly frequent in palaeoclimatology. Yet, there is so far no general framework for handling the continuous wavelet transform when the time sampling is irregular.

Here we provide such a framework. To this end, we define the scalogram as the continuous-wavelet-transform equivalent of the extended Lomb–Scargle periodogram defined in Part 1 of this study (Lenoir and Crucifix, 2018). The signal being analysed is modelled as the sum of a locally periodic component in the time–frequency plane, a polynomial trend, and a background noise. The mother wavelet adopted here is the Morlet wavelet classically used in geophysical applications. The background noise model is a stationary Gaussian continuous autoregressive-moving-average (CARMA) process, which is more general than the traditional Gaussian white and red noise processes. The scalogram is smoothed by averaging over neighbouring times in order to reduce its variance. The Shannon–Nyquist exclusion zone is however defined as the area corrupted by local aliasing issues. The local amplitude in the time–frequency plane is then estimated with least-squares methods. We also derive an approximate formula linking the squared amplitude and the scalogram. Based on this property, we define a new analysis tool: the weighted smoothed scalogram, which we recommend for most analyses. The estimated signal amplitude also gives access to band and ridge filtering. Finally, we design a test of significance for the weighted smoothed scalogram against the stationary Gaussian CARMA background noise, and provide algorithms for computing confidence levels, either analytically or with Monte Carlo Markov chain methods. All the analysis tools presented in this article are available to the reader in the Python package WAVEPAL.

1 Introduction

The continuous wavelet transform (CWT) is widely used for the time–frequency analysis of geophysical time series, mainly through its scalogram, which is the squared modulus of the CWT. The CWT relies on a probing function, called the mother wavelet. A common choice for the mother wavelet is the Morlet wavelet (Grossmann and Morlet, 1984), which is well suited for the analysis of signals whose components have a time-varying frequency and/or a time-varying amplitude. The scalogram is then often smoothed to reduce its variance, and significance testing against a stationary Gaussian white or red noise is commonly applied. State-of-the-art references in climate for the analysis of regularly sampled time series include Torrence and Compo (1998), who provide the basis for the subsequent works; Torrence and Webster (1999), who define a smoothing method for the scalogram (which is a particular case of the wavelet coherency developed in there); and Maraun and Kurths (2004) and Maraun et al. (2007), who build more reliable significance tests for the smoothed scalogram. A non-exhaustive list of the applications, in climatology, of the scalogram of the CWT with the Morlet wavelet, includes the following.

- Studies in climate and weather: analysis of the El Niño Southern Oscillation in Torrence and Compo (1998), analysis of the Arctic oscillation in Grinsted et al. (2004), or the analysis of daily precipitation in the Alps in Schaefli et al. (2007).

- Studies in palaeoclimate: analysis of the astronomical forcing in Berger et al. (1998), analysis of the mid-Pleistocene transition in Elderfield et al. (2012), or the analysis of the equatorial Pacific thermocline over the last eight glacial periods in Regoli et al. (2015).

Most of these studies use the algorithms provided by the papers cited above or similar algorithms, and all of them require the data to be regularly spaced. However, it may happen that the time series be intrinsically irregularly sampled (this actually happens in some of the above examples) and it is then interpolated on a regularly spaced grid in order to apply the algorithms of the CWT and its scalogram. But the interpolation procedure may seriously affect the analysis with unpredictable consequences for the scientific interpretation, especially when performing significance testing. This is illustrated in Appendix F.

A solution to this problem was addressed by Foster in a series of articles which share a common thread with our two papers, in the sense that it first generalises the Lomb–Scargle periodogram, based on orthogonal projection methods, in Foster (1996a) and Foster (1996b), and then extends the formalism to the continuous wavelet transform, in Foster (1996c), allowing the time series not to be interpolated. Foster's formulas were motivated by the astronomical study of the light curves of variable stars, which are unevenly sampled time series with large gaps. The methods presented in this article are influenced by Foster's theory and it is shown in Appendix E2 that most of its formulas can actually be deduced from our general framework. The main limitations of Foster's theory are the following (see Appendix E2 for detailed explanations): significance testing is only performed for the white noise background case, it only deals with the unsmoothed scalogram, and the areas in the time–frequency plane corrupted by aliasing are underestimated. It therefore suffers from a limited interest in geophysical applications[1]. Another study tackling the problem of the estimation of the scalogram of irregularly sampled time series, without interpolating the data, is given in Mathias et al. (2004). In there, the authors propose an estimator of the amplitude of the signal, locally in the time–frequency plane, but no significance testing is performed. The other limitations of their algorithm are basically the same as for Foster's formulas. This is detailed in Appendix E3.

In this article, we extend the analysis tools that we derived in the first part of this study (Lenoir and Crucifix, 2018) in the case of the frequency analysis of irregularly sampled time series. They are based on a similar model, which is a locally periodic component in the time–frequency plane, plus a polynomial trend, plus a stationary Gaussian continuous autoregressive-moving-average (CARMA) process. Let us sketch the main points of the article. First, the taper of the periodogram, derived in Lenoir and Crucifix (2018, Sect. 4.4), is chosen here to be a time-dependent Gaussian function with a variance depending on the scale in order to define the Morlet wavelet-based scalogram. This is detailed in Sect. 3.2 of this work. Second, the scalogram is smoothed in order to reduce its variance, by averaging over neighbouring times. To this end, we apply the same formula as in Cohen and Walden (2010). This is explained in Sect. 3.4. Third, in Sect. 3.5, we estimate the amplitude of the locally periodic component, extending the results obtained in Sect. 6 of Part 1, and define, in Sect. 3.6 of this article, the weighted smoothed scalogram as the time–frequency analogue of the weighted WOSA periodogram defined in the first part of this study (Lenoir and Crucifix, 2018). Fourth, we define in Sect. 3.8 the Shannon–Nyquist exclusion zone (SNEZ) as the area of the time–frequency plane which must be excluded from the analysis because of the local aliasing issues. Fifth, we design a test of significance for the weighted smoothed scalogram, against the stationary Gaussian CARMA background noise. This is based on the theory developed in Sect. 5 of Part 1. More specifically, we define a null and an alternative hypothesis, and estimate the distribution of the weighted smoothed scalogram under the null hypothesis, either analytically, conserving the first moments of the distribution, or with Markov Chain Monte Carlo (MCMC) methods. The latter approach allows the uncertainty in the parameters of the CARMA background process to be fully taken into account. This is presented in Sect. 4. Sixth, we provide, in Sect. 5, formulas for filtering the signal in a band delimited by two scales, or with the ridges, which are the lines going through the maxima of the estimated amplitude, in the time–frequency plane. Ridge filtering is based on state-of-the-art algorithms provided in Lilly and Olhede (2010) and (https://github.com/jonathanlilly/jLab). Seventh, we define in Sect. 6 the global scalogram as the time-averaged weighted smoothed scalogram, resulting in a periodogram-like analysis tool with a frequency-varying bandwidth. Eighth, we illustrate, in Sect. 7, the theory on the same palaeoclimate data set as in our first article (Lenoir and Crucifix, 2018). Finally, a Python package named WAVEPAL is available to the reader and is presented in Sect. 8. Before tackling the problem of irregularly sampled time series, the paper starts with the theory of the CWT applied to continuous-time signals. This gives the bases for the subsequent developments.

Most of the mathematical concepts and notations are introduced in the first part of this study (Lenoir and Crucifix, 2018), and the reader is invited to revise them. Throughout this article, we will denote the equations of the preceding paper by, for example, "Eq. (I,30)", meaning "the equation (30) of Part 1", and will refer to the paper itself by "Part 1".

[1] An application of Foster's formulas on palaeoclimate data is found in Witt and Schumann (2005).

2 The continuous wavelet transform of continuous-time processes

2.1 The continuous wavelet transform and its scalogram

The mathematical background to Fourier analysis is given in Appendix A. Let $\mathcal{S}$ denote the Schwartz space. The continuous wavelet transform of a function $x \in \mathcal{S}(\mathbb{R})$ is

$$S_x(\tau, a) = \langle \psi_{\tau,a} | x \rangle, \tag{1}$$

where $\psi_{\tau,a} \in \mathcal{S}(\mathbb{R})$ is defined by

$$\psi_{\tau,a}(t) = c(a)\psi\left(\frac{t-\tau}{a}\right). \tag{2}$$

Here, ψ is called the *mother wavelet*, $\tau \in \mathbb{R}$ is the *translation time*, $a \in \mathbb{R}_0^+$ is the *scale*, and $c(a) \sim a^m$ with $m \in \mathbb{Q}$. We can write the CWT as a convolution product,

$$S_x(\tau, a) = \left(\psi_a^\sharp \star x\right)(\tau), \tag{3}$$

where

$$\psi_a^\sharp(t) = c(a)\overline{\psi\left(\frac{-t}{a}\right)}, \tag{4}$$

in which $\bar{\cdot}$ denotes the complex conjugate. From the convolution theorem,

$$\widehat{S_x}(\omega, a) = \sqrt{2\pi}\,\widehat{\psi_a^\sharp}(\omega)\widehat{x}(\omega) = \sqrt{2\pi}\,a\,c(a)\overline{\widehat{\psi}(a\omega)}\widehat{x}(\omega), \tag{5}$$

and $S_x(\tau, a)$ is then obtained by taking the inverse Fourier transform.

$|S_x(\tau, a)|^2$ gives the local power in the time–scale plane, and is called the *scalogram* by analogy with the *periodogram*.

2.2 The wavelet power spectrum

The wavelet power spectrum (WPS) of a continuous-time stochastic process $\{x(t)\}_{t\in\mathbb{R}}$ is defined by the following (see Li and Oh, 2002; Maraun and Kurths, 2004):

$$\mathrm{WPS}_x(\tau, a) = E\{|S_x(\tau, a)|^2\}, \tag{6}$$

where the expectation is taken over the samples of the stochastic process. A simple example is the WPS of a real-valued stationary white noise. Define $\{\eta(t)\}_{t\in\mathbb{R}}$ satisfying the following covariance property:

$$E\{\eta(t)\eta(t')\} = \sigma^2\delta(t-t'). \tag{7}$$

Its WPS is then

$$\mathrm{WPS}_\eta(\tau, a) = a\,c(a)^2||\psi||^2\sigma^2. \tag{8}$$

2.3 The Morlet wavelet as the mother wavelet

In this article, we choose the mother wavelet ψ to be the Morlet wavelet (Grossmann and Morlet, 1984):

$$\psi(t) = \tag{9}$$
$$\pi^{-1/4}\sigma_0^{-1/2}\left[\exp(i\omega_0 t) - \exp\left(-\omega_0^2\sigma_0^2/2\right)\right]\exp\left(-t^2/2\sigma_0^2\right).$$

This mother wavelet is a complex plane wave weighted by a Gaussian, to which is added a correction term to make it *admissible*[2], i.e. satisfying $\int_{-\infty}^{+\infty}\mathrm{d}\omega|\widehat{\psi}(\omega)|^2|\omega|^{-1} < \infty$. This correction term is negligible[3] for $\sigma_0\omega_0 \geq 5.5$. If this inequality is satisfied, and with the variable change $a' = a/\omega_0$, the CWT with the Morlet wavelet is

$$S(\tau, a') = \tag{10}$$
$$c(a')\int_{-\infty}^{+\infty}\mathrm{d}t\exp\left(-\frac{i(t-\tau)}{a'}\right)\exp\left(-\frac{(t-\tau)^2}{2\sigma_0^2\omega_0^2a'^2}\right)x(t),$$

where $c(a') \sim (a')^m$, $m \in \mathbb{Q}$, and $c(a')$ holds all the multiplicative constants. Without loss of generality, we impose $\sigma_0 = 1$ and assume that

$$\omega_0 \geq 5.5 \tag{11}$$

is fulfilled in the following of this article. Therefore,

$$S(\tau, a) = c(a)\int_{-\infty}^{+\infty}\mathrm{d}t\exp\left(-\frac{i(t-\tau)}{a}\right)\exp\left(-\frac{(t-\tau)^2}{2\omega_0^2a^2}\right)x(t)$$
$$= (\psi_a^\sharp \star x)(\tau), \tag{12}$$

where

$$\psi_a^\sharp(t) = c(a)\exp(it/a)\exp\left(-t^2/2\omega_0^2a^2\right). \tag{13}$$

Under this form, interpreting Eq. (12) is straightforward: the CWT is the inner product between the signal x and a Gaussian wave packet centred in $\tau = t$, of period $2\pi a$, and with numerical support[4] of length $6\omega_0 a$. As the scale increases (decreases), the support becomes wider (narrower).

2.4 On the parameter $c(a)$

There are two common choices for $c(a)$ (see Maraun and Kurths, 2004, Sect. 3). The first one is $c(a)$ proportional to $1/\sqrt{a}$,

$$c(a) \sim \frac{1}{\sqrt{a}}, \tag{14}$$

[2]The admissibility criteria is required for ψ to be a wavelet (Holschneider, 1995, p. 5).

[3]We have $\exp(-(5.5)^2) = 7.288 \times 10^{-14}$ and $\exp(-(5.5)^2/2) = 2.700 \times 10^{-7}$.

[4]The length of the support of the Gaussian may be approximated by 6 times its standard deviation.

and gives a constant $\mathbf{L}^2$ norm for $\psi_{\tau,a}$, namely $||\psi_{\tau,a}|| = ||\psi||$. This implies that the wavelet power spectrum of a white noise is flat, as we can see in Eq. (8). The second choice is

$$c(a) \sim \frac{1}{a}, \tag{15}$$

which gives a constant $\mathbf{L}^1$ norm for $\psi_{\tau,a}$ and, most importantly, gives the same maximal power for sines of the same amplitude and with different frequencies. Indeed, from the Fourier transform of $\psi_a^\sharp$,

$$\widehat{\psi_a^\sharp}(\omega) = c(a) a \omega_0 \exp\left(-\frac{\omega_0^2(\omega a - 1)^2}{2}\right), \tag{16}$$

and applying Eq. (5), we must require $c(a)a$ to be constant to have the maxima of the scalogram of a sum sine waves (all with the same amplitude but with different frequencies) invariant with the scale.

2.5 The parameter ω_0 and the time–frequency resolution

The parameter ω_0 controls the time–frequency resolution, as it can be seen from the standard deviations of the Gaussian weights in $\psi_a^\sharp$, Eq. (13), and in its Fourier transform, Eq. (16). The standard deviations are equal to $\omega_0 a$ and $1/\omega_0 a$ respectively. Consequently, for a fixed scale, increasing (decreasing) the value of ω_0 will generate a CWT with a better (worse) frequency resolution and a worse (better) time resolution. This property is of primary importance for the applications to time series, as illustrated in Sect. 7. Note that, for any time–frequency transform, there is always a trade-off between time and frequency localisation according to the Fourier uncertainty principle. The Morlet wavelet exhibits the best trade-off, thanks to its Gaussian shape. We provide further details on this topic in Appendix B.

2.6 Scale-to-period conversion

The Morlet wavelet is often used to detect the periodicities in a signal, and it is therefore suitable to convert scales a into periods T (Meyers et al., 1993). In practice, take a signal $x(t) = A\exp(i\omega t) = A\exp(i2\pi t/T)$. Its scalogram writes

$$|S(\tau, a)|^2 = 2\pi A c(a)^2 \omega_0^2 a^2 \exp\left(-\omega_0^2(\omega a - 1)^2\right), \tag{17}$$

and is independent of τ. Scale-to-period conversion is performed with the value of the scale for which $|S(\tau, a)|^2$ is at its maximum (as a function of a). We find the following:

$$T = \begin{cases} 2\pi a & \text{if } c(a) \sim 1/a, \\ \dfrac{4\pi\omega_0 a}{\omega_0 + \sqrt{\omega_0^2 + 2}} & \text{if } c(a) \sim 1/\sqrt{a}. \end{cases} \tag{18}$$

For a fixed scale, and while $\omega_0 \geq 5.5$, the difference between both never exceeds 2 %.

2.7 Reconstruction with the amplitude ridges

Reconstruction of a signal can be performed with the CWT along the *amplitude ridges*[5] (Carmona et al., 1997), which are the lines going through the maxima of the scalogram. Indeed, take the signal $x(t) = A\exp(i\omega t)$ and $c(a) \sim 1/a$. Its scalogram is maximum at $a = 1/\omega$ (from Eq. 17) and we can therefore easily recover the amplitude A at each time τ, going through the ridge $a(\tau) = 1/\omega$ in the scalogram, on which we have $|S(\tau, 1/\omega)| = \alpha A \forall \tau$, where $\alpha \in \mathbb{R}$ is a multiplicative constant. Jointly with the amplitude, the full signal $x(t)$ can be exactly recovered from the CWT along the ridge.

This can be extended to signals with slowly varying amplitude and phase (see Delprat et al., 1992; Carmona et al., 1997), namely,

$$x(t) = A(t)\exp(i\phi(t)), \text{ such that } \left|\frac{\mathrm{d}\phi}{\mathrm{d}t}\right| \gg \left|\frac{1}{A}\frac{\mathrm{d}A}{\mathrm{d}t}\right|, \tag{19}$$

for which the CWT taken along the ridge, i.e. at the maxima of its modulus, can approximately reconstruct $x(t)$. The inequality in Eq. (19), called the *asymptoticity* condition, means that the instantaneous frequency inside the wave packet must be much smaller than the frequency of the amplitude of the wave packet. The analysis can be further extended to a sum of asymptotic signals plus noise, which can be detected by multiple ridges (Carmona et al., 1999). When considering a real signal like $x(t) = A(t)\cos(\phi(t))$, we have to work with its *analytic* counterpart, which is built from the Fourier transform of x, $\widehat{x}$, for which we impose $\widehat{x}(\omega < 0) = 0$ and then take the inverse Fourier transform. Analyticity ensures that the phase and amplitude of a signal are uniquely determined (see Lilly and Olhede, 2010, and the references therein for more details). State-of-the-art algorithms for ridge detection are developed in Lilly and Olhede (2010) and are available for use in the package jLab (https://github.com/jonathanlilly/jLab), in which the ridge-finding algorithm is general enough to be applied to various mother wavelets, such as the Morlet wavelet.

By construction, ridge filtering is well-adapted for filtering a multi-periodic signal, even if it is plunged in a noisy environment (Lilly and Olhede, 2010). In such conditions, it outperforms the techniques based on the Hilbert transform. As mentioned in Lilly and Olhede (2010, p. 4135), "[...] the Hilbert transform can lead to disastrous results as the amplitude and phase will then reflect the aggregate properties of the multi-component signal".

[5] Another type of ridge is the *phase ridge*, defined in Delprat et al. (1992), but we consider only the amplitude ridges in this study since they are easier to generalise to irregularly time series. A comparison of both the amplitude and phase ridges is found in Lilly and Olhede (2010).

2.8 Writing the scalogram under the formalism of orthogonal projections

Finally, we mention that the scalogram can be written under the formalism of orthonormal projections. Indeed, defining

$$y_{\tau,a}(t) = \frac{\pi^{-1/4}}{\sqrt{\omega_0 a}} \exp\left(\frac{i(t-\tau)}{a}\right) \exp\left(-\frac{(t-\tau)^2}{2\omega_0^2 a^2}\right), \quad (20)$$

which has a unit norm, the scalogram can be formulated as

$$\begin{aligned} |S(\tau,a)|^2 &= \gamma(a)|\langle y_{\tau,a}|x\rangle|^2 \\ &= \gamma(a)||\frac{|y_{\tau,a}\rangle\langle y_{\tau,a}|}{\langle y_{\tau,a}|y_{\tau,a}\rangle}|x\rangle||^2, \end{aligned} \quad (21)$$

where $\gamma(a) = \alpha$ ($\alpha \in \mathbb{R}$) if $c(a) \sim 1/\sqrt{a}$, or $\gamma(a) \sim 1/a$ if $c(a) \sim 1/a$.

3 The continuous wavelet transform of irregularly sampled time series

3.1 The model for the data

We consider the same model as in Part 1:

$$\begin{aligned} |X\rangle &= |\text{Trend}\rangle + E_{\tau,a}\cos(\Omega|t\rangle + \phi_{\tau,a}) + |\text{Noise}\rangle \\ &= |\text{Trend}\rangle + A_{\tau,a}|c_\Omega\rangle + B_{\tau,a}|s_\Omega\rangle + |\text{Noise}\rangle, \end{aligned} \quad (22)$$

where $|X\rangle = [X_1, \ldots, X_N]'$ and is real, $|t\rangle = [t_1, \ldots, t_N]'$, $A_{\tau,a} = E_{\tau,a}\cos(\phi_{\tau,a})$, $B_{\tau,a} = -E_{\tau,a}\sin(\phi_{\tau,a})$, $E^2_{\tau,a} = A^2_{\tau,a} + B^2_{\tau,a}$, $|c_\Omega\rangle = [\cos(\Omega t_1), \ldots, \cos(\Omega t_N)]'$ and $|s_\Omega\rangle = [\sin(\Omega t_1), \ldots, \sin(\Omega t_N)]'$. We have added subscripts (τ, a) since all the subsequent analyses will be done in the time–scale plane. The trend is a polynomial of degree m,

$$|\text{Trend}\rangle = \sum_{k=0}^{m} \gamma_k |t^k\rangle, \quad (23)$$

and the background noise term, $|\text{Noise}\rangle$, is a zero-mean stationary Gaussian CARMA process sampled at the times of $|t\rangle$, as defined in Sect. 3.2 of Part 1. As stated in Part 1, considering or not the presence of a trend in the model for the data is left to the user, given that we can always interpret a polynomial trend of low order as a very low-frequency oscillation.

3.2 The scalogram

When applying the CWT to finite discrete time series, a choice for the discretisation must be made. In the influential paper of Torrence and Compo (1998), which deals with regularly sampled time series, the expression under the form of a convolution product in the Fourier space, Eq. (5), is conserved, and computed with the discrete Fourier transform (DFT) of the data. The CWT is then the inverse DFT of the convolution product. Unfortunately, we cannot extend the convolution theorem[6] to irregularly spaced time series and we cannot therefore follow the same computational procedure as in Torrence and Compo (1998). Alternatively, we can conserve the squared norm of the orthogonal projection, Eq. (21). The advantage of such a formalism is that it can be applied to irregularly sampled time series, as shown in Part 1. Similarly to Part 1, we work with cosines and sines instead of working with complex exponentials. Very little difference is observed between both choices. Based on the results of Sect. 2.8, our Morlet wavelet scalogram for irregularly sampled time series is therefore

$$||\mathbf{P}_{\overline{\text{sp}}\{|\mathbf{G}_{\tau,a}c_{\tau,a}\rangle, |\mathbf{G}_{\tau,a}s_{\tau,a}\rangle\}}|X\rangle||^2, \quad (24)$$

where $\mathbf{G}_{\tau,a}$ is a diagonal matrix with diagonal elements

$$\mathbf{G}_{\tau,a_{ii}} = \exp\left(-\frac{(t_i-\tau)^2}{2\omega_0^2 a^2}\right), \quad \forall i \in \{1, \ldots, N\}, \quad (25)$$

and

$$|c_{\tau,a}\rangle = \cos((|t\rangle - \tau)/a), \quad |s_{\tau,a}\rangle = \sin((|t\rangle - \tau)/a), \quad (26)$$

are vectors of length N. We can impose $\tau = 0$ into the cosine and sine terms, since $\overline{\text{sp}}\{|\mathbf{G}_{\tau,a}c_{\tau,a}\rangle, |\mathbf{G}_{\tau,a}s_{\tau,a}\rangle\}$ is invariant with respect to the variable τ appearing in the cosine and sine, and the scalogram becomes

$$||\mathbf{P}_{\overline{\text{sp}}\{|\mathbf{G}_{\tau,a}c_a\rangle, |\mathbf{G}_{\tau,a}s_a\rangle\}}|X\rangle||^2. \quad (27)$$

In the following, the notations $|\mathbf{G}_{\tau,a}c_a\rangle$ and $\mathbf{G}_{\tau,a}|c_a\rangle$ refer to the same vector. Our wavelet scalogram is similar to the tapered periodogram defined in Sect. 4.4 of Part 1, and its properties and generalisations will therefore be similar as well. In particular, the variables a and τ are considered as continuous variables, similarly to the continuous frequency variable of Part 1.

When the time series is regularly sampled, the scalogram, given by Eq. (27), is extremely close to what is obtained with the traditional approach based on the convolution theorem, e.g. in Torrence and Compo (1998), Maraun and Kurths (2004) or Cohen and Walden (2010). This is illustrated with a regularly sampled time series representing a noisy version of the caloric summer insolation[7] at 65° N, in Fig. 1b and c. We also show the effect of randomly removing a large amount of data points, with the resulting time series being irregularly sampled. This is illustrated in Fig. 1d and we observe that our algorithms still do a good job at estimating significant regions, although there are some artefacts and the power tends to be overestimated.

[6] The convolution theorem for continuous-time functions is given in Appendix A, and its counterpart for regularly sampled time series is given in Mallat (2009, p. 74).

[7] It is computed from the R package PALINSOL (https://CRAN.R-project.org/package=palinsol) from the formulas of Berger (1978) and the data of Laskar et al. (2004).

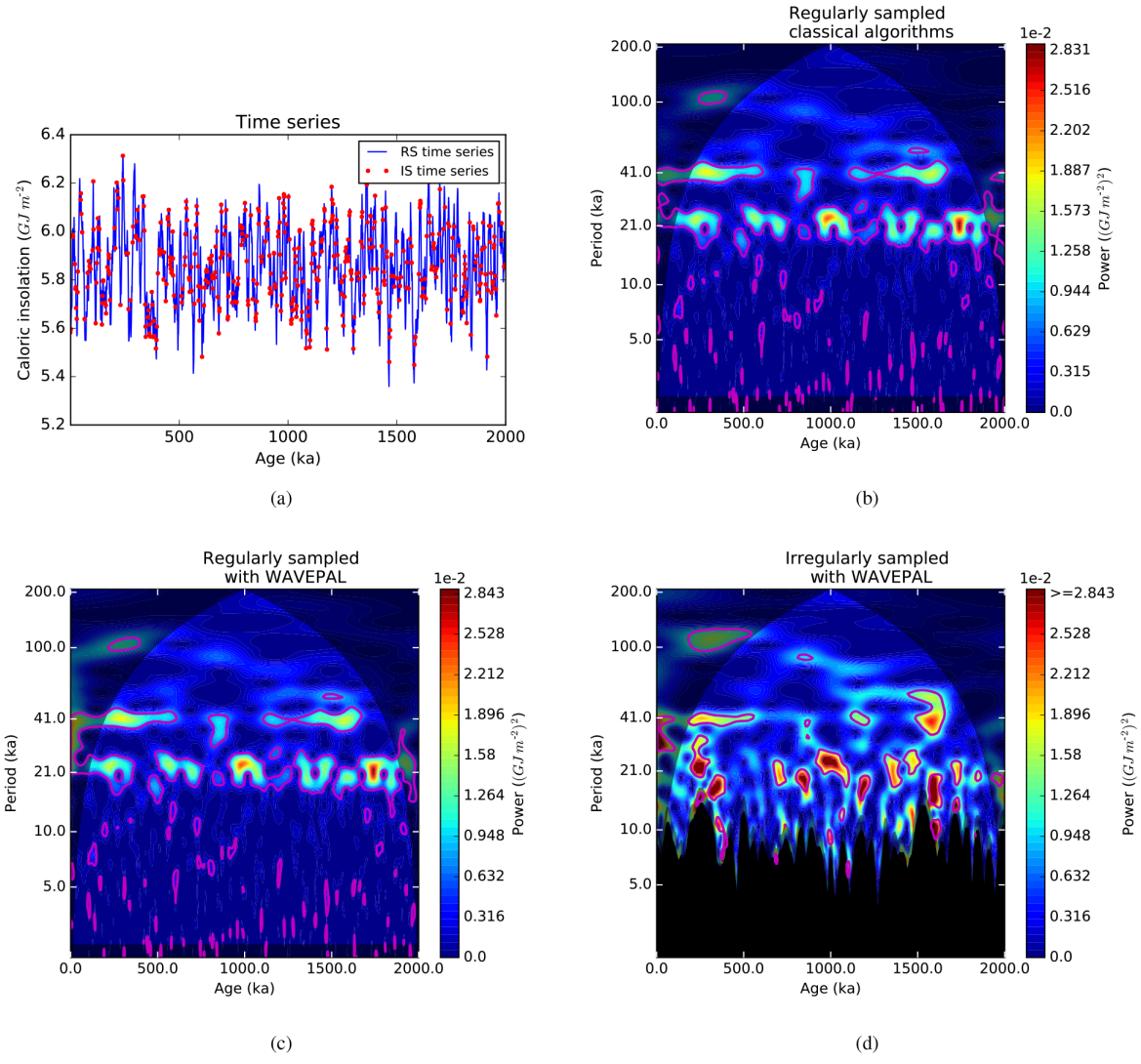


Figure 1. (a) The regularly sampled (RS) time series, with $\Delta t = 1$ ka, which represents the caloric summer insolation at 65° N to which is added a realisation of a Gaussian red noise process with $\alpha = 0.1$ and σ equal to half of the standard deviation of the original time series (these parameters are defined in Sect. 3.2.3 of Part 1). The red dots are obtained from randomly removing 75 % of the data points of the RS time series, resulting in an irregularly sampled (IS) time series with 500 data points. Panels **(b)**, **(c)** and **(d)** compare of the scalograms with $\omega_0 = 10$, jointly with their 95 % analytical confidence levels against a red noise background. **(b)** Scalogram of the RS time series computed with the classical approach. **(c)** Scalogram of the RS time series computed with WAVEPAL. **(d)** Scalogram of the IS time series computed with WAVEPAL. The black zone, called the Shannon–Nyquist exclusion zone (SNEZ) and defined in Sect. 3.8, is the area where the sampling is not sufficient to probe the lowest periods. In panels **(b)**, **(c)** and **(d)**, the two lateral shaded areas are the half-cones of influence (see Sect. 3.7), and the bottom shaded area is the refinement of the SNEZ (defined in Sect. 3.10). The bounds of the colour scale in the panels **(b)**, **(c)** and **(d)** are the extrema of the scalogram in panel **(c)** over the non-shaded area in order to make a meaningful visual comparison. Technical details about the computation of the scalogram and its confidence levels are given in Sects. 3 and 4.

Equation (27) reduces to the Lomb–Scargle periodogram, defined in Eq. (I,40), if the weight $\mathbf{G}_{\tau,a}$ is replaced by the identity matrix. Similarly to the Lomb–Scargle periodogram, we rescale $|\mathbf{G}_{\tau,a}c_a\rangle$ and $|\mathbf{G}_{\tau,a}s_a\rangle$ such that they are orthonormal. This can be done by defining

$$|c_a^\sharp\rangle = \frac{\cos\left(|t\rangle/a - \beta_{\tau,a}\right)}{\sqrt{\Sigma_{i=1}^N \mathbf{G}_{\tau,a_{ii}}^2 \cos^2\left(t_i/a - \beta_{\tau,a}\right)}},$$
$$|s_a^\sharp\rangle = \frac{\sin\left(|t\rangle/a - \beta_{\tau,a}\right)}{\sqrt{\Sigma_{i=1}^N \mathbf{G}_{\tau,a_{ii}}^2 \sin^2\left(t_i/a - \beta_{\tau,a}\right)}}, \tag{28}$$

where $\beta_{\tau,a}$ is the solution of

$$\tan(2\beta_{\tau,a}) = \frac{\Sigma_{i=1}^{N} \mathbf{G}^2_{\tau,a_{ii}} \sin(2t_i/a)}{\Sigma_{i=1}^{N} \mathbf{G}^2_{\tau,a_{ii}} \cos(2t_i/a)}. \quad (29)$$

The scalogram is then

$$||\mathbf{P}_{\overline{\mathrm{sp}}\{|\mathbf{G}_{\tau,a}c_a\rangle,|\mathbf{G}_{\tau,a}s_a\rangle\}}|X\rangle||^2 = \langle \mathbf{G}_{\tau,a}c_a^\sharp|X\rangle^2 + \langle \mathbf{G}_{\tau,a}s_a^\sharp|X\rangle^2. \quad (30)$$

3.3 Scalogram and trend

Analogously to Sect. 4.3 of Part 1, we extend the scalogram to take into account the presence of a polynomial trend of degree m in the data. Indeed, the scalogram defined in Sect. 3.2 applies well to data which can be modelled as $|X\rangle = A_{\tau,a}|c_\Omega\rangle + B_{\tau,a}|s_\Omega\rangle + |\mathrm{Noise}\rangle$. If we want to work with the full model, Eq. (22), holding a polynomial trend of degree m, we define a new scalogram as

$$||\left(\mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle,|\mathbf{G}_{\tau,a}c_a\rangle,|\mathbf{G}_{\tau,a}s_a\rangle\}} - \mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle\}}\right)|X\rangle||^2, \quad (31)$$

which is invariant with respect to the parameters of the trend. This is the analogue of Eq. (I,57).

3.4 Smoothing the scalogram

The scalogram suffers from the same inconsistency issue as the periodogram, in the sense that it remains very noisy regardless of the number of data points we have at our disposal[8]. Smoothing techniques must therefore be applied, and we proceed like in Part 1, extending the formulas used with regularly sampled time series. Note that the disadvantage of any smoothing procedure is that the resolution (in time, frequency or both, depending on the smoothing choice) is reduced. Consequently, there is always a trade-off between variance reduction and resolution.

Smoothing is traditionally performed by averaging the scalogram over neighbouring points in the time–scale plane, either by averaging over times followed by averaging over scales (Torrence and Webster, 1999; Grinsted et al., 2004), or simply by averaging over time (Cohen and Walden, 2010). In this work, we apply the latter technique because, even for very simple signals like $|X\rangle = \sin(\omega|t\rangle)$, the correlations in the scalogram between neighbouring scales, for a fixed time, are highly irregular when the time series is irregularly sampled, unlike the correlations between neighbouring times, for a fixed scale, which are driven by the Gaussian shape of the wave packets $|\mathbf{G}_{\tau,a}c_a\rangle$ and $|\mathbf{G}_{\tau,a}s_a\rangle$. Smoothing over time must be carried out in accordance with the length of the support of the wave packets, which is proportional to the scale and to parameter ω_0 (Eq. 25). This choice also implies that the number of oscillations over which smoothing is performed is constant throughout the time–scale plane. This results from Eq. (26).

We adopt here the smoothing procedure of Cohen and Walden (2010) for which they derived analytical asymptotic results in the case of regularly sampled time series. The averaging window is a square window with a length proportional to the scale. Our smoothed scalogram is

$$||\mathbf{P}_{\mathrm{smoothed}}(\tau,a)|X\rangle||^2 = \frac{1}{2\gamma\omega_0 a}\int_{\tau-\gamma\omega_0 a}^{\tau+\gamma\omega_0 a} \mathrm{d}\tau' ||\left(\mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle,|\mathbf{G}_{\tau',a}c_a\rangle,|\mathbf{G}_{\tau',a}s_a\rangle\}} - \mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle\}}\right)|X\rangle||^2, \quad (32)$$

in which γ is called the *smoothing coefficient*. Appendix D provides further details on the practical implementation of the bounds of integration.

3.5 The amplitude scalogram

3.5.1 Definition

We want to estimate the amplitude $E_{\tau,a} = \sqrt{A^2_{\tau,a} + B^2_{\tau,a}}$ of our model, Eq. (22), at a given point (τ, a) in the time–scale plane. The estimation of $E^2_{\tau,a}$ is called the *amplitude scalogram* and is denoted by $\widehat{E}^2_{\tau,a}$. We start with a trendless signal and derive an approximate formula linking the amplitude scalogram and the scalogram.

3.5.2 Trendless signal

Formula (I,118) is applied with the left-hand-side term changed to encompass wavelet formalism. $\widehat{A}$ and $\widehat{B}$ are determined by projecting the data onto the tapered cosine and sine:

$$\mathbf{P}_{\overline{\mathrm{sp}}\{|\mathbf{G}_{\tau,a}c_a\rangle,|\mathbf{G}_{\tau,a}s_a\rangle\}}|X\rangle = \widehat{A}_{\tau,a}|c_\omega\rangle + \widehat{B}_{\tau,a}|s_\omega\rangle = \mathbf{V}_{\omega_2}|\widehat{\Phi}_{\tau,a}\rangle, \quad (33)$$

where the taper $\mathbf{G}_{\tau,a}$ is defined in Eq. (25),

$$\mathbf{V}_{\omega_2} = \begin{pmatrix} | & | \\ |c_\omega\rangle & |s_\omega\rangle \\ | & | \end{pmatrix}, \text{ and } |\widehat{\Phi}_{\tau,a}\rangle = \begin{pmatrix} \widehat{A}_{\tau,a} \\ \widehat{B}_{\tau,a} \end{pmatrix}. \quad (34)$$

Conversion from the angular frequency ω to the scale a is performed with the formula $\omega = 1/a$ (justification is given in Sect. 3.9). Using the same development as in Sect. 6.2.2 of Part 1, we obtain

$$|\widehat{\Phi}_{\tau,a}\rangle = (\mathbf{V}'_{a_2}\mathbf{G}_{\tau,a}\mathbf{V}_{a_2})^{-1}\mathbf{V}'_{a_2}\mathbf{G}_{\tau,a}|X\rangle. \quad (35)$$

The amplitude scalogram is then

$$\widehat{E}^2_{\tau,a} = ||\,|\widehat{\Phi_{\tau,a}}\rangle||^2. \quad (36)$$

[8] The scalogram often looks *smooth* because neighbouring points in the time–frequency plane are strongly correlated, but it nevertheless remains inconsistent (see the discussion in Maraun and Kurths, 2004, Sect. 4.2).

The approximations made in Sect. 6.2.2 of Part 1 are valid in this work, and applying Eq. (I,128) to our case gives an approximate formula linking the scalogram and the amplitude scalogram, namely

$$\widehat{E}^2_{\tau,a} \approx \frac{2\mathrm{tr}\left(\mathbf{G}^2_{\tau,a}\right)}{\mathrm{tr}\left(\mathbf{G}_{\tau,a}\right)^2} ||\mathbf{P}_{\overline{\mathrm{sp}}\{|\mathbf{G}_{\tau,a}c_a\rangle,|\mathbf{G}_{\tau,a}s_a\rangle\}}|X\rangle||^2. \tag{37}$$

Let us compare this equation with its continuous counterpart, Eq. (21), in which the weight must be $\gamma(a) \sim 1/a$ to get an estimation of the local squared amplitude, as explained in Sect. . The comparison is made by analysing the weight of the right-hand-side term of Eq. (37) in the continuous limit:

$$\frac{1}{\overline{\Delta t}} \frac{2\mathrm{tr}\left(\mathbf{G}^2_{\tau,a}\right)}{\mathrm{tr}\left(\mathbf{G}_{\tau,a}\right)^2} \longrightarrow \frac{2\int_{-\infty}^{+\infty} \mathrm{d}t \exp\left(-\frac{(t-\tau)^2}{\omega_0^2 a^2}\right)}{\left(\int_{-\infty}^{+\infty} \mathrm{d}t \exp\left(-\frac{(t-\tau)^2}{2\omega_0^2 a^2}\right)\right)^2} = \frac{1}{\sqrt{\pi}\omega_0 a}, \tag{38}$$

where $\overline{\Delta t}$ is the average time step. This is proportional to $1/a$ and it is therefore consistent with the continuous case.

3.5.3 Signal with a trend

Formula (I,129) is applied with the left-hand-side term changed to encompass wavelet formalism:

$$\begin{aligned}\mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle,|\mathbf{G}_{\tau,a}c_a\rangle,|\mathbf{G}_{\tau,a}s_a\rangle\}}|X\rangle \\ = \sum_{k=0}^{m} \widehat{\gamma_k}|t^k\rangle + \widehat{A}|c_\omega\rangle + \widehat{B}|s_\omega\rangle \\ = \mathbf{V}_{\omega_{m+3}}|\widehat{\Phi}\rangle,\end{aligned} \tag{39}$$

where

$$\mathbf{V}_{\omega_{m+3}} = \begin{pmatrix} | & & | & | & | \\ |t^0\rangle & \ldots & |t^m\rangle & |c_\omega\rangle & |s_\omega\rangle \\ | & & | & | & | \end{pmatrix}, \tag{40}$$

and

$$|\widehat{\Phi}\rangle = \begin{pmatrix} \widehat{\gamma_0} \\ \vdots \\ \widehat{\gamma_m} \\ \widehat{A} \\ \widehat{B} \end{pmatrix}. \tag{41}$$

Conversion from the angular frequency ω to the scale a is performed with the formula $\omega = 1/a$ (justification is given in Sect. 3.9). Using the same development as in Sect. 6.3 of Part 1, we obtain

$$|\widehat{\Phi}_{\tau,a}\rangle = \left(\mathbf{W}'_{\tau,a_{m+3}}\mathbf{V}_{a_{m+3}}\right)^{-1}\mathbf{W}'_{\tau,a_{m+3}}|X\rangle, \tag{42}$$

where $\mathbf{W}_{\tau,a_{m+3}}$ is identical to $\mathbf{V}_{a_{m+3}}$ except in the last two columns, where the cosine and sine are tapered by $\mathbf{G}_{\tau,a}$. This gives

$$\widehat{A}_{\tau,a} = \widehat{\Phi}_{\tau,a}(m+2), \quad \widehat{B}_{\tau,a} = \widehat{\Phi}_{\tau,a}(m+3), \tag{43}$$

where $\widehat{\Phi}_{\tau,a}(m+2)$ and $\widehat{\Phi}_{\tau,a}(m+3)$ are the two last components of vector $|\widehat{\Phi}_{\tau,a}\rangle$. The amplitude scalogram is then

$$\widehat{E}^2_{\tau,a} = \widehat{A}^2_{\tau,a} + \widehat{B}^2_{\tau,a}. \tag{44}$$

3.5.4 With smoothing

Like in Part 1, estimating the amplitude is more robust against noise when a smoothing procedure is performed. We apply to the squared amplitude, Eq. (44), the same kind of smoothing as for the scalogram (see Eq. 32), giving

$$\widehat{E}^2_{\tau,a} = \frac{1}{2\gamma\omega_0 a}\int_{\tau-\gamma\omega_0 a}^{\tau+\gamma\omega_0 a} \mathrm{d}\tau' \left(\widehat{\Phi}_{\tau,a}(m+2)^2 + \widehat{\Phi}_{\tau,a}(m+3)^2\right). \tag{45}$$

Appendix D provides further details on the practical implementation of the bounds of integration.

3.6 The weighted smoothed scalogram

The weighted smoothed scalogram is the analogue of the *weighted WOSA periodogram*, defined in Sect. 7 of Part 1, and its objectives are the same, i.e. to keep the advantages of both the amplitude scalogram and the scalogram, namely,

- provide an estimation of the squared amplitude of a signal, locally in the time–frequency plane, by weighting the scalogram like in Eq. (37);
- conserve the advantage of the formalism of orthogonal projections in order to avoid the matrix inversions required for the computation of the amplitude scalogram (see, for example, Eq. 45, relying on Eq. 42 which requires a matrix inversion).

The last item is useful for building confidence levels when performing a test of significance (see Sect. 4). The disadvantage of weighting the smoothed scalogram is that it no longer provides a flat *pseudo-wavelet power spectrum* for a white noise signal (see Sect. 4.2), analogously to its frequency counterpart (see Sect. 7 of Part 1). The weighted smoothed scalogram is derived from Eq. (32), in which the integrand is weighted by the right-hand-side weight of Eq. (37), namely,

$$||\mathbf{P}_{\mathrm{smoothed}}(\tau,a)|X\rangle||^2 = \frac{1}{2\gamma\omega_0 a}\int_{\tau-\gamma\omega_0 a}^{\tau+\gamma\omega_0 a} \mathrm{d}\tau' \frac{2\mathrm{tr}(\mathbf{G}^2_{\tau',a})}{\mathrm{tr}(\mathbf{G}_{\tau',a})^2} ||\left(\mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle,|\mathbf{G}_{\tau',a}c_a\rangle,|\mathbf{G}_{\tau',a}s_a\rangle\}} - \mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle\}}\right)|X\rangle||^2. \tag{46}$$

Appendix D provides further details on the practical implementation of the bounds of integration. We recommend the use of the weighted smoothed scalogram in most time–frequency analyses under irregular sampling.

3.7 Cone of influence

When the wave packets $|\mathbf{G}_{\tau,a}c_a\rangle$ and $|\mathbf{G}_{\tau,a}s_a\rangle$ intersect the borders of the time series, a part of their support can stand after the last point of the time series, or before the first point of the time series. Consequently, one has to remove two half-cones from the area under analysis. From Eq. (25), the support of the wave packets is approximately equal to $2\beta\omega_0 a$, so that the excluded areas are given by

$$\{\tau, a\} \text{ such that } |\tau - t_1| \leq \beta\omega_0 a \text{ and } |\tau - t_N| \leq \beta\omega_0 a, \quad (47)$$

with $\beta = 3$ (conservative choice) or $\beta = \sqrt{2}$ (choice in Torrence and Compo, 1998). We recommend the conservative choice. When smoothing is performed, Eq. (47) becomes

$$\begin{aligned} &\{\tau, a\} \quad \text{such that} \quad (48)\\ &|\tau - t_1| \leq (\beta+\gamma)\omega_0 a \quad \text{and} \quad |\tau - t_N| \leq (\beta+\gamma)\omega_0 a, \end{aligned}$$

where γ controls the smoothing length; see Eqs. (32), (45) and (46).

This has another implication: the maximal scale available by the analysis is

$$a_{\max} = \frac{t_N - t_1}{2(\beta+\gamma)\omega_0}. \quad (49)$$

3.8 Aliasing and Shannon–Nyquist exclusion zone (SNEZ)

When probing the irregularly sampled time series with the wavelet packet, it may happen that the period of the oscillation inside the packet, $2\pi a$, is too low compared to the local time step in the time series, therefore causing aliasing issues according to the Shannon–Nyquist theorem, locally in the time–scale plane. As stated in Part 1, this issue also happens with the WOSA periodogram. We adapt formulas (I,72), (I,73) and (I,74) and define the local time step as

$$\overline{\Delta t}_{\tau,a} = \max\{\overline{\Delta t}_{\mathbf{G}_{\tau,a}}, \overline{\Delta t}_{\mathbf{H}_{\tau,a}}\}, \quad (50)$$

where

$$\begin{aligned} &\overline{\Delta t}_{\mathbf{G}_{\tau,a}} = \frac{\sum_{k=1}^{N} \mathbf{G}_{\tau,a_{k,k}} \Delta t_{c_k}}{\mathrm{tr}(\mathbf{G}_{\tau,a})},\\ &\overline{\Delta t}_{\mathbf{H}_{\tau,a}} = \frac{\sum_{k=1}^{N-1} \mathbf{H}_{\tau,a_{k,k}} \Delta t_k}{\mathrm{tr}(\mathbf{H}_{\tau,a})}, \quad (51)\\ &\Delta t_k = t_{k+1} - t_k, \ \forall k \in \{1, \ldots N-1\},\\ &\Delta t_{c_k} = \frac{t_{k+1} - t_{k-1}}{2}, \ \forall k \in \{2, \ldots N-1\}, \end{aligned}$$

$$\Delta t_{c_1} = t_2 - t_1, \Delta t_{c_N} = t_N - t_{N-1}, \quad (52)$$

and $\mathbf{H}_{\tau,a}$ is a diagonal matrix with

$$\mathbf{H}_{\tau,a_{k,k}} = \exp\left(-\frac{\left(\frac{t_k+t_{k+1}}{2} - \tau\right)^2}{2\omega_0^2 a^2}\right), \ k \in \{1, \ldots, N-1\}. \quad (53)$$

We then apply the Shannon–Nyquist theorem to this local time step, namely:

$$\text{Compute the scalogram at } (\tau, a) \text{ if } a \geq a_{\mathrm{SNEZ}}(\tau), \quad (54)$$

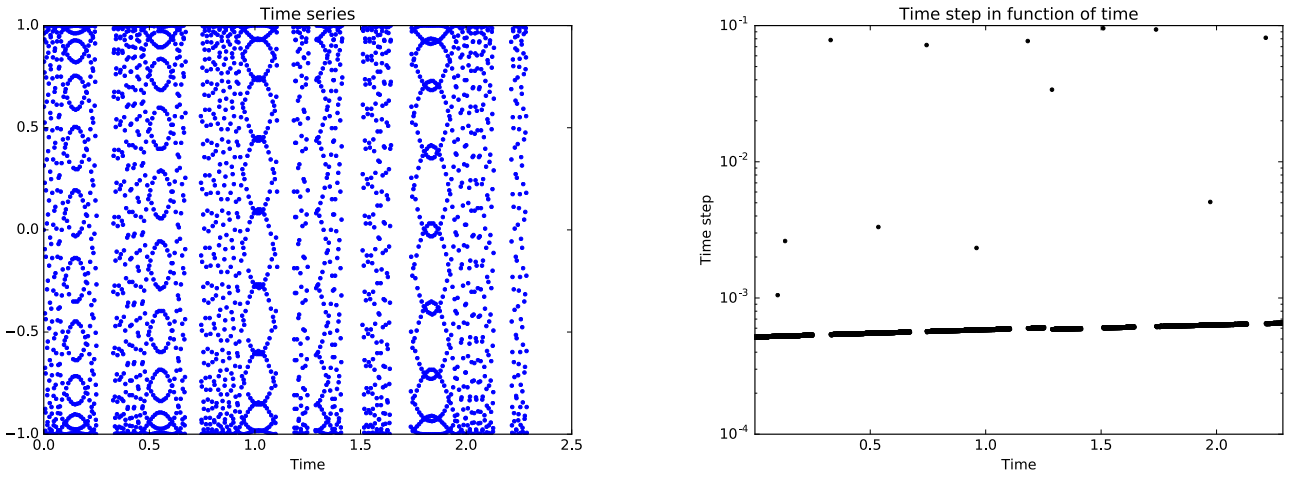
where

$$a_{\mathrm{SNEZ}}(\tau) \text{ is the largest solution of } a = \frac{\overline{\Delta t}_{\tau,a}}{\pi}. \quad (55)$$

We call *Shannon–Nyquist exclusion zone*, the area in the scalogram that does not satisfy Eq. (54) and which is therefore delimited by a_{SNEZ}. Note that matrix $\mathbf{H}_{\tau,a}$ is similar to matrix $\mathbf{G}_{\tau,a}$, defined in Eq. (25), but with elements taken at $(t_k+t_{k+1})/2$ instead of t_k. Quantity $\overline{\Delta t}_{\tau,a}$ is equal to the maximum between the average weighted time step and the average weighted central time step.

We now justify Formula (50) with an example. Consider the function $X(t) = \sin(2\pi t/0.01)$, sampled on an irregular grid. This is drawn in Fig. 2a. The time step is represented in Fig. 2b. These two figures show that the time series exhibits intervals where it is more or less regularly sampled, separated by large gaps. The weighted (unsmoothed) scalogram is drawn in Fig. 3a. We remind the reader that the weighted scalogram is supposed to estimate the local squared amplitude in the time–frequency plane. Since $X(t)$ has an amplitude equal to 1, we expect that the maximal power of the scalogram be equal to 1, along a scale corresponding to the period of x, for all τ. Because of the large gaps in the time series, extended regions corrupted by aliasing occur in Fig. 3a, resulting in a maximal power for the scalogram which is much greater than 1. Figure 3b, c and d present the weighted scalogram corrected by the SNEZ. In Fig. 3b the SNEZ is computed with $\overline{\Delta t}_{\tau,a} = \overline{\Delta t}_{\mathbf{G}_{\tau,a}}$. We observe that it does a good job of rejecting the areas where aliasing occur, although it is desirable that the black areas peak on higher scales. In Fig. 3c, the SNEZ is computed with $\overline{\Delta t}_{\tau,a} = \overline{\Delta t}_{\mathbf{H}_{\tau,a}}$. We observe that most of the aliasing-related areas are rejected, although we would rather the black areas be wider. Finally, the SNEZ computed with $\overline{\Delta t}_{\tau,a} = \max\{\overline{\Delta t}_{\mathbf{G}_{\tau,a}}, \overline{\Delta t}_{\mathbf{H}_{\tau,a}}\}$ is drawn in Fig. 3d and we observe that it does a very satisfactory job at rejecting the areas where aliasing occur.

The SNEZ is applied to all the analysis tools defined above. When smoothing is to be applied, it is performed on the areas outside of the SNEZ, since the scalogram is not computed in the SNEZ. In the neighbourhood of the SNEZ, adjustments of the smoothing procedure are therefore necessary, as explained in Appendix D.



(a) (b)

Figure 2. (a) The time series $|X\rangle = \sin(2\pi|t\rangle/0.01)$ and **(b)** its time step, with the vertical axis in log-scale. The time vector $|t\rangle$ is taken from a real palaeoclimate time series (Liviu Giosan, WHOI, personal communication, 2017).

3.9 From the scale to the period

Scale-to-period conversion is performed in the continuous limit, with Eq. (18). The first case of Eq. (18), with $c(a) \sim 1/a$, corresponds to estimators of the *amplitude*, and is then used for scale-to-period conversion with the amplitude scalogram (all the formulas of Sect. 3.5) and for the weighted smoothed scalogram, Eq. (46). The second case of Eq. (18), with $c(a) \sim 1/\sqrt{a}$, is used for scale-to-period conversion with the unweighted scalogram, that is, the formulas appearing in Sect. 3.2, 3.3 and 3.4.

3.10 Refining the Shannon–Nyquist exclusion zone

As illustrated in Fig. 4, the Shannon–Nyquist exclusion zone may not to be sufficient to avoid all the patches due to aliasing, because of the correlations between neighbouring scales in the scalogram. We therefore extend the Shannon–Nyquist exclusion zone by considering the continuous limit case for the simple periodic signal $x(t) = \exp(i2\pi t/T_{\mathrm{SNEZ}}(\tau))$, where $T_{\mathrm{SNEZ}}(\tau)$ is the period at the border of the SNEZ, determined by Eqs. (55) and (18). Its scalogram is given in Eq. (17). In order to make the correspondence with all the above formulas, three cases are considered.

1. $c(a) \sim 1/a$: in this case, we have $|S(\tau,a)|^2 \sim \exp(-\omega_0^2(2\pi a/T_{\mathrm{SNEZ}}(\tau)-1)^2)$, and the standard deviation for the scale is then $\sigma_{a,1}(\tau) = T_{\mathrm{SNEZ}}(\tau)/2\sqrt{2}\pi\omega_0$. The border of the *extended* Shannon–Nyquist exclusion zone at time τ is therefore on scale $a_{\mathrm{SNEZ}}(\tau)+\beta\sigma_{a,1}(\tau)$, where β is a coefficient estimating the half-support of Gaussian-shaped functions (it is defined in Sect. 3.7).

2. $c(a) \sim 1/a$ and work with $|S(\tau,a)|$: in this case, we have $|S(\tau,a)| \sim \exp(-\omega_0^2(2\pi a/T_{\mathrm{SNEZ}}(\tau)-1)^2/2)$, and the standard deviation for the scale is then $\sigma_{a,2}(\tau) = T_{\mathrm{SNEZ}}(\tau)/2\pi\omega_0$. The border of the extended Shannon–Nyquist exclusion zone at time τ is therefore on scale $a_{\mathrm{SNEZ}}(\tau)+\beta\sigma_{a,2}(\tau)$.

3. $c(a) \sim 1/\sqrt{a}$: in this case, we have $|S(\tau,a)|^2 \sim a\exp(-\omega_0^2(2\pi a/T_{\mathrm{SNEZ}}(\tau)-1)^2)$. We know from Eq. (18) that the scalogram is at its maximum on the scale $a_{\mathrm{SNEZ}}(\tau) = T_{\mathrm{SNEZ}}(\tau)(\omega_0+\sqrt{\omega_0^2+2})/4\pi\omega_0$. The pseudo-standard deviation is computed such that $a\exp(-\omega_0^2(2\pi a/T_{\mathrm{SNEZ}}(\tau)-1)^2)$ decreases from its maximum by the same percentage as in case 1, namely, $\beta\sigma_{a,3}(\tau)$ is equal to the largest of the two solutions of

$$a\exp\left(-\omega_0^2(2\pi a/T_{\mathrm{SNEZ}}(\tau)-1)^2\right) = a_{\mathrm{SNEZ}}(\tau)\exp\left(-\omega_0^2(2\pi a_{\mathrm{SNEZ}}(\tau)/T_{\mathrm{SNEZ}}(\tau)-1)^2\right)\exp\left(-\beta^2/2\right),$$

in which the unknown is a. The border of the extended Shannon–Nyquist exclusion zone at time τ is therefore on scale $a = a_{\mathrm{SNEZ}}(\tau)+\beta\sigma_{a,3}(\tau)$.

Case 1 is used with formulas giving the squared amplitude $\widehat{E}^2_{\tau,a}$ in Sect. 3.5 and with the weighted smoothed scalogram, Eq. (46). The (unsquared) amplitude $\widehat{E}_{\tau,a}$ can also be of interest, and case 2 is therefore used. Case 3 is used with formulas arising in Sect. 3.2, 3.3 and 3.4. Finally, note that the refinement of the SNEZ is performed after the smoothing procedure, because an extension of the SNEZ may result from the smoothing, as explained in Appendix D.

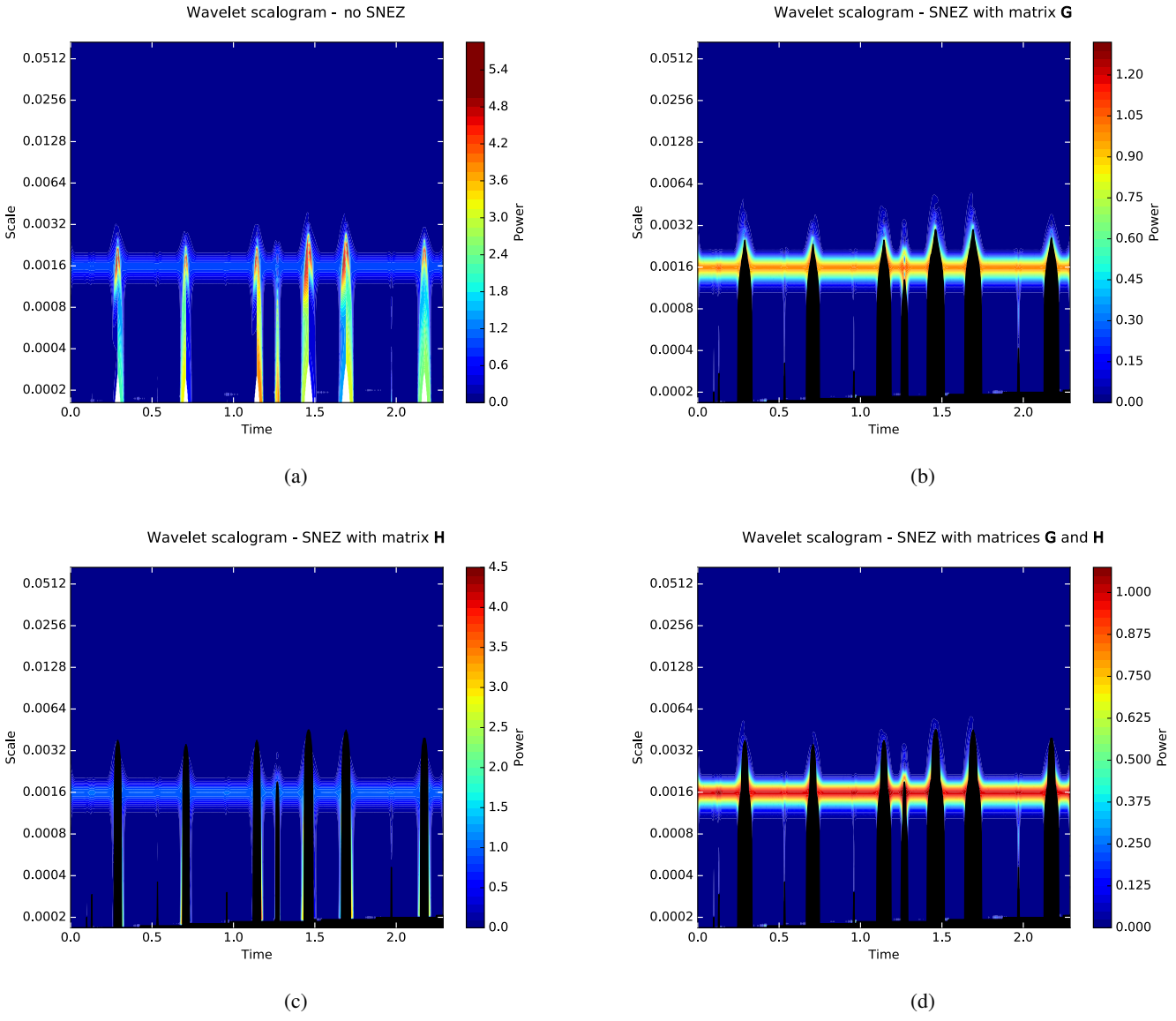


Figure 3. Weighted (unsmoothed) scalogram of the time series presented in Fig. 2a. **(a)** No correction for aliasing. **(b)** Corrected with $\overline{\Delta t}_{\tau,a} = \overline{\Delta t}_{\mathbf{G}_{\tau,a}}$. **(c)** Corrected with $\overline{\Delta t}_{\tau,a} = \overline{\Delta t}_{\mathbf{H}_{\tau,a}}$. **(d)** Corrected with $\overline{\Delta t}_{\tau,a} = \max\{\overline{\Delta t}_{\mathbf{G}_{\tau,a}}, \overline{\Delta t}_{\mathbf{H}_{\tau,a}}\}$.

3.11 Discretising τ and a

With regularly sampled data, the discretised variable τ is usually equal to $|t\rangle$ (like in Torrence and Compo, 1998), or a subset of $|t\rangle$ with regularly spaced elements. For irregularly spaced time series, we opt for the same type of grid as in the regularly sampled case, i.e. a linear regular grid, namely

$$\tau_k = \tau_0 + k\Delta\tau, \quad k \in \{0, \ldots, K\}, \quad \text{with } \tau_0 \geq t_1 \text{ and } \tau_K \leq t_N. \tag{56}$$

The scales are commonly discretised as fractional powers of two (Torrence and Compo, 1998), namely

$$a_j = a_{\min} 2^{j\delta j}, \quad j = 0, \ldots, J, \tag{57}$$

where

$$J = \log_2 \left(a_{\max}/a_{\min}\right)/\delta j. \tag{58}$$

Here, $a_{\min}$ is the minimum over τ of a_{SNEZ} (defined in Eq. 55), and $a_{\max}$ is defined in Eq. (49). Discretisation as a power law comes from the geometry of the wavelet transform, and is justified in Appendix C.

The integrals in Eqs. (32), (45) and (46) are discretised with the rectangle method. In particular, the discretisation of the integrals in Eqs. (32) and (46) allows these formulas to be written as finite-size matrices. To this end, we apply a Gram–Schmidt orthonormalisation to the orthogonal projec-

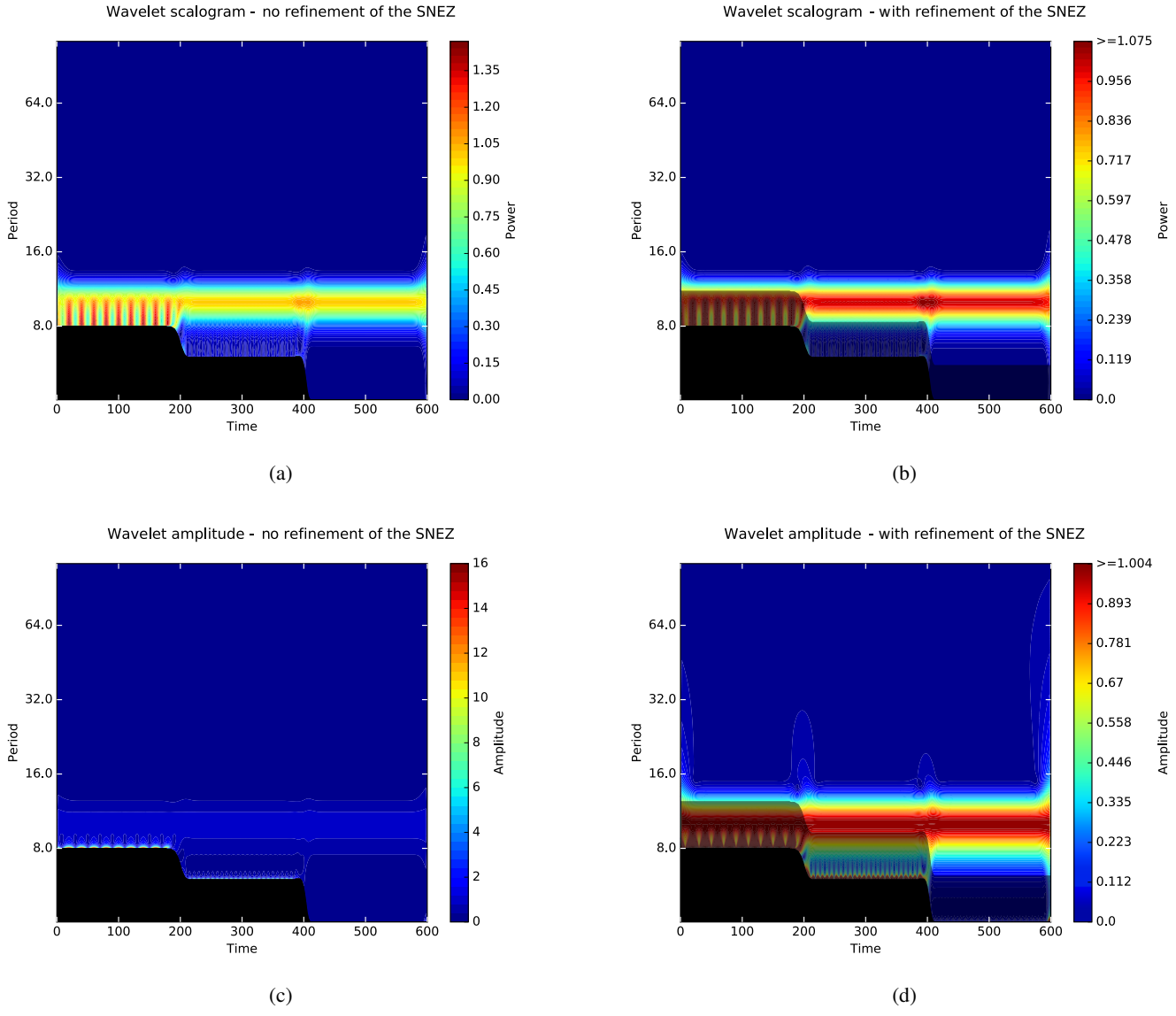


Figure 4. Scalogram of the time series $|X\rangle = \sin(2\pi|t\rangle/10)$, where $|t\rangle$ has a piecewise constant time step. From $t = 0$ to $t = 200$, $\Delta t = 4$. From $t = 200$ to $t = 400$, $\Delta t = 3$. From $t = 400$ to $t = 600$, $\Delta t = 2$. **(a)** Weighted (unsmoothed) scalogram. The black area is the SNEZ. **(b)** Same as panel **(a)** with the addition of the refinement of the SNEZ, which is the shaded area on the top of the SNEZ. **(c)** Amplitude scalogram (unsmoothed). The black area is the SNEZ. **(d)** Same as panel **(c)** with the addition of refinement of the SNEZ, which is the shaded area on the top of the SNEZ. In the four panels, the bounds of the colour scale are the extrema of the scalogram over the non-shaded area. Thanks to the refinement of the SNEZ, the upper bound of the colour scale is close to 1, which is the value of the squared amplitude of the signal $|X\rangle$.

tions, like in Eq. (I,67). This gives

$$||\mathbf{P}_{\text{smoothed}}(\tau,a)|X\rangle||^2 = \langle X|\mathbf{M}_{\tau,a}\mathbf{M}'_{\tau,a}|X\rangle, \tag{59}$$

which is the analogue of Eq. (I,68). $\mathbf{M}_{\tau,a}$ is a matrix of size $(N, 2n_{\text{col}}(\tau,a))$, $n_{\text{col}}(\tau,a) \geq 1$, where n_{col} is a non-trivial function depending on the scale and on the closeness of (τ,a) with the SNEZ and with edges of the time–frequency plane.

4 Significance testing with the scalogram

4.1 Hypothesis testing

We test for the presence of periodic components, locally in the time–frequency plane. Significance testing is mathematically expressed as a hypothesis testing. Taking our model, Eq. (22), the null hypothesis is

$$H_0 : A_{\tau,a} = B_{\tau,a} = 0. \tag{60}$$

Therefore, $|X\rangle = |\text{Trend}\rangle + |\text{Noise}\rangle$. The alternative hypothesis is

$$H_1 : A_{\tau,a} \text{ and } B_{\tau,a} \text{ are not both zero.} \quad (61)$$

The decision of accepting or rejecting the null hypothesis is based on the scalogram (Eq. 46), independently for each couple (τ, a) (this is called *pointwise testing*). Concretely, for each couple (τ, a), we compute the distribution of the scalogram under the null hypothesis, and then see if the *data* scalogram at (τ, a) is above or below a given percentile of that distribution. The percentile is called *level of confidence*. If the data scalogram is above the Xth percentile of the reference distribution, we reject the null hypothesis with $X\,\%$ of confidence. The *level of significance* is equal to $(100 - X)\,\%$, e.g. a 95 % confidence level is equivalent to a 5 % significance level.

To perform significance testing, we thus need

1. to estimate the parameters of the process under the null hypothesis (this is studied in Sect. 5.2 of Part 1);

2. to estimate the distribution of the scalogram under the null hypothesis (this is studied in Sect. 4.2 below).

Finally, we mention that, for regularly sampled time series, the pointwise significance test can be supplemented with an *area-wise* significance test, which takes into account the correlations between neighbouring points in the time–frequency plane. This is introduced in Maraun and Kurths (2004) and studied in detail in Maraun et al. (2007). Applying this method to irregularly sampled series is way beyond the scope of this work, since the correlations between neighbouring points in the time–frequency plane are highly irregular.

4.2 Estimation of the distribution of the scalogram under the null hypothesis

The results obtained for the periodogram in Sect. 5.3 of Part 1 are valid for the scalogram, with minor changes that we detail below.

1. Monte Carlo approach: the same procedure as in Part 1 is applied to the (weighted) smoothed scalogram, Eq. (32) or (46). We can thus estimate the confidence levels for the (weighted) smoothed scalogram taking into account the uncertainty in the parameters of the background noise.

2. Analytical approach (with a unique set of CARMA parameters):

   – Theorem 1 of Part 1 can be applied to the (weighted) smoothed scalogram, as follows.

   Theorem 1. *The (weighted) smoothed scalogram, defined in Eq. (59), under the null hypothesis (60), is*[9]

   $$||\mathbf{P}_{smoothed}(\tau,a)|X\rangle||^2 \stackrel{d}{=} \sum_{k=1}^{2n_{col}(\tau,a)} \lambda_k(\tau,a)\chi^2_{1_k}, \quad (62)$$

   where $|X\rangle = \sum_{k=0}^{m}\gamma_k|t^k\rangle + \mathbf{K}|Z\rangle$ *and* $\mathbf{K}$ *is the CARMA matrix defined in Eq. (I,20) or (I,38). The* $\chi^2_{1_k}$ *distributions are iid, and* $\lambda_1(\tau,a)$, ..., $\lambda_{2n_{col}(\tau,a)}(\tau,a)$ *are the eigenvalues of* $\mathbf{M}'_{\tau,a}\mathbf{K}\mathbf{K}'\mathbf{M}_{\tau,a}$ *and are non-negative. Matrix* $\mathbf{M}_{\tau,a}$ *is defined in Eq. (59).*

   – The *pseudo-wavelet power spectrum*, $\widehat{\text{WPS}}$, is the analogue of the pseudo-spectrum defined in Eq. (I,92). It is defined as the expected value of the (weighted) smoothed scalogram distribution, namely

   $$\begin{aligned}\widehat{\text{WPS}}(\tau,a) &= \sum_{k=1}^{2n_{\text{col}}(\tau,a)} \lambda_k(\tau,a) \\ &= \text{tr}\left(\mathbf{M}'_{\tau,a}\mathbf{K}\mathbf{K}'\mathbf{M}_{\tau,a}\right). \end{aligned} \quad (63)$$

   – For a Gaussian white noise background with variance σ^2, the unweighted pseudo-wavelet power spectrum is flat, and is equal to $2\sigma^2$, for all (τ, a). Moreover, if the scalogram is not smoothed, it is exactly chi-square distributed with 2 degrees of freedom:

   $$\begin{aligned}&||\left(\mathbf{P}_{\overline{\text{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle,|\mathbf{G}_{\tau,a}c_a\rangle,|\mathbf{G}_{\tau,a}s_a\rangle\}} - \mathbf{P}_{\overline{\text{sp}}\{|t^0\rangle,|t^1\rangle,\ldots,|t^m\rangle\}}\right)\sigma|Z\rangle||^2 \\ &\quad\stackrel{d}{=} \sigma^2\chi^2_2, \end{aligned} \quad (64)$$

   where $|Z\rangle$ is a standard Gaussian white noise.

   – The variance of the distribution of the (weighted) smoothed scalogram at (τ, a) is equal to $2||\mathbf{M}'_{\tau,a}\mathbf{K}\mathbf{K}'\mathbf{M}_{\tau,a}||^2_F$, where $||\cdot||_{\text{F}}$ is the Frobenius norm.

   – We approximate the linear combination of the independent chi-square distributions, appearing in Eq. (62), by a gamma-polynomial distribution conserving its first d moments, based on the theory developed in Provost et al. (2009). The formulas are given in Sect. 5.3.3 of Part 1.

   We observe, however, that the convergence of the percentiles (as the number of conserved moments grows) strongly depends on the smoothing coefficient γ, defined in Eqs. (32) and (46). As a general rule, the larger γ is, the faster the convergence. Moreover, it turns out that the gamma-polynomial

[9] The symbol $\stackrel{d}{=}$ means "is equal in distribution".

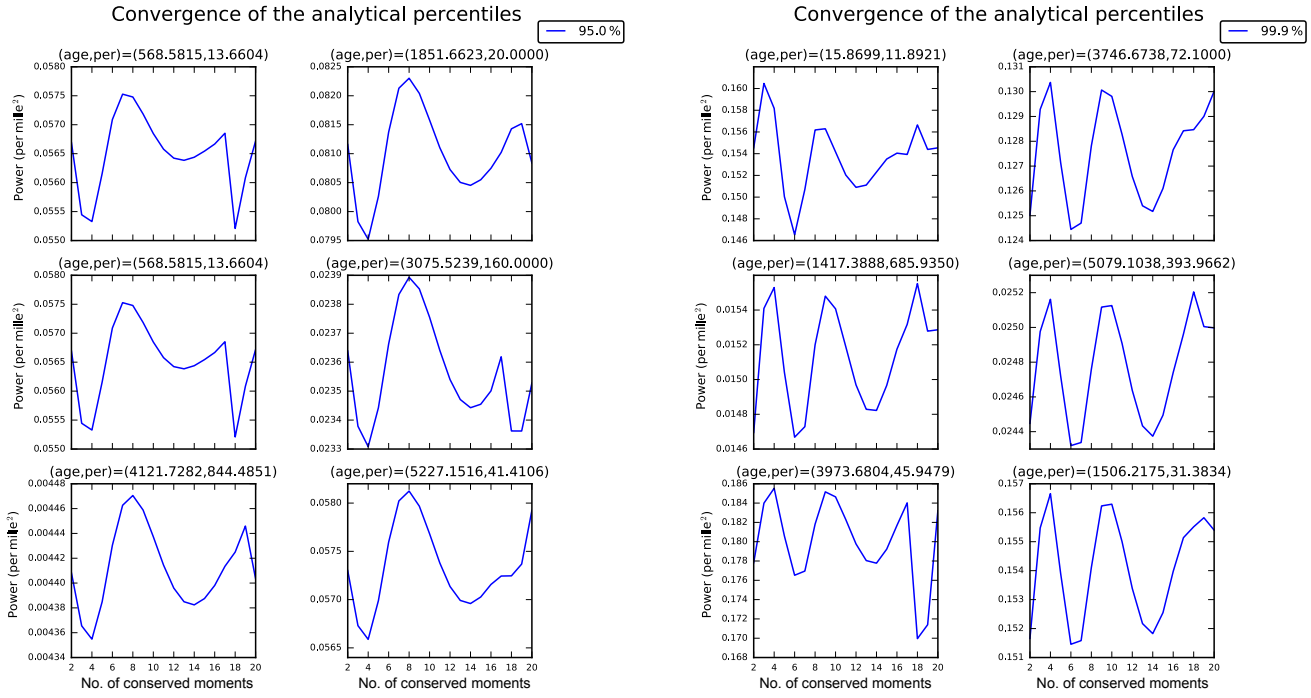


Figure 5. Analytical confidence levels in function of the number of conserved moments, at six particular couples (τ, a), for the scalogram of the time series presented in Sect. 7. Parameter γ is equal to 0.5. **(a)** 95th percentile. **(b)** 99.9th percentile. Slow convergence as well as numerical instabilities (spurious peaks) at high numbers of conserved moments are observed. Convergence cannot therefore be numerically guaranteed.

approximation becomes numerically unstable at large numbers of conserved moments, because the matrix in Eq. (I,100) becomes singular. Consequently, for relatively small values of γ, convergence cannot be numerically guaranteed. This is illustrated in Fig. 5. In such cases, a simple 2-moment approximation is therefore a reasonable choice since it is always numerically stable, it is much quicker than with higher numbers of conserved moments from a computational point of view, and it provides a satisfactory approximation.

3. A comparison between the computing times of the Monte Carlo approach and the analytical approach is presented in Appendix G.

5 Filtering with the amplitude scalogram

5.1 Band filtering

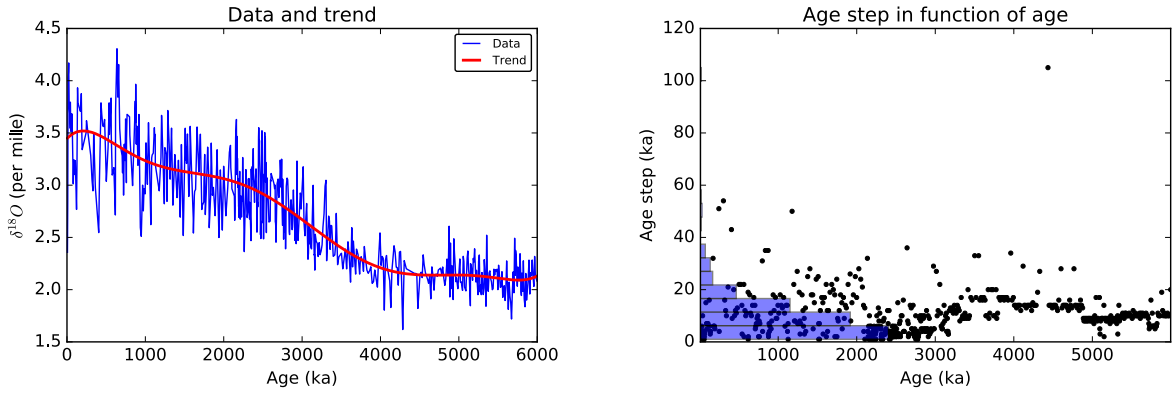
From Sect. 3.5.3, Eq. (43) gives $\widehat{A}_{\tau,a}$ and $\widehat{B}_{\tau,a}$. We can therefore reconstruct the signal $\widehat{A}_{\tau,a}|c_a\rangle + \widehat{B}_{\tau,a}|s_a\rangle$ over the whole time–scale plane, i.e. for all (τ, a). Band filtering is performed by averaging the reconstructed signal between scales $a_{j_{\min}}$ and $a_{j_{\max}}$, namely

$$X_{\text{filt}}(\tau) = \frac{1}{j_{\max} - j_{\min} + 1} \sum_{j=j_{\min}}^{j_{\max}} \widehat{A}_{\tau,a_j}|c_{a_j}\rangle + \widehat{B}_{\tau,a_j}|s_{a_j}\rangle, \quad (65)$$

where the discretised scale is defined in Eq. (57). Such a filtering procedure is a generalisation of the *scale-averaged wavelet power* of Torrence and Compo (1998) which deals with trendless regularly sampled signals. Note that we use the formulas for which there is no smoothing. Indeed, the smoothing procedure in Eq. (45) does not give access to $\widehat{A}_{\tau,a}$ and $\widehat{B}_{\tau,a}$ (only the sum of their squared values is available). An example of band filtering is shown in Figs. 9 and 10.

5.2 Ridge filtering

Consider a signal $|X\rangle = E\cos(\omega|t\rangle + \phi)$. We can easily reconstruct the signal from the estimated amplitudes $\widehat{A}_{\tau,a}$ and $\widehat{B}_{\tau,a}$, given by Eq. (43), taken at the maximum of the scalogram, in this case at $a = 1/\omega$. More generally, we can reconstruct more complex signals relying on the theory of the *amplitude ridges*, developed for continuous-time signals (Sect. 2.7) and which can approximately be applied to irregularly sampled time series. An example of ridge filtering is shown in Figs. 9 and 10.



(a) (b)

Figure 6. (a) The time series and its 7th-degree polynomial trend. **(b)** The age step, $[t_k - t_{k-1}]$, $\forall k \in 2, \ldots, N$, and its distribution.

6 The global scalogram

Analogously to the *global wavelet spectrum* of Torrence and Compo (1998) for trendless regularly sampled time series, we define here the global scalogram as the scalogram averaged over time. Technically, it is nothing but the smoothed scalogram (Eqs. 32, 45 or 46) with integration over the whole interval of the analysis time τ. We can write the discretised global scalogram under a similar matrix form as in Eq. (59), and find the confidence levels according to Sect. 4. Compared to the periodogram defined in Part 1, which has a fixed bandwidth, the global scalogram has a varying bandwidth with the frequency. From Fig. C1 of Appendix C, we deduce that the global scalogram exhibits a frequency resolution that gets better when the frequency decreases. Examples of global scalograms are given in Sect. 7.

7 Application on palaeoceanographic data

7.1 Preliminary analysis

The time series we use to illustrate the theoretical results is the benthic foraminiferal δ^{18}O record from Jian et al. (2003), which holds 608 data points with distinct ages and covers the last 6 million years. The choice of a CARMA(1,0) process as the background stationary noise, as well as the choice of $m = 7$ for the degree of the polynomial trend, are justified in Sect. 9 of Part 1, in which the same data set is used as an example of frequency analysis. The time series, its trend and its time step are drawn in Fig. 6. We remind the reader that the time series is not detrended before computing the scalogram of the data, but it is detrended before estimating the confidence levels.

7.2 Time–frequency analysis

The weighted smoothed scalogram (Sect. 3.6) and its 95 % analytical and MCMC confidence levels are presented in Fig. 7 with parameter $\omega_0 = 5.5$, and in Fig. 8 with parameter $\omega_0 = 15$. As explained in Sect. 2, increasing ω_0 results in a better frequency resolution and a worse time resolution. In our example, the scalogram with $\omega_0 = 15$ exhibits more clearly the period band around 40 kyr and the changes in amplitude along that band. The form of the SNEZ, which is the black region at the bottom in Figs. 7 and 8, follows from the sampling of the time series presented in Fig. 6b.

The parameters are $\gamma = 0.5$ (smoothing coefficient), 2 is the number of conserved moments in the gamma-polynomial approximation (see the discussion in Sect. 4.2), a fixed-length smoothing per scale (see Appendix D), $\beta = 3$ (half-support of a standard Gaussian function $\exp(-x^2/2)$) and $\delta j = 0.05$ (coefficient for the scale resolution).

7.3 Filtering

As explained in Sect. 5, band and ridge filtering are performed on the unsmoothed amplitude scalogram. This is illustrated in Fig. 9, with a filtering band in the interval [35, 45] kyr and with the ridges. From the whole set of the ridges, we select those in the band [35, 45] kyr in order to make a comparison with the band filtering. Band and ridge filtered signals are shown in Fig. 10. We can see that the amplitude modulations in Fig. 10a and b are consistent. Compared to band filtering, the ridge filtering method has the advantage of representing the signal for which the amplitude scalogram is locally maximal, and also allows the time-varying amplitude of the filtered signal to be reconstructed (in red in Fig. 10b). The drawback is that it rarely delivers a continuous reconstruction with climate data.

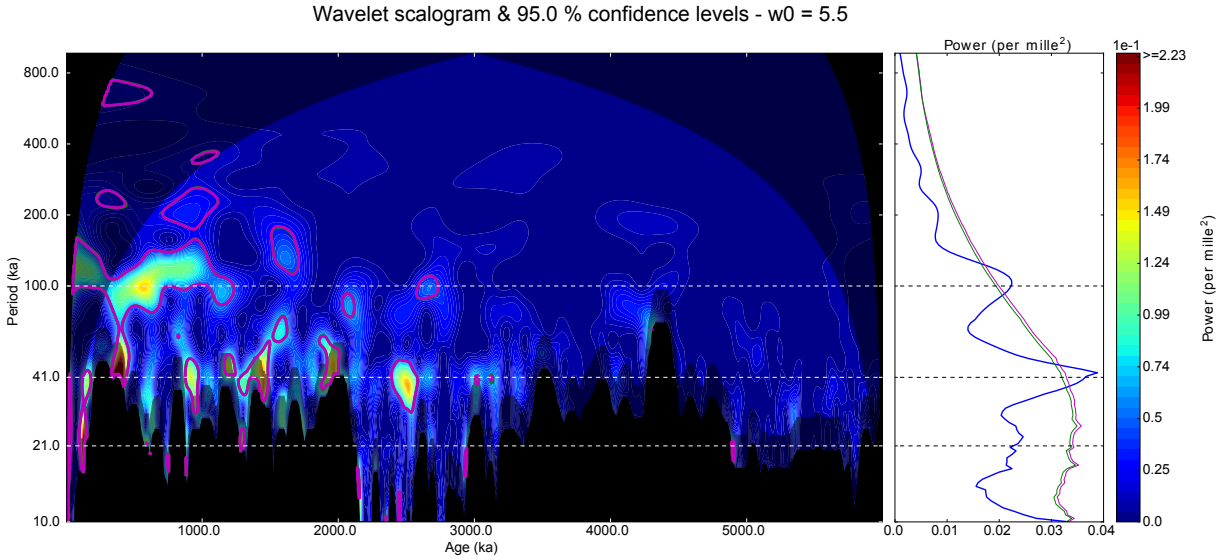


Figure 7. Weighted smoothed scalogram (left) and its global scalogram (right) with $\omega_0 = 5.5$. The 95 % analytical confidence levels (green) and 95 % MCMC confidence levels (magenta), against a red noise background, are also drawn. Note that the green and magenta contours are almost superposed. The two lateral shaded areas are the half-cones of influence, the bottom black area is the SNEZ, and the shaded area above the SNEZ is the refinement of the SNEZ. There are also two lateral black areas, where the scalogram is not computed, because of the fixed-length smoothing per scale. The bounds of the colour scale are the extrema of the scalogram over the non-shaded area. As we work with the weighted scalogram, the power is an estimation of the local squared amplitude. Dashed lines at usual palaeoclimate periods are also drawn.

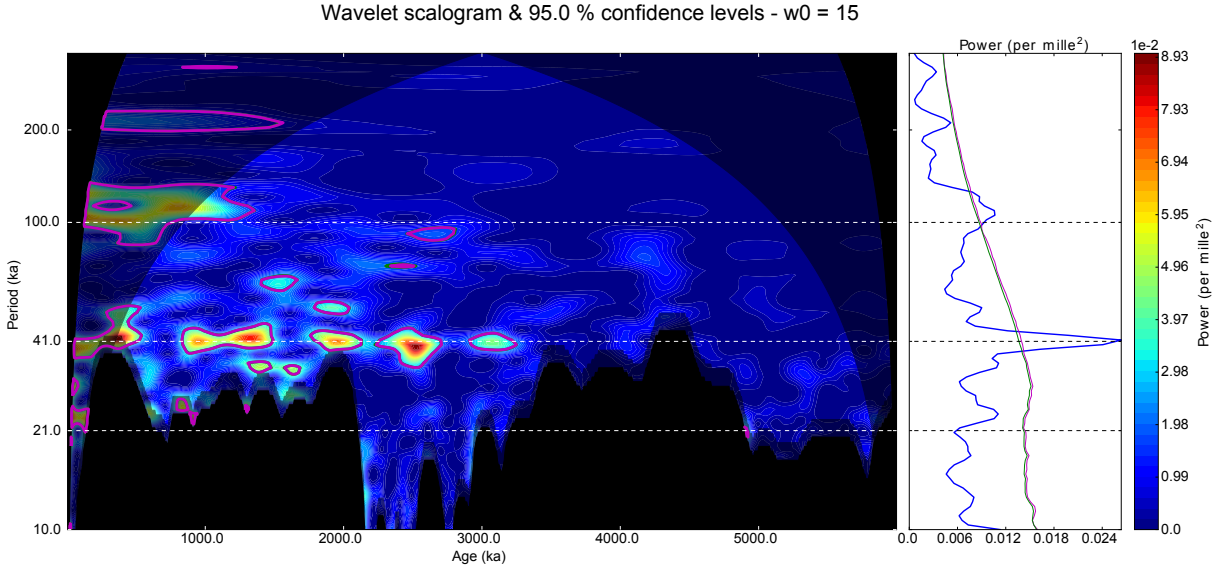


Figure 8. Weighted smoothed scalogram (left) and its global scalogram (right) with $\omega_0 = 15$. The 95 % analytical confidence levels (green) and 95 % MCMC confidence levels (magenta), against a red noise background, are also drawn. Note that the green and magenta contours are almost superposed.

8 WAVEPAL Python package

WAVEPAL is a package, written in Python 2.X, that performs frequency and time–frequency analyses of irregularly sampled time series, significance testing against a stationary Gaussian CARMA(p, q) process, and filtering. Frequency analysis is based on the theory developed in Part 1, and time–frequency analysis relies on the theory developed in this article. It is available at https://github.com/guillaumelenoir/WAVEPAL.

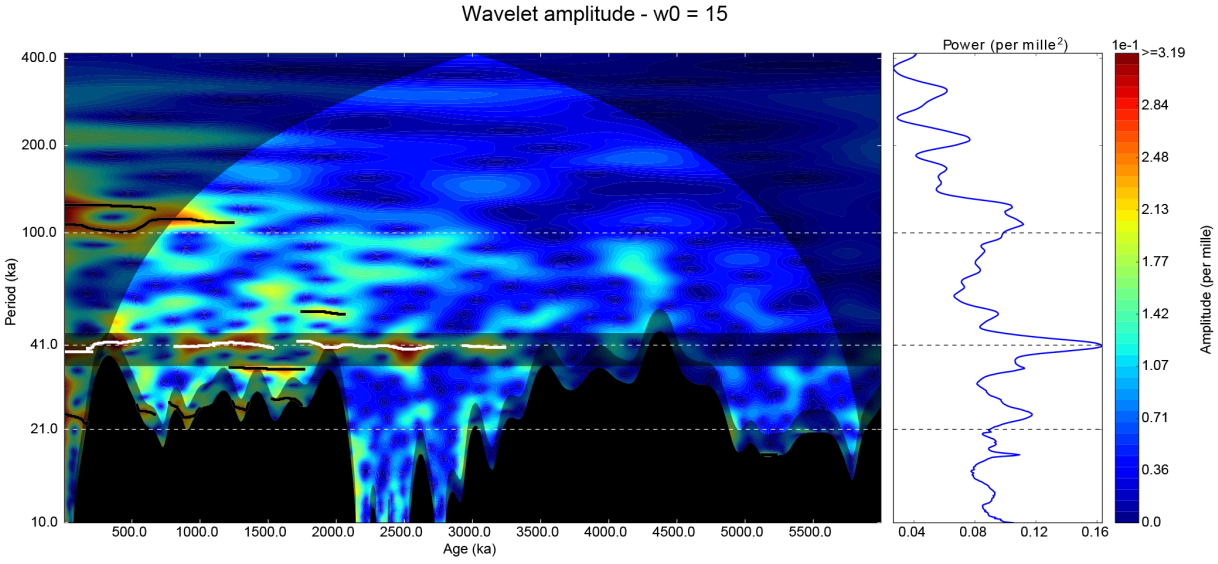


Figure 9. The unsmoothed estimated amplitude (which is the square root of the amplitude scalogram, Eq. 44), jointly with the filtering band in the interval [35, 45] kyr (shaded). Black and white curves are the ridges. They go through the local maxima of the amplitude scalogram. The white ones are the ridges in the band [35, 45] kyr. Parameters are $\omega_0 = 15$, $\beta = 3$ and $\delta j = 0.01$.

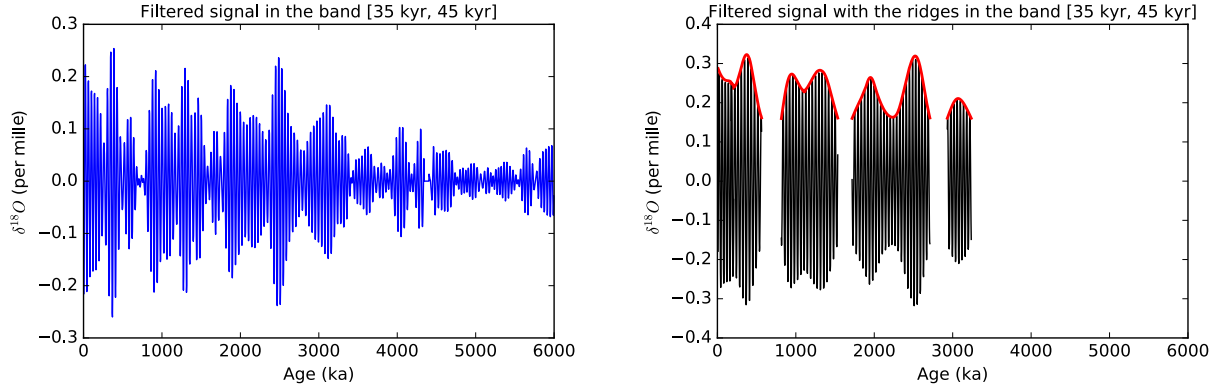


Figure 10. Filtered signal in the band [35, 45] kyr. **(a)** Band filtering. **(b)** Ridge filtering. The red curve is the amplitude of the filtered signal, which is only available with ridge filtering.

9 Conclusions

We defined the scalogram as an extension of the generalised Lomb–Scargle periodogram developed in Part 1. This analysis tool is well suited for irregularly sampled time series which can be modelled as a locally periodic component in the time–frequency plane, plus a polynomial trend, plus a Gaussian CARMA stochastic process. In the particular case of trendless regularly sampled times series, we showed that the unsmoothed scalogram gives the same results as with the traditional algorithms such as in Torrence and Compo (1998). A smoothing procedure, by averaging over neighbouring points in time, was then applied to the scalogram in order to reduce its variance. Besides, we derived estimators of the amplitude of the locally periodic component, based on the general results of Part 1, and proposed an approximate formula linking the scalogram and the squared amplitude. The latter result is at the basis of the weighted smoothed scalogram, which is the analysis tool that we recommend for most time–frequency analyses. We then showed that local aliasing issues may occur in the analysis tools previously derived, implying the delimitation of a forbidden area for the analyses, called the Shannon–Nyquist exclusion zone. Moreover, a test of significance for the scalogram was designed, similarly to its counterpart for the frequency analysis developed in Part 1. Finally, the classical filtering procedures, namely band and ridge filtering, were made available for use with our operator of the estimated amplitude.

Code availability. The Python code generating the figures of this article is available in the Supplement.

Appendix A: Fourier analysis of functions

$\mathbf{L}^2(\mathbb{R})$ is the space of measurable functions on $\mathbb{R}$ with finite energy:

$$||f||^2_{\mathbf{L}^2} = \int_{-\infty}^{+\infty} \mathrm{d}t |f(t)|^2 < \infty. \tag{A1}$$

This defines the squared norm for such functions, that we denote simply by $||f||^2$ in Sect. 2. We provide the $\mathbf{L}^2$ space with the usual inner product:

$$\langle f|g\rangle_{\mathbf{L}^2} = \int_{-\infty}^{+\infty} \mathrm{d}t \overline{f(t)} g(t), \tag{A2}$$

which makes it a Hilbert space; $\overline{f(t)}$ denotes the complex conjugate of $f(t)$, and $\langle f|g\rangle_{\mathbf{L}^2}$ is denoted by $\langle f|g\rangle$ in Sect. 2.

Strictly speaking, the Fourier transform and the convolution product cannot be defined on $\mathbf{L}^2(\mathbb{R})$. We therefore restrict to the Schwartz space, $\mathcal{S}(\mathbb{R})$, which is a subspace of $\mathbf{L}^2(\mathbb{R})$, and on which the Fourier transform and the convolution product can be defined. The Schwartz space is defined as follows:

$$f \in \mathcal{S} \iff \begin{cases} f \in C^\infty, \\ \forall m, j \in \mathbb{N} : |t|^m |f^{(j)}(t)| \to 0 \\ \text{when } |t| \to \infty, \text{ i.e. } f \text{ and all its} \\ \text{derivatives are rapidly decreasing.} \end{cases} \tag{A3}$$

The Fourier transform of $f \in \mathcal{S}(\mathbb{R})$ is defined by

$$\widehat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}t f(t) \exp(-i\omega t), \tag{A4}$$

and $\widehat{f}$ is also in $\mathcal{S}(\mathbb{R})$. The inverse Fourier transform is

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{d}\omega \widehat{f}(\omega) \exp(i\omega t). \tag{A5}$$

Some properties of the Fourier transform are listed below.

- Parseval–Plancherel identities: $\langle f|g\rangle = \langle \widehat{f}|\widehat{g}\rangle$ and $||f||^2 = ||\widehat{f}||^2$.
- Convolution theorem: $[f \star g]\widehat{\ }(\omega) = \sqrt{2\pi}\,\widehat{f}(\omega)\widehat{g}(\omega)$, where the convolution product between f and g is $(f \star g)(t) = \int_{-\infty}^{+\infty} \mathrm{d}t' f(t-t')g(t')$.
- Translation–modulation: the Fourier transform of $f(t-b)$ is $\exp(-i\omega b)\widehat{f}(\omega)$.
- Dilation: the Fourier transform of $f(at)$, $a \neq 0$, is $\frac{1}{|a|}\widehat{f}(\omega/a)$.

Appendix B: Fourier uncertainty principle for the Morlet wavelet

The Fourier uncertainty principle states that the temporal variance and the frequency variance of a function $f \in \mathcal{S}(\mathbb{R})$ satisfy

$$\sigma_t^2 \sigma_\omega^2 \geq \frac{1}{4}, \tag{B1}$$

where

$$\sigma_t^2 = \frac{1}{\sqrt{2\pi}||f||^2} \int_{-\infty}^{+\infty} \mathrm{d}t (t-u)^2 |f(t)|^2, \tag{B2}$$

and

$$\sigma_\omega^2 = \frac{1}{\sqrt{2\pi}||f||^2} \int_{-\infty}^{+\infty} \mathrm{d}t (\omega-\xi)^2 |\widehat{f}(\omega)|^2. \tag{B3}$$

Here, μ and ξ are the average time and average frequency and are defined with the same densities as for the variances. For the Morlet wavelet, the densities are $|\psi_a^\sharp(t)|^2$ (from Eq. 13) and $|\widehat{\psi_a^\sharp}(\omega)|^2$ (from Eq. 16), up to a normalising multiplicative factor. As they are Gaussian functions, their variances are trivial, and we have

$$\sigma_t^2 \sigma_\omega^2 = \frac{\omega_0^2 a^2}{2} \frac{1}{2\omega_0^2 a^2} = \frac{1}{4}. \tag{B4}$$

This is equal to the lower bound of the inequality, as expected for Gaussian functions (see Mallat, 2009, p. 43, for additional details[10]). It is in that sense that the Morlet wavelet is said to be ideally localised.

Appendix C: Uncertainty boxes and scale discretisation

C1 Time–frequency resolution and uncertainty boxes

We saw in Appendix B that the standard deviations of the continuous-time density $|\psi_a^\sharp(t)|^2$ and continuous-frequency density $|\widehat{\psi_a^\sharp}(\omega)|^2$ are $\sigma_t = \omega_0 a/\sqrt{2}$ and $\sigma_\omega = 1/\sqrt{2}\omega_0 a$ respectively. Moreover, the centre angular frequency of $|\widehat{\psi_a^\sharp}(\omega)|^2$ is $\omega = 1/a$, from Eq. (16). With all these coefficients and Eqs. (3) and (5), we can draw rectangles, that we call *uncertainty boxes* (Mallat, 2009, p. 109[11]), in the time–frequency plane indicating the energy spread around each couple (t, ω), or equivalently, the time–frequency resolution at each couple (t, ω). This is illustrated in Fig. C1. Note that their area is equal to $\sigma_t \sigma_\omega = 1/2$ and is therefore constant.

[10] In that book, the Fourier uncertainty principle is called *Heisenberg uncertainty theorem*, but we will not use this misattribution in a non-quantum context.

[11] In that book, the author calls them *Heisenberg boxes*, but as previously mentioned, this is a misattribution in a non-quantum context.

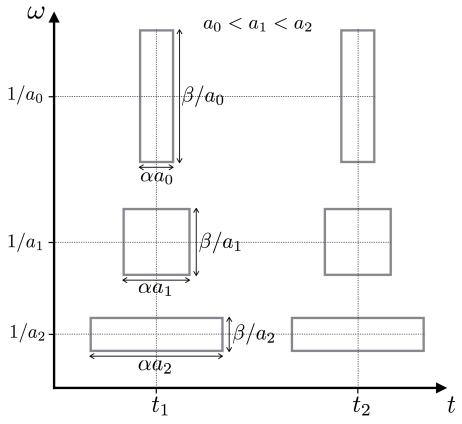


Figure C1. Uncertainty boxes for the Morlet wavelet, with $\alpha = \omega_0/\sqrt{2}$ and $\beta = 1/\sqrt{2}\omega_0$.

C2 Scale discretisation

Scale discretisation is naturally based on the geometry of the boxes. We can, for example, require that the frequency component of the centre of mass of the box corresponding to scale a_j be at the frequency of the border of the box corresponding to scale a_{j+1}. This is illustrated in Fig. C2. We obtain

$$\frac{1}{a_j} = \frac{1}{a_{j+1}} + \frac{\beta}{2a_{j+1}}, \tag{C1}$$

where β is defined in Fig. C1, giving

$$a_{j+1} = \left(\frac{2+\beta}{2}\right) a_j, \tag{C2}$$

and by recurrence,

$$a_{j+1} = \left(\frac{2+\beta}{2}\right)^j a_1, \quad j \in \mathbb{N}. \tag{C3}$$

Multiplying β by a positive factor, γ, allows to control the density of the discretised scales. With variable change $\delta j = \log_2[(2+\beta\gamma)/2]$, we obtain

$$a_{j+1} = 2^{j\delta j} a_1, \quad \delta j > 0,\ j \in \mathbb{N}. \tag{C4}$$

Appendix D: Smoothing the (amplitude) scalogram: technical details

In the formulas of the smoothed (amplitude) scalogram, Eqs. (32), (45) and (46), the integration is in principle performed over the interval $[\tau - \gamma\omega_0 a, \tau + \gamma\omega_0 a]$. When this interval intersects the edges of the time–frequency plane or the SNEZ, we are no longer able to integrate over the full interval. Two choices are then possible.

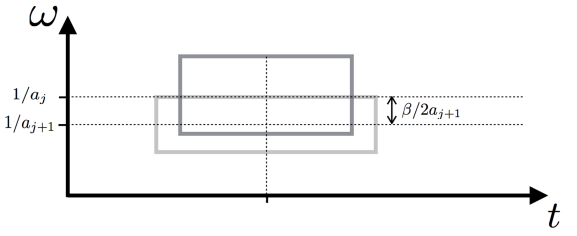


Figure C2. Example of rule for the discretisation of scales taking into account the geometry of the uncertainty boxes. $\beta = 1/\sqrt{2}\omega_0 > 0$.

1. Keep the length of integration equal to $2\gamma\omega_0 a$, and therefore exclude from the analysis some areas of the time–frequency plane. This results in two excluded zones at the time borders of the scalogram and in an extension of the SNEZ.

2. Shorten the interval of integration in order to not exclude from the analysis any extra region of the time–frequency plane.

Both options are available in WAVEPAL and we recommend the first one in order to keep a consistent degree of smoothing at each point (τ, a) in the time–scale plane.

Appendix E: Other studies on the scalogram for irregularly sampled time series

E1 Introduction

We review the only two rigorous studies that we have found in the literature about the estimation of the scalogram for irregularly sampled time series. Like our theory, they are based on the Lomb–Scargle periodogram, and define a kind of scalogram of the CWT for the Morlet mother wavelet. However, these theories are too restrictive to have an interest in geophysical applications.

E2 Foster's theory

E2.1 Introduction

In this section, we derive and comment on the formulas published in Foster (1996c), and based on developments published in Foster (1996a, b). Foster's theory is restricted to the case of the unsmoothed scalogram applied to signals with an additive Gaussian white noise and a piecewise trend for which the shape is the envelope of the Morlet wavelet. It also defines something similar to our amplitude scalogram and generalises the F periodogram of Part 1. We show that some of its formulas can be deduced from our general theory Foster (see 1996c, for the original derivation of the formulas, which is rather different from our approach). Foster's formulas are available for use in a Fortran code provided by the American

Association of Variable Star Observers (AAVSO); see https://www.aavso.org/sites/default/files/software/wwz.tar.gz.

E2.2 Foster's approximation and weighted inner products

Let us start with the approximation made in Foster (1996b) and used in Foster (1996c). Define $\mathbf{U}$ as equal to a full rank real matrix, whose columns are the vectors generating the vector space on which we project the data vector $|X\rangle$, the latter belonging to $\mathbb{R}^N$. Define $\mathbf{G}$ as equal to a real diagonal square matrix of size N with positive elements. Foster's approximation (Foster, 1996b, Eq. 7.9) writes[12]

$$\mathbf{U}'\mathbf{G}^2\mathbf{U} \approx \frac{\mathrm{tr}(\mathbf{G}^2)}{\mathrm{tr}(\mathbf{G})}\mathbf{U}'\mathbf{G}\mathbf{U}. \tag{E1}$$

Note that, when $\mathbf{U}$ is a 2-column matrix holding a cosine vector and a sine vector, the above approximation can also be obtained from Eq. (I,127). The orthogonal projection on the span of $\mathbf{GU}$ thus becomes

$$\begin{aligned}\mathbf{P}_{\overline{\mathrm{sp}}\{\mathbf{GU}\}} &= \mathbf{GU}(\mathbf{U}'\mathbf{G}^2\mathbf{U})^{-1}\mathbf{U}'\mathbf{G} \\ &\approx \frac{\mathrm{tr}(\mathbf{G})}{\mathrm{tr}(\mathbf{G}^2)}\mathbf{GU}(\mathbf{U}'\mathbf{GU})^{-1}\mathbf{U}'\mathbf{G},\end{aligned} \tag{E2}$$

and, for any pair of vectors $|Y\rangle$ and $|W\rangle$ in $\mathbb{R}^N$, we have

$$\langle Y|\mathbf{P}_{\overline{\mathrm{sp}}\{\mathbf{GU}\}}|W\rangle \approx N_{\mathrm{eff}}\frac{\langle Y|\mathbf{GU}}{\mathrm{tr}(\mathbf{G})}(\mathbf{U}'\mathbf{GU})^{-1}\mathrm{tr}(\mathbf{G})\frac{\mathbf{U}'\mathbf{G}|W\rangle}{\mathrm{tr}(\mathbf{G})}, \tag{E3}$$

where $N_{\mathrm{eff}} = \frac{\mathrm{tr}(\mathbf{G})^2}{\mathrm{tr}(\mathbf{G}^2)}$ is defined in Foster (1996c, Eq. 7.7) and called the *effective number of data points*. We can actually rewrite the right-hand side of Eq. (E3) as $N_{\mathrm{eff}}\langle Y|\mathbf{P}_{\overline{\mathrm{sp}}\{\mathbf{U}\}}|W\rangle_{\mathrm{Weighted}}$, where the weighted inner product is defined by:

$$\langle Y|W\rangle_{\mathrm{Weighted}} = \frac{\langle Y|\mathbf{G}|W\rangle}{\mathrm{tr}(\mathbf{G})}, \tag{E4}$$

for any $|Y\rangle$ and $|W\rangle$ in $\mathbb{R}^N$. The term $\langle\cdot|\cdot\rangle_{\mathrm{Weighted}}$ satisfies the requirements of an inner product since the elements of $\mathbf{G}$ are positive (see Brockwell and Davis, 1991, p. 43). Foster's theory is developed in a vector space provided with this weighted inner product.

E2.3 WWT

Now, we derive Foster's scalogram from our theory. The diagonal elements of the weight matrix $\mathbf{G}$ are as follows (Foster, 1996c, Eq. 5-3):

$$\mathbf{G}_{\tau,\omega_{kk}} = \exp\left(-c\omega^2(t_k-\tau)^2\right). \tag{E5}$$

[12] In Foster (1996b), the author works with tensor notations, so that the equivalence is not direct.

Correspondence with our weight matrix, defined in Eq. (25), is performed with the variable changes $\omega = 1/a$ and $c = 1/2\omega_0^2$. Next, consider the formula of the unsmoothed scalogram, Eq. (31), with $a = 1/\omega$, and transformed to accommodate for a trend given by $|\mathbf{G}_{\tau,\omega}t^0\rangle$. This results in

$$||\mathbf{P}_{\overline{\mathrm{sp}}\{|\mathbf{G}_{\tau,\omega}t^0\rangle,|\mathbf{G}_{\tau,\omega}c_\omega\rangle,|\mathbf{G}_{\tau,\omega}s_\omega\rangle\}}|X\rangle||^2 - ||\mathbf{P}_{\overline{\mathrm{sp}}\{|\mathbf{G}_{\tau,\omega}t^0\rangle\}}|X\rangle||^2. \tag{E6}$$

We then make use of the approximation of Eq. (E2) with $\mathbf{U} = [|t^0\rangle\,|c_\omega\rangle\,|s_\omega\rangle]$ for the first projection and $\mathbf{U} = |t^0\rangle$ for the second projection, resulting in the following formula:

$$N_{\mathrm{eff}}\left(||\mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|c_\omega\rangle,|s_\omega\rangle\}}|X\rangle||^2_{\mathrm{Weighted}} - ||\mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle\}}|X\rangle||^2_{\mathrm{Weighted}}\right), \tag{E7}$$

for which we now work in a vector space provided with the weighted inner product. If $|X\rangle$ is a zero-mean Gaussian white noise, Formula (E6) follows exactly a chi-square distribution with 2 degrees of freedom multiplied by the variance of the white noise, namely $\sigma^2\chi^2_2$. Consequently, under the null hypothesis that the process is a Gaussian white noise, the following expression,

$$\begin{aligned}&\mathrm{WWT} = \\ &\frac{N_{\mathrm{eff}}}{2\sigma^2}\left(||\mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle,|c_\omega\rangle,|s_\omega\rangle\}}|X\rangle||^2_{\mathrm{Weighted}} - ||\mathbf{P}_{\overline{\mathrm{sp}}\{|t^0\rangle\}}|X\rangle||^2_{\mathrm{Weighted}}\right),\end{aligned} \tag{E8}$$

approximately follows a chi-square distribution with 2 degrees of freedom and expected value 1. Formula (E8) is rigorously the same[13] as the *weighted wavelet transform* (WWT) of Foster (1996c), in which the author estimates σ^2 as

$$\widehat{\sigma}^2 = \frac{N_{\mathrm{eff}}}{N_{\mathrm{eff}}-1}\left(\langle X|X\rangle_{\mathrm{Weighted}} - \langle t^0|X\rangle^2_{\mathrm{Weighted}}\right). \tag{E9}$$

Significance testing against a Gaussian white noise can be therefore be performed with the WWT.

Below, we comment on the WWT and make a comparison with our formulas.

- The WWT is built on the assumption that the time series holds a Gaussian-shaped trend centred at the probed translation time τ, the support of which varies with the probed frequency. This is equivalent to a constant trend in the vector space provided with the weighted inner product. This contrasts with our choice for the trend, Eq. (23), which is independent of the analysis function.

- The WWT under the null hypothesis is only *approximately* chi-square distributed, compared to Formula (E6) which is *exactly* chi-squared-distributed.

[13] Note that Eq. (5-10) of Foster (1996c), which is prerequisite for the formula of the WWT, is probably erroneous, making unclear the correspondence with our Eq. (E8). However, the formula given here in Eq. (E8) is strictly the same as the WWT encoded at https://www.aavso.org/sites/default/files/software/wwz.tar.gz.

– The estimation of the variance of the white noise, $\widehat{\sigma}^2$, which is part of the WWT formula, depends on the sampling. However, two realisations of a white noise are uncorrelated regardless of the time step separating them, and the estimation of its variance should thus be independent of the sampling, like in Sect. 5.2.2 of Part 1.

– To our point of view, working with weighted inner products, approximations like in Eq. (E1) and complicated tensor notations (see Foster, 1996b) does not bring a simple and unified view of the problematic.

E2.4 WWA

The *weighted wavelet amplitude* (WWA), defined in Foster (1996c, Eq. 5–14), is similar to our amplitude scalogram defined in Eq. (44). The former is obtained from the latter taking the trend to be $|\mathbf{G}_{\tau,\omega}t^0\rangle$, where $\mathbf{G}_{\tau,\omega}$ is defined in Sect. E2.3. For practical applications, we note that computing the inverse of a matrix is needed for the computation of the WWA (this is also the case for our amplitude scalogram). But Foster's theory lacks an in-depth consideration of aliasing issues, and the WWA at some points of the time–frequency plane may be numerically infinite due to the occurrence of singular matrices caused by aliasing.

E2.5 WWZ

Under the null hypothesis that the data $|X\rangle$ is a Gaussian white noise, its squared norm in the vector space provided with the weighted inner product is approximately chi-square distributed with N_{eff} degrees of freedom, as this follows from the 2-moment approximation of Sect. 5.3.3 of Part 1, in which formula (I,98) is applied to matrix $\mathbf{G}$. Consequently, under the null hypothesis, the following formula,

$$\frac{(N_{\text{eff}}-3)\left[||\mathbf{P}_{\overline{\text{sp}}\{|t^0\rangle,|c_\omega\rangle,|s_\omega\rangle\}}|X\rangle||^2_{\text{Weighted}} - ||\mathbf{P}_{\overline{\text{sp}}\{|t^0\rangle\}}|X\rangle||^2_{\text{Weighted}}\right]}{2\left[||X||^2_{\text{Weighted}} - ||\mathbf{P}_{\overline{\text{sp}}\{|t^0\rangle,|c_\omega\rangle,|s_\omega\rangle\}}|X\rangle||^2_{\text{Weighted}}\right]}, \quad \text{(E10)}$$

is approximately equal to the Fisher-Snedecor distribution with 2 and $N_{\text{eff}}-3$ degrees of freedom. Formula (E10) is defined in Foster (1996c, Eq. 5–12) and called the *weighted wavelet Z-transform* (WWZ). It generalises the F periodogram that we defined in Sect. 5.4 of Part 1.

E3 Mathias et al. (2004)'s theory

Mathias et al. (2004) present a formula similar to our amplitude scalogram in the case of a trendless signal. The difference with our Eq. (35) is that they work with a complex exponential, $\exp(i\omega|t\rangle)$, instead of sine and cosine. Switching these terms in our Eq. (35) and taking $a = 1/\omega$ gives the Eq. (17) of Mathias et al. (2004). The authors then approximate the Gaussian shape of the Morlet wavelet by a function with a finite support. Based on that approximation, they develop a fast algorithm for the computation of the scalogram. Apart from this advantage, this study is in fact more restrictive than Foster's theory, since it does not perform significance testing, a zero trend is assumed, no smoothing procedures are considered, and it does not tackle the problem of aliasing issues explained in Sect. 3.8 and 3.10.

Appendix F: Warning about interpolating the time series

This appendix compares the scalograms and their confidence levels in the case of interpolated and non-interpolated time series. The time series we consider is the δ^{18}O signal from the GISP2 ice core (Grootes and Stuiver, 1997), for which the first 11 kyr are removed in order to facilitate the detrending procedure. The time series is drawn in Fig. F1a and b, and its time step is given in Fig. F2. The interpolated time series is built on a time grid with $\Delta t = 30$ yr (this is the smallest time step of the raw time series) in Fig. F1a, or $\Delta t = 300$ yr in Fig. F1b. The (unsmoothed) scalograms with $\omega_0 = 15$ of the raw and interpolated time series are shown in Fig. F3, jointly with the 95 % analytical confidence levels against a red noise. We observe that significance testing is strongly dependent on the interpolation procedure. This is because the parameters of the red noise are badly estimated when the time series is interpolated. Consequently, in general, we cannot rely on interpolated time series to perform significance testing. In particular, we draw the attention on the geological *stacks* (such as in Lisiecki and Raymo, 2005, or Huybers, 2007), which are composed of multiple interpolated time series and averaged together. Significance testing or analysis of the background noise for such time series may therefore be strongly biased.

Finally, we observe that, in this example, the power of the scalogram of the data is weakly affected by the interpolation.

Appendix G: Computing time: analytical versus Monte Carlo significance levels

A comparison between the computing times, for generating the scalogram, with the analytical and with the MCMC confidence levels, based on the hypothesis of a red noise background, is presented in Fig. G1. The computing times are expressed in function of the number of data points, which are disposed on a regular time grid in order to make a meaningful comparison. Confidence levels with the analytical approach are estimated with a 2-moment approximation. The number of samples for the MCMC approach is 10 000 for the 95th percentiles and 100 000 for the 99th percentiles. The smoothing coefficient is $\gamma = 0.5$, and the other parameters are default parameters of WAVEPAL. All the runs were performed on the same computer[14].

[14] CPU type: SandyBridge 2.3 GHz. RAM: 64 GB.

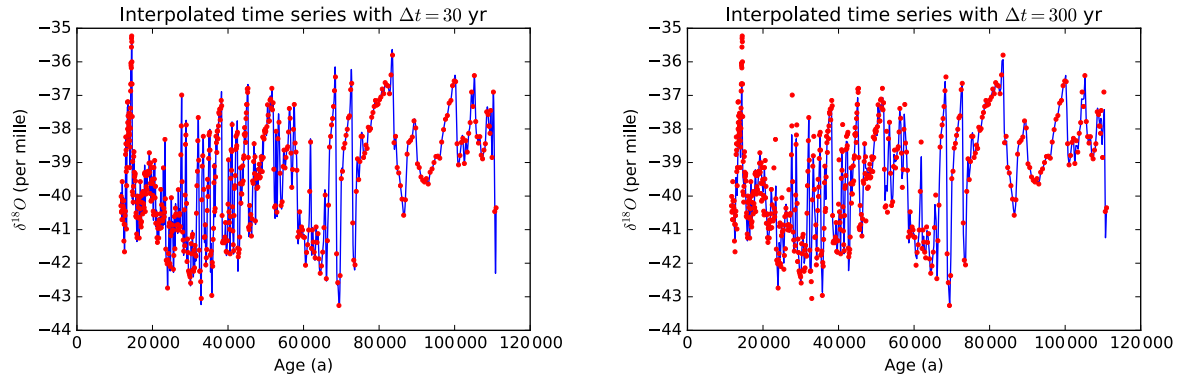


Figure F1. δ^{18}O signal from the GISP2 ice core (Grootes and Stuiver, 1997), for which the first 11 kyr are removed. Raw (red dots) and interpolated (blue line) time series, with **(a)** $\Delta t = 30$ yr and **(b)** $\Delta t = 300$ yr.

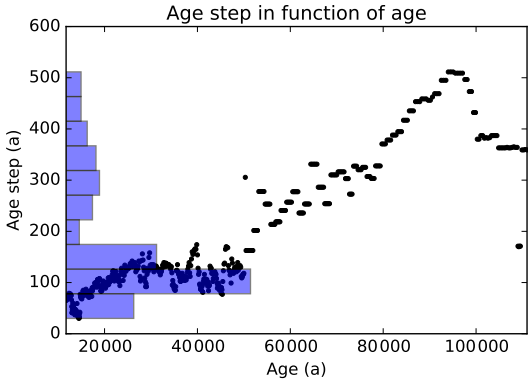
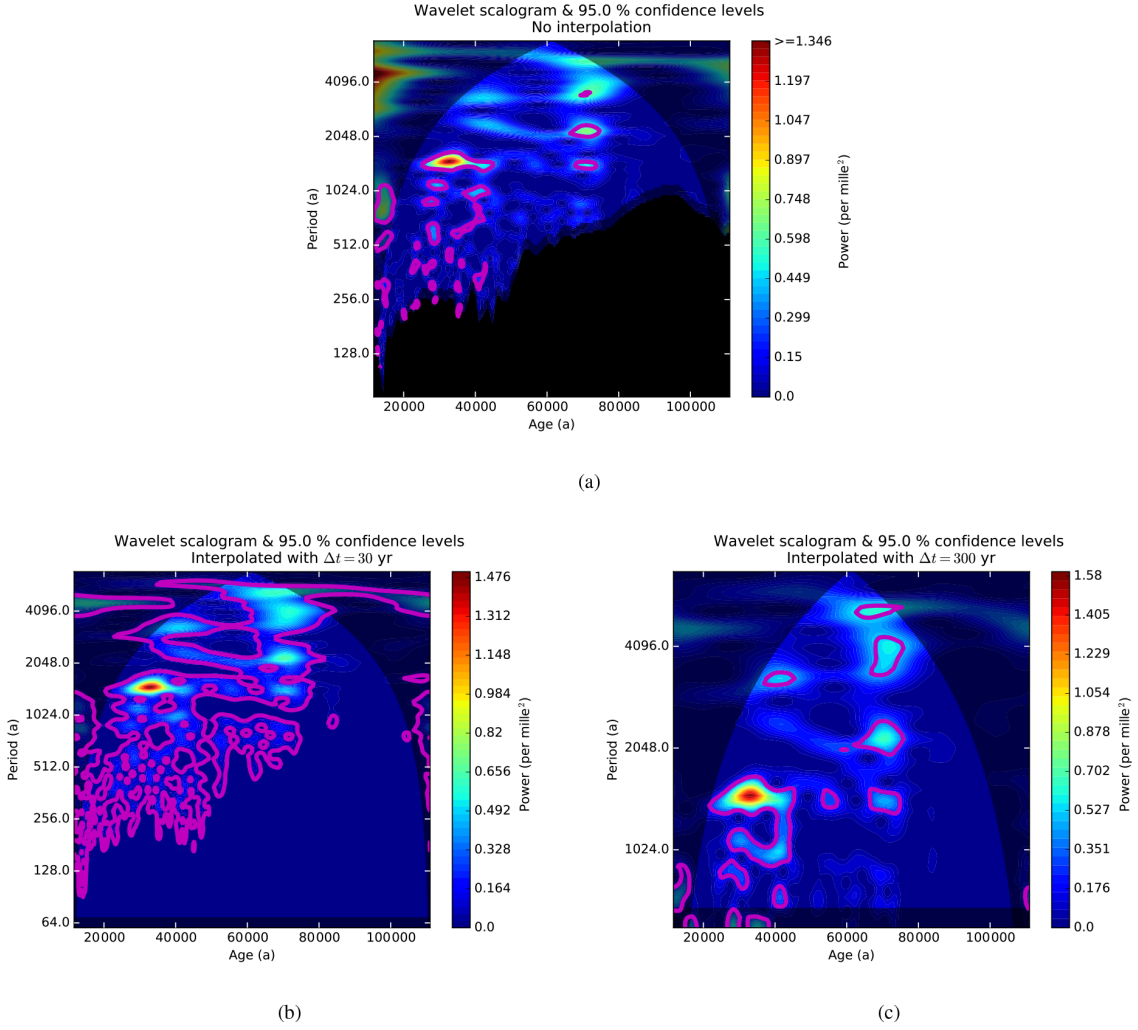


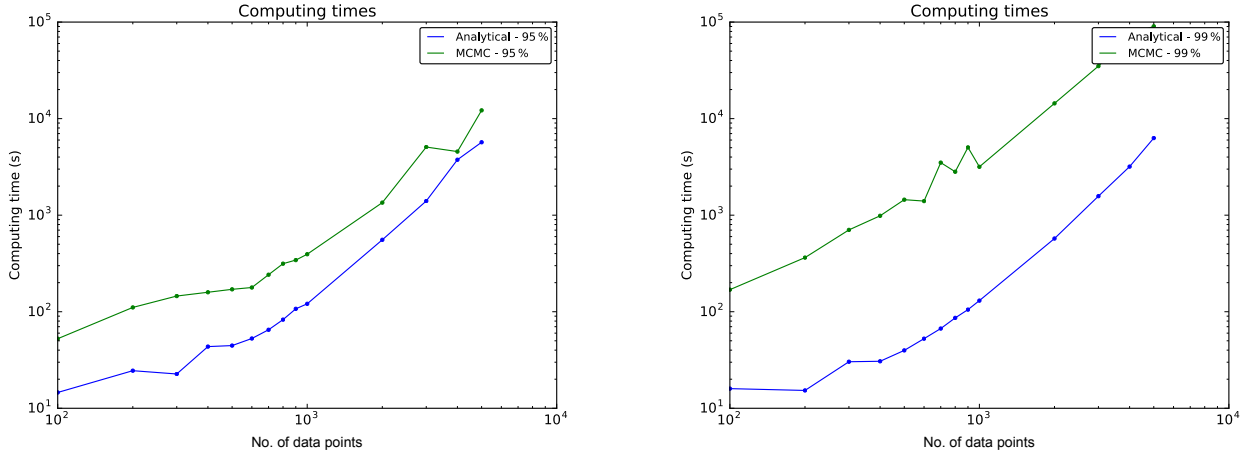
Figure F2. Time step of the δ^{18}O signal from the GISP2 ice core.

With this parametrisation, and within this interval of the number of data points, we see that the analytical approach is faster than the MCMC approach. The analytical approach delivers computing times of the same order of magnitude regardless of the percentile (the two blue curves in Fig. G1a and b are of the same order of magnitude), unlike the MCMC approach, which must require more samples as the level of confidence increases in order to keep a sufficient accuracy. The difference between both computing times therefore increases as the level of confidence increases. Note, however, that the 2-moment approximation, for the estimation of the analytical confidence levels, is very fast from a computational point of view. Increasing the number of conserved moments may considerably increase the computational cost associated with the analytical approach. But this configuration is rarely used in practice because it often results in numerical instabilities and badly estimated percentiles, as explained in Sect. 4.2.



(a) (b) (c)

Figure F3. Scalogram of the time series presented in Fig. F1 and the 95 % analytical confidence levels against a red noise. **(a)** Raw time series. **(b)** Interpolated time series with $\Delta t = 30$ yr. **(c)** Interpolated time series with $\Delta t = 300$ yr.



(a) (b)

Figure G1. Computing times for generating the scalogram with analytical (blue) and MCMC (green) confidence levels, in function of the number of data points (disposed on a regular time grid). Log–log scale. **(a)** 95th percentiles. **(b)** 99th percentiles.

The Supplement related to this article is available online at https://doi.org/10.5194/npg-25-175-2018-supplement.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors are very grateful to Jean-Pierre Antoine, Reik Donner, Laurent Jacques, Lilian Vanderveken, and Samuel Nicolay, for their comments on a preliminary version of the paper. This work is supported by the Belgian Federal Science Policy Office under contract BR/12/A2/STOCHCLIM. Guillaume Lenoir is currently supported by the FSR-FNRS grant PDR T.1056.15 (HOPES).

Edited by: Jinqiao Duan
Reviewed by: two anonymous referees

References

Berger, A. L.: Long-Term Variations of Caloric Insolation Resulting from the Earth's Orbital Elements, Quaternary Res., 9, 139–167, https://doi.org/10.1016/0033-5894(78)90064-9, 1978.

Berger, A. L., Loutre, M. F., and Mélice, J. L.: Instability of the atsronomical periods from 1.5 Myr BP to 0.5 Myr AP, Paleoclimates, 2, 239–280, 1998.

Brockwell, P. and Davis, R.: Time Series: Theory and Methods, Springer Series in Statistics, Second edn., Springer, New York, USA, 1991.

Carmona, R., Hwang, W., and Torresani, B.: Characterization of signals by the ridges of their wavelet transforms, IEEE T. Signal Proces., 45, 2586–2590, https://doi.org/10.1109/78.640725, 1997.

Carmona, R., Hwang, W., and Torresani, B.: Multiridge detection and time-frequency reconstruction, IEEE T. Signal Proces., 47, 480–492, https://doi.org/10.1109/78.740131, 1999.

Cohen, E. and Walden, A.: A Statistical Study of Temporally Smoothed Wavelet Coherence, IEEE T. Signal Proces., 58, 2964–2973, https://doi.org/10.1109/TSP.2010.2043139, 2010.

Delprat, N., Escudie, B., Guillemain, P., Kronland-Martinet, R., Tchamitchian, P., and Torresani, B.: Asymptotic wavelet and Gabor analysis: extraction of instantaneous frequencies, IEEE T. Inform. Theory, 38, 644–664, https://doi.org/10.1109/18.119728, 1992.

Elderfield, H., Ferretti, P., Greaves, M., Crowhurst, S., McCave, I. N., Hodell, D., and Piotrowski, A. M.: Evolution of Ocean Temperature and Ice Volume Through the Mid-Pleistocene Climate Transition, Science, 337, 704–709, https://doi.org/10.1126/science.1221294, 2012.

Foster, G.: Time Series Analysis by Projection. I Statistical Properties of Fourier Analysis, Astron. J., 111, 541–554, https://doi.org/10.1086/117805, 1996a.

Foster, G.: Time Series Analysis by Projection. II. Tensor Methods for Time Series Analysis, Astron. J., 111, 555–566, https://doi.org/10.1086/117806, 1996b.

Foster, G.: Wavelets for period analysis of unevenly sampled time series, Astron. J., 112, 1709–1729, https://doi.org/10.1086/118137, 1996c.

Grinsted, A., Moore, J. C., and Jevrejeva, S.: Application of the cross wavelet transform and wavelet coherence to geophysical time series, Nonlin. Processes Geophys., 11, 561–566, https://doi.org/10.5194/npg-11-561-2004, 2004.

Grootes, P. M. and Stuiver, M.: Oxygen 18/16 variability in Greenland snow and ice with 10^{-3} to 10^{-5} year time resolution, J. Geophys. Res.-Oceans, 102, 26455–26470, https://doi.org/10.1029/97JC00880, 1997.

Grossmann, A. and Morlet, J.: Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape, SIAM J. Math. Anal., 15, 723–736, https://doi.org/10.1137/0515056, 1984.

Holschneider, M.: Wavelets – An analysis tool, Oxford Mathematical Monographs, Oxford University Press, New York, USA, 1995.

Huybers, P.: Glacial variability over the last two million years: an extended depth-derived agemodel, continuous obliquity pacing, and the Pleistocene progression, Quaternary Sci. Rev., 26, 37–55, https://doi.org/10.1016/j.quascirev.2006.07.013, 2007.

Jian, Z., Zhao, Q., Cheng, X., Wang, J., Wang, P., and Su, X.: Pliocene-Pleistocene stable isotope and paleoceanographic changes in the northern South China Sea, Palaeogeogr. Palaeocl., 193, 425–442, https://doi.org/10.1016/S0031-0182(03)00259-1, 2003.

Laskar, J., Robutel, P., Joutel, F., Gastineau, M., Correia, A. C. M., and Levrard, B.: A long-term numerical solution for the insolation quantities of the Earth, Astron. Astrophys., 428, 261–285, https://doi.org/10.1051/0004-6361:20041335, 2004.

Lenoir, G. and Crucifix, M.: A general theory on frequency and time–frequency analysis of irregularly sampled time series based on projection methods – Part 1: Frequency analysis, Nonlin. Processes Geophys., 25, 145–173, https://doi.org/10.5194/npg-25-145-2018, 2018.

Li, T.-H. and Oh, H.-S.: Wavelet spectrum and its characterization property for random processes, IEEE T. Infor. Theory, 48, 2922–2937, https://doi.org/10.1109/TIT.2002.804046, 2002.

Lilly, J. and Olhede, S.: On the Analytic Wavelet Transform, IEEE T. Inform. Theory, 56, 4135–4156, https://doi.org/10.1109/TIT.2010.2050935, 2010.

Lisiecki, L. and Raymo, M.: A Pliocene-Pleistocene stack of 57 globally distributed benthic δ^{18}O records, Paleoceanography, 20, PA1003, https://doi.org/10.1029/2004PA001071, 2005.

Mallat, S.: A Wavelet Tour of Signal Processing, Third edn., Academic Press, Boston, USA, 2009.

Maraun, D. and Kurths, J.: Cross wavelet analysis: significance testing and pitfalls, Nonlin. Processes Geophys., 11, 505–514, https://doi.org/10.5194/npg-11-505-2004, 2004.

Maraun, D., Kurths, J., and Holschneider, M.: Nonstationary Gaussian processes in wavelet domain: Synthesis, estimation, and significance testing, Phys. Rev. E, 75, 016707, https://doi.org/10.1103/PhysRevE.75.016707, 2007.

Mathias, A., Grond, F., Guardans, R., Seese, D., Canela, M., and Diebner, H.: Algorithms for Spectral Analysis of Irregularly Sampled Time Series, J. Stat. Softw., 11, 1–27, https://doi.org/10.18637/jss.v011.i02, 2004.

Meyers, S. D., Kelly, B. G., and O'Brien, J. J.: An Introduction to Wavelet Analysis in Oceanography and Meteorology: With Application to the Dispersion of Yanai Waves, Mon. Weather Rev., 121, 2858–2866, https://doi.org/10.1175/1520-0493(1993)121<2858:AITWAI>2.0.CO;2, 1993.

Provost, S., Ha, H.-T., and Sanjel, D.: On approximating the distribution of indefinite quadratic forms, Statistics, 43, 597–609, https://doi.org/10.1080/02331880902732123, 2009.

Regoli, F., de Garidel-Thoron, T., Tachikawa, K., Jian, Z., Ye, L., Droxler, A., Lenoir, G., Crucifix, M., Barbarin, N., and Beaufort, L.: Progressive shoaling of the equatorial Pacific thermocline over the last eight glacial periods, Paleoceanography, 30, 439–455, https://doi.org/10.1002/2014PA002696, 2015.

Schaefli, B., Maraun, D., and Holschneider, M.: What drives high flow events in the Swiss Alps? Recent developments in wavelet spectral analysis and their application to hydrology, Adv. Water Resour., 30, 2511–2525, https://doi.org/10.1016/j.advwatres.2007.06.004, 2007.

Torrence, C. and Compo, G.: A Practical Guide to Wavelet Analysis, B. Am. Meteorol. Soc., 79, 61–78, https://doi.org/10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2, 1998.

Torrence, C. and Webster, P.: Interdecadal Changes in the ENSO-Monsoon System, J. Climate, 12, 2679–2690, https://doi.org/10.1175/1520-0442(1999)012<2679:ICITEM>2.0.CO;2, 1999.

Witt, A. and Schumann, A. Y.: Holocene climate variability on millennial scales recorded in Greenland ice cores, Nonlin. Processes Geophys., 12, 345–352, https://doi.org/10.5194/npg-12-345-2005, 2005.