Estimation of the significance of the Foster's wavelet spectrum by means of a permutation test and its application for paleoclimate records

J.M. Polanco-Martínez[1,2] and S.H. Faria[1,3]

(1) Basque Centre for Climate Change (BC3),
Edificio Sede N° 1, Planta 1ª, Parque Científico de la UPV/EHU, Barrio Sarriena s/n, 48940, Leioa, Spain.
(2) EPHE, UMR CNRS 5805 EPOC, University of Bordeaux, 33615 Pessac, France.
josue.m.polanco@gmail.com
(3) IKERBASQUE, Basque Foundation of Science, 48013 Bilbao, Spain.
Sergio.faria@bc3research.org

ABSTRACT

Here we propose a permutation test -a non-parametric computing-intensive test- to evaluate the statistical significance of the Foster's wavelet spectrum by means of Monte Carlo simulations. A procedure (algorithm) is introduced in order to carry out this aim. The Foster's wavelet is an adequate method to cope directly with unevenly spaced paleoclimatic time series. We have conducted time series simulations to study the performance of the Foster's wavelet spectrum and applied the permutation test to localize periodic signals known a priori by randomly increasing the fraction of missing data (from 25% to 75% of the total amount of data of an evenly spaced time series). We found that the periodic signals are progressively lost as larger amounts of data are removed. This loss becomes noticeable at circa 50% of the missing data and gets more evident when 75% of the data is removed. The signal loss is more pronounced in the high-frequency range, due to an intrinsic bias in the wavelet spectrum. Notwithstanding, the Foster's wavelet spectrum and the permutation test succeeded in locating the periodic signals, even with a moderate amount of missing data. Finally, we applied our procedure to the oxygen isotope ratio ($\delta^{18}O$) data of the GISP2 deep ice core (Greenland) and we were able to detect the main spectral signature of the unevenly spaced time series of the GISP2 $\delta^{18}O$ record, i.e., the spectral peak around 1,470 years.

Key words: irregular sampling, multi-scale phenomena, non-stationarity, permutation test, weighted wavelet Z-transform

Estimación de la significación estadística del espectro wavelet de Foster mediante el test de permutación y aplicaciones en paleoclima

RESUMEN

En este trabajo se propone un test (no paramétrico pero de cálculo intensivo) de permutación para evaluar la significación estadística del espectro wavelet de Foster mediante simulaciones de Monte Carlo. Un procedimiento (algoritmo) se introduce para llevar a cabo esta tarea. Los wavelet de Foster son un método adecuado para analizar directamente series paleoclimáticas irregulares. Se simularon series temporales con conocida información espectral (eventos periódicos) con el objeto de evaluar el desempeño en la detección de eventos periódicos conocidos a priori, incrementando el porcentaje de datos faltantes de manera aleatoria (desde un 25% a un 75% de la cantidad total de datos de una serie de control equiespaciada). Se encontró que ante la presencia de una gran cantidad de datos faltantes (lo cuál es evidente desde un 50% y más pronunciado con un 75%), las señales periódicas no se detectan en el espectro wavelet. Esto es más evidente en las señales de corta frecuencia debido a la existencia de un sesgo en el espectro wavelet de Foster. Sin embargo, a pesar de lo anterior, el espectro wavelet de Foster en combinación con el test de permutación detectan adecuadamente señales periódicas en series sintéticas (simuladas) y reales con una moderada cantidad de datos faltantes. Finalmente, se aplicó este método a valores medios de $\delta^{18}O$ de los testigos de hielo del GISP2 (Groenlandia) y se encontró la principal característica espectral de estos registros, esto es, la señal espectral ca. 1470 años.

Palabras clave: fenómenos multiescala, muestreo irregular, no estacionariedad, transformada Z ponderada de wavelet, test de permutación.

<div align="center">VERSIÓN ABREVIADA EN CASTELLANO</div>

Introducción y metodología

Una herramienta estadística adecuada que enfrenta varias de las particularidades presentes en datos paleo-climáticos (v. gr., no estacionariedad, información multiescala, señal periódica o cuasi-periódica contaminada por ruido, etc.), es la transformada de wavelet (TW). La TW expande una serie temporal en el dominio del tiempo–frecuencia y es capaz de localizar eventos periódicos así como señales transitorias (Torrence and Compo, 1998; Grinsted et al., 2004). Existen básicamente dos tipos de transformada de wavelet: la primera se realiza mediante la transformada continua de wavelet (TCW) y la segunda a través de la transformada discreta de wavelet (TDW) (Torrence and Compo, 1998; Polanco-Martínez and Fernández-Macho 2014; Polanco-Martínez and Faria 2014, 2015). La TDW es una representación compacta de los datos y es útil para la reducción de ruido y compresión de datos, mientras que la TCW es más adecuada para localizar información señal/ruido (Torrence and Compo, 1998; Grinsted et al., 2004). Debido a que en este estudio estamos interesados en la identificación de potenciales eventos periódicos, nos enfocaremos en la TCW.

A día de hoy, existen varias implementaciones computacionales para realizar análisis espectral mediante la TCW, pero la mayoría de estas herramientas no son capaces de trabajar directamente con series temporales no equiespaciadas en el tiempo (Torrence and Compo, 1998; Maraun et al., 2007; Cazelles et al., 2008). Una forma muy común para enfrentar la irregularidad en tiempo de las series paleoclimáticas, es mediante la interpolación de estos datos. Sin embargo, una buena cantidad de trabajos han demostrado que la interpolación en el dominio del tiempo altera el espectro estimado de una serie y debería evitarse (Schultz and Stattegger, 1997; Schultz and Mudelsee, 2002; Mudelsee, 2014). Una forma más adecuada para enfrentar este problema es mediante la transformada wavelet de Foster (transformada Z ponderada de wavelet), la cuál es capaz de trabajar directamente con series temporales irregulares (Foster, 1996a,b).

La transformada wavelet de Foster fue desarrollada originalmente para analizar datos astronómicos (Foster, 1996a,b). Sin embargo, existe una implementación para ser utilizada con datos paleoclimáticos (Witt and Schumman 2005), aunque el software y los detalles técnicos de la implementación no están disponibles. Por otro lado, Witt and Schumann (2005) no incluyeron un operador de suavizado en el dominio del tiempo o de la escala (frecuencia) en el escalograma wavelet (el escalograma wavelet al no estar suavizado en tiempo o escala podría contener falsa información espectral; cf. Percival and Walden, 2006). En un par de trabajos previos, Polanco-Martínez and Faria (2014, 2015) presentaron resultados preliminares de una implementación estadística-computacional para estimar las amplitudes wavelet mediante la transformada wavelet de Foster, pero ellos usaron un test paramétrico de significación estadística (mediante el uso de un modelo AR1), el cuál exige que el ruido de fondo de las series tenga un comportamiento de tipo "rojo" y esto no siempre es así (Braun et al., 2009). Por todas estas razones, en este trabajo, se propone utilizar un test no-paramétrico de significación estadística, conocido como test de permutación, siguiendo la metodología de Pardo-Igúzquiza and Rodríguez-Tovar (2000) and Pardo-Igúzquiza and Rodríguez-Tovar (2012; 2013).

Resultados y discusión

Se realizaron dos experimentos de Monte Carlo para estudiar el desempeño del espectro wavelet de Foster y del test de permutación para localizar señales periódicas conocidas a priori en series simuladas. Para ello, en el primer experimento se creó un conjunto de series temporales sinusoidales equiespaciadas (de longitud N = 500, espaciado entre las muestras de un año y amplitudes unitarias) con periodos de 4, 10 y 20 años y fases aleatorias. Después de ello, se removieron aleatoriamente un 10, 25, 50 y 75% de datos de las series sinusoidales y se estimaron las amplitudes wavelet para cada nueva serie irregular utilizando el algoritmo presentado en la Sect. Procedure to compute the wavelet spectrum. Los resultados más importantes fueron los siguientes: 1) se encontró que, ante la presencia de una alta cantidad de datos faltantes (algo que es evidente desde un 50% y mucho más marcado con un 75% de datos faltantes), las señales periódicas se pierden. 2) Esta pérdida es más notoria en las cortas frecuencias (desde los 4 años hasta cerca de la frecuencia máxima) debido a la existencia de un sesgo en el espectro wavelet de Foster (véanse las Figs. 1-4). Nótese que algún tipo de sesgo siempre está presente en otros tipos de espectros wavelet (Liu et al., 2007). 3) El test de permutación trabaja adecuadamente aunque identifica algunas áreas espectrales como significativas, pero estas áreas están originadas por azar (las cuáles podrían disminuir si se incrementa el nivel de confianza, por ejemplo, a un 99%).

En el segundo experimento se construyó un conjunto de series temporales equiespaciadas (de longitud N = 1000, espaciado entre las muestras de un año y amplitudes unitarias) con periodos de 4 y 10 años y fases aleatorias más ruido blanco gaussiano (con media cero y varianza uno). Se removió aleatoriamente un 25, 50 y 75% de datos de las series sinusoidales equiespaciadas (Fig. 5) y se estimaron la amplitudes espectrales de las series irregulares (Fig. 6). Los resultados se presentan en la Fig. 7. Los eventos periódicos de 4 y 10 años son detectados adecuadamente para todo el intervalo de estudio (Fig. 6 y 7). Nótese que el sesgo del espectro wavelet de Foster también está presente (Fig 6, abajo). Por otro lado, en las amplitudes wavelet para el caso del 25% de datos faltantes (Fig. 7, arriba) se detecta bastante bien el periodo de 10 años, pero el de 4 años es detectado intermitentemente. Con un 50% de datos faltantes (Fig. 7 centro), el único periodo detectado es el de 10 años y, para el caso de un 75% de datos faltantes (Fig. 7, abajo), no se detectan los eventos periódicos conocidos a priori. Esto significa que, a pesar de que los wavelet de Foster fueron diseñados para analizar y detectar eventos periódicos en series temporales no equiespaciadas. Esta técnica estadística tiene algunas limitaciones, relacionadas principalmente con la cantidad de datos disponibles y la presencia de un sesgo en el espectro wavelet. Por lo cual, esto debería de tenerse en cuenta en trabajos futuros.

En el caso de estudio con datos reales (Figs. 8, 9 y 10), se utilizan valores medios de $\delta^{18}O$ de GISP2 (Grootes and Stuiver, 1997). Una de las razones del porqué de ello, es que nos permite comparar nuestros resultados con los obtenidos por Witt and Schumann (2005). Como se puede observar en las amplitudes de wavelet (Fig. 11), el primer resultado notable es la reconocida marca espectral de diversos testigos de hielo de GIPS2, i.e., el prominente pico espectral de 1,470 años, el cuál está centrado en el intervalo 37,000–27,000 años B.P. (antes del presente). Este pico espectral está relacionado con procesos climáticos a escala milenaria conocidos como eventos de Dansgaard-Oschager y Heinrich y que tuvieron lugar durante el último periodo glacial (Dansgaard et al. 1993). El pico espectral de 1,470 años centrado entre 37,000–27,000 años B.P. también fue obtenido por Witt and Schumann (2005). Hay otros intervalos en nuestro espectro wavelet donde se localiza señal significativa cercana a los 1,470 años, pero las amplitudes wavelet no son tan intensas y los intervalos temporales son cortos. Otros resultados interesantes de nuestro trabajo es la localización de dos señales espectrales milenarias. La primera, centrada en una banda de 2,000–5,000 años y localizada en el intervalo temporal 48,000–28,000 años B.P., mientras que la segunda centrada en una banda de 2,500–6,250 años cubriendo el intervalo 24,000–14,000 años B.P. Estos resultados, aunque bastante parecidos con los de Witt and Schumann (2005), no siempre están en concordancia. Las diferencias obtenidas entre nuestro trabajo con las de Witt and Schumann (2005), podrían deberse principalmente al tipo de test de significación estadística utilizado y al sesgo que está presente en el espectro wavelet de Foster (Andronov, 1998). Debido a que algún tipo de sesgo puede estar presente en los espectros wavelet, algunos autores sugieren una rectificación del espectro (Liu et al., 2007). Aunque esta tarea va más allá del objetivo de este trabajo, pero debería tenerse en cuenta al estimar las amplitudes wavelet de Foster en trabajos futuros.

Introduction

One of the main approaches to study the climates of the past is the statistical analysis of paleoclimate time series obtained from diverse sources (e.g. marine and lake sediments, speleothems, ice cores, etc.). Paleoclimate time series analysis is a useful mathematical tool for identifying past climate signatures archived in paleodata (Weedon 2003; Mudelsee, 2014). However, paleoclimatic data interpretation is not always straightforward, because paleoclimate time series are generally short and noisy, do not have many elements, are usually unevenly spaced (it is not always possible to control the sampling intervals), may contain periodic and quasi-periodic events or transient signals, and are a composition of numerous packages of information in time-scales ranging from days to millennia (multiscale phenomena) (Weedon, 2003; Grinsted *et al.*, 2004; Mudelsee, 2010; Polanco-Martínez, 2012; Polanco-Martínez and Faria 2014, 2015). A powerful statistical tool for overcoming many of

these complications is the wavelet transform (WT). The WT expands time series into the time–frequency domain and can therefore reveal periodicities and localized–intermittent signals (Torrence and Compo, 1998; Grinsted *et al.*, 2004). There are essentially two approaches to wavelet transforms: the first by means of the continuous wavelet transform (CWT) and the second through the discrete wavelet transform (DWT) (Torrence and Compo, 1998; Polanco-Martínez and Fernández-Macho 2014; Polanco-Martínez and Faria 2014, 2015). The DWT is a compact representation of the data and is particularly useful for noise reduction and data compression, whilst the CWT is better for extracting low signal/noise ratio in time series (Torrence and Compo, 1998; Grinsted *et al.*, 2004). As we are interested in extracting climatic signals contained in time series, in this study we have focused on the CWT.

Currently, there are several computational implementations of wavelet spectral analysis techniques using the CWT (Torrence and Compo, 1998; Maraun *et*

al., 2007; Cazelles *et al.*, 2008). However, these tools are not able to directly handle unevenly spaced time series. A frequently used method to cope with this problem is to interpolate in time the original, unevenly spaced time series, in order to obtain equidistance. However, a great deal of research has shown that interpolation in time domain alters the estimated spectrum of a time series and should be avoided (Schultz and Stattegger, 1997; Schultz and Mudelsee, 2002; Mudelsee, 2014). A more effective way to tackle this problem would be to use the Morlet weighted wavelet Z-transform (MWWZ; Foster's wavelet hereafter), which is able to directly handle unevenly spaced time series (Foster, 1996 a,b).

The Foster's wavelet was originally developed to analyse astronomical data (Foster, 1996 a,b). There is an implementation of MWWZ in the literature that can be used for paleoclimate data (Witt and Schumann 2005), which includes a parametric significance test (1996), but this implementation has several drawbacks. For instance, neither the details of this computational implementation nor the supporting software are currently available, and it does not appear to include an averaging in the time or scale directions (the wavelet scalogram is a raw measurement of the time–frequency variability and could contain false spectral information; cf. Percival and Walden, 2006). In a couple of previous studies, Polanco-Martínez and Faria (2014, 2015) presented a preliminary statistical-computational implementation to estimate the wavelet amplitudes of unevenly spaced paleoclimate time series by means of the MWWZ, but they used a paraδmetric (a first-order auto-regressive model) method for testing the wavelet statistical significance and they assume a red noise background in paleoclimatic time series and this is not always the case (Braun *et al.* 2009). For all these reasons, in this work we propose to use a non-parametric method, known as permutation test, for testing the significance of the Foster's wavelet amplitudes following the methodology of Pardo-Igúzquiza and Rodríguez-Tovar (2000; 2012; 2013). Then we conducted time series simulations to study the performance of the Foster's wavelet spectrum and the permutation test to localize periodic signals known a priori by increasing a percentage of missing data. Finally, we applied our algorithm to a well-known unevenly spaced paleoclimate time series: the oxygen isotope ratio ($\delta^{18}O$) data from the GISP2 deep ice core, Greenland (Grootes and Stuiver, 1997).

Methodological and computational aspects

In this section, we describe the wavelet analysis method via the weighted wavelet Z-transform and introduce an algorithm to estimate the wavelet amplitudes directly from unevenly spaced paleoclimate time series. In order to evaluate the statistical significance of the Foster's wavelet spectrum a permutation test is introduced.

A) The continuous wavelet spectral analysis

The spectral analysis via the continuous wavelet transform (CWT) is a powerful mathematical tool for studying multi-scale phenomena and, more particularly, for dealing with non-stationarity (viz. the time change of statistical properties, such as the mean or the variance), both commonly found in many geophysical, financial, medical, and paleoclimatic time series, amongst others. This tool performs a localized spectral decomposition of a time series by determining the dominant modes of variability, and how these modes change with time and scale (frequency) (Grossman and Morlet, 1984; Torrence and Compo, 1998). The CWT of a time series $f(t)$ with a dilation a and a translation parameter (time shift) τ, with respect to a wavelet mother function $\psi(z)$, is defined through the integral transform (Haubold, 1998)

$$W_f(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t)\psi * ([t - \tau]/a)dt, \qquad a > 0 \tag{1}$$

where $\psi*(z)$ is the complex conjugate of $\psi(z)$.

The most common wavelet function used to perform spectral analysis via CWT is the Morlet wavelet, which is essentially a complex exponential function with a Gaussian decay profile (Grossman and Morlet, 1984). Foster (1996b) proposed a useful simplified version of the Morlet wavelet, called the abbreviated Morlet wavelet

$$\psi(z) = e^{iz - cv^2} \rightarrow \psi([t - \tau]/a) = \exp\left[1\frac{(t - \tau)}{a} - c\frac{(t - \tau)^2}{a^2}\right] \tag{2}$$

The constant c determines how quickly the wavelet decays. In the literature, its value is usually chosen to be $c = (8\pi^2)^{-1}$, which ensures that the exponential term decreases significantly in a single cycle $2\pi a$ (Foster, 1996b; Witt and Schumann, 2005).

The choice of a particular wavelet mother function can influence the time and scale (frequency) resolutions of the time series decomposition (Torrence and Compo, 1998; Mallat, 1999). However, the Morlet wavelet provides a good balance between the dilata-

tion *a* and time localizations (Grinsted *et al.*, 2004; Mi *et al.*, 2005; Polanco *et al.*, 2011), so that it is considered one of the best mother functions in terms of reproducing the frequency decomposition of a time series (Kirby, 2005; Li *et al.*, 2007).

B) The weighted wavelet Z-transform (WWZ)

Following Foster (1996b) and Haubold (1998), we start by proposing a straightforward, discretized version of (1), viz.

$$W_f(a,\tau) = \frac{1}{a^{1/2}} \sum_{\alpha=1}^{N} f(t_\alpha)\psi^* \left(\left[t_\alpha - \tau\right]/a\right) \tag{3}$$

for a time series with N data points $f(t_\alpha)$ recorded at a discrete set of times t_α ($\alpha = 1, 2, \ldots N$). However, it turns out that, for uneven time sampling, the sum (3) is a too naive approximation to the continuous integral (1) and offers unsatisfactory performance. In order to overcome this drawback, Foster (1996b) suggested following an approach similar to that used in discrete Fourier transform (Foster, 1996a; Lomb, 1976; Scargle, 1982) by interpreting the wavelet transform (3) with Morlet wavelet (2) as a weighted projection onto the trial function

$$\phi(t) = e^{i\frac{(t-\tau)}{a}} \tag{4}$$

with statistical weights

$$\omega_\alpha = e^{-c\frac{(t_\alpha - \tau)^2}{a^2}} \tag{5}$$

In the theory of discrete Fourier transform, a projection determines the coefficients y_k of a set of r trial functions $\phi_k(t)$, with $k = 1, 2, \ldots r$, for which the model function

$$y(t) = \sum_{k=1}^{r} yk\phi k(t) \tag{6}$$

fits most closely the time series under study. The best-fit coefficients y_k are determined by the formula

$$yk = \sum_{l=1}^{r} S_{kl}^{-1} < \phi_l \,|\, f > \tag{7}$$

where $S_{kl} = < \phi_k |\, \phi_l >$ is the (super-)S-matrix of the trial functions, S_{kl}^{-1} is its inverse, and

$$< u \,|\, v >= \frac{\sum_{\alpha=1}^{N} \omega_\alpha u(t_\alpha)v(t_\alpha)}{\sum_{\beta=1}^{N} \omega\beta} \tag{8}$$

is the inner product of two functions $u(t)$ and $v(t)$, with ω_α denoting the statistical weight assigned to the α-th data point.

Based on the above procedure, Foster (1996b) introduces the concept of weighted wavelet transform as follows. First, he replaces (4) by a set of three trial functions ($r = 3$) that are better suited for dealing with irregular time spacing, viz.

$$\phi_1(t) = 1(t) \tag{9}$$

$$\phi_2(t) = con\left(\left[t - \tau\right]/a\right) \tag{10}$$

$$\phi_3(t) = \sin\left(\left[t - \tau\right]/a\right) \tag{11}$$

He also defines the power for evaluating the projection statistically as

$$P = \frac{N_{eff}}{(r-1)s_w^2} \left[\sum_{k,l} s_{kl}^{-1} < \phi_k \,|\, f >< \phi_l \,|\, f > - < 1 |\, f >^2\right] \tag{12}$$

with ϕ_k ($k = 1, 2, 3$) given by (9)–(11). The number $r - 1$ (= 2 in this case) describes the degrees of freedom of the power (which is chi-squared distributed), N_{eff} is the effective number of data points, and S_ω^2 is the weighted estimated variance of the time series. The last two are defined by Foster (1996b) as

$$N_{eff} = \frac{\left(\sum_{\alpha=1}^{N} \omega_\alpha\right)^2}{\sum_{\alpha=1}^{N} \omega_\alpha^2} \tag{13}$$

$$S_\omega^2 = \frac{N_{eff}V_f}{N_{eff} - 1} \tag{14}$$

where ω_α is given by (5) and V_f, the weighted variation of the time series $f(t)$, is computed via

$$V_f = <f \,|\, f > - < 1 |\, f >^2 \tag{15}$$

Finally, the weighted wavelet transform (WWT) is defined by

$$WWT(a, \tau) = \frac{(N_{eff} - 1)V_y}{2V_f}$$

(16)

where V_y, the weighted variation of the model function $y(t)$, is calculated in a similar fashion as (15) through the expression

$$V_y = <y \,|\, y> - <1 \,|\, y>^2$$

(17)

For fixed scale factor a and time shift τ, the WWT can be treated as a chi-square statistic with two degrees of freedom and expected value of one (Foster, 1996b; Haubold, 1998). It turns out, however, that the WWT in (16) still has a serious shortcoming: it is very sensitive to the effective number of data N_{eff}, which leads to a shift of the WWT peaks to lower frequencies (at lower frequencies the window is wider, so that more data points can be sampled and N_{eff} becomes larger). To compensate for this fact, Foster (1996b) introduced an improved version of the WWT (16), which is less sensitive to the effective number of data, called the the weighted wavelet Z-transform

$$WWZ(a, \tau) = \frac{(N_{eff} - 3)V_y}{2(V_f - V_y)}$$

(18)

As remarked by Haubold (1998), the statistical behaviour of the WWZ (18) is derived for a projection where the statistical weights are inversely proportional to the variance of the data.

In practice, the WWZ is an excellent locator of the signal frequency, but it is not a suitable measure of amplitude. Fortunately, with a projection it is easy to define the amplitude of the corresponding periodic fluctuation as the square root of the sum of the squares of the expansion coefficients for the sine and cosine functions (cf. Eqs. 6, 7, 10 and 11), which Foster (1996b) calls the weighted wavelet amplitude

$$WWA(a, \tau) = \sqrt{y_2^2 + y_3^2}$$

(19)

Similarly, the wavelet phase can be computed by

$$\phi(a, \tau) = \arctan\left(\frac{y_2^2}{y_3^2}\right)$$

(20)

C) Procedure to compute the wavelet spectrum

In this section we present a numerical procedure to estimate the wavelet amplitudes of unevenly spaced paleoclimate time series. In this study we use just the WWA (Eq. 19) to estimate the wavelet amplitudes as a good approximation of the wavelet spectrum (Witt and Schumann 2005). We do not apply an averaging operator or the WOSA (Welch overlapped segment averaging) procedure to smooth the wavelet spectrum. For this reason, we should be cautious with the estimation of the wavelet spectrum because this could contain false spectral information. The statistical significance of the wavelet amplitudes was tested by means of permutation test at 95% confidence level following Pardo-Igúzquiza and Rodríguez-Tovar (2000; 2012; 2013). We used this non-parametric test since it is a distribution free method for testing hypotheses, and although it is computer-intensive, the permutation test is applicable with any kind of spectral estimator, it can be applied to short and long time series and it is easy to program (Pardo-Igúzquiza and Rodríguez-Tovar 2005).

The procedure (algorithm) is described as follows:
1. Input: $f(t)$ (for $t=1,\dots,N$) the unevenly spaced paleoclimate time series under study (with the mean and linear trend removed).
2. Define some key parameters:
   a) $c = (8\pi^2)^{-1}$, cf. (2);
   b) the frequency range, defined by the minimum *(lofreq)* and maximum *(hifreq)* frequencies. The *lofreq* is defined as $1/(t_N - t_1)$. There are several ways to define a maximum frequency for unevenly spaced time series (Pardo-Igúzquiza and Rodríguez-Tovar 2012; 2013; Mudelsee 2014; Polanco-Martínez 2012). We have used the average Nyquist frequency (Schulz and Stattegger 1997; Schulz and Mudelsee 2002), i.e., *hifreq* $= 1 / (<\Delta t>)$, where $<\Delta t>$ is the mean of the differences of the times of $f(t)$;
   c) the distance $\delta freq$ between successive frequencies can be chosen by practice or experience. There are not independent frequencies in the spectrum estimation for unevenly spaced time series (Schulz and Mudelsee 2002; Mudelsee 2014).
3. Compute the WWA of $f(t)$. For doing this, we used a program in Fortran from the American Association of Variable Star Observers (AAVSO), available at the AAVSO website http://www.aavso.org/software-directory.
4. From the original time series $f(t)$ a random permutation is performed, i.e., $f^*(t)$.
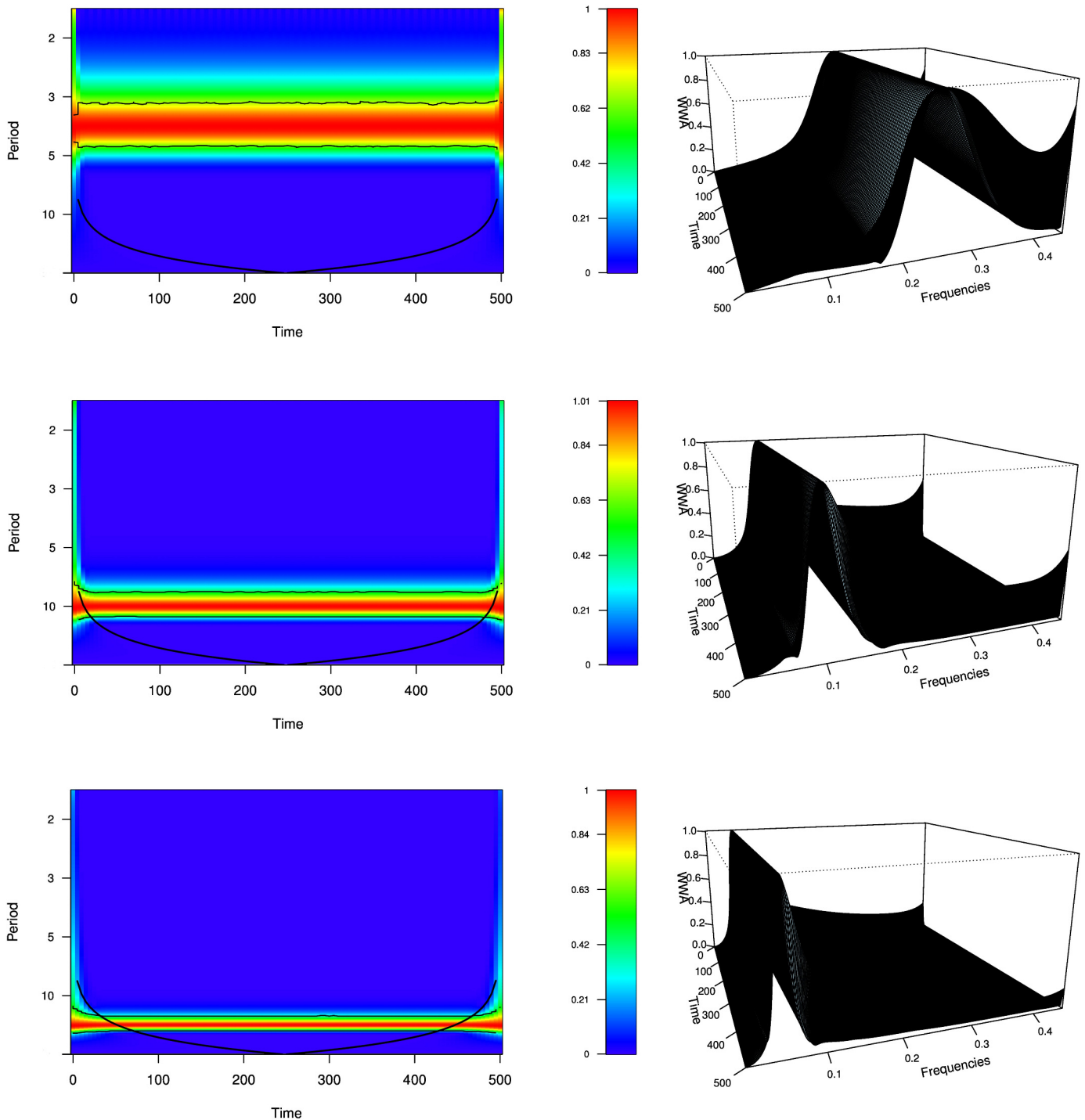
Figure 1. Wavelet spectrum (left) and wavelet amplitudes (right) for the sinusoidal control time series (Δt = 1 year) with periods of 4 (above), 10 (centre) and 20 (below) years. Contour lines indicate wavelet amplitudes that are statistically significant (95% confidence level). The thick black line draws the cone of influence (COI) that delimits the region not influenced by edge effects. The vertical box on the right side indicates the colour code of the spectral amplitudes (values not normalized).
Figura 1. Espectro wavelet (izquierda) y amplitudes wavelet (derecha) para las series sinusoidales (Δt = 1 año) con periodos de 4 (arriba), 10 (centro) y 20 (abajo) años. Las líneas de contorno indican las amplitudes que son estadísticamente significativas (con un 95% de nivel de confianza) de las series sinusoidales simuladas. La curva negra gruesa indica el cono de influencia (COI) y delimita la región no influenciada por los efectos de los bordes. El rectángulo vertical, en el lado derecho, indica el código de colores de las amplitudes wavelet (valores no normalizados).

5. The WWA is estimated for $f^*(t)$.
6. Steps (4) nd (5) are repeated a large number of times N_{sim} (at least 1000).
7. Compute the 95th percentile of the ensemble: WWA[$f^*(t)$; $t = 1$; N_{sim}]
8. Determine whether every spectral point of WWA[$f(t)$] is significantly different from zero (to the 95% confidence level), by comparing it with the 95th percentile of the ensemble.
9. *Output:* The wavelet amplitudes with its significant spectral points.

The fact that paleoclimatic time series have a finite length implies that discontinuities at data boundaries generally distort the wavelet power spectrum in terminal regions, which define a frequency-dependent cone of influence (COI). Within the COI the statistical significance of the wavelet amplitudes (or spectral points) is unreliable. Therefore, the COI imposes a serious limitation in wavelet analysis of time series that must be taken into account. There are many ways of fitting a COI (e.g. Zhang and Moore, 2011; Torrence and Compo, 1998; Meyers *et al.*, 1993). In this study we have used the COI proposed by Torrence and Compo (1998) and used in other computational implementations for wavelet spectrum estimation, such as SOWAS (Maraun 2007; Maraun *et al* (2007).

This algorithm to estimate the wavelet amplitudes via the MWWZ was programmed in the R language (R Development Core Team, 2011) and it currently runs under Linux only. An improved version is now in preparation as an R package to be submitted to the Comprehensive R Archive Network (CRAN) repository. The great advantage of building an R package is that it can be used on most operative systems (R is multiplatform). The Foster's wavelet transform was computed with a program in Fortran (named *wwz11.f*) available on the AAVSO website http://www.aavso.org/software-directory.
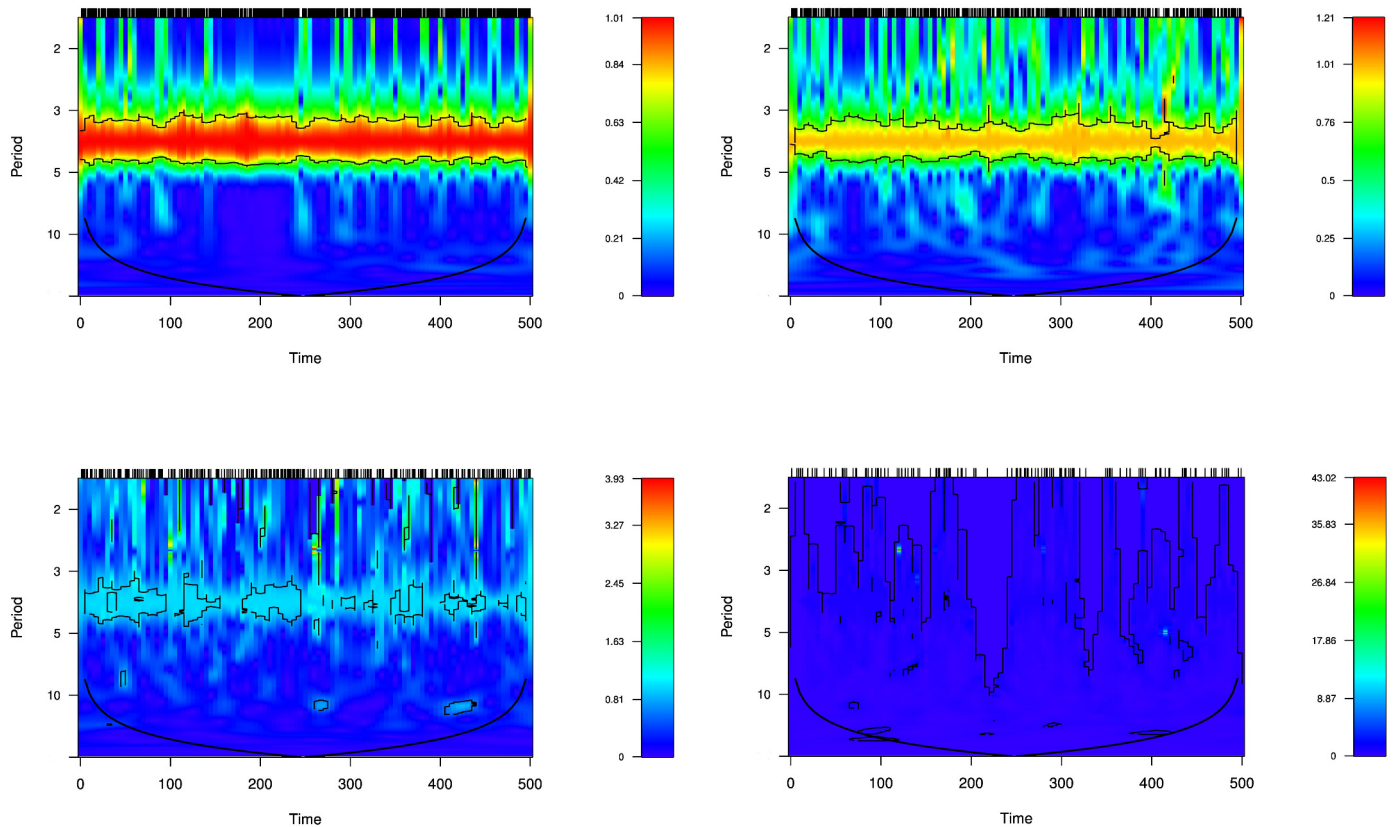

Figure 2. Heat maps of the wavelet amplitudes for the unevenly spaced time series, which have been built by randomly removing 10% (above, left), 25% (above right), 50% (below, left) and 75% (below, right) of data from the sinusoidal time series (Δt = 1 year) with a period of 4 years, amplitude equal to 1 and random phase. The tick marks in the X axis (above) indicate the sampling times.
Figura 2. Mapa de calor de las amplitudes wavelet de las series temporales no equiespaciadas, las cuales se les ha removido aleatoriamente un 10% (arriba izquierda), 25% (arriba derecha), 50% (abajo izquierda) y 75% (abajo derecha) de datos de las series sinusoidales (Δt = 1 year) con un periodo de 4 años, amplitud unitaria y fase aleatoria. Las marcas en los bordes del eje X (arriba) indican los tiempos de muestreo.
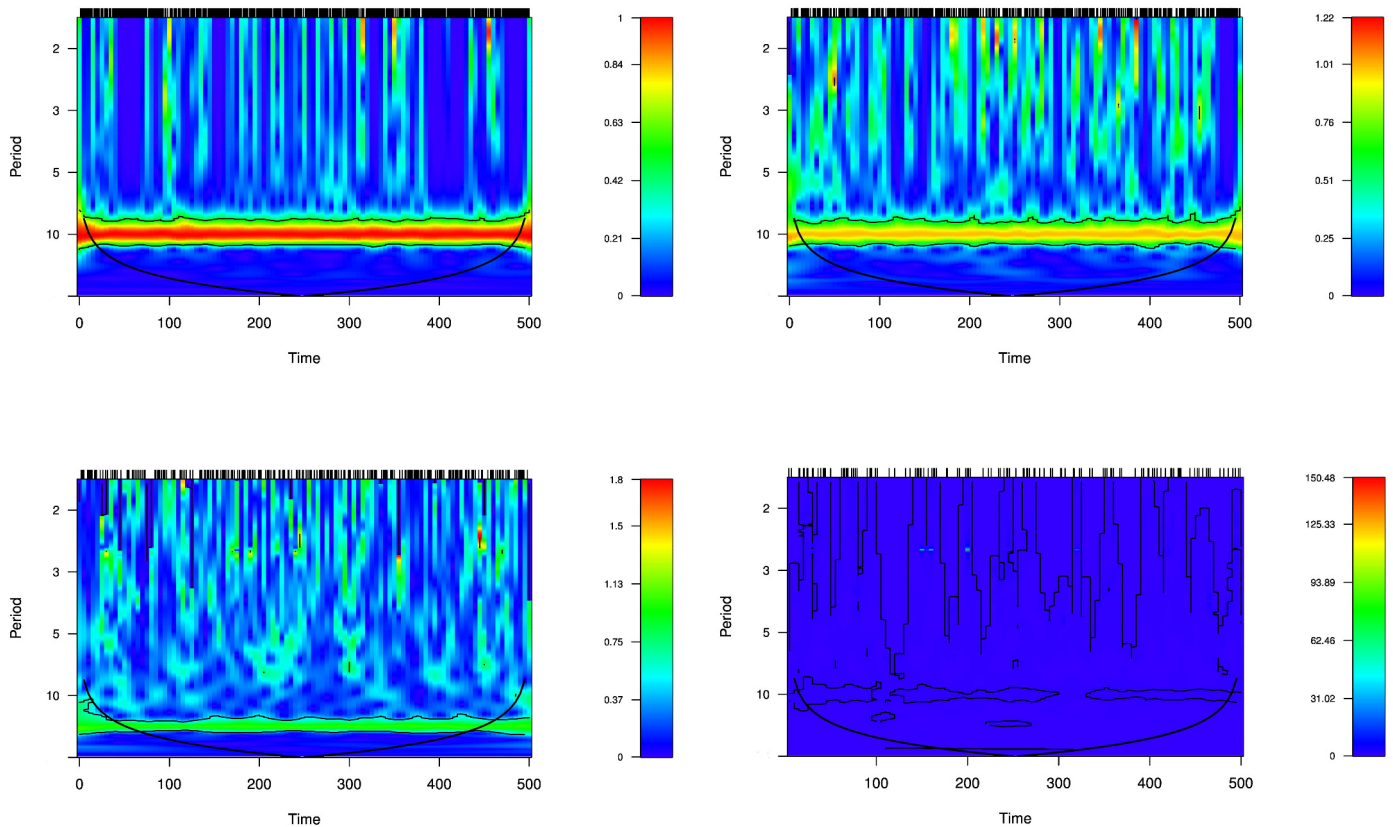
Figure 3. Heat maps of the wavelet amplitudes for the unevenly spaced time series, which have been built by randomly removing 10% (above, left), 25% (above right), 50% (below, left) and 75% (below, right) of data from the sinusoidal time series (Δt = 1 year) with a period of 10 years, amplitude equal to 1 and random phase.
Figura 3. Mapa de calor de las amplitudes wavelet de las series temporales no equiespaciadas, las cuales se les ha removido aleatoriamente un 10% (arriba izquierda), 25% (arriba derecha), 50% (abajo izquierda) y 75% (abajo derecha) de datos de las series sinusoidales (Δt = 1 year) con un periodo de 10 años, amplitud unitaria y fase aleatoria.

Monte Carlo simulation experiments

We conducted Monte Carlo simulation experiments in order to study the performance of the Foster's wavelet spectrum and the permutation test to localize periodic signals known a priori by increasing the fraction of missing data. We generated two sets of simulated cases of study. In the first experiment, we created evenly spaced time series ($N = 500$ and $\Delta t = 1$ year) as control cases, with a unique sinusoidal component (amplitude = 1), periods of 4, 10 and 20 years, and random phases. After that, we randomly removed 10, 25, 50 and 75% of data from the control time series and estimated the wavelet amplitudes for each time series with the algorithm explained in the section *Procedure to compute the wavelet spectrum*. The results are as follows:

1) We found that the periodic signals are progressively lost as larger amounts of data are removed.

This loss becomes noticeable at circa 50% of missing data and gets more evident when 75% of the data is removed. The signal loss is more pronounced in the high-frequency range (from 4 years to higher frequencies).

2) There is an intrinsic bias in the wavelet spectrum, which is most evident in the highest frequencies (Figs. 1–4). The existence of bias is also observed in the estimation of other kinds of wavelet spectrum (Liu *et al.*, 2007).

3) The permutation test performs the identification of periodic signals properly, although some wavelet areas identified as statistical significant are actually created by chance (these areas diminish when a higher confidence level is used, e.g. 99%).

The second experiment was configured in the following way. As a control case we created an evenly spaced synthetic time series ($N = 1000$ and $\Delta t = 1$ year) with two sinusoidal components (with ampli-
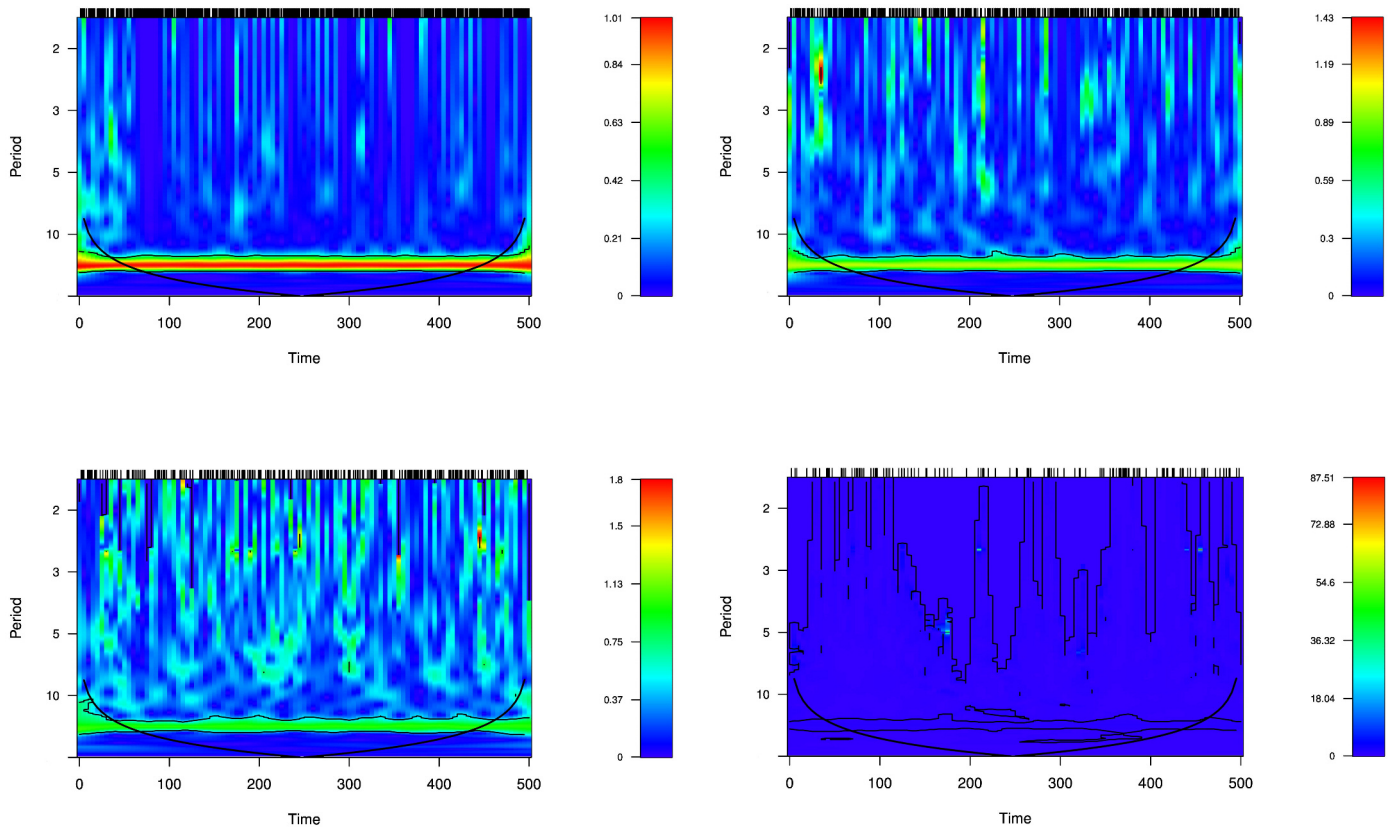
Figure 4. Heat maps of the wavelet amplitudes for the unevenly spaced time series, which have been built by randomly removing 10% (above, left), 25% (above right), 50% (below, left) and 75% (below, right) of data from the sinusoidal time series (Δt = 1 year) with a period of 20 years, amplitude equal to 1 and random phase.

Figura 4. *Mapa de calor de las amplitudes wavelet de las series temporales no equiespaciadas, las cuales se les ha removido aleatoriamente un 10% (arriba izquierda), 25% (arriba derecha), 50% (abajo izquierda) y 75% (abajo derecha) de datos de las series sinusoidales (Δt = 1 year) con un periodo de 20 años, amplitud unitaria y fase aleatoria.*

tudes $A_1 = A_2 = 1$), with periods of 4 and 10 years and with random phases plus Gaussian white noise (mean equal to zero and standard deviation equal to one). After that, we randomly removed 25, 50 and 75% of the data points from the control time series (Fig. 5) and estimated the wavelet amplitudes (Fig. 6). The results are presented in Figure 7. The periodic events 4 and 10 for the control case are detected properly in the whole time interval of this study (Figs. 6 and 7). However, as in the first case of study, we found a conspicuous bias in the wavelet spectrum, that is, the signal at short periods is over-represented, which is more evident in Figure 6 (below), although the wavelet amplitudes have practically the same value for the periods 4 and 10 (Fig. 6, below). On the other hand, in the case of 25% of missing data (Fig. 7, above), the periodic event at 10 years is detected quite well, whereas the period at 4 years is detected intermittently. With 50% of missing data (Fig. 7, cen-

tre), the only period detected is that of 10 years, while in the case of 75% of missing data (Fig. 7, below), all the periodic signals are lost. This means that, despite the ability of the Foster's wavelet to analyse unevenly spaced time series and detect periodic signals in simulated data, there are some caveats, mainly related to the amount of missing data and the presence of a bias in the spectrum. These issues should be considered in future studies.

Case study: δ^{18}O record from GISP2

In order to demonstrate the capability of the permutation test to detect periodical signals in unevenly spaced paleoclimatic records, we applied it to a well-known unevenly spaced paleoclimate time series: the oxygen isotope ratio (δ^{18}O) data of the GISP2 deep ice core, Greenland (Grootes and Stuiver, 1997). The δ^{18}O
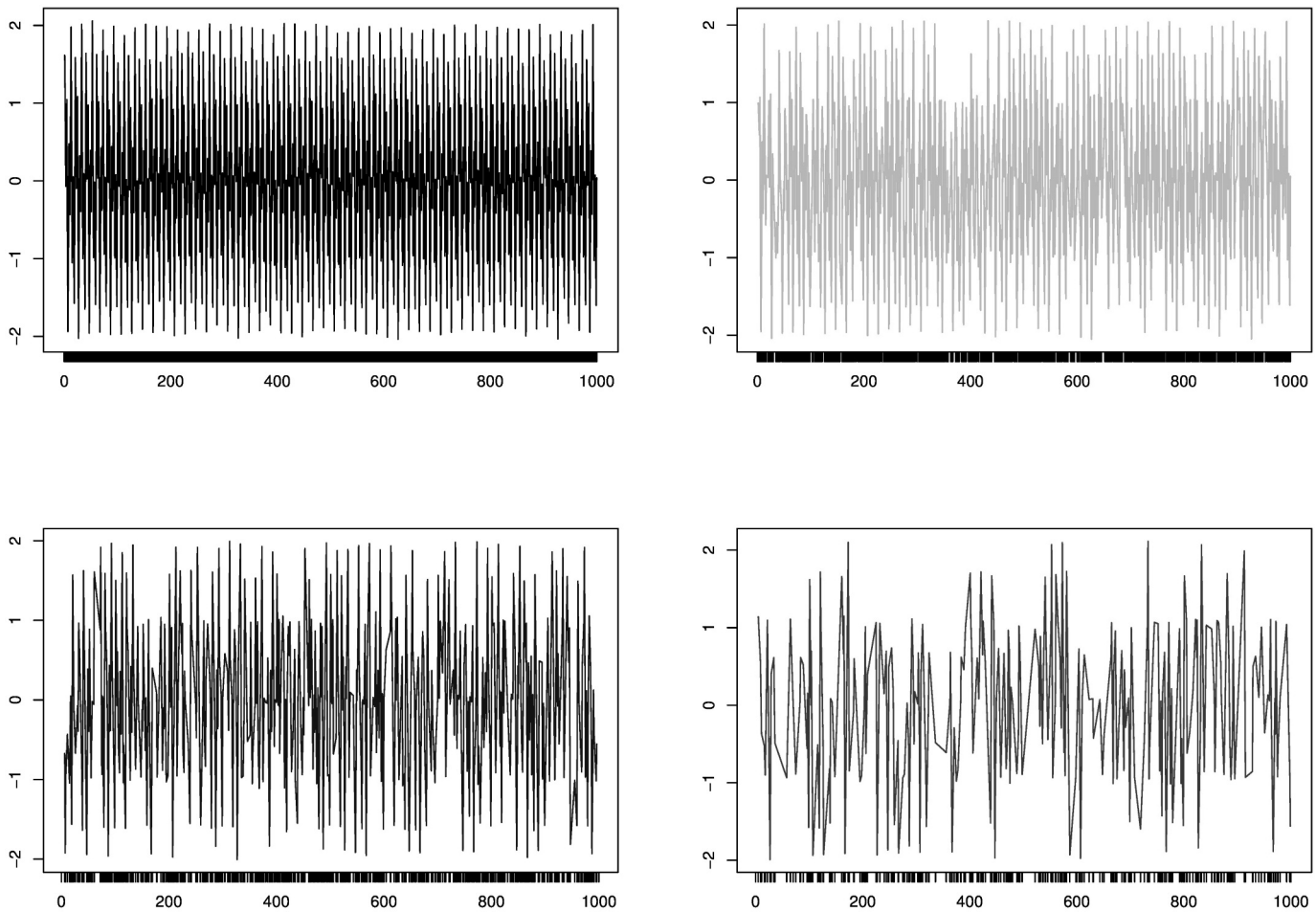
Figure 5. Sinusoidal time series (Δ*t* = 1 year) with periods of 4 and 10 years and random phases (above, left) with 25% (above, right), 50% (below, left) and 75% (below, right) of missing data. The tick marks in the X axis (below) indicate the sampling times.
Figura 5. Series sinusoidales (Δt = 1 año) con periodos de 4 y 10 años y fases aleatorias (arriba izquierda) con un 25% (arriba derecha), 50% (abajo izquierda) y 75% (abajo derecha) de datos faltantes. Las marcas en los bordes del eje X (abajo) indican los tiempos de muestreo.

record from GISP2 allows us to corroborate the relia-bility of our procedure (algorithm) and to compare our results with those by Witt and Schumann (2005). The record was obtained from the National Climatic Data Center, NOAA and covers the last 50,000 years (Fig. 8). It shows strong variability (standard deviation of 2.48), with values oscillating between -43.05‰ and -33.41‰ and a mean value of -36.44‰. Moreover, one of the most remarkable characteristics of this time series is the part that covers the last 10,000 years (covering the major part of the Holocene), which seems to reflect a bi-modal distribution (Fig. 9) and a non-stationary behaviour. Fortunately, wavelet analy-sis is able to tackle both stationary and non-stationary time series.

On the other hand, due to the fact that the GISP2 $\delta^{18}O$ record is an unevenly spaced time series, the dis-tance among elements is not regular: they can take values from 4.74 to 174 years, with a mean of 41.98 years. Figure 10 shows a bimodal distribution, which can be explained by different sampling rates or tem-poral resolutions. There are more samples and much better temporal resolution in the Holocene (last 10,000 years) than in the time period 20,000-50,000 years B.P. These sampling and statistical conditions are a challenge for the computation of the wavelet amplitudes via the MWWZ, as will become clear in the following lines (one of the advantages of this sta-tistical method is that it makes no use of a mean value of the distance among time samplings; rather, it needs just the lowest and highest frequencies, and an "arbitrary" constant resolution among frequencies).
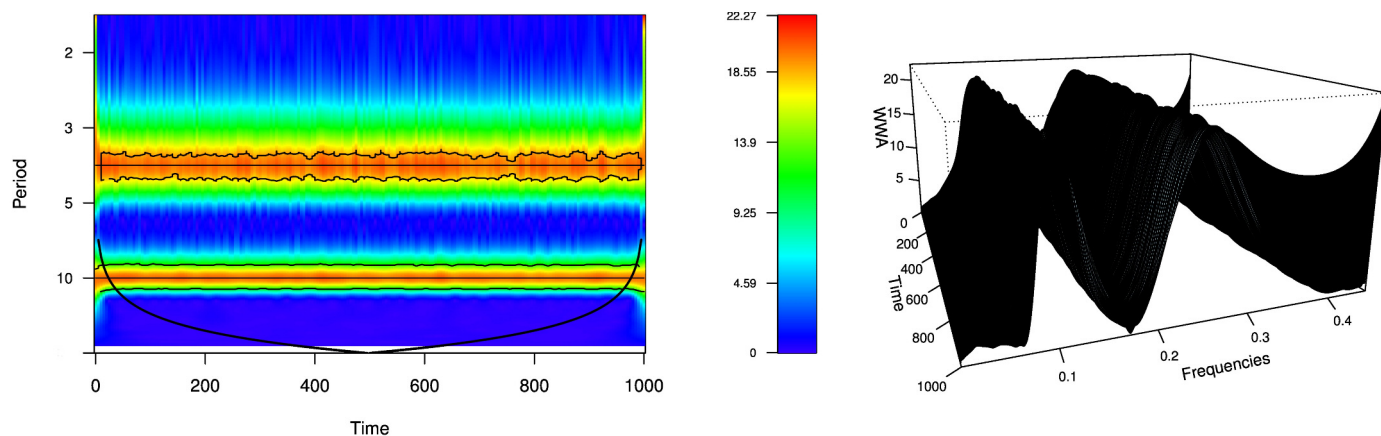
Figure 6. Heat map of the wavelet amplitudes (above) for the sinusoidal control time series (Δ*t* = 1 year) with periods of 4 and 10 years and random phases. Contour lines indicate wavelet amplitudes that are statistically significant (95% confidence level) of the simulated sinusoidal time series. The thick black line draws the cone of influence (COI) that delimits the region not influenced by edge effects. The vertical box on the right side indicates the colour code of the spectral amplitudes (values not normalized). Wavelet amplitudes (WWA) in 3D (below).

Figura 6. Mapa de calor de las amplitudes wavelet (arriba). Las líneas de contorno indican las amplitudes que son estadísticamente significativas (con un 95% de nivel de confianza) de las series sinusoidales simuladas. La curva negra gruesa indica el cono de influencia (COI) y delimita la región no influenciada por los efectos de los bordes. El rectángulo vertical, en el lado derecho, indica el código de colores de las amplitudes wavelet (valores no normalizados). Amplitudes wavelet (WWA) en 3D (abajo).

Figure 11 shows the wavelet amplitudes (wavelet spectrum, not smoothed) via the MWWZ of the GISP2 δ¹⁸O record. The inputs to the program are presented in Table 1. The first remarkable result is that the wavelet amplitudes are able to detect the prominent 1,470 year spectral peak centred within the 37,000–27,000 years B.P. time interval (note that the Fourier spectral analysis is not able to provide the time intervals for the spectral peaks), a well-known spectral signature related to the millennial climate variability associated with Dansgaard–Oschager events (Dansgaard *et al*. 1993) and recorded in the GISP2 ice cores (Grootes and Stuiver, 1997). This spectral signature was also obtained by Witt and Schumann (2005) using the Foster's wavelet transform. There are other time intervals where the 1,470 year cycle is statistically significant, but their wavelet amplitudes are not as strong as in this time interval and the time intervals are not sufficiently long to contain a signal of 1,470 years.

Another interesting result is a couple of conspicuous spectral structures that are statistically significant, located at different frequency bands and time intervals. The first is located in the time interval 48,000–28,000 years B.P. within a period range between 2,000 and 5,000 years, whilst the second lies in the time interval 24,000–14,000 years B.P. within a period range between 6,250 and 2,500 years. These spectral structures are far from being identical, although they share some similarities with the corre-

sponding spectral areas found by Witt and Schumann (2005).

The conspicuous differences obtained with our procedure, in comparison to Witt and Schumann (2005), can be explained fundamentally by two factors: (1) the kind of the statistical significance test used and (2) the existence of an inherent bias in the wavelet spectrum (Andronov, 1998). This kind of bias is also present in any kind of estimation of the wavelet spectrum by means of the continuous wavelet transform. Therefore, it requires a wavelet spectrum bias correction (Liu *et al.*, 2007). In the case of the estimation of the wavelet amplitudes via MWWZ, Andronov (1998) tackled this inconvenience by introducing additional weighting factors in the estimation of the wavelet amplitudes. However, this task is beyond the scope of this study and it will be added to the next version of our statistical-computational software.

The former result and probably the most intriguing is that our procedure was not able to find any spectral structure statistically significant in the time interval corresponding to the Holocene (last 10,000 years). This result is not in concordance with Witt and Schumann (2005), who found centennial and millennial scale variability for some time intervals. A possible explanation could be due to the bi-modality of the δ¹⁸O time series. However, Witt and Schumann (2005) analysed the same time series and time interval. Thus, this is probably not the cause. A more probable expla-
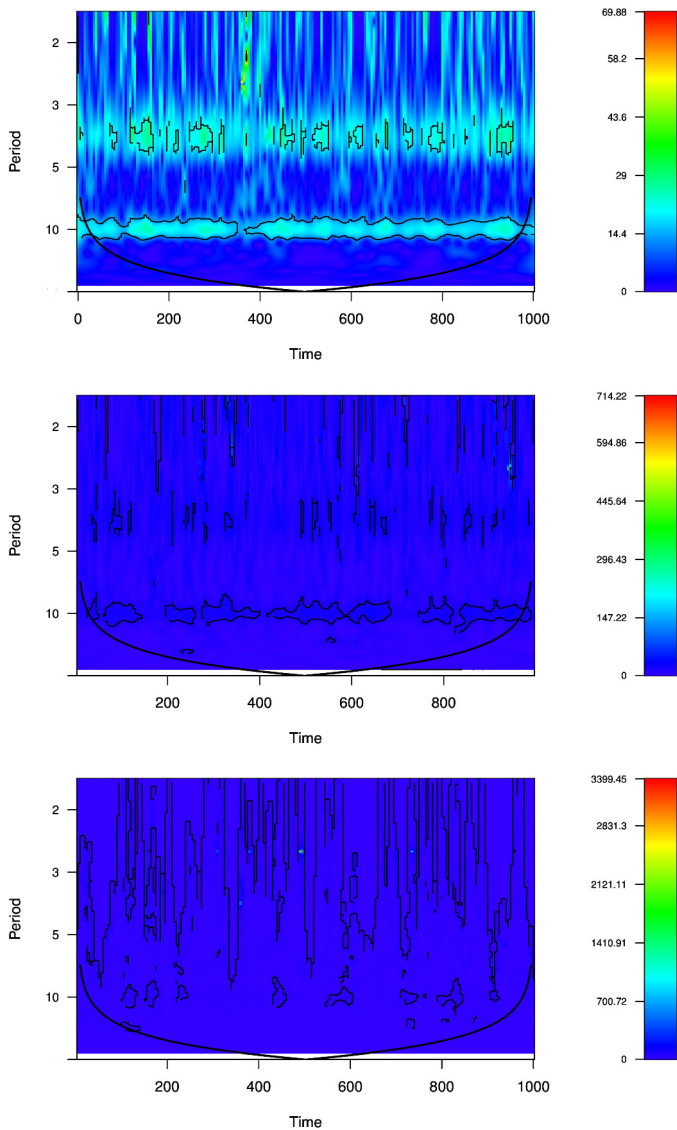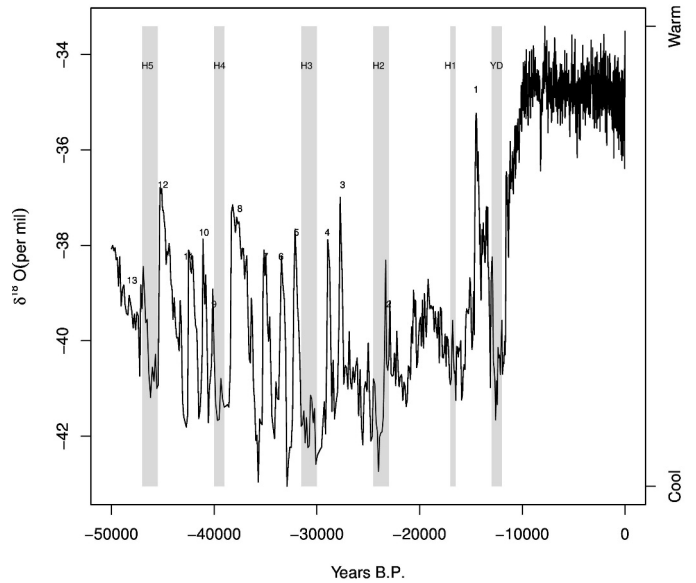
Figure 8. Time series of the GISP2 δ^{18}O (‰) record with 1,192 elements. The grey vertical bars highlight the last five Heinrich events (H1–H5) and the Younger Dryas (YD). The small numbers from 1 to 13 over certain peaks identify Dansgaard–Oeschger (DO) events. The labels "warm" and "cool" at the right side indicate the range of past climates.
Figura 8. Series temporales de los valores de δ^{18}O (‰) de GISP2 (N = 1,192 elementos). Las barras verticales grises identifican los eventos Heinrich (H1–H5) y Younger Dryas (YD). Los pequeños números del 1 al 13 son los eventos Dansgaard–Oeschger (DO). Las etiquetas "warm" y "cool" indican eventos climáticos cálidos y fríos, respectivamente.

Figure 7. Heat maps of the wavelet amplitudes for the simulated sinusoidal time series (Δt = 1 year) with periods of 4 and 10 years and random phases with 25% (above), 50% (centre) and 75% (below) of missing data.
Figura 7. Mapas de calor de las amplitudes wavelet para las series simuladas sinusoidales (Δt = 1 año) con periodos de 4 y 10 años y fases aleatorias con un 25% (arriba), 50% (centro) y 75% (abajo) de datos faltantes.

nation could be related directly with the estimation of the spectral amplitudes and the statistical significance test, but further research is needed to solve this question.

Conclusions

The permutation test —a non-parametric computing-

intensive test— is an adequate tool to detect periodic events registered in simulated data. However, periodic signals are lost in the presence of a large amount of missing data (something that becomes noticeable at 50% of missing data, taking as benchmark an evenly spaced time series). The Foster's wavelet spectrum has an inherent bias, which mainly affects the high frequencies. Our procedure to estimate the wavelet amplitudes via MWWZ successfully detected the main spectral signature of the unevenly spaced time series of the GISP2 δ^{18}O record, viz. the well-known spectral peak around 1,470 years. However, despite this encouraging result, our estimation of the wavelet amplitudes of the GISP2 δ^{18}O record is still not in full concordance with the wavelet amplitudes previously estimated by Witt and Schumann (2005). This means that our current procedure needs some improvements mainly related with the implementation of a window smoothing at the time or scale domain to obtain a smoothed wavelet spectrum, as well as the introduction of extra weighting factors in the estimation of the wavelet spectral amplitudes via the MWWZ (Andronov, 1998). Once these improvements
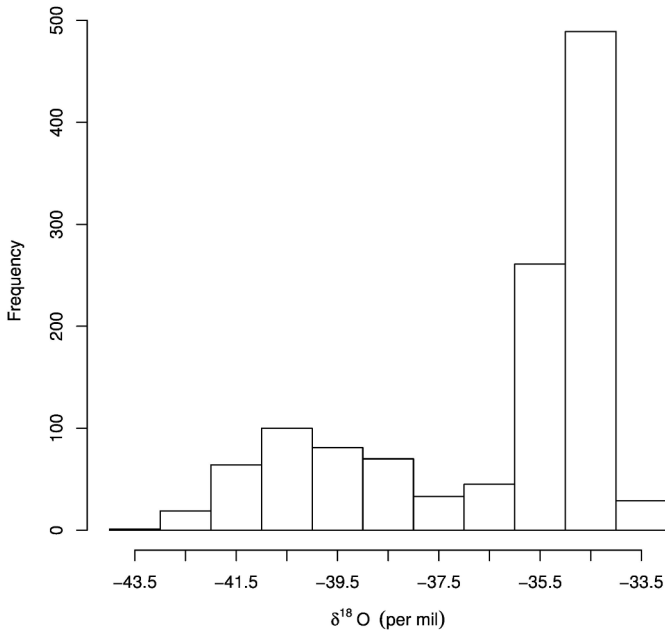
Figure 9. Histogram of the GISP2 δ18O (‰) record.
Figura 9. Histograma del registro de valores medios de δ18O (‰) de GISP2.

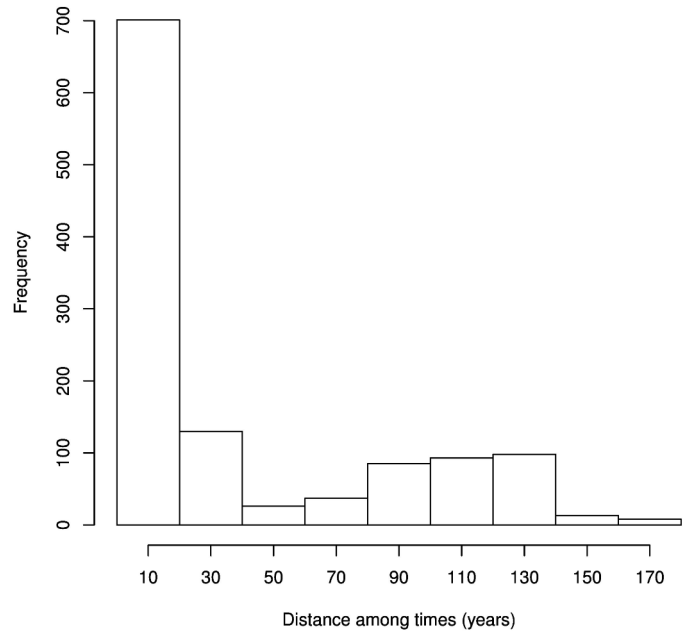**Figure 10.** Histogram of the distances between sampling times in the time series of the GISP2 δ18O (‰) record.
Figura 10. Histograma de las distancias de las razones de muestreo de las series temporales de δ18O (‰) de GISP2.

Time

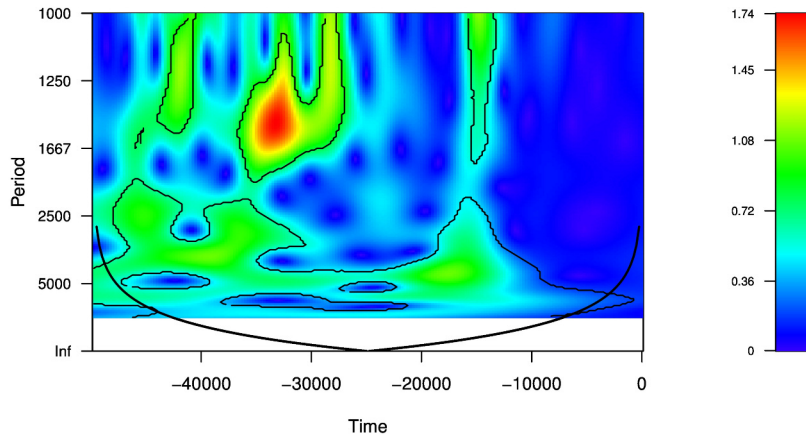**Figure 11.** Wavelet amplitudes of the GISP2 δ18O (‰) record. Contour lines indicate wavelet amplitudes that are statistically significant (95% confidence level). The thick black line draws the cone of influence (COI) that delimits the region not influenced by edge effects. The vertical box on the right side indicates the colour code of the spectral amplitudes (values not normalized).
Figura 11. Amplitudes wavelet de los valores medios de δ18O (‰) de GISP2. Las líneas de contorno indican las amplitudes que son estadísticamente significativas (con un 95% de nivel de confianza). La curva negra gruesa indica el cono de influencia (COI) y delimita la región no influenciada por los efectos de los bordes. El rectángulo vertical, en el lado derecho, indica el código de colores de las amplitudes wavelet (valores no normalizados).

have been carried out, we will include the whole statistical-computational implementation in an R package to be published in the Comprehensive R Archive Network (CRAN), in order to make this implementation freely available for the paleoclimate community around the world.

lofreq (years^{-1})	hifreq (years^{-1})	$\delta freq$ (years^{-1})	c	N_{sim}
10^{-4}	10^{-3}	5×10^{-6}	$(8\pi^2)^{-1}$	2,000

Table 1. The input parameters of the program to compute the wavelet amplitudes of the GISP2 δ^{18}O record (Fig. 8). The labels *lofreq* and *hifreq* indicate the minimum and maximum frequencies of the wavelet amplitudes. The parameter $\delta freq$ defines the distance between frequencies, c is the decay constant of the abbreviated Morlet wavelet (cf. Eq. 2) and N_{sim} is the number of Monte-Carlo simulations (cf. Sect. *Procedures to compute the wavelet spectrum*).

Tabla 1. Parámetros de entrada del programa utilizado para estimar las amplitudes wavelet de los valores medios de δ¹⁸O (‰) de GISP2. lofreq y hifreq indican la frecuencia mínima y máxima donde se estima el espectro wavelet. El parámetro δfreq define las distancias entre las frecuencias, c es la constante de decaimiento (cf. Eq. 2) y N$_{sim}$ es el número de simulaciones de Monte Carlo (cf. Sec. Procedures to compute the wavelet spectrum).

References

Andronov, I.L. 1998. Wavelet analysis of time series by the least-squares method with supplementary weights. *Kinematics and Physics of Celestial Bodies,* 14, 374–392.

Braun, H., Ditlevsen, P., Kurths, J. and Mudelsee, M. 2010. Limitations of red noise in analysing Dansgaard-Oeschger events. *Climate of the Past*, 6(1), 85–92.

Cazelles, B, M. Chavez, D. Berteaux, F. Ménard, J.O. Vik, S. Jenouvrier, and N.C. Stenseth. 2008. Wavelet analysis of ecological time series. *Oecologia,* 156(2), 287–304.

Dansgaard, W, Johnsen, S.J., Clausen, H.B., Dahl-Jensen, D., Gundestrup, N.S., Hammer, C.U. Hvidberg, C.S., Steffensen, J.P., Sveinbjörnsdottir, A.E., Jouzel, J. and Bond, G. 1993. Evidence for general instability of past climate from a 250-kyr ice-core record. *Nature*, 364, 218–220.

Foster, G. 1996a. Time series analysis by projection. I. Statistical properties of Fourier analysis. *The Astronomical Journal,* 111, 541–554.

Foster, G. 1996b. Wavelets for period analysis of unevenly sampled time series. *The Astronomical Journal,* 112, 1709–1729.

Grinsted, A., Moore, J.C., and Jevrejeva, S. 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics,* 11, 561–566.

Grootes, P.M. and Stuiver, M. 1997. Oxygen 18/16 variability in Greenland snow and ice with 10^{-3}- to 10^5-year time resolution. *Journal of Geophysical Research: Oceans,* 102(C12), 26455–26470.

Grossman A. and Morlet, J. 1984. Decomposition of Hardy functions into square integrable wavelets of constant shape. *The SIAM Journal on Mathematical Analysis* 15, 723–736.

Hasselmann, K. 1976. Stochastic climate models Part I. Theory. *Tellus,* 28(6), 473–485.

Haubold, H.J. 1998. Wavelet analysis of the new solar neutrino capture rate data for the homestake experiment. *Astrophysics and Space Science*, 258(1-2), 201–218.

Kirby, J.F. 2005. Which wavelet best reproduces the Fourier power spectrum? *Computers and Geoscience,* 31(7), 846–864.

Kumar, P. and Foufoula-Georgiou, E. 1994. *Wavelet Analysis in Geophysics: an Introduction.* Wavelets in Geophysics, London: Academic Press.

Liu, Y., Liang, X.S. and Weisberg, R.H. 2007. Rectification of the bias in the wavelet power spectrum. *Journal of Atmospheric and Oceanic Technology.,* 24(12), 2093–2102.

Lomb, N.R. 1976. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science,* 39(2), 447–462.

Mallat, S. 1999. *A Wavelet Tour of Signal Processing.* London: Academic press.

Mann, M.E. and Lees, J. M. 1996. Robust estimation of background noise and signal detection in climatic time series. *Climatic Change,* 33(3), 409–445.

Maraun, D.J. 2007. Sowas: Software for Wavelet Analysis and Synthesis, R package version 0.93, http://tocsy.pik-potsdam.de/wavelets/ (Last accessed 22 February 2016).

Maraun, D., Kurths, J., and Holschneider, M. 2007. Nonstationary Gaussian processes in wavelet domain: synthesis, estimation, and significance testing. *Physical Review E,* 75(1), 16707 (1–14).

Meyers, S.D., Kelly, B.G. and O'Brien, J.J. 1993. An introduction to wavelet analysis in oceanography and meteorology: with application to the dispersion of Yanai waves. *Monthly Weather Review,* 121(10), 2858–2866.

Mi, X., Ren, H., Ouyang, Z., Wei, W. and Ma, K. 2005. The use of the Mexican Hat and the Morlet wavelets for detection of ecological patterns. *Plant Ecology,* 179(1), 1–19.

Mudelsee, M. 2002. TAUEST: a computer program for estimating persistence in unevenly spaced weather/climate time series. *Computers & Geosciences,* 28(1), 69–72.

Mudelsee, M. 2014. *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods.* Second Edition. Berlin: Springer.

Pardo-Igúzquiza, E. and Rodríguez-Tovar, F. J. 2000. The permutation test as a non-parametric method for testing

the statistical significance of power spectrum estimation in cyclostratigraphic research. *Earth and Planetary Science Letters*, 181(1), 175–189.

Pardo-Igúzquiza, E. and Rodríguez-Tovar, F. J. 2005. MAX-ENPER: a program for maximum entropy spectral estimation with assessment of statistical significance by the permutation test. *Computers & Geosciences*, 31(5), 555–567.

Pardo-Igúzquiza, E. and Rodríguez-Tovar, F. J. 2012. Spectral and cross-spectral analysis of uneven time series with the smoothed Lomb–Scargle periodogram and Monte Carlo evaluation of statistical significance. *Computers & Geosciences*, 49,φ 207–216.

Pardo-Igúzquiza, E. and Rodríguez-Tovar, F. J. 2013. Análisis espectral de series temporales de variables geológicas con muestreo irregular. *Boletín Geológico y Minero*, 124(2), 319–333.

Percival, D.B. and Walden, A.T.. 2006. *Wavelet Methods for Time Series Analysis.* Cambridge: Cambridge Univ. Press.

Polanco, J., Ganzedo, U., Sáenz, J., Caballero-Alfonso, A.M. and Castro-Hernandez, J.J. 2011. Wavelet analysis of correlation among canary islands octopus captures per unit effort, sea-surface temperatures and the north Atlantic oscillation. *Fisheries Research,* 107(1), 177–183.

Polanco-Martínez, J.M. 2012. *Aplicación de Técnicas Estadísticas en el Estudio de Fenómenos Ambientales y Ecosistémicos.* PhD thesis (in Spanish). Leioa: University of Basque Country (UPV/EHU).

Polanco-Martínez, J.M. and Fernández-Macho, J. 2014. The package W2CWM2C: description, features and applications. *Computing in Science & Engineering*, 16(6), 68–78.

Polanco-Martínez, J.M. and Faria, S.H. 2014. Hunting spectro-temporal information in unevenly spaced paleoclimate time series. BC3 Working Paper Series 2014-07. Basque Centre for Climate Change (BC3). Bilbao, Spain.

Polanco-Martínez, J.M. and Faria, S.H. 2015. Towards a new statistical tool for analyzing unevenly space paleoclimate time series. Proceedings of the International Work-Conference on Time Series Analysis ITISE 2015, pp. 98–107.

R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

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

Scargle, J.D. 1989. Studies in astronomical time series analysis. III- Fourier transforms, autocorrelation functions, and cross-correlation functions of unevenly spaced data. *The Astrophysical Journal,* 343, 874–887.

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

Schulz, M. and Stattegger, K. 1997. SPECTRUM: spectral analysis of unevenly spaced paleoclimatic time series. *Computers and Geosciences,* 23(9), 929–945.

Torrence C. and Compo, G.P. 1998. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society,* 79(1), 61–78.

Weedon, G.P. 2003. *Time-Series Analysis and Cyclostratigraphy: Examining Stratigraphic Records of Environmental Cycles.* Cambridge: Cambridge Univ Press.

Witt, A. and Schumann, A.Y. 2005. Holocene climate variability on millennial scales recorded in Greenland ice cores. *Nonlinear Processes in Geophysics*, 12(3), 345–352.

Zhang, Z. and Moore, J.C. 2011. Improved significance testing of wavelet power spectrum near data boundaries as applied to polar research. *Advances in polar science*, 22(3), 192–198.