Computers & Geosciences 91 (2016) 11–18

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

REDFIT-X: Cross-spectral analysis of unevenly spaced paleoclimate time series

Kristín Björg Ólafsdóttir [a,b,*], Michael Schulz [b], Manfred Mudelsee [a]

[a] Climate Risk Analysis, Heckenbeck, 37581 Bad Gandersheim, Germany
[b] MARUM - Center for Marine Environmental Sciences, University of Bremen, 28334 Bremen, Germany

ARTICLE INFO

Article history:
Received 13 July 2015
Received in revised form
23 December 2015
Accepted 1 March 2016
Available online 4 March 2016

Keywords:
Cross-spectral analysis
Coherency
Phase spectrum
Irregular sampling
Monte Carlo simulations
Paleoclimate time series

ABSTRACT

Cross-spectral analysis is commonly used in climate research to identify joint variability between two variables and to assess the phase (lead/lag) between them. Here we present a Fortran 90 program (REDFIT-X) that is specially developed to perform cross-spectral analysis of unevenly spaced paleoclimate time series. The data properties of climate time series that are necessary to take into account are for example data spacing (unequal time scales and/or uneven spacing between time points) and the persistence in the data. Lomb–Scargle Fourier transform is used for the cross-spectral analyses between two time series with unequal and/or uneven time scale and the persistence in the data is taken into account when estimating the uncertainty associated with cross-spectral estimates. We use a Monte Carlo approach to estimate the uncertainty associated with coherency and phase. False-alarm level is estimated from empirical distribution of coherency estimates and confidence intervals for the phase angle are formed from the empirical distribution of the phase estimates. The method is validated by comparing the Monte Carlo uncertainty estimates with the traditionally used measures. Examples are given where the method is applied to paleoceanographic time series.

1. Introduction

Cross-spectral analysis is often used to estimate the relationship between two time series as a function of frequency. Of particular importance are the coherency and phase spectrum. Coherency is a dimensionless measure on how well two time series co-vary at different frequencies while the phase spectrum shows if the variations happen synchronously at each frequency or if there is a phase difference between them. Cross-spectral analysis is used in climate research to identify joint variability between two variables and to assess the phase (lead/lag) between them.

Paleoclimate proxy time series come from various archives such as marine sediments, ice caps, lake sediments, speleothems, tree rings or corals. Usually the sampling is carried out at constant length intervals and then transferred into the time domain, either by direct dating, or by aligning with other dated time series. Most cross-spectral analysis methods require that the two time series are sampled at identical times and have constant spacing between time points (evenly spaced). This is rarely the case with paleoclimate time series as the archives do normally not accumulate at constant rate, which makes in many cases some kind of interpolation necessary prior to the analysis. Unfortunately, interpolation can bias the spectral results substantially as the spectral power may be shifted from higher to lower frequencies, that is, the spectrum becomes redder (Schulz and Stattegger, 1997). The interpolation can be avoided by estimating the spectrum directly from unevenly spaced time series with the Lomb–Scargle Fourier transform (Lomb, 1976; Scargle, 1982) as done for example in the computer programs by Schulz and Stattegger (1997), Schulz and Mudelsee (2002) and Pardo-Igúzquiza and Rodríguez-Tovar (2012).

Generally, climate time series include persistence (serial correlation) or memory as there is natural inertia in the climate system. Due to the persistence, the spectra of climate time series are characterized by greater amplitude values at lower frequencies (red noise). To distinguish the signals (spectral peaks) in the spectrum of climate time series from background variability they need to be tested against red noise. First-order autoregressive or AR(1) process can be used to model the climate noise (Hasselmann, 1976). The model is normally fitted to the observed time series and the estimated AR(1) parameter is used to form the red noise spectrum (Allen and Smith, 1996). In REDFIT (Schulz and Mudelsee, 2002), the AR(1) parameter is estimated directly from the unevenly spaced time series, so there is no need to interpolate the time series, which can bias the estimated value. The estimated

* Corresponding author at: Climate Risk Analysis, Heckenbeck, 37581 Bad Gandersheim, Germany.
E-mail address: olafsdottir@climate-risk-analysis.com (K. Björg Ólafsdóttir).

http://dx.doi.org/10.1016/j.cageo.2016.03.001
0098-3004/

AR(1) parameter is used to form a theoretical AR(1) spectrum and false-alarm level for testing the significance of spectral peaks via the χ^2 distribution. In addition Monte Carlo simulations can be used, where a large number of red noise processes are generated with the same estimated AR(1) parameter, to form false-alarm level as percentiles of the Monte Carlo ensemble (Schulz and Mudelsee, 2002).

The REDFIT program can only be used for univariate spectral analysis or the autospectrum. Given the need for cross-spectral analysis for unevenly spaced data where the significance is evaluated with Monte Carlo simulations, we present a computer program REDFIT-X, in which cross-spectral analysis has been implemented. Until now the computer program SPECTRUM (Schulz and Stattegger, 1997) has been available for cross-spectral analysis for unevenly spaced climate time series. However the significance measurements in SPECTRUM do not allow for the persistence included in paleoclimate time series. Therefore it is necessary to combine the two approaches, to perform both auto- and cross-spectral analysis with reliable uncertainty estimates.

2. Method

2.1. Cross-spectral analysis – background

2.1.1. Coherency and phase spectrum

The most important features of the cross-spectrum are coherency spectrum and phase between the two signals. Coherency is defined as:

$$c_{xy}^2(f) = \frac{|G_{xy}(f)|^2}{G_{xx}(f)\, G_{yy}(f)}, \tag{1}$$

where $G_{xx}(f)$ and $G_{yy}(f)$ are the autospectra of the signals $x(t)$ and $y(t)$ (with t being time), respectively and $G_{xy}(f)$ is the cross-spectrum between them (e.g., Bendat and Piersol, 2010). It is a dimensionless measure that informs about the degree of linear relationship between two time series, as a function of frequency (f). Coherency is in the range from 0 (no relationship) to 1 (perfect relationship) and can be thought of as a squared correlation coefficient depending on frequency (von Storch and Zwiers, 2003).

Coherency is estimated as:

$$\hat{c}_{xy}^2(f_k) = \frac{\left|\hat{G}_{xy}(f_k)\right|^2}{\hat{G}_{xx}(f_k)\, \hat{G}_{yy}(f_k)}, \tag{2}$$

where $\hat{G}_{xx}(f_k)$ and $\hat{G}_{yy}(f_k)$ are the estimated autospectra and $\hat{G}_{xy}(f_k)$ is the estimated cross-spectrum of two weakly stationary time series $\{t_x(i), x(i)\}_{i=1}^{n_x}$ and $\{t_y(i), y(i)\}_{i=1}^{n_y}$ (n_x and n_y are numbers of data points in each series) (e.g., Bendat and Piersol, 2010). The frequency f_k is in the range from the fundamental frequency $\bar{f} = 1/(n\bar{d})$ to the average Nyquist frequency $f_{\mathrm{Nyq}} = 1/(2\bar{d})$, where n is the number of data points and $\bar{d} = [t(n) - t(1)]/(n-1)$ is the average spacing of the time series (Mudelsee, 2010). Apparently when two time series do not have the same sampling points, the average spacing ($\bar{d}_x$, $\bar{d}_y$) and the fundamental frequency ($\bar{f}_x$, $\bar{f}_y$) for each time series can differ. To ensure that the time series with the lower resolution determines these variables, we use $\bar{d}_{xy} = \max(\bar{d}_x, \bar{d}_y)$, $\bar{f}_{xy} = \max(\bar{f}_x, \bar{f}_y)$ and the average Nyquist frequency is determined as $f_{\mathrm{Nyq}} = 1/(2\bar{d}_{xy})$ (Schulz and Stattegger, 1997).

The auto- and cross-spectra are estimated with the Lomb–Scargle Fourier transform (Lomb, 1976; Scargle, 1982) in combination with the "Welch's Overlapped Segment Averaging" (WOSA) procedure (Welch, 1967) as done in Schulz and Stattegger (1997).

The WOSA segmenting is used to smooth the estimated raw spectrum and make it consistent (the raw spectrum is an inconsistent estimator as the variance does not decrease with increasing data size). The time series of length n_x and n_y are split into a number n_{50} of overlapping segments of length n_{seg}^x and n_{seg}^y (with 50% overlap)

$$\begin{aligned} x_i(j) &= x[(i-1)(n_{\mathrm{seg}}^x/2) + j], \quad j = 1, \ldots, n_{\mathrm{seg}}^x, \\ y_i(j) &= y[(i-1)(n_{\mathrm{seg}}^y/2) + j], \quad j = 1, \ldots, n_{\mathrm{seg}}^y, \end{aligned} \tag{3}$$

where $i = 1, \ldots, n_{50}$. A linear trend is subtracted from each segment to avoid possible artifacts at low frequencies (i.e., resulting from periods, which exceed the segment length). Of course, interpretation of the low-frequency part of a spectrum requires sufficiently long segments. The segments are multiplied by a taper $w(j)$, $j = 1, \ldots, n_{\mathrm{seg}}$ (see different types of spectral windows in Harris, 1978) to reduce spectral leakage and then Fourier transformed

$$\begin{aligned} X_i(f_k) &= \mathcal{F}_{\mathrm{LS}}\{x_i(j)w(j)\}, \\ Y_i(f_k) &= \mathcal{F}_{\mathrm{LS}}\{y_i(j)w(j)\}, \end{aligned} \tag{4}$$

where $\mathcal{F}_{\mathrm{LS}}$ denotes the Lomb–Scargle Fourier transform (Lomb, 1976; Scargle, 1982). Finally the n_{50} segments are averaged to form consistent auto- and cross-spectrum

$$\hat{G}_{xx}(f_k) = \frac{2}{n_{50}\bar{f}_{xy}\, n_{\mathrm{seg}}^x} \sum_{i=1}^{n_{50}} |X_i(f_k)|^2, \tag{5}$$

$$\hat{G}_{yy}(f_k) = \frac{2}{n_{50}\bar{f}_{xy}\, n_{\mathrm{seg}}^y} \sum_{i=1}^{n_{50}} |Y_i(f_k)|^2, \tag{6}$$

$$\hat{G}_{xy}(f_k) = \frac{2}{n_{50}\bar{f}_{xy}\sqrt{n_{\mathrm{seg}}^x n_{\mathrm{seg}}^y}} \sum_{i=1}^{n_{50}} \left[X_i(f_k) Y_i^*(f_k)\right], \tag{7}$$

where $*$ denotes the complex conjugate (Schulz and Stattegger, 1997).

The coherency estimate is biased, where the coherency between two uncoupled time series are expected to be greater than zero (Benignus, 1969). We use the bias approximation from Bendat and Piersol (2010) to form a bias-corrected coherency estimate

$$\begin{aligned} \mathrm{bias}\left[\hat{c}_{xy}^2(f_k)\right] &\approx \left[1 - \hat{c}_{xy}^2(f_k)\right]^2 / n_{\mathrm{eff}}, \\ \hat{c}_{xy}'^2(f_k) &= \hat{c}_{xy}^2(f_k) - \mathrm{bias}\left[\hat{c}_{xy}^2(f_k)\right], \end{aligned} \tag{8}$$

where n_{eff} is the effective number of segments (defined in Section 2.1.2).

The phase spectrum is estimated as

$$\hat{\phi}_{xy}(f_k) = \tan^{-1}\left(\frac{\hat{Q}_{xy}(f_k)}{\hat{C}_{xy}(f_k)}\right), \tag{9}$$

where $\hat{Q}_{xy}(f_k)$ and $\hat{C}_{xy}(f_k)$ are the real and imaginary parts of the estimated cross-spectrum $\hat{G}_{xy}(f_k)$, respectively (e.g., Bendat and Piersol, 2010). We use the four-quadrant inverse tangent function in Fortran (atan2), which returns the result in appropriate quadrant. Therefore the estimated phase angle falls in the range from $[-180°, 180°]$, where zero value means that the two time series are in phase while non zero value means out of phase.

2.1.2. Theoretical uncertainty measurements

A false-alarm level for estimated coherency has been derived from the statistical distribution of coherency estimate, when true coherency equals zero

$$z_{xy}^2 = 1 - \alpha^{1/(n_{\text{eff}}-1)}, \tag{10}$$

where α is the significant level and n_{eff} is the effective number of segments (cf. Carter, 1977; Koopmans, 1995). The statistical distribution of coherency estimate (Goodman, 1957) was derived under the assumptions that the two signals are stationary Gaussian random processes and the data segments are independent (non-overlapping) (Carter, 1977). The overlapping segments are known to reduce the bias and variance of the coherency estimate (Carter et al., 1973), but it also introduces correlation between segments and changes the statistical distribution of the coherency estimate. Therefore it is necessary to use effective number of segments (n_{eff}) instead of the real number of segments (n_{50}) to calculate the false-alarm level for coherency. The effective number of segments for 50% overlap is calculated according to Welch (1967) as $n_{\text{eff}} = n_{50}/(1 + 2c_{50}^2 - 2c_{50}^2/n_{50})$, where c_{50} is a constant depending on the applied spectral window (Harris, 1978). The estimated phase is approximately normally distributed and the standard deviation of the phase estimate is used to form confidence interval for the phase angle. The standard deviation (in radians) is approximated as:

$$\sigma[\hat{\phi}_{xy}(f_k)] \approx \frac{\left(1 - \hat{c}_{xy}^{\prime 2}\left(f_k\right)\right)^{1/2}}{\left|\hat{c}_{xy}^{\prime}\left(f_k\right)\right|\sqrt{2n_{\text{eff}}}} \tag{11}$$

(Bendat and Piersol, 2010). Uncertainties for phase values, $\hat{\phi}_{xy}(f_k) \pm z(\alpha)\cdot\sigma[\hat{\phi}_{xy}(f_k)]$, are derived from percentage points of the normal distribution $z(\alpha)$. Due to the dependency of the standard deviation on coherency, the uncertainty for a phase angle can be very large if coherency is close to zero. Accordingly the phase should only be analysed at frequencies with significant coherency.

2.2. Monte Carlo procedure

2.2.1. Monte Carlo false-alarm level for coherency

We use a Monte Carlo simulation technique to approximate the additional false-alarm level to detect whether there is significant coherency between the two time series at a given frequency. The Monte Carlo false-alarm level is formed from an empirical sampling distribution of the coherency estimator. The sampling distribution is obtained by estimating coherency between many pairs of uncoupled time series formed by Monte Carlo simulations, which mimic key properties of the observed signals, such as autocorrelation and time spacing. The detailed steps of the simulations are as follows:

1. The persistence times τ_x and τ_y are estimated from the two observed time series. They are estimated with the least-squares algorithm TAUEST (Mudelsee, 2002), which fits a first-order autoregressive or AR(1) persistence model to unevenly spaced time series.
2. Monte Carlo simulation loop, repeated n_{sim} times (where n_{sim} is the number of simulations, usually around 1000).
   - Two uncoupled AR(1) processes r_x and r_y are generated as

$$\begin{aligned} &r_x(1) = \mathcal{E}_x(1) \sim N(0, 1), \\ &\text{and, for } i = 2, \ldots, n_x \\ &r_x(i) = \exp\left\{-\left[t_x(i) - t_x(i-1)\right]/\tau_x\right\}\cdot r_x(i-1) + \mathcal{E}_x(i), \\ &\mathcal{E}_x(i) \sim N\left(0, 1 - \exp\left\{-2\left[t_x(i) - t_x(i-1)\right]/\tau_x\right\}\right), \end{aligned} \tag{12}$$

and

$$\begin{aligned} &r_y(1) = \mathcal{E}_y(1) \sim N(0, 1), \\ &\text{and, for } i = 2, \ldots, n_y \\ &r_y(i) = \exp\left\{-\left[t_y(i) - t_y(i-1)\right]/\tau_y\right\}\cdot r_y(i-1) + \mathcal{E}_y(i), \\ &\mathcal{E}_y(i) \sim N\left(0, 1 - \exp\left\{-2\left[t_y(i) - t_y(i-1)\right]/\tau_y\right\}\right), \end{aligned} \tag{13}$$

   see Mudelsee (2010). We include the estimated τ_x and τ_y from Step 1, sampling times $(t_x(i))$ and $(t_y(i))$ (unevenly or evenly spaced) from each of the observed time series and unrelated set of $\mathcal{E}_x$ and $\mathcal{E}_y$ for each process, which are purely random processes drawn from normal distribution $N(\mu, \sigma^2)$, with zero mean μ and unit variance σ^2. The variance of the innovation term $\mathcal{E}$ is set to, for example $\sigma^2 = 1 - \exp(-2[t_y(i) - t_y(i-1)]/\tau_y)$, to have the AR(1) process stationary with unit variance.
   - The coherency $\hat{c}_{r_x r_y}^2(f_k)$ between the two generated AR (1) processes is estimated using Eq. (2).
3. The false-alarm level for $\hat{c}_{xy}^{\prime 2}(f_k)$ is determined as the $n_{\text{sim}}(1-\alpha)$ th point of the empirical sampling distribution of the $\hat{c}_{r_x r_y}(f_k)$ estimates. This is done at all frequencies f_k. Coherency values above this false-alarm level can be taken as significantly different from zero at given significance level α.

2.2.2. Monte Carlo phase confidence interval

Monte Carlo confidence intervals for the phase are estimated at frequencies f_k where the coherency exceeds the false-alarm level. The confidence interval is formed from the empirical distribution of the phase estimates formed by Monte Carlo simulations. The uncertainty in the phase angle depends on the coherency value and effective number of segments n_{eff}, which means that we need to generate two time series with prescribed coherency. The method described in Carter et al. (1973) is used to generate two white noise processes with prescribed coherency independent of frequency (similar approaches are also found in Foster and Guinzy, 1967; Miles, 2011)

$$\mathcal{E}_x^*(i) = \mathcal{E}_x(i) + G\,\mathcal{E}_y(i), \quad i = 1, \ldots, n, \tag{14}$$

$$\mathcal{E}_y^*(i) = \mathcal{E}_y(i) + G\,\mathcal{E}_x(i), \quad i = 1, \ldots, n, \tag{15}$$

where $\mathcal{E}_x$ and $\mathcal{E}_y$ are unrelated random processes with normal distribution, zero mean and unit variance and n is the number of data points. G can be selected from $c_{xy}^2(f_k) = 4G^2/(1+G^2)^2$ for desired value of c_{xy}^2 (Carter et al., 1973). This contains a fourth-order polynomial equation for the calculation of G, leading to $c_{xy}^2 G^4 + 2c_{xy}^2 G^2 + c_{xy}^2 - 4G^2 = 0$, which has four solutions for G. Here we use $G = \sqrt{-2(\sqrt{1 - c_{xy}^2}/c_{xy}^2) + 2/c_{xy}^2 - 1}$ to generate the coupled time series. This method works only if the two time series have identical time steps. Therefore to include both time scales, the Monte Carlo simulation loop is performed twice. In the first round the sample times $\{t(i)\}_{i=1}^{n}$ are set equal to the times from the first time series $\{t_x(i)\}_{i=1}^{n_x}$ and the persistence time τ equal to τ_x. In the second round the times are set equal to the times from second time series $\{t_y(i)\}_{i=1}^{n_y}$ and the persistence time equal to τ_y. Confidence intervals are formed from the empirical distributions of phase estimates for both cases and the mean value is taken for

final result. The detailed steps of the simulations are the following:

1. For each frequency (f_k) with significant coherency a Monte Carlo simulation loop is repeated n_{sim} times with $n_{\text{sim}} = 1000$:
   - Two coupled AR(1) processes r_x^* and r_y^* are generated as

$$r_x^*(1) = \mathcal{E}_x^*(1),$$
$$\text{and, for } i = 2, \ldots, n$$
$$r_x^*(i) = \exp\{-\left[t(i) - t(i-1)\right]/\tau\} \cdot r_x^*(i-1) + \mathcal{E}_x^*(i), \quad (16)$$

$$r_y^*(1) = \mathcal{E}_y^*(1),$$
$$\text{and, for } i = 2, \ldots, n$$
$$r_y^*(i) = \exp\{-\left[t(i) - t(i-1)\right]/\tau\} \cdot r_y^*(i-1) + \mathcal{E}_y^*(i). \quad (17)$$

     The white noise terms $\mathcal{E}_x^*$ and $\mathcal{E}_y^*$ are coupled with the method explained above Eqs. (14) and (15), where the bias corrected coherency estimate $\hat{c}_{xy}^{\prime 2}(f_k)$ is used to select appropriate G value.
   - The phase $\hat{\phi}_{r_x^* r_y^*}(f_k)$ is estimated according to (Eq. (9)).
2. The estimated phase $\hat{\phi}_{r_x^* r_y^*}(f_k)$ is centered around zero by subtracting the average phase (average of $\hat{\phi}_{r_x^* r_y^*}(f_k)$ over the n_{sim} simulations) from each phase value.
3. Confidence intervals for the phase are determined as $\hat{\phi}_{r_x^* r_y^*}(f_k) + 100(\alpha)$ th and $\hat{\phi}_{r_x^* r_y^*}(f_k) + 100(1-\alpha)$th percentage point of the empirical distribution of the estimated phase values $\hat{\phi}_{r_x^* r_y^*}(f_k)$.

The Monte Carlo simulation loop is performed twice for each frequency with significant coherency, first with the time scale t_x and persistence time τ_x and second with the times t_y and persistence time τ_y. The phase confidence interval at given frequency consists of the mean value from the two Monte Carlo simulation loops.

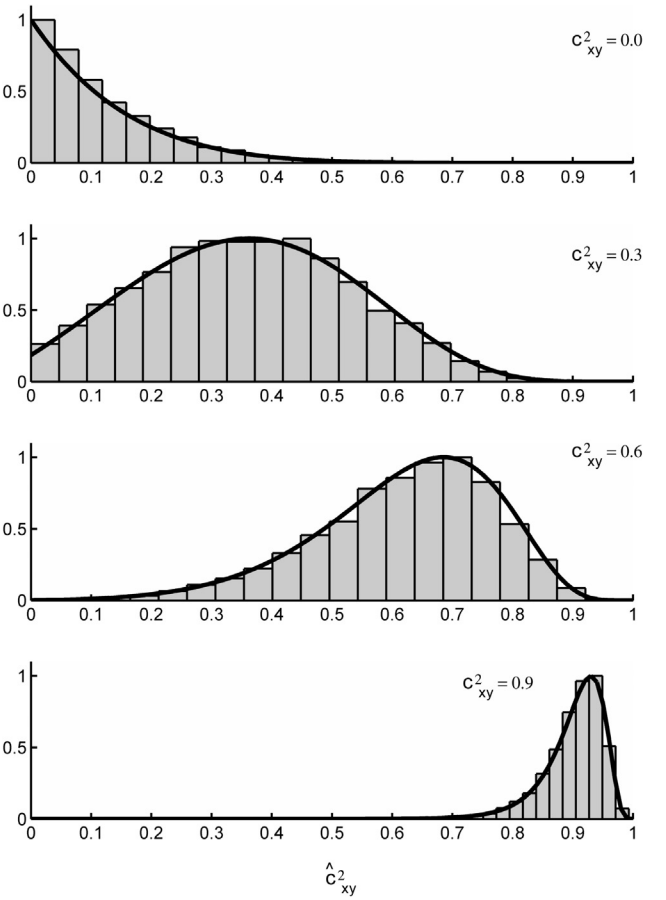

Fig. 1. Empirical distribution of coherency estimate at single frequency for four different prescribed coherency values. White noise processes, $n_{\text{eff}} = 8.24$, $n=300$, $n_{\text{sim}} = 10,000$. Black lines show the theoretical distributions (Eq. (18)).

3. Comparison of theoretical and Monte Carlo results

3.1. Coherency

We compare the empirical sample distribution of coherency formed by the Monte Carlo simulations with the theoretical distribution of coherency estimate from Carter et al. (1973), as the false-alarm levels for coherency are formed from these distributions.

The Monte Carlo simulation loop in the program is used to form the empirical distribution of coherency (see Section 2.2.1). In the first example 10,000 pairs ($n_{\text{sim}} = 10000$) of white noise processes were generated, where $\{r_x\}_{i=1}^{n_x}$ and $\{r_y\}_{i=1}^{n_y}$ in Eqs. (12) and (13) are set as Gaussian random processes $\mathcal{E}_x$ and $\mathcal{E}_y$, with mean $\mu = 0$ and variance $\sigma^2 = 1$. The number of data points in the generated time series is $n=300$ and the two series have identical and evenly spaced time points. Several experiments were done for different predefined coherency value between the two processes, independent of frequency. The white noise processes were coupled as in Eqs. (14) and (15), with four different prescribed coherency values $c_{xy}^2 \in \{0.0, 0.3, 0.6, 0.9\}$. Welch window was used in the spectral analysis, number of segments n_{50} was set to 10 which gives $n_{\text{eff}} = 8.24$. Histogram of the coherency estimates, from the Monte Carlo simulations, for single frequency is compared with the probability density function of the coherency estimate from Carter et al. (1973),

$$p\left(\hat{c}_{xy}^2 \middle| n_{\text{eff}}, c_{xy}^2\right) = (n_{\text{eff}} - 1)\left(1 - c_{xy}^2\right)^{n_{\text{eff}}}\left(1 - \hat{c}_{xy}^2\right)^{n_{\text{eff}}-2} \cdot \left(1 - c_{xy}^2 \hat{c}_{xy}^2\right)^{1-2n_{\text{eff}}} 2F_1\left(1 - n_{\text{eff}}, 1 - n_{\text{eff}}; 1; c_{xy}^2 \hat{c}_{xy}^2\right). \quad (18)$$

Matlab code from Pearson (2009) was used to compute the hypergeometric function $2F_1(1 - n_{\text{eff}}, 1 - n_{\text{eff}}; 1; c_{xy}^2 \hat{c}_{xy}^2)$. It is clear from Fig. 1 that we are able to reproduce the theoretical distribution of coherency estimate with the Monte Carlo simulations, in the case of two evenly spaced white noise processes.

The false-alarm level (Eq. (10)) is used to test if the estimated coherency values are different from zero, that is different from estimated coherency between two independent processes. The false-alarm level is derived from the probability function of the coherency estimator when $c_{xy}^2 = 0$, which can be written as

$$p\left(\hat{c}_{xy}^2 \middle| n_{\text{eff}}, c_{xy}^2 = 0\right) = (n_{\text{eff}} - 1)\left(1 - \hat{c}_{xy}^2\right)^{n_{\text{eff}}-2} \quad (19)$$

(Carter, 1977). The estimated coherency between two completely unrelated time series is expected to be greater than zero with the expected value of $E[\hat{c}_{xy}^2 | n_{\text{eff}}, c_{xy}^2 = 0] = 1/n_{\text{eff}}$ (which is also the bias) (Miles, 2006). This is correspondingly observed in the empirical distribution from the simulations above. We can determine the empirical bias for the estimated coherency by averaging $(\hat{c}_{r_x r_y}^2(f_k) - c_{xy}^2)$ over the $n_{\text{sim}} = 10,000$ simulations. The empirical

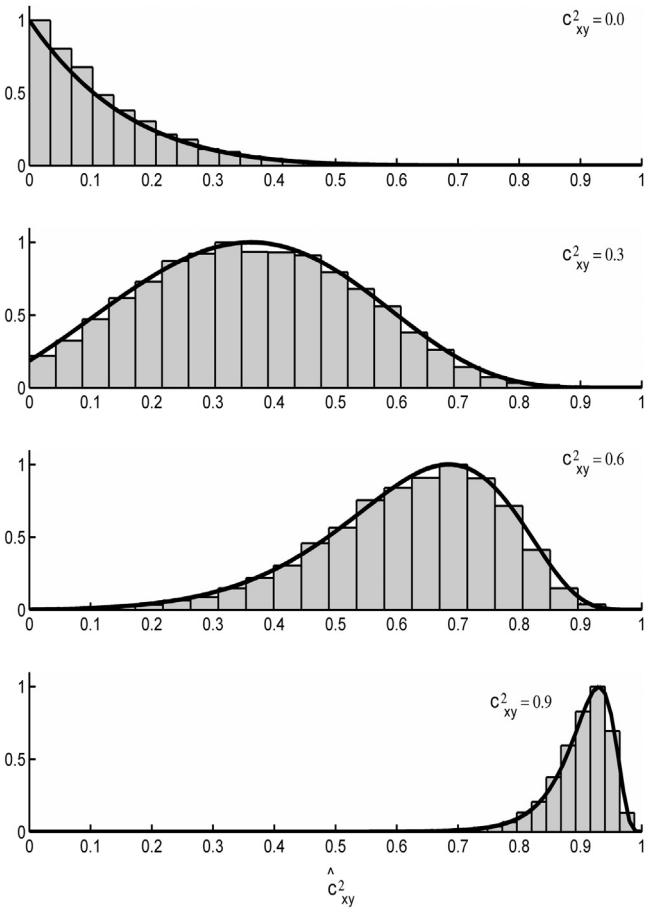

Fig. 2. As Fig. 1 but for red noise processes ($\tau_x = 3.0$, $\tau_y = 5.0$) with uneven spacing. $n_{\text{eff}} = 8.24$, $n = 300$, $n_{\text{sim}} = 10,000$. Black lines show the theoretical distributions (Eq. (18)).

bias for estimated coherency when $c^2_{xy} = 0$ and $n_{\text{eff}} = 8.24$ equals 0.12 which agrees with the theoretical bias. The theoretical false-alarm level is derived from the theoretical distribution of coherency estimator when the true coherency is zero. There is no bias correction done beforehand. According to that we do not correct for bias in the coherency estimates formed in the Monte Carlo loop, as it causes mismatch between the theoretical and Monte Carlo false-alarm level (the false-alarm level estimated with the Monte Carlo simulations would be lower than the false-alarm level from Eq. (10)).

In the second example the same settings were used in the spectral analysis but the generated processes were more complex. Now we generated 10,000 pairs of unevenly spaced Gaussian AR(1) processes. Several experiments were done with different predefined coherency values. Two coupled AR(1) processes r^*_x and r^*_y were generated as in Eqs. (16) and (17), where the τ values were set as 3.0 for the first time series r^*_x and as 5.0 for the second time series r^*_y. The number of data points in the time series is $n = 300$. The two time series have the same sampling times but the time scale is unevenly spaced. The uneven time spacing was drawn from Gamma distribution and the average time spacing set to 1.0. Fig. 2 shows the distribution of the coherency estimates at single frequency with four different prescribed coherency values $c^2_{xy} \in \{0.0, 0.3, 0.6, 0.9\}$. The figure shows that the empirical distribution explicitly agrees with the theoretical distribution (Eq. (18)). Neither the persistence in the data nor the uneven spacing seems to influence the distributional shape of the coherency

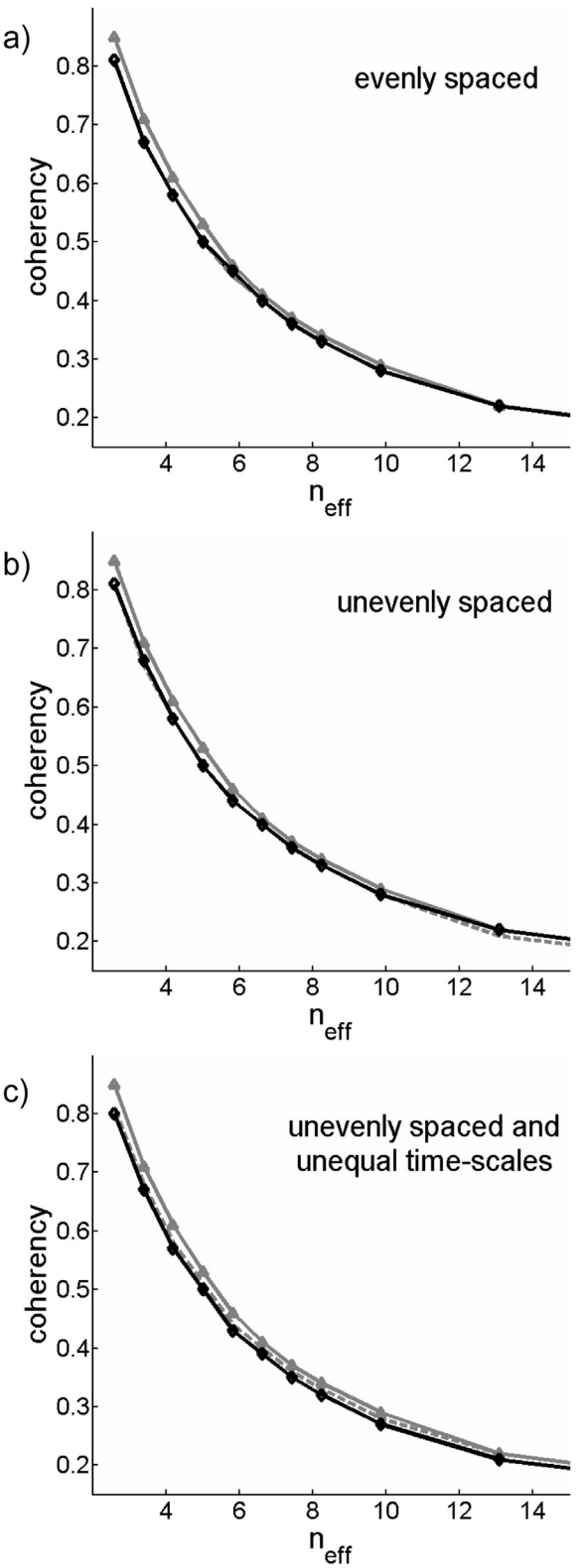

Fig. 3. 95% false-alarm level for coherency plotted against effective number of segments (n_{eff}). Grey line with triangle shows the theoretical false-alarm level, grey dashed line shows Monte Carlo false-alarm level generated with white noise processes, black line with diamonds shows Monte Carlo false-alarm level generated with red noise processes ($\tau_x = 3.0$, $\tau_y = 5.0$). Number of simulations were 1000. (a) Monte Carlo false-alarm level formed with evenly spaced time series, $n = 300$. (b) Monte Carlo false-alarm level formed with unevenly spaced time series, $n = 300$. (c) Monte Carlo false-alarm level formed with unevenly spaced time series with unequal time-scale, $n_x = 305$, $n_y = 293$.

estimates.

To analyse this further and to see if the Monte Carlo false-alarm level follows the theoretical false-alarm level for any number of segments, we plot the Monte Carlo false-alarm level against effective number of segments. As the Monte Carlo false-alarm level is more or less constant over all frequencies, we use the mean value over all frequencies. This was done both for Monte Carlo false-alarm levels formed with white noise and red noise processes ($\tau_x = 3.0$, $\tau_y = 5.0$) and compared with the theoretical false-alarm level for different numbers of segments (Fig. 3). The generated processes are either evenly spaced (Fig. 3a), unevenly spaced on equal timescales (Fig. 3b) or unevenly spaced on unequal timescale (Fig. 3c). The results show again that there is practically no difference between the Monte Carlo false-alarm levels formed with white noise processes or red noise processes. There is a small difference between false-alarm level formed with the Monte Carlo simulations (generated with both white and red noise processes) and the theoretical false-alarm level when the number of segments is less than 7 ($n_{\text{eff}} = 5.82$). The uneven spacing does not have any influence but there seems to be a little bit more mismatch between the theoretical and Monte Carlo false-alarm level in case of unequal timescale. In this experiment we used $n_{\text{sim}} = 1000$. The results were identical if the number of simulations were increased up to 10,000, which shows that there is no reason to use more than 1000 simulations to form the Monte Carlo false-alarm level.

3.2. Phase

The Monte Carlo phase confidence intervals are compared with the theoretical phase confidence intervals. For the comparison we generated two time series with identical periodicities and known phase difference. The first time series consists of two sine waves (frequencies $f_1 = 0.03$ and $f_2 = 0.2$; amplitudes: $A_1 = 1.0$ and $A_2 = 1.4$; phase values $\phi_1 = 0°$ and $\phi_2 = 60°$). Additionally Gaussian noise was added to the signal (variance $\sigma^2 = 1.0$). The time scale is unevenly spaced (drawn from Gamma distribution) with average time spacing $\bar{d} = 1.0$ year and number of data points $n = 300$. The second time series consists of the same periodicities (frequencies $f_1 = 0.03$ and $f_2 = 0.02$; amplitudes: $A_1 = 1.4$ and $A_2 = 1.0$; phase values $\phi_1 = 90°$ and $\phi_2 = 0°$) plus Gaussian noise (variance $\sigma^2 = 1.0$). The time scale is evenly spaced with time spacing $d = 1.0$ year and number of data points $n = 300$. Coherency and phase spectrum were estimated along with Monte Carlo and theoretical phase confidence intervals. Welch window was used in the spectral analysis and number of segments $n_{50} = 7$ ($n_{\text{eff}} = 5.82$).

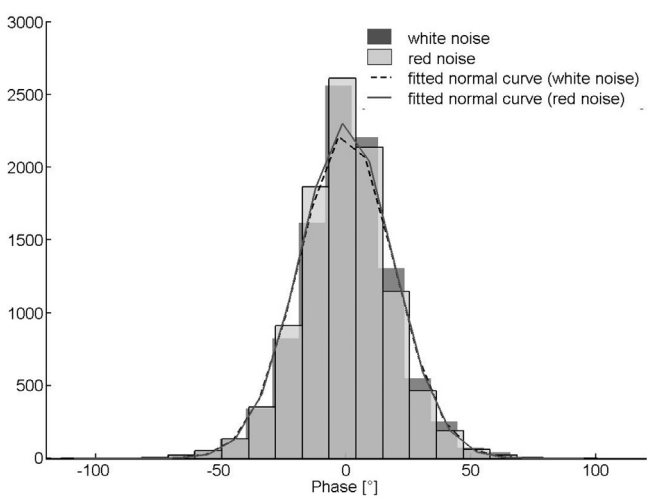

Fig. 4. Empirical distributions of phase estimates at single frequency with significant coherency for two cases, white noise and red noise. The lines show normal curves with estimated mean and standard deviation from the data (black dashed line for white noise and grey line for red noise).

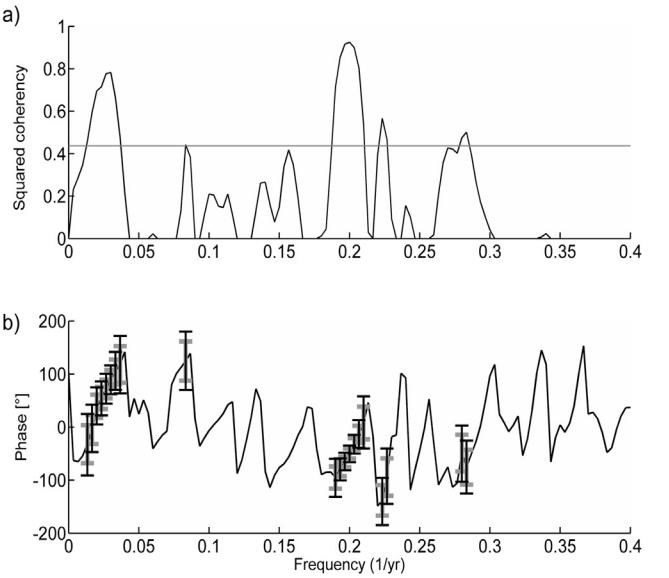

Fig. 5. Comparison between Monte Carlo phase confidence intervals and theoretical phase confidence intervals. (a) Coherency spectrum for two generated time series with 95% Monte Carlo false-alarm level (the mean value over all frequencies). (b) Phase spectrum with 95% confidence intervals estimated at frequencies with significant coherency. Monte Carlo phase confidence intervals (black narrow error bars) estimated with $n_{\text{sim}} = 1000$, theoretical confidence intervals (grey thick error bars).

Histogram of the phase estimates, from the Monte Carlo simulation loop (with $n_{\text{sim}} = 10,000$), for single frequency with a significant coherency is used to infer if the phase estimates follow a normal distribution. In the first case white noise processes (τ_x and $\tau_y = 0$) were generated to form the empirical distribution of the phase estimates, while in the second case red noise processes ($\tau_x = 5.0$ and $\tau_y = 3.0$) were generated (Fig. 4). χ^2-test was used to test the null hypothesis at 5% significance level that the phase estimates come from a normal distribution. The χ^2-test statistics in both cases are above the critical value, which means that the sample distribution of the phase estimates are significantly different from a normal distribution. The color of the noise does not influence the distributional shape of the phase estimates (Fig. 4). The phase values are in general shifted along the x-axis when persistence is included but that is corrected for in the REDFIT-X program as the phase values are centered around zero by subtracting the average phase value from the Monte Carlo ensemble from each phase value.

Fig. 5 shows a comparison of the Monte Carlo phase confidence intervals estimated with $n_{\text{sim}} = 1000$ and the theoretical phase confidence intervals. It shows that the Monte Carlo phase confidence intervals are slightly wider than the theoretical ones.

4. Application

As an example we applied the method to the same time series used in the SPECTRUM paper (Schulz and Stattegger, 1997). The first time series is the SPECMAP stack (oxygen isotope data, $\delta^{18}O$) (Imbrie et al., 1984) and the second time series is a proxy record of North-Atlantic sea-surface temperature (SST) (Ruddiman et al., 1989). The example investigates the relation between global climate and regional SST changes at the Milankovitch periods of Earth orbital variations. The intention is not to answer the underlying climatological questions. But to allow comparison

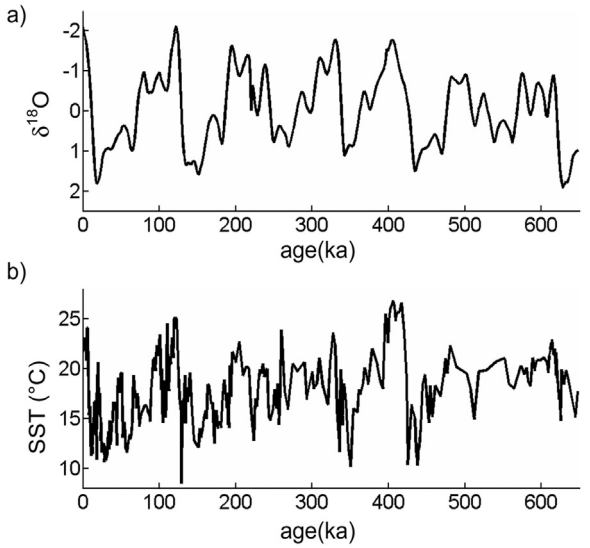

Fig. 6. (a) SPECMAP oxygen isotope data (δ_{18}O). (b) North Atlantic sea-surface temperatures (SST).

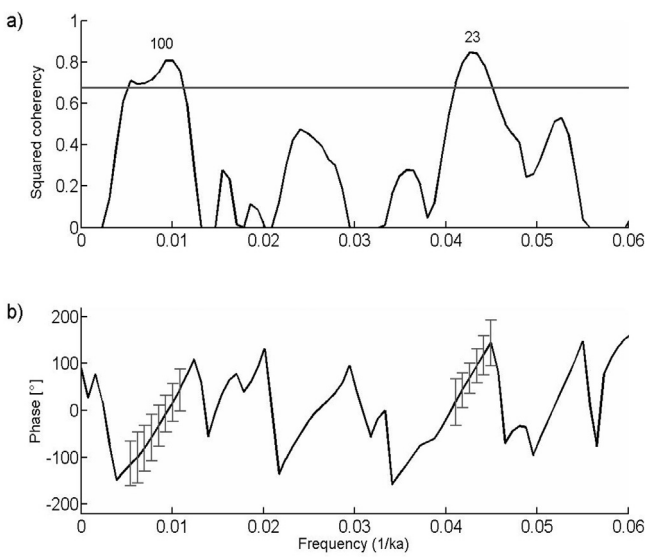

Fig. 7. Estimated coherency and phase spectrum on the SPECMAP and SST time series from Fig. 4. (a) Coherency spectrum, the grey line indicates 95% false-alarm level (the mean value over all frequencies) formed with 1000 Monte Carlo simulations. (b) Phase spectrum, 95% Monte Carlo confidence intervals were estimated at frequencies where the coherency exceeds the false-alarm level.

between the outcomes from the new program with the previous SPECTRUM results. Therefore we used the identical age model as in the previous paper, although this has now been revised (Lisiecki and Raymo, 2005). The oxygen isotope time series is evenly spaced, with time spacing=2 ka and number of data points n=325 (Fig. 6a). The SST time series is unevenly spaced with number of data points n=322 and average spacing $\bar{d}$ = 2 ka (Fig. 6b). The settings used in the cross-spectral analysis were the same as used in the example in Schulz and Stattegger (1997), as follows, number of segments n_{50} = 4, Welch window, OFAC=5 and HIFAC=1.0. The parameters OFAC and HIFAC are used to set the frequency grid in the spectra. OFAC is the oversampling factor that determines how many frequencies are investigated in the Lomb–Scargle Fourier transform and HIFAC determines the highest frequency to analyse in the Fourier transform or HIFAC $*$ f_{Nyq} (Schulz and Stattegger, 1997). The minimum values in the δ^{18}O record are expected to correlate with maxima in the SST record. Therefore the sign of the δ^{18}O record was changed prior to the analysis to prevent an artificial offset in the phase by 180°. The coherency spectrum indicated significant coherencies at frequencies 1/100 and 1/23 ka^{-1} (Fig. 7a). The Monte Carlo false-alarm level (α=0.05) was estimated with 1000 simulations. The phase spectrum showed that the ice-volume minima lead SST maxima by 17°[− 23.5°; 57.1°] (f_k = 1/99.26 ka^{-1}) and 95°[58.0°; 131.1°] (f_k = 1/23.04 ka^{-1}) (Fig. 7b). The confidence interval for the phase was estimated with Monte Carlo simulations (n_{sim} = 1000). The results for the coherency spectrum were in harmony with previous results from Schulz and Stattegger (1997). The Monte Carlo false-alarm level was slightly lower than the theoretical one, but there is a small difference between them when the number of segments is less than 7, as mentioned in the previous section. The confidence interval for the phase angles was wider than in the previous paper as the Monte Carlo phase confidence interval is in general wider than the theoretical one. The seesaw shape of the phase spectrum is due to the fact that the mean sampling time in the two time series differs. This was discussed in Schulz and Stattegger (1997) and in the SPECTRUM program there is an option to estimate aligned phase spectrum to make it easier to measure the phase angles from the phase spectrum plot. In the case of alignment, the phase angles need to be corrected back, before interpreting the results. For simplification, this alignment option is not included in the new version of the program.

5. Discussion

Simulation approaches have been previously applied to estimate false-alarm level for coherency. Pardo-Igúzquiza and Rodríguez-Tovar (2012) use Monte Carlo simulations to form statistical significance for coherency spectrum estimated via Lomb–Scargle cross-periodogram. They use a permutation test which preserves the distribution of the underlying data but they do not take into account the persistence of the observed time series as they generate white noise processes. Faes et al. (2004) define threshold value for zero coherence with three different surrogate analysis type, where the surrogates preserve either the distributional shape or the spectral properties of the observed data. Huybers and Denton (2008) analyse the significance of the coherency with Monte Carlo simulation technique as well as confidence intervals for the phase, but they do not consider unevenly spaced data.

We show that the Monte Carlo false-alarm level fits the theoretical false-alarm, when using more than 6 segments. This is in agreement with results from Gallet and Julien (2011), who modified the false-alarm level equation (Eq. (10)) for 50% overlapping to take into account the effect of overlapping segmenting. This is similar to what is done here with the effective number of segments. The persistence of the generated data in the Monte Carlo simulations does not influence either the distributional shape of the coherency estimate or, consequently, the false-alarm level for coherency. The color of the noise has also little impact on significance level for wavelet coherence formed with Monte Carlo simulations by Grinsted et al. (2004). Although there is a little difference between the Monte Carlo false-alarm level and the theoretical false-alarm level, we highly recommend the usage of the Monte Carlo false-alarm level to add more confidence to the result. It can be especially convenient when the method is applied on for example noisy or strange data.

The Monte Carlo phase confidence intervals are slightly wider than the theoretical confidence intervals used in SPECTRUM. The sample distribution of the Monte Carlo phase estimates does not follow a normal distribution. The phase estimates are known to be approximately normally distributed (Bloomfield, 2000). When the approximation is good, variance is small, coherency is not small the

theoretical confidence interval from SPECTRUM can be suitable. Hannan (1970) uses for example percentage point of the t distribution to form confidence intervals for the phase angle as it is thought to give a better approximation of the distribution of the phase estimates than the normal distribution.

6. Conclusion

Here we present a computer program REDFIT-X that performs cross-spectral analysis of unevenly spaced paleoclimate time series. We use a Monte Carlo approach to estimate the uncertainty associated with coherency and phase. The program is an updated version of the existing programs REDFIT (Schulz and Mudelsee, 2002) and SPECTRUM (Schulz and Stattegger, 1997). Now both auto- and cross-spectral analysis can be performed under the same program where the significance measurements are estimated with Monte Carlo simulation technique.

A comparison between the empirical sample distribution of coherency formed by the Monte Carlo simulations and the theoretical distribution of coherency shows that we are able to reproduce the theoretical distribution with the Monte Carlo simulations. The color of the noise does not have large influence on the distribution or the false-alarm level for coherency. The worked example demonstrates the applicability of the method to real paleoclimate time series and allows for exact comparison between the outcomes from the new program with the previous SPECTRUM results. The Monte Carlo phase confidence intervals are slightly wider than the theoretical confidence intervals for phase used in SPECTRUM. The program is suitable for unevenly spaced climate time series where the persistence of the data is taken into account when estimating the uncertainty associated with the coherency and phase estimates.

Acknowledgments

This study has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) - Marie-Curie ITN, under Grant agreement no. 238512, GATEWAYS project. Funding also came from the DFG-Research Center MARUM - The Ocean in the Earth System and GLOMAR - Bremen International Graduate School for Marine Sciences. We thank Heather Johnstone for helpful comments regarding the English. We thank two anonymous reviewers for constructive comments.

Appendix A. Supplementary data

Supplementary data associated with this paper can be found in the online version at http://dx.doi.org/10.1016/j.cageo.2016.03.001.

References

Allen, M.R., Smith, L.A., 1996. Monte Carlo SSA: detecting irregular oscillations in the presence of colored noise. J. Clim. 9, 3373–3404.
Bendat, J.S., Piersol, A.G., 2010. Random Data Analysis and Measurement procedures, 4th Edition. Wiley, New Jersey.
Benignus, V.A., 1969. Estimation of the coherence spectrum and its confidence interval using the fast Fourier transform. IEEE Trans. Audio Electroacoust. 17, 145–150.
Bloomfield, P., 2000. Fourier Analysis of Time Series: An Introduction, 2nd Edition. Wiley, New York.
Carter, G.C., 1977. Receiver operating characteristics for a linearly thresholded coherence estimation detector. IEEE Trans. Acoust. Speech Signal Process. 25, 90–92.
Carter, G.C., Knapp, C.H., Nuttall, A.H., 1973. Estimation of the magnitude-squared coherence function via overlapped fast Fourier transform processing. IEEE Trans. Audio Electroacoust. AU-21, 337–344.
Faes, L., Pinna, G.D., Porta, A., Maestri, R., Nollo, G., 2004. Surrogate data analysis for assessing the significance of the coherence function. IEEE Trans. Biomed. Eng. 51, 1156–1166.
Foster, M.R., Guinzy, N.J., 1967. The coefficient of coherence; its estimation and use in geophysical data processing. Geophysics 32, 602–616.
Gallet, C., Julien, C., 2011. The significance threshold for coherence when using the Welch's periodogram method: effect of overlapping segments. Biomed. Signal Process. Control 6, 405–409.
Goodman, N.R., 1957. On the Joint Estimation of the Spectra, Cospectrum, and Quadrature Spectrum of a Two-dimensional Stationary Gaussian Process. Scientific Paper No.10 (AD 134919). New York University, New York.
Grinsted, A., Moore, J.C., Jevrejeva, S., 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. Nonlinear Precess. Geophys. 11, 561–566.
Hannan, E.J., 1970. Multiple Time Series. Wiley, New York.
Harris, F.J., 1978. On the use of windows for harmonic analysis with the discrete Fourier transform. Proc. IEEE 66, 51–83.
Hasselmann, K., 1976. Stochastic climate models: Part I. Theory. Tellus 28, 473–485.
Huybers, P., Denton, G., 2008. Antarctic temperature at orbital timescales controlled by local summer duration. Nat. Geosci. 1, 787–792.
Imbrie, J., Hays, J.D., Martinson, D.G., McIntyre, A., Mix, A.C., Morley, J.J., Pisias, N.G., Prell, W.L., Shackleton, N.J., 1984. The orbital theory of Pleistocene climate: support from a revised chronology of the marine δ^{18}O record. In: Berger, A., Imbrie, J., Hays, H., Kukla, G., Saltzman, B. (Eds.), Milankovitch and Climate, Part I. D. Reidel Publishing, Dordrecht, pp. 269–305.
Koopmans, L.H., 1995. The Spectral Analysis of Time Series. Academic Press, New York.
Lisiecki, L.E., Raymo, M.E., 2005. A Pliocene–Pleistocene stack of 57 globally distributed benthic δ^{18}O records. Paleoceanography 20. http://dx.doi.org/10.1029/2004PA001071, PA1003.
Lomb, N.R., 1976. Least-squares frequency analysis of unequally spaced data. Astrophys. Space Sci. 39, 447–462.
Miles, J.H., 2006. Aligned and Unaligned Coherence: A New Diagnostic Tool, AIAA Paper No. 2006-10. Technical Report NASA/TM2006-214112, NASA, Cleveland, OH.
Miles, J.H., 2011. Estimation of signal coherence threshold and concealed spectral lines applied to detection of turbofan engine combustion noise. J. Acoust. Soc. Am. 129, 3068–3081.
Mudelsee, M., 2002. TAUEST: a computer program for estimating persistence in unevenly spaced weather/climate time series. Comput. Geosci. 28, 69–72.
Mudelsee, M., 2010. Climate Time Series Analysis: Classical Statistical and Bootstrap Methods, 1st Edition. Springer, Dordrecht.
Pardo-Igúzquiza, E., 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. Comput. Geosci. 49, 207–216.
Pearson, J., 2009. Computation of Hypergeometric Functions [Master's thesis]. University of Oxford.
Ruddiman, W.F., Raymo, M.E., Martinson, D.G., Clement, B.M., Backman, J., 1989. Pleistocene evolution: northern hemisphere ice sheets and North Atlantic Ocean. Paleoceanography 4, 353–412.
Scargle, J.D., 1982. Studies in astronomical time series analysis. II. Statistical aspects of spectral analysis of unevenly spaced data. Astrophys. J. 263, 835–853.
Schulz, M., Mudelsee, M., 2002. REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series. Comput. Geosci. 28, 421–426.
Schulz, M., Stattegger, K., 1997. SPECTRUM: spectral analysis of unevenly spaced paleoclimatic time series. Comput. Geosci. 23, 929–945.
von Storch, H., Zwiers, F.W., 2003. Statistical Analysis in Climate Research. Cambridge University Press, Cambridge.
Welch, P.D., 1967. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. IEEE Trans. Audio Electroacoust. AU-15, 70–73.