Physics of the Earth and Planetary Interiors 130 (2002) 103–116

PHYSICS OF THE EARTH AND PLANETARY INTERIORS

www.elsevier.com/locate/pepi

Spectral analysis of unevenly spaced climatic time series using CLEAN: signal recovery and derivation of significance levels using a Monte Carlo simulation

D. Heslop [a,*], M.J. Dekkers [a,b]

[a] Paleomagnetic Laboratory 'Fort Hoofddijk', Faculty of Earth Sciences, Utrecht University, Budapestlaan 17, 3584 CD Utrecht, The Netherlands

[b] CEREGE, BP80, 13545 Aix en Provence cedex 4, France

Received 2 July 2001; received in revised form 26 November 2001; accepted 3 December 2001

Abstract

We present a Monte Carlo based method for the determination of errors associated with frequency spectra produced by the CLEAN transformation of Roberts et al. (1987). The Monte Carlo procedure utilises three different types of simulation involving a data stripping operation and the addition of white and red noise to the analysed time series. The simulations are tested on both synthetic and real data sets demonstrating the ability of the procedures to extract coherent information from time series characterised by the low signal-to-noise-ratio that is typical of many palaeoclimatic records. Significance levels derived for the Monte Carlo spectra of four time series from the Vostok ice core are utilised in the study of eccentricity components contained within the palaeoclimatic archive since ∼420 ka. Inversion of the Vostok frequency spectra into the time domain reveals the differing influence of orbital parameters in the palaeoclimatic proxy records as well as the relative magnitudes of the eccentricity components contained in the time series of greenhouse gas concentration, ice volume and local temperature.

Keywords: CLEAN transformation; Unevenly spaced time series; Palaeoclimate

1. Introduction

In many cases palaeoclimatic data sets are unevenly sampled in the time domain. This is often a product of the non-linear relationship that commonly exists between depth and time, resulting in the transformation of a sampling regime that is equidistant in the depth domain into a non-uniformly spaced series in the time domain. Many palaeoclimatic investigations not only analyse data in the depth and time domains, but also in the frequency domain. A considerable number of spectral analysis techniques, however, are based on the discrete Fourier transform (DFT), which require evenly sampled data and irregularly spaced time series must therefore undergo preprocessing before investigation of frequency content can be undertaken. The simplest form of this preprocessing is to linearly interpolate the dataset onto an evenly spaced time array. Unless performed carefully such an interpolation procedure can lead to aliasing of the signal (Schulz and Stattegger, 1997; Smith, 1997) resulting in the introduction of spurious components that may influence or even dominate the signal in the frequency domain.

* Corresponding author. Tel.: +31-30-2531-672; fax: +31-30-2531-677.
E-mail address: heslop@geo.uu.nl (D. Heslop).

0031-9201/02/$ – see front matter
PII: S0031-9201(01)00310-7

Roberts et al. (1987) developed the CLEAN algorithm to address the problem of distortion of frequency spectra caused by the incomplete sampling of effectively continuous signals. Their procedure utilised knowledge of the sampling function to perform an iterative deconvolution of the spectral window in the frequency domain (see Roberts et al., 1987; Baisch and Bokelmann, 1999 for derivations of the algorithm). Using generated time series, Roberts et al. (1987) demonstrated the robustness of their CLEAN procedure in the analysis of unevenly sampled sequences and data sets containing missing values.

Baisch and Bokelmann (1999) tested the algorithm further using both synthetic signals (with a 31% noise contribution) and real seismological data sets. In addition, the studies of Roberts et al. (1987) and Baisch and Bokelmann (1999) both showed that through the use of the inverse Fourier transform (IFT), a "CLEANed" signal could be successfully reconstructed in the time domain from the CLEAN frequency spectra.

A disadvantage of the CLEAN technique is that, unlike some other spectral techniques designed for unevenly spaced time series (i.e. the Lomb–Scargle periodogram, Lomb, 1976; Scargle, 1982, 1989; Schulz and Stattegger, 1997), there is no simple calculation to determine the significance of the frequency peaks contained within the spectrum. Here, we present a technique for assessing the significance levels of CLEAN spectra based upon a series of Monte Carlo simulations. We demonstrate the Monte Carlo process using both synthetic and real data and show that reconstruction of signals in the time domain from frequency components above specified significance levels can provide important insights into palaeoclimatic time series. The Monte Carlo CLEAN software (and user manual) written to perform the analyses presented in this paper can be freely downloaded from http://www.geo.uu.nl/~forth/software/soft.html. The software is a *MATLAB* GUI and utilises modified versions of the *clean.m* function of Baisch and Bokelmann (1999).

2. Monte Carlo simulations

The previously demonstrated robustness of the CLEAN algorithm in the treatment of unequally spaced (Roberts et al., 1987) and noisy (Baisch and Bokelmann, 1999) signals makes it suitable for use in Monte Carlo simulations. We employ three simulation procedures involving data stripping and the addition of noise (white and red) to the input time series.

2.1. Data stripping procedure

The power of the CLEAN algorithm in the analysis of unevenly spaced data allows it to be used in a data stripping procedure where datasets become inherently unevenly sampled in time. In order to perform the stripping procedure a random sample of r data points was taken from the total dataset of length N. Using such a method produces $N!/[r!(N-r)!]$ different data combinations, each of which can be processed with the CLEAN algorithm.

The stripping of data points from the input signal will introduce artificial gaps into the dataset, the size of which will be controlled by the relative magnitudes of N and r. To investigate the effects of the stripping procedure on the input data sets we produced a 500 point equally-spaced time array that could be stripped to a specified level and the characteristics of the resultant gaps analysed. For each analysed value of r the data set was stripped 1000 times and the mean length of the point spacing in the series was calculated. This procedure demonstrated that the mean lengths of the artificial gaps follow the path expected, i.e. approximating to N/r. Analysis of the maximum gap length throughout all 1000 iterations indicates that in some of the runs gaps approximately an order of magnitude greater than the mean were produced (Fig. 1).

The data stripping process can be repeated for a specified number of loops, Iter, with a CLEAN frequency spectrum being produced from the newly selected r data points in each cycle. This results in Iter different spectra of modulus verses frequency for which confidence intervals and significance limits can be calculated.

The calculation of 95% confidence limits for the modulus distributions at each frequency value allows a confidence interval to be displayed about the mean of the frequency spectrum (using $[\bar{x} - 1.96\sigma/\sqrt{\mathrm{Iter}},\ \bar{x} + 1.96\sigma/\sqrt{\mathrm{Iter}}]$). Sorting the power levels across all frequencies from all spectra allows a significance limit, α, to be determined for the mean spectrum. For example, if the Monte Carlo CLEAN procedure outputs 1000 spectra composed of 250 points

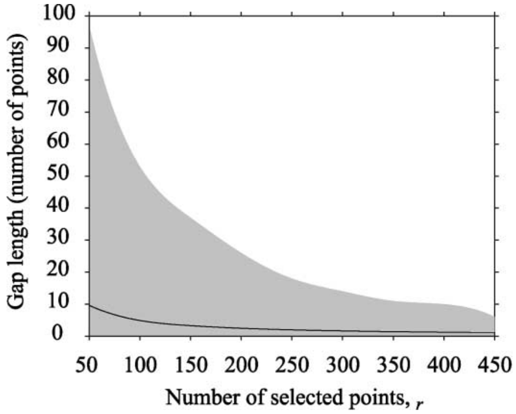

Fig. 1. Analysis of the length of the gaps introduced into an equally-spaced 500 point time series by the procedure utilised in the data stripping analysis. The black line shows the mean gap length determined over 1000 simulations at each value of r. The shaded regions represents the full range of gap lengths produced by the procedure.

each (yielding a total of 250,000 modulus values) determining the point below which 95% of the sorted modulus array occurred would provide the α_{95} value.

2.2. White noise addition

Through the addition of white noise it is possible to reduce the signal-to-noise ratio of a data set and test further the robustness of any frequency content contained within the time series. White noise has a flat frequency spectrum and was constructed in this study by producing a pseudo-random data array sampled from a normal distribution (using a table lookup algorithm, with a mean of zero and variance equal to one). The random data set was then scaled to the maximum and minimum values of the input time series before being multiplied by a predefined coefficient in order to control the amplitude of the white noise signal with respect to the amplitude of the input time series. Finally, the noise array was added to the original time series and the composite signal was processed by CLEAN (no stripping was involved in this procedure). After repeating the procedure a total of Iter times with a new noise array for each iteration, a mean spectrum, confidence limits and significance level (α) could be calculated for the frequency-modulus matrix.

2.3. Red noise addition

To generate a discrete finite red noise series we utilised a first-order autoregressive process (hereafter referred to as the AR(1) process, Mann and Lees, 1996). A normally distributed white noise array (ω_n) was constructed and scaled to have the same variance as the input time series. The magnitude of the initial white noise series therefore remained constant for all iterations in a simulation and a red noise model could be constructed with a predefined lag-one autocorrelation coefficient, ρ (where $0 < \rho \leq 1$), according to

$$r_n = \rho r_{n-1} + \omega_n \tag{1}$$

where $n = 1, \dots, N$ and represents the sampling increment (Δt) of an evenly sampled time vector. Mann and Lees (1996) show that the characteristic noise decay time scale of the series can be determined as

$$\tau = -\frac{\Delta t}{\log \rho} \tag{2}$$

Therefore the dominant periodicity in the red noise spectrum can be controlled using ρ. Once a red noise series had been generated, the principles of the Monte Carlo simulation followed those of the white noise addition procedure. The constructed noise was added to the original (non-stripped) data series and the composite signal was analysed. Again a mean spectra, confidence limits, and significance level (α) were calculated after the procedure had been repeated a total of Iter times.

2.4. Randomisation test

For each of the three methods described above an additional randomisation test was performed during the individual iterations of the Monte Carlo procedures. The test involved the randomisation of the data array (stripped or containing introduced noise) whilst the time array remained unchanged. The CLEAN transformation is then performed upon the randomised series. This procedure produces Iter spectra for the input data arrays and Iter spectra for the randomised data arrays at the end of each simulation, α significance levels can then be calculated for the two sets of spectra. If the α_{95} modulus value of the mean spectrum obtained from the randomised data arrays was found to be greater than that of the non-randomised data then it was taken as an indication that the frequency peaks contained within the Monte Carlo CLEAN spectrum were unstable.

2.5. Inversion of the mean spectrum

Performing an IFT on the mean CLEAN spectrum allows reconstruction of a CLEANed signal in the time domain (see Roberts et al., 1987, for a detailed methodology). Such an inversion can also be performed using only the frequencies of the spectrum above the defined α significance. In practice this was achieved by performing the IFT on a modified spectrum in which the power levels of the significant frequencies remained unchanged whilst those of the non-significant frequencies were reduced to zero. By reconstructing the time series at different values of α it becomes possible to identify how specific periodic components of the input signal behave in the time domain.

3. Analysis of a synthetic time series

To investigate the applicability of our Monte Carlo simulations to the analysis of palaeoclimatic data, we produced a synthetic time series based upon the 65°N insolation data (0–500 ka) of Laskar (1990). The synthesis procedure involved the construction of a new time array consisting of 400 randomly distributed ages between 0 and 500 ka, onto which the original La_{90} insolation curve could be interpolated. A normally distributed white noise data set was generated and scaled to the same variance as the La_{90} series. The synthetic palaeoclimatic record was formed by addition of the La_{90} and white noise arrays to produce a time series that thus was known to be 50% signal and 50% noise. The record was then analysed using the Monte Carlo CLEAN procedure to investigate if the original La_{90} insolation curve could be successfully extracted from the noise contribution.

3.1. Data stripping simulations

The first task of the analysis was to determine a magnitude for Iter (the number of loops in the simulation), which would be sufficient to produce a representative sample of all the possible data combinations

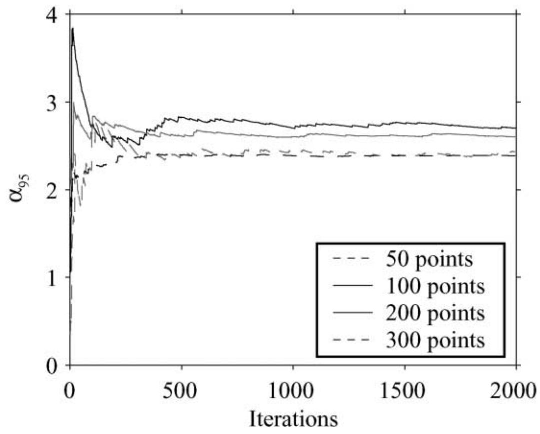

Fig. 2. Progress of a Monte Carlo CLEAN simulation involving the stripping of a 400 point synthetic palaeoclimatic time series into a 300, 200, 100 and 50 point input array. Variations in α_{95} of the 300, 200, 100 and 50 point simulations are shown with an increasing number of iterations. Each of the simulations appears to become stable by 500 iterations.

and thus yields a stable spectrum. To achieve this aim we ran four stripping simulations (arbitrarily chosen at 50, 100, 200 and 300 points) on the synthetic palaeoclimatic time series and determined the value of the α_{95} level after each iteration. Fig. 2 shows the results of this investigation and demonstrates that the α_{95} values became stable around the 500 iteration point in all four simulations. In all subsequent simulations we set Iter equal to 1000.

The mean spectrum of the 200 point simulation (representing data stripping to 50% of the original signal) was used to investigate the characteristics of the modulus distributions produced during the procedure. The output Monte Carlo CLEAN spectrum for the 200 point simulation is shown in Fig. 3a. The modulus array in Fig. 3b, obtained by sorting all the modulus values across all the analysed frequencies for all the iterations was utilised in the calculation of the α_{95} confidence limit.

The shape of the distribution of modulus values for single frequency increments was found to vary within the spectrum. Fig. 3c–e show the distribution of modulus values obtained for frequencies of 0.0424, 0.0625 and 0.0066 kyr^{-1}, respectively. The frequency $f = 0.0424$ has the highest mean modulus of the entire spectrum shown in 3(a) and in this case the values are normally distributed. The other two distributions have

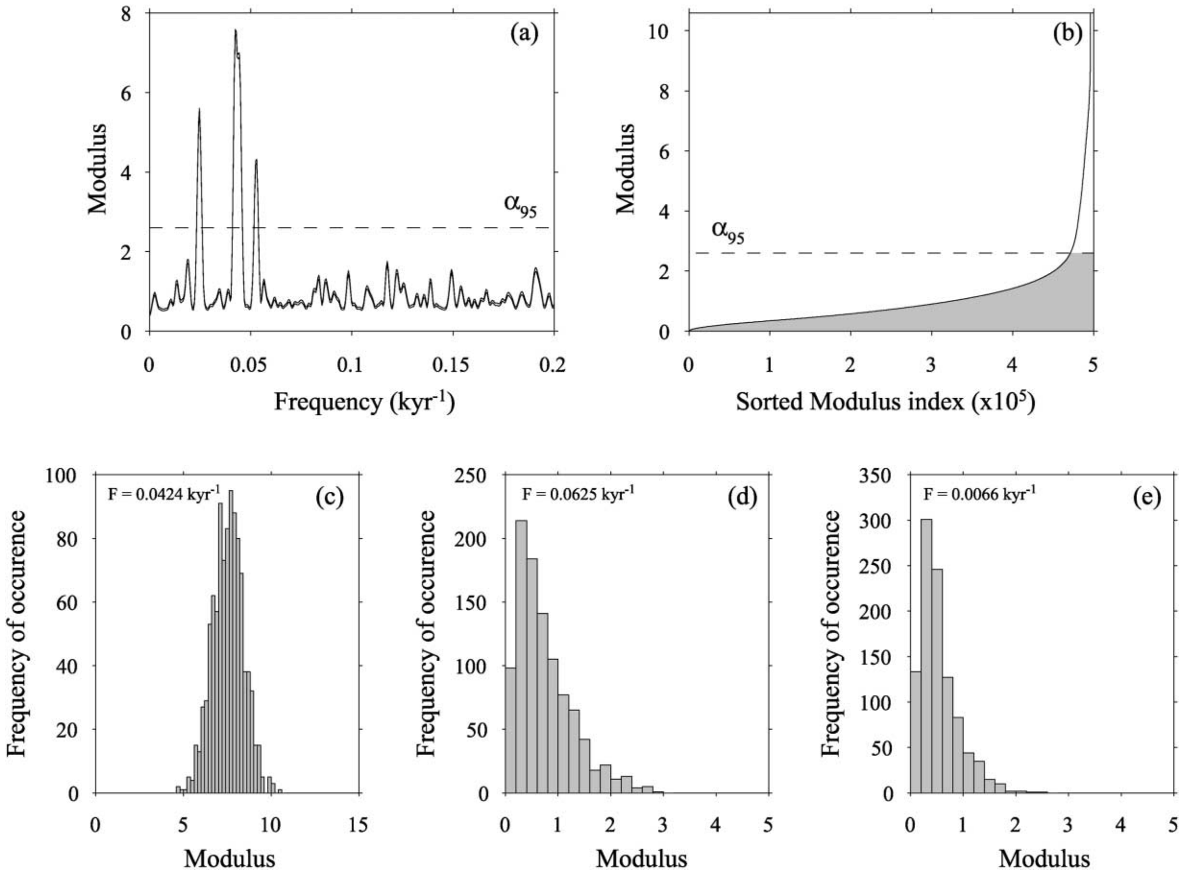

Fig. 3. **(a)** A typical spectrum produced by the Monte Carlo CLEAN procedure (stripping of the synthetic palaeoclimatic time series from 400 to 200 points). The varying thickness of the shaded line represents the 95% confidence interval determined from the distribution of Itermodulus values produced at each frequency interval during the simulation. (b) Determination of the α_{95} significance level for the mean spectrum utilising all the modulus values across all frequencies in all of the spectra produced during the simulation. Distributions of modulus results at the frequencies: (c) 0.0424; (d) 0.0625; (e) 0.0066 kyr^{-1}, respectively.

low mean modulus values and are characterised by positively skewed distributions resulting from the truncation of the baseline values by zero (negative moduli are not possible).

Fig. 4 shows construction of the synthetic palaeoclimatic time series and the analysis of its frequency content using the CLEAN algorithm in its traditional non-Monte Carlo form. Fig. 4a and d show the (400 out of 500 pts) 65°N insolation series and the combined insolation-noise data sets, respectively. Performing the CLEAN transformation on the unmodified insolation series produces a smooth frequency spectrum (Fig. 4b) and the reconstructed signal (Fig. 4c) yields a linear correlation coefficient of $R^2 = 0.995$, when compared to original insolation curve. The frequency spectrum (Fig. 4e) derived from the synthetic palaeoclimatic series reveals the effect of high-level white noise in the frequency domain. The main frequency peaks observed in Fig. 4b are still present in Fig. 4e, however, reconstruction of the signal (Fig. 4f) does not successfully isolate the insolation series from the noise component ($R^2 = 0.373$).

The Monte Carlo data stripping simulations shown in Fig. 5 resampled the 400 point signal at the 50, 100, 200 and 300 point levels for a total of 1000 iterations. The panels on the left of the figure show the final spectra with the thickness of the shaded line representing

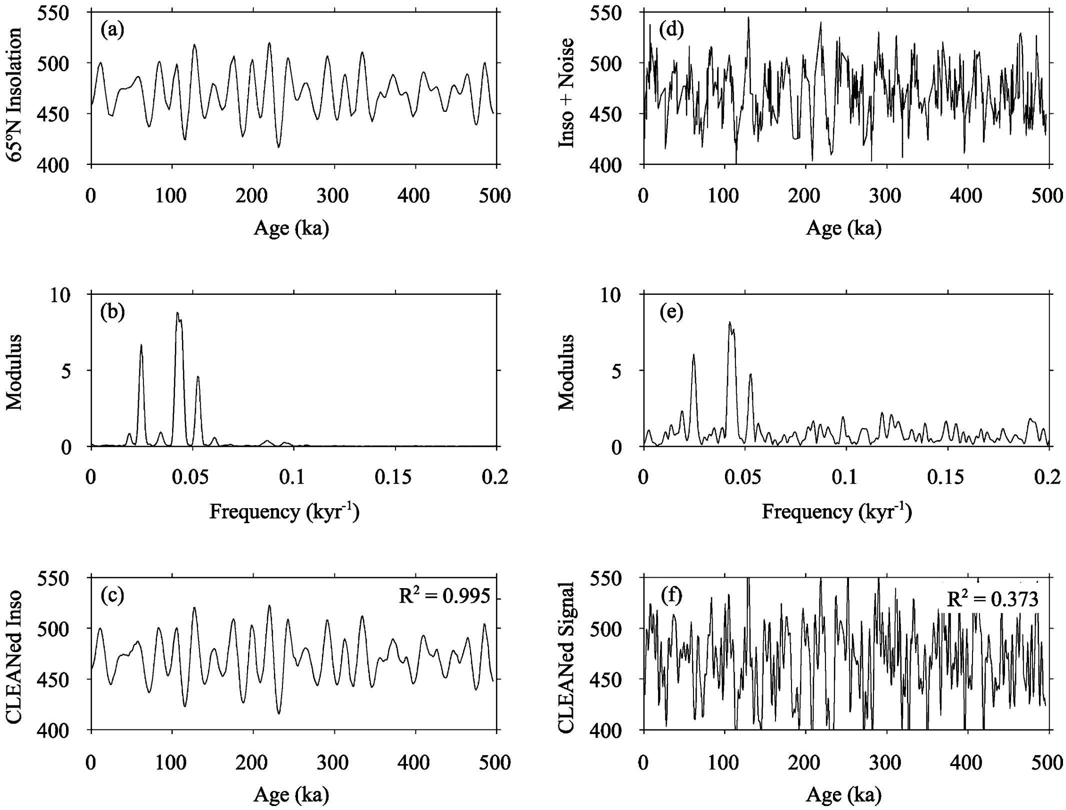

Fig. 4. (a) The 65°N insolation time series (0–500 ka) of Laskar (1990) interpolated onto a randomly spaced 400 point time array. (b) CLEAN frequency spectrum of the insolation curve shown in (a). (c) Inversion of the (b) frequency spectrum into the time domain. In a noise-free situation the CLEANed insolation curve is closely correlated to the input signal. (d) Production of a synthetic palaeoclimatic record using the insolation time series shown in (a) with the addition of normally distributed white noise contribution (scaled to 100% of the insolation magnitude). (e) CLEAN frequency spectrum of the synthetic palaeoclimatic time series shown in (d). (f) Reconstruction of the CLEANed synthetic palaeoclimatic time series in the time domain. This reconstruction shows that the CLEAN transformation, when performed in its traditional manner, cannot recover the original insolation curve to a reasonable extent from such a large noise contribution.

the 95% confidence limits about the mean modulus of each frequency increment. In the mean spectra the magnitudes of the peaks corresponding to the noise component (i.e. those appearing in Fig. 4e but not in Fig. 4b) are reduced. The ability to define a 95% significance value via the Monte Carlo procedure allows reconstruction of the signal using only frequencies above this level. Reconstruction of the signal above the 95% level for the 100, 200 and 300 point simulations produces time series that, while not identical to the original insolation curve, does show a very strong correlation to the Laskar time series (R^2 values of 0.957, 0.966 and 0.964, respectively). The successful extraction of such a good representation of the input insolation signal from a noise contribution of equal magnitude demonstrates the power of the CLEAN algorithm and its effectiveness when utilised in a Monte Carlo simulation.

3.2. White noise simulation

The analysis of the synthetic palaeoclimatic time series (irregular, 400 point, La_{90} insolation curve with an introduced 50% noise contribution) was repeated using the white noise addition procedure. Four simulations (each of 1000 iterations) were performed with noise scaled to 10, 25, 50 and 100% of the input signal magnitude (Fig. 6). In the first three

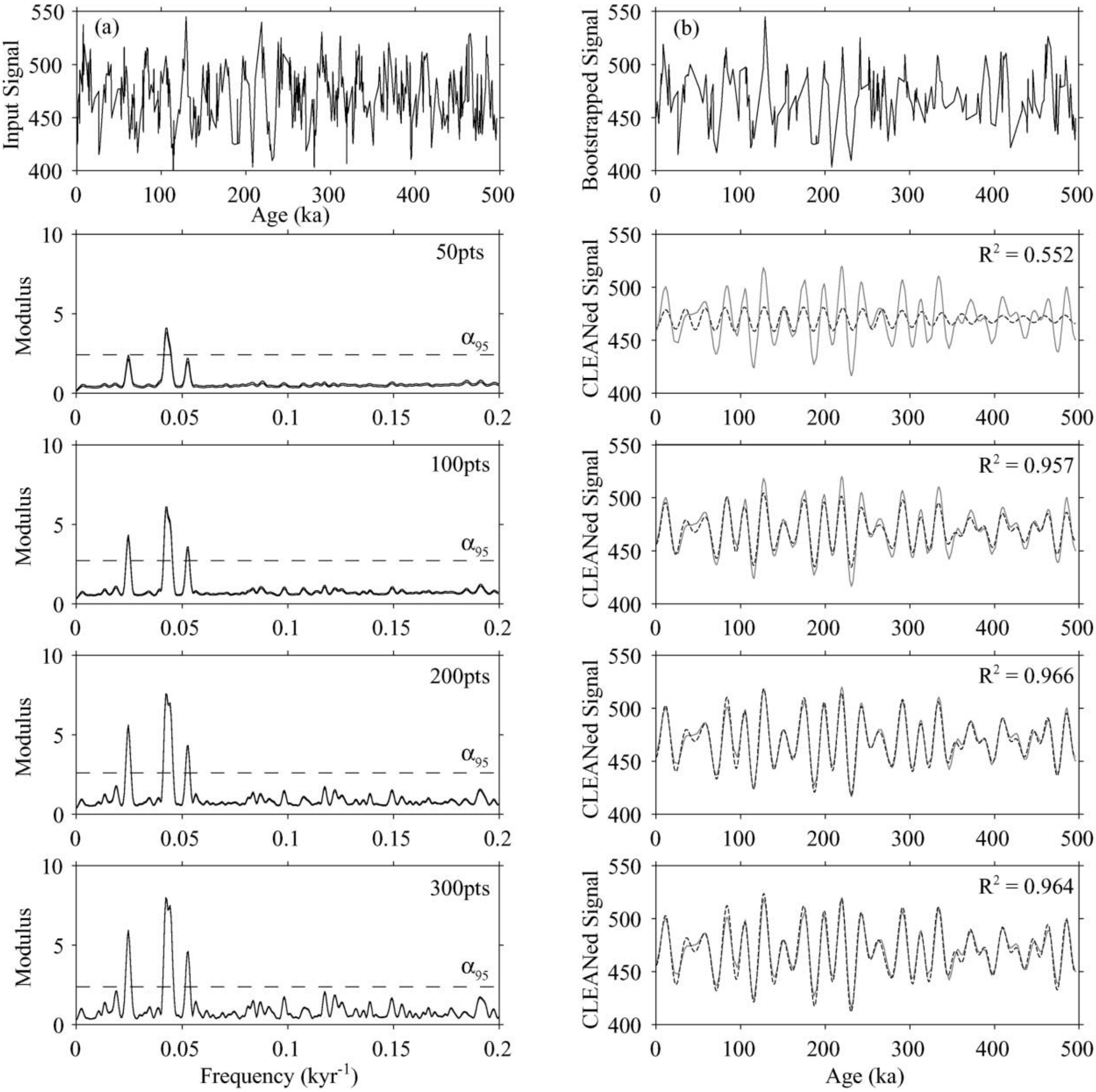

Fig. 5. (a) Synthetic palaeoclimatic time series utilised in the data stripping simulation. (b) Typical example of the synthetic palaeoclimatic time series after stripping to the 200 point level. Remaining panels: Monte Carlo CLEAN frequency spectra of the synthetic palaeoclimatic curve produced by stripping the data set to 50, 100, 200 and 300 points for 1000 iterations. The thickness of the shaded line indicates the 95% confidence limits about the mean spectrum at each frequency increment and the α_{95} significance levels are shown as dashed lines. Reconstructions of the frequency spectra in the time domain performed with only the frequency components above the α_{95} level included in the inversion process are shown on the right hand side of the diagram. The dashed lines correspond to the reconstructed signals, whilst the grey lines represent the noise-free insolation signal shown in 4(a).

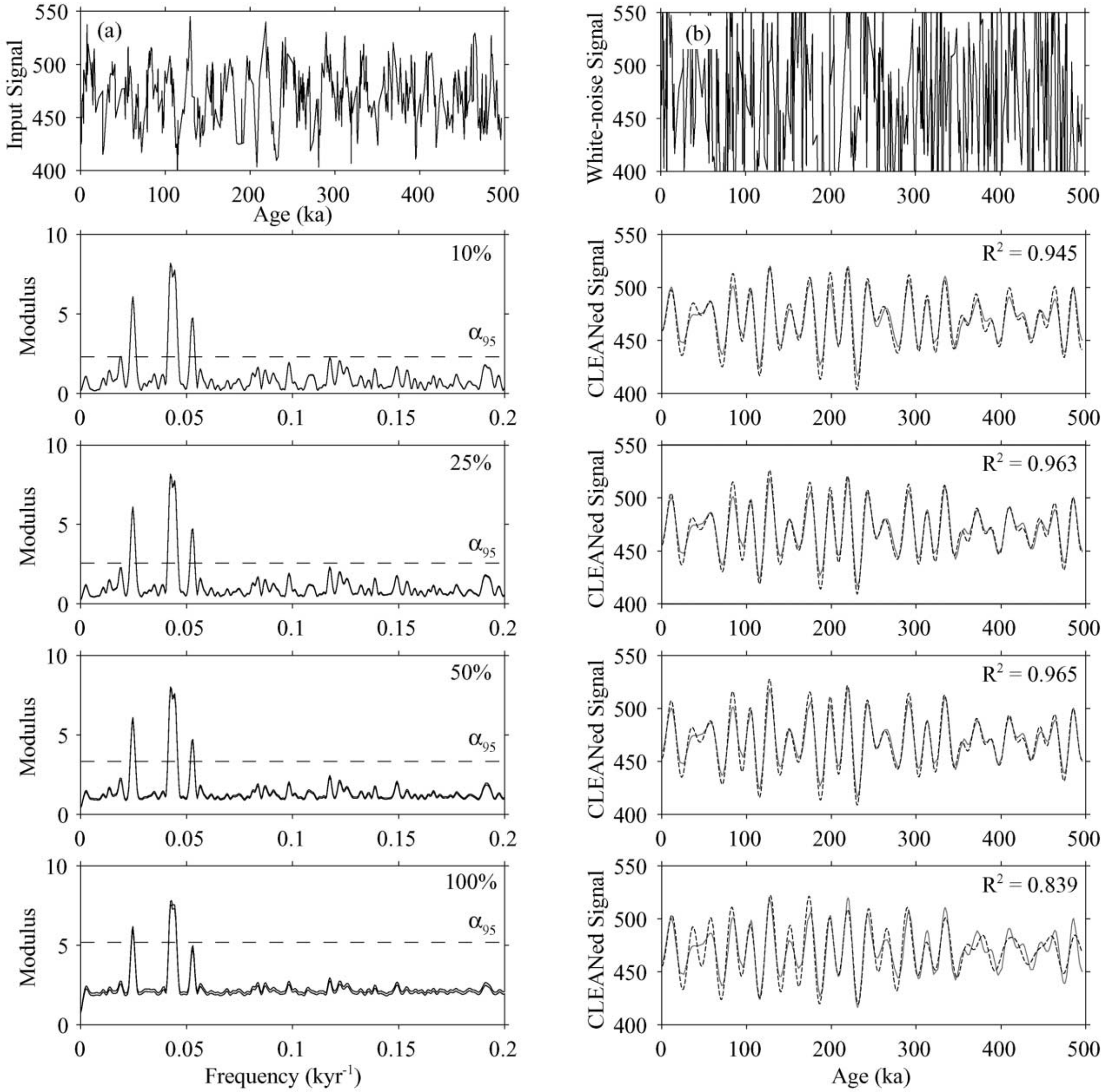

Fig. 6. (a) Synthetic palaeoclimatic time series utilised in the white noise simulation. (b) Typical example of the synthetic palaeoclimatic time series after addition of white noise at the 100% level. Remaining panels: results of the white noise simulations for the synthetic palaeoclimatic time series in the same format as Fig. 5 (noise scaled to 10, 25, 50 and 100%).

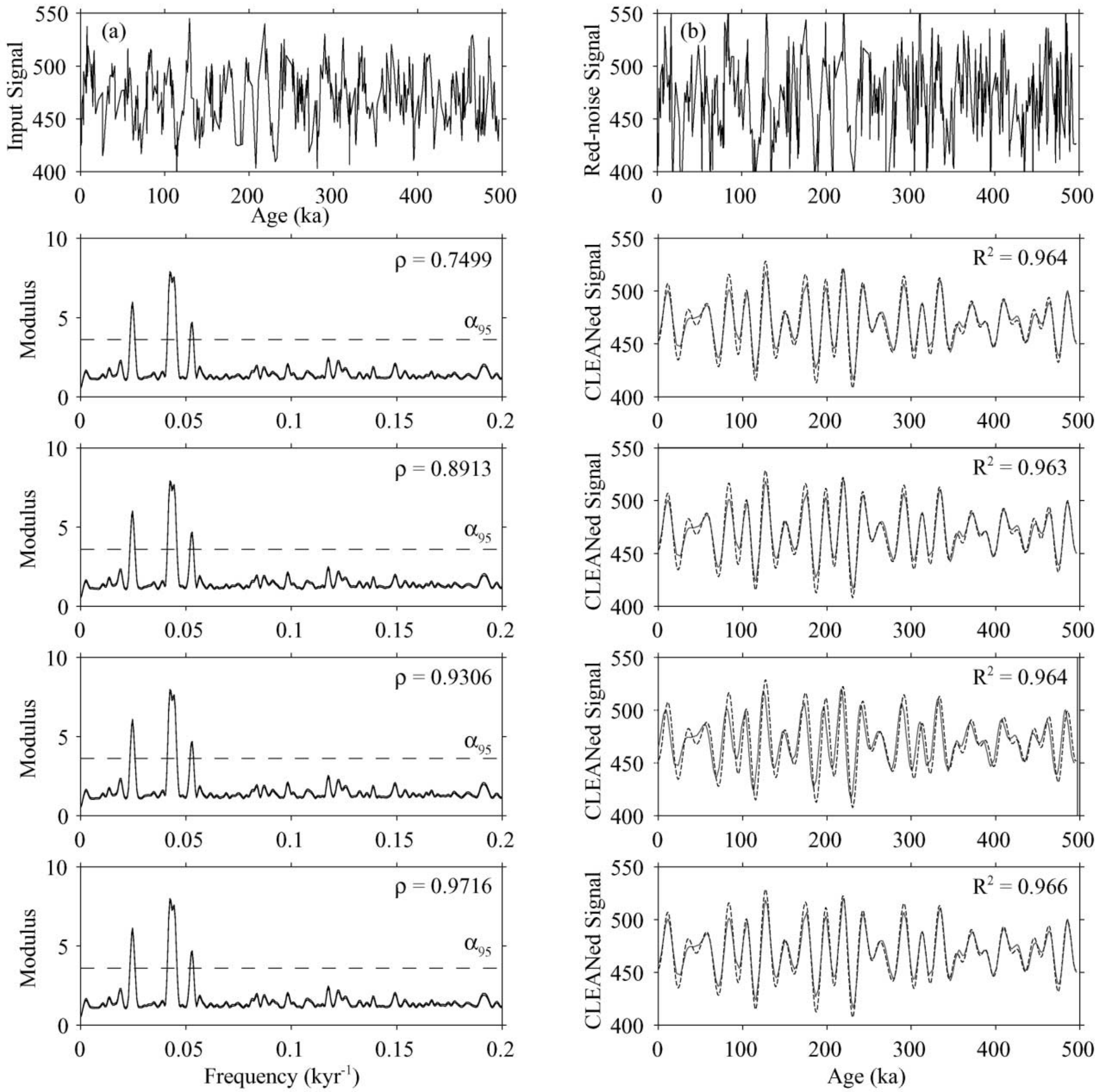

Fig. 7. (a) Synthetic palaeoclimatic time series utilised in the Red Noise simulation. (b) Typical example of the synthetic palaeoclimatic time series after addition of red noise for $\rho = 0.9306$. Remaining panels, results of the red noise simulations for the synthetic palaeoclimatic time series in the same format as Fig. 5 (using lag-one autocorrelation coefficients of 0.7499, 0.8913, 0.9306 and 0.9716).

simulations the signal reconstructed from the portions of the mean spectra above the α_{95} level show a very strong correlation with the insolation time series and the artificial noise introduced into the input signal has been successfully removed. In the case of the 100% noise addition, however, the significance limits about each of the points in the frequency spectrum are broadened and the highest frequency peak of the three main spectral components falls below the α_{95} level, resulting in a decrease in the correlation coefficient (reconstruction could be performed with a reduced significance level, e.g. α_{90}, to encompass the high frequency peak, however this would reduce the certainty of the recovered signal).

3.3. Red noise simulation

Analysis of the synthetic palaeoclimatic data was performed using the red noise simulation with ρ control values of 0.7499, 0.8913, 0.9306 and 0.9716 corresponding to dominant periodicites of 10, 25, 40 and 100 kyr^{-1}, respectively. Fig. 7 shows the results of the each of the four simulations determined after 1000 iterations. The mean spectra are extremely consistent between the four different simulations and each yields a robust reconstruction of the La_{90} insolation series in the time domain.

4. Real palaeoclimatic data

To demonstrate the application of the Monte Carlo CLEAN method to real data we investigated palaeoclimatic time series from the Vostok ice core (78°S, 106°E) spanning the last four glacial–interglacial cycles. Through the compositional analysis of gas bubbles trapped within the ice core, Petit et al. (1999) reconstructed variations in atmospheric O_2, CO_2, CH_4 and ^2H, and placed them within the chronological framework of their GT4 glaciological time scale. By the comparison of deuterium levels in the ice core (δD_{ice}) to marine $\delta^{18}O$ levels, Petit et al. (1999) reconstructed a proxy record of local atmospheric temperature levels (hereafter D_{temp}). Variations in the ^{18}O content of the ice core O_2 (hereafter $\delta^{18}O_{atm}$) have been shown to correspond to the changes in

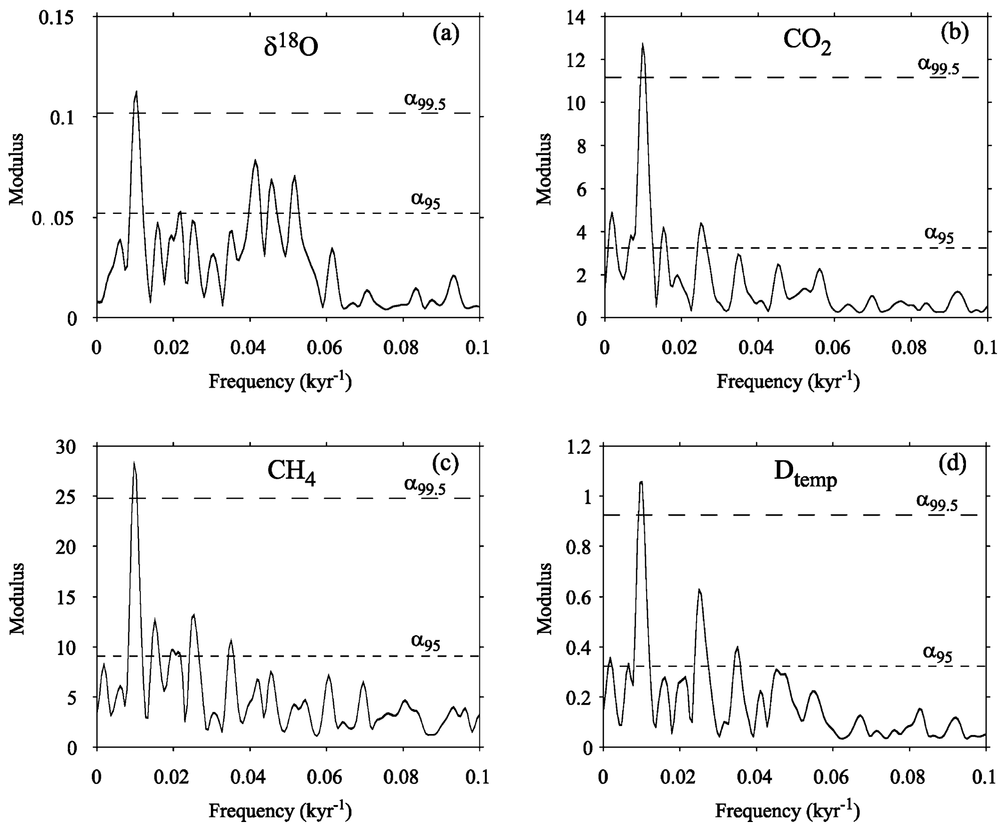

Fig. 8. Monte Carlo CLEAN frequency spectra for the Vostok $\delta^{18}O_{atm}$, CH_4, CO_2 and D_{temp} time series produced using the data stripping technique (data sets stripped from 318, 454, 283 and 3303 points, respectively to 50% of their original size, Iter = 1000). Significance levels at 95 and 99.5% are shown.

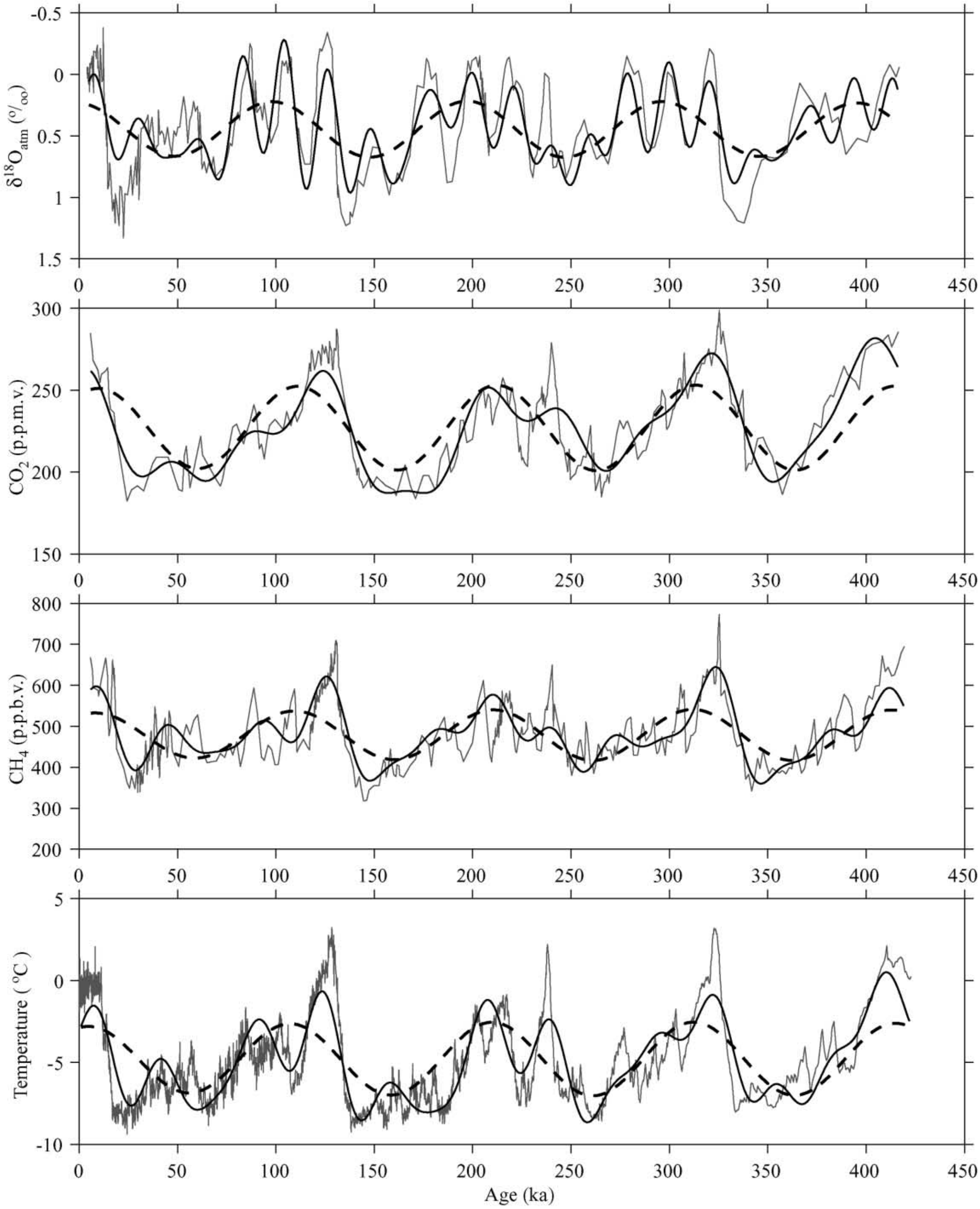

Fig. 9. Reconstruction of the Monte Carlo CLEAN frequency components above the 95 (solid black lines) and 99.5% significance levels (dashed lines). In each of the proxies, the reconstructed signal corresponds to eccentricity and is compared to the original time series (gray line). Labels ppmv and ppbv correspond to parts per million by volume and parts per billion by volume, respectively.

marine $\delta^{18}O$ that are controlled by global ice sheet volume (Shackleton, 2000). Shackleton suggested that the records from the Vostok core would demonstrate that eccentricity-driven variations in the global carbon cycle (controlling atmospheric CO_2 and CH_4 concentrations) were responsible for the $100\,kyr^{-1}$ glacial cyclicity of the Late Pleistocene.

To investigate the eccentricity variations recorded in the greenhouse gas concentration, global ice volume, and local temperature records contained in the Vostok core, we performed a data stripping simulation on each of the time series. Sampling density in the Vostok records is extremely variable both between records and within the individual records themselves. The effects of such a variability in a stripping procedure can lead to under and oversampling in certain regions of the curve, for this reason the magnitude of the stripping was not chosen to reflect the eccentricity signal but instead was arbitrarily set at 50% (Iter = 1000). The frequency spectra from each of the Monte Carlo simulations are shown in Fig. 8, with significance levels marked at both the 95 and 99.5% levels. In each

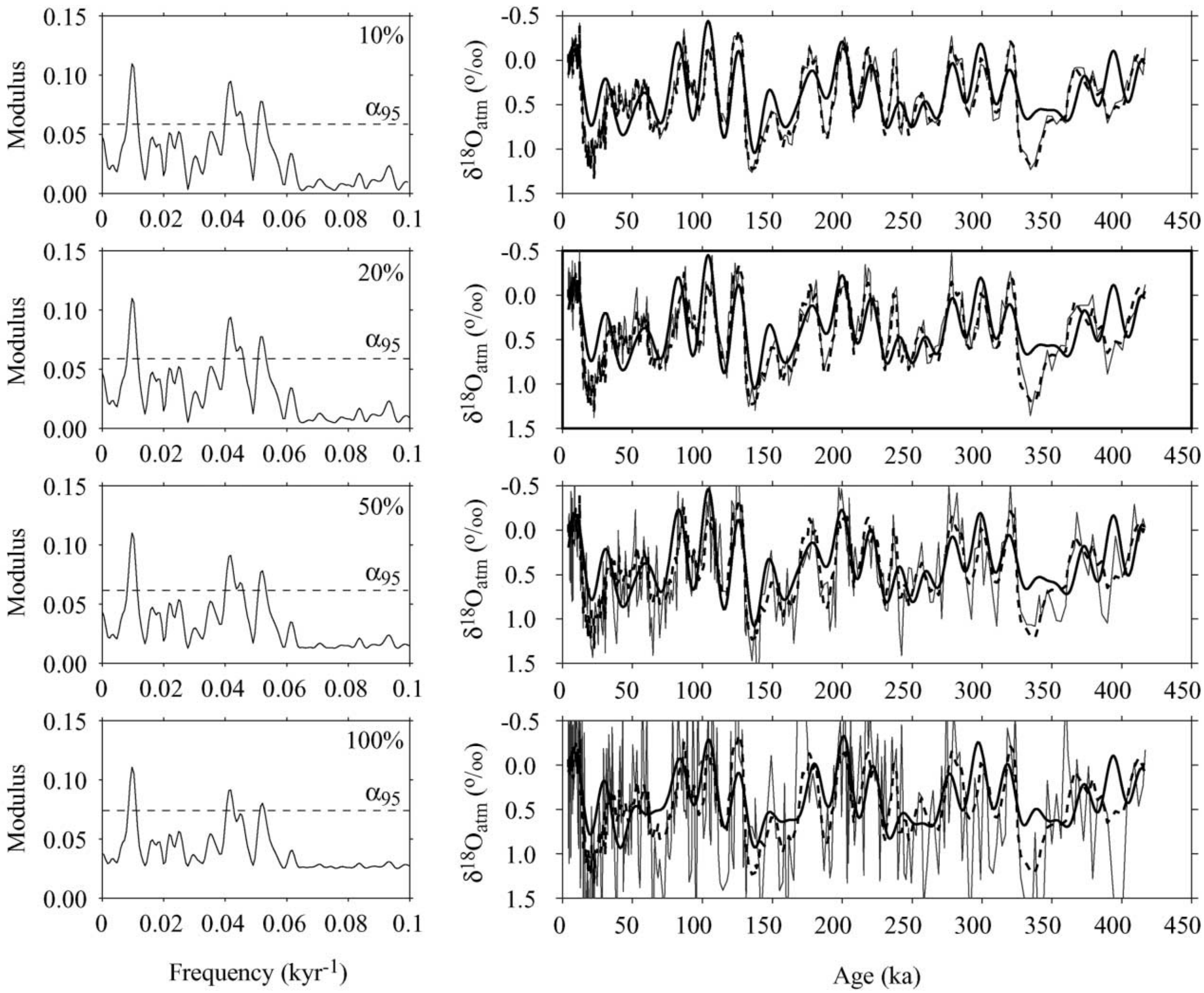

Fig. 10. Spectra and reconstructed time series produced by white noise simulations performed upon the Vostok $\delta^{18}O_{atm}$ record. The solid black lines represent the reconstructed signal whilst the dashed line is the original Vostok $\delta^{18}O_{atm}$ record. The series in gray gives an example of the data processed by CLEAN after addition of the white noise component (at the 10, 20, 50 and 100% levels).

of the frequency spectra the peak corresponding to eccentricity is isolated above the 99.5% level providing a clear threshold for reconstruction of the signals in the time domain.

Analysis of the time series after reconstruction at the 95 and 99.5% significance levels (Fig. 9) shows the different natures of the palaeoclimatic records. At the 95% level the $\delta^{18}O_{atm}$ signal shows a precession component that is substantially stronger than observed in the other three records. In contrast, the D_{temp} time series contains a large obliquity component that is present at a much weaker level in the eccentricity dominated CO_2 record. The CH_4 record reveals a combined influence of eccentricity and obliquity, which is most closely paralleled in the CO_2 signal. Reconstruction at the 99.5% level appears to successfully isolate the eccentricity contributions contained within the different proxy records. As with the high frequency components identified in the α_{95} reconstruction the contribution of eccentricity to overall signal variation does not appear to be constant between the different proxy parameters.

When analysed with a series of white noise simulations the $\delta^{18}O_{atm}$ spectrum appears to be extremely robust (Fig. 10). When noise is added with amplitudes equal to 10, 20 and 50% of that of the ice volume record, the Monte Carlo CLEAN procedure successfully extracts the eccentricity and precessional components of the record. In the case of 100% noise addition the higher frequency portion of the double precession peak centered at $\sim$0.045 kyr^{-1} falls below the α_{95} level. This results in a reduction in the amplitude of the reconstructed time series in certain portions of the record (e.g. at $\sim$65, 150 and 250 ka).

5. Conclusions

The development and application of the Monte Carlo CLEAN procedure to both real and synthetic data sets has allowed us to draw following four main conclusions:

1. Use of the CLEAN transformation in a Monte Carlo simulation allows confidence limits to be determined for each frequency value in the mean spectrum, and significance levels to be determined for the entire spectrum.
2. The three simulation types: data stripping, and white and red noise addition, produce consistent spectra for a synthetic palaeoclimate record, demonstrating that the Monte Carlo procedure produces stable results in the frequency domain.
3. The analysis of a synthetic palaeoclimatic time series effectively recovered a coherent insolation record that was hidden within a white noise contribution of equal magnitude. This analysis demonstrates the applicability of the Monte Carlo CLEAN procedure to the analysis of periodic signals contained within noisy palaeoclimatic records.
4. The reconstruction of a CLEANed signal in the time domain using the significance levels determined by the Monte Carlo technique allows specific periodic components to be isolated in the time series. Comparisons of the reconstructed Vostok time series demonstrate that the relative contributions of eccentricity, obliquity and precession are not constant between the different climatic proxies.

Acknowledgements

MJD thanks CEREGE for support during a sabbatical stay. Comments on the manuscript made by two anonymous reviewers are greatly appreciated. The authors are grateful to Dr. Andrew Shovlin for his assistance during the writing of the *MATLAB* software. This work was conducted under the programme of the Vening Meinesz Research School of Geodynamics.

References

Baisch, S., Bokelmann, G.H.R., 1999. Spectral analysis with incomplete time series: an example from seismology. Comps. Geosci. 25, 739–750.

Laskar, J., 1990. The chaotic motion of the solar system: a numerical estimate of the size of the chaotic zones. Icarus 88, 266–291.

Lomb, N.R., 1976. Least-squares frequency analysis of unequally spaced data. Astrophys. Space Sci. 39, 447–462.

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

Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J.M., Basile, I., Bender, M., Chappellaz, J., Davis, M., Delaygue, G., Delmotte, M., Kotlyakov, V.M., Legrand, M., Lipenkov, V.Y., Lorius, C., Pepin, L., Ritz, C., Saltzman, E., Stievenard, M., 1999. Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica. Nature 399, 429–436.

Roberts, D.H., Lehar, J., Dreher, J.W., 1987. Time series analysis with CLEAN. Part I. Derivation of a spectrum. Astron. J. 93 (4), 968–989.

Scargle, J.D., 1982. Studies in astronomical time series analysis. Part II. Statistical aspects of spectral analysis of unevenly spaced data. Astrophys. J. 263 (2), 835–853.

Scargle, J.D., 1989. Studies in astronomical time series analysis. Part III. Fourier transforms autocorrelation functions and cross-correlation functions of unevenly spaced data. Astrophys. J. 343 (2), 874–887.

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

Shackleton, N.J., 2000. The 100,000-year ice-age cycle identified and found to lag temperature, carbon dioxide, and orbital eccentricity. Science 289, 1897–1902.

Smith, S.W., 1997. The Scientist and Engineer's Guide to Digital Signal Processing. California Technical Publishing, California, 650 pp.