ELSEVIER

Earth and Planetary Science Letters 186 (2001) 363–370

EPSL

www.elsevier.com/locate/epsl

Scaling of power spectrum of extinction events in the fossil record

V.P. Dimri *, M. Ravi Prakash

National Geophysical Research Institute, Hyderabad 500 007, India

Received 22 September 2000; received in revised form 7 December 2000; accepted 10 January 2001

Abstract

The fast Fourier transform, maximum entropy method, Lomb's method of spectral analysis and rescaled range analysis are applied to the study of extinction patterns. Using a database of marine families from mid-Permian to Pleistocene, it is shown that a long-range correlation is present. Since the data record is non-stationary and unevenly spaced, linear interpolation is carried out for obtaining evenly spaced data. The data are also de-trended from their mean to obtain a stationary time series. Scaling behavior is observed in both interpolated and detrended unevenly spaced data. Application of the randomization test on both interpolated and de-trended data revealed that the interpolated data lost the randomness of the original record due to a smoothing effect, while the de-trended data retained the randomness property and hence are reliable for drawing information. The most popular method, the fast Fourier transform spectral method based on stationarity assumption, yields a contradictory result, and is independent of the interpolation technique used to fill gaps in the discontinuous fossil record. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: scale factor; spectral analysis; fossil record

1. Introduction

The topic of extinction (particularly of mass extinction) has been an important issue in paleobiology. The non-linear dynamics underlying extinction and diversification might have played an important role in the evolution of life [1,2]. The possible role of self-organization and criticality in the observed dynamics of extinction events of fossil records is a controversial topic. Recent papers [3–6] have explored the presence of trends in the fossil record by using different methods and reaching rather different results. Some analyses suggest the presence of long-range trends with fractal properties and others suggest random-like fluctuations. Newman and Eble [1] observed two different frequency regimes with different spectral properties in their analysis, one shows scaling, the other doesn't. In [3–5] the most common fast Fourier transform method has been applied to estimate power spectral density of the fossil record. It should be noted that the fossil record presents problems for spectral methods, which demand evenly spaced data, and a statistically stationary process. Most fossil time series do not

* Corresponding author. Tel.: +91-40-717-4433; Fax: +91-40-717-1564 and +91-40-717-0491; E-mail: postmast@csngri.ren.nic.in

0012-821X/01/$ – see front matter © 2001 Elsevier Science B.V. All rights reserved.
PII: S0012-821X(01)00246-1

fulfill the above requirements [5]. In such situations, one way of getting evenly spaced data from unevenly spaced data is by interpolating the data at desired points of time. For unevenly sampled data a completely different method of spectral analysis called Lomb–Scargle method with some desirable properties has been developed by Lomb and Scargle [7–9]. This method is used in the present study to examine long term correlation in the fossil record of marine families.

2. Data

In the present study the primary database for the analysis of fossil record is the record used by Sepkoski and Raup [10] and is derived from the compendium of fossil marine families [11]. Approximately 1800 marine animal families (excluding soft bodied taxa) have been described, of these 970 families are extinct. The data span 268 million yr (My) and consist of 43 stages from the mid-Permian (Leonardian) to Pleistocene providing an average resolution of 6.23 My on familial extinction. Four basic metrics of the intensity of familial extinction are considered. These are (a) simple number of extinction, (b) percent extinction, that is number of extinction relative to standing familial diversity, (c) total rate of extinction, that is the number of extinctions relative to stage duration, and (d) per family rate of extinction, that is total rate relative to standing diversity. Here we present only the results of the simple number of extinction (Fig. 1a) for 268 My, since the accuracy of ancient fossil record is questionable in the literature and the other metrics are derivatives of simple number and would yield the same results. The data points are unevenly spaced; the first and

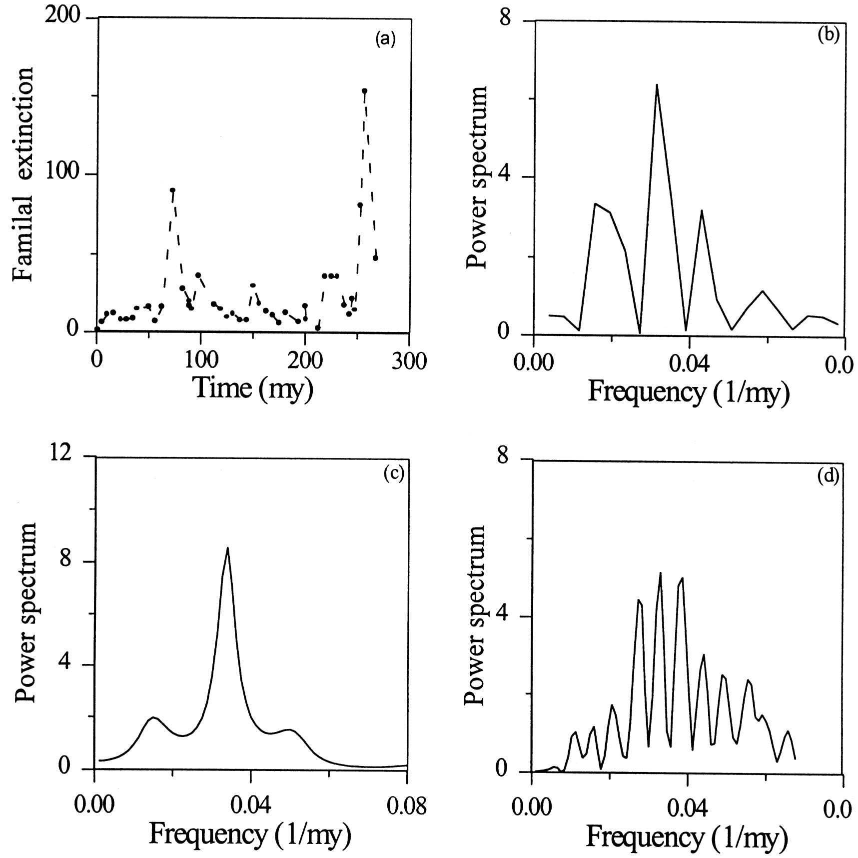

Fig. 1. (a) Variation in the extinction of families by the simple number over the last 268 My. The dashed line indicates the gaps in the data record. Power spectrum using (b) the FFT, (c) the MEM, and (d) the Lomb's method for de-trended data.

second order statistics (i.e. mean, variance, autocorrelation functions) are varying as a function of time, hence the data are non-stationary in nature. Using the available 43 unevenly spaced data points of the record, 67 evenly spaced data points are linearly interpolated.

3. Spectral methods

Spectrum estimation for evenly spaced data based on the fast Fourier transform (FFT) suffers from a major problem of resolution and is discussed by Dimri [12]. In this method an invalid assumption, namely zero data outside the duration of observation, is made. In the FFT one uses a window function and applies a finite Fourier transformation. The estimated spectrum is a convolution between the true spectrum and the spectrum of window function. The estimated spectrum will be a smeared version of the true spectrum. In the process of smearing, closely spaced spectral peaks get merged into a broader peak and the spectral power from strong peak leaks into its neighborhood, often leading to false estimation of spectral power. To minimize this effect, a high-resolution method known as the maximum entropy method (MEM) has been proposed by Burg [13]. The MEM estimates the spectrum that is most random or has the maximum entropy of any power spectrum and is consistent with the measured data. The MEM method does not make any assumption about the data outside the duration of observations. It completely eliminates the use of a window function for spectral estimation. This is achieved by fitting an auto-regressive model to the data and extrapolating the available data outside the duration of observations. Once the model parameters are estimated the spectrum may be derived from the model parameters themselves or from the extrapolated data. The MEM gives excellent result (better resolution) for short data length [14]. A high-resolution spectral method is defined as its ability to resolve two close frequency components. The advantage of the MEM over the FFT method regarding the correct frequency information and higher resolution has been demonstrated by Fouguer [15]. For unevenly spaced data, Lomb and Scargle developed a method for computing the power spectrum, that gives results superior to conventional methods [7–9]. The Lomb–Scargle method does not interpolate data points to fill the gaps and it weighs the data on a 'per point' basis instead of on a 'per time interval' basis, when uneven sampling can render the later seriously in error [8].

4. Trend analysis

A stationary time series is homogeneous and self-repeating in time. Its statistical properties, viz. mean, variance, and all higher moments, remain invariant in time, whereas a non-stationary time series is such that its statistical properties change with time. A non-stationary time series is often modeled as a sum of two components, viz. drift and residual:

$$X(t) = D(t) + R(t) \tag{1}$$

where $D(t)$ is the drift and $R(t)$ is the residual. The drift consists of the average value of the variable within a neighborhood, which varies slowly and forms the non-stationary part and reflects the large-scale tendencies of the phenomenon. The drift at a point 't' may be expressed as a polynomial. The residual is the difference between the actual measurement and the drift. It is the random aspect and accounts for short-scale variations. If the drift is removed from a non-stationary variable the residual becomes stationary. The residuals are not errors but contain full fledge features of the phenomenon. In the present study trend analysis is performed to ensure stationarity, a prerequisite for obtaining a reliable spectrum. The trend in the data is computed by fitting a third order orthogonal polynomial of the type:

$$D(t) = B_0 f_0(t) + B_1 f_1(t) + B_2 f_2(t) + B_3 f_3(t) \tag{2}$$

where $f_0(t) = 1$, $f_1(t) = t$, $f_2(t) = 2t^2 - 1$ and $f_3(t) = 4t^3 - 3t$ are Chebyshev polynomials and $B_0, \ldots, B_3$ are coefficients of the polynomial. The Chebyshev polynomials are orthogonal polynomials and are more efficient in terms of computational time and convergence rate. These are also the minimax polynomials, which (among all polynomials of the same degree) have the smallest maximum deviation from the true function [8,16]. The trend in mean is removed from both evenly and unevenly spaced data. The residual, which is stationary, is used for computing the power spectrum. The principle of parsimony (minimum model parameters) is adopted in selecting the order of the polynomial so that the problem of over-fitting the data is avoided.

5. Randomization test

When a transformation such as interpolation or the de-trending is applied to data and is subjected for drawing some conclusion, then the conclusion is valid only if the transformed data retain all information and properties of the original record. A run test for randomization is one such procedure [17] that can tell us whether the transformed data retain the randomness property of the original data or not. It is based on the sequence in which the observations are obtained. We closely follow [17] to describe the randomization test. A run is defined as a successive of identical symbols, which are followed and preceded by different symbols. For example, suppose a series of plus and minus scores occurring in the order +++ −− ++ −−−. This sample of scores begins with a run of three pluses. A run of two minuses follows, then comes another run of two pluses followed by a run of three minuses. In practice the median of the data is computed, then all observation values falling below median are designated as minus and all above the median are designated as plus. The null hypothesis H_0 (the order of observations above and below median in the sequence is random) is tested against its alternative hypothesis H_1 (the order of observation above and below median in the sequence is not random). Let n_1 be the number of pluses, n_2 be the number of minuses, then n_1+n_2 will be the total number of observations, and let r be the number of runs. For large samples say either n_1 or n_2 is larger than 20, a good approximation to the sampling distribution of 'r' is the normal distribution with:

$$\text{mean} = \mu = (2n_1n_2)/(n_1+n_2)+1 \quad (3)$$

$$\text{standard deviation} = \sigma = \sqrt{\,}\ [(2n_1n_2(2n_1n_2-n_1-n_2))/(n_1+n_2)^2(n_1+n_2)] \quad (4)$$

$$\text{standard normal variate} = Z = (r-\mu)/\sigma \quad (5)$$

The values of Z under H_0 are approximately normally distributed with zero mean and unit variance. The significance of the Z value may be determined by referring to the normal table. Randomization test is carried out both for interpolated and de-trended (residuals) data (Table 1). The computed Z value for the interpolated data is −6.81 and is greater than the table value of Z = Abs|2.33| at 1% of level of significance (l.o.s.), hence the null hypothesis H_0 is rejected. The computed Z value for the de-trended data is −2.00 and is less than the table value of Z at 1% l.o.s., hence the alternative hypothesis H_1 is rejected in this case. The test clearly confirms that the interpolated data have lost the randomness of the original data due to smoothing effect, while the de-trended data contain all the randomness of the data and are reliable for drawing the information of the original population.

6. Spectral analysis

Power spectrum is computed (Figs. 1b,c, 2a,b)

Table 1
Randomization test for interpolated and de-trended data

Data	n_1	n_2	R	Median	μ	σ	Z value
Interpolated	33	31	6	20.34	32.97	3.96	−6.81
De-trended	22	21	16	−4.50	22.49	3.28	−2.00

Z value from normal table at 1% l.o.s. is Abs|2.33|.

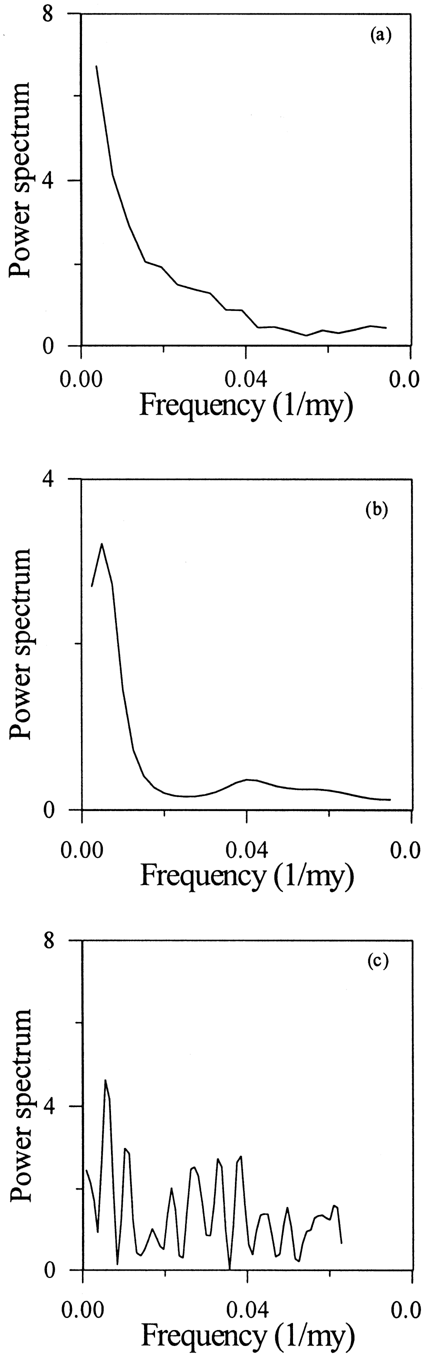

Fig. 2. Power spectrum using (a) the FFT, (b) the MEM for interpolated data, and (c) the Lomb's method for actual data.

using the FFT and the MEM [13,14] both for interpolated data and their residuals (evenly spaced). The Lomb–Scargle method [7–9] is adopted for the actual data (unevenly spaced) and their residuals (Figs. 2c, 1d). Since the interpolated data are no longer random, we discuss the power spectra of the de-trended and the actual data only. The power spectra of de-trended data have better resolution and are more reliable as they are based on stationary data compared to the actual data. A comparison is made between the power spectra obtained for residuals by different methods. The spectrum obtained by the FFT (Fig. 1b) has a peak at a frequency of 0.0313 corresponding to a periodicity ($1/f$) of 32 My. For the spectrum obtained by the MEM (Fig. 1c) the peak is at 0.03375 corresponding to a periodicity of 29 My. The spectrum obtained by Lomb's method (Fig. 1d) has two peaks which have nearly equal power and appear at frequencies of 0.03285 and 0.03848 respectively, suggesting an intermediate frequency equal to 0.03566 which corresponds to a period around 28 My. Plots of the logarithm of power spectrum and logarithm of frequency (Figs. 3b,c, 4a–c) reveal a power law relationship between power spectrum $P(f)$, and frequency f, of the form $P(f) \propto f^{-\beta}$, where β is scaling exponent. Scaling exponents obtained from the slopes of the linear least squares fits of the power spectra based only on de-trended data are analyzed. A non-significant scaling exponent value of 0.06 is obtained for the power spectrum by the FFT method (Fig. 3a). This result confirms the earlier finding [5] and also holds for the higher frequencies [3]. The other two spectra obtained by the MEM (Fig. 3b) and the Lomb–Scargle method (Fig. 3c) gave similar results yielding significant scaling exponent values of -0.80 ± 0.12 and -1.09 ± 0.12 respectively. The MEM and Lomb's method are more appropriate and are more reliable for evenly and unevenly spaced data respectively as they give better resolution. However, the scaling exponents for interpolated data using different spectral methods are shown in Fig. 4a–c.

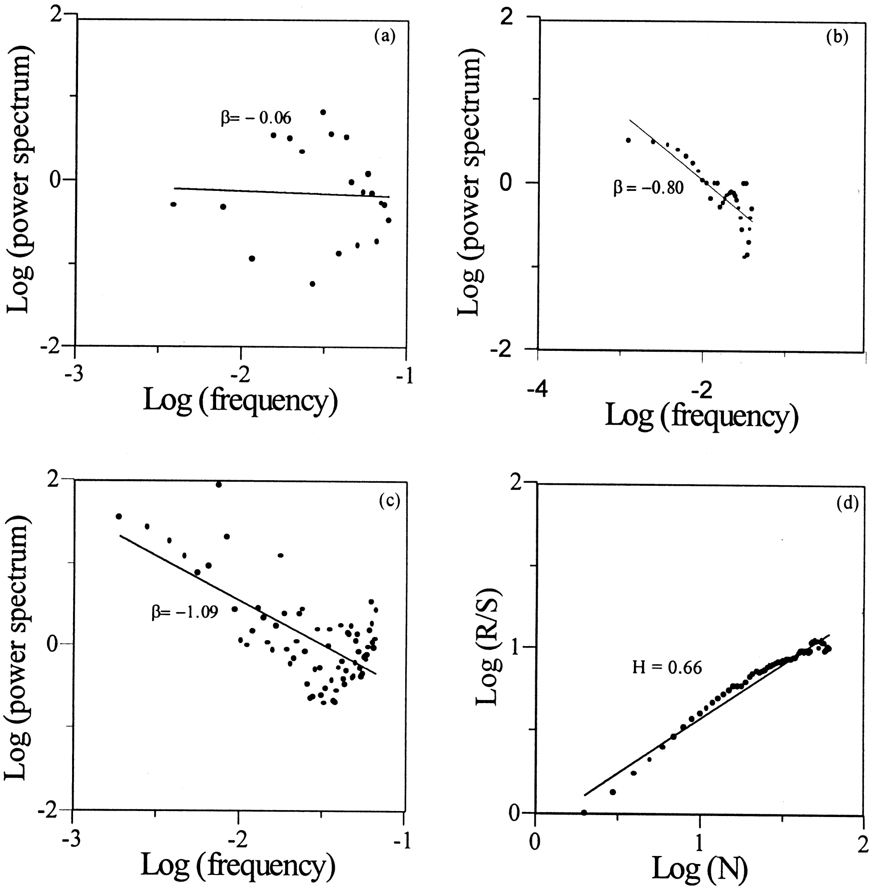

Fig. 3. Graphs on log–log scale for the *P(f)* versus *f*, for estimating scaling exponent using the least squares fit for the power spectrum obtained by (a) the FFT, (b) the MEM, (c) the Lomb's method, and (d) the rescaled range analysis for estimating Hurst coefficient for the de-trended data.

7. Rescaled range analysis

Finally, the rescaled range (*R/S*) analysis has also been carried out both for the actual and interpolated data. A plot between the range/standard deviation (*R/S*) and logarithm of number of data points N (Figs. 3d, 4d) reveals a scaling law of the form $(R/S) \propto N^H$, where H is the Hurst coefficient which is obtained from the slope of the linear least squares fit of the above plot. The value of H determines the scaling behavior of a process and if $H > 0.5$, the process exhibits persistence [18] or long term correlation [19]. For drawing inference we consider the Hurst coefficient value computed only from the de-trended data. A significant value of Hurst coefficient ($H = 0.66 \pm 0.02$) further supports that the fossil data exhibit long term correlation. The typical value of $H = 0.7$ suggests that the time series (fossil record) is non-stationary but consistent with long term correlation [20], which agrees to our assumptions of fossil record.

8. Conclusions

1. The fractal pattern in the fossil record is thus properly exposed by our method, which consists of de-trending the data. The de-trended data retain the randomness of the original data and hence are reliable for drawing information.

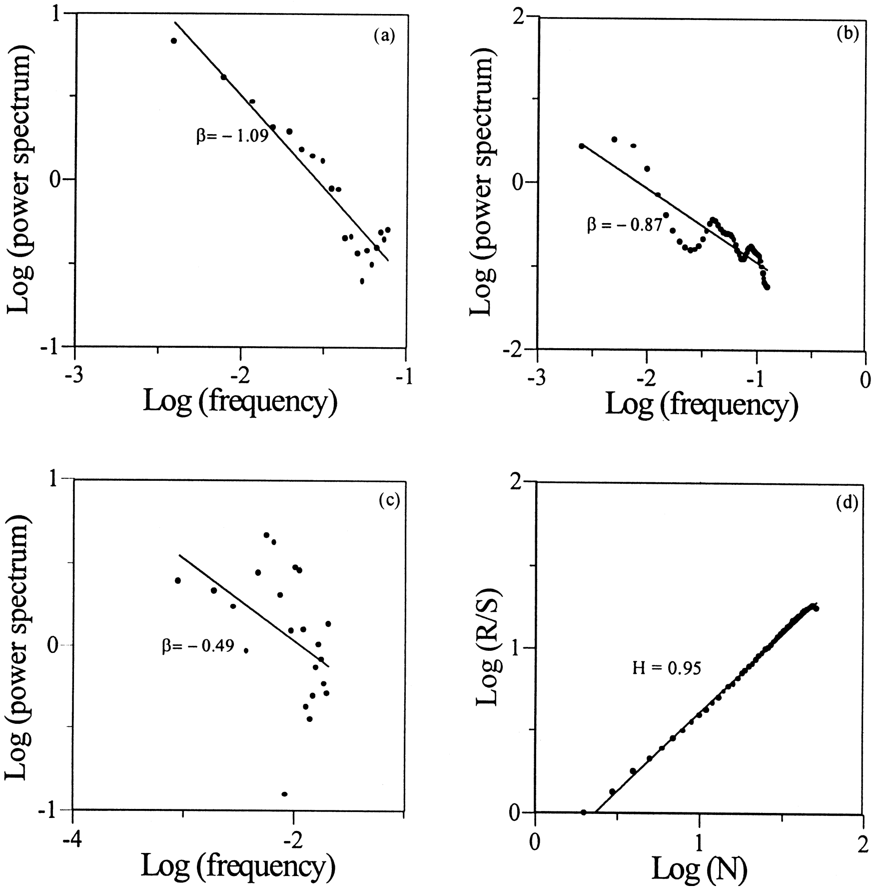

Fig. 4. Graphs on log–log scale for the $P(f)$ versus f for estimating scaling exponent using the least squares fit for the power spectrum obtained by (a) the FFT, (b) the MEM, (c) the Lomb's method, and (d) the rescaled range analysis for estimating Hurst coefficient for the interpolated data.

2. The study supports the theory, but more such studies have to be carried out to provide an insight into the mechanism for the non-linear dynamics of extinction events.
3. The Lomb–Scargle method is essentially developed for the uneven astronomical time series data [7,8] and is also applied to climate data [9]. Our results focus the attention of a large number of scientists engaged in periodicity problems such as the problem of mass extinction, to the application of the Lomb–Scargle method to naturally occurring uneven geological data.

Acknowledgements

We are grateful to Dr. H.K. Gupta, FNA, Director N.G.R.I., Hyderabad for his kind permission to publish the paper. Thanks are also due to our colleague Mr. Abhey Ram Bansal in preparing the manuscript. Our sincere thanks to all anonymous reviewers for their constructive comments, particularly for the use of the randomization test for the de-trended data, which has improved the manuscript. ***[SK]***

References

[1] S. Conway Morris, The evolution of diversity in ancient

ecosystems, a review, Philos. Trans. R. Soc. Lond. B 353 (1999) 327–345.

[2] R.V. Sole, S.C. Manrubia, J. Bascompte, Extinctions: bad genes or bad luck?, Proc. R. Soc. Lond. B 26 (1996) 1407–1413.

[3] M.E.J. Newman, G.J. Eble, Power spectra of extinction in the fossil record, Proc. R. Soc. 266 (1999) 1267–1270.

[4] R.V. Sole, S.C. Manrubia, M. Benton, P. Bak, Self-similarity of extinction statistics in the fossil record, Nature 388 (1997) 764–766.

[5] J.W. Kirchner, A. Weil, No fractals in fossil extinction statistics, Nature 395 (1998) 337–338.

[6] M.E.J. Newman, G. Eble, Decline in extinction rates and scale invariance in the fossil record, Paleobiology 25 (1999) 434–439.

[7] N.R. Lomb, Least-square frequency analysis of unequally spaced data, Astrophys. Space Sci. 29 (1976) 447–462.

[8] W.H. Press et al., Spectral analysis of uneven sampled data, in: W.H. Press, S.A. Teukolsky, W.T. Vettering, B.P. Flannery (Eds.), Numerical Recipes, Cambridge University Press, London, 1992, pp. 569–573.

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

[10] J.J.Sepkoski Jr., D.M. Raup, Periodicity in marine extinction events, in: D.K. Eliot (Ed.), Dynamics of Extinction, Wiley, New York, 1986, pp. 3–35.

[11] J.J.Sepkoski Jr., A compendium of fossil marine families, Milwaukee Public Museum Contributions in Biology and Geology no. 51, 1982, 125 pp.

[12] V.P. Dimri, Deconvolution and Inverse Theory, Elsevier Science Publishers, Amsterdam, 1992, pp. 185–203.

[13] J.P. Burg, Maximum entropy spectral analysis, Ph.D. thesis (unpublished), Standford University, 1975, 123 pp.

[14] T.J. Ulrych, T.N. Bishop, Maximum entropy spectral analysis and autoregressive decomposition, Rev. Geophys. Space Phys. 13 (1975) 183–200.

[15] F.P. Fouguer, On the accuracy of spectrum analysis of red noise processes using maximum entropy and periodogram methods: Simulation studies and application to geophysical data, J. Geophys. Res. 90 (1985) 4355–4366.

[16] E. Isaacson, H.B. Keller, Chebyshev polynomials, in: E. Isaacson, H.B. Keller (Eds.), Analysis of Numerical Methods, Wiley, New York, 1966, p. 535.

[17] S. Siegel, The one sample runs test, in: S. Siegel (Ed.), Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill, New York, 1956, p. 303.

[18] G. Rangarajan, Fractals, in: V.P. Dimri (Ed.), Application of Fractals in Earth Sciences, A.A. Balkema, Rotterdam, 2000, p. 238.

[19] G. Korvin, Fractal Models in the Earth Sciences, Elsevier Science Publishers, Amsterdam, 1992, 396 pp.

[20] D.C. Boes, Schemes exhibiting Hurst behavior, in: J.N. Srivastava (Ed.), Probability and Statistics Assay, Elsevier Science Publishers, Amsterdam, 1988, pp. 21–42.