IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 119, No. 1, January – March, 2015, pp. 53–68

Estimating spectra of unevenly spaced climatological time series

István Matyasovszky

Department of Meteorology, Eötvös Loránd University,
Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary,
E-mail: matya@caesar.elte.hu

(Manuscript received in final form May 3, 2014)

Abstract—Spectral analysis is often based on a comparison of the periodogram and the spectral density of a so-called background noise. This spectral density is estimated by fitting a first order autoregressive (AR(1)) process to data, as climatological time series generally exhibit red noise spectra that can be approximated by AR(1) models. When periodogram exceeds some threshold at a frequency, the spectrum is said to differ from this background noise, and the frequency is characteristic for the time series in question. The traditional periodogram, however, must not be used without modifications for unevenly spaced data. Additionally, red noise, characterized by spectral densities monotone increasing to low frequencies, covers a much wider class of processes than the AR(1) processes. Our purpose is (1) to introduce a new periodogram (ELSP) based on a least square (LS) fit for an entire set of frequencies instead of using the well-known Lomb-Scargle periodogram (LSP) based on individual LS fits for individual frequencies; (2) to estimate the spectral density without any assumption on its analytical form using the nearly isotonic regression (NIR) method with either ELSP or LSP. As NIR allows the possibility of deviations from red noise, comparison of the periodogram with a background noise is unnecessary. Note that ELSP has never been used before as is a new concept for defining the periodogram for unevenly spaced data. NIR is more or less known for curve fitting problems but has not been applied yet to spectral density estimation. Three examples show that although ELSP does not radically differ from LSP, NIR-ELSP and NIR-LS spectra can exhibit distinct shapes.

Key-words: spectra, unevenly spaced data, Lomb-Scargle periodogram, red noise, nearly isotonic regression

1. Introduction

Literature of spectral analysis of climatological time series is extremely broad. The task of the spectral analysis is to identify sets of frequencies that essentially contribute to the behavior of time series. A common way is to calculate the

periodogram and then to fit a first order autoregressive (AR(1)) process to data in order to model the so-called background noise. When periodogram exceeds some threshold at a frequency, the spectrum is said to differ from this background noise, and the frequency is characteristic for the time series in question. The threshold depends on the AR(1) model and the significance level selected.

Let $x(t_1),...,x(t_n)$ be a stationary time series observed at $t_1,...,t_n$. Usually, the data set is evenly spaced, and $t_i = i$ can be taken. Hence, the time series can be written as $x_1,...,x_n$. The background noise is taken as red noise, and is generally described with the spectral density

$$f(\lambda) = (\sigma_e^2 / \pi)/(1 + a^2 - 2a\cos(\lambda)),\ \sigma_e^2 = \sigma^2(1 - a^2) \qquad (1)$$

of an AR(1) process with substituting the autoregressive parameter a and variance σ^2 with their consistent estimates $\hat{a}$ and $\hat{\sigma}^2$ obtained from $x_1,...,x_n$. As Eq. (1) provides red noise spectra under positive a, red noise and AR(1) spectra are seldom used as synonyms. But red noise, characterized by spectral densities monotone increasing to low frequencies, represents a much wider class of processes than the AR(1) processes, and the usage of AR(1) spectra can thus fail to properly detect frequencies mainly contributing to spectra. Therefore, we propose a method that ignores the comparison of periodograms with background noise models.

In some cases, time series are unevenly spaced, and hence the periodogram defined for evenly spaced data must not be used without modifications. There are two main ways to handle the problem. The first one is based on an interpolation of the data onto an equispaced time grid, and this new regularly spaced data set is analyzed with traditional techniques. Such data manipulations, however, always deform the true spectra (*Broersen*, 2006). The other way, which is addressed in this paper, produces a reformulated periodogram directly from data. Evidently, such a periodogram is always affected by temporal distributions of data spacing. Generally, the Lomb-Scargle periodogram (LSP) (*Lomb*, 1976; *Scargle*, 1982), based on a simple least squares (LS) estimation procedure, is used for the purpose. However, important statistical properties of LSP, e.g., its probability distribution are known only for white background noise. Additionally, the bias of LSP for unevenly spaced data can be substantially higher than that of the periodogram for evenly spaced data (*Vio et al.*, 2010), principally at high frequencies (*Schulz* and *Mudelsee*, 2002). Therefore, we will examine first the properties of LSP. Then we will propose a periodogram that is based on a so-called entire least squares (ELS) technique (*Matyasovszky,* 2013a). Properties of this newly introduced periodogram will also be discussed.

The methodology is described in Sections from 2 to 4. As paleo-climatological records represent a typical case of unevenly spacing, our technique is demonstrated by three paleoclimatic records in Section 5. Finally, a section for discussion and conclusions is provided.

2. Periodograms

2.1. Lomb-Scargle periodogram (LSP)

Let $x(t_1), x(t_2),..., x(t_n)$ with $t_1 = 1, t_n = N$ be a time series coming from a stationary stochastic process with mean zero. The LSP for any frequency λ_j in the interval $(0, \pi]$ is based on an LS procedure as follows. Parameters a_j and b_j that minimize

$$\sum_{i=1}^{n} (x(t_i) - a_j \cos(\lambda_j t_i) - b_j \sin(\lambda_j t_i))^2$$

are obtained with the solution of the system of equations

$$\underline{\underline{D}}\,\underline{c} = \underline{w}, \tag{2}$$

where $\underline{\underline{D}} = \underline{\underline{Z}}^T \underline{\underline{Z}}$, and the elements of $\underline{\underline{Z}}$ are $z_{i1} = \cos(\lambda_j t_i), z_{i2} = \sin(\lambda_j t_i), i = 1,..., n$, furthermore $\underline{w} = \underline{\underline{Z}}^T \underline{x}$, and $\underline{x} = (x(t_1),..., x(t_n))^T$, $\underline{c} = (a_j, b_j)^T$ with superscript T denoting transpose. The quadratic form

$$I_{LS}(\lambda_j) = 1/(2\pi)\,\underline{c}^T \underline{\underline{D}}\,\underline{c} \tag{3}$$

defines LSP. For evenly spaced data, Eq. (3) becomes to the well-known expression

$$I(\lambda_j) = n/(4\pi)(a_j^2 + b_j^2) \tag{4}$$

with

$$a_j = 2/n \cdot \sum_{i=1}^{n} x_i \cos(\lambda_j i), \quad b_j = 2/n \cdot \sum_{i=1}^{n} x_i \sin(\lambda_j i).$$

It is known from the LS procedure that $\underline{c}$ is asymptotically normally distributed under very general conditions. Supposing that $\underline{x}$ comes from a white

noise process, the covariance matrix of $\underline{c}$ is $\sigma^2 \underline{\underline{D}}^{-1}$, where σ^2 is the variance of the process. Hence, the random variable $I_{LS}(\lambda_j)$ defined with the quadratic form in Eq. (3) follows an exponential distribution (*Scargle*, 1982). The probability distribution of LSP when $\underline{x}$ does not come from a white noise process is, however, an open question. Therefore, we will discuss this issue together with the expected value of $I_{LS}(\lambda_j)$ in Section 4. Although λ_j can be any of frequencies in an interval $[\lambda_{\min}, \lambda_{\max}]$, it is advisable to define a grid $2\pi j/(n\Delta), j = 1,...,L$, where L is the largest integer not larger than $n/2$, $\lambda_{\max} = \pi/\Delta$ is the average Nyquist frequency (*Stoica et al.,* 2009), and Δ is the average of time steps $\delta_i = t_i - t_{i-1}, i = 2,...,n$. As the frequency range of $I(\lambda)$ is $(0,\pi]$, the range $[\lambda_{\min}, \lambda_{\max}]$ of $I_{LS}(\lambda)$ is generally rescaled into $[2\pi/n, \pi]$ for convenience. $I_{LS}(\lambda)$ can thus be viewed as $I(\lambda)$ of a time series sampled evenly at time steps Δ. An important difference is, however, that elements of $I_{LS}(\lambda_j)$ are correlated in contrast to elements of $I(\lambda_j)$. Additionally, the bias of $I_{LS}(\lambda_j)$ can be higher than the bias of $I(\lambda_j)$ (*Vio et al*., 2010). This is due to the interrelationship between unevenly spacing and the effect of the omission of frequencies different from λ_j when calculating $I_{LS}(\lambda)$ at λ_j. Furthermore, it is known that $I(\lambda)$ integrates to the sample variance $\hat{\sigma}^2$ in the sense that

$$\frac{2\pi}{n}\sum_{j=1}^{L} I(\lambda_j) = \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x(t_i) - \bar{x})^2,$$

while $I_{LS}(\lambda)$ does not integrate to $\hat{\sigma}^2$.

2.2. Entire least squares periodogram (ELSP)

The deficiency of LSP is that it handles the different frequencies separately. Therefore, we propose an entire least squares (ELS) procedure by calculating the constants a_j, b_j at frequencies $2\pi j/(n\Delta), j = 1,...,L$ simultaneously. This results in a system of equations Eq. (2), but with

$$z_{ij} = \left.\begin{cases}\cos(\lambda_j t_i), j = 1,...,L \\ \sin(\lambda_j t_i), j = L+1,...,2L\end{cases}\right\}, \quad i = 1,...,n, \quad \underline{c} = (a_1,...,a_L,b_1,...,b_L)^T.$$

It is easy to see that $\underline{c}^T \underline{\underline{D}}\underline{c} = n\hat{\sigma}^2$. As $\underline{\underline{D}}\underline{c} = \underline{w}$, hence $\underline{c}^T \underline{w} = c_1 w_1 + ... + c_{2L} w_{2L} = n\hat{\sigma}^2$. Therefore, we define ELS periodogram (ELSP) as the contribution of frequencies to the sample variance, i.e.,

$$I_{ELS}(\lambda_j) = 1/(2\pi)(c_j w_j + c_{j+L} w_{j+L}). \quad (5)$$

Although every frequency λ_j is affected by frequencies not involved in the estimation procedure, the accuracy of ELSP is expected to be higher than the accuracy of LSP. This is because ELSP is defined for the entire set of frequencies $\lambda_j, j = 1,...,L$ and not for particular frequencies separately. Properties of ELSP will be discussed in Section 4.

3. Estimating spectra

A stationary stochastic process exhibits red noise spectrum when its spectral density function satisfies $f(\lambda) \geq f(\omega), \lambda < \omega$ for every $\lambda, \omega \in (0, \pi]$. Estimation of a spectral density corresponding to this definition is now based on an LS technique. The solution of the LS problem

$$\min\left\{\sum_{j=1}^{L}\left(I_*(\lambda_j) - \hat{f}(\lambda_j)\right)^2\right\}, \quad \hat{f}(\lambda_j) \geq \hat{f}(\lambda_k),\ j < k$$

can be obtained using a procedure called isotonic regression (IR). Namely,

$$\hat{f}(\lambda_j) = \min_{i \leq j} \max_{j \leq k} \frac{I_*(\lambda_i) + ... + I_*(\lambda_k)}{k - i + 1}$$

for $j = 1,...,L$, and $\hat{f}(\lambda)$ is left-continuous otherwise (*Zhao* and *Woodroofe*, 2012), where I_* can be either I, I_{LS}, or I_{ELS}. Note that $\hat{f}(\lambda)$ is monotone decreasing and is stepwise constant over certain frequency ranges.

Behavior of periodogram elements at frequencies close to possibly existing discrete frequencies (frequencies contributing to discrete spectra) substantially differs from the behavior of the majority of periodogram elements. Thus, periodogram elements at these frequencies should be taken as outliers, and an IR, robust against outliers has to be found. *Álvarez* and *Yohai* (2011) proposed a robust IR technique that can thus be used as a method to estimate the spectral density of the red background noise without any assumption on its analytical form. A possibility to find essential frequencies is to detect significant deviations of the periodogram from this background noise utilizing the statistical properties of robust IR (*Matyasovszky*, 2013b). Another way is to give up the background noise concept and estimate the spectral density non-robustly and without monotonicity constraint. The nearly-isotonic regression (NIR) introduced by *Tibshirani et al.* (2011) permits the possibility of deviations from

monotonicity when necessary. The necessity of monotonicity violations is controlled via a parameter β that is estimated within the procedure. The task is to find $\hat{f}(\lambda_1),...,\hat{f}(\lambda_L)$ minimizing the quantity

$$\left\{\sum_{j=1}^{L}\left(I_*(\lambda_j)-\hat{f}(\lambda_j)\right)^2+\beta\sum_{j=1}^{L-1}\left(\hat{f}(\lambda_j)-\hat{f}(\lambda_{j+1})\right)_+\right\},$$

where $(u)_+=-u$ when u is negative, and zero otherwise. When $\beta=0$, the solution is $\hat{f}(\lambda_j)=I_*(\lambda_j)$, and letting $\beta\to\infty$, we obtain the isotonic regression. The optimal value of β, which can be estimated (*Tibshirani et al.,* 2011), gives a trade off between monotonicity and goodness-of-fit. The resulting $\hat{f}(\lambda)$ can thus correspond to red noise or colored noise with certain local peaks according to the shape of $f(\lambda)$.

4. Properties of periodograms

4.1. Lomb-Scargle periodogram (LSP)

Now we examine the probability distribution and the expected value of LSP. For this reason, take first a two-dimensional normal random vector $\underline{\xi}$ with expected value zero and covariance matrix $\underline{\underline{P}}$. Let $\underline{\underline{Q}}$ be a positive definite matrix of size (2x2), and take the random variable $q=\underline{\xi}^T\underline{\underline{Q}}\underline{\xi}$. This quadratic form can be rewritten as a linear combination of two independent chi-squared random variables with one degree of freedom, and the coefficients in this linear combination are the eigenvalues of $\underline{\underline{Q}}\underline{\underline{P}}$. If these coefficients are identical, q is distributed exponentially, but when difference between the coefficients is not too large, the probability distribution is also closely exponential (*Yuan* and *Bentler,* 2010).

It is known from the LS theory that covariance matrix of $\underline{c}$ is $\underline{\underline{D}}^{-1}\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}\underline{\underline{D}}^{-1}$ (*Nielsen*, 2011), where *(i,j)*th element of $\underline{\underline{B}}$ is the covariance between $x(t_i)$ and $x(t_j)$. As $\underline{c}$ is distributed asymptotically normally, we recognize that the quadratic form Eq. (3) asymptotically corresponds to the above mentioned random variable q with $\underline{\underline{P}}=\underline{\underline{D}}^{-1}\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}\underline{\underline{D}}^{-1}$ and $\underline{\underline{Q}}=1/(2\pi)\underline{\underline{D}}$. When $\underline{x}$ comes from a white noise process, the eigenvalues of $\underline{\underline{Q}}\underline{\underline{P}}=1/(2\pi)\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}\underline{\underline{D}}^{-1}$ are identical, and this is why LSP follows an exponential distribution under a white noise process. For other processes, the eigenvalues depend on length of the data set, spacing, autocovariances of the underlying process, and frequency. However, except for some degenerate cases, the difference between these eigenvalues is not too

large. Therefore, our final conclusion is that probability distribution of LSP is closely exponential even for processes different from white noise.

It can be shown (*Mathai* and *Provost*, 1992) that the expected value of *q* is the trace of $\underline{\underline{Q}}\underline{\underline{P}}$, and therefore, the expected value of Eq. (3) is the trace of

$$1/(2\pi)\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}\underline{\underline{D}}^{-1}.$$

Remember that trace of a quadratic matrix is the sum of diagonal elements of this matrix, which is identical with the sum of its eigenvalues. Utilizing basic trigonometric identities we obtain

$$E\left[I_{LS}(\lambda_j)\right]=trace\left(1/(2\pi)\sum_{k=-(N-1)}^{N-1}\underline{\underline{D}}_k\underline{\underline{D}}^{-1}B(k)\cos(\lambda_j k)\right), \tag{6}$$

where $B(k)$ is the autocovariance function of the process that generates the time series, $\underline{\underline{D}}_k=\underline{\underline{S}}_k^T\underline{\underline{S}}_k$ with *(i,j)*th element of $\underline{\underline{S}}_k$ $s_{k,ij}=\alpha_i\alpha_{i+|k|}(\cos(\lambda_j i),\sin(\lambda_j i))^T$, $i=1,...,N-|k|$, and $\alpha_i, i=1,..,N$ is an indicator series that equals with zero when no data is available at time *i*, while equals with one when data is available at time *i*. Evidently, Eq. (6) for evenly sampled data has the well-known form

$$E\left[I(\lambda_j)\right]=B(0)/\pi+2/\pi\sum_{k=1}^{n-1}(1-k/n)B(k)\cos(\lambda_j k), \text{ or} \tag{7}$$

$$E\left[I(\lambda_j)\right]=\int_{-\pi}^{\pi}g(\omega)K_n(\omega-\lambda_j)d\omega$$

with $g(\omega)=f(\omega),\omega\geq 0$, $g(\omega)=f(-\omega),\omega<0$, and

$$K_n(\omega)=1/(2\pi)\sum_{k=-(n-1)}^{n-1}u(k)\cos(\omega k),$$

where $u(k)=1-|k|/n$. Hence, $K_n(\omega)$ is the Fejér-kernel. In the frequency domain representation of

$$E\left[I_{LS}(\lambda_j)\right]=\int_{-\pi}^{\pi}g(\omega)K_n(\omega-\lambda_j)d\omega,$$

$u(k)$ corresponding to Eq. (6) is $u(k)=trace(\underline{\underline{D}}_k\underline{\underline{D}}^{-1})/2$. Note that LSP is an unbiased estimator for white noise processes (see Eq. (7)).

4.2. Entire least squares periodogram (ELSP)

As $\underline{\underline{D}}^{-1}\underline{w}=\underline{c}$, therefore $c_i w_i=(\underline{\underline{D}}^{-1}\underline{w})_i w_i$. Additionally, the covariance matrix of $\underline{w}$ is $\underline{\underline{Z}}^T\underline{\underline{B}}\underline{\underline{Z}}$. Utilizing these facts and basic trigonometric identities we obtain after *Mathai* and *Provost* (1992) that

$$E\left[I_{ELS}(\lambda_j)\right]=trace\left(1/(2\pi)\sum_{k=-(N-1)}^{N-1}\left\{\underline{\underline{D}}^{-1}\underline{\underline{D}}_k\right\}_j B(k)\cos(\lambda_j k)\right), \tag{8}$$

where $\left\{\underline{\underline{U}}\right\}_j$ denotes a matrix of size (2×2) consisting of *(j,j)*th, *(j,j+L)*th, *(j+L,j+L-1)*th *(j+L,j+L)*th elements of $\underline{\underline{U}}$. Here,

$$s_{k,ij}=\left\{\begin{matrix}\alpha_i\alpha_{i+|k|}\cos(\lambda_j i), j=1,...,L\\ \alpha_i\alpha_{i+|k|}\sin(\lambda_j i), j=L+1,...,2L\end{matrix}\right\}, i=1,...N-|k|.$$

Following *Yuan* and *Bentler* (2010) we observe that ELSP is also a linear combination of two independent chi-squared random variables with one degree of freedom, and if difference between coefficients in this linear combination is not very large, the probability distribution of ELSP is closely exponential. Eq. (8) for evenly sampled data also has the well-known form Eq. (7). In the frequency domain representation of

$$E\left[I_{ELS}(\lambda_j)\right]=\int_{-\pi}^{\pi}g(\omega)K_n(\omega-\lambda_j)d\omega,$$

$u(k)$ corresponding to Eq. (8) is $u(k)=trace(\left\{\underline{\underline{D}}^{-1}\underline{\underline{D}}_k\right\}_j)/2$. Note that for both the LSP and ELSP, $K_n(\omega)$ depends not only on *n* but also on data spacing and λ_j. However, ELSP is an unbiased estimator for white noise processes (see Eq. (8)).

5. Examples

5.1. Hallet Lake

The Hallet Lake temperature proxy record (*Mc Kay et al.*, 2008) from south-central Alaska available for a period of AD 492-2005 is based on biogenic silica preserved in the lacustrina sediments. Its data spacing varies from 1 to 35 years with an average spacing of 10.15 years. Data (*n*=150) are standardized to have zero mean and unit variance.

Fig. 1 shows no substantial differences between LSP and ELSP. However, many small differences count for much, as NIR performed with LSP (NIR-LSP) provides a red noise spectral density, while NIR with ELSP (NIR-ELSP) delivers a colored noise density. In this latter case, a moderate but wide peak of the spectral density reinforces the importance of multi-decadal oscillations shown in other Alaskan proxy records (e.g., *Wilson et al*., 2007).

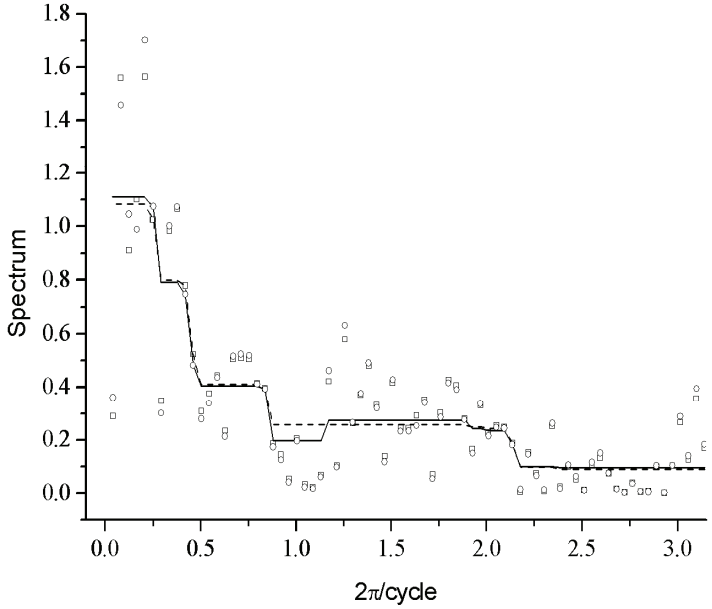

Fig. 1. ELSP (circle), LSP (square), NIR-ELSP spectral density (solid line), and NIR-LSP spectral density (dashed line) of Hallet Lake temperature proxy record AD 492-2005.

5.2. Vostok

Vostok ice core deuterium content data are available for the last 422,766 years (*Petit et al*., 1999). Data spacing varies from 20 to 664 years with a mean spacing of 127.8 years. Data (n=3303) are standardized to have zero mean and unit variance. Not surprisingly, the highest peak of both the NIR-LSP and NIR-ELSP spectra appears at 105,500 years corresponding to the Earth eccentricity cycle (*Fig. 2*). The cycle related to obliquity can be seen at somewhat lower frequency (closer to the 41,000-year astronomical cycle) for NIR-ELSP than for NIR-LSP. More importantly, dominance of the eccentricity cycle is much clearer from NIR-ELSP, since NIR-ELSP peak at this cycle is substantially higher than the NIR-LSP peak. The ratio of the magnitude of peak at eccentricity to peak at obliquity is 1.63 for NIR-LSP, while it is 4.88 for NIR-ELSP. Peaks in an interval of 21,000-28,000 years corresponding to the precession are considerably more modest for NIR-ELSP than for NIR-LSP. The third largest peak of NIR-LSP is around a 60,000-year cycle which cannot be explained by a direct astronomical forcing but is probably a

side-effect of the aforementioned cycles (e.g., *Rial* and *Anaclerio*, 2000). Note, however, that this period is essentially missing in NIR-ELSP. *Fig. 3* shows peaks at high frequencies, too. Cycles around 270 years are substantially stronger with NIR-ELSP than with NIR-LSP. Similar fact can be mentioned for cycles corresponding to 500-525 years (almost double of the 270-year period). These cycles are clearly related to solar cycles listed in *Schove* (1983).

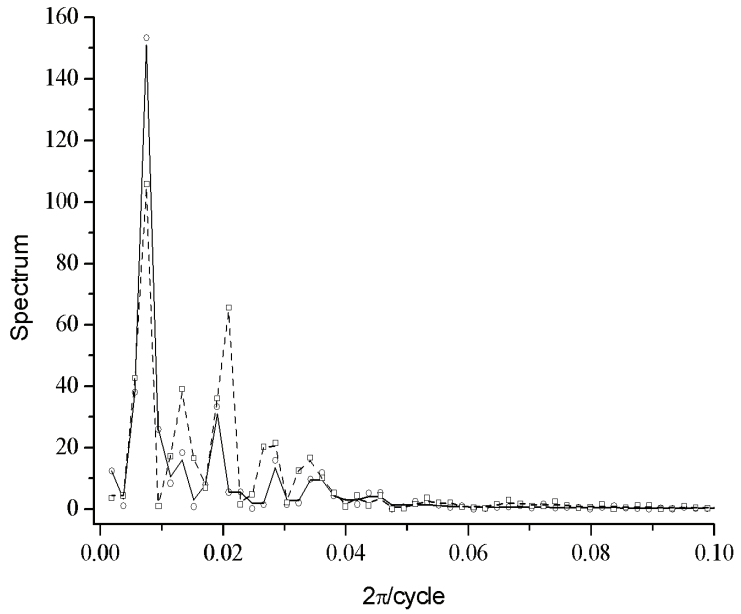

Fig. 2. ELSP (circle), LSP (square), NIR-ELSP spectral density (solid line), and NIR-LSP spectral density (dashed line) for Vostok ice core deuterium content data for the last 422,766 years at low frequencies.

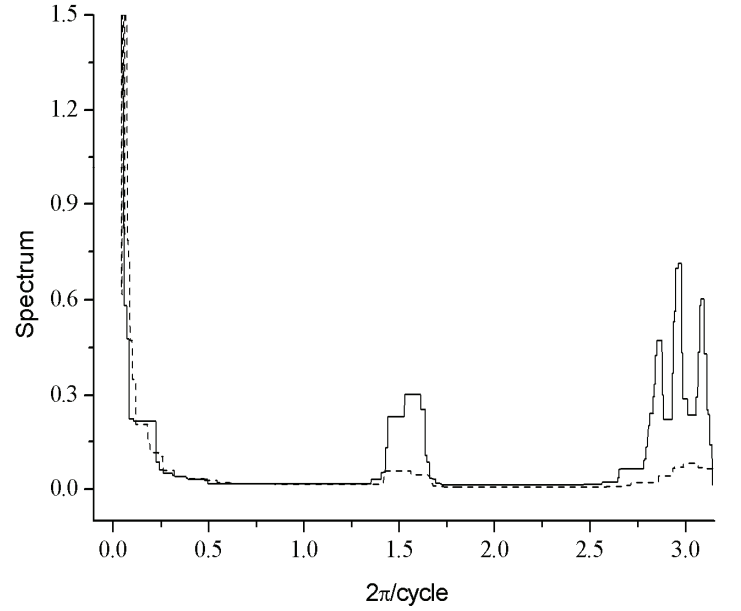

Fig. 3. NIR-ELSP spectral density (solid line) and NIR-LSP spectral density (dashed line) for Vostok ice core deuterium content data for the last 422,766 years at high frequencies.

5.3. GISP2

Oxygen-isotope data from GISP2 ice core from Greenland (*Groots* and *Stuvier*, 1997) are unevenly spaced in time, varying from 68 to 257 years with a mean spacing of Δ=125.8 years in the period between 15,000 and 60,000 BP. Data (n=358) are standardized to have zero mean and unit variance. *Schulz* and *Mudelsee* (2002) analyzed this record in order to detect a spectral peak at a 1470-year cycle corresponding to the spacing of the well-known Dansgaard-Oeschger events. Although the difference between ELSP and LSP does not seem substantial, NIR-LSP does not detect any spectral peak but does provide a red noise spectral density. In contrast, a peak at the above mentioned 1470-year cycle clearly appears when using the NIR method with ELSP (*Fig. 4*).

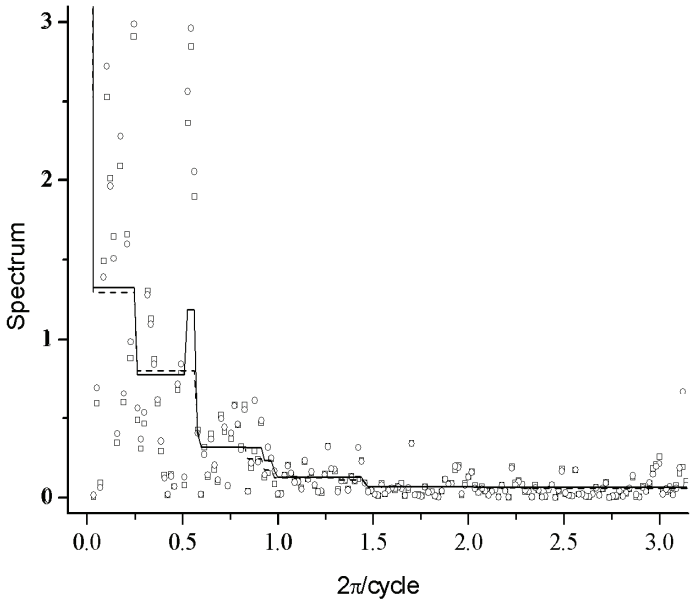

Fig. 4. ELSP (circle), LSP (square), NIR-ELSP spectral density (solid line), and NIR-LSP spectral density (dashed line) for GISP2 oxygen-isotope data for period between 15,000 and 60,000 BP. A big spectral peak around 40,000 years is not visible due to the vertical axis scale.

6. Discussion and conclusions

In order to demonstrate the ability of NIR and the drawback of AR(1) fitting, a simple example is taken by a stochastic process

$$Y_t = 0.5\cos(0.05\pi t) + X_t, \tag{9}$$

where the red background noise comes from a first order moving average (MA(1)) process $X_t = e_t - 0.5e_{t-1}$, and e_t is a white noise Gaussian process with $\sigma_e = 1$. Note that the variance corresponding to the discrete cycle $0.5\cos(0.05\pi t)$ is only 10% of the background noise variance. A time series of Y_t with n=400 is simulated and the spectrum is estimated with both the AR(1) fitting and the NIR method. The procedure is repeated 1,000 times. *Fig. 5* shows that the mean of the 1,000 NIR spectra exhibits a very sharp peak recognizing the discrete frequency at 0.05π. At the rest of frequencies, NIR spectrum reproduces well the background noise spectrum as compared to the background noise obtained with AR(1) fitting. Note that AR(1) spectral density is around two times higher than the true background spectral density at low frequencies causing difficulties in detecting the discrete frequency with traditional techniques. When omitting the discrete frequency from Eq. (9), the AR(1) spectral density is almost the same as in the previous case involving discrete frequency, while the NIR spectral density essentially coincides with the true background noise spectral density.

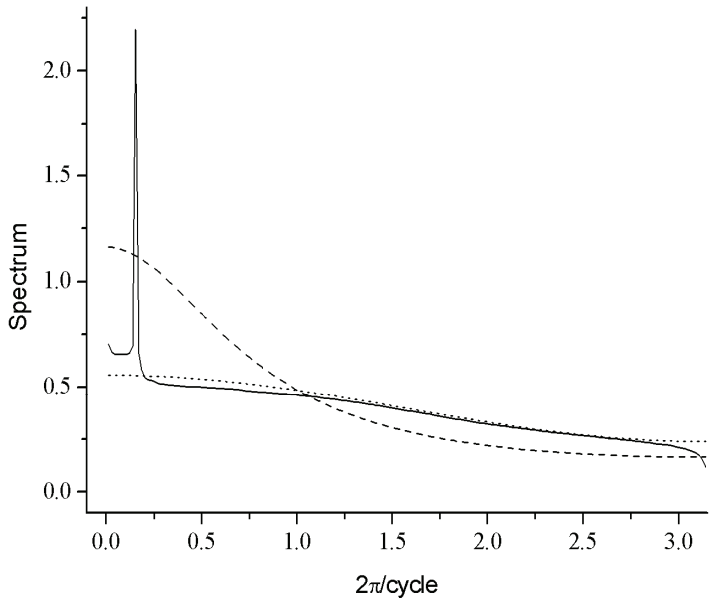

Fig. 5. Mean of NIR-ELSP spectral densities (solid line) and AR(1) spectral densities (dashed line) obtained from 1,000 simulated time series of length n=400 according to Eq. (9), and AR(1) spectral density (dotted line) of the background noise in Eq. (9).

Comparison of properties of LSP and ELSP is illustrated with the third data set in Section 5.3. As it was mentioned earlier, both the LSP and ELSP can be written as

$$I_*(\lambda_j) = h_1\xi_1 + h_2\xi_2, \tag{10}$$

where $I_*(\lambda_j)$ is either LSP or ELSP, and ξ_1 and ξ_2 are independent chi-squared random variables with one degree of freedom. For simplicity, dependence on the frequency of h_1 and h_2 is not indicated. If coefficients h_1, h_2 are identical, $I_*(\lambda_j)$ is distributed exponentially, but when the ratio $r = \max\{h_1, h_2\}/\min\{h_1, h_2\}$ is not too large, the probability distribution is closely exponential. Approximating the true probability distribution with an exponential distribution is highly accurate for ratios from *r*=1 to at least *r*=2-3 (*Yuan* and *Bentler*, 2010). Calculation of *r* (*Yuan* and *Bentler*, 2010) requires the autocorrelations of the underlying process. These are here substituted by autocorrelations corresponding to the AR(1) model fitted to data with a procedure described in *Schulz* and *Mudelsee* (2002). *Fig. 6* shows this ratio against frequencies. It is obvious that both periodograms can be taken as they are distributed exponentially. At very high frequencies, the distribution tends to deviate from the exponential one, but with smaller degree for ELSP than for LSP. Note that exponential approximation of the distribution of ELSP is accurate even at highest frequencies.

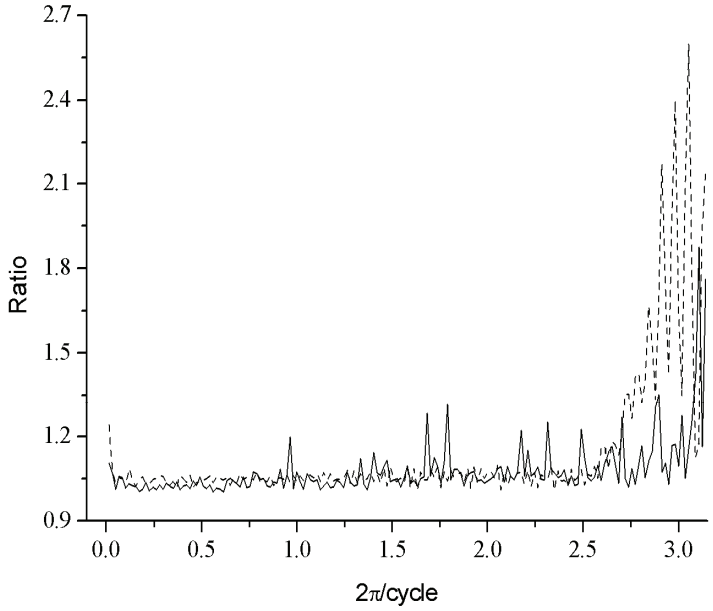

Fig. 6. Ratio $\max\{h_1, h_2\}/\min\{h_1, h_2\}$ against frequencies for ELSP (solid line) and LSP (dotted line), where coefficients h_1, h_2 are defined in Eq. (10).

Another important property of both the LSP and ELSP is that these periodogram elements at different frequencies are correlated for unevenly spaced data. *Fig. 7* shows the correlation between $I_{LS}(\lambda_j)$ and $I_{LS}(\lambda_i - \lambda_j)$ against the frequency λ_j and frequency shift $\lambda_i - \lambda_j$ for $i, j = 1,...,L$. It is apparent

that correlations are essentially zero at any frequencies and any frequency shifts except for the highest frequencies. At highest frequencies, the correlations are not negligible but only within narrow frequency shift intervals. The overall picture for ELSP (the corresponding figure is not shown) is the same but with slightly lower correlations. For instance, the largest correlation under every combination of λ_j and $\lambda_i - \lambda_j$ is 0.47 for LSP, while it is 0.43 for ELSP. These results are consistent with findings obtained for exponential approximation to the distribution of ELSP and LSP. The distribution of ELSP and LSP tends to deviate from the exponential distribution, when correlation between sinusoid and cosinusoid parts of the periodogram at a given frequency increases. Somewhat similar phenomenon can be found in *Vio et al*. (2010) but only for LSP and for time series simulated from white noise processes. It is to be mentioned that calculating the correlation between two periodogram elements utilizes that $\underline{c}$ has an asymptotic multivariate normal distribution. Hence, the mentioned correlation consists of fourth-order central moments of $\underline{c}$. These moments, due to the normality of $\underline{c}$, can be expressed via second order central moments, ie., via the covariance matrix of $\underline{c}$. Finally, this covariance matrix can be approximated using the autocorrelations corresponding to the AR(1) model fitted to data.

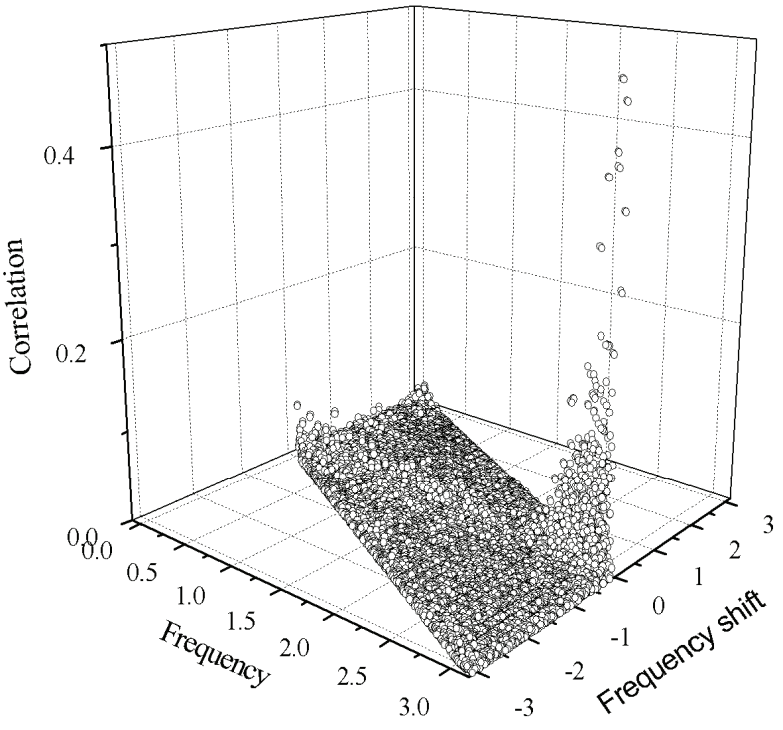

Fig. 7. Correlation between LSP elements at different frequencies and different frequency shifts.

Similar calculations shows that main conclusions discussed in the last two paragraphs are applicable for data sets of Sections 5.1 and 5.2.

Originally, LSP has been developed for time series generated by stochastic processes consisting of a certain number of periodic components plus a white noise process. Later, it has been using to estimate not only discrete spectra but spectral densities, too (e.g., *Schulz* and *Mudelsee*, 2002). Recognizing that LSP at a given frequency can be highly affected by other frequencies, *Stoica et al.* (2009) introduced a weighted least square fit at every separate frequency, where the weights are related to other LSP elements. The procedure thus necessitates an iterative technique requiring bigger computational effort than ELSP. More importantly, the weights are chosen with an approximation that holds accurately only for evenly spaced data. Such a simplification appears also in *Nygrén* and *Ulich* (2010). Here, after performing the entire least squares technique, the matrix $\underline{\underline{D}}$ is taken diagonal with elements $n/2$. Hence, their periodogram $n/(4\pi)(c_j^2 + c_{j+L}^2)$ provides a biased estimator even for white noise processes and does not integrate to $\hat{\sigma}^2$ (except for evenly spacing data) in contrast to our Eq. (5).

The matrix $\underline{\underline{D}}$ is generally close to being singular for large values of n, and solving Eq. (2) for $\underline{c}$ to form ELSP is not easy. Our experience is that traditional techniques such as Gauss-Seidel, successive over-relaxation, or conjugate gradient methods might be unsuccessful. Therefore, Eq. (2) was solved with a Monte Carlo technique (e.g., *Liu*, 2001), as this procedure is carried out numerically with a totally different scheme than the previous techniques.

References

Álvarez, E.E., and *Yohai, V.J.,* 2011: M-estimators for Isotonic Regression. arXiv: 1105.5065v1stat.ME.

Broersen, P.M.T., 2006: Autoregressive Order Selection for Irregularly Sampled Data. *IEEE Trans. Instrum. Meas. 54*, 1004–1009.

Groots, P.M. and *Stuvier, M.*, 1997: Oxygen 18/16 variability in Greenland snow and ice with 10^{-3}–10^{5}-year time resolution. J. *Geophys. Res. 102 (C12)*, 26455–26470.

Liu, J.S., 2001: Monte Carlo strategies in Scientific Computing. Springer, New York.

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

Mathai, A.M. and *Provost, S.B.,* 1992: Quadratic Forms in Random Variables. Taylor & Francis.

Matyasovszky, I., 2013a: Spectral analysis of unevenly spaced climatological time series. *Theor. Appl. Climatol 111*, 371–378.

Matyasovszky, I., 2013b: Estimating red noise spectra of climatological time series. *Időjárás 117*, 187–200.

Mc Kay, N.P., Kaufman, D.S., and *Michelutti, N.*, 2008: Biogenic-silica concentration as a high-resolution, quantitative temperature proxy at Hallet Lake, south-central Alaska. *Geophys. Res. Lett. 35*, LO5709.

Nielsen, A.A. ,2011: Least Squares Adjusments: Linear and Nonlinear Weighted Regression Analysis. Lecture Note, Lyngby, Denmark: IMM, DTU, http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=2804.

Nygrén, T., and *Ulich, Th.*, 2010: Calculation of signal spectrum by means of stochastic inversion. *Ann Geophys 28*, 1409–1418.

Petit, J.R., Jouzel, J., Raynaud, D., Barkov, N.I., Barnola, J.M., Basile, I., Bender, M., Chappellaz, J., Davis, J., Delaygue, G., Delmotte, M., Kotlyakov, V.M., Legrand, M., Lipenkov, V., Lorius, C., Pépin, L., Ritz, C., Saltzman, E., and *Stievenard M.,* 1999: Climate and Atmospheric History of the Past 420,000 years from the Vostok Ice Core, Antarctica. Nature 399, 429-436.

Rial, J.A., and *Anaclerio, C.A.*, 2000: Understanding nonlinear responses of the climate system to orbital forcing. *Quat. Sci. Rev. 19*: 1709–1722.

Scargle, J.D., 1982: Studies in astronomical time series analysis II: statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J. 261*, 835–853.

Schove, D.J. (Ed.), 1983: Sunspot cycles. *Benchmark Papers in Geology 68*, Hutchinson.Ross Publ.Co.

Schulz, M., and *Mudelsee, M.*, 2002: REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatological time series. *Comput. Geosci. 28*, 421–426.

Stoica, P., Li, J., and *He, H.*, 2009: Spectral Analysis of Nonuniformly Sampled Data: A New approach Versus the Periodogram. *IEEE Trans. Signal Process. 57*, 1415–1425.

Tibshirani, R.J., Hoefling, H., and *Tibshirani, R.*, 2011: Nearly-Isotonic Regression. *Technometrics 53*, 54–61.

Vio, R., Adreani, P., and *Biggs, A.*, 2010: Unevely-sampled signals: a general formalism of the Lomb-Scargle periodogram. *Astron. Astrophys. 519*, A86, 12.

Wilson. R., Wiles, G., D'Arrigo, R., and *Zweck, C.*, 2007: Cycles and shifts: 1,300 years of multi-decadal temperature variability in the Gulf of Alaska. *Clim. Dyn. 28*, 425–440.

Yuan, K-H., and *Bentler, P.M.,* 2010: Two simple approximations to the distributions of quadratic forms. *Br J Math Stat Psychol* 63, 273–291.

Zhao, O., and *Woodroofe, M.*, 2012: Estimating a monotone trend. *Stat Sinica 22*, 359–378.