Ann. Geophysicae 16, 356–358 (1998) © EGS – Springer-Verlag 1998

Annales Geophysicae

Letter to the Editor

Phase estimation with the Lomb-Scargle periodogram method

K. Hocke

Communications Research Laboratory, Tokyo, Japan

Received: 25 August 1997 / Accepted: 24 September 1997

Abstract. The Lomb-Scargle periodogram method was developed for the spectral analysis of unevenly sampled data. To date program routines exist for the calculation of the so-called Lomb normalized periodogram (normalized spectral power as function of frequency). In geophysical applications the phase and amplitude spectra are also of interest. Commonly, phases and amplitudes of spectral components of unevenly sampled data are determined by least squares fit procedures which are independently performed after calculation of the Lomb normalized periodogram. In this paper, it is shown how the Lomb-Scargle periodogram method can be used for estimation of the phase and amplitude spectra. The extended program version of the Lomb-Scargle periodogram provides the same spectral information as the Fourier transformation in the case of evenly sampled data.

1 Introduction

Before the initiation of the notion "Lomb-Scargle periodogram", named after N.R. Lomb and J.D. Scargle, Lomb called it just "least squares spectrum". In his fundamental work (Lomb, 1976), Lomb derived the statistics and behaviour of the least-squares frequency analysis of unequally spaced data and explained the method as following: "The sole justification for the use of periodogram analysis is that, as will be shown later, it provides a reasonably good approximation to the spectrum obtained by fitting sine waves by least-squares to the data and plotting the reduction in the sum of the residuals against frequency. This least squares (or LS) spectrum (Barning, 1963) provides the best measure of the power contributed by the different frequencies to the overall variance of the data. It reduces to the Fourier power spectrum in the limit of equal spacing."

This article was processed by the author using the LATEX style file *cljour2* from Springer-Verlag.

The most contributions to the development of the Lomb-Scargle periodogram method were made by astronomers (e.g., Scargle, 1982; Horne and Baliunas, 1986; Press and Rybicki, 1989). It is likely that the detection of a periodic signal hidden in noise is a more important goal in astronomical data analysis than the estimation of the phase and amplitude spectra of time series, which is for example of great interest in geophysical data analysis of gravity waves. This is probably the reason why the existing program routines only calculate the Lomb normalized power spectrum. In recent years the Lomb-Scargle periodogram method was often applied in atmospheric data analysis (e.g., Hall and Hoppe, 1997; Salah *et al.*, 1997; Forbes and Zhang, 1997).

In the present work, the program 'period.f' (Press *et al.*, 1992), which derives the Lomb normalized periodogram, is extended for the calculation of the phase and amplitude spectra. Thus, the full information of the Lomb-Scargle periodogram method is used, so that additional least squares fit procedures are not necessary.

2 Phase and amplitude spectra

Sine waves of the form

$$y_f(t_i) = a\cos\omega(t_i - t_{ave} - \tau) + b\sin\omega(t_i - t_{ave} - \tau) \quad (1)$$

are fitted to the data y_i (with zero mean) at times t_i, $i = 1, ..., n$. ω is the angular frequency and $t_{ave} = (t_1 + t_n)/2$. (The redundant constant t_{ave} was considered in order to be consistent with the already mentioned numerical recipes program 'period.f'.) The parameter τ is defined by (Scargle, 1982)

$$\tan(2\omega\tau) = \frac{\sum_{i=1}^{n} \sin 2\omega(t_i - t_{ave})}{\sum_{i=1}^{n} \cos 2\omega(t_i - t_{ave})} . \quad (2)$$

a and b are

$$a = \frac{\sqrt{\frac{2}{n}} \sum_{i=1}^{n} y_i \cos \omega (t_i - t_{ave} - \tau)}{\left(\sum_{i=1}^{n} \cos^2 \omega (t_i - t_{ave} - \tau) \right)^{1/2}} \tag{3}$$

$$b = \frac{\sqrt{\frac{2}{n}} \sum_{i=1}^{n} y_i \sin \omega (t_i - t_{ave} - \tau)}{\left(\sum_{i=1}^{n} \sin^2 \omega (t_i - t_{ave} - \tau) \right)^{1/2}} \tag{4}$$

The Lomb normalized periodogram (e.g., Press *et al.*, 1992) is

$$P_N(\omega) = \frac{1}{2\sigma^2} \frac{n}{2} (a^2 + b^2) \quad \text{with } \sigma^2 = \frac{1}{n-1} \sum_{i=1}^{n} y_i^2. \tag{5}$$

The amplitude spectra $A(\omega)$ is easily derived by

$$A(\omega) = \sqrt{\frac{2}{n} 2\sigma^2 P_N(\omega)} \tag{6}$$

(or $A(\omega) = \sqrt{a^2 + b^2}$)

($P_N(\omega)$ is the same power spectral density as in the program 'period.f'. If a windowing function was applied to the time series, $P_N(\omega)$ must be multiplied with a further factor, e.g. factor 4 for the Hanning windowing function.)

According to the addition theorems the fitted sine wave $y_f(t_i)$ from Eq. (1) can be also expressed by

$$y_f(t_i) = A(\omega) \cos [\omega(t_i - t_{ave} - \tau) + \phi] \tag{7}$$

$$\phi = -\text{atan}(b, a). \tag{8}$$

The phase spectrum shall be defined by the cosine argument of $y_f(t_i)$ at the time $t_i = 0$

$$\varphi(\omega) = \omega(-t_{ave} - \tau) + \phi \tag{9}$$

$$= -\omega t_{ave} - \omega \tau + \phi. \tag{10}$$

In the program 'period.f' the periodogram is calculated at the frequencies

$$\omega_j = \frac{2\pi}{(t_n - t_1)\, ofac} j, \qquad j = 1, 2, 3, ... \tag{11}$$

$ofac$ is the so-called oversampling factor and is usually chosen as integer greater than 4 (Press *et al.*, 1992). The discrete phase spectrum is then

$$\varphi_j = -\omega_j t_{ave} - \omega_j \tau - \text{atan}(b, a), \quad j = 1, 2, 3... \tag{12}$$

($\omega_j \tau = \underline{wtau}$ and $\text{atan}(b, a) = \text{atan}(sumsy/\sqrt{sums}, sumcy/\sqrt{sumc})$ in the special notion of program 'period.f'.)

Finally, the phase and amplitude spectra are shown for one example, related to atmospheric tidal analysis. The time interval of the artificial data set is 4 days, and the sampling time is 20 min. The time series in Fig. 1a) is the superposition of 4 cosine waves with frequencies (0.55, 1.1, 2.2, 3.3) cycles/day, amplitudes (6, 7, 6, 4), and phases (315, 90, 180, 252) degree. In addition, a normally-distributed noise with a standard deviation of 2 was superposed and a Hanning windowing function was used.

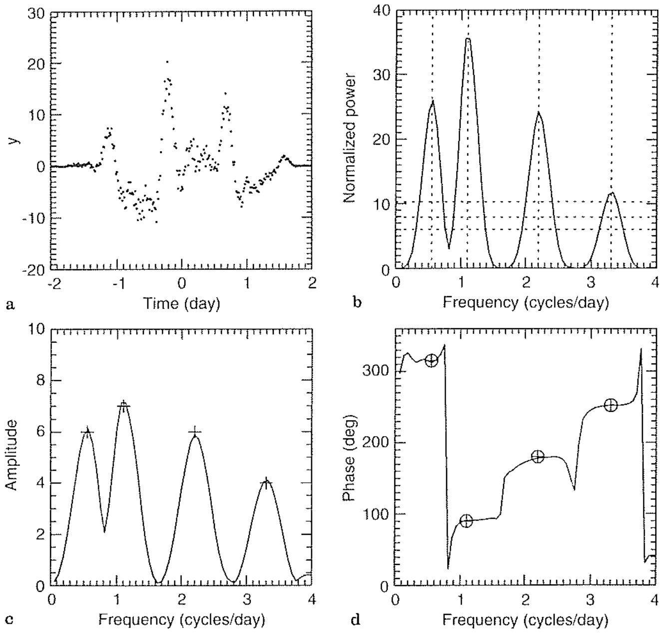

Fig. 1. Lomb-Scargle peridogram: a) time series y_i b) normalized power periodogram c) amplitude spectrum d) phase spectrum. The correct amplitude and phase values of the 4 superposed sine waves of y_i are denoted by the crosses in c) and d). Further explanations are in the text.

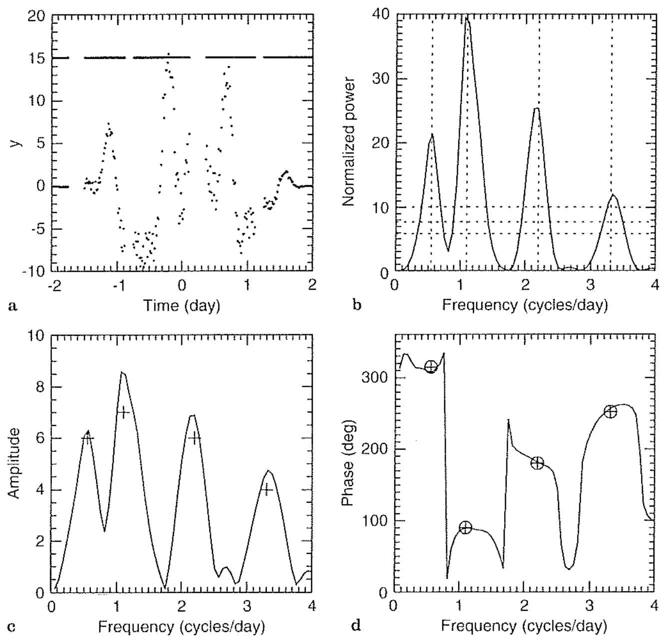

Fig. 2. Same as Fig. 1a-d) but this time with data gaps. The positions of the data gaps can be seen by means of the interrupted horizontal line in the top of a)

The frequencies were chosen at 'uneven' values in order to demonstrate the advantage of oversampling ($ofac = 4$). Contrary to the Fast Fourier transformation, the Lomb-Scargle periodogram can estimate the spectrum at these frequencies. The Lomb normalized power spectrum is shown in Fig. 1b). The vertical dotted lines denote the frequencies of the 4 waves, and the horizontal dotted lines indicate the significance levels of 99, 90, and 50 % from top to bottom. The estimated amplitude spectra is shown in Fig. 1c), and the phase spectrum in Fig. 1d), where the crosses indicate the correct values of amplitude and phase, respectively. In Fig. 2 the phase and amplitude estimation is depicted for the same time series, but now the time series has gaps as shown by the thick horizontal line at the top of Fig. 2a). The gaps are quite typical for radar data of the mesosphere: they have a distance of around 1 day and a length of around 3 - 5 hours. This kind of gaps can be caused, for example, by insufficient backscatter of radar signals in the mesosphere due to the lack of electrons and ions at these heights during nighttime. We see that the estimate of the amplitude is of lower quality than in Fig. 1 while the phase spectrum is still quite good. The quality of the phase and amplitude estimation by the Lomb-Scargle periodogram method depends on the time series, e.g. on the data gaps, but this is the same problem as in the case of conventional least squares fit methods.

3 Concluding remarks

The phase and amplitude spectra can be calculated by the Lomb-Scargle periodogram method and the use of Eqs. (12) and (6). The extension of the already existing program 'period.f' (Press *et al.*, 1992) is simple and requires only less additional computing time. It is now possible to estimate phase and amplitude spectra of unequally spaced data by calling a subroutine, such like 'period.f'. An application of the program can be the determination of the phase difference of a gravity wave observed at locally spaced stations in order to derive the wave propagation velocity. An IDL-version (Image Data Language) of the extended program 'period.f' is available by the author upon request. This IDL subroutine is sufficiently fast for a convenient use since many scalar operations in loops of the Fortran program 'period.f' were replaced by array operations in the IDL program.

Acknowledgement. The author thanks the Science and Technology Agency of Japan for a fellowship. M. Kunitake from CRL is thanked for discussions.

References

Barning, F.J.M., *Bull. Astron. Inst. Neth.*, **17**, 22-28, 1963.

Forbes, J.M. and X. Zhang, Quasi 2-day oscillation of the ionosphere: A statistical study, *J. Atmos. Solar-Terr. Phys.*, **59**, 1025-1034, 1997.

Hall, C.M. and U.-P. Hoppe, Characteristic vertical wavenumbers for the polar mesosphere, *Geophys. Res. Lett.*, **24**, 837-840, 1997.

Horne, J.H. and S.L. Baliunas, A prescription for periodic analysis of unevenly sampled time series, *Astrophys. J.*, **302**, 757-763, 1986.

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

Press, W.H. and G.B. Rybicki, Fast algorithm for spectral analysis of unevenly sampled data, *Astrophys. J.*, **338**, 277-280, 1989.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical recipes in Fortran*, Cambridge Univ. Press, 569-577, 1992.

Salah, J.E., W. Deng, and R.R. Clark, Observed dynamical coupling through tidal wave propagation in the mesosphere and lower thermosphere at midlatitudes, *J. Atmos. Solar-Terr. Phys.*, **59**, 641-654, 1997.

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