Annu. Rep. Nishi-Harima Astron. Obs. 3, 57 - 64 (1993)

Implementation and Performance Tests of Period Analysis Methods I. The Phase Dispersion Minimization method

Toshihito ISHIDA

Nishi-Harima Astronomical Observatory, Sayo-cho, Hyogo 679-53

Abstract

I have implemented a period analysis method, the phase dispersion minimization method, as a library of Nishi-Harima Astronomical Observatory. Usage and data formats of the library are presented in this note. Furthermore, preliminary tests for the following effects are carried out for the library, viz., (1) form of the wave function, (2) double-periodicity, (3) unevenly sampled data, and (4) introduction of noise. Results of above tests indicate that the performace of the library is fairly well.

Key words: Nishi-Harima software Library; Period Analysis; Phase Dispersion Minimization Method

1. Introduction

Many astronomical measurements indicate time-dependency. For examples, brightness of variable stars, magnitude of active galactic nuclei, and sunspot numbers are well known to change its values. Usual first step to extract physical insight from observed time-series is to search for a periodicity. This situation is the same as that in the laboratory experiments. However, there is a difference between the astronomical data and the data obtained by experiments, namely, data spacing.

In the usual laboratory experiments, we can get data whenever we want. Therefore, an equally spaced time series is available. In contrast, we can not obtain the astronomical data at sometimes due to many reasons. In visual light, data are not available on a bad weather day. Moreover, from a ground-based observatory, siderial data gaps are unavoidable. Many popular period analysis methods, as the fast Fourier transform and the maximum entropy method, are usually assuming that data are sampled at equally spaced times. Therefore, we need a period search method applicable to unevenlly sampled data. Many authors have devised and revised some methods for unevenly sampled data, some of which are comprehensively summerized in Fullerton(1986). As one of these methods, the phase dispersion minimization (PDM) method of Stellingwerf(1978) is implemented as a Nishi-Harima software Library(NHL).

In this note, usage of the library and preliminary performance tests are presented. In section 2, the principle of the PDM method is briefly indicated. Usages are shown in section 3. Results of preliminary performance tests on the wave function, double-periodicity, data sampling, and introduction of noise are presented in section 4. Future prospects are discussed in section 5.

2. Principle

Original idea of PDM is rather simple. With a completely wrong trial period, observed points scatter widely. But, at the right period, the scatter of observations around the mean light curve becomes small. Namely, in every short interval of phase, dispersion of data points is minimum at the right period. Simultaneously, dispersion of all means of phase interval takes its maximum value. This is a classical method searching for the periodicity.

The total variance of all data points are given by

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{N-1},$$

where $\bar{x} = \sum x_i/N$. For any subset $(x_i, i = 1, n_j)$, we can define the sample variance as the same as total variance. When we have M distinct subsets of data, we can get a set of sample variances $(s_j^2, j = 1, M)$. The overall variace of all the susets is given by

$$s^2 = \frac{\sum(n_j - 1)s_j^2}{\sum n_j - M}.$$

Using above two variances, we define the statistic as

$$\Theta = s^2/\sigma^2.$$

The right period is searched as the period, which minimizes Θ.

In PDM, the phase interval is divided into N_b bins, so the length of every bin is $1/N_b$. When we select a trial period, we can determine which bin a observation point belongs. Furthermore, we take N_c covers for each bins, which are shifted by $1/(N_b N_c)$ phase each other. This bin structure is denoted by a pair (N_b, N_c).

When we fix the bin structure $N_b, N_c)$, midpoints of $N_b N_c$ covers distribute in unit interval with uniform spacing. Thus, at the right period, we can get simultaneously the mean light curve with $N_b N_c$ points.

In above description, we make no assumption of the shape of the wave function. Therefore, PDM method is well suited to the variation with non-sinusoidal form.

3. Usages

3.1. Input data format

Data points are read by a list-directed read statement of Fortran. Input data ought to be a list of pairs of observed time and observed value, which are divided with some spaces. Namely, input data are as follows

$$\begin{array}{cc} time_1 & data_1 \\ time_2 & data_2 \\ time_3 & data_3 \\ \vdots & \vdots \end{array}$$

3.2. How to run the library

The PDM library is implemented to workstation of NHAO. At first, add `/use/local/bin` directory to your path. Type in `PDM`, then PDM library asks you to input a file name to be analysed.

After reading data, the library print out default trial period range and spacing between trial periods, and asks you to be continued or not. You can choice "n" here to change parameters manually. However, using default values is recommended, because of the following reason. The default values of the shortest trial period and period spacing are chosen to be inverse of a Pseudo-Nyquist frequency $\nu_{Ny} = (2 < \delta t >)^{-1}$, where $< \delta t >$ is the average of all the data sampling intervals. On the other hand, the default value of the longest trial period is chosen to be 20% of the total time span. These selection seems to be very near to the optimum one, so that when you change parameters to the very different values from default values you will get a meanigless result.

3.3. Output data

Results obtained by PDM library is printed out in a file named `Theta.data`. In this file, we can find a list of trial periods, statistics θ, and error estimates of θ. The errors of θ in the last column is estimated by the standard deviation of 29 bin structures, which are the same as those used in Saitou(1989).

4. Results of the Performance Tests

To confirm the performance of the implemented library, some preliminary test data files are generated. All of the test data have the same period near 3.2.

4.1. Wave function

(a) Sinusoidal function

As the simplest test, sinusoidal function with evenly spaced data is generated. The test data are shown in figure 1a. The result of PDM analysis is in figure 1b. Clearly, we can find a deep minimum at the right period 3.2. We can also see two subharmonics near 6.5 and 9.7, at which there are less deep minima than the right period.

(b) Saw-tooth function

In the section 2, I have poited out that the PDM method is useful to nonsinusoidal curves. So, we test a data with saw-tooth function. The test data, which are shown in figure 2a, linearly increase in phase $(0.0, 0.9)$ and linearly decrease in phase $(0.9, 1.0)$.

The result of PDM analysis are shown in figure 2b. Comparing with figure 1b, we can find some additional shallow minima. These minima seem to be included as harmonics in saw-tooth function. The main period near 3.2 is the most deep minimum. This result indicate clearly that the PDM method is useful for the saw-tooth function.

4.2. Double periodicity

There are some objects that oscillates with two or more period simultaneously. As a simplest case, double preiodic data are generated. The secondary period is set to be about 0.7 of the principal period. The amplitude of this period is set to be about 0.3 of that of the principal period. The test data are shown in figure 3a.

The result of PDM analysis are shown in figure 3b. Although the principal period has a clear minimum, the minimum for the secondary period is unclear. At phase 4.4, the first subharmonics of the secondary period has a rather clear minimum. These results seem to indicate that we need

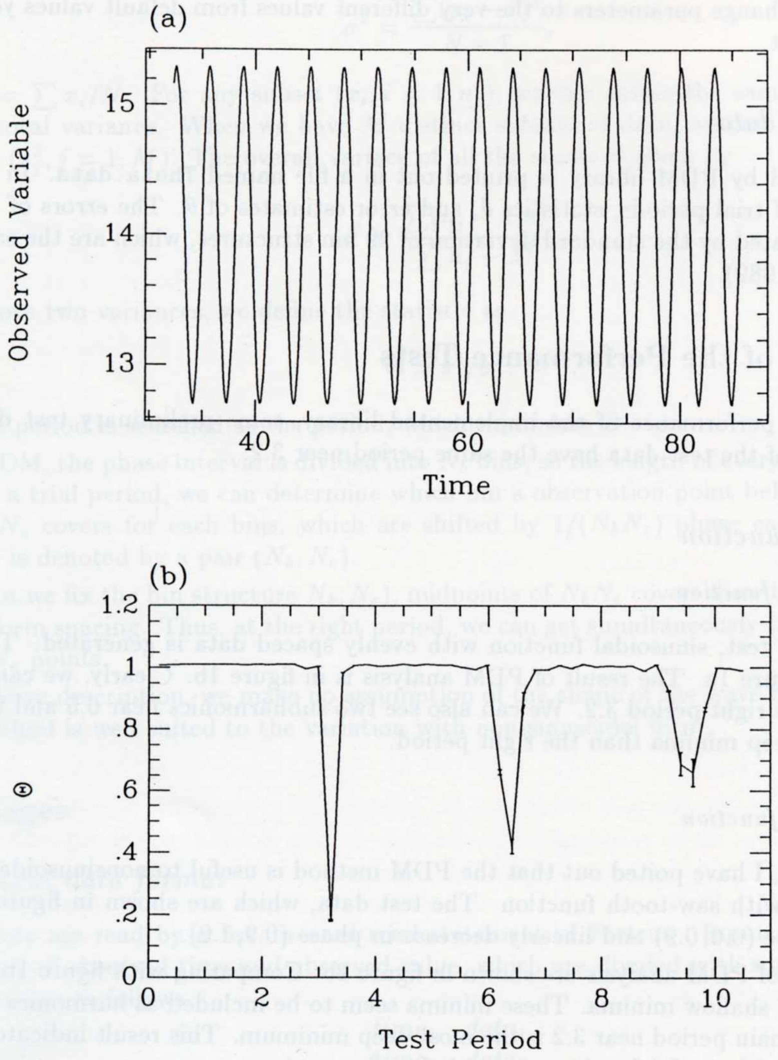

Fig. 1. A test result of PDM library. (a) A test data with sinusoidal function and (b) the result of PDM library are shown.

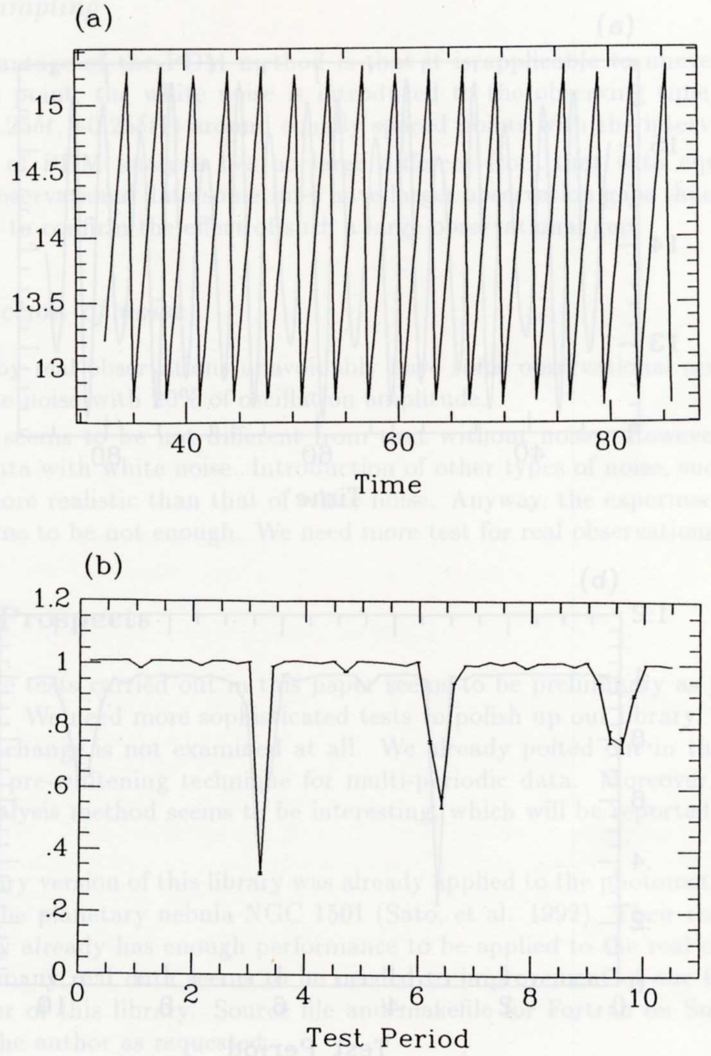

Fig. 2. A test result of PDM library. (a) A test data with saw-tooth function and (b) the result of PDM library are shown.

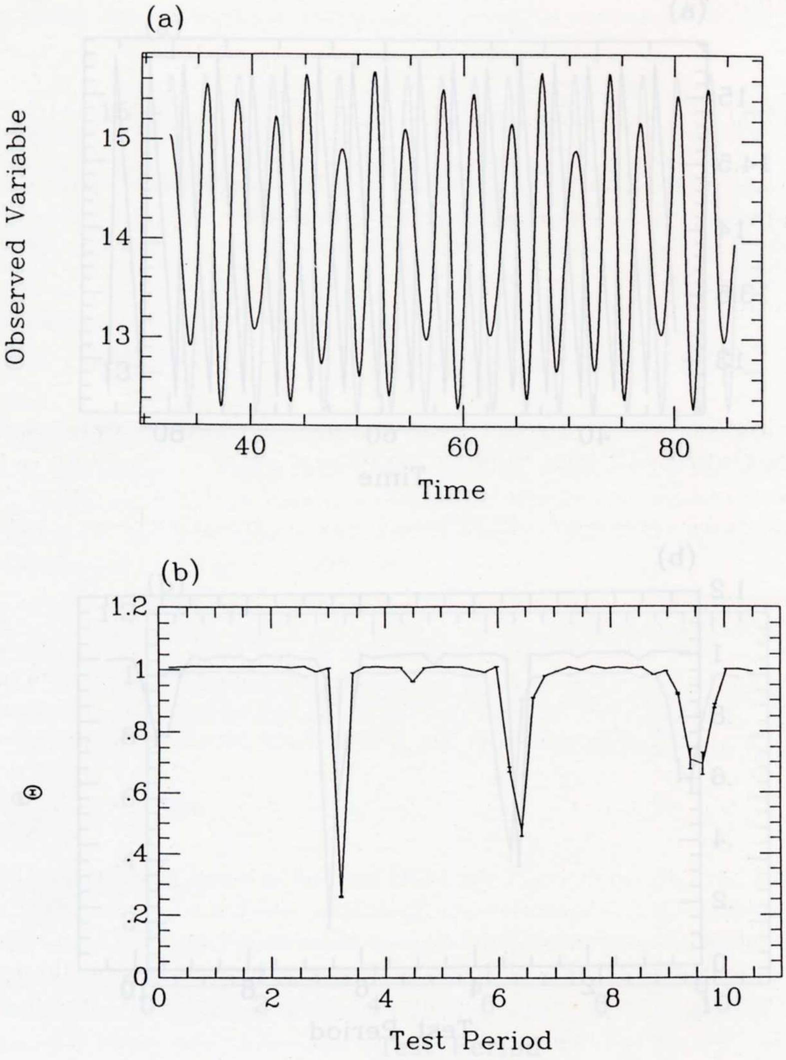

Fig. 3. A test result of PDM library. (a) A test data with double-periodicity and (b) the result of PDM library are shown.

a pre-whitenig technique to clearly distinguish such a low amplitude secondary periodicity. We need more tests with different amplitude to find the amplitude, at which the scondary periodicity is distingushed by the PDM library.

4.3. *Data sampling*

One of the advantage of the PDM method is that it is applicable to unevenly sampled data. To confirm this point, the white noise is introduced to the observing time. Data points are distributed $(-0.25\delta t, +0.25\delta t1)$ around equally spaced points with the interval δt.

The result of PDM analysis has no large different from that with equally spaced data. However, real observational data sometimes have larger observation gaps than tested above. We need more tests to confirm the effect of such a large observational gap.

4.4. *Introduction of noise*

Data obtained by real observations unavoidably have some observational noise. Therefore, we introduced white noise with 20% of oscillation amplitude.

The result seems to be not different from that without noise. However, we will not get observational data with white noise. Introduction of other types of noise, such as the Gaussian noise, will be more realistic than that of white noise. Anyway, the experiments with the effect of the noise seems to be not enough. We need more test for real observational data.

5. Future Prospects

The performance tests carried out in this paper seems to be preliminary as pointed out in the previous section. We need more sophisticated tests to polish up our library. For examples, the effect of period change is not examined at all. We already poited out in the previous section that we need a pre-whitening technique for multi-periodic data. Moreover, comparison with other period analysis method seems to be interesting, which will be reported in the next paper of this series.

A preliminary version of this library was already applied to the photometry of the pulsating central star of the planetary nebula NGC 1501 (Sato, et al. 1992). Their results indicate that the PDM library already has enough performance to be applied to the real observational data. Application to many real data seems to be needed to improvement of our library. Therefore, I welcome a user of this library. Source file and makefile for Fortran on Sun workstation are available from the author as requested.

I am very imdebted to Mr. M. Saitou for sending me his master thesis of Tôhoku University, which is very useful to understand and implement the PDM method.

References

Fullerton, A.W., 1986, in “The Study of Variable Stars Using Small Telescoples”, (ed. J.R. Percy), Cambridge University Press.

Saitou, M., 1989, “Studies of Pulsating Stars”(Text in Japanese), Master thesis of Astronomical Institute, Tôhoku University.

Sato, T., Okyudo, M., Tokimasa, N., Kuroda, T., and Ishida, T., 1992, Annu. Rep. Nishi-Harima Aston. Obs., **2**, 16.

Stellingwerf, 1978, *Astrophys. J.*, **224**, 953.