Bull. Astron. Belgrade № 155 (1997), 27 – 34

UDC 52–17
Original scientific paper

COMPARISON OF THE SMOOTHING METHODS: PRELIMINARY RESULTS

G. Damljanović, P. Jovanović and B. Jovanović

Astronomical Observatory, Volgina 7, 11000 Belgrade, Yugoslavia

(Received: February 12, 1997)

SUMMARY: The least-squares collocation method (LSC), the Whittaker - Robinson - Vondrak (WRV), the cubic spline (SPL) and the least - squares method (LSM) are compared. The smoothing by these methods was applied to the 5-day Polar Motion (PM) data (IERS, 1993) in the period 2443989 JD (April 25, 1979) - 2445984 JD (October 10, 1984); the number of the input data was $N = 400$. Some results of smoothing and comparison are presented.

1. INTRODUCTION

The real observations or the observed values l_i are composed of the signal s_i and the random noise n_i. In our interest is to filter the noise and to find estimation of the signal $\hat{s}_i$ which is close to s_i as much as possible. The LSC is a linear transform $\hat{s}_i = Fl_i$ and sometimes appropriate to filtering because it is a method of stochastic filtering. If we know the autocovariances of the signal we can filter the noise. The theoretical base is expounded in Moritz (1980).

The LSC algorithm was taken from Gubanov and Petrov (1994), but a few formulae were corrected for mistakes. We used the results presented in Titov (1995).

Let the time-series l_i (i=1,2,...,N) correspond to the equidistant time moments. The series l_i and s_i are centered. Their autocovariance functions are estimated by the formulae:

$$q_{ll}(j) = \frac{1}{N} \sum_{i=1}^{N-j} l_i l_{i+j},$$

$$q_{ss}(j) = \frac{1}{N} \sum_{i=1}^{N-j} s_i s_{i+j},$$

where $j = 0, 1, \ldots, N-1$. If we consider the series l_i and s_i as vectors $\mathbf{l}$ and $\mathbf{s}$ the problem of filtering is to find the operator F which satisfies the condition $||\hat{\mathbf{s}} - \mathbf{s}||^2 = min$ (Gubanov and Petrov, 1994). We need no mathematical model of the signal but only some of its statistic characteristics (which can be known a priori). Let the vector $\mathbf{n}$ (errors of observations) represent the white noise. Therefore, the signal and the noise are not correlated with each other. The variance of the white noise σ_n^2 can be calculated from the observed data without any assumption, and if we know the parameter σ_n^2 we could perform the filtering (Gubanov and Petrov, 1994). The covariance function of the data and that of the signal differ only when index $j = 0$. We can calculate σ_n^2 as a difference between the calculated value of the covariance function $q_{ll}(j)$ of the vector $\mathbf{l}$ and the covariance function $q_{ss}(j)$ of the signal s, when $j = 0$ (Gubanov and Petrov, 1994); $\sigma_n^2 = q_{ll}(0) - q_{ss}(0)$. In the same paper, it was shown that the basic formula of LSC filtering is:

$$\hat{\mathbf{s}} = Q_{ss} Q_{ll}^{-1} \mathbf{l},$$

where Q_{ss} and Q_{ll} are covariance matrices of the signal and raw data which have the following simple and symetric form:

$$Qu = \begin{pmatrix} qu(0) & qu(1) & \dots & qu(N-1) \\ qu(1) & qu(0) & \dots & qu(N-2) \\ \vdots & \vdots & \ddots & \vdots \\ qu(N-1) & qu(N-2) & \dots & qu(0) \end{pmatrix}.$$

One of the basic problems of LSC is to provide the autocovariance function q_{ss}.

2. FILTERING OF THE IERS PM SERIES

For filtering of the Earth's Polar Motion by LSC the model of its covariance function can be used, as described in Petrov et al. (1995):

$$C_p(\tau) = P_c e^{-\frac{\omega_c \tau}{2Q_c}} \begin{pmatrix} \cos\omega_c\tau & -\sin\omega_c\tau \\ \sin\omega_c\tau & \cos\omega_c\tau \end{pmatrix},$$

where C_p is the steady-state covariance function of the Polar Motion vector $\mathbf{p}(x,y)$, P_c is the power of the Chandler wobble, ω_c is the angular Chandler freqency and Q_c is dimensionless quality factor. In this paper the annual, semiannual term and linear trend were added to C_p. The values $P_{ch} = 432^d$ (about $1^y.181$), $P_a = 365^d.25$ and $P_s = 182^d.62$ (the periods of the Chandler, annual and semiannual wobbles, respectively) were used. To determine P_{ch}, the periods of seasonal terms (P_a and P_s) were fixed, and P_{ch} was varied from $1^y.170$ to $1^y.200$, by the lag $0^y.001$, with the purpose to obtain the best fitting by LSM. The linear trend has also been included. The minimum standard deviation for the x - component σ_x, obtained by this method, was $0''.0097$, for the y - component $\sigma_y = 0''.0152$.

The standard deviations computed from the successive data differences $l_i - l_{i-1}$ $(i = 2, 3, \dots, N)$ are $\sigma_{Ax}/\sqrt{2} = 0''.0096$, for x - component, and $\sigma_{Ay}/\sqrt{2} = 0''.0113$ for y - component, where

$$\sigma_A = \sqrt{\frac{\sum_{i=2}^{N}(l_i - l_{i-1})^2}{N-2}}.$$

The values Q_c (in the LSC method) were chosen to minimize the standard deviation. They were $Q_c = 10$ for the x - component and $Q_c = 12$ for the y - component (with the standard deviations $0''.0127$ and $0''.0276$, respectively).

The smoothed curve by LSC, the residuals and amplitude periodograms computed by direct Fourier transforms (FT) for the x - component are presented in Fig. 1. , Fig. 3. and Fig. 5., for the y - component in Fig. 2., Fig. 4. and Fig. 6. respectively.

The value of the smoothing paeameter ε in WRV method was determined in usual way, as explained in Vondrák (1977): it has been varied to obtain $M = \sqrt{\frac{\sum_{i=1}^{N}(l_i - s_i)^2}{N-1}}$ (standard deviation of residuals) close to σ_x (as much as possible) for x - component and σ_y for y - component. The obtained value was $\varepsilon = 10^{-8}$, for x - component and $\varepsilon = 10^{-9}$, for y - component. For that case, the smoothing curves of x and y - components are presented in Fig. 1. and Fig. 2., the corresponding residuals in Fig. 3. and Fig. 4., and the amplitude periodograms (FT) of residuals in Fig. 5. and Fig. 6. respectively.

On the other hand, when the values of ε were determined to obtain the smoothing residuals whose standard deviations M were equal to the standard deviations obtained from the successive data differences ($\sigma_{Ax}/\sqrt{2}$ for x - component and $\sigma_{Ay}/\sqrt{2}$ for y - component), ε was changed only by y - component (the new value was $\varepsilon = 10^{-8}$). The amplitude periodogram (FT) of y - residuals is shown in Fig. 9. Evidently, the peaks for Chandler and annual periods are remarkably less than ones in the case $\varepsilon = 10^{-9}$. The peak for semiannual period is still less than one in the case $\varepsilon = 10^{-9}$ but not significantly. It is because the residual systematic errors depend on the frequences and the number N of input data.

Similarly, the smoothing parameter S (in SPL method) determined so that the standard deviations of x and y residuals be close to σ_x and σ_y respectively are: 135 and 360. The smoothing curve, the residuals and the amplitude periodograms (FT) for x - component are presented in Fig. 1. , Fig. 3. and Fig. 5.; for y - component in Fig. 2., Fig. 4. and Fig. 6. respectively.

Also, the parameter S can be chosen in usual way, from confidence interval $[N - \sqrt{2N}, N + \sqrt{2N}]$ (Reinsch, 1967). The limit and the mean values of S are 372, 400, 428. The corresponding amplitude periodograms of residuals of smoothing for x and y - components are presented in Fig. 7. and Fig. 8. It may be seen that systematic errors of residuals remain in all cases. For y - component the changes of amplitude periodogram are negligible, but for x - component they are large (because $S = 135$ is fairly less than the lower limit of the confidence interval). As in the WRV method, the greatest differences of the amplitude periodograms are at the peaks for Chandler and annual periods, but for the semiannual peak the corresponding differences are less (see Fig. 7. and Fig. 8.). The reason is the same as in the case of the WRV; the residual systematic errors depend on the frequences and the number N of input data.

Also, we used LSM for smoothing and found the following results:

	A_{ch}	Ph_{ch}	A_a	Ph_a	A_s	Ph_s
x	$0''.172$	216^o	$0''.085$	101^o	$0''.018$	247^o
y	$0''.166$	126^o	$0''.086$	4^o	$0''.006$	184^o

where: A_{ch}, A_a, A_s are the amplitudes and Ph_{ch}, Ph_a, Ph_s are the phases (for Chandler, annual and semiannual wobble, respectively). The phases are obtained for the instant 2443989 JD.

Evidently, by the above described smoothing the relative large systematic errors are not eliminated from the residuals. The largest ones vary with the Chandler, annual and semiannual periods. Disadvantage of LSM is the necessity of knowledge of appropriate class of functions for the best fitting of the input data. The periods of P_{ch}, P_a and P_s and the

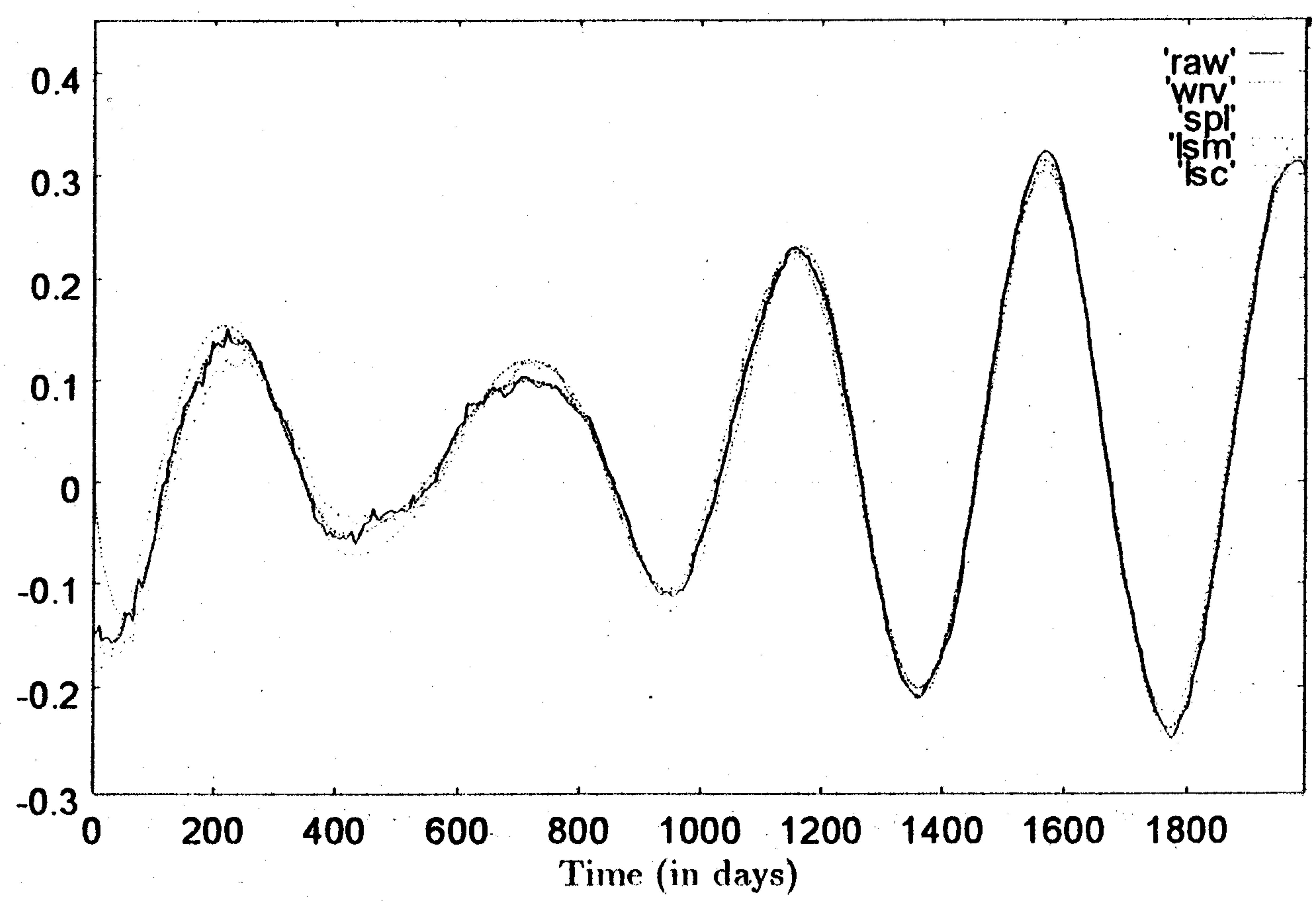

Fig. 1. Raw and smoothed data (x - component) (″).

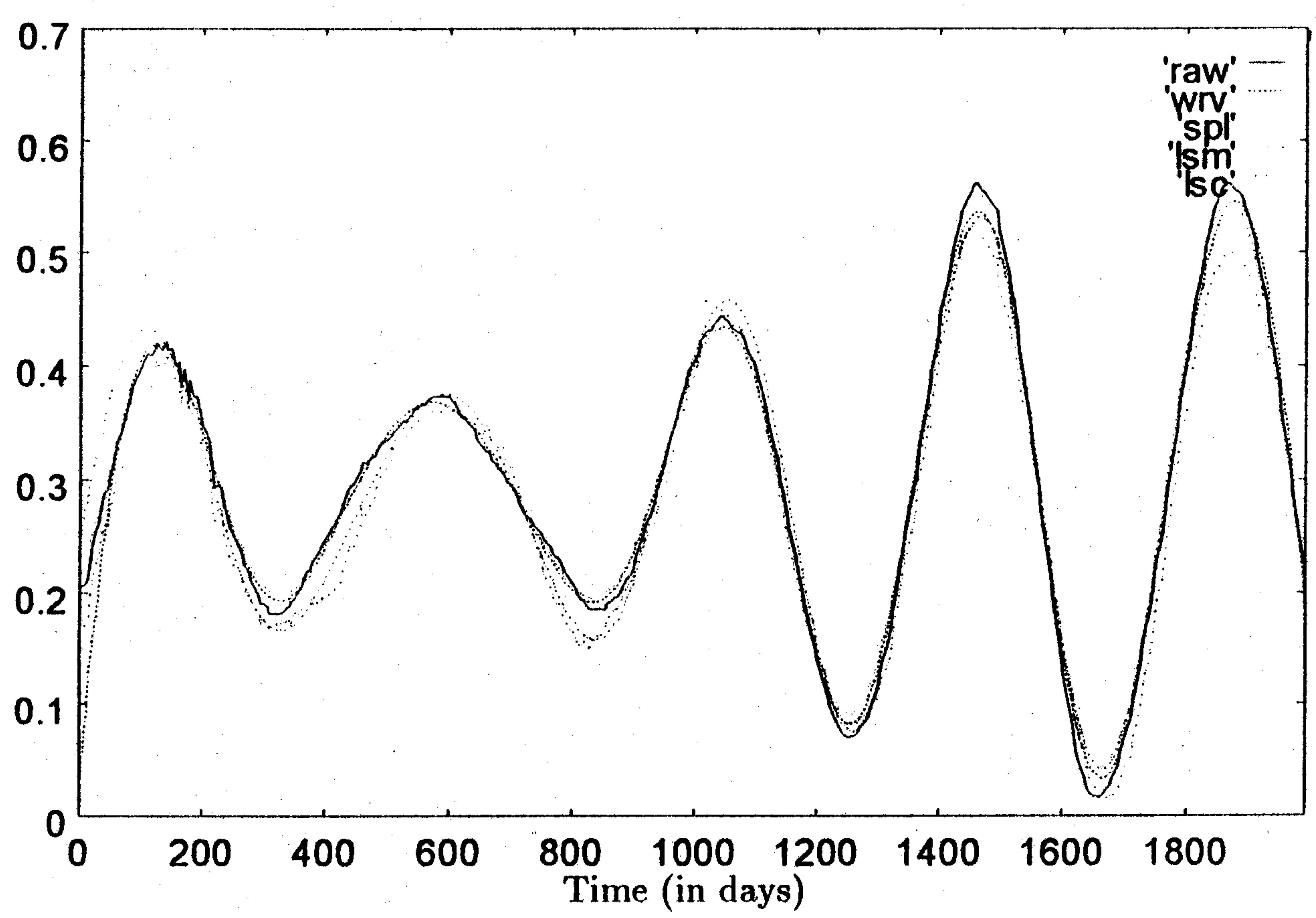

Fig. 2. Raw and smoothed data (y - component) (″).

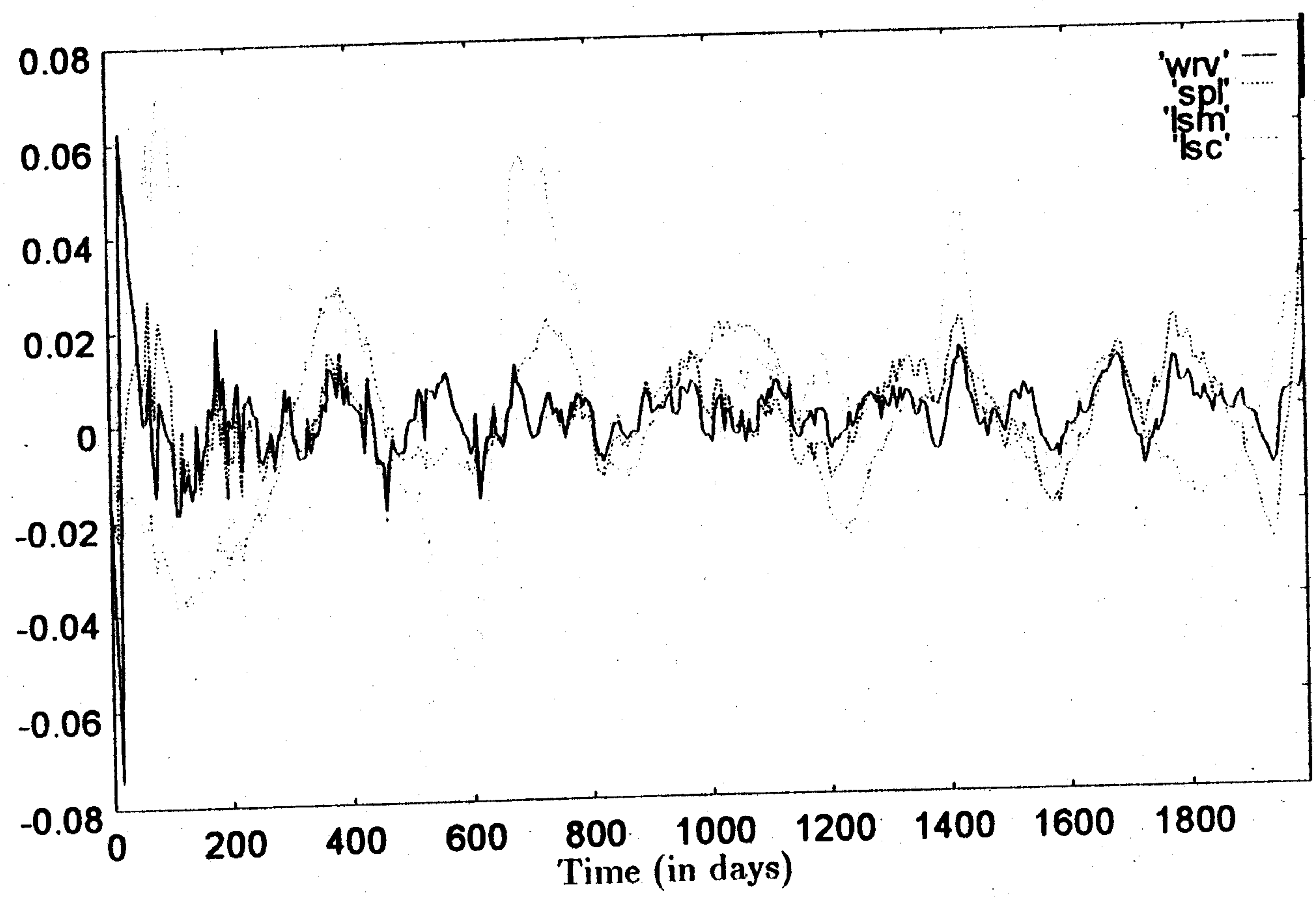

Fig. 3. Residuals of x - component (″).

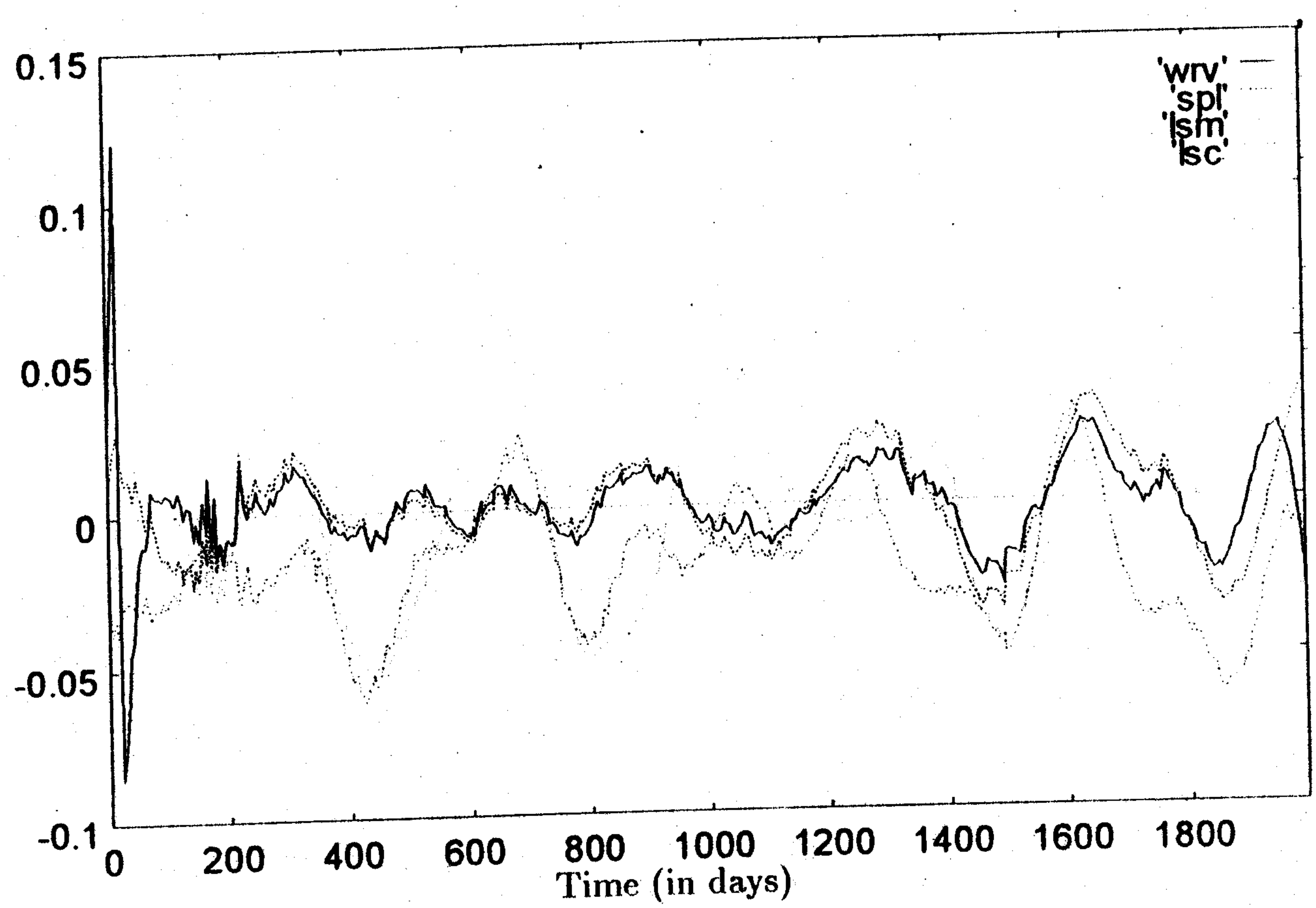

Fig. 4. Residuals of y - component (″).

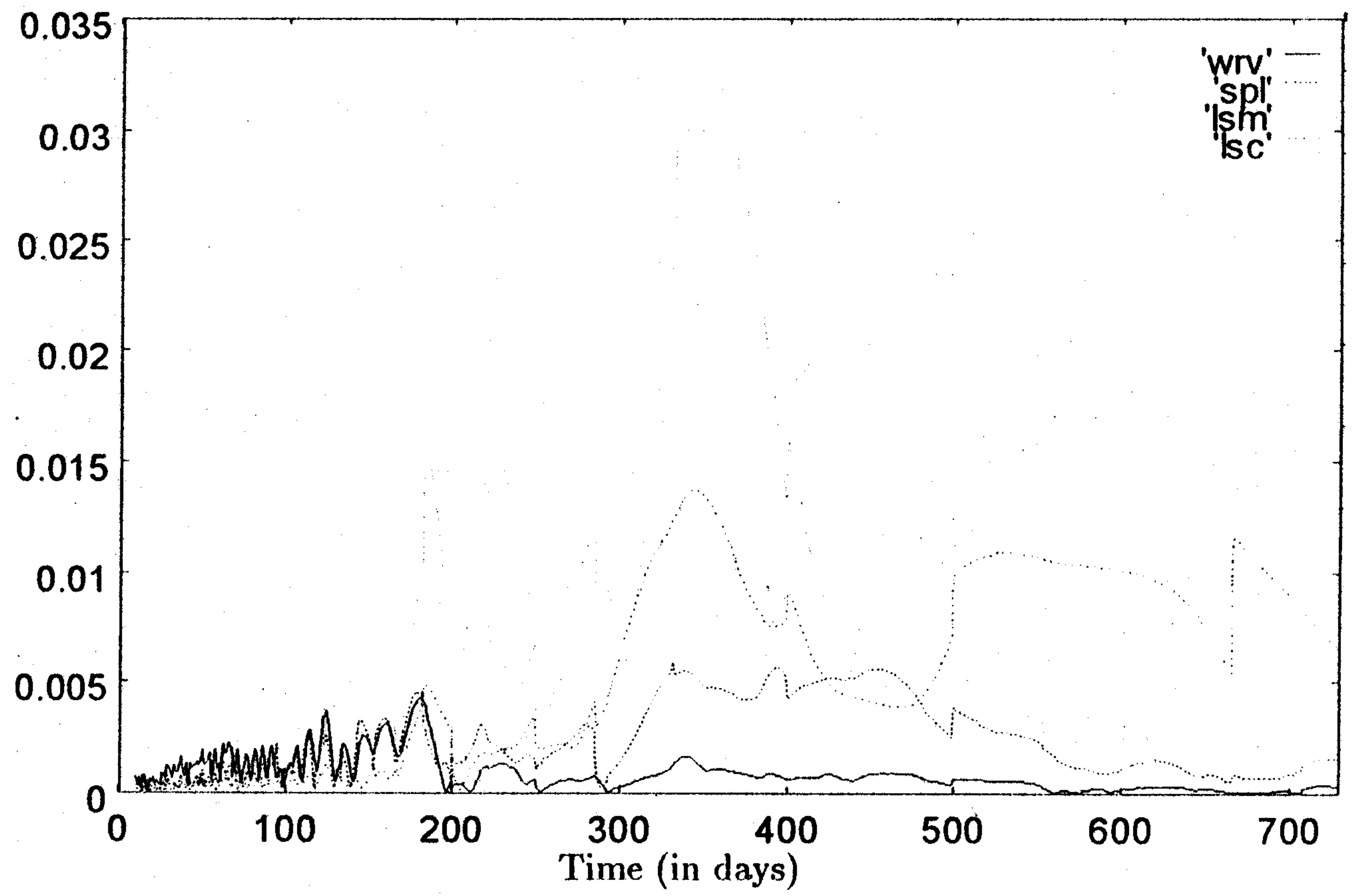

Fig. 5. Periodograms (FT) of residuals of x - component (″).

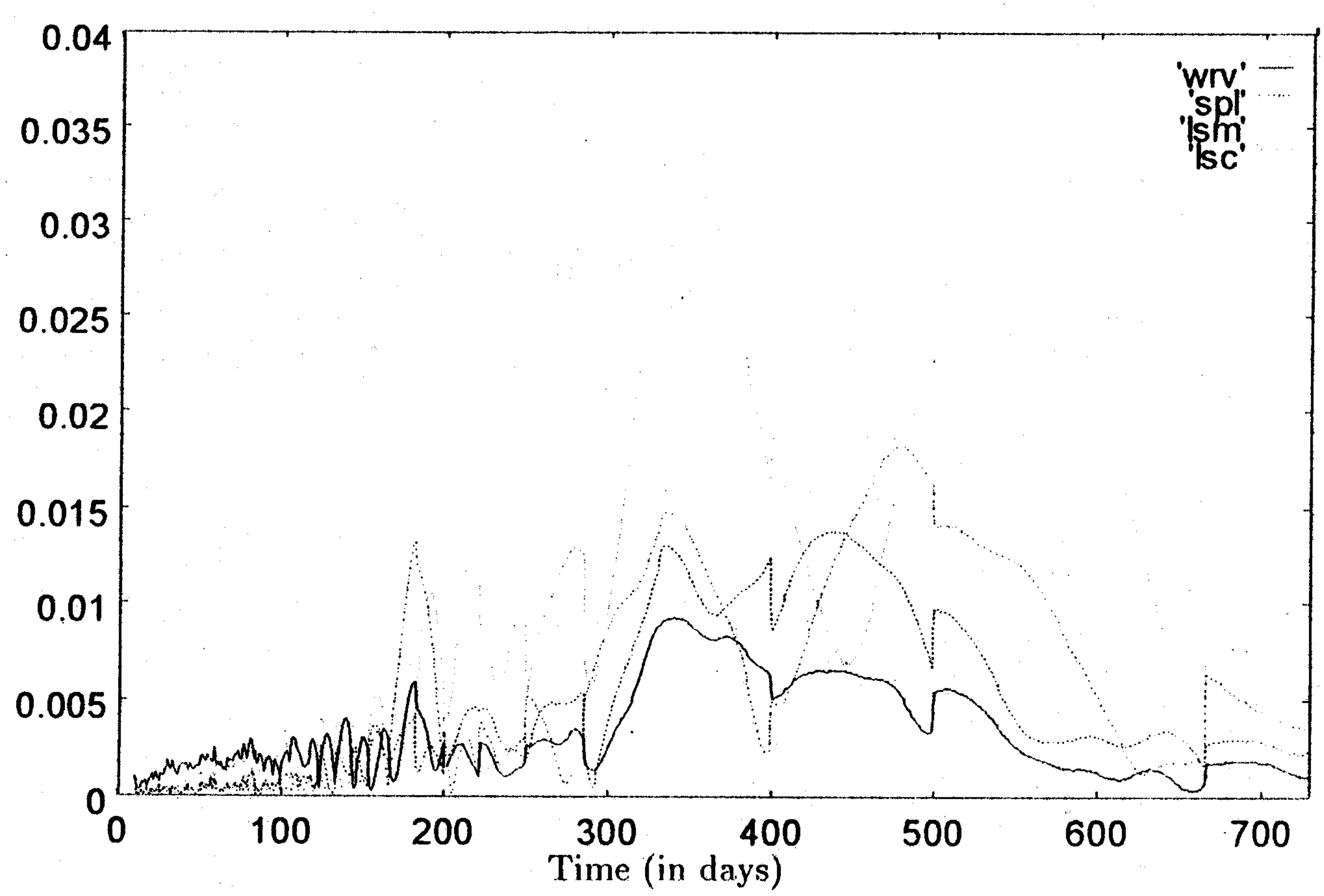

Fig. 6. Periodograms (FT) of residuals of y - component (″).

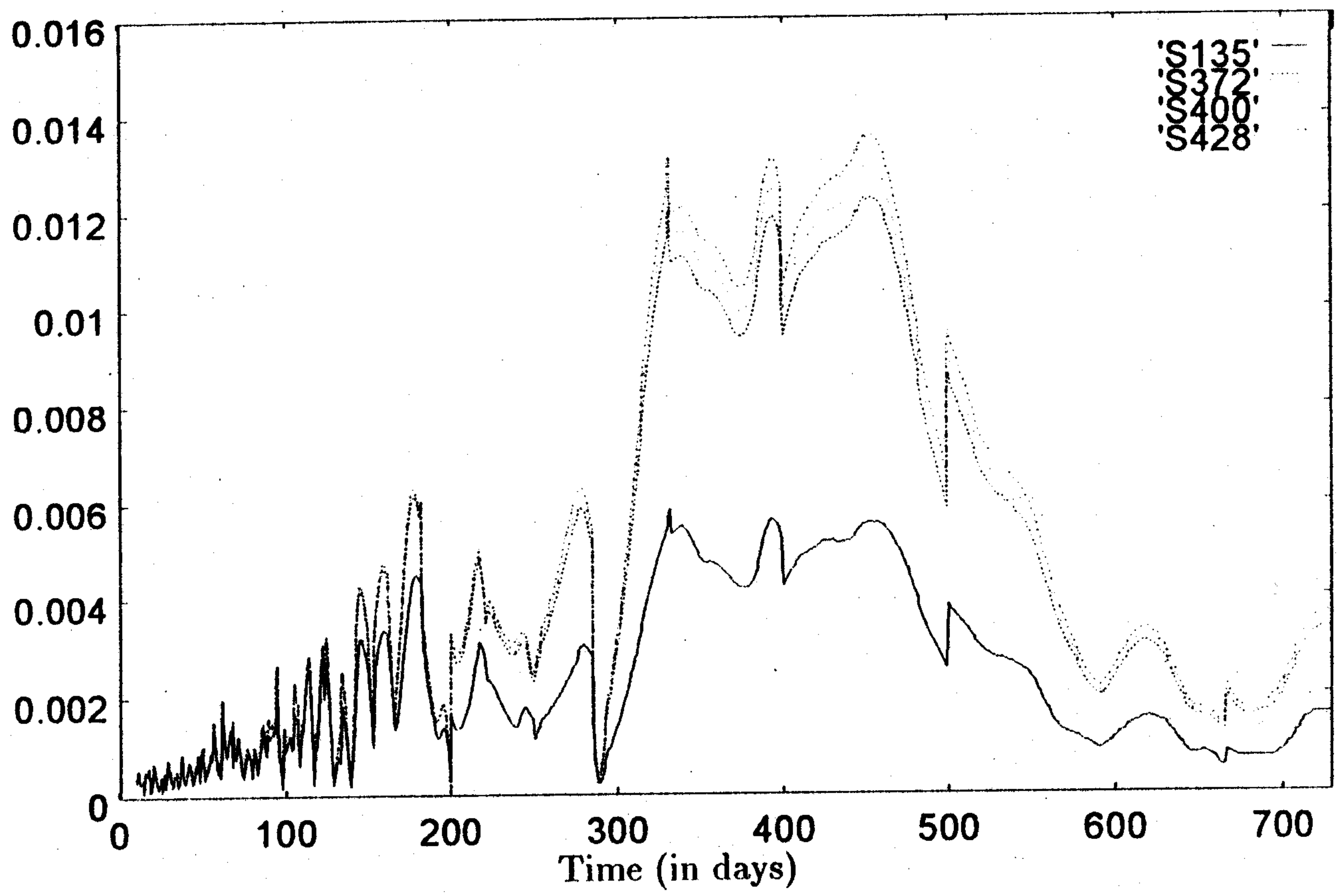

Fig. 7. Periodograms (FT) of residuals (SPL) of x - component (″).

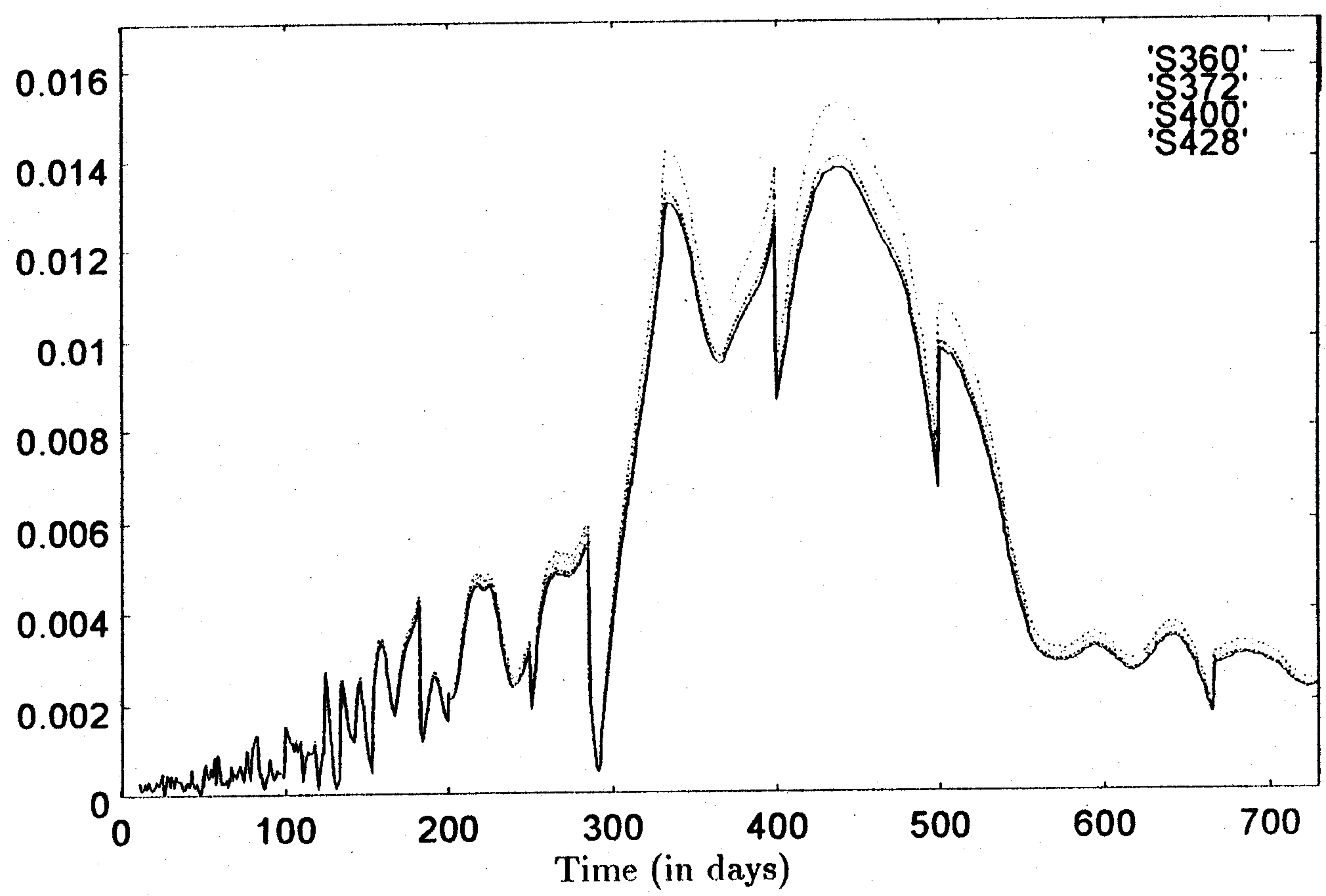

Fig. 8. Periodograms (FT) of residuals (SPL) of y - component (″).

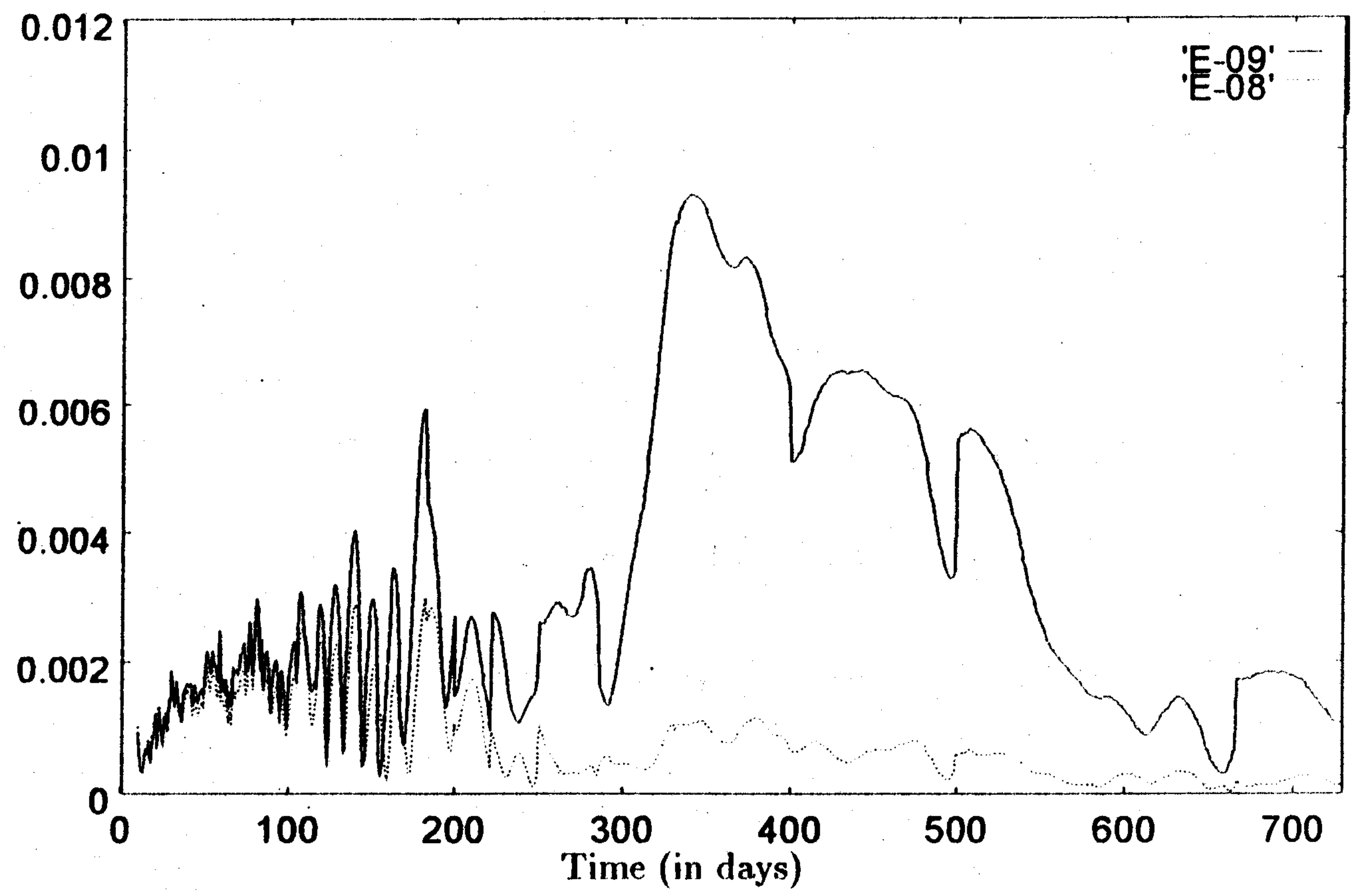

Fig. 9. Periodograms (FT) of residuals (WRV) of y - component (″).

amplitudes (A_{ch}, A_a, A_s) were assumed constant in the given time interval, but the real values vary with time.

Because of the short interval (about 6 years), it was not possible to separate the Chandler and annual wobbles by FT (see Figures 5. - 9.).

3. CONCLUSION

The smoothing methods, which use the polinomials (like WRV and SPL) or the models (LSC for autocovariance function and LSM for signal), are often not appropriate to PM data smoothing. As a consequence the systematic discrepances remain in the residuals.

For using LSC in any real case it is necessary to have a good model of autocovariance function. In the LSM case good model of signal is required. The LSC, WRV and SPL are flexible methods but the LSM is not. It could be noticed that the largest systematic errors of residuals (after applying LSM) represent the residual Chandler, annual and semiannual oscilations (see Fig. 5. and Fig. 6.).

Besides the end faults present in the WRV and SPL methods (see Fig. 3. and Fig. 4.), the base of these methods are the third order polynomials which are not efficient for the approximations of the harmonic oscillations, present in PM data series. The existing systematic errors (and the values of their amplitudes) in the residuals depend on the frequences and the number N of input data.

Acknowledgements – This work is a part of the project "Astrometrical, Astrodynamical and Astrophysical investigations", supported by Ministry of Science and Technology of Serbia.

REFERENCES

*** : 1993, *IERS Annual Report 1992*, Published by IERS, Observatoire de Paris.

Gubanov, V., Petrov, S.: 1994, *Filtering and Forecasting of EOP-Series by Least-Square Collocation Technique*, Proc. of V Russian Symposium "Time and Space Metrology", Mendeleevo, Russia, Oct. 11-13, 1994, IMVP GP "VNIIFTRI", 208-211.

Moritz, H.: 1980, *Advanced Physical Geodcsy*, Herbert Wichmann Verlag Karlsruhe, Abacus Press Tunbrige Wells Kent.

Petrov, S., Brzeziński, A., Gubanov, S. P.: 1995, *Comparison of the Kalman Filter and the Least Squares Collocation as Applied in Earth Rotation Investigations*, Presented at Proc. of Journées 1995, "Systémes de Référence Spatio-Temporels", Space Research Centre, PAS, Warsaw, 125-128.

Reinsch, C.: 1967, *Smoothing by spline functions*, Numer. Math. Bd. **10**, 177-183.

Titov, O.: 1995, *Least-Squares Collocation Method for VLBI Data Analysis*, Proc. Journeés 1995 "Systemes de Reference Spatio- Temporels", Warsaw, Poland, Sep. 18-20, 1995, Published by the Space Research Centre, PAS, 129-130.

Vondrák, J.: 1977, *Problem of Smoothing Observational Data II*, Bull. Astron. Inst. Czech., **28**, 84-89.

ПОРЕЂЕЊЕ МЕТОДА ИЗРАВНАЊА: ПРЕЛИМИНАРНИ РЕЗУЛТАТИ

Г. Дамљановић, П. Јовановић и Б. Јовановић

Астрономска опсерваторија, Волгина 7, 11000 Београд, Југославија

УДК 52–17
Оригинални научни рад

Примењене су следеће методе изравнања: најмања квадратна колокација (LSC), Витакер - Робинсон - Вондрак-ова (WRV), кубни сплајн (SPL) и метода најмањих квадрата (LSM) на компоненте PM (IERS, 1993) из интервала 2443989 JD (25. IV 1979) - 2445984 JD (10. X 1984) на ком je 400 тачака са еквидистантним размаком од 5 дана. Приказани су резултати поређења наведених метода.