Statistica Sinica 2(1992), 605-618

CONTRIBUTION OF TIME SERIES ANALYSIS TO DATA PROCESSING OF ASTRONOMICAL OBSERVATIONS IN CHINA

Dawei Zheng and Shifang Luo

Shanghai Observatory, Chinese Academy of Sciences

Abstract: In this paper, studies and applications of Time Series Analysis in China to data processing of astronomical observations are described in detail. They include maintaining the stability of an astronomical measurement system, prediction of Earth Rotation Parameters, identification of disconnected data, improvement of the edge effects of the series data, detection of implicit periodic terms and narrow-band filtering of astronomical data.

Key words and phrases: Time series analysis, leap-step autoregressive, multi-stage filter, astronomical observations, earth rotation.

1. Introduction

Since publication of Box and Jenkins (1970), methods in time series analysis (TSA) have been applied widely to scientific and economic fields, and have won great success. Chinese astronomers have used and studied various techniques of TSA for data processing. The status of this work is summarized in Table 1.

2. AR Model for Maintaining Stability of a Measured System

The accurate measurement of a physical state is made by using several instruments in order to reduce the random errors in measurements. However systematic deviations of instruments caused by differences between various instruments must be eliminated so that the stability of a measured system can be maintained.

In the Chinese Joint Universal Time (UT) System consisting of several optical instruments, the Auto-Regressive (AR) model has been used by Zheng et al. (1980) and Luo et al. (1981) in the determination of systematic deviations of instruments. The adopted AR model for a data series Z_n $(n = 1, 2, \ldots, N)$ is described by

$$Z_n = \sum_{i=1}^{P} a_i \, Z_{n-i} + E_n \qquad (1)$$

where a_i is the coefficient of the AR model and is derived by the Yule-Walker equation consisting of autocorrelation estimates with different time lags of the data series Z_n, E_n is white noise, and P is the order for the AR model determined by the FPE or AIC criteria presented by Akaike (1973).

The results from systematic deviations of instruments by AR models have been compared with those reduced by the method of Bureau International de l'Heure (BIH) in Paris, in which systematic deviations are fitted with constant, annual and semi-annual terms. The differences between real and predicted systematic deviations obtained by the two methods mentioned above are drawn in Figure 1.

It is shown that for various instruments (perhaps for various observational techniques), the predictions of systematic deviations obtained by the AR model are more coincident with real systematic deviations. In conseqence, AR models are valid for maintaining the stability of a measured system. Wang and Jin (1988) constructed a new global solution of Earth Rotation Parameters (ERP) by using 10 data series obtained from new observational techniques. AR models were also successfully adopted in this work to predict systematic deviations of these data.

3. TAR Model for Predicting Earth Rotation Parameters

The Threshold Auto-Regressive (TAR) model (Tong and Lim (1980)) is a class of models for nonlinear time series such as

$$Z_n = a_0^{(j)} + \sum_{i=1}^{P_j} a_i^{(j)} Z_{n-i} + E_n^{(j)} \quad Z_{n-d} \in R_{j=1,2,\ldots,L} \tag{2}$$

where d is a lag parameter for the TAR model, R_j is the jth threshold domain of the data series Z_n, P_j is the order for the AR model of the series Z_n belonging to the R_j, $a_i^{(j)}$ is the coefficient for the jth AR model, $E_n^{(j)}$ is the white noise of the jth AR model, but $E_n^{(j)}$ and $E_n^{(j')}$ are independent when $j \neq j'$. These parameters of d, R_j, P_j and $a_i^{(j)}$ in (2) can be simultaneously estimated by the method of minimal AIC.

Using the data series for 1968 through 1981 which was reduced by BIH from optical observations, Zheng et al. (1982, 1986a) adopted the TAR model to forecast three components of Earth Rotation Parameters. The three components are the X and Y components characterizing the motion of earth's pole as well as the UT component describing the motion of the earth's rotation. The dispersions between the predicted values calculated by the three TAR models and the real ones for X, Y, and UT are listed in Table 2. The results obtained by using the fitting functions proposed by United States Naval Observatory are also given in Table 2 for comparison purpose.

It is seen from Table 2 that the predicted values of ERP by the TAR models have higher accuracy, especially for the predicted UT. There are sufficient reasons to believe that better forecasts could be obtained by TAR models if the data series from the space astronomical techniques, such as Very Long Baseline Interferometer (VLBI), Satellite Laser Ranging (SLR), and etc. are used.

X and Y Components from 1963 to 1967 are predicted by the same TAR models mentioned above. The dispersions of the predicted results are listed in Table 3. It is seen that the predicted values of ERP by TAR models are not only of higher accuracy but also of more stability in longer periods than that of the existing other methods.

4. Identification of Disconnected Astronomical Data by the AIC Method

Disconnected astronomical data may occur when the observation session is long. For example, the observed data of VLBI are broken into several sections due to the changes of frequency or phase of the VLBI clock in a 24-hour observation session (Robertson (1975)). The AIC (Akaike Information Criterion) method is used to identify and estimate the parameters for the number and the instant of broken clock and for the models of clock behaviour in separate sections in order to improve the processing procedure of VLBI (Zheng et al. (1986b)).

In regard to the practical problem for the disconnected astronomical data, the followed information criterion is adopted:

$$\mathrm{AIC}(\hat{B},\hat{T},\hat{K}) = \min \sum_{j=1}^{L}\left(N_j' \ln\left(\frac{\mathrm{RSS}\,(K_j)}{N_j'}\right)+2(K_j+1)\right), \quad L=1,2,\ldots,L_m, \tag{3}$$

where $\hat{B}$, $\hat{T}$, and $\hat{K}$ are the optimal estimates for the number, the instant of clock discontinuities and the orders of clock behaviour models, respectively. L_m is the maximum number of clock behaviour models possibly existing in the VLBI observations, RSS is the residual sum of squares after fitting the jth model and N_j' is the number of equations for solving the jth model. From formula (3), the optimal estimates of B, T, and K must simultaneously satisfy the condition for a minimum of AIC.

The AIC method is tested by using the data of VLBI observations in 6 different days of 1983 when the changes of the VLBI clock occurred. The results of the estimated clock models by the AIC method and with the manual method adopted by the National Geodetic Survey, NOAA, USA are listed in Table 4, in which N is the sample size of data in each day. Both the data and the results for two days are also presented in Figure 2.

From Table 4 it is clear that the clock models and the broken instant estimated by the above two methods are different, and that most of the clock models have lower orders and the residual mean squares (RMS) are smaller for the AIC method.

The AIC method is currently in operation in the VLBI software of Shanghai Observatory, resulting in saving the processing time and in improving the accuracy of the results (Luo et al. (1987)).

5. LSAR Model for Limiting the Edge Effects of Data Series

In the processing procedure of observation series, distortion near both ends of data series may occur when several methods, such as the smoothing, the digital filter, the fitting function and so on, are used in the data processing. The Leap-Step Auto-Regressive (LSAR) model has been suggested and utilized by Dong and Zheng (1985) in order to reduce the edge effects of the data series. The LSAR model is described as

$$Z_n = \sum_{i=1}^{K_j} A_i^{(j)} Z_{n-i} + E_n^{(j)} \quad Z_n \in R_j \quad j = 1, 2, \ldots, L. \tag{4}$$

where, Z_n $(n = 1, 2, \ldots, N)$ — data series.

R_j denote the jth leap-step domain of data series. If $N = L \times M$, then $Z_{j+(L\times K)} \in R_j$ and $K = 0, 1, \ldots, M-1$.

K_j and $A_i^{(j)}$ denoted the order and the coefficients of autoregressive model in the jth leap-step domain of data series.

$E_n^{(j)}$ denoted the white noise of data in the jth leap-step domain.

An artificial simulated series is considered, which consists of 10 sine waves that have the periods of 0.5, 0.6,..., 1.4 years, respectively. The values of the amplitudes for most of the terms are taken as 1, but the amplitudes for the terms of 1.0 and 1.2 years as 10 and 15, respectively. In addition, a normal noise with $N(0,2)$ is mixed in the series. The edge effects (solid curve) of the smoothing curve which is obtained from the smoothing artificial simulated series with Vondrak's method (Vondrak (1969, 1977)) are shown in Figure 3. The edge effects (dotted curve) will decrease, obviously, if the LSAR model is adopted in the smoothing procedure.

It is clear that for denser data series, such as Atmospheric Angular Momentum (AAM) and VLBI intensive data, the accuracy of prediction by the LSAR model might be better than that by other models of time series analysis. When this paper was written, we noted, with great interest, the paper of Cleveland and Tiao (1979) in which the seasonal (or periodic) time series was studied.

6. AR Spectral Analysis for Detecting the Periodic Signals in Astronomical Data

Since the 1970's, Chinese astronomers have used AR spectral analysis (i.e., MEM) for studies of Earth rotation (Zheng (1978), Zheng and Zhao (1979), Gu (1986), and Li and Wilson (1987)), polar motion (Zhao and Zheng (1980), Zhang et al. (1982)), solar activity (Yang and Zhao (1988)), planet motion (Yang (1988)), and the analysis of systematic errors arising in the data from various observation techniques (Zheng et al. (1980, 1984)). The resolution and the adopted order in the AR spectral estimate were also discussed in their work (Zheng and Zhao (1979), Zhang et al. (1986)).

Recently, Zheng and Dong (1987) compared the Marple (1980) and Burg algorithm (Smylie et al. (1973)) in the AR spectral estimate by detecting astronomical tidal monthly (Mm) and fortnightly (Mf) terms from UT data of VLBI intensive observations from April 1985 to March 1986. The spectral results of Mm and Mf waves are shown in Figures 4 and 5, respectively. From the deviations between the theoretical spectral lines and the spectral peaks obtained by the two algorithms, it is confirmed that more precise estimates of frequencies will be from the Marple algorithm. Moreover, it is also noted from their work that the Marple algorithm will decrease the splitting of spectral peaks.

Using the VLBI intensive UT data for 3 months during the period of April-June 1984 in international MERIT campaign, the tidal fine signals of high frequency terms are detected by Luo et al. (1987) and are shown in Figure 6. The tidal waves with period of 9.13-day is found clearly besides the tidal waves of both Mm and Mf because the noise is obviously limited in the AR spectral estimate.

7. MSF Method for Separating the Signal Process from Astronomical Data

According to the digital filter theory of time series (Koopmans (1974)) and the Vondrak smoothing method (Vondrak (1969, 1977)), the Multi-Stage Filter (MSF) was suggested and utilized by Zheng and Dong (1985, 1986c). The theoretical formula for the frequency response function of the MSF derived by Zheng and Dong (1986c) is given by

$$R = c\left(1 - A(f,e)^L\right)^M, \tag{5}$$

where c is a real constant, L and M are positive integers, and $A(f, e)$ is the frequency response of the Vondrak filter provided by Huang and Zhou (1981), namely,

$$A(f, e) = \left(1 + e^{-1}(2\pi f)^6\right)^{-1} \tag{6}$$

where f and e are the frequency component and the filtering factor respectively.

The frequency response function for the MSF is characterized by small bandwidth of truncated frequency, so it can be treated as narrow-band filtering, Chinese astronomers have separated the Chandler component with the period of 14 months and the annual one from polar motion series successfully (Zheng and Dong (1986c)), and stidied the problem about the mechanism of Chandler Wobble (Zhao (1988)) by using the MSF method.

The band-pass filter has also been used to study the relationship between the Length Of Day (LOD) and El Nino events (Zheng et al. (1988)). The interannual variation after filtering the LOD data from astronomical observations by MSF is shown in the top part of Figure 7. The monthly departure of Sea Surface Temperature (SST) in the equatorial eastern Pacific area (180° – 80°W, 5°S –5°N) is shown in the bottom part of Figure 7.

From Figure 7, it is apparent that the deceleration and acceleration of the interannual variation of Earth rotation are clearly consistent with the warming and cooling of SST in the equatorial area. If the interannual variation in LOD is calculated in time by the MSF method and the minimum of interannual variation is monitored during routine work, then El Nino events can be predicted in advance of one year, which corresponds to long range forecast in meteorology (Zheng et al. (1990a,b)).

8. Conclusions

Astronomical data and other geophysical data, which have been recorded over time for several hundreds of years, provide a rich source of sample groups for Times Series Analysis. As mentioned above, Time Series Analysis has been successful in improving astronomical data processing in China and has become a very important means of data analysis. With the accumulation of data and the increasing need of high precision, Time Series Analysis will play a more and more important role in prediction, identification and control, parameter estimation, digital filtering, frequency detection and other studies in Astronomy.

Table 1. Applications of Time Series Analysis to astronomical studies in China

Methods of Time Series Analysis		Applications
AR	model	Keep the stability of measured system
TAR	model	Predict the Earth Rotation Parameters
AIC	method	Identify the disconnected data series
LSAR	model	Limit the edge effects of the data series
AR	spectrum	Detect the various periodic terms implied in astronomical data series
MSF	method	Treat the narrow-band filtering of the data series

Table 2. Dispersions of predicted ERP in 1981 by TAR and fitting function models. The units of dispersion are seconds of angle for X, Y and seconds of time for UT.

Predicted period	TAR models			Fitting function models		
(days)	X	Y	UT	X	Y	UT
20	0".006	0".006	$0^s.0028$	0".024	0".016	$0^s.0065$
40	0".016	0".016	$0^s.0042$	0".032	0".026	$0^s.0117$

Table 3. Dispersions of predicted X and Y from 1963 to 1967 by TAR models. The units of dispersions are seconds of angle for X and Y.

	X			Y		
Predicted year	20-day	40-day	60-day	20-day	40-day	60-day
1963	0".008	0".019	0".032	0".008	0".019	0".030
1964	0".007	0".019	0".031	0".007	0".014	0".015
1965	0".007	0".017	0".025	0".007	0".015	0".020
1966	0".007	0".017	0".025	0".006	0".016	0".025
1967	0".009	0".019	0".030	0".008	0".020	0".032

Table 4. Estimated clock models of VLBI by AIC and manual methods. N denotes the sample size of the data for each day and the unit of RMS is nanoseconds of time.

code	date	N	AIC k1 k2	AIC broken time	AIC RMS(ns)	Manual k1 k2	Manual broken time	Manual RMS(ns)
OC122	AUG. 1 83	172	1 2	$2^d01^h24^m$	0.511	2 2	$2^d06^h20^m$	0.544
OC136	OCT.17 83	72	1 1	$17^d23^h27^m$	0.399	1 1	$17^d23^h15^m$	0.399
OC137	OCT.22 83	193	2 2	$23^d10^h16^m$	0.351	1 2	$23^d10^h20^m$	0.378
OC144	NOV.26 83	184	2 1	$27^d10^h01^m$	0.211	2 2	$27^d06^h00^m$	0.216
OC146	DEC. 6 83	165	2 1	$7^d05^h42^m$	0.205	2 2	$7^d06^h00^m$	0.205
OC147	DEC.11 83	196	2 1	$12^d08^h56^m$	0.296	2 2	$12^d04^h15^m$	0.303

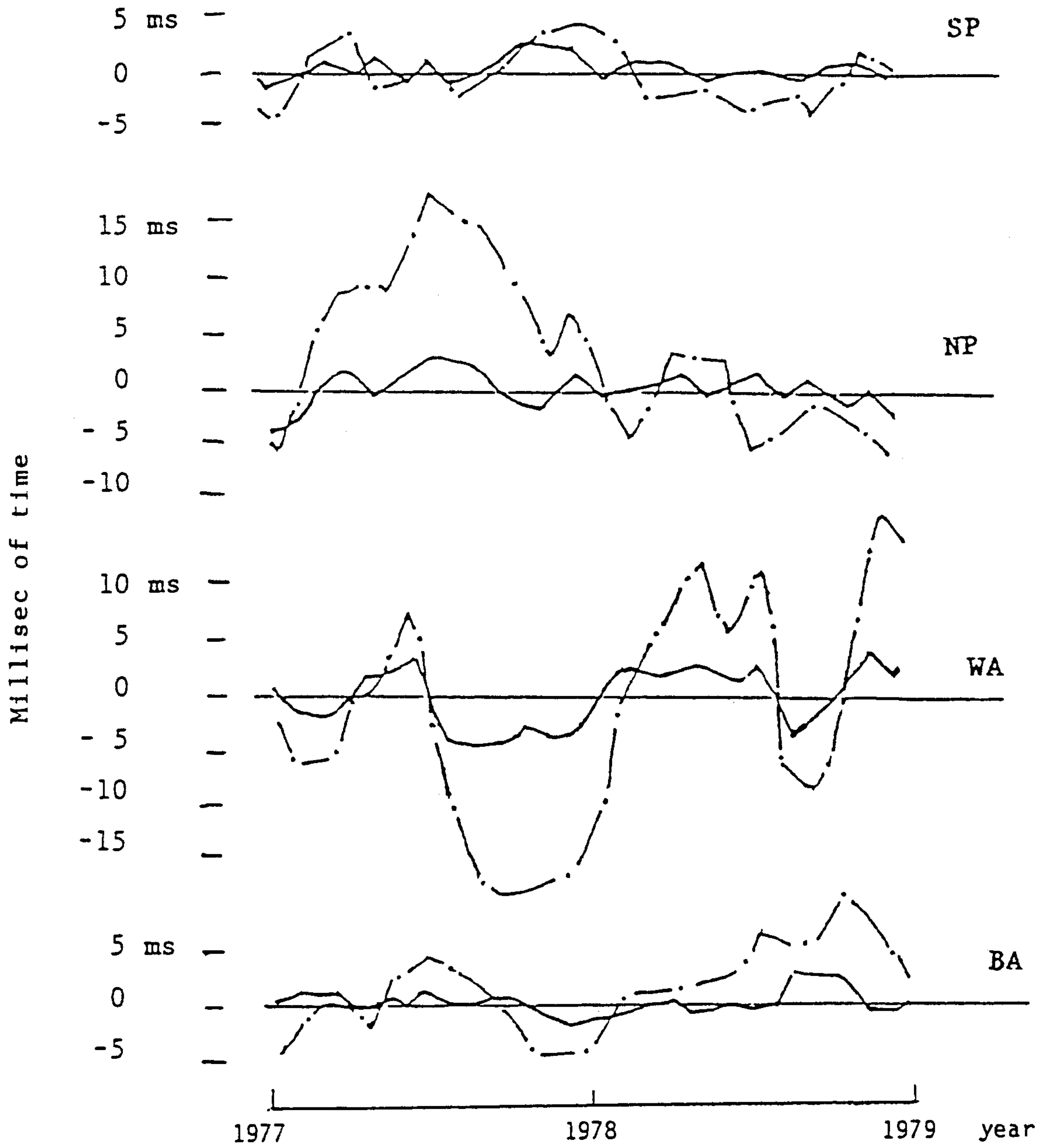

Figure 1. Comparison of the systematic deviations of instruments for two years obtained by the AR models as solid curves and by the BIH methods as dashed curves. Zero lines represent the real systematic deviations for various instruments located in Shanghai, Nanjing, Wuham and Beijing of China and denoted by the symbols of SP, NP, WA and BA, respectively.

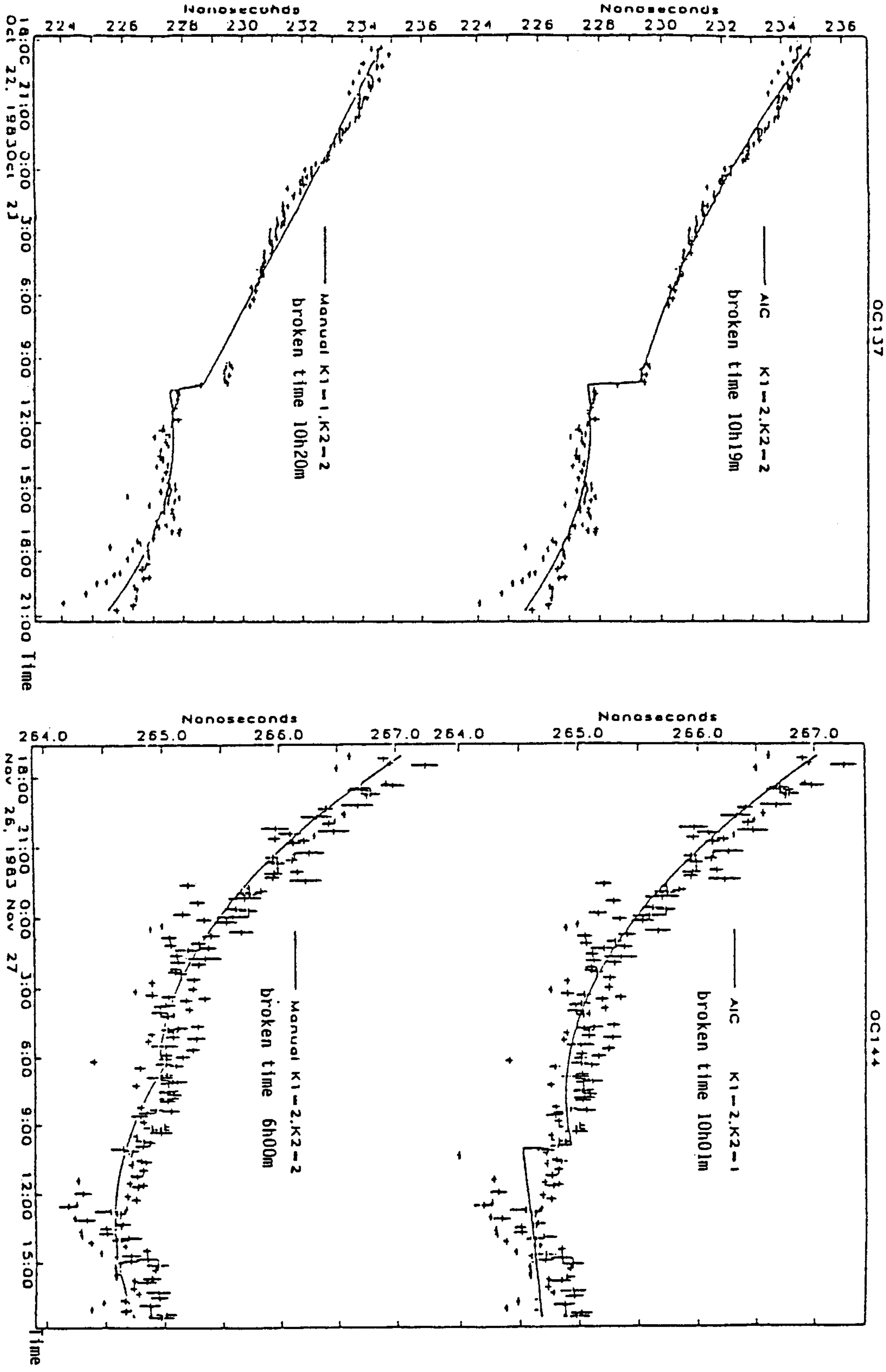

Figure 2. The estimates of clock behaviour models for VLBI observations with the codes of OC137 and OC144 by AIC and manual methods. k1 and k2 represent the orders of the clock models.

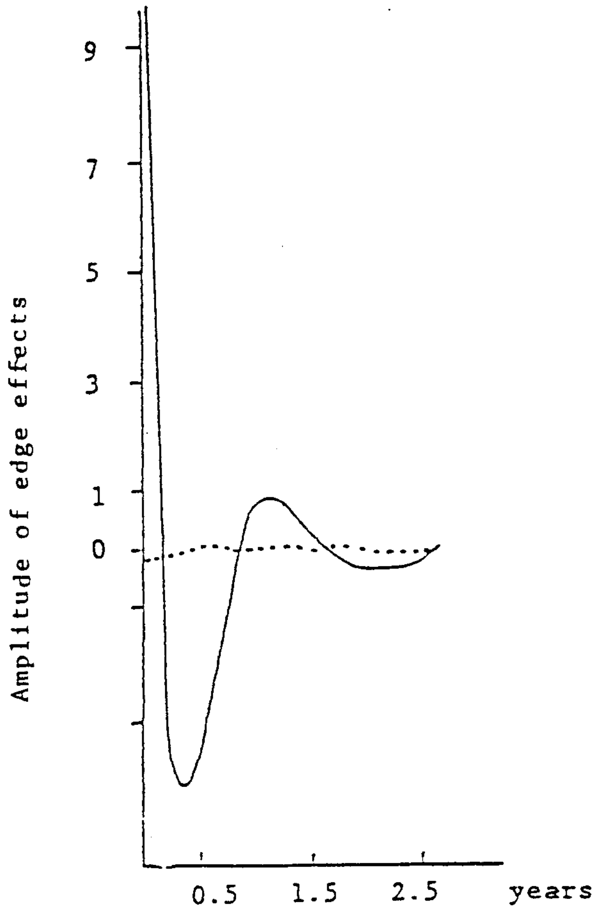

Figure 3. The edge effects of Vondrak's smoothing curve. The solid and dotted curves are obtained without and with LSAR model, respectively.

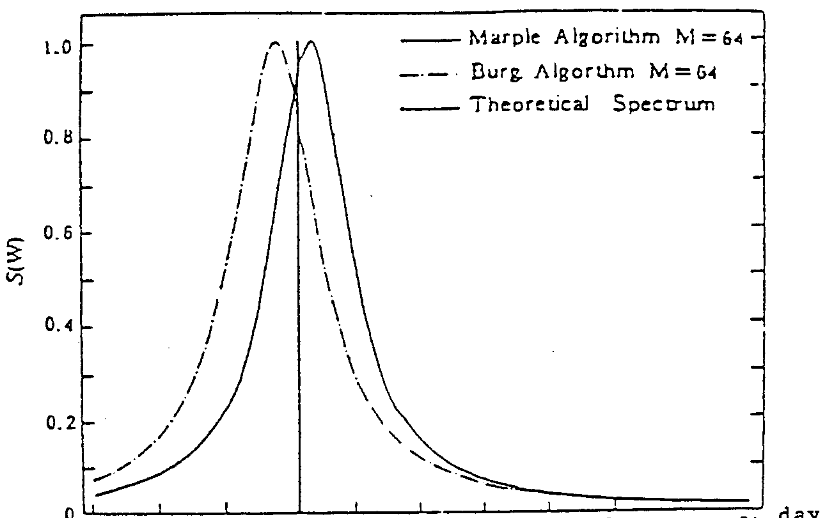

Figure 4. AR spectral estimates of the monthly term of astronomical tidal waves in UT data calculated by Marple and Burg algorithms.

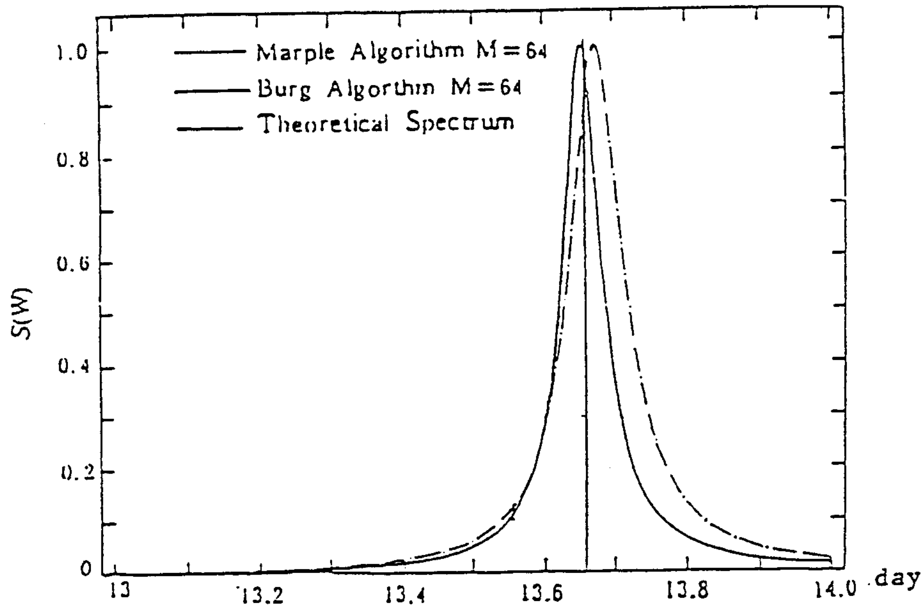

Figure 5. AR spectral estimates of the fortnightly term of astronomical tidal waves in UT data calculated by Marple and Burg algorithms.

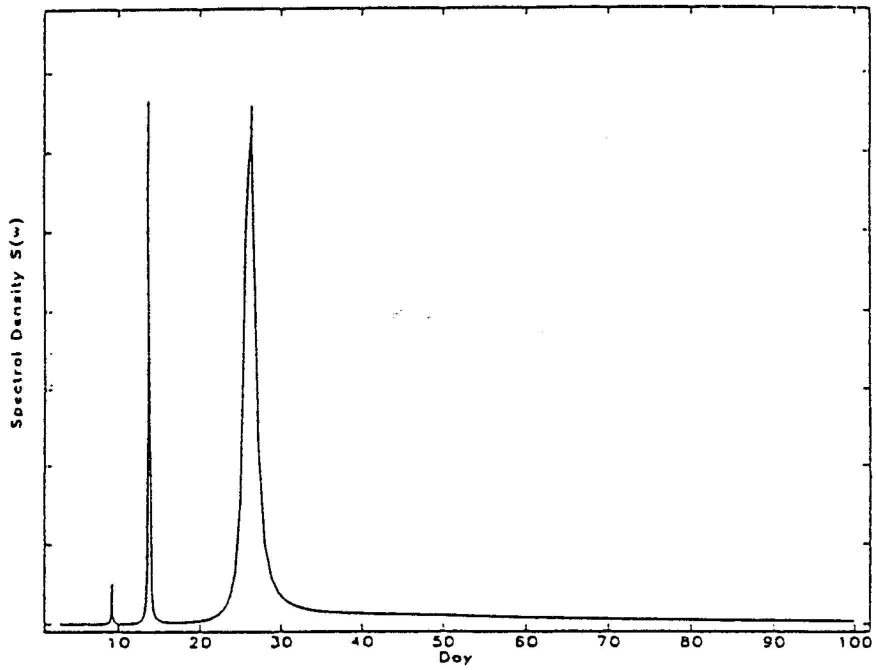

Figure 6. The tidal high-frequency components in UT data detected with AR spectral estimate of Marple algorithm from the VLBI intensive series during April-June, 1984.

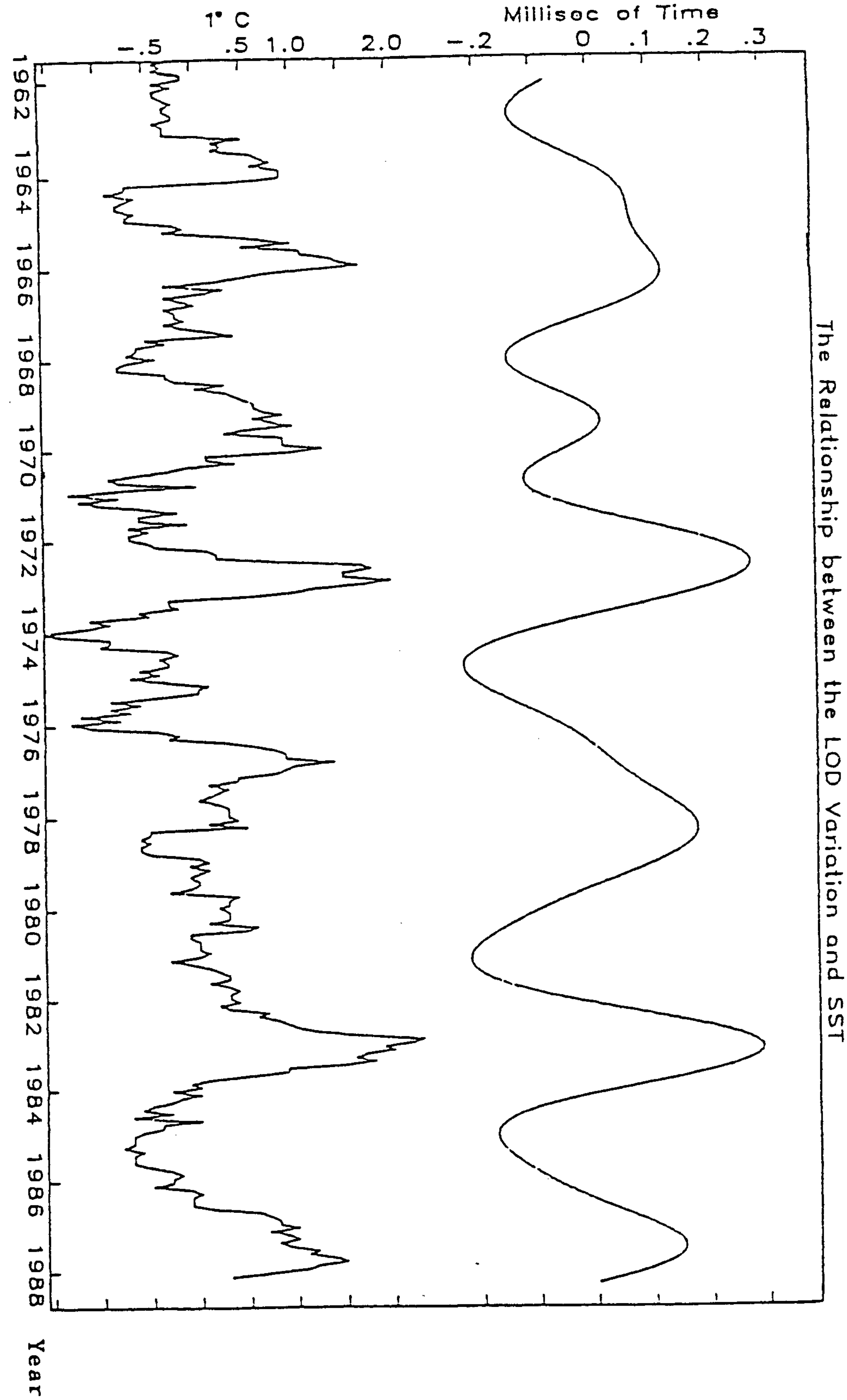

Figure 7. The interannual variation of LOD (in the top part) and the monthly departures of sea surface temperature in the equatorial eastern Pacific area (in the bottom part) during 1962.0 to 1988.0.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *2nd Internat. Symp. Inform. Theory* (Edited by B. N. Petrov and F. Csaki), 267–281. Akademi Kiado, Budapest.

Box, G. E. P. and Jenkins, G. M. (1970). *Times Series Analysis, Forecasting and Control.* Holden-Day, California.

Cleveland, W. P. and Tiao, G. C. (1979). Modeling seasonal time series. *Economic Appliquee* 107–129.

Dong, D. and Zheng, D. (1985). The ends effects of Vondrak filter. *Ann. Shanghai Observatory* **7**, 13–25.

Feissel, M. and Lewandowski, W. (1984). A comparative analyses of Vondrak and Gaussian smoothing techniques. *Bull. Geod.* **58**, 464–474.

Gu, Z. (1986). Analyses of irregular variation of the earth's rotation. *Ann. Shanghai Observatory* **8**, 99–104.

Huang, K. and Zhou, X. (1981). On essentiality of the Whittaker-Vondrak method as a filter, and estimations of standard deviation and correlation for digital filter. *Acta Astron. Sin.* **22**, 120–130.

Koopmans, L. H. (1974). *The Spectral Analysis of Time Series.* Academic press, New York.

Li, Z. and Wilson, C. R. (1987). A damped oscillation model of the 50 day oscillation in the length of day. *Acta Astron. Sin.* **28**, 29–38.

Luo, S., Zheng, D., Liao, D. and Huang, H. (1981). New chinese joint UT1 system. *Acta Geodetica et Cartographica Sin.* **10**, 263–269.

Luo, S., Zheng, D., Zhou, R., Xue, Z. and Chen, J. (1988). Polar motion and UT1 from VLBI. *Ann. Report of BIH for 1987*, 29–31.

Luo, S., Zheng, D., Robertson, D. S. and Carter, W. E. (1987). Short-period variation in the length of day: atmospheric angular momentum and tidal components. *J. Geophys. Res.* **92**, 11657–11661.

Marple, L. (1980). A new autoregressive spectral algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **28**, 441–454.

Robertson, D. S. (1975). M. I. T. (Ph.D. Thesis for VLBI).

Smylie, D. E., Clarke, G. K. and Ulrych, T. J. (1973). *Method in Computational Physics* **13**, 391–430. Geophysics Academic Press, New York.

Tong, H. and Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *J. Roy. Statist. Soc. Ser.B* **42**, 245–292.

Vondrak, J. (1969). Problem of smoothing observational data. *Bull. Astron. Inst. Czech* **20**, 349–360.

Vondrak, J. (1977). Problem of smoothing observational data II. *Bull. Astron. Inst. Czech* **28**, 84–93.

Wang, Q. and Jin, W. (1988). Combined solution of ERP during 1983–1985. *Acta Astron. Sin.* **29**, 79–87.

Yang, Z. and Zhao, M. (1988). An investigation for the cause of formation of the major period of the relative sunspot numbers with bipolarity. *Acta Astron. Sin.* **29**, 297–304.

Yang, Z. (1988). The orbital motion of major planets and the long periodic change of the relative numbers of sunspot. *Acta Astrophysica Sin.* **8**, 294–300.

Zhang, H., Han, Y. and Zheng, D. (1982). Chandler Wobble: A result on analysis of observations. *Scientia Sin. Ser.A*, 837–846.

Zhang, H., Han, Y. and Li, Z. (1986). Checking of the double-frequency feature of Chandler main peak for different periods. *Acta Geopgys. Sin.* **29**, 16–27.

Zhao, M. and Zheng, D. (1980). On the discussion of secular polar motion. *Acta Astron. Sin.* **21**, 69–72.

Zhao, M. (1988). On the variation of parameters of Chandler Wobble. *Science in China Ser.A*, 741–748.

Zheng, D. (1978). An analysis of the short period terms in the universal time. *Acta Astron. Sin.* **19**, 103–108.

Zheng, D. and Zhao, M. (1979). Application of autoregressive technique to astronomy and geodynamics. *Acta Astron. Sin.* **20**, 301–307.

Zheng, D., Huang, H., Liao, D. and Luo, S. (1980). Improvement of the universal time series by means of AR series model. *Acta Astron. Sin.* **21**, 123–129.

Zheng, D. and Chen, Z. (1982). Prediction of Earth rotation parameters. *Ann. Shanghai Observatory* **4**, 116–120.

Zheng, D., Wang, S. and Gu, Z. (1984). Comparison between the Earth rotation parameters determined by different observational techniques. *Scientia Sin. Ser.A*, 292–302.

Zheng, D. and Dong, D. (1985). Research on the fine structure in polar motion with multi-stage filter. In *Proc. Internat. Conf. Earth Rotation and Terrestrial Reference Frame,* 55–66, Columbus, Ohio, USA.

Zheng, D. (1986a). Applications of time series analysis to astronomical geophysica. *Math. Statist. Appl. Probab.* **1**, 15–20.

Zheng, D., Luo, S. and Mackay, J. R. (1986b). An optimal identification of clock behaviour model for VLBI. *The Impact of VLBI on Astrophysics and Geophysics,* 525–528, Boston, USA.

Zheng, D. and Dong, D. (1986c). Realization of narrow band filtering of the polar motion data with multi-stage filter. *Acta Astron. Sin.* **27**, 368–376.

Zheng, D. and Dong, D. (1987). Marple algorithm of autoregressive spectrum estimate and its application to analysing astronomical data. *Acta Astron. Sin.* **28**, 364–373.

Zheng, D., Luo, S. and Song, G. (1988). Interannual variations of earth rotation, El Nino events and atmospheric angular momentum. *Science in China Ser.B*, 332–337.

Zheng, D., Song, G. and Luo, S. (1990a). Prediction of El Nino events from observations of the length of day. *Kexue Tongbao* **35**, 1244–1246.

Zheng, D., Song, G. and Luo, S. (1990b). El Nino prediction. *Nature* **348**, 119.

Shanghai Observatory, Chinese Academy of Sciences, 80 Nandam, Shanghai.
Shanghai Observatory, Chinese Academy of Sciences, 80 Nandam, Shanghai.

(Received February 1991; accepted January 1992)