J. Univ. Kuwait (Sci.) **17**, 1990

A threshold model with piece-wise non-linear dynamics for sunspot series

BASIL YOUNIS THANOON

Department of Mathematics and Statistics, College of Science, University of Mosul, Mosul, Iraq.

ABSTRACT

A model with piece-wise non-linear dynamics is proposed for the sunspot time series 1700–1920 as a modification of a model fitted by Tong & Lim (1980). The proposed model provides a description of the solar cycle similar to that of a different type of model proposed by Yoshimura (1979).

1. INTRODUCTION

In 1843 the German astronomer Samuel Heinrich Schwabe reported the discovery of the solar cycle. Six years later, Professor Rudolf Wolf of Zurich introduced the Wolf's sunspot number (or Wolf's relative index) which is an index to measure sunspot activity. It is defined as $R = k(10g + f)$, where g is the number of groups of sunspots, f is the total number of sunspots and k is a constant for the observatory where the observations are made. In the historical records, annual means of Wolf's sunspot numbers are available dating back to 1700.

The astrophysical explanation for the formation of sunspots, as well as their periodicity, is still subject to intensive research (see e.g. Parker (1979), Yoshimura (1979), Newkirk & Frazier (1982) and Gurman *et al.* (1982)).

Statistically, Wolf's sunspot time series is a much-analyzed set of data for which there is no completely satisfactory explanation (Morris 1977). The earliest linear model built for these data is probably due to Yule (1927) who introduced the class of linear autoregressive models as a result. Since then the literature on linear time series analysis of these data has been growing almost exponentially (Moran 1954; Box & Jenkins 1970; Akaike 1978). Non-linear time series modelling has entered the scene recently (Tong & Lim 1980; Gabr & Subba-Rao 1981).

2. SOME PRELIMINARIES

It is well known that the solar cycle has an approximate period of 11 years (Figs 1 & 2). The solar cycle might be viewed as an engine in which differential rotation and convection drive an oscillation between two (toroidal and poloidal) field geometries

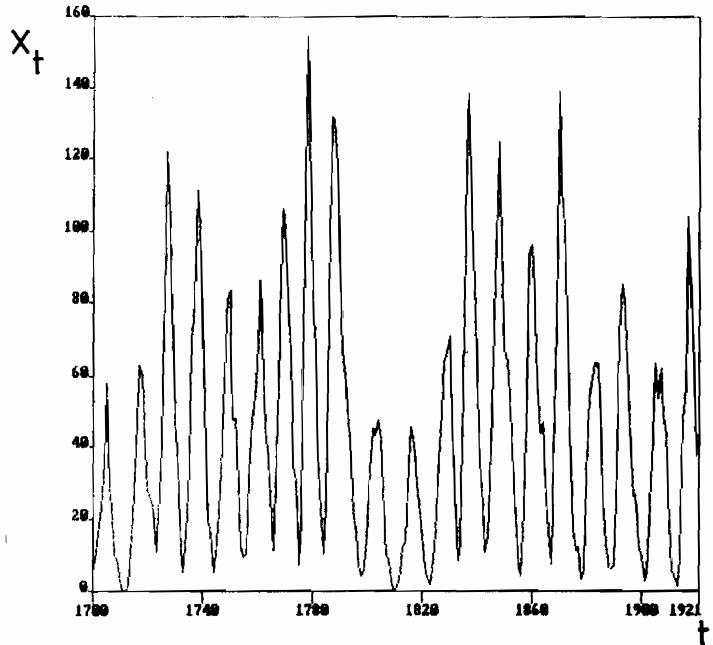

Fig. 1. Sunspot series 1700–1920.

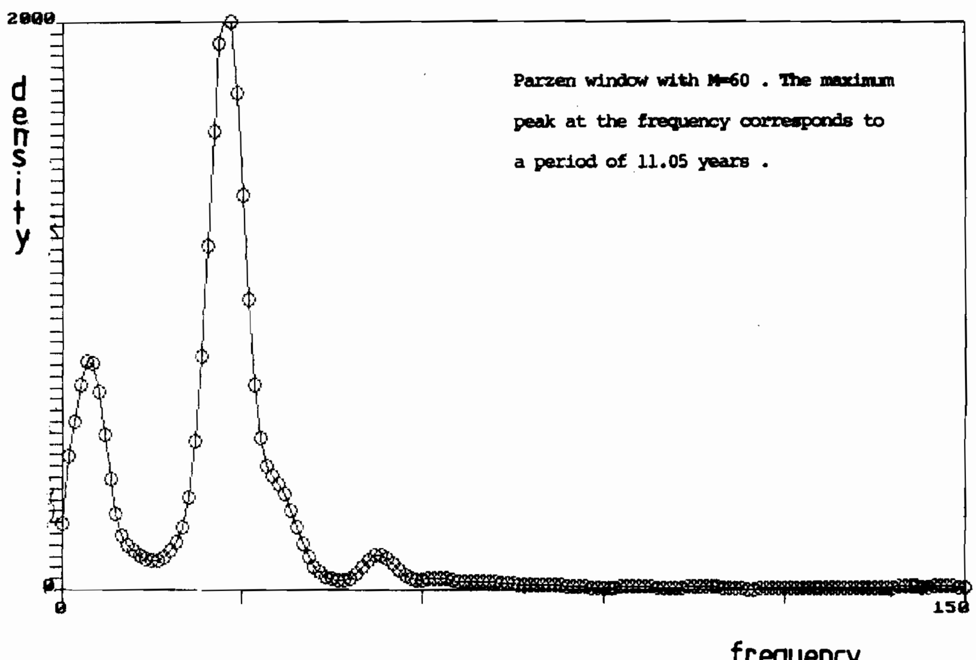

Fig. 2. Power spectral density function of sunspot series 1700–1920.

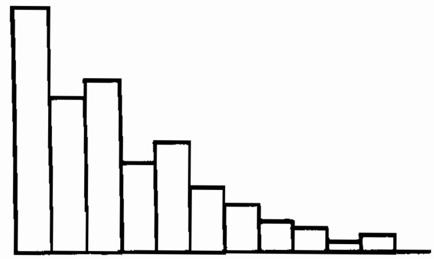

Fig. 3. Histogram of sunspot series 1700–1920.

and magnetic polarities (Newkirk & Frazier 1982). The dynamo theory is now the most accepted theory to explain the solar cycle (Yoshimura 1979; Newkirk & Frazier 1982). According to this theory, sunspots appear on the surface of the sun when the toroidal field intensity exceeds some threshold. Yoshimura (1979) represents the poloidal and toroidal fields by some non-linear equations which share the basic concepts of threshold and time delay with threshold autoregressive (TAR) models developed recently by Tong & Lim (1980) and Tong (1983).

It is natural, therefore, to represent the sunspot series by a TAR model. Let X_t denote Wolf's relative index at time t. Then at any instant t, X_t will correspond to one of the two states, say the state of energy absorption or the state of energy dissipation. Assuming that there is a well-defined variable J_t which can indicate these two states, and writing $J_t = 1$ for the absorption state and $J_t = 2$ for the dissipation state, X_t may be represented as

$$X_t = \begin{cases} h^{(1)}(X_{t-1}, X_{t-2}, \ldots) + e_t^{(1)} & \text{if } J_t = 1 \text{ (absorption)} \\ h^{(2)}(X_{t-1}, X_{t-2}, \ldots) + e_t^{(2)} & \text{if } J_t = 2 \text{ (dissipation)} \end{cases} \tag{2.1}$$

where $h^{(1)}$ and $h^{(2)}$ are some functions and $\{e_t^{(1)}\}$ and $\{e_t^{(2)}\}$ are uncorrelated white noise sequences. In Tong's approach $h^{(1)}$ and $h^{(2)}$ are usually approximated by linear functions. A more general approximation will be described in the next section.

First we list some of the main physical and statistical features of sunspot time series:

(i) The 11-year period might be regarded as a "fractional period" of a "grand-cycle" which consists of "local cycles" with varying amplitudes (see Newkirk & Frazier (1982) for an explanation of the irregularity of the cycle).

(ii) The variation in the heights of the peaks is much more than the variation in the troughs (Fig. 1).

(iii) There is a fairly clear evidence of irreversibility; the descent period, from a local maximum to the next local minimum, exceeds the ascent period, from a local minimum to the next local maximum.

(iv) The sunspot data are neither linear nor Gaussian nor stationary (Subba-Rao & Gabr 1980; Akaike 1978). Moreover, there is a point of "condensation" (i.e. a dominant peak near the origin in the histogram) which causes certain problems (Fig. 3).

(v) Some authors suggested transformations (e.g. the square root transformation) to this series. According to some studies, it seems almost impossible to reduce the data distribution to Gaussian (Hadid 1988). Hence, it is not surprising that linear models fail to offer a good statistical and physical description of this series. In fact, Hokstand (1983) emphasizes the need for a non-linear model.

3. A THRESHOLD AUTOREGRESSIVE MODEL WITH PIECE-WISE NON-LINEAR DYNAMICS

Let $\{X_t; t = 1, 2, \ldots, 221\}$ denote the yearly average of sunspot numbers in the year $1699 + t$. Gabr & Subba-Rao (1981) fitted a subset bilinear model, denoted by SBL, for this part of the series (Equation 5.3 of Gabr & Subba-Rao (1981)). Although this non-linear model represents a good fit and good predictors, it is difficult to explain the cyclical structure of the series through this class of models.

The dynamo theory tends to suggest fitting a model of the form (2.1). Tong & Lim (1980) fitted a TAR model of this form where $h^{(j)}$ approximated a linear function $L^{(j)}$ say, $(j = 1, 2)$ (Equation 9.2 of Tong & Lim (1980)). To distinguish this linear approximation from a non-linear one, we call the model of Tong & Lim (1980) the threshold autoregressive model with piece-wise linear dynamics and we denote it by TARL.

From the statistical point of view, the TARL model is not completely adequate, as mentioned by Tong (1983), since it gives non-Gaussian residuals (see the Appendix). Moreover, the variance of its residuals is 23% greater than that of the SBL model.

Define the indicator variable

$$J_t = \begin{cases} 1 & \text{if } X_{t-3} \leqslant 36{\cdot}6 \\ 2 & \text{if } X_{t-3} > 36{\cdot}6 \end{cases} \tag{3.1}$$

We can write the TARL model as $X_t = L^{(J_t)} + e_t^{(J_t)}$, where $L^{(1)}$ and $L^{(2)}$ denote linear functions of the past values of X_t and $\mathrm{Var}(e_t^{(1)}) = 252{\cdot}64$ and $\mathrm{Var}(e_t^{(2)}) = 66{\cdot}55$. Taking into account that for $J_t = 1$ the real data range over $[0, 36{\cdot}6]$ and for $J_t = 2$ range over $(36{\cdot}6, 145{\cdot}5]$, we discover immediately that the reason for inadequacy of this model is due to $L^{(1)}$ since $\mathrm{Var}(e_t^{(1)})$ is much bigger than the range of the data in this case ($\mathrm{Var}(e_t^{(1)}) \simeq 4\ \mathrm{Var}(e_t^{(2)})$). This means that the linear approximation of $h^{(1)}$ in (2.1), i.e. $L^{(1)}$, is not very satisfactory. This may be due to the fact that there is a high non-linear structure in the data corresponding to the state $J_t = 1$. Hence, a non-linear approximation is needed here. On the other hand, it is quite clear that the quality of linear approximation corresponding to $J_t = 2$ is satisfactory. Hence, a modification of the model will be restricted to the case $J_t = 1$ only.

Assume now that $h^{(1)}$ of (2.1) is analytic. Following Priestley (1980), the expansion of $h^{(1)}$ will lead to a "Volterra series" expansion, namely,

$$h^{(1)}(X_{t-1}, X_{t-2}, \ldots) = a_0^{(1)} + T_1^{(1)} + T_2^{(1)} + \cdots, \tag{3.2}$$

where $T_1^{(1)} = \sum_{u=1}^{\infty} a_u^{(1)} X_{t-u}$, $T_2^{(1)} = \sum_{u=1}^{\infty} \sum_{v=1}^{\infty} a_{uv}^{(1)} X_{t-u} X_{t-v}$, etc. and $a_0^{(1)}, a_u^{(1)}, a_{uv}^{(1)}$ are constants.

The linear approximation of (3.2) can be obtained by setting $T_i^{(1)} = 0$, $i = 2, 3, \ldots$. This approximation is provided by $L^{(1)}$ of the TARL model which estimated the order of $T_1^{(1)}$ at 4. To obtain a quadratic approximation, we set $T_i^{(1)} = 0$,

$i = 3, 4, \ldots$. Set the maximum order of $T_1^{(1)}$ and $T_2^{(1)}$ at 4, and let $Q^{(1)}$ denote the quadratic approximation of $h^{(1)}$ which is written as

$$Q^{(1)} = b_0^{(1)} + \phi_1^{(1)}(\mathbf{X}_{t-1})X_{t-1} + \phi_2^{(1)}(\mathbf{X}_{t-1})X_{t-2} + \phi_3^{(1)}(\mathbf{X}_{t-1})X_{t-3} + \phi_4^{(1)}(\mathbf{X}_{t-1})X_{t-4}, \tag{3.3}$$

where $b_0^{(1)}$ is a constant and $\phi_i^{(1)}(\mathbf{X}_{t-1})$ are some linear functions of X_{t-1}, X_{t-2}, X_{t-3}, X_{t-4}.

The AIC is now a well-known criterion for time series model selection (Akaike 1983). For a fitted model containing k parameters and residual variance $\hat{\sigma}^2$, the corresponding AIC value is defined as

$$\text{AIC} = n \log_e \hat{\sigma}^2 + 2k, \tag{3.4}$$

where n is the number of data points used in fitting the model. The normalized AIC, denoted by NAIC, is then defined as

$$\text{NAIC} = \text{AIC}/n. \tag{3.5}$$

In the sense of the AIC, the best model will be that which has the minimum NAIC value.

Using the AIC procedure, the estimated parameters (least squares) of (3.3) are listed below with the bracketed entries giving the approximate standard errors.

$$\hat{b}_0^{(1)} = \underset{(4\cdot981)}{13\cdot136}, \quad \hat{\phi}_1^{(1)}(\mathbf{X}_{t-1}) = \underset{(0\cdot148)}{1\cdot536},$$

$$\hat{\phi}_2^{(1)}(\mathbf{X}_{t-1}) = \underset{(0\cdot007)}{-0\cdot035}X_{t-1} + \underset{(0\cdot012)}{0\cdot052}X_{t-2},$$

$$\hat{\phi}_3^{(1)}(\mathbf{X}_{t-1}) = \underset{(0\cdot611)}{-1\cdot706} + \underset{(0\cdot015)}{0\cdot051}X_{t-1} - \underset{(0\cdot025)}{0\cdot122}X_{t-2} + \underset{(0\cdot018)}{0\cdot070}X_{t-3},$$

$$\hat{\phi}_4^{(1)}(\mathbf{X}_{t-1}) = \underset{(0\cdot006)}{0\cdot011}X_{t-1} + \underset{(0\cdot011)}{0\cdot016}X_{t-2}.$$

Recall that for the Tong & Lim linear approximation, namely,

$$L^{(1)} = c_0^{(1)} + c_1^{(1)}X_{t-1} + c_2^{(1)}X_{t-2} + c_3^{(1)}X_{t-3} + c_4^{(1)}X_{t-4},$$

the estimated parameters are

$$\hat{c}_0^{(1)} = \underset{(4\cdot04)}{10\cdot54}, \quad \hat{c}_1^{(1)} = \underset{(0\cdot12)}{1\cdot69}, \quad \hat{c}_2^{(1)} = \underset{(0\cdot25)}{-1\cdot16}, \quad \hat{c}_3^{(1)} = \underset{(0\cdot32)}{0\cdot24}, \quad \hat{c}_4^{(1)} = \underset{(0\cdot18)}{0\cdot15}$$

We find that $\hat{\phi}_1^{(1)}$ is similar to $\hat{c}_1^{(1)}$. This might be because the relationship between X_t and X_{t-1} is linear (the autocorrelation function of X_t and X_{t-1} is 0.8). The largest coefficients in $\hat{\phi}_i^{(1)}$'s are those of $\hat{\phi}_3^{(1)}$. In the linear approximation, when compared with $\hat{c}_1^{(1)}$ and $\hat{c}_2^{(1)}$, $\hat{c}_3^{(1)}$ is clearly much less significant. It seems that the non-linear relation between X_t and X_{t-3} is of a much more complicated nature, which is perhaps better approximated by $\hat{\phi}_3^{(1)}$. We would like to point out that a study of the regression functions, based on indices of linearity, has indicated similar

Table 1

	Model		
Statistic	SBL	TARL	TARNL
$\hat{\sigma}^2$	124·33	152·65	108·19
NAIC	4·93	4·99	4·82
MSE1	123·77	148·21	146·37

results (Thanoon 1984). Since higher order approximations of (3.2) will lead to a large number of parameters, we decided to stop at the quadratic approximation.

Now since the linear approximation of $h^{(2)}$ in (2.1) which is provided by the TARL model is satisfactory, we keep it as in Tong & Lim (1980). That is, the approximation of $h^{(2)}$ is

$$L^{(2)} = c_0^{(2)} + c_1^{(2)} X_{t-1} + c_2^{(2)} X_{t-2} + \cdots + c_{12}^{(2)} X_{t-12}. \tag{3.6}$$

Least-squares estimates of the parameters of (3.6) with their standard errors are listed below.

$$\underset{(3\cdot292)}{\hat{c}_0^{(2)} = 7\cdot804}, \quad \hat{c}_1^{(2)} = \underset{(0\cdot063)}{0\cdot743}, \quad \hat{c}_2^{(2)} = \underset{(0\cdot073)}{-0\cdot041}, \quad \hat{c}_3^{(2)} = \underset{(0\cdot077)}{-0\cdot202},$$

$$\hat{c}_4^{(2)} = \underset{(0\cdot082)}{0\cdot173}, \quad \hat{c}_5^{(2)} = \underset{(0\cdot085)}{-0\cdot227}, \quad \hat{c}_6^{(2)} = \underset{(0\cdot090)}{0\cdot019}, \quad \hat{c}_7^{(2)} = \underset{(0\cdot106)}{0\cdot161},$$

$$\hat{c}_8^{(2)} = \underset{(0\cdot124)}{-0\cdot256}, \quad \hat{c}_9^{(2)} = \underset{(0\cdot126)}{0\cdot320}, \quad \hat{c}_{10}^{(2)} = \underset{(0\cdot121)}{-0\cdot389}, \quad \hat{c}_{11}^{(2)} = \underset{(0\cdot101)}{0\cdot431},$$

$$\hat{c}_{12}^{(2)} = \underset{(0\cdot063)}{-0\cdot040}.$$

Hence, the proposed model is

$$X_t = \begin{cases} Q^{(1)} + Z_t^{(1)} & \text{if } X_{t-3} \leqslant 36\cdot6 \\ L^{(2)} + Z_t^{(2)} & \text{if } X_{t-3} > 36\cdot6 \end{cases} \tag{3.7}$$

where $Q^{(1)}$ is given by (3.3) and $L^{(2)}$ is given by (3.6), $\text{Var}(Z_t^{(1)}) = 156\cdot4$ and $\text{Var}(Z_t^{(2)}) = 66\cdot55$. We call the above model the threshold autoregressive model with piece-wise non-linear dynamics, and we denote it by TARNL.

Comparing $\text{Var}(Z_t^{(1)})$ of (3.7) with $\text{Var}(e_t^{(1)})$ of the TARL model we note that $\text{Var}(Z_t^{(1)})$ is 38% less than $\text{Var}(e_t^{(1)})$. Of course $Z_t^{(2)} \equiv e_t^{(2)}$. A comparison between pooled variances of the fitted residuals, $\hat{\sigma}^2$, and the NAIC values obtained from the SBL model, the TARL model and the TARNL model is provided in Table 1. Also, the mean square errors of the one-step-ahead forecast errors (MSE1) for the period 1921–1955 are given in the table.

It is quite clear that the fitting is greatly improved by this non-linear approximation. Moreover, the fitted residuals of the TARNL model are now accepted as Gaussian white noise (see the Appendix).

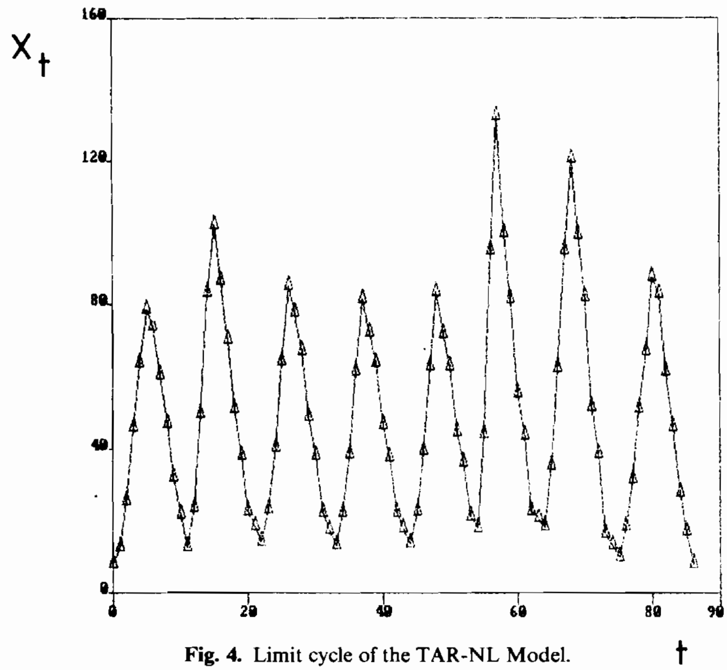

Fig. 4. Limit cycle of the TAR-NL Model.

4. CONCLUSION AND DISCUSSION

We have proposed a non-linear time series model for the sunspot series 1700–1920. The model is of a piece-wise non-linear structure and provides better fitting to the data than a model with piece-wise linear structure proposed by Tong & Lim (1980). We have modified the model of Tong & Lim (1980) in order to interpret the cyclical structure of the series rather than prediction. The increased complexity of the modified model makes more than one-step-ahead prediction almost impossible.

The proposed model shares the basic concepts of threshold and time delay with the "dynamo" model of Yoshimura (1979). Fig. 4 shows the systematic part of the proposed TARNL model with the white noise suppressed and with the last few observations as initial values of the recursion. This is known as the eventual forecasting function, by an abuse of a terminology used by Box & Jenkins (1970). The eventual forecasting function of the TARNL model gives a stable limit cycle of period 86 years which consists of 8 subcycles, the average period being 86/8 = 10·75 years. The average ascent period is 4·125 and the average decent period is 6·625 years. Clearly, this limit cycle is of amplitude modulation type and the variation in the maxima and minima is similar to that of the observed data (see Figs 1 and 4). This seems to agree with the dynamo equations of Yoshimura (1979). Results of his approach suggest that the eventual forecasting function consists of "grand cycles" with each "grand cycle" consisting of subcycles.

ACKNOWLEDGEMENT

The author is very grateful to Professor H. Tong and the referees for their valuable comments which led to an improved version of the paper.

REFERENCES

Akaike, H. 1978. On the likelihood of a time series. The Statistician **27**: 217–35.

Akaike, H. 1983. Information measures and model selection. Proceedings of the 44th Session of ISI (Madrid, Spain) **1**: 277–90.

Box, G.E.P. & Jenkins, G.M. 1970. Time series analysis forecasting and control. Holden-Day, San Francisco.

Gabr, M.M. & Subba-Rao, T. 1981. The estimation and prediction of subset bilinear time series models with applications. Journal of Time Series Analysis **2**: 155–71.

Gurman, J.B., Leibacher, J.W., Shine, R.A., Woodgate, B.E. & Henze, W. 1982. Transition region oscillations in sunspots. The Astrophysical Journal **253**: 939–48.

Hadid, S.K. 1988. Analysis and modelling of sunspot time series. Unpublished M.Sc. thesis, University of Mosul, Iraq.

Hokstand, P. 1983. A method for diagnostic checking of time series models. Journal of Time Series Analysis **4**: 177–83.

Lin, C.C. & Mudholkar, G.G. 1980. A simple test for normality against asymmetric alternatives. Biometrika **67**: 455–61.

Moran, P.A.P. 1954. Some experiments on the prediction of sunspot numbers. Journal of the Royal Statistical Society, Series B, **16**: 112–17.

Morris, M.J. 1977. Forecasting the sunspot cycle. Journal of the Royal Statistical Society, Series A, **140**: 437–68.

Newkirk, G. & Frazier, K. 1982. The solar cycle. Physics Today, April: 25–34.

Parker, E.N. 1979. Cosmic magnetic fields: their origin and their activity. Clarendon Press, Oxford.

Priestley, M.B. 1980. State dependent models: a general approach to non-linear time series analysis. Journal of Time Series Analysis **1**: 47–72.

Subba-Rao, T. & Gabr, M.M. 1980. A test of linearity of stationary time series. Journal of Time Series Analysis **1**: 145–58.

Thanoon, B.Y. 1984. A study of threshold autoregressive models with some applications. Unpublished Ph.D. Thesis, Manchester University, U.K.

Tong, H. 1983. Threshold models in non-linear time series analysis. Lecture Notes in Statistics, No. 21. Springer Verlag.

Tong, H. & Lim, K.S. 1980. Threshold autoregression, limit cycles and cyclical data. Journal of the Royal Statistical Society, Series B, **42**: 245–92.

Yoshimura, H. 1979. The solar-cycle period-amplitude relation as evidence of hypothesis of the solar-cycle non-linear magnetic oscillation and the long term (55) cycle modulation. Astrophysical Journal **227**: 1047–58.

Yule, G.U. 1927. On a method of investigating periodicities in disturbed series with special reference to Wolf's sunspot numbers. Philosophical Transactions of the Royal Society of London, Series A, **226**: 267–98.

(*Received 19 September 1988, revised 10 April 1989*)

APPENDIX

Let $e_1, e_2, \ldots, e_n$ denote the fitted residuals and let r_k denote the sample autocorrelation function of e_t and e_{t-k}.

The hypothesis of white noise was tested by looking to r_k's; $k = 1, 2, \ldots, 20$. For the TARL model it is found that r_{13} and r_{17} lie outside the band $\pm 1{\cdot}96/\sqrt{n}$, while only r_{12} of e_t's of the TARNL model lies outside this band.

The fitted residuals of the TARNL model are shown in Fig. A1 and a histogram of these residuals is shown in Fig. A2. The hypothesis of normality of these e_t's was tested by using the test developed by Lin & Mudolkar (1980), where under the hypothesis of normality Z is $N(0, 1)$. Estimates of the skewness coefficient g_1 and the

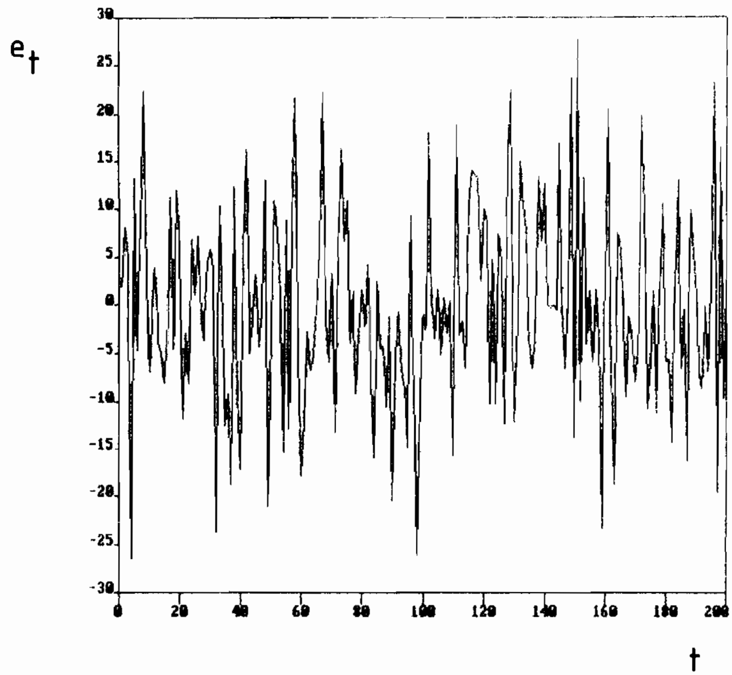

Fig. A1. Residuals of the TARNL Model.

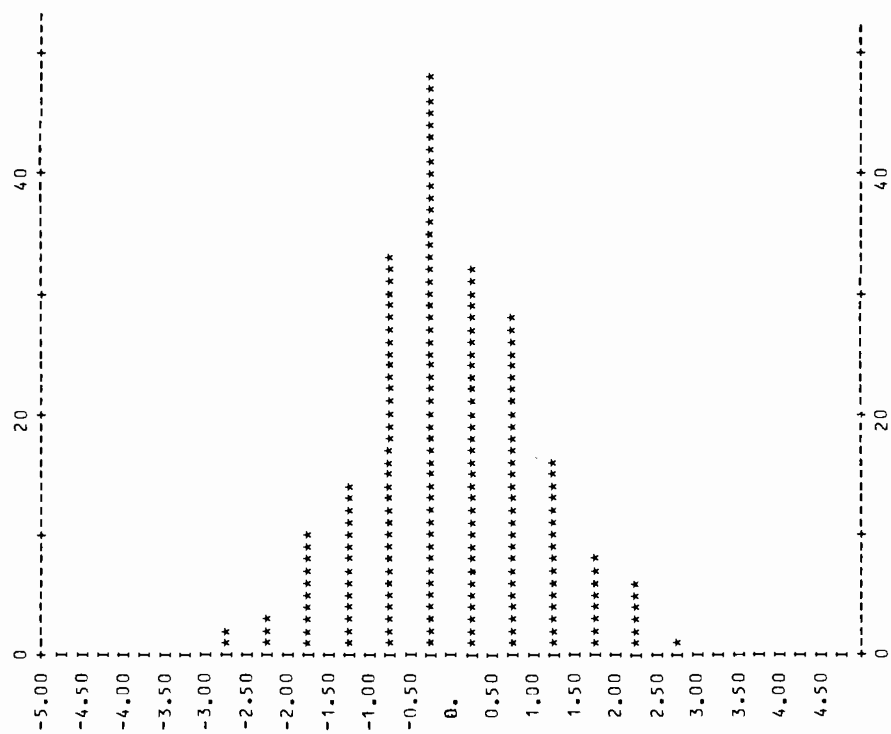

Fig. A2. Histogram of the residuals of the TARNL Model.

Table A. Some of the relevant statistics of the fitted residuals

	Model	
Statistic	TARL	TARNL
g_1	0·67	0·14
g_2	2·35	−0·06
Z	−2·72	−0·82

kurtosis coefficient g_2 of e_t's are shown in Table A with the Z values obtained. (For normal distributions, $g_1 = g_2 = 0$, and the 5% tailed points of $N(0, 1)$ are $\pm 1{\cdot}96$.)

Clearly e_t's of the TARL model cannot be accepted as normal white noise while it seems reasonable to accept e_t's of the TARNL model as normal white noise.

نموذج عتبة ذو ديناميكية بشكل قطع غير خطية للسلسلة الزمنية للبقع الشمسية

باسل يونس ذنون
قسم الرياضيات والاحصاء بكلية العلوم ، جامعة الموصل ،
الموصل ، العراق

خلاصة

في هذا البحث نقترح نموذجا ذا ديناميكية بشكل قطع غير خطية للسلسلة الزمنية للبقع الشمسية ١٧٠٠-١٩٢٠ ، كتطوير لنموذج آخر كان قد وضعه الباحثان تونج وَلِيم (١٩٨٠) . إن النموذج المقترح يزودنا بوصف للدورة الشمسية مشابه لذلك الذي يعطيه نموذج رياضي آخر سبق أن اقترحه الباحث يوشيمورا (١٩٧٩) .