A Method of the Variable Transformation for the Time Series Analysis

A Method of the Variable Transformation for the Time Series Analysis

Ken'ichi MORI* and Toshirou KUROZAWA**

(Recieved June 15, 1985)

The transformation of time series is considered for the minimization of the fitted ARIMA model. An AIC criterion is developed for estimating the optimum model size, transformation parameter λ and ARIMA model parameters, jointly. This criterion evaluates the badness of fitted model and consists of twice of sum of a negative likelihood and number of model parameters. Use of the AIC criterion is illustrated in analysing the sunspot number and the seasonal bank deposits time series.

1. Introduction

ARMA (Autoregressive Moving Average) model has been widely used as the basic model for time series forecasting, which explains the correlation between serial error terms. In addition, Box and Jenkins[1] has proposed a little more extended ARIMA model which resolves some kind of the nonstationarity by differencing the original time series data.

One of the problems in estimation of time series model is the derivation of minimum sufficient order model to describe the treated data. This is sometimes called the principle of parsimony. In this study, an application of a method of transformation of data is considered to attain the minimum size model. The power transformation method proposed by Box and Cox[2] is used to this object. This one contains the wide variety of transformation class, including the square and the square root transformation and is porved to be efficient for the transformation of the object variable in regression analysis[2]. Ansley and others[3] has studied its application to the time series analysis. But they used the maximization of the likelihood function as the evaluation criterion and didn't take the size of model highly into account.

On the other hand, the estimation criterion considering both size and fitting has been proposed by Akaike[4),6]. This one is called AIC and consisted of the sum of likelihood and number of parameters. This criterion has been shown efficient in many types of estimation problems. We propose a model estimation method based on AIC to get the transformed time series model with minimum sufficient order. Using this method, two examples from real situations are estimated to show the efficiency.

2. The Formulation of The Time Series Transformation

Let y_t be a time series value at time t, its Box-Cox power transformation $z_t^{(\lambda)}$ is defined as

$$z_t^{(\lambda)} = \begin{cases} (y_t^{\lambda} - 1) / \lambda & (\lambda \neq 0) \\ ln y_t & (\lambda = 0) \end{cases}$$

* Department of Industrial Engineering, College of Engineering.
** Department of Industrial Engineering, College of Engineering, Setsunan University.

where λ is power transformation parameter and some noted cases are non-transformation ($\lambda = 1$), a square root ($\lambda = 0.5$) and a reciprocal ($\lambda = -1$). ARIMA (p, d, q) model is given as

$$\phi(B) w_t = \theta(B) a_t \tag{1}$$

where B is the backward shift operator defined by $By_t = y_{t-1}$, w_t the dth differenced series

$$w_t = (1 - B)^d z_t^{(\lambda)} = \nabla^d z_t^{(\lambda)}$$

and a_t is a sequence of mutually independent identically distributed random variable following $N(0, \sigma_a^2)$. Also, in Eq. (1), $\phi(B)$ and $\theta(B)$ are autoregressive and moving average operators defined as

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots\cdots - \phi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots\cdots - \theta_q B^q$$

respectively. For this model, the problem is reduced to the estimation of transformed ARIMA model which minimize AIC criterion.

The LF (Likelihood Function) for the transformed ARIMA model is

$$L_z = (2\pi\sigma_a^2)^{-n/2} \left| M_n \right|^{1/2} \exp(-S/2\sigma_a^2)$$

where

$$S = S(\boldsymbol{\phi}, \boldsymbol{\theta}, \lambda, \sigma_a^2)$$

$$= \sum_{t=-\infty}^{\infty} E[a_t \mid z^{(\lambda)}, \boldsymbol{\phi}, \boldsymbol{\theta}, \lambda]^2 \tag{2}$$

is an estimation error sum of square and $\sigma_a^2 M_n^{-1}$ a variance-covariance matrix of $z_t^{(\lambda)}$. The LF for original series y_t, (t = 1, 2, ..., n) is defined as

$$L_y = (2\pi\sigma_a^2)^{-2/n} \left| M_n \right|^{1/2} \exp(-S/2\sigma_a^2) J$$

where J is a Jacobian of transformation estimated as

$$J = \prod_{t=1}^{n} \left| \frac{\partial z_t^{(\lambda)}}{\partial y_t} \right| = \prod_{t=1}^{n} y_t^{\lambda-1} \quad .$$

The logarithmic LF is also defined as

$$L = -\frac{n}{2} \ln 2 - \frac{n}{2} \ln \sigma_a^2 + \frac{1}{2} \ln \left| M_n \right| - S/2\sigma_a^2 + \ln J \,. \tag{3}$$

Generally, $\frac{1}{2} \ln \left| M_n \right|$ is so small compared with $S/2\sigma_a^2$ that this term can be neglected from Eq. (3) and LF is reduced to

$$L = const. -\frac{n}{2} ln\sigma_a^2 - S/2\sigma_a^2 + ln\, J\,. \tag{4}$$

From Eq. (4), an estimate of σ_a^2 is derived as ML (Maximum Likelihood) solution S/n, and LF becomes

$$L = const. -\frac{n}{2} ln\, S + ln\, J\,. \tag{5}$$

Considering the ML value $L\,(\hat{\boldsymbol{\beta}})$, AIC criterion is given by

$$AIC = -2ln\, L(\hat{\boldsymbol{\beta}}) + 2m \tag{6}$$

where $\hat{\boldsymbol{\beta}}$ is ML solution of the model parameters $\boldsymbol{\beta}$ and m the number of parameters.[6] As is usual, this AIC gives the badness of the estimated model and value of AIC decrease as LF become larger and m smaller. So, AIC has to be minimized to estimate the best model.

For this formulation, $\boldsymbol{\beta} = (\boldsymbol{\phi}, \boldsymbol{\theta}, \lambda, \sigma_a^2)$ and substituting Eq. (7) into Eq. (6), the AIC of transformed ARIMA (p, d, q) model becomes

$$\begin{aligned} \mathrm{AIC} &= n\, ln\, S - 2ln\, J + 2\,(p+q+1) \\ &= n\, ln\,(S/J^{\frac{n}{2}}) + 2\,(p+q+1) \end{aligned} \tag{7}$$

discarding the constant terms. Finally, the AIC optimum model estimation problem is reduced to the simultaneous estimation of its order and parameters $\boldsymbol{\beta}$.

3. Estimation of The Actual Time Series

By substituting $a_t = \theta^{-1}(B)\phi(B)w_t$ from Eq. (1) into Eq. (2), sum of square of estimation error becomes

$$S\,(\boldsymbol{\phi}, \boldsymbol{\theta}\,|\lambda) = \sum_{t=-\infty}^{n} [\theta^{-1}(B)\phi(B)w_t\,|z^{(\lambda)}, \boldsymbol{\phi}, \boldsymbol{\theta}, \lambda]^2. \tag{8}$$

A procedure for the estimation is given in the next three steps:

Step 1) Identify the possible models by autocorrelation and partial autocorrelation functions;

Step 2) Estimate the parameters $(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}, \hat{\lambda})$ which minimize Eq. (8), and the corresponding AIC values.;

Step 3) Repeat the procedure of step 2 for all possible models and select the best one which minimizes AIC criterion.

3.1. Estimation for AR(p) model

As a special case of ARIMA model, the estimation of AR(p) model is considered. Because of the linear relation among parameters, the estimates are obtained easily by least square method. For a given λ, the sum of square of estimation error is defined as

$$S(\boldsymbol{\phi} \mid \lambda) = \sum_{t=p+1}^{n} [w_t - \phi_0 - \phi_1 w_{t-1} - \cdots\cdots - \phi_p w_{t-p}]^2$$

by putting $\theta(B) = 1$ in Eq. (8). The estimates $\hat{\boldsymbol{\phi}}$ which minimize this $S(\boldsymbol{\phi}|\lambda)$ are obtained as

$$\hat{\boldsymbol{\phi}} = (\boldsymbol{w}'\boldsymbol{w})^{-1} \boldsymbol{w}' A$$

where $\boldsymbol{\phi}$, $\boldsymbol{w}$, A are respectively

$$\boldsymbol{\phi} = \begin{bmatrix} \phi_0 \\ \phi_1 \\ \cdot \\ \cdot \\ \cdot \\ \phi_p \end{bmatrix}, \quad \boldsymbol{w} = \begin{bmatrix} 1 & w_p & & w_1 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 1 & w_{n-1} & & w_{n-p} \end{bmatrix}, \quad A = \begin{bmatrix} w_{p+1} \\ \cdot \\ \cdot \\ \cdot \\ w_n \end{bmatrix}$$

As $\hat{\boldsymbol{\phi}}$ depend on λ through $\boldsymbol{w}$, LF can be considered to depend only on λ and p. So AIC is given as

$$\text{AIC}\,(\lambda) = n\, ln\, [S\,(\hat{\boldsymbol{\phi}} \mid \lambda) \,/\, J^{2/n}] + 2\,(p+1)\ . \tag{9}$$

Thus, the AIC optimum transformed AR model is estimated from Eq. (9) by minimizing it about λ and p. Actually, this estimation is made by numerical method and following Steps 1) to 3).

For one example of this estimation, the time series of sun spot number from 1770 to 1869 shown in Fig. 1 is fitted to the transformed AR(p) model. In non-transformation case ($\lambda = 1$), this series is identified as AR(2) or AR(3) model. AIC criterion is estimated as in Table 1 for these models, and AR(3) is selected to be a little bit better than AR(2). The estimated equation is

Table 1. AIC for AR(p) models

P	Error Var.	AIC
2	227.500	548.715
3	219.314	547.050

$$y_t = 13.322 + 1.553 y_{t-1} - 1.007 y_{t-2} + 0.208 y_{t-3} + a_t\ . \tag{10}$$

For transformed case, we tentatively assume the AR(1) – AR(5) models as the candidates for the optimum one. The estimates of λ's and corresponding AIC's are listed in Table 2. From this result, the optimal model is estimated as AR(2) which minimizes AIC and gives $\lambda = 0.43$. In this case, the estimated model is

$$z_t^{(\lambda)} = 2.962 + 1.419 z_{t-1}^{(\lambda)} - 0.708 z_{t-2}^{(\lambda)} + a_t\ . \tag{11}$$

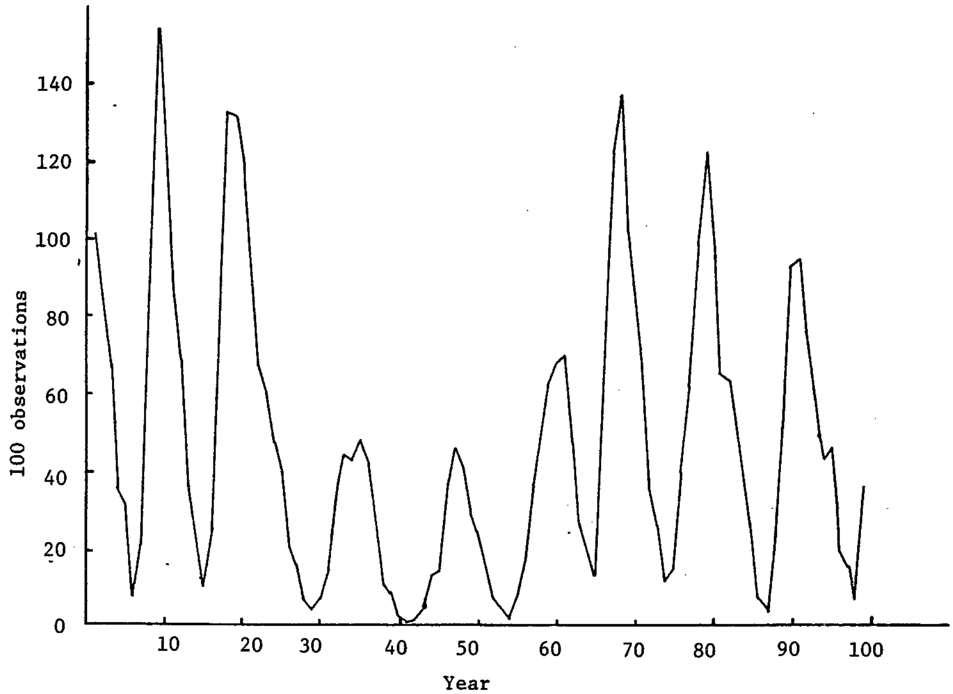

Fig. 1. Yearly sun spot numbers data

Table 2. AIC of transformation for AR(p)

p	λ	L(λ)	AIC
1	0.23	−288.550	583.100
2	0.43	−255.821	519.642
3	0.47	−255.224	520.449
4	0.48	−255.061	522.122
5	0.48	−254.888	523.775

This model of Eq. (11) is shown to give not only minimum AIC value but also better satisfy the conditions assumed for a_t, i.e., i) independency, ii) normality and iii) equi-variance.

When we use the LF criterion, the higher order model (with larger p) gives better fit to data as is shown in Table 2. But the higher model brings about more complexity and difficulty in the estimation and implementation.

3.2. Seasonality model

The another time series example for model estimation is the bank deposit in West Germany[5] illustrated in Fig. 2. This series consists of 72 terms which are taken quarterly from 1957 to 1974 and has the steep trend and seasonality of 1 year cycle. The seasonal ARIMA model is called ARIMA (p, d, q) X (P, D, Q)$_s$ and defined as

$$\phi(B)\Phi(B^s)\nabla^d\nabla_s^D z_t^{(\lambda)} = \theta(B)\Theta(B^s)\, a_t$$

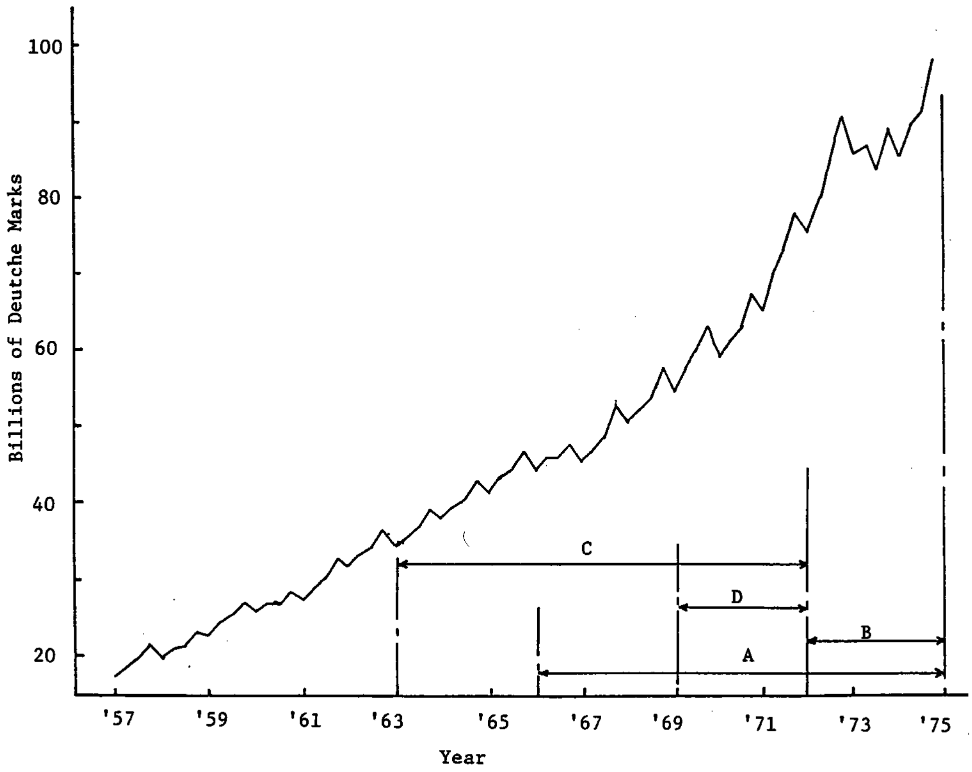

Fig. 2. Saving deposit data in West Germany

where $\Phi\ (B^s)$ and $\Theta\ (B^s)$ are respectively Pth order autoregressive and Qth order moving average model associated with seasonality of cycle length s defined as

$$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_P B^{Ps}$$

$$\Theta(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs} .$$

s takes 12 for monthly data, 4 for quarterly and so on. For time series of Fig. 2, the possible models are IMA (0, 1, 1) X $(0, 1, 1)_4$ and IMA (0, 1, 1) X $(0, 2, 1)_4$ which do not have the autoregressive term. These models are written respectively as

$$(1 - B)(1 - B^4) z_t^{(\lambda)} = (1 - \theta B)(1 - \Theta B^4) a_t$$

$$(1 - B)(1 - B^4)^2 z_t^{(\lambda)} = (1 - \theta B)(1 - \Theta B^4) a_t .$$

So, the three parameters θ, Θ, λ are needed to be estimated. The data sets used for the estimation are classified as follows;
(A) the latter 36 of 72 terms, (B) the latest 12 of 72 terms, (C) the latter 36 of 64 terms from 1957 to 1972 and (D) the latest 12 of 64 terms. The estimation results are given in Fig. 3 and 4. For IMA (0, 1, 1) X $(0, 1, 1)_4$ model, we get the estimates $\theta = 0.6$ and $\Theta = -0.1$. But the optimal estimates of the transformation parameter λ is dif-

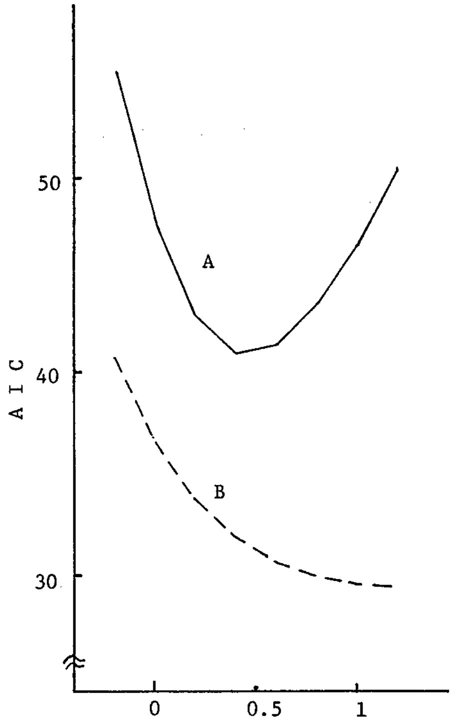

Fig. 3. Estimation in 72 series

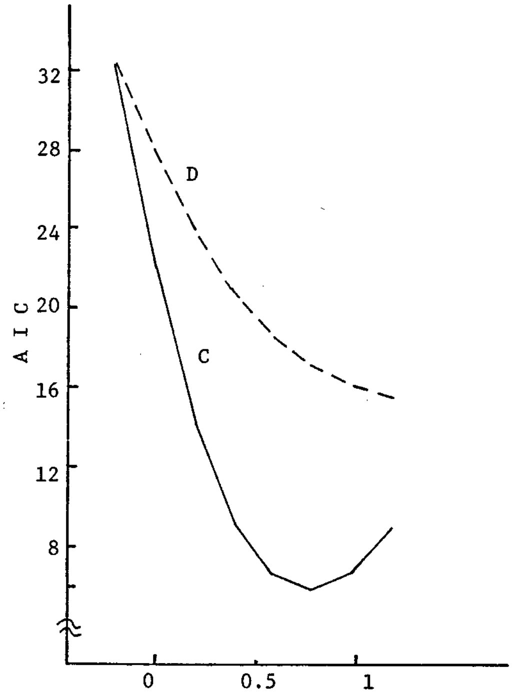

Fig. 4. Estimation in 64 series

ferent for each data which gives the minimum AIC. The optimal values of λ are 0.4 for A series, 1 for B, 0.8 for C and 1.2 for D. This fact shows that the difference in estimation time point could give the different parameter estimates even for the same time series.

The estimation for the IMA (0, 1, 2) X (0, 1, 1)$_4$ model is summerized in Fig. 5. The estimates for this series is $\theta = 0.75$, $\Theta = -0.1$, but the transformation parameter λ is estimated as different for A and C time series, that is, $\hat{\lambda} = 0$, log transformation for series A and $\hat{\lambda} = 1$, non transformation for series C.

The relation between AIC and transformation parameter λ is given in Fig. 6 for fitting of non seasonal IMA (0, 2, 2) model

$$(1 - B)^2 z_t^{(\lambda)} = (1 - \theta_1 B - \theta_2 B^2)\, a_t$$

to this series. For this model, λ = 1 is shown to be the optimal one and estimates of θ_i's are $\theta_1 = 0.8$ and $\theta_2 = 0.01$. But, AIC of this estimate is bigger than that of the seasonable model for any λ. This fact shows that this model is not so well fitted for this time series.

The estimation results reveal that IMA (0, 1, 1) X (0, 1, 1)$_4$ model gives the best fit to this time series from AIC point of view.

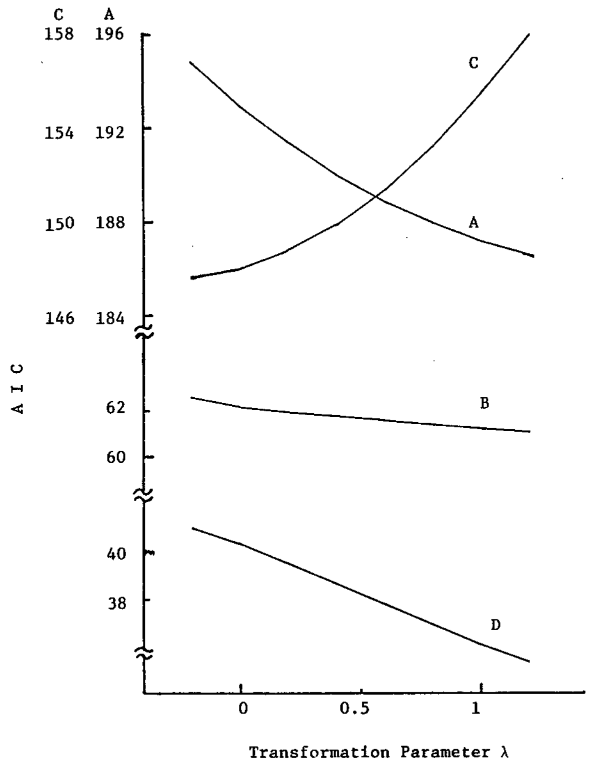

Fig. 5.
Fitting to IMA (0, 1, 1) × (0, 2, 1)$_4$

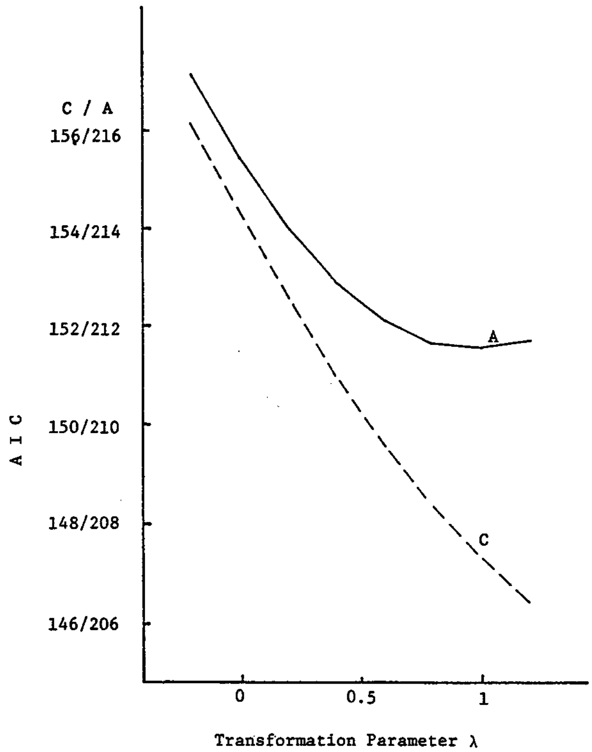

Fig. 6.
Estimation by IMA (0, 2, 2) model

4. Conclusion

The estimation based on LF or least square method can determine the best fitted model, when the order of model is given. In these methods, the higher order model gives the better fit. But the smaller model is desirable for the practical use. These requests contradict each other.

This contradiction is resolved by AIC criterion and data transformation. Concretely, AIC consists of linear combination of LF and order of model, and presents the badness of model to be minimized. In this paper, we propose an AIC method which also take the transformation into account. The proposed estimation method is proved to provide good results in the estimation of two practical time series.

References

1) G.E.P. Box and G.M. Jenkins, "Time Series Analysis — forecasting and control", Holden-Day (1970).
2) G.E.P. Box and D.R. Cox, J. Roy. Statistical Society B, 26 (2), 211 (1977).
3) C.S. Ansley and W.A. Spivey, Appl. Stat., 26 (2), 173 (1977).
4) H. Akaike, 2nd Int. Symp. on Information Theory, Eds. B.N. Petrov and F. Csâki, 267 (1973).
5) M.N. Bhattacharyya, "Comparison of Box-Jenkins and Bonn Monetary Model Prediction Performance", Springer Varlag (1980).
6) H. Akaike, Suurikagaku (in Japanese), No 153, 5 (1976).
7) K. Takeuchi, Suurikagaku (in Japanese), No. 153, 12 (1976).