

31,
202

TIME SERIES STOCHASTIC PROCESS
AND FORECASTING

by

TONY LEE-CHUIN CHIEN

Diploma, Tamkang College of Arts and Sciences, 1965

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

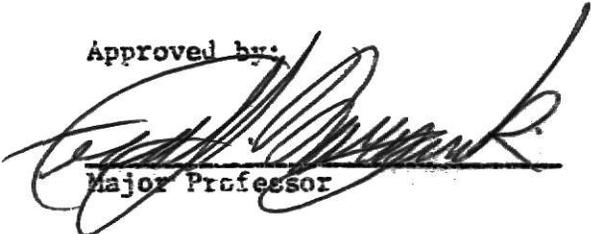
MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1979

Approved by:


Major Professor

Spec. Coll.
LD
2668
.R4
1979
C516
C.2

To my parents and my wife Nancy
for
their love and patience

ILLEGIBLE

**THE FOLLOWING
DOCUMENT (S) IS
ILLEGIBLE DUE
TO THE
PRINTING ON
THE ORIGINAL
BEING CUT OFF**

ILLEGIBLE

ACKNOWLEDGEMENTS

I wish to take this space to acknowledge my major professor and the members of my committee. I extend to Dr. Fred J. Maryanski my deepest appreciation for his guidance, inspiration and patience to make this report possible to complete. A special thanks is extended to Dr. Nasir Ahmed who is responsible for providing me with this topic in the early stage. For their supportive efforts I extend my thanks to Dr. S. K. Perng and Dr. E. Unger.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.	iii
LIST OF FIGURES	vii
LIST OF TABLES.	viii
CHAPTER	
I. FUNDAMENTAL CONCEPTS IN TIME SERIES AND STOCHASTIC PROCESS. .	1
1.1 Introduction	1
1.2 The Concept of Stationarity.	2
1.3 Autocorrelation.	3
1.4 Stochastic Process	5
II. STOCHASTIC MODELS AND THEIR FORECASTING	7
2.1 Linear Stationary Model.	7
2.1.1 Linear Stochastic Process	8
2.1.2 Autoregressive Model.	12
2.1.3 Moving Average Model.	20
2.1.4 Mixed Autoregressive-Moving Average Process .	22
2.1.5 Summary	27
2.2 Linear Nonstationary Models.	29
2.2.1 Difference and Homogeneous Nonstationarity. .	29
2.2.2 The General Form of the ARIMA Process.. . . .	31
2.2.3 Three Explicit Forms for the ARIMA Model. . .	32
III. STOCHASTIC MODEL BUILDING	36
3.1 Introduction	36
3.2 Identification	39
3.3 Estimation	42
3.4 Diagnostic Checking.	43

CHAPTER	Page
3.4.1 Autocorrelation check.	44
3.4.2 A Portmanteau Lack of Fit Test	45
3.5 Forecasting	45
3.5.1 Minimum Mean-Square Error Forecast	46
3.5.2 Calculation of the ψ Weights	51
3.5.3 Calculation of the Probability Limit of the Forecasts	53
IV. EMPIRICAL APPLICATIONS	55
4.1 Gross National Product (G.M.P.)	56
4.2 Dow Jones Utility Index (8/28/72 - 12/18/72).	74
4.3 Wolfer Sunspot Number (1770 - 1860)	120
4.4 Summary of Result	141
V. CONCLUSION	143
REFERENCES	145
APPENDIX A	146
A.1 Univariate Stochastic Model Identification - MAIN1.	147
A.2 Univariate Stochastic Model Preliminary Estimation - MAIN2	148
A.3 Program Descriptions.	149
A.3.1 Routine MAIN1	149
A.3.2 Subroutine BJDIF.	152
A.3.3 Subroutine BJSTAT.	153
A.3.4 Subroutine BJRES	154
A.3.5 Subroutine BJPLOT.	155
A.3.6 Routine MAIN2 - Preliminary Estimation	156
A.3.7 Subroutine BJARMA.	159

	Page
A.3.8 Subroutine BJMASR.	160
A.4 Program Listing.	161
A.4.1 Identification - MAIN1	162
A.4.2 BJDIFF	163
A.4.3 BJSTAT	165
A.4.4 BJRES.	168
A.4.5 BJPLOT	169
A.4.6 Estimation - MAIN2	170
A.4.7 BJMASR	171
A.4.8 BJARMA	173
APPENDIX B.	174
B.1 Table relating ρ_1 to θ for a first-order moving average process	175
B.2 Chart relating ρ_1 and ρ_2 to ϕ_1 and ϕ_2 for a second- order autoregressive process.	176
B.3 Chart relating ρ_1 and ρ_2 to θ_1 and θ_2 for a second- order moving average process.	177
B.4 Chart relating ρ_1 and ρ_2 to ϕ and θ for a mixed first- order autoregressive-moving average process	178

LIST OF FIGURES

Figure	Page
2.1. Realizations from first-order autoregressive processes and their corresponding theoretical autocorrelation functions.	19
2.2. Autocorrelation and partial autocorrelation functions ρ_k and ϕ_{kk} for various ARMA(1,1) models.	26
2.3. Summary of autocorrelation and partial autocorrelation functions for pure AR and MA processes	27
3.1. Stages in the iterative approach to model building	38

LIST OF TABLES

Table	Page
1. Summary of properties of autoregressive, moving-average, and mixed ARMA processes.	28
2. Admissible region and initial estimates for various process . .	41
3. Approximate variance for parameter estimates of different models.	43
4. Summary of results of Portmanteau test applied to residuals of various model fitted to the series in the report	142

CHAPTER I

FUNDAMENTAL CONCEPTS IN TIME SERIES AND STOCHASTIC PROCESS

1.1 Introduction

A time series is a set of observations, discrete or continuous, generated sequentially in time. They are observed and analyzed with respect to the statistics of their relation in time. They may be simple or multiple.

An example of a simple, discrete time series is the closing price of IBM on the New York Stock Exchange, tabulated daily. A simple continuous time series would be daily temperature readings. An example of a multiple, discrete time series is the closing price of all stocks on the New York Stock Exchange, tabulated daily. A multiple, continuous time series might be an infinite number of slightly different radar frequencies.

Many forecasting and data analysis techniques are available, through the exponential smoothing to the spectral analysis and general autoregressive integrated moving average stochastic process method. The specific problem under consideration here is the application of the general autoregressive integrated moving average stochastic method to simulate the G.N.P. (1939-1976), Dow Jones Utility Index (8/28/72-12/18/72) and Wölfer Sunspot Numbers (1770-1860).

The technique of parameter driven stochastic modeling has been developed to represent and forecasting many practical time series, in a business situation, in inventory control, in a scientific phenomenon, in an industrial production process.

The general ARIMA process of parametric stochastic modeling can include most possible models. In practice, it is frequently true that an adequate

representation of an actual time series can be obtained with a lower order model. The order of the model is usually not greater than two.

1.2 The Concept of Stationarity

Let the sequence of observations Z_1, Z_2, \dots, Z_N be taken at the discrete and equally spaced time intervals $1, 2, \dots, N$, then the joint distribution is invariant with regard to a displacement in time, that is,

$$P(Z_t, \dots, Z_{t+k}) = P(Z_{t+m}, \dots, Z_{t+k+m}) \quad (1.2.1)$$

where t is any point in time and k and m are any pair of positive integers. The subscripts on the distribution function have been dropped since location in time is no longer relevant. The property defined by Eq. (1.2.1) is known as "stationarity".

If we let $k = 0$, then Eq. (1.2.1) becomes

$$P(Z_t) = P(Z_{t+m}) \quad m = \underline{+1}, \underline{+2}, \dots \quad (1.2.2)$$

that is, the marginal distribution function for any two observations are the same. It follows directly that their expected values are the same

$$E(Z_t) = E(Z_{t+m}) \quad (1.2.3)$$

and that their variances are the same

$$V(Z_t) = V(Z_{t+m}) \quad (1.2.4)$$

Similarly, for $k = 1$ we have

$$P(Z_t, Z_{t+1}) = P(Z_{t+m}, Z_{t+m+1}) \quad m = \underline{+1}, \underline{+2}, \dots \quad (1.2.5)$$

which means that the covariance between Z_t and Z_{t+1} and between Z_{t+m} and

Z_{t+m+1} must be the same. These covariances may be denoted simply by $\sqrt{1}$ since their value depends only on the fact that the observations in question are separated by one period.

1.3 Autocorrelation

Another implication of stationarity which has an important interpretation in terms of the behavior of a time series derives from the fact that the autocovariance between any two observations depends only on the number of time periods separating them. Recalling the definition of autocovariance, we have

$$\begin{aligned}\gamma_j &= C(Z_t, Z_{t+j}) \\ &= E[(Z_t - EZ_t)(Z_{t+j} - EZ_{t+j})] \\ &= E[(Z_t - \mu)(Z_{t+j} - \mu)]\end{aligned}\tag{1.3.1}$$

which indicates that the covariance between observations Z_t and Z_{t+j} is the expected product of their derivations from the mean of the process. In other words if a higher-than-average observation tends to be followed by another higher-than-average observation j periods later, and likewise for lower-than-average observation, the autocovariance between Z_t and Z_{t+j} is positive. But if a higher-than-average observation tends to be followed by a lower-than-average observation j periods later and vice versa, then the autocovariance is negative.

The fact that the autocovariance γ_j seem to determine the appearance of a time series suggests that a stationary process will display the same general pattern of behavior no matter when we observe it. The realization $(Z_{100}, \dots, Z_{200})$ will not be exactly the same as $(Z_{400}, \dots, Z_{500})$, but its general appearance will be very similar. Thus in a probabilistic sense history repeats itself in stationary time series. It might seem appropriate then to characterize

a process simply by displaying the set of covariances $\gamma_0, \gamma_1, \dots$. For purpose of comparing different series, however, this is not entirely satisfactory since a difference in the dispersion of the processes, perhaps caused by different scale of measurement, would lead to very different autocovariance. For example, if sales are measured in hundred of thousands instead of millions of dollars, all second moments are increased by a factor of 100. Because variance is a measure of dispersion, comparability is achieved if we standardized the autocovariance by dividing them all by γ_0 , that is, by transforming them to correlations. Such correlations are referred to as "autocorrelation". If we denote the correlation between Z_t and Z_{t+j} by ρ_j , then the set of autocorrelations, often referred to collectively as the "autocorrelation function", is given by

$$\begin{aligned} \rho_0 &\equiv \frac{\gamma_0}{\gamma_0} = 1 \\ \rho_1 &\equiv \frac{\gamma_1}{\gamma_0} \\ \rho_2 &\equiv \frac{\gamma_2}{\gamma_0} \\ &\dots\dots\dots \\ &\dots\dots\dots \end{aligned} \tag{1.3.2}$$

Estimates of the autocorrelations of any time series are readily computed from a data sample, applying first the sample analog of Eq. (1.3.1) to obtain estimates of the autocovariances. Denoting this estimate of γ_j by C_j , we have then from data Z_1, Z_2, \dots, Z_T the estimates

$$C_j = \frac{1}{T} \sum_{t=1}^{T-j} [(Z_t - \bar{Z})(Z_{t+j} - \bar{Z})] \quad j = 1, 2, \dots \tag{1.3.3}$$

where \bar{Z} denote the sample mean

$$\bar{Z} = \frac{1}{T} \sum_{t=1}^T Z_t \quad (1.3.4)$$

Using relations Eq. (1.3.2), estimates of the autocorrelations γ_j are given then by

$$\gamma_j = \frac{C_j}{C_0} \quad j = 1, 2, \dots \quad (1.3.5)$$

A graph of the γ_j is referred to as the "sample correlogram". The sample correlogram serves much the same function in time series analysis as does the histogram in sampling problems; namely, it is not the final objective of the analysis but rather provides the basis for choice of a model, suitable to the data at hand.

1.4 Stochastic Process

A model which describes the probability structure of a sequence of observations is called a "stochastic process". A time series of N successive observations $Z' = (Z_1, Z_2, \dots, Z_N)$ is regarded as a sample realization, from an infinite population of such samples, which could have been generated by the process. A major objective of the population from those of statistical investigations is to infer properties of the population from those of the sample. For example, to make a forecast is to infer the probability distribution of a future observation from the population, given a sample, Z , of past values. To do this we need ways of describing stochastic processes and time series, and also need classes of stochastic models which are capable of describing practically a given occurring situation.

An important class of stochastic models for describing time series which has received a great deal of attention, is the so called stationary models

which assume that the process remains in equilibrium about a constant mean level. However, forecasting has been of particular importance in industry, business and economics, which many time series are represented as "non-stationary" and, in particular, as having no natural mean.

CHAPTER II

STOCHASTIC MODELS AND THEIR FORECASTING

Now, we shall introduce some simple operators which we are going to use them later on.

- 1) Backward shift operator B which is defined by $BZ_t = Z_{t-1}$; hence

$$B^m Z_t = Z_{t-m}$$

- 2) Forward shift operator F which is the inverse operation of B , and is performed by $F = B^{-1}$ giving by $FZ_t = Z_{t+1}$; hence $F^m Z_t = Z_{t+m}$

- 3) Backward difference operator ∇ which can be written in terms of B , since $\nabla Z_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1 - B)Z_t$

- 4) Summation operator S which is the inverse of the difference operator

$$\begin{aligned} \nabla \text{ is given by } \nabla^{-1} Z_t = S Z_t &= \sum_{j=0}^{\infty} Z_{t-j} \\ &= Z_t + Z_{t-1} + Z_{t-2} + \dots \\ &= Z_t + BZ_t + B^2 Z_t + \dots \\ &= (1 + B + B^2 + \dots) Z_t \\ &= (1 - B)^{-1} Z_t \end{aligned}$$

2.1 Linear Stationary Model

A general linear stochastic model is described, which suppose a time series to be a linear aggregation of random shocks. For practical representation, it is described to employ models which use parameters parsimoniously. Parsimony may often be achieved by representation of the linear process in terms of a small number of autoregressive and moving average terms. All the models for stationary time series that we shall study belong to the general class of discrete linear stochastic processes.

2.1.1 Linear Stochastic Process. A stochastic process is a linear discrete process if each observation Z_t may be expressed in the form

$$Z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots \quad (2.1.1)$$

where μ and ψ_i are fixed parameters and time series $(\dots, a_{t-1}, a_t, a_{t+1}, \dots)$ is a sequence of identically and independently distributed random disturbance with mean zero and variance σ_a^2 , often referred to as "white noise". The process is linear because the Z_t are a linear combination of the current and past disturbance.

The model in Eq. (2.1.1) implies that $\tilde{Z}_t = Z_t - \mu$ can be written alternately as a weighted sum of past values of the Z 's, plus an added shock a_t , that is

$$\begin{aligned} \tilde{Z}_t &= \pi_1 \tilde{Z}_{t-1} + \pi_2 \tilde{Z}_{t-2} + \dots + a_t \\ &= \sum_{j=1}^{\infty} \pi_j \tilde{Z}_{t-j} + a_t \end{aligned} \quad (2.1.2)$$

The alternative form Eq. (2.1.2) may be thought of as one where the current deviation Z_t , from the level μ , is "regressed" on past deviations Z_{t-1} , Z_{t-2} , \dots of the process. The relationship between the π weights and ψ weights can be expressed by equation

$$\pi(B) = \psi^{-1}(B) \quad (2.1.3)$$

The relationship in Eq. (2.1.3) may be used to derive the π weights, knowing the ψ weights, and vice versa.

Given a particular linear process, that is, particular values for the parameters, how can we verify that the process is stationary? First of all, the mean and variance-covariance matrix must exist and be invariant with respect to time. The mean of the process is simply given by

$$E(Z_t) = \mu + E(a_t + \psi_1 a_{t-1} + \dots) \quad (2.1.4)$$

Now, it is attempting to evaluate the expectation of the infinite sum by taking the sum of expectations of individual terms as one would take the expectation of a finite sum of random variables. This, however, is not a generally valid procedure; rather, it requires (from integral calculus) that

$$\sum_{i=0}^{\infty} \psi_i = K$$

$$\text{where } \psi_0 = 1 \quad (2.1.5)$$

$K = \text{some finite number}$

if condition on eq. (2.1.5) is satisfied, then the summation $\sum_{i=0}^{\infty} \psi_i$ is said to converge and the mean of the process is $E(Z_t) = \mu$

If we observed the evolution of the process over time, we could notice that it fluctuated around the value μ taking trips away from that value but always returning to that neighborhood. Note that the mean of the process does not depend on t , which satisfied one requirement of stationarity.

The variance of the process is easily derived directly from its definition as follows:

$$\begin{aligned}
\gamma_0 &= E[Z_t - E(Z_t)]^2 \\
&= E[a_t + \psi_1 a_{t-1} + \dots]^2 \\
&= E[a_t^2 + \psi_1^2 a_{t-1}^2 + \dots] + E(\text{cross-product term}) \\
&= \sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2 \tag{2.1.6}
\end{aligned}$$

The derivation is meaningful only if the mean of the process exists and the sum $\sum_{i=0}^{\infty} \psi_i^2$ converges. It is important to see why only the square terms contribute to the final answer and expectation of the cross-product term is zero. Since the mean of any a_{t-i} is zero, $E(a_{t-i}^2)$ is the variance of a_{t-i} , namely σ_a^2 , and $E(a_{t-i}a_{t-j})$, when $i \neq j$, is the covariance between a_{t-i} and a_{t-j} and therefore zero by the independence of the disturbances.

The covariance between Z_t and, say, Z_{t-j} , is easily derived in a similar fashion;

$$\begin{aligned}
\gamma_j &= E[Z_t - E(Z_t)][Z_{t-j} - E(Z_{t-j})] \\
&= E[(a_t + \psi_1 a_{t-1} + \dots)(a_{t-j} + \psi_1 a_{t-j-1} + \dots)] \\
&= E[(\psi_j a_{t-j}^2 + (\psi_1 \psi_{j+1} a_{t-j-1}^2) + \dots)] + E(\text{cross-product}) \\
&= \sigma_a^2 (\psi_j + \psi_1 \psi_{j+1} + \dots) \\
&= \sigma_a^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+j} \tag{2.1.7}
\end{aligned}$$

which is meaningful only if $\sum_{i=0}^{\infty} \psi_i \psi_{i+j}$ exists. Note that neither the variance nor the covariances depends on t , again a requirement of stationarity.

An example will clarify the points made so far. Consider the process

$$Z_t = \mu + a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \dots \tag{2.1.8}$$

where ϕ is some fraction, that is, $|\phi| < 1$. Eq. (2.1.8) satisfies the stationarity

condition $\sum_{i=0}^{\infty} \psi_i = K$, since $\psi_i = \phi^i$ and

$$\sum_{i=0}^{\infty} \psi_i = \sum_{i=0}^{\infty} \phi^i = \frac{1}{1-\phi} \quad (2.1.9)$$

The mean of the process then is μ . The variance is readily shown to be

$$\gamma^2 = \frac{\sigma_a^2}{1-\phi^2} \quad (2.1.10)$$

and similarly the autocovariances are

$$\gamma_j = \frac{\phi^j \sigma_a^2}{1-\phi^2} \quad (2.1.11)$$

Now suppose $|\phi|$ is 1 or greater, say equal to 1 exactly, so that (letting $\mu = 0$ for convenience)

$$Z_t = a_t + a_{t-1} + a_{t-2} + \dots \quad (2.1.12)$$

Clearly, the process is nonstationary since the sum $\sum_{i=0}^{\infty} \psi_i = (1 + 1 + 1 + \dots)$ does not converge, and it is also clear that there is no point in attempting to compute variances or autocovariances since $\sum_{i=0}^{\infty} \psi_i \psi_{i+j} = (1 + 1 + 1 + \dots)$ also does not converge.

We now consider a restriction applied to the π weights to ensure what is called "invertibility". The invertibility condition is independent of the stationarity condition and is applicable also to the nonstationary linear models which will be discussed later on.

To illustrate the basic idea of invertibility, consider again the model

$$\tilde{Z}_t = (1 - \theta B)a_t \quad (2.1.13)$$

Expressing the a 's in terms of the Z 's, eq. (2.1.13) becomes

$$\begin{aligned}
 a_t &= (1 - \theta B)^{-1} \tilde{z}_t \\
 &= (1 + \theta B + \theta^2 B^2 + \dots + \theta^k B^k) (1 - \theta^{k+1} B^{k+1})^{-1} \tilde{z}_t
 \end{aligned}$$

that is

$$\tilde{z}_t = -\theta \tilde{z}_{t-1} - \theta^2 \tilde{z}_{t-2} - \dots - \theta^k \tilde{z}_{t-k} + a_t - \theta^{k+1} a_{t-k-1} \quad (2.1.14)$$

and, if $|\theta| < 1$, on letting k tend to infinity, we obtain the infinite series

$$\tilde{z}_t = -\theta \tilde{z}_{t-1} - \theta^2 \tilde{z}_{t-2} - \dots + a_t \quad (2.1.15)$$

The π weights of the model in the form of eq. (2.1.2), are $\pi_j = -\theta^j$. Whatever the value of θ , eq. (2.1.13) defines a perfectly proper stationary process. However if $|\theta| > 1$, the current deviation Z_t in eq. (2.1.14) depends on \tilde{z}_{t-1} , \tilde{z}_{t-2} , \dots , \tilde{z}_{t-k} , with weights which increase as k increases. We avoid this situation by requiring that $|\theta| < 1$. We shall then say that the series is invertible. We see that this condition is satisfied if the series

$$\pi(B) = (1 - \theta B)^{-1} = \sum_{j=0}^{\infty} \theta^j B^j$$

coverage for all $|B| < 1$, that is on or within the unit circle.

In summary, a linear process is "stationary" if $\psi(B)$ converges on, or within the unit circle.

2.1.2 Autoregressive Model. A stochastic model which the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a shock a_t . Let Z_t, Z_{t-1}, Z_{t-2} be values of a process, $\tilde{z}_t, \tilde{z}_{t-1}, \tilde{z}_{t-2}$ be deviations from μ . $t, t-1, t-2$ are equally spaced time intervals. Then

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t \quad (2.1.16)$$

is referred to as an "auto-regressive process" of order p or AR(p).

If we define an autoregressive operator of order p by

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

then the model may be written economically as

$$\phi(B)\tilde{Z}_t = a_t$$

The model contains $p + 2$ unknown parameters μ , ϕ_1 , ϕ_2 , ..., ϕ_p , σ^2 which in practice have to be estimate from data. The additional parameter σ_a^2 is the variance of white noise process a_t .

It is not difficult to see that the autoregressive model is a special case of eq. (2.1.1). For example, we can eliminate Z_{t-1} from right hand side of eq. (2.1.16) by substituting

$$\tilde{Z}_{t-1} = \phi_1 \tilde{Z}_{t-2} + \phi_2 \tilde{Z}_{t-3} + \dots + \phi_p \tilde{Z}_{t-p} + a_t$$

we can likewise substitute for \tilde{Z}_{t-2} , and so on, to yield eventually an infinite series in the a 's. Symbolically, we have

$$\phi(B)\tilde{Z}_t = a_t$$

is equivalent to

$$\tilde{Z}_t = \psi(B)a_t$$

with

$$\psi(B) = \phi^{-1}(B)$$

Autoregressive process can be stationary or non-stationary. For the process to be stationary, the ϕ 's must be chosen so that the weight ψ_1 , ψ_2 ,

in $\psi(B) = \phi^{-1}(B)$ form a convergent series. For illustration, consider AR(1) model

$$(1 - \phi_1 B)\tilde{Z}_t = a_t$$

may be written

$$\tilde{Z}_t = (1 - \phi_1 B)^{-1} a_t = \sum_{j=0}^{\infty} \phi_1^j a_{t-j}$$

Hence

$$\psi(B) = (1 - \phi_1 B)^{-1} = \sum_{j=0}^{\infty} \phi_1^j B^j \quad (2.1.17)$$

we have seen in section 2.1.1 that for stationarity, $\psi(B)$ must converge for $|B| < 1$. From eq. (2.1.17) we see that this implies that parameter ϕ_1 , of an AR(1) process, must satisfy the condition $|\phi_1| < 1$ to ensure stationarity, no restrictions are required on the parameters of AR process to ensure invertibility.

An important recurrence relation for the autocorrelation function for a stationary autoregressive process is found by multiplying throughout in

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t$$

by \tilde{Z}_{t-k} to obtain

$$\tilde{Z}_{t-k} \tilde{Z}_t = \phi_1 \tilde{Z}_{t-k} \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-k} \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-k} \tilde{Z}_{t-p} + \tilde{Z}_{t-k} a_t \quad (2.1.18)$$

On taking expected values in eq. (2.1.16), we obtain the difference equation

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \quad k > 0 \quad (2.1.19)$$

Note $E[Z_{t-k} a_t]$ vanishes when $k > 0$, since Z_{t-k} can only involve the shocks a_j up to time $t-k$, which are uncorrelated with a_t . On dividing through in eq. (2.1.19) by γ_0 , it is seen that the autocorrelation function satisfies the same form of difference equation

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \quad k > 0 \quad (2.1.20)$$

If we substitute $k = 1, 2, \dots, p$ in eq. (2.1.20), we obtain a set of linear equations for $\phi_1, \phi_2, \dots, \phi_p$ in terms of $\rho_1, \rho_2, \dots, \rho_p$, that is

$$\begin{aligned} \rho_1 &= \phi_1 & + \phi_2 \rho_1 & + \dots + \phi_p \rho_{p-1} \\ \rho_2 &= \phi_1 \rho_1 & + \phi_2 & + \dots + \phi_p \rho_{p-2} \\ &\vdots & & \vdots & \dots & \vdots \\ &\vdots & & \vdots & \dots & \vdots \\ &\vdots & & \vdots & \dots & \vdots \\ \rho_p &= \phi_1 \rho_{p-1} & + \phi_2 \rho_{p-2} & + \dots + \phi_p \end{aligned} \quad (2.1.21)$$

These are usually called the Yule-Walker equations [20], [21]. We obtain Yule-Walker estimates of the parameters by replacing the theoretical autocorrelation ρ_k by the estimated autocorrelations γ_k . If we write

$$\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \vdots \\ \phi_p \end{bmatrix} \quad \rho_p = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \vdots \\ \rho_p \end{bmatrix} \quad P_p = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & 1 \end{bmatrix}$$

the solution of eq. (2.1.21) for the parameters in terms of autocorrelations may be written

$$\phi = P_p^{-1} \rho_p \quad (2.1.22)$$

when $k = 0$ in eq. (2.1.18), we have

$$\gamma_0 = \phi_1 \gamma_{-1} + \phi_2 \gamma_{-2} + \dots + \phi_p \gamma_{-p} + \sigma_a^2$$

On dividing throughout by $\gamma_0 = \sigma_z^2$ and substituting $\gamma_k = \gamma_{-k}$, the variance σ_z^2 may be written

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \rho_1 \phi_1 - \rho_2 \phi_2 - \dots - \rho_p \phi_p} \quad (2.1.23)$$

In general, the autocorrelation function of a stationary autoregressive process will consist of a mixture of damped exponentials and damped sine waves.

Initially, we may not know which order of autoregressive process to fit to an observed time series. The partial autocorrelation function denote by ϕ_{kk} , is a device which exploits the fact that whereas an AR(p) process has an autocorrelation function which is infinite in extent, it can by its very nature be described in terms of p non-zero functions of the autocorrelation. Denote by ϕ_{kj} , the jth coefficient in an autoregressive process of order k, so that ϕ_{kk} is the last coefficient. Eq. (2.1.20), ϕ_{kk} satisfies the set of equations

$$\rho_j = \phi_{k1} \rho_{j-1} + \dots + \phi_{k(k-1)} \rho_{j-k-1} + \phi_{kk} \rho_{j-k} \quad j = 1, 2, \dots, k \quad (2.1.24)$$

leading to the Yule-Walker equation (2.1.21) which may be written

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & & \rho_{k-2} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \cdot \\ \cdot \\ \phi \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdot \\ \cdot \\ \rho_k \end{bmatrix} \quad (2.1.25)$$

Solving these equations for $k = 1, 2, 3, \dots$, successively, we obtain

$$\begin{aligned} \phi_{11} &= \rho_1 \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} \\ \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \end{aligned} \quad (2.1.26)$$

In general, for ϕ_{kk} , the determinant in the numerator has the same elements as that in the denominator, but with the last column replaced by ρ_k . The quantity ϕ_{kk} , regarded as a function of the lag k , is called the partial autocorrelation function. For an autoregressive process of order p the partial autocorrelation function ϕ_{kk} will be non-zero for k less than or equal to p and zero for k greater than p . In other words, the partial autocorrelation function of a p^{th} order autoregressive process has a cutoff after lag p .

We now discuss one of the two particularly important autoregressive processes, namely the first-order autoregressive (Markov) process

$$\begin{aligned} \tilde{z}_t &= \phi_1 \tilde{z}_{t-1} + a_t \\ &= a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \dots \end{aligned} \quad (2.1.27)$$

where ϕ_1 must satisfy the condition $-1 < \phi_1 < 1$ for the process to be stationary

$$\rho_k = \phi_1^k \rho_{k-1} \quad k > 0$$

which, with $\rho_0 = 1$, has the solution

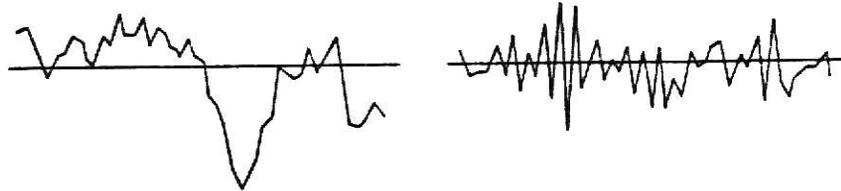
$$\rho_k = \phi_1^k \quad k > 0$$

As shown in Figure 2.1, the autocorrelation function decays exponentially to zero when ϕ_1 is positive, but decays exponentially to zero and oscillates in sign when ϕ_1 is negative, it will be noted that

$$\rho_1 = \phi_1$$

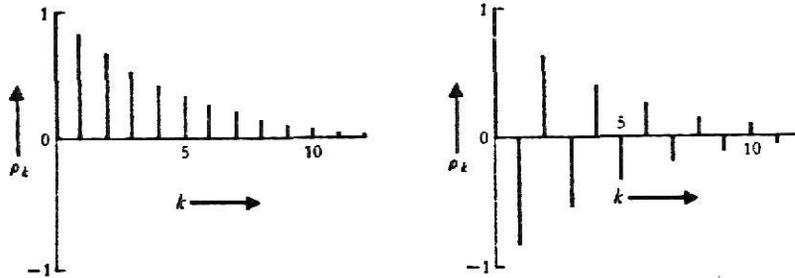
The variance of the process is

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \rho_1 \phi_1} = \frac{\sigma_a^2}{1 - \phi_1^2} \quad (2.1.28)$$



$$\tilde{z}_t = 0.8\tilde{z}_{t-1} + a_t$$

$$\tilde{z}_t = -0.8\tilde{z}_{t-1} + a_t$$



Theoretical Autocorrelation Functions

Figure 2.1. Realizations from first-order autoregressive processes and their corresponding theoretical autocorrelation functions

For AR(2) process, to achieve the stationarity, the roots of

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 0$$

must lie outside the unit circle, which implies that the parameters ϕ_1 and ϕ_2 must lie in the triangular region

$$\phi_2 + \phi_1 < 1$$

$$\phi_2 - \phi_1 < 1$$

$$-1 < \phi_2 < 1$$

Substituting $p = 2$ in Eq. (2.1.21), the Yule-Walker equations are

$$\rho_1 = \phi_1 + \phi_2 \rho_1$$

$$\rho_2 = \phi_1 \rho_1 + \phi_2$$

which,

1) when solved for ϕ_1 and ϕ_2 , give

$$\phi_1 = \frac{\rho_1(1-\rho_2)}{1-\rho_1^2}$$

$$\phi_2 = \frac{\rho_2-\rho_1^2}{1-\rho_1^2}$$

Chart B.2 in Appendix B allows values ϕ_1 and ϕ_2 to be read off for any given values of ρ_1 and ρ_2 .

- 2) When solved for ρ_1 and ρ_2 in terms of ϕ_1 and ϕ_2 , for a stationary AR(2) process, must lie in the region

$$\begin{aligned} -1 < \rho_1 < 1 \\ -1 < \rho_2 < 1 \\ \rho_1^2 < \frac{1}{2}(\rho_2 + 1) \end{aligned}$$

2.1.3 Moving Average Model. If we made \tilde{z}_t linearly dependent on a finite number q of previous a 's. Thus

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.1.29)$$

is called a "moving average process or order q " or simply MA(q).

If we defined a "moving average operator of order q " by

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

then the model may be written economically as

$$\tilde{z}_t = \theta(B)a_t$$

It contains $q + 2$ unknown parameters μ , $\theta_1 \dots \theta_q$, σ_a^2 , which in practice have to be estimated from the data. The parameters $\theta_1 \dots \theta_q$ must ensure the invertibility of the MA(q) process, no restrictions are needed on the parameters of the moving average process to ensure stationarity.

The autocorrelation function for a MA(q) process is

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \dots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & k = 1, 2, \dots, q \\ 0 & k > q \end{cases} \quad (2.1.30)$$

From eq. (2.1.29) we see that the autocorrelation function of a MA(q) process is zero, beyond the order q, of the process. In other words, the autocorrelation function of a MA(q) has cut-off at lag q.

For MA(1), θ_1 must lie in the range $-1 < \theta_1 < 1$ for the process to be invertible. However, the process is of course stationary for all values of θ_1 . Table B.1 in Appendix B allows for θ_1 to be read off over the whole range of possible values $-.5 < \rho_1 < .5$. For MA(2), it is stationary for all values of θ_1 and θ_2 . However, it is invertible only if the roots of the characteristic equation

$$1 - \theta_1 B - \theta_2 B^2 = 0$$

lie outside the unit circle, that is

$$\theta_2 + \theta_1 < 1$$

$$\theta_2 - \theta_1 < 1$$

$$-1 < \theta_2 < 1$$

These are parallel to the conditions required for stationarity of an AR(2) process. Using eq. (2.1.30), the autocorrelation function is

$$\rho_1 = \frac{-\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2}$$

$$\rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}$$

$$\rho_k = 0 \quad k > 3$$

Thus the autocorrelation function has a cut-off at lag 2.

It follows that the first two autocorrelations of an invertible MA(2) process must lie with the area bound by segment of curves

$$\rho_2 + \rho_1 = -.5$$

$$\rho_2 - \rho_1 = -.5$$

$$\rho_1^2 = 4\rho_2(1 - 2\rho_2)$$

Chart B.2 in Appendix B allows θ_1 and θ_2 to be read off over the whole range of ρ_1 and ρ_2 .

2.1.4 Mixed Autoregressive-Moving Average Process. To achieve greater flexibility in fitting of actual time series, it is sometimes advantageous to include both autoregressive and moving average term in model. This leads to the mixed autoregressive-moving average model

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.1.31)$$

or

$$\phi(B)Z_t = \theta(B)a_t$$

where $\phi(B)$ and $\theta(B)$ are polynomials of degree p and q , in B . This is referred to as an ARMA(p, q) process.

It is obvious that moving average terms on the right of eq. (2.1.31) will not affect the argument of Section 2.1.2, which establishes condition for stationarity of an autoregressive process. Thus, $\phi(B)\tilde{Z}_t = \theta(B)a_t$ will define a stationary process, provided that the $\phi(B) = 0$ has all its roots lying outside the unit circle. Similarly, the roots of $\theta(B) = 0$ must lie outside the unit circle if the process is to be invertible.

The autocorrelation function of ARMA process may be derived by a method similar to that used for the AR process in Section 2.1.2. On multiplying the throughout by \tilde{Z}_{t-k} in eq. (2.1.31) and computing the expectation, we see that the autocovariance function satisfies the difference equation

$$\gamma_k = \phi_1\gamma_{k-1} + \dots + \phi_p\gamma_{k-p} + \gamma_{za}(k) - \theta_1\gamma_{za}(k-1) - \dots - \theta_q\gamma_{za}(k-q) \quad (2.1.32)$$

where $\gamma_{za}(k) = 0$ when $k > 0$
 $\gamma_{za}(k) \neq 0$ when $k < 0$

We see that eq. (2.1.32) implies

$$\gamma_k = \phi_1\gamma_{k-1} + \phi_2\gamma_{k-2} + \dots + \phi_p\gamma_{k-p} \quad k > q + 1$$

and hence

$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} + \dots + \phi_p\rho_{k-p} \quad k < q + 1 \quad (2.1.33)$$

Thus, for the ARMA(p,q) process, there will be q autocorrelations $\rho_q, \rho_{q-1}, \dots, \rho_1$ whose values depend directly, through eq. (2.1.32), on the choice of the q moving average parameters θ , as used as on the p autoregressive parameters ϕ . Also, the p values $\rho_q, \rho_{q-1}, \dots, \rho_{q-p+1}$ provide the necessary

starting values for the difference equation $\phi(B)\rho_k = 0$, where $k > q + 1$, which then entirely determines the autocorrelations at higher lags. If $q - p < 0$, the whole autocorrelation function ρ_j , for $j = 0, 1, 2, \dots$, will consist of a mixture of damped exponentials and/or damped sine waves, where nature is dictated by the polynomial $\phi(B)$ and the starting value. If however, $q - p > 0$ there will be $q - p + 1$ initial values $\rho_0, \rho_1, \dots, \rho_{q-p}$, which do not follow this general pattern. These facts are useful in identifying mixed series.

The partial autocorrelation function of an ARMA process is infinite in extent. It behaves eventually like the partial autocorrelation function of an MA process, being dominated by a mixture of damped exponentials and/or damped sine waves, depending on the order of the moving average and the values of the parameters it contains.

A mixed process of considerable practical importance is the first-order autoregressive--first-order moving average ARMA(1,1) process

$$\tilde{Z}_t - \phi_1 \tilde{Z}_{t-1} = a_t - \theta_1 a_{t-1} \quad (2.1.34)$$

that is

$$(1 - \phi_1 B) \tilde{Z}_t = (1 - \theta_1 B) a_t$$

We now derive some of its important properties.

- 1) The process is stationary if $-1 < \phi_1 < 1$, and invertible if $-1 < \theta_1 < 1$.
- 2) Autocorrelation function. From eq. (2.1.32), we obtain

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_a^2 - \theta_1 \gamma_{za} (-1)$$

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_a^2$$

$$\gamma_k = \phi_1 \gamma_{k-1} \quad k > 2$$

On multiplying by a_{t-1} throughout eq. (2.1.34), and taking expectations, we obtain

$$\gamma_{za}(-1) = (\phi_1 - \theta_1) \sigma_a^2$$

Hence,

$$\gamma_0 = \frac{1 + \theta_1^2 - 2\phi_1\theta_1}{1 - \phi_1^2} \sigma_a^2$$

$$\gamma_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_a^2$$

$$\gamma_k = \phi_1 \gamma_{k-1} \quad k > 2$$

Then, we have

$$\rho_1 = \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1}$$

$$\rho_2 = \phi_1 \theta_1 \quad (2.1.35)$$

Thus, the autocorrelation function decays exponentially from the starting value ρ_1 , which depends on ϕ_1 , as well as on θ_1 . As shown in Figure 2.3, this exponential decay is smooth if ϕ_1 is positive and oscillatory if ϕ_1 is negative. Furthermore, the sign of ρ_1 is determined by the sign of $(\phi_1 - \theta_1)$ and dictate from which side of zero the exponential decay takes place. Using eq. (2.1.35) and the stationarity and invertibility conditions it may be shown that ρ_1 and ρ_2 must lie in the region

$$\begin{aligned} |\rho_2| &< |\rho_1| \\ \rho_2 &> \rho_1(2\rho_1 + 1) \quad \rho_1 < 0 \\ \rho_2 &> \rho_1(2\rho_1 - 1) \quad \rho_1 > 0 \end{aligned}$$

- 3) The partial autocorrelation function of the mixed ARMA(1,1) process consists of a single initial value $\phi_{11} = \rho_1$. Thereafter it behaves like it in MA(1) process, and is dominated by a damped exponential. Thus as shown in Figure 2.2, when θ_1 is positive it is dominated by a smoothly damped exponential which decays from ρ_1 , with sign determined by the sign of $(\phi_1 - \theta_1)$. Similarly, when θ is negative, it is dominated by an exponential which oscillates as it decays from ρ_1 , with sign determined by the sign of $(\phi_1 - \theta_1)$.

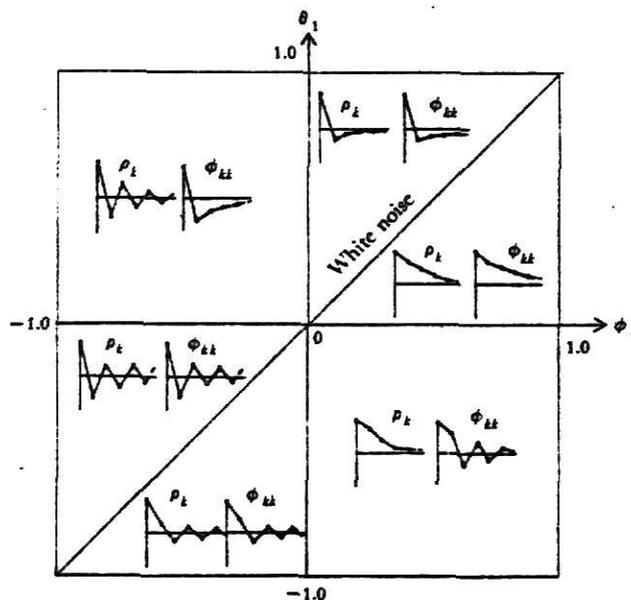


Figure 2.2 Autocorrelation and partial autocorrelation functions ρ_k and ϕ_{kk} for various ARMA(1,1) models

2.1.5 Summary. The pure autoregressive and moving-average processes are duals of one another and opposite behavior of autocorrelation and partial autocorrelation functions indicates a pure autocorrelation or moving-average process as illustrated in Figure 2.3. This is a very important concept of identification stage in Chapter 3.

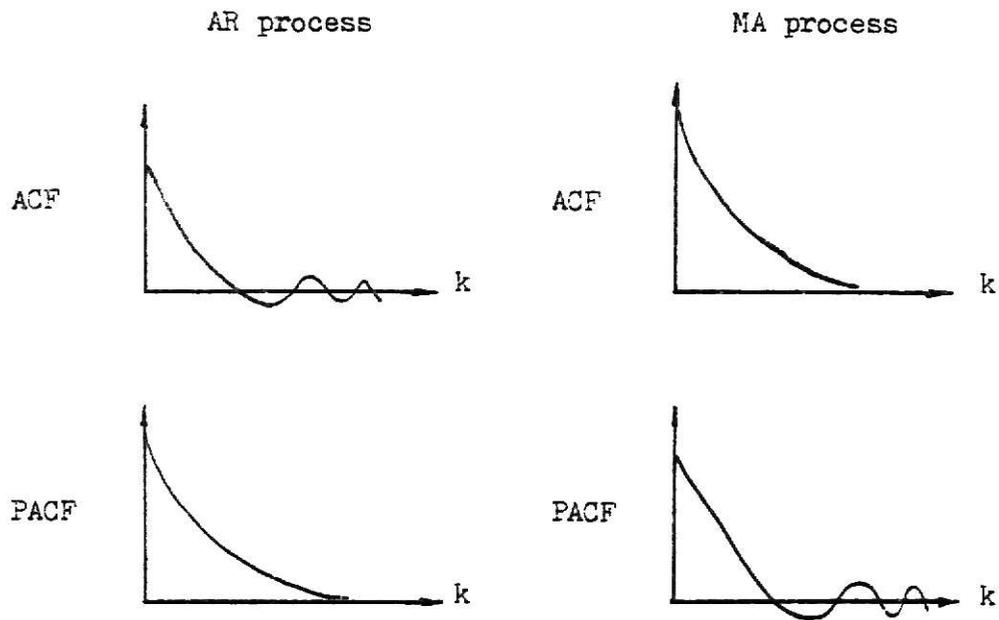


Figure 2.3 Summary of autocorrelation and partial autocorrelation functions for pure AR and MA processes

Table 1 summarizes the properties of mixed autoregressive-moving-average processes, and brings together all the important results for autoregressive, moving-average and mixed processes, which will be needed in the next chapter to identify models for observed time series.

Table 1. Summary of properties of autoregressive, moving-average, and mixed ARMA processes.

	AR processes	MA processes	Mixed processes
Model in terms of previous Z_t 's	$\phi(B)Z_t = a_t$	$\theta^{-1}(B)Z_t = a_t$	$\theta^{-1}(B)\phi(B)Z_t = a_t$
Model in terms of previous a_t 's	$Z_t = \phi^{-1}(B)a_t$	$Z_t = \theta(B)a_t$	$Z_t = \phi^{-1}(B)\theta(B)a_t$
π weights	finite series	infinite series	infinite series
ψ weights	infinite series	finite series	infinite series
Stationarity condition	roots of $\phi(B) - 0$ lie outside unit circle	always stationary	roots of $\phi(B) = 0$ lie outside unit circle;
Invertibility condition	always invertible	roots of $\theta(B) = 0$ lie outside unit circle	roots of $\theta(B) = 0$ lie outside unit circle
Autocorrelation function	infinite (damped exponentials and/or damped sine waves)	finite	infinite (damped exponentials and/or damped sine waves after first $q - p$ lags)
Partial autocorrelation function	finite	infinite (dominated by damped exponentials and/or sine waves)	infinite (dominated by damped exponentials and/or sine waves after first $p - q$ lags)
	tails off	cuts off	tails off
	cuts off	tails off	tails off

2.2 Linear Nonstationary Models

Stationarity is a very strong condition to impose on a time series and is presumably never literally true of series encountered in practice; rather, it should be viewed as a useful working assumption. There are any number of ways in which departures from stationarity might occur, but the most apparent one is the lack of affinity for a mean value.

The kind of nonstationarity displayed by such series may be characterized as being "homogeneous" in the sense that although the series moves freely without affinity for a particular location, its behavior at different periods in time is essentially the same. Fortunately, "homogeneous nonstationarity is displayed by series whose successive changes or difference are stationarity". Thus, the models for stationary series studied in Section 2.1 provide the basis for a highly flexible class of models for nonstationary series if we follow the simple expedient of working with their differences.

2.2.1 Difference and Homogeneous Nonstationarity. The motivation for focusing on differencing as a means of accommodating homogeneous nonstationarity becomes evident by considering the AR(1) process

$$Z_t = \phi_1 Z_{t-1} + a_t \quad (2.2.1)$$

and the values that might be taken by the parameters ϕ_1 . If $|\phi_1| < 1$, then the process is stationary. On the other hand, if $|\phi_1| > 1$, then the behavior of the series will be explosive. That is, if we were to start the process off at, say 0, the disturbances would be important in determining the first few values of the series; however, after a time the series would "take off" growing exponentially. The disturbances would become negligible relative to the level of the series, and hence the series would become essentially deterministic in

its evolution. Consider the case $\phi_1 = 1$. This is the random walk process that, as we have seen in Section 2.1, is nonstationary. It is homogeneous because the distribution of change or differences in the process is unchanging; that is the time series of differences is stationary because the difference are just

$$Z_t - Z_{t-1} = a_t \quad (2.2.2)$$

and the distribution of a_t is fixed.

A natural generalization of the random walk is to consider the whole class of stationary ARMA processes as potential generating mechanisms for the differences of a nonstationary series. Thus if we define w_t to be the sequence of differences

$$w_t = Z_t - Z_{t-1} \quad (2.2.3)$$

then, the general model may be written

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.2.4)$$

Replacing w_t with $(Z_t - Z_{t-1})$, we see that the observed series Z_t is given directly by

$$\begin{aligned} Z_t = & Z_{t-1} + \phi_1(Z_{t-1} - Z_{t-2}) + \dots + \phi_p(Z_{t-p} - Z_{t-p-1}) \\ & + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \end{aligned} \quad (2.2.5)$$

Note that Z_t is just the sum of all past changes; that

$$Z_t = w_t + w_{t-1} + w_{t-2} + \dots \quad (2.2.6)$$

Hence, Z_t is referred to as the integration of w_t series and the process in eq. (2.2.4) as an integrated autoregressive--moving average process.

In some case, it may be that even the first difference are nonstationary but that the second differences are stationary. The second differences are the differences of the first differences; that is if the y_t are the second differences of the Z_t , then

$$\begin{aligned} Y_t &= w_t - w_{t-1} \\ &= (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2}) \\ &= Z_t - 2Z_{t-1} + Z_{t-2} \end{aligned} \quad (2.2.7)$$

Denoting the degree of differencing by d , then an ARIMA process may be described by the dimensions p , d , and q . As a matter of convenience, an ARIMA process with no moving average part may be referred to as ARI(p,d) and one with no autoregressive part as IMA(d,q).

2.2.2 The General Form of the ARIMA Process.

$$(B)Z_t = \phi(B)\nabla^d Z_t = \theta_0 + \theta(B)a_t \quad (2.2.8)$$

where

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

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

In the following discussion

- 1) $\phi(B)$ will be called the autoregressive operator, it is assumed to be stationary, that is the roots of $\phi(B) = 0$ lie outside the unit circle.
- 2) $(B) = \nabla^d \phi(B)$ will be called the generalized autoregressive operator; it is nonstationary operator with d of the roots of $(B) = 0$ equal to unity.

- 3) $\theta(B)$ will be called the moving average operator; it is assumed to be invertible, that is, the root of $\theta(B) = 0$ lie outside the unit circle.

When $d = 0$ the model in eq. (2.2.8) represents a stationary process. The requirement of stationary and invertibility apply independently and in general the operators $\phi(B)$ and $\theta(B)$ will not be of the same order.

When the constant term θ_0 is omitted, the model in eq. (2.2.8) is capable of representing series which have stochastic trends, as typified for example, by random changes in the level and slope of the series.

An alternative way of expressing this more general model in eq. (2.2.8) is in the form of a stationary invertible ARMA process in $\tilde{w}_t = w_t - \mu_w$. That is

$$\phi(B)\tilde{w}_t = \theta(B)a_t \quad (2.2.9)$$

In many applications, where no physical reason for a deterministic component exists, the mean of w can be assumed to be zero unless such an assumption proves contrary to facts presented by the data. It is clear that, for many applications, the assumption of a stochastic trend is often more realistic than the assumption of a deterministic trend. This is of special importance in forecasting a time series, since a stochastic trend does not necessitate the series to follow the identical pattern which it has developed in the past. In what follows, when $d > 0$, we shall often assume that $\mu_w = 0$, or equivalently that $\theta_0 = 0$, unless it is clear from the data or from the nature of the problem that a nonzero mean, or more generally a deterministic component of known form, is needed.

2.2.3 Three Explicit Forms for the ARIMA Model. In this section (2.2) we are mainly concerned with nonstationary models, in which $\nabla^d z_t$ is a stationary

process and $d > 0$. For such models we can, without loss of generality, omit μ from the specification or equivalently replace \tilde{Z}_t by Z_t . The result of this chapter will, however, apply to stationary models for which $d = 0$, provided Z_t is then interpreted as the deviation from the mean.

We now consider three different "explicit" forms for the general model in eq. (2.2.8). Each of these allows some special aspect to be appreciated.

- 1) The difference-equation form of the ARIMA process. The current value of the process Z_t can be expressed in terms of previous values of the Z 's and current and previous values of the a 's (disturbances) by directly use of the difference equation; as in eq. (2.2.5) for the case $d = 1$. Rearranging the terms in eq. (2.2.5) we have

$$\begin{aligned} Z_t = & (1 + \phi_1)Z_{t-1} + (\phi_2 - \phi_1)Z_{t-2} + \dots + (\phi_p - \phi_{p-1})Z_{t-p} \\ & - \phi_p Z_{t-p-1} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \end{aligned} \quad (2.2.10)$$

which is referred to as the difference-equation form of the ARIMA $(p,1,q)$ model. It is this form of the model that will ultimately be used in computing forecasts.

As an example, the difference-equation form of the very important ARIMA(0,1,1), or simply IMA(1,1), process is

$$Z_t = Z_{t-1} + a_t - \theta_1 a_{t-1} \quad (2.1.11)$$

- 2) The random-shock form of the ARIMA process. Z_t can be expressed in terms of previous a 's (disturbances) only

$$Z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \psi(B)a_t \quad (2.2.12)$$

It is referred to as the "random-shock" form of the process. It may be obtained by successive substitution for Z_{t-1} , Z_{t-2} , and so forth, using the difference equation of the process. The specific values of weights ψ_i then depend on the order of differencing and values of the θ and coefficients. The randomshock form of a process will be very important in forecasting since it is the ψ_i weights along with σ_a^2 that will be used for computation of standard errors and confidence intervals for future observation.

For example, the random-shock form of the IMA(1,1) model may be obtained by successive substitution for Z_{t-1} , Z_{t-2} , and so forth; hence

$$\begin{aligned}
 Z_t &= Z_{t-1} + a_t - \theta_1 a_{t-1} \\
 &= (Z_{t-2} + a_{t-1} - \theta_1 a_{t-2}) + a_t - \theta_1 a_{t-1} \\
 &\dots\dots\dots \\
 &= a_t + (1 - \theta_1)a_{t-1} + (1 - \theta_1)a_{t-2} + \dots
 \end{aligned} \tag{2.2.13}$$

so that for IMA(1,1) we have

$$\psi_i = (1 - \theta_1) \quad i = 1, 2, \dots \tag{2.2.14}$$

As expected, the IMA(1,1) process does not satisfy the stationary condition for linear process since $1 = (1 - \theta_1) + (1 - \theta_1) + \dots$ does not converge.

- 3) The inverted form of the ARIMA process. Starting with the difference-equation form of the process and substituting successively for a_{t-1} , a_{t-2} , \dots , Z_t may be expressed in terms of the current shock a_t and previous Z 's only; that is

$$Z_t = \pi_1 Z_{t-1} + \pi_2 Z_{t-2} + \dots + a_t \quad (2.2.15)$$

This form of the process is referred to as the inverted form of the process. For the IMA(1,1) process, the substitution lead to

$$\begin{aligned} Z_t &= Z_{t-1} + a_t - \theta_1 a_{t-1} \\ &= Z_{t-1} + a_t - \theta_1 [(Z_{t-1} - Z_{t-2}) + \theta_1 a_{t-2}] \\ &= a_t + (1 - \theta_1) Z_{t-1} + \theta_1 Z_{t-2} - \theta_1^2 [(Z_{t-2} - Z_{t-3}) + \theta_1 a_{t-3}] \\ &\dots\dots\dots \\ &= (1 - \theta_1) Z_{t-1} + \theta_1 (1 - \theta_1) Z_{t-2} + \theta_1^2 (1 - \theta_1) Z_{t-3} + \dots\dots + a_t \end{aligned} \quad (2.2.16)$$

so that the weight π_i are given by

$$\pi_i = \theta_1^{i-1} (1 - \theta_1) \quad i = 1, 2, \dots \quad (2.2.17)$$

CHAPTER III

STOCHASTIC MODEL BUILDING

The main purpose in constructing a model for the given input-output data is to obtain an understanding of the process. Typically, any analytical expression that explains the nature and extent of the dependency of the present observation on past history can be said to increase our understanding of the process. Another way of looking at a model of an empirical process is that the model is a convenient way of summarizing the entire available set of observations; i.e., the important characteristics of the data can be recovered from the model by analyzing the model or simulating it.

3.1 Introduction

All models contain constants or parameters whose values must be estimated from the data. It is important that we employ the smallest possible number of parameters for adequate representation. The central role played by this principle of parsimony [25] in the use of parameters will become clear as we proceed. As a preliminary illustration, we consider the model

$$Y_t = (\eta_0 + \eta_1 \nabla + \eta_2 \nabla^2 + \dots + \eta_s \nabla^s) X_t \quad (3.1)$$

when dealing with system which are adequately represented by

$$(1 + \xi \nabla) Y_t = X_t \quad (3.2)$$

The model in eq. (3.2) contains only one parameter but, for s sufficiently large, could be represented approximately by the model in eq. (3.1). Because of experimental error, we could easily fail to recognize the relationship between

the coefficients in the fitted equation. Thus, we may needlessly fit a relationship like eq. (3.1), containing s parameters, where the much simpler form in eq. (3.2), containing only one, would have been adequate. This could, for example, lead to a poor estimate of the output Y_t for given values of the input X_t, X_{t-1}, \dots

An approach to the modeling of stationary and nonstationary time series, such as those that commonly occur in economic situations and control problems is discussed by Box and Jenkins [1], involves iterative use of the three stage processes of identification, estimation, and diagnostic checking (refer to Figure 3.1). Our objective then, must be to obtain adequate but parsimonious models. Much care and effort is needed in selecting the model.

Summarizes the Box-Jenkins iterative approach to model building for forecasting and control [1].

- 1) From the interaction of theory and practice, a useful class of model for the purpose at hand is considered.
- 2) Because this class of model is too extensive to be conveniently fitted directly to data, rough methods for identifying subclasses of these models are developed. Such methods of model identification employ data and knowledge of the system to suggest an appropriate parsimonious subclass of models which may be tentatively entertained. In addition, the identification process can be used to yield rough preliminary estimate of the parameters in the model.
- 3) The tentatively entertained model is fitted to data and its parameters estimated. The rough estimates obtained during the identification stage can now be used as starting values in more refined iterative methods for estimating the parameters.

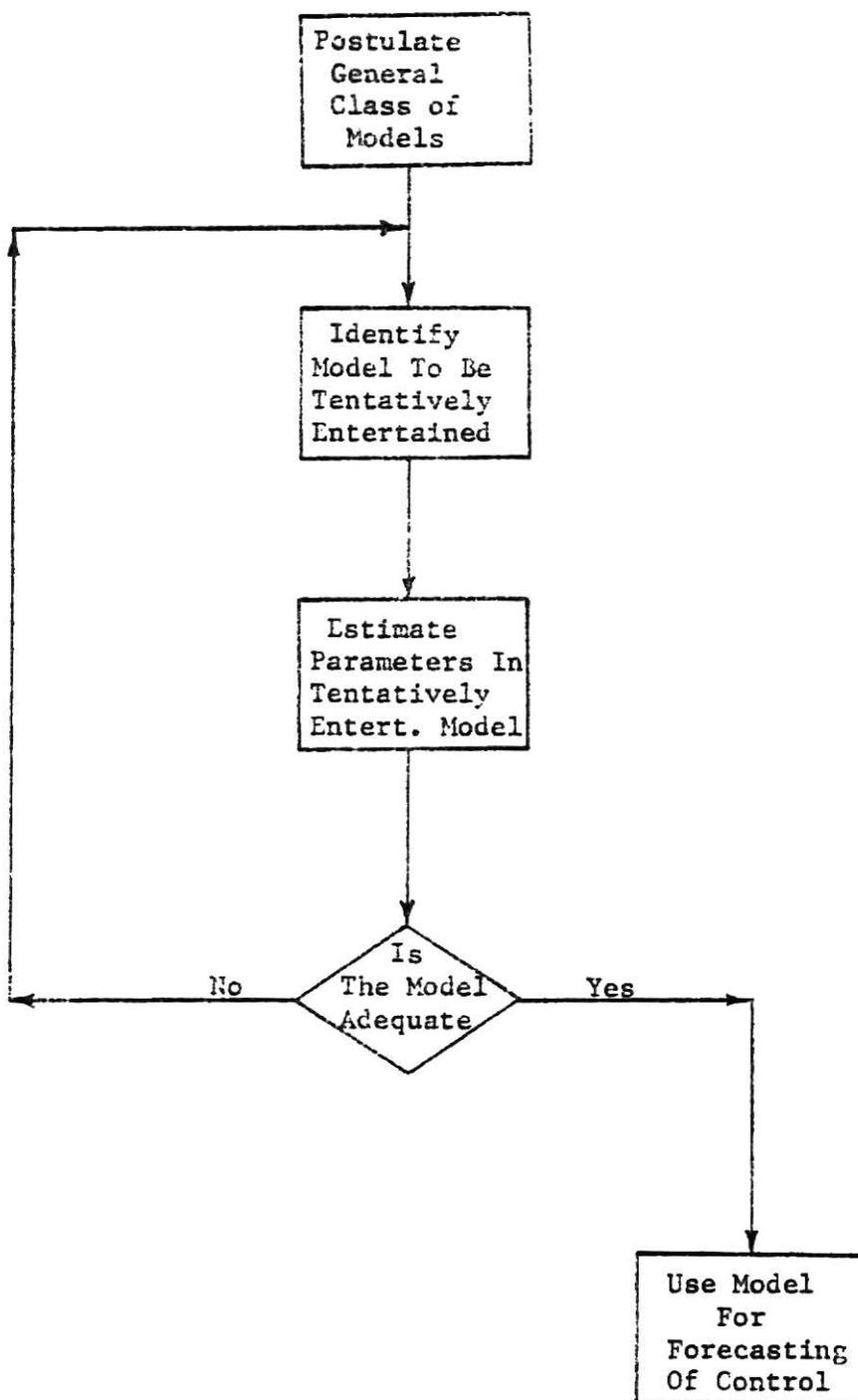


Figure 3.1 Stages in the iterative approach to model building

- 4) Diagnostic checks are applied with the object of uncovering possible lack of fit and diagnosing the cause. If no lack of fit is indicated, the model is ready to use. If any inadequacy is found, the iterative cycle of identification, estimation, and diagnostic checking is repeated until a suitable representation is found.

3.2 Identification

Identification methods are roughly applied to a set of data to indicate the kind of representational model which is worthy of further investigation. The specific aim here is to obtain some idea of the values of p , d , and q needed in the general linear ARIMA model.

It should be said that identification and estimation necessarily overlap. Thus, we may estimate the parameters in a model, which is more elaborate than that which we expect to find, so as to decide at what point simplification is possible. It should also be explained that identification is necessarily inexact. It is inexact because the question of the types of models occur in practice and under which circumstances they occur, is a property of the behavior of the physical world and cannot, therefore, be decided by pure mathematical argument. It is a stage at which graphical methods are particularly useful and judgment must be exercised.

Our task then, is to identify an appropriate subclass of models from the general ARIMA family

$$\phi(B)\nabla^d Z_t = \theta(B)a_t \quad (3.1.1)$$

Our approach will be

- 1) to difference Z_t as many times as needed to produce stationarity, hopefully reducing the process under study to the ARIMA process

$$\phi(B)w_t = \theta(B)a_t$$

where

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

- 2) To identify the resulting ARMA process. Our principal tool for putting 1) and 2) into effect will be the autocorrelation function and the partial autocorrelation function. They are used not only to help guess the form of the model, but also to obtain approximate estimates of the parameters. Such approximations are often useful at the estimation stage to provide starting values for iterative procedures employed at that stage.
- 3) Having obtained a tentative model, rough values for the appropriate parameters can be formed using Table 2, checking first that the ρ_1 and ρ_2 are admissible for the particular model. It is recently discovered by O. D. Anderson [3], an interesting set of inequalities, which must be obeyed by all MA process. For any MA(q) process,

$$|\rho_1| < \cos\left(\frac{\pi}{q+2}\right)$$

and for MA(1)

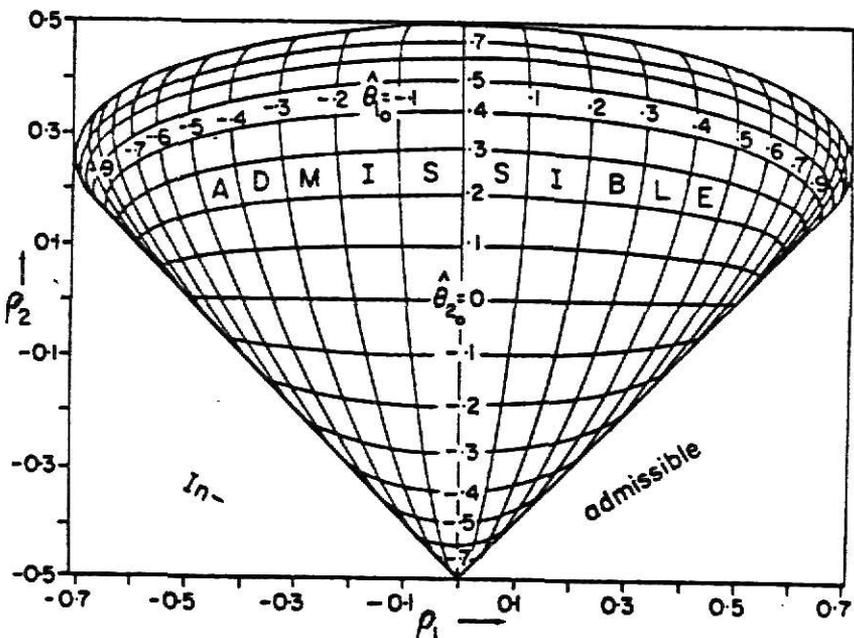
$$|\rho_1| < 1/2$$

which for MA(2)

$$|\rho_1| < 1/\sqrt{2}$$

$$|\rho_2| < 1/2$$

Table 2. Admissible Region and Initial Estimates for Various Process.

Process	Admissible region	Initial estimates
AR(1)	$-1 < \rho_1 < 1$	$\hat{\phi}_1 = \rho_1$
AR(2)	$-1 < \rho_1 < 1$ $-1 < \rho_2 < 1$ $\rho_1^2 < \frac{1}{2}(\rho_2 + 1)$	$\hat{\phi}_1 = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2}$ $\hat{\phi}_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$
MA(1)	$-.5 < \rho_1 < .5$	$\hat{\theta}_1 = \frac{[1 - (1 - 4\rho_1^2)^{1/2}]}{2\rho_1}$
MA(2)		
ARMA(1,1)	$2\rho_1^2 - \rho_1 < \rho_2 < \rho_1 $	$\hat{\phi}_1 = \frac{\rho_2}{\rho_1}$ $\hat{\theta}_1 = \frac{b \pm (b^2 - 4)}{2}$ where $b = \frac{1 - 2\rho_2 + \hat{\phi}_1^2}{\rho_1 - \hat{\phi}_1}$

3.3 Estimation

Having identified a tentative model, the next stage is to estimate its parameters efficiently using the rough values, calculated at identification stage, as initial values. Inadequacy of fit may simply arise because of insufficient fitting and not because the form of the model is inadequately. A non-linear least square procedure is used to obtain the vector of parameters estimates

$$(\hat{\phi}, \hat{\theta}) \equiv (\hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q)$$

which minimizes the shock sum of squares

$$S(\phi, \theta) = \sum_{i=1}^N a_i^2$$

where the $a_i = \theta^{-1}(B)\phi(B)Z_i$ are the estimated shocks given the model and the series. Box and Jenkins give a vague description of such a program, but satisfactory packages are now become available, for instance I.C.L. 31.

Alternately, a grid-search method can be used. A neighborhood of the initial estimates vector, within the stationary invertible region, is first chosen, and a fine mesh of points superimposed. The estimated shocks, for models with parameters at the mesh points, are calculated, and the S contains plotted. From these, the point corresponding to the minimum of the S surface can generally be determined with precision, and this is taken as vector estimate.

For a single parameter model S surface is merely a curve, and for not more than two parameters in the model (which covers most the cases we are usually considering), the surface can be readily represented. The advantage of the method is that the S surface, for varying parameters, can be studied visually, and it often provides useful information. Approximate standard error for the parameter estimates of the various simple models can be obtained from Table 3.

Table 3. Approximate variance for parameter estimates of different models.

AR(1)	$\text{VAR}[\hat{\phi}] = \frac{1 - \phi^2}{N}$
AR(2)	$\text{VAR}[\hat{\phi}_1], \text{VAR}[\hat{\phi}_2] = \frac{1 - \phi^2}{N}$
MA(1)	$\text{VAR}[\hat{\theta}] = \frac{1 - \theta^2}{N}$
MA(2)	$\text{VAR}[\hat{\theta}_1], \text{VAR}[\hat{\theta}_2] = \frac{1 - \theta^2}{N}$
ARMA(1,1)	$\text{VAR}[\hat{\phi}] = \frac{(1 - \phi^2)(1 + \phi\theta)^2}{N(\phi + \theta)^2}$
	$\text{VAR}[\hat{\theta}] = \frac{(1 - \theta^2)(1 + \phi\theta)^2}{N(\phi + \theta)^2}$

3.4 Diagnostic Checking

Following the parametric estimation procedure, a diagnostic checking step is performed, which involves examination of the white noise residuals a_t from the fitted model. This results in an indication of model adequacy, or model inadequacy, together with information on how to better describe the series. Thus, the residuals a_t , which should be white noise, would be examined for any lack of randomness.

One technique which can be used for diagnostic is "overfitting". Having identified what is believed to be a correct model, are actually fit a more elaborate one. This puts the identified model in jeopardy, because the more elaborate model contains additional parameters covering feared direction of

discrepancy. Careful thought should be given to the question of how the model should be augmented. In particular, in accordance with the model redundancy, it would be foolish to add factors simultaneously to both sides of the ARMA model. By extending the model in a particular direction, we assume that we know what kind of discrepancies are to be feared. Procedures less dependent upon knowledge are based on the analysis of "residuals", these procedures provide more opportunity for the data themselves to suggest modifications.

3.4.1 Autocorrelation check. Suppose a model

$$\phi(B)\tilde{w}_t = \theta(B)a_t$$

with $w_t = \nabla^d Z_t$ has been fitted and estimates $(\hat{\phi}, \hat{\theta})$ obtained for the parameters. Then we shall refer to the quantities

$$\hat{a}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)\tilde{w}_t$$

as the "residuals". Now, it is possible to show that, if the model is adequate,

$$\hat{a}_t = a_t + O\left(\frac{1}{\sqrt{n}}\right)$$

As the series length increases, the \hat{a}_t 's become close to white noise a_t 's. Therefore, study \hat{a}_t 's could indicate the existence and nature of model inadequacy. In particular, recognizable patterns in the estimated autocorrelation function of the \hat{a}_t 's could point to appropriate modification in the model.

Now, suppose the form of the model were correct and that we knew the true parameter ϕ and θ . Then, using Bartlett's approximation

$$\text{VAR}[\gamma_k] = \frac{1}{N} \left\{ 1 + 2 \sum_{v=1}^q \rho_v^2 \right\}, \quad k > q$$

and a result of Anderson [10], the estimated autocorrelation $\gamma_k(a)$, of the a 's, would be uncorrelated and distributed approximately $N(0, N^{-1})$, and hence, with a standard error of $N^{-1/2}$. We could use these facts to assess approximately the statistical significance of apparent departures of the autocorrelation from zero.

3.4.2 A Portmanteau Lack of Fit Test. Rather than consider the $\gamma_k(a)$'s individually, an indication is often needed of whether, the first k ($k = 20$, will be large enough) autocorrelations of the \hat{a} 's, taken as a whole, indicate inadequacy of the model.

Suppose we have the first k autocorrelations $\gamma_k(a)$ ($k = 1, 2, \dots, k$) from any ARIMA(p, d, q) process, then it is possible to show [16] that if the fitted model is appropriate, $n \sum_{k=1}^M \gamma_k^2(\hat{a})$ would possess a χ^2 distribution with M degree of freedom. If M is taken sufficiently large so that the elements after the M^{th} in the

$$Q = n \sum_{k=1}^M \gamma_k^2(\hat{a})$$

is approximately distributed as $\chi^2(M - p - q)$, where $n = N - d$ is the number of w 's used to fit the model. On the other hand, if the model is inappropriate, the average values of Q will be inflated. Therefore, an approximate, general, or "Portmanteau" test of the hypothesis of model adequacy, designed to take account of the difficulties discussed above, may be made by referring an observed value of Q to a table of the percentage points of χ^2 .

3.5 Forecasting

In Chapter I, the concept of stochastic processes as models for time series was introduced. It was the ultimate objective of using the model to

infer from the past history of a series its likely course in the future. More precisely, we want to derive from a model the conditional distribution of future observation given the past observation. This final step in the modeling building process is what we refer to as "forecasting".

Forecasts are usually needed over a period known as the lead time, which varies with each problem.

The accuracy of the forecasts may be expressed by calculating probability limit on either side of each forecast. These limits may be calculated for any convenient set of probabilities, for example 50% and 95%. They are such that the analyzed value of the time series, when it eventually occurs, will be included within these limits with the stated probability.

It should be noted that in practice the model in hand is never the hypothetical "true" process generating the data we have observed. Rather, it is an approximation to the generating process and is subject to errors in both identification and estimation step. Thus, although we shall discuss forecasting as if we knew the generating process, it is clear that our success in practice will depend in part on the adequacy of our empirical model and therefore on success in the preceding stage of identification and estimation.

3.5.1 Minimum Mean-Square Error Forecast. In Section 2.2.3, we discussed three explicit forms for the general ARIMA model;

$$\varphi(B)Z_t = \theta(B)a_t \quad (3.5.1)$$

where $\varphi(B) = \phi(B)\nabla^d$. We begin by recalling these three forms, since each one throws light on a different aspect of the forecasting problem.

We shall be concerned with forecasting a value Z_{t+l} , $l > 1$, when we are currently standing at time t . This forecast is said to be made at origin t for

lead time 1. We now summarize the results of Section 2.2.3, but writing $t + 1$ for t .

An observation Z_{t+1} generated by the process in eq. (3.5.1) may be expressed:

- 1) Directly in terms of the difference equation by

$$Z_{t+l} = \varphi_1 Z_{t+l-1} + \dots + \varphi_{p+d} Z_{t+l-p-d} - \theta_1 a_{t+l-1} - \dots - \theta_q a_{t+l-q} + a_{t+l} \quad (3.5.2)$$

- 2) As an infinite weighted sum of current and previous shocks a ,

$$Z_{t+l} = \sum_{-\infty}^{t+l} \psi_{t+l-j} a_j = \sum_{j=0}^{\infty} \psi_j a_{t+l-j} \quad (3.5.3)$$

where $\psi_0 = 1$ and, the weights may be obtained by equating coefficients in

$$\varphi(B)(1 + \psi_1 B + \psi_2 B^2 + \dots) = \theta(B) \quad (3.5.4)$$

- 3) As an infinite weighted sum of previous observations, plus a random shock

$$Z_{t+l} = \sum_{j=1}^{\infty} \pi_j Z_{t+l-j} + a_{t+l} \quad (3.5.5)$$

Also, if $d > 1$

$$Z_{t+l-1}(\pi) = \sum_{j=1}^{\infty} \pi_j Z_{t+l+j} \quad (3.5.6)$$

will be a weighted average, since then $\sum_{j=1}^{\infty} \pi_j = 1$, and the weights may be obtained from

$$\varphi(B) = (1 - \pi_1 B - \pi_2 B^2 - \dots) (B)$$

Now suppose, standing at origin t , we are to make a forecast $\hat{Z}_t(\ell)$ of $Z_{t+\ell}$ which is to be a linear function of current and previous observations Z_t, Z_{t-1}, \dots . Then it will also be a linear function of current and previous shocks a_t, a_{t-1}, \dots .

Suppose then, that the best forecast is

$$\hat{Z}_t(\ell) = \psi_\ell^* a_t^* + \psi_{\ell+1}^* a_{t-1} + \psi_{\ell+2}^* a_{t-2} + \dots$$

where the weights $\psi_\ell^*, \psi_{\ell+1}^*, \dots$ are determined. Then using eq. (3.5.3) the mean square error of the forecast is

$$E[Z_{t+\ell} - \hat{Z}_t(\ell)]^2 = (1 - \psi_1^2 + \dots + \psi_{\ell-1}^2) \sigma_a^2 + \sum_{j=0}^{\infty} \{\psi_{\ell+j} - \psi_{\ell+j}^*\}^2 \sigma_a^2 \quad (3.5.7)$$

which is minimized by setting $\psi_{\ell+j}^* = \psi_{\ell+j}$, a conclusion which is a special case of more general results in prediction theory due to Wiener [29] and Whittle [30].

Then

$$\begin{aligned} Z_{t+\ell} &= (a_{t+\ell} + \psi_1 a_{t+\ell-1} + \dots + \psi_{\ell-1} a_{t+1}) + (\psi_\ell a_t + \psi_{\ell+1} a_{t-1} + \dots) \\ &= e_t(\ell) + \hat{Z}_t(\ell) \end{aligned} \quad (3.5.8)$$

where $e_t(\ell)$ is the error of the forecast $\hat{Z}_t(\ell)$ at time t .

Certain important facts emerge. As before denote $E[Z_{t+\ell} | Z_t, Z_{t-1}, \dots]$ the conditional expectation of $Z_{t+\ell}$ given knowledge of all the Z 's up to time t , by ${}_t Z_{t+\ell}$. Then

$$1) \quad \hat{Z}_t(\ell) = \psi_\ell a_t + \psi_{\ell+1} a_{t-1} + \dots \equiv {}_t Z_{t+\ell} \quad (3.5.9)$$

Thus, the minimum mean square error forecast at origin t , for lead time ℓ , is the conditional expectation of $E[Z_{t+\ell}]$, at time t . When

$\hat{Z}(\ell)$ is regarded as a function of ℓ for fixed t , it will be called the "forecast function" for origin t .

- 2) The forecast error for lead time ℓ is

$$e_t(\ell) = a_{t-\ell} + \psi_1 a_{t+\ell-1} + \dots + \psi_{\ell-1} a_{t-1} \quad (3.5.10)$$

since

$$E_t[e_t(\ell)] = 0 \quad (3.5.11)$$

the forecast is unbiased. Also, the variance of the forecast error is

$$V(\ell) = \text{var}[e_t(\ell)] = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{\ell-1}^2) \alpha_a^2 \quad (3.5.12)$$

- 3) It is readily shown that, not only is $\hat{Z}_t(\ell)$ the minimum mean square error forecast of $Z_{t+\ell}$, but that only linear function $\sum_{\ell=1}^L W_\ell \hat{Z}_t(\ell)$, of the forecasts, is a minimum square error forecast of the corresponding linear function $\sum_{\ell=1}^L W_\ell Z_{t+\ell}$ of the future observations. [For example, suppose that using eq. (3.5.9), we have obtained, from monthly data, minimum mean square error forecast $\hat{Z}_t(1)$, $\hat{Z}_t(2)$, and $\hat{Z}_t(3)$ of the sales of a product, one, two, and three months ahead. Then it is true that $\hat{Z}_t(1) + \hat{Z}_t(2) + \hat{Z}_t(3)$ is the minimum mean square error forecast of the sales $Z_{t+1} + Z_{t+2} + Z_{t+3}$ during the next quarter.]
- 4) The residuals are one step ahead forecast errors. Using eq. (3.5.10), the one step ahead forecast error is

$$e_t(1) = Z_{t+1} - \hat{Z}_t(1) = a_{t+1} \quad (3.5.13)$$

Hence, the residuals a_t which generate the process, are which so far it has been introduced merely as a set of independent random variables

or shocks, turn out to be the one step ahead forecast errors. [It follows that, for a minimum mean square error forecast, the one step ahead forecast errors must be uncorrelated. This is reasonable, for if one step ahead errors were correlated, then the forecast error a_{t+1} could, to some extent, be predicted from available forecast errors $a_t, a_{t-1}, a_{t-2}, \dots$. If the prediction so obtained was \hat{a}_{t+1} , then $\hat{Z}_t(1) + \hat{a}_{t+1}$ would be a better forecast of Z_{t+1} than was $\hat{Z}_t(1)$.]

- 5) Correlation between the forecast errors. Although the optimal forecast error at lead time 1 will be uncorrelated, the forecast errors for longer lead time in general will be correlated.

Now using the fact, that the minimum mean square error forecast $\hat{Z}_t(l)$ for lead time l is the conditional expectation $E_t[Z_{t+l}]$, of Z_{t+l} , at origin t , to write down expressions for the forecast in any one of three different ways, summarized earlier in this section. For $l > 0$, the three different ways of expressing the forecasts are:

- 1) Forecasts from difference equation. Taking conditional expectations at time t in eq. (3.5.2), we obtain

$$\begin{aligned} E_t[Z_{t+l}] &= \hat{Z}_t(l) \\ &= \varphi_1 E_t[Z_{t+l-1}] + \dots + \varphi_{p+d} E_t[Z_{t+l-p-d}] \\ &\quad - \theta_1 E_t[a_{t+l-1}] - \dots - \theta_q E_t[a_{t+l-q}] + E_t a_{t+l} \end{aligned} \quad (3.5.14)$$

- 2) Forecast in integrated form. Using eq. (3.5.3)

$$\begin{aligned} E_t[Z_{t+l}] &= \hat{Z}_t(l) \\ &= \psi_1 E_t[a_{t+l-1}] + \dots + \psi_{l-1} E_t[a_{t+1}] \\ &\quad + \psi_l E_t[a_t] + \psi_{l+1} E_t[a_{t-1}] + \dots + E_t[a_{t+l}] \end{aligned} \quad (3.5.15)$$

- 3) Forecast as weighted average of previous observations and forecasts made at previous lead times from the same origin. Taking conditional expectation in eq. (3.5.5)

$$\begin{aligned} E[Z_{t+l}] &= \hat{Z}_t(l) \\ &= \sum_{j=1}^{\infty} \pi_j E[Z_{t+l-j}] + E[a_{t+l}] \end{aligned} \quad (3.5.16)$$

It is to be noted that the minimum mean square error forecast is defined in term of the conditional expectation

$$E_t[Z_{t+l}] = E[Z_{t+l} | Z_t, Z_{t-1}, \dots]$$

which theoretically requires knowledge of the Z's stretching back into the infinite past. However, the requirement of invertibility, which we have imposed on the general ARIMA model, ensures that the π weights in eq. (3.5.16) from a convergent series. Hence, for the computation of forecast to a given degree of accuracy, for some k , the dependence on Z_{t-j} for $j > k$ can be ignored. In practice, the π weights usually decay rather quickly, so that whatever form of the model is employed in the computation, only a moderate length of series $Z_t, Z_{t-1}, \dots, Z_{t-k}$ is needed to calculate the forecasts to sufficient accuracy.

3.5.2 Calculation of the ψ Weights. Suppose forecast at lead time 1, 2, ..., L are required to obtain probability limits for these forecasts and also to allow now forecasts to be calculated by a process of updating the old, it is necessary to calculate the weights $\psi_1, \psi_2, \dots, \psi_{L-1}$. This is accomplished using eq. (3.3.4)

$$\varphi(B)\psi(B) = \theta(B)$$

$\psi_2 = - .42$, we obtain

$$\psi_0 = 1$$

$$\psi_1 = \varphi_1 - \theta_1 = 1.42 \quad \theta_1 = 0$$

$$\psi_2 = \varphi_1\psi_1 + \varphi_2\psi_0 - \theta_2 \quad \theta_2 = 0$$

$$= 1.42 \times 1.42 - .42 = 2.0164 - .42 = 1.5964$$

3.5.3 Calculation of the Probability Limit of the Forecasts. The expression in eq. (3.5.12) shows that, in general, the variance of the 1 steps ahead forecast error for any origin t is the expect value of

$$e_t^2(l) = \{Z_{t+l} - \hat{Z}_t(l)\}^2$$

and is given by

$$V(l) = \{1 + \sum_{j=1}^{l-1} \psi_j^2\} \sigma_a^2$$

Assuming that the a 's are Normal, it follows that, given information up to time t , the conditional probability distribution $P(Z_{t+l} | Z_t, Z_{t-1}, \dots)$ of a future value Z_{t+l} of the process will be Normal with mean $\hat{Z}_t(l)$ and standard deviation $\{1 + \sum_{j=1}^{l-1} \psi_j^2\}^{1/2} \sigma_a$. When the number of observation is, say, at least

50 and approximate $1 - \epsilon$ probability limits $Z_{t+l}(-)$ and $Z_{t+l}(+)$ for Z_{t+l} will be given by

$$Z_{t+l}(\pm) = \hat{Z}_t(l) \pm u_{\epsilon/2} \{1 + \sum_{j=1}^{l-1} \psi_j^2\}^{1/2} \sigma_a \quad (3.5.17)$$

where $u_{\epsilon/2}$ is the deviate exceed by a proportion $\epsilon/2$ of the Normal distribution,

For example for Dow Jones series, $S_a = .38$; hence the 50% and 95% limits, for $Z(2)$, are given by

50% limit:

$$\begin{aligned} Z_{t+1}(\pm) &= \hat{Z}_t(2) \pm .674(1 + 1.42^2)^{\frac{1}{2}}(.38) \\ &= \hat{Z}_t(2) \pm .4448 \end{aligned}$$

95% limit:

$$\begin{aligned} Z_{t+1}(\pm) &= \hat{Z}_t(2) \pm 1.96(1 + 1.42^2)^{\frac{1}{2}}(.38) \\ &= \hat{Z}_t(2) \pm 1.2936 \end{aligned}$$

CHAPTER IV

EMPIRICAL APPLICATIONS

The purpose of this chapter is to illustrate the use of the BOXJEN programs in constructing a stochastic model and forecasting algorithm. The series using in illustrations are:

- 1) G.M.P. (1939 - 1976) in billions of current dollars [22]
- 2) Dow Jones utility index [3]
- 3) Wolfer sunspot number [1]

The following three models are going to try in the identification phase for all the series used in this chapter.

Model	NDIFF1	NDIFF2	IVAG	KMAX	LMAX
1	0	0	1	12	12
2	1	0	0	12	12
3	1	1	0	12	12

Model 1: Calculate the statistical properties of the original data and fits an AR model of order 12.

Model 2: Calculate the statistical properties of the first difference.

Model 3: Calculate the statistical properties of the second difference.

These three models are typical of ones that can be test on many practical time series. All the results generated either by MAIN1 - Identification or by MAIN2 - Estimation are in computer printout form.

4.1 Gross National Product (G.M.P.)

ID phase: (p. 57 - 64)

Step 1: The autocorrelations of the original series dies off slowly indicate the original series is nonstationary. Use the standard deviation of the residuals as a criterion to determine the proper degree of differencing. In this series, Model 3 (p. 62 - 63) has the smallest standard deviation, indicates second order degree of differencing operation is suggested.

Step 2: The autocorrelations of the second differenced series (p. 57) are in damped sine wave, indicate an autoregressive model is suggested.

Step 3: The partial autocorrelations are insignificant after second lag indicate that second order autoregressive model is suggested.

Step 4: From step 1 - 3, we have a tentative ARIMA(2,2,0) model on hand. The parameters for this model are calculated in subroutine BJSTAT.

80XJFN PRELIMINARY IDENTIFICATION
INPUT DATA

1	90.80	2	100.00	3	124.90	4	158.30	5	192.50	6	210.50
7	212.30	8	209.60	9	232.80	10	259.10	11	258.00	12	286.20
13	330.20	14	347.20	15	366.10	16	366.30	17	399.30	18	420.70
19	442.80	20	448.90	21	486.50	22	506.00	23	523.30	24	563.00
25	594.70	26	635.70	27	688.10	28	753.00	29	796.30	30	849.50
31	935.50	32	982.40	33	1063.40	34	1171.10	35	1306.60	36	1412.90
37	1528.80	38	1706.50	39	1890.40						

BJDIFF

ITERM = 1
 LTERM = 39
 NDIFF1 = 0
 NDIFF2 = 0
 IAVG = 1
 AVG = 612.05

Model 1

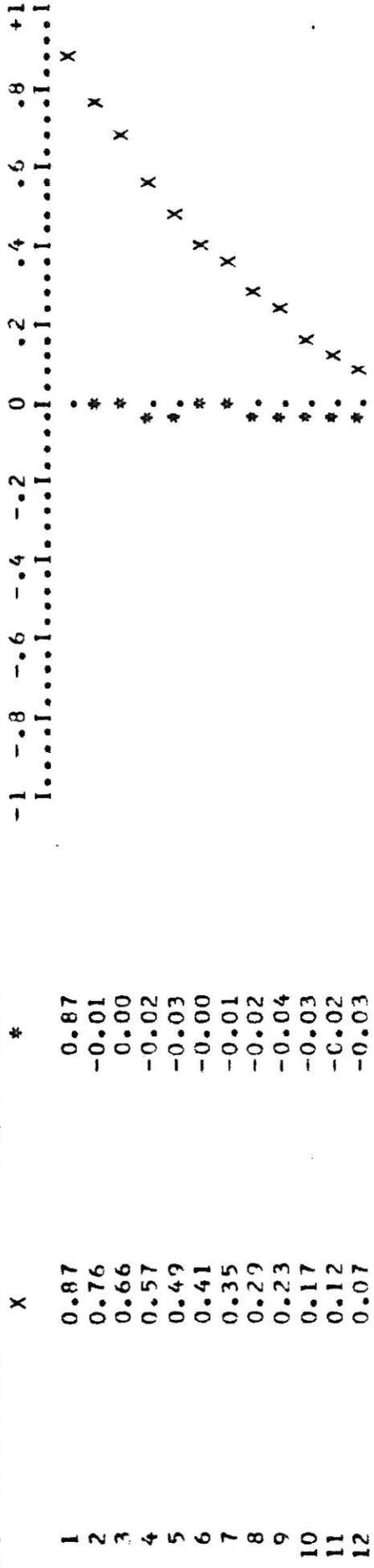
BJSTAT

INPUT PARAMETERS ITERM = 1
 LTERM = 39
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = -521.250
 LTERM = 1278.350

CALCULATED RESULTS MFAN = 0.00
 ABS. DEV. = 367.72
 VAR. = 211828.69
 STD. DEV. = 460.25

LAG AUTOCORRELATION PARTIAL ACORREL *



PARAMETERS OF AN AUTOREGRESSIVE MODEL

PARAMETER	LAG	VALUE
1	0.88	3
2	-0.02	10
3	0.02	0
4	-0.00	4
5	-0.02	12
6	0.00	0
7	0.01	7
8	0.01	0
9	-0.02	10
10	-0.00	11
11	0.00	12
12	-0.03	0

BJRES

BJSTAT

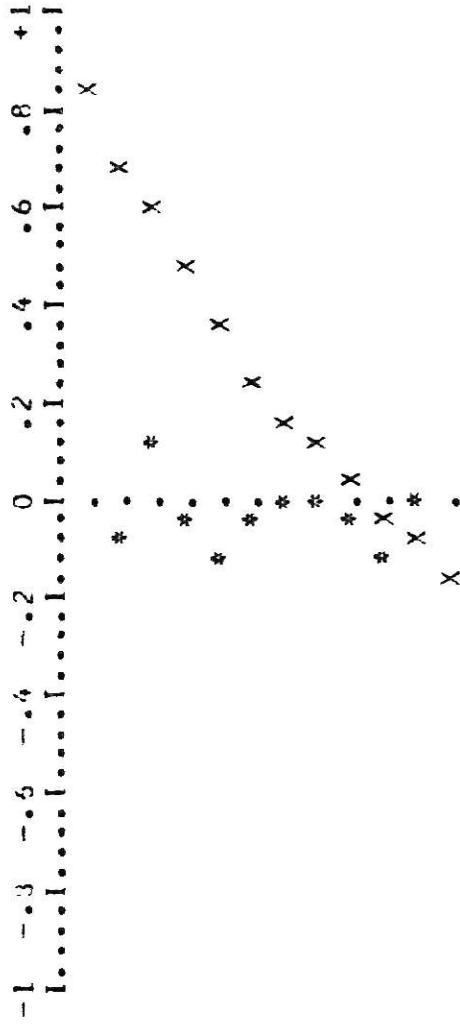
INPUT PARAMETERS ITERM = 13
 LTFRM = 39
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = -14.724
 LTERM = 318.669

CALCULATED RESULTS MEAN = 65.60
 ABS. DEV. = 81.04
 VAR. = 9968.82
 STD. DEV. = 99.84

LAG AUTOCORRELATION PARTIAL ACCORREL
 X

1	0.83		0.83
2	0.68		-0.06
3	0.58		0.11
4	0.48		-0.06
5	0.36		-0.14
6	0.25		-0.02
7	0.17		-0.02
8	0.11		0.02
9	0.05		-0.06
10	-0.03		-0.10
11	-0.08		0.00
12	-0.15		-0.14



PARAMFTERS OF AN AUTOREGRESSIVE MODEL

1	0.88	2	-0.16	3	0.14	4	0.06	5	-0.13	6	0.01	7	-0.04
8	0.06	9	0.06	10	-0.13	11	0.13	12	-0.14				

Model 2

BJDIFF

ITERM = 1
 LTERM = 39
 NDIFF1 = 1
 NDIFF2 = 0
 IAVG = 0
 AVG = 47.36

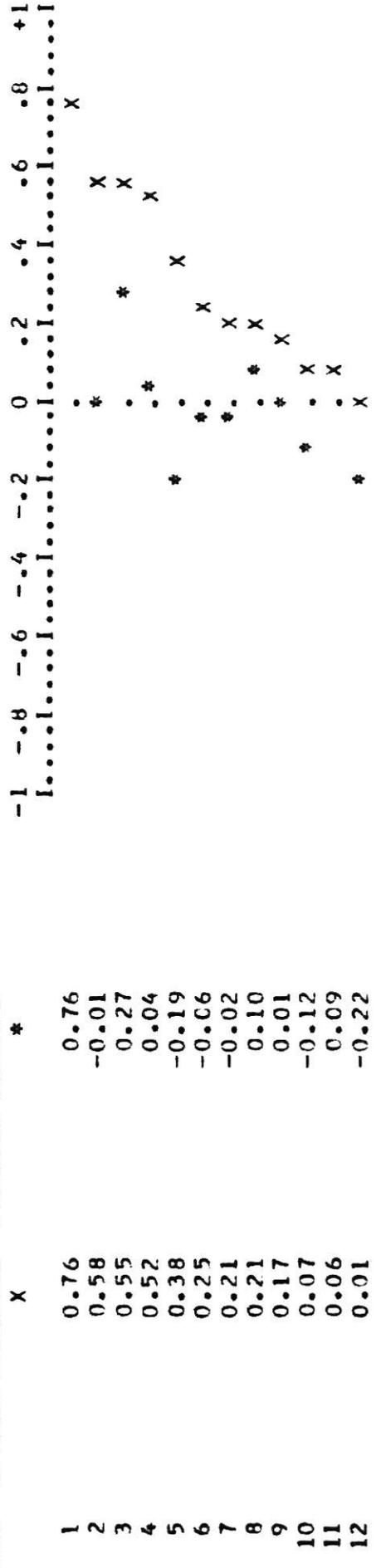
BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 39
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = 9.200
 LTERM = 183.900

CALCULATED RESULTS MEAN = 47.36
 ARS. DEV. = 33.87
 VAR. = 2066.66
 STD. DEV. = 45.46

LAG AUTOCORRELATION PARTIAL ACORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.79 2 -0.19 3 0.23 4 0.18 5 -0.20 6 0.01 7 -0.12
 8 0.09 9 0.16 10 -0.22 11 0.26 12 -0.22

BJDIFF

Model 3

ITERM = 1
 LTERM = 39
 NDIFF1 = 1
 NDIFF2 = 1
 IAVG = 0
 AVG = 4.72

RJSTAT

INPUT PARAMETERS ITERM = 3
 LTERM = 39
 KMAX = 12
 LMAX = 12

 INPUT SERIES VALUES ITERM = 15.700
 LTERM = 6.200

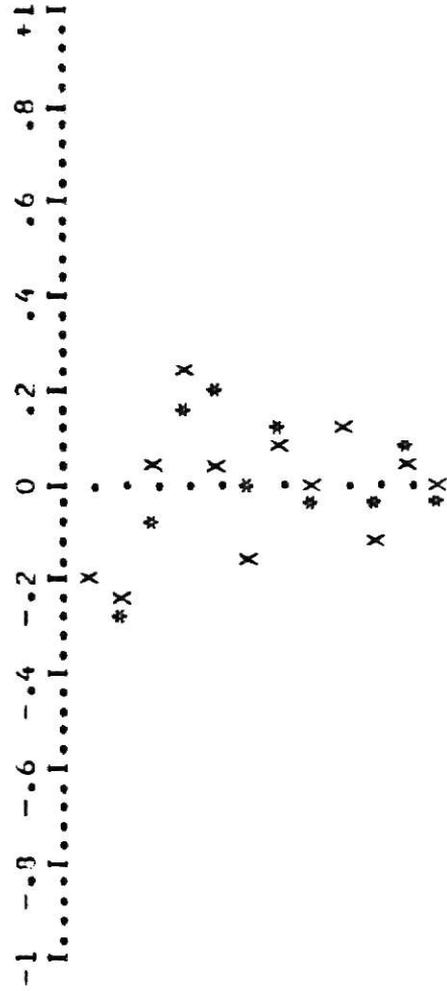
CALCULATED RESULTS MEAN = 4.72
 ABS. DEV. = 17.67
 VAR. = 450.14
 STD. DEV. = 21.22

LAG AUTOCORRELATION PARTIAL ACCORREL *

1	-0.21	-0.21
2	-0.24	-0.30
3	0.06	-0.08
4	0.22	0.16
5	0.05	0.19
6	-0.16	0.00
7	0.09	0.11
8	-0.00	-0.06
9	0.14	0.14
10	-0.12	-0.06
11	0.04	0.06
12	0.02	-0.04

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	-0.29	2	-0.30	3	0.02	4	0.19	5	0.20	6	0.02	7	C.12
8	-0.03	9	0.14	10	-0.05	11	0.05	12	-0.04				



BJRES

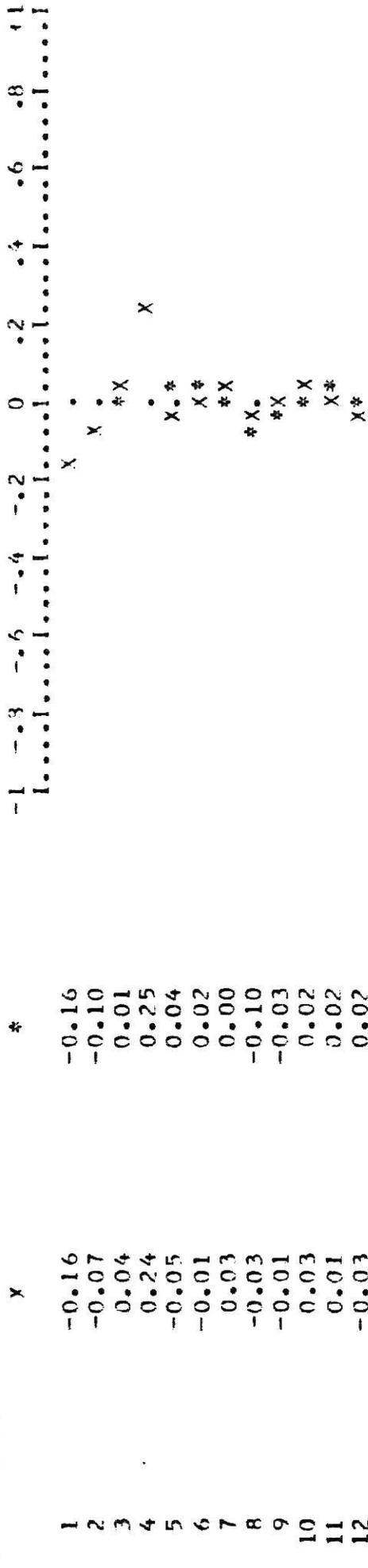
BJSTAT

INPUT PARAMETERS ITERM = 15
 LTERM = 39
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = 5.002
 LTERM = 13.093

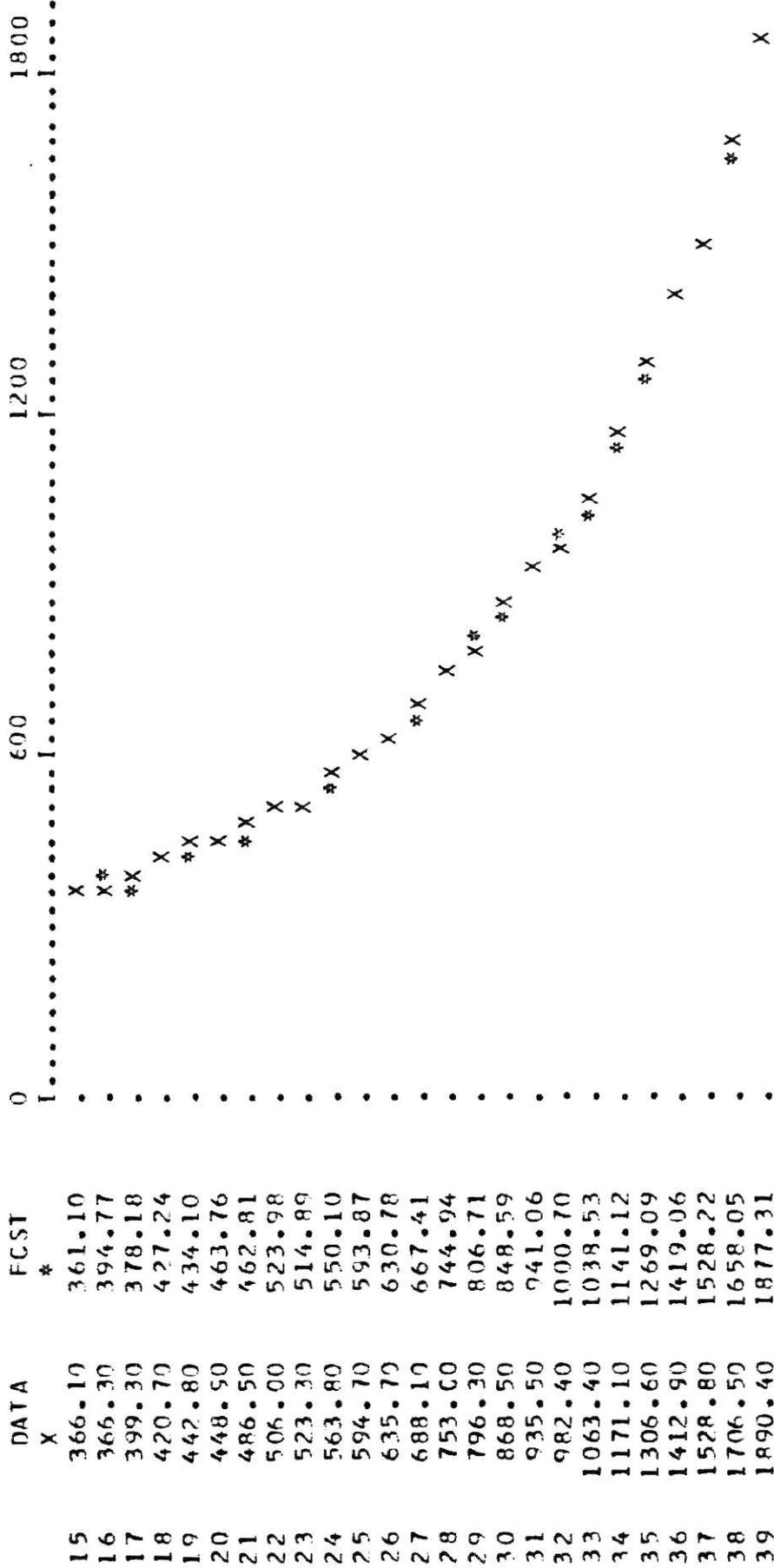
CALCULATED RESULTS MEAN = 7.25
 ABS. DEV. = 14.71
 VAR. = 332.79
 STD. DEV. = 18.24

LAG AUTOCORRELATION PARTIAL ACORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	-0.19	2	-0.07	3	0.06	4	0.29	5	0.05	6	0.01	7	-0.03
8	-0.11	9	-0.02	10	0.02	11	0.03	12	0.02				



Estimation phase: (p. 66 - 73) Estimate parameter values for ARIMA(2,2,0) model, and verify its goodness of fit.

Step 1: Initialization

BJDIFF1 = 1

BJDIFF2 = 1

I AVG = 0

KMAX = 12

LMAX = 2

NSRT1 = 1

NSRT2 = 1

T1MIN = 0

T1MAX = 0

T2MIN = 0

T2MAX = 0

LSEA = 1

Step 2: Since $N = 39$, the approximate upper bound for the standard error of a single autocorrelation is $1/\sqrt{39} = .16$. Then compare with this standard error bound, the autocorrelation of the residuals (p. 72) are rather small except for $\gamma_4(\hat{a}) = .26$, but taking these results as a whole, certainly, there is no lack of fit of the model.

Step 3: In the estimation phase the program also does forecasts. Print out the original and the forecast values.

BOXJEN PRELIMINARY ESTIMATION
INPUT DATA

1	90.80	2	100.00	2	124.90	4	153.30	5	192.50	6	210.50
7	212.30	8	209.60	9	232.80	10	259.10	11	258.00	12	286.20
13	330.20	14	347.20	15	366.10	16	366.30	17	399.30	18	420.70
19	442.80	20	448.90	21	486.50	22	506.00	23	523.30	24	563.10
25	594.70	26	635.70	27	688.10	28	753.00	29	796.30	30	868.50
31	935.50	32	982.40	33	1063.40	34	1171.10	35	1306.60	36	1412.90
37	1528.80	38	1706.50	39	1890.40						

BJDIFF

ITERM = 1
 LTERM = 39
 NDIFF1 = 1
 NDIFF2 = 1
 IAVG = 0
 AVG = 4.72

JMASR

NSRT1 = 1 T1MIN = 0.0 T1MAX = 0.0
 NSRT2 = 1 T2MIN = 0.0 T2MAX = 0.0

BJARMA

T1 = 0.0
T2 = 0.0
LSEA = 1
INTERM = 3

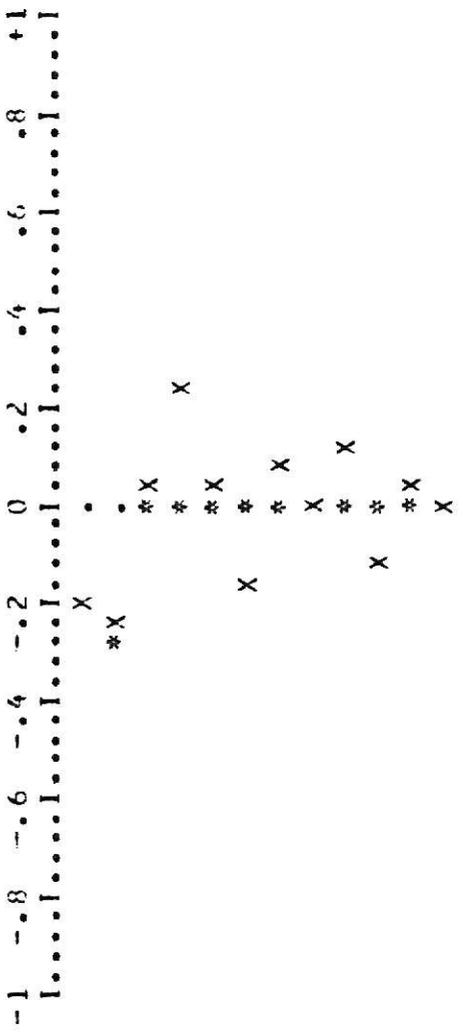
BJSTAT

INPUT PARAMETPS ITERM = 3
LTERM = 39
KMAX = 12
LMAX = 2

INPUT SERIES VALUES ITERM = 15.700
LTERM = 6.200

CALCULATED RESULTS MEAN = 4.72
ABS. DEV. = 17.67
VAR. = 450.14
STD. DEV. = 21.22

LAG AUTOCORRELATION PARTIAL ACORREL *



PARAMETPS OF AN AUTOREGRESSIVE MODEL

1 -0.27 2 -0.30

BJRES

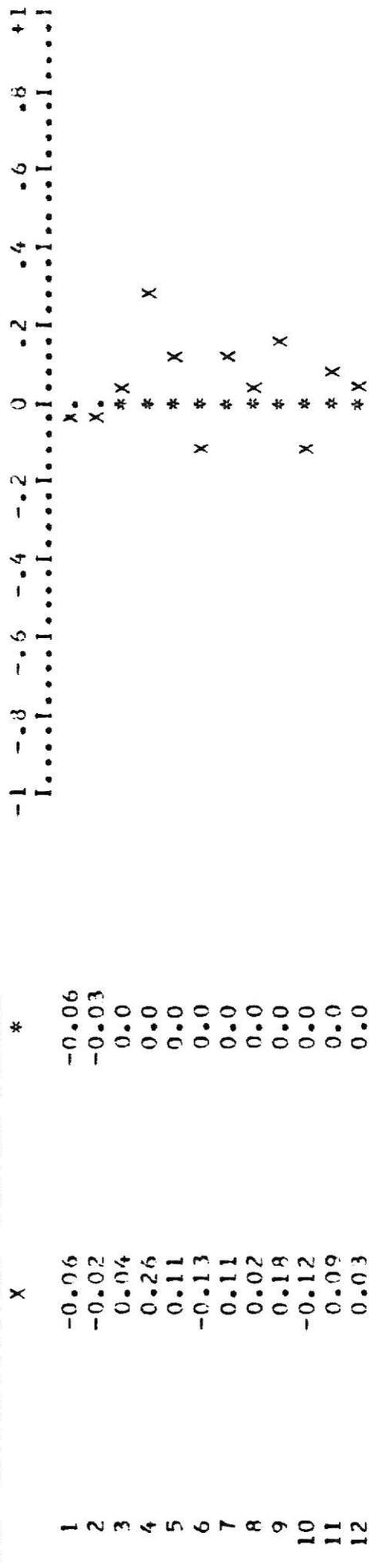
BJSTAT

INPUT PARAMETERS ITERM = 5
 LTERM = 39
 KMAX = 12
 LMAX = 2

INPUT SERIES VALUES ITERM = 7.833
 LTERM = 25.981

CALCULATED RESULTS MEAN = 6.41
 ABS. DEV. = 17.04
 VAR. = 399.33
 STD. DEV. = 19.98

LAG AUTOCORRELATION PARTIAL ACORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.06 2 -0.03

SEARCH NUMBER 1
 BT1 = 0.0 RT2 = 0.0 BRMS = 19.98 T1 = 0.0 T2 = 0.0 SIGMA = 19.98

SEARCH GRID

TABLE OF SIGMA VALUES

T2 PARAMETER VALUES

* 0.0

T1 PARAMETER

VALUES

0.0 * 19.98

BEST PARAMETERS

T1 = 0.0
T2 = 0.0
BRMS = 19.98

BJARMA

T1 = 0.0
T2 = 0.0
LSFA = 1
INTERM = 3

3JSTAT

INPUT PARAMFTERS ITERM = 3
 LTERM = 39
 KMAX = 12
 LMAX = 2

INPUT SERIES VALUES ITERM = 15.700
 LTERM = 6.200

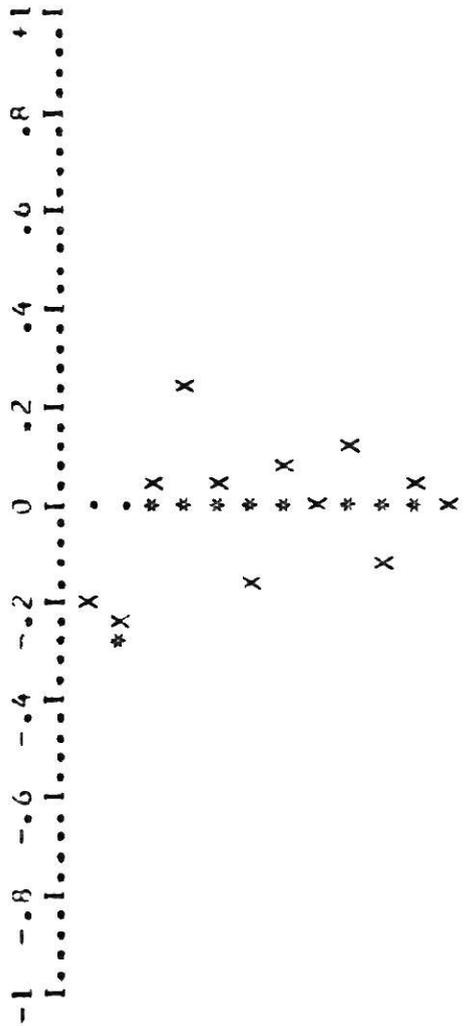
CALCULATED RESULTS MEAN = 4.72
 ABS. DEV. = 17.67
 VAR. = 450.14
 SID. DEV. = 21.22

LAG AUTOCORRELATION PARTIAL ACORREL *

1	-0.21		-0.21
2	-0.24		-0.30
3	0.06		0.0
4	0.22		0.0
5	0.05		0.0
6	-0.16		0.0
7	0.09		0.0
8	-0.07		0.0
9	0.14		0.0
10	-0.12		0.0
11	0.04		0.0
12	0.02		0.0

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.27 2 -0.30



RJRES

RJSTAT

INPUT PARAMETERS ITERM = 5
 LTERM = 39
 KMAX = 12
 LMAX = 2

INPUT SERIES VALUES ITERM = 7.833
 LTERM = 25.981

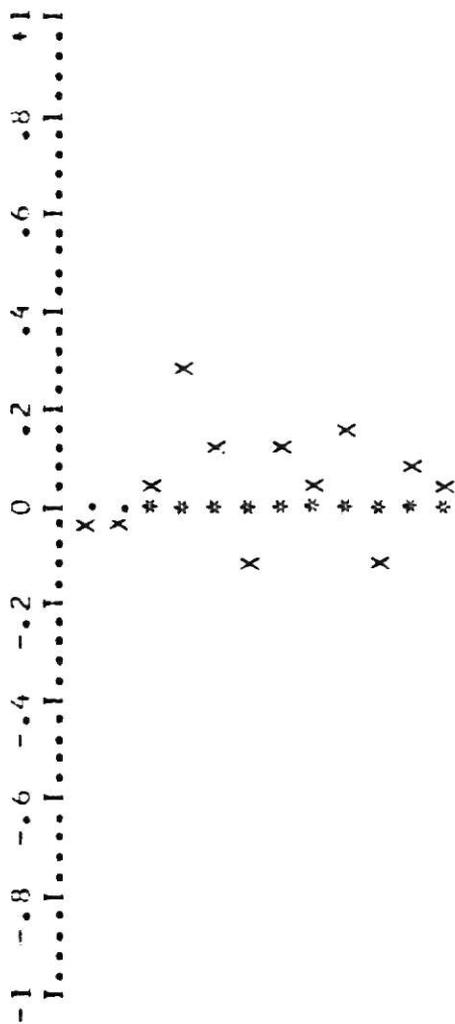
CALCULATED RESULTS MEAN = 6.41
 ABS. DEV. = 17.04
 VAR. = 399.33
 STD. DEV. = 19.98

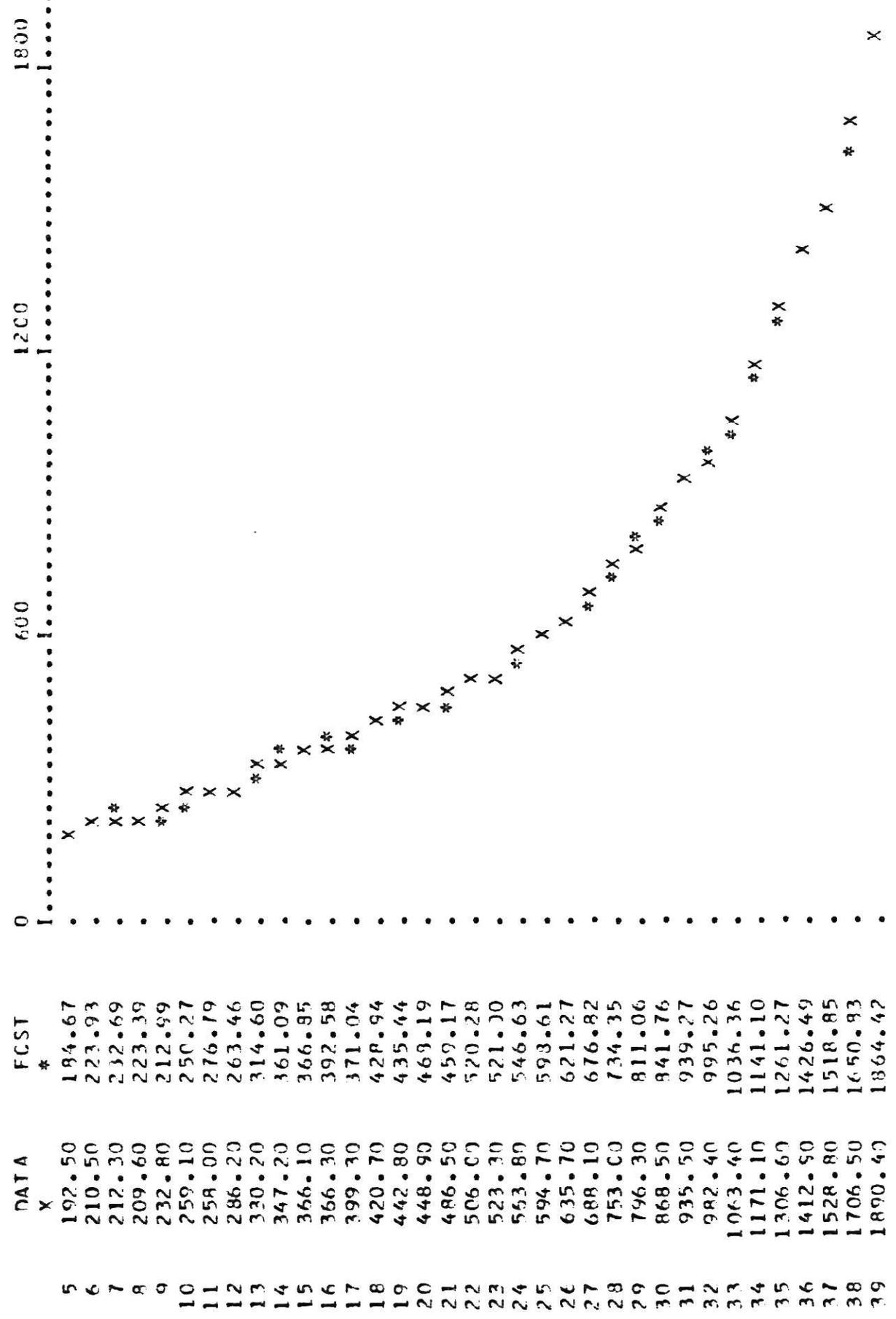
LAG AUTOCORRELATION PARTIAL ACORREL *

LAG	AUTOCORRELATION	PARTIAL	ACORREL	*
1	-0.06	X	-0.06	
2	-0.02		-0.03	
3	0.04		0.0	
4	0.26		0.0	
5	0.11		0.0	
6	-0.13		0.0	
7	0.11		0.0	
8	0.02		0.0	
9	0.18		0.0	
10	-0.12		0.0	
11	0.09		0.0	
12	0.03		0.0	

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.06 2 -0.03





4.2 Dow Jones Utility Index (8/28/72 - 12/18/72)

Identification phase: (p. 75 - 83)

- Step 1: The autocorrelations of the original series (Model 1, p. 76 - 77) decays very slowly indicate the original series is nonstationary. Go through model 2 (p. 78 - 79) and 3 (p. 80 - 81) indicate model 2 has the smaller standard deviation. There, a first order degree of differencing operation is suggested.
- Step 2: The autocorrelations of the first differenced series, there are in mixed damped exponential and damped sine waves indicate that a autoregressive model or a mixed ARMA model is suggested.
- Step 3: The partial autocorrelations after first lag are insignificant indicate that first order autoregressive model is suggested. Also, for the mixed model, the autocorrelations are insignificant after first lag indicate that a first moving average model is suggested.
- Step 4: From step 1 - 3, we have tentative ARIMA(1,1,0) and ARIMA(1,1,1) models on hand.

BOXJEN PRELIMINARY IDENTIFICATION
INPUT DATA

1	110.54	2	110.69	3	110.43	4	110.56	5	110.75	6	110.84
7	110.46	8	110.56	9	110.46	10	110.65	11	109.60	12	109.31
13	109.31	14	109.25	15	109.02	16	108.54	17	108.77	18	109.02
19	109.44	20	109.33	21	109.53	22	109.89	23	110.56	24	110.56
25	110.72	26	111.23	27	111.48	28	111.58	29	111.90	30	112.19
31	112.06	32	111.96	33	111.68	34	111.36	35	111.42	36	112.00
37	112.22	38	112.70	39	113.15	40	114.36	41	114.65	42	115.06
43	115.86	44	116.40	45	116.44	46	116.88	47	113.07	48	113.51
49	119.28	50	119.79	51	119.70	52	119.28	53	119.66	54	120.14
55	120.97	56	121.13	57	121.55	58	121.56	59	122.25	60	123.79
61	124.11	62	124.14	63	123.37	64	123.02	65	122.86	66	123.02
67	123.11	68	123.05	69	123.05	70	122.83	71	123.18	72	122.67
73	122.73	74	122.86	75	122.67	76	122.05	77	122.00	78	121.23

Model 1

BJDIFF

ITRM = 1
 LTERM = 78
 NDIFF1 = 0
 NDIFF2 = 0
 IAVG = 1
 AVG = 115.68

BJSTAT

INPUT PARAMETERS ITERM = 1
 LTERM = 78
 KMAX = 12
 LMAX = 12
 INPUT SERIES VALUES ITERM = -4.742
 LTERM = 5.548

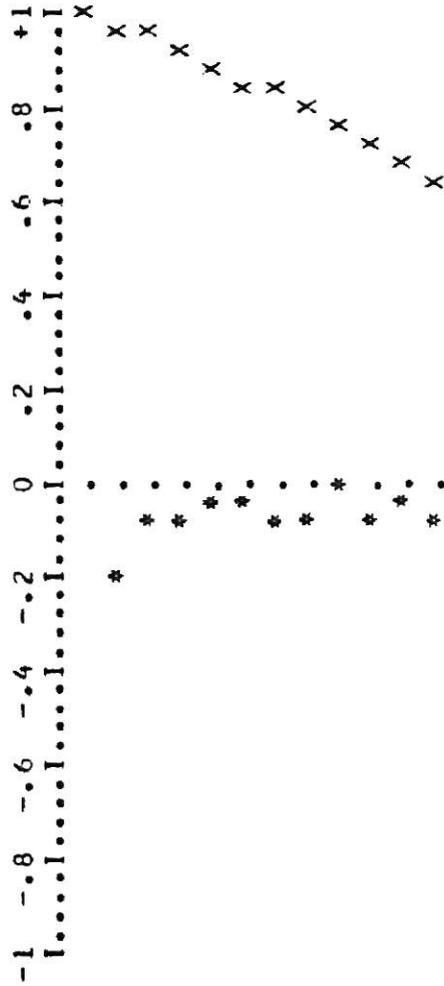
CALCULATED RESULTS MEAN = 0.00
 ABS. DEV. = 5.10
 VAR. = 29.93
 STD. DEV. = 5.47

LAG AUTOCORRELATION PARTIAL ACCORREL *
 X

1	0.99	0.99
2	0.97	-0.20
3	0.94	-0.08
4	0.92	-0.08
5	0.89	-0.06
6	0.86	-0.02
7	0.83	-0.09
8	0.79	-0.08
9	0.76	0.01
10	0.72	-0.09
11	0.68	-0.06
12	0.64	-0.09

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	1.13	2	-0.11	3	0.00	4	-0.02	5	-0.04	6	0.07	7	0.01
8	-0.11	9	0.11	10	-0.03	11	0.04	12	-0.04				



BJRES

BJSTAT

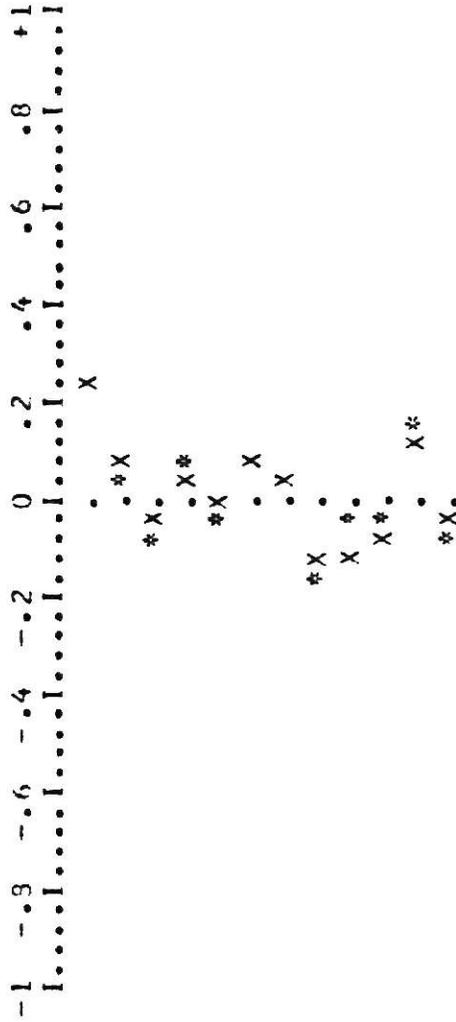
INPUT PARAMETERS ITERM = 13
 LTERM = 78
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = -0.054
 LTERM = -0.512

CALCULATED RESULTS MEAN = 0.08
 ABS. DEV. = 0.29
 VAR. = 0.15
 STD. DEV. = 0.38

LAG AUTOCORRELATION PARTIAL ACOPPEL
 X

1	0.23		0.23
2	0.10		0.04
3	-0.03		-0.07
4	0.04		0.06
5	-0.01		-0.03
6	0.06		0.07
7	0.06		0.04
8	-0.13		-0.17
9	-0.12		-0.05
10	-0.09		-0.03
11	0.13		0.17
12	-0.03		-0.10



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	0.25	2	0.06	3	-0.06	4	0.05	5	-0.06	6	0.03	7	0.06
8	-0.14	9	-0.06	10	-0.06	11	0.19	12	-0.10				

Model 2

BJDIFF

ITRM = 1
 LTERM = 78
 NDIFF1 = 1
 NDIFF2 = 0
 IAVG = 0
 AVG = 0.13

BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 12
 INPUT SERIES VALUES ITERM = -0.250
 LTERM = -0.770

CALCULATED RESULTS MEAN = 0.13
 ABS. DEV. = 0.33
 VAR. = 0.18
 STD. DEV. = 0.42

LAG AUTOCORRELATION PARTIAL ACORREL

LAG	AUTOCORRELATION	PARTIAL	ACORREL
1	0.42		0.42
2	0.27		0.11
3	0.16		0.01
4	0.23		0.16
5	0.15		-0.01
6	0.20		0.11
7	0.17		0.05
8	0.03		-0.15
9	0.04		0.04
10	0.05		0.01
11	0.18		0.15
12	0.01		-0.13

PARAMETERS OF AN AUTOREGRESSIVE MODEL	1	2	3	4	5	6	7	0.06	0.13	0.2	0.4	0.6	0.8	+1
1	0.40													
8	-0.14	9	0.02	10	-0.03	11	0.20	12	-0.13					

BJDIFF

ITERM = 1
LTERM = 78
NDIFF1 = 1
NDIFF2 = 1
I AVG = 0
AVG = -0.01

Model 3

BJSTAT

INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = -0.010
 LTERM = -0.680

CALCULATED RESULTS MEAN = -0.01
 ABS. DEV. = 0.35
 VAR. = 0.20
 STD. DEV. = 0.45

LAG AUTOCORRELATION PARTIAL ACORREL *

LAG	AUTOCORRELATION	PARTIAL	ACORREL	*
1	-0.35			-0.35
2	-0.06			-0.21
3	-0.14			-0.28
4	0.15			-0.05
5	-0.14			-0.18
6	0.05			-0.11
7	0.15			0.15
8	-0.19			-0.13
9	0.01			-0.06
10	-0.10			-0.19
11	0.28			0.11
12	-0.31			-0.22

LAG	AUTOCORRELATION	PARTIAL	ACORREL	*
-1	1.000			
0	1.000			
1	0.35			
2	0.21			
3	0.28			
4	0.05			
5	0.18			
6	0.11			
7	0.15			
8	0.13			
9	0.06			
10	0.19			
11	0.11			
12	0.22			

PARAMETERS OF AN AUTOREGRESSIVE MODEL
1 -0.46 2 -0.42 3 -0.39
8 -0.22 9 -0.18 10 -0.21

4 -0.21 5 -0.26 6 -0.12 7 -0.06
11 0.01 12 -0.22

BJRES

BJSTAT

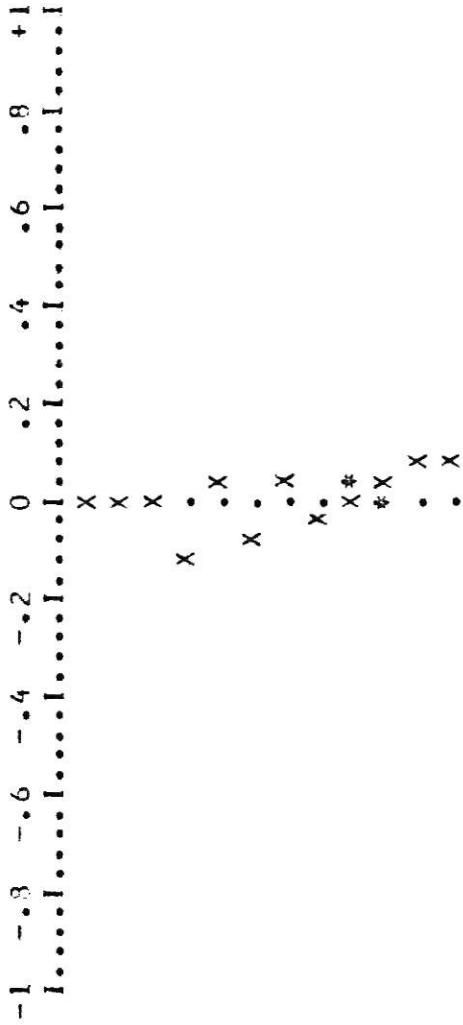
INPUT PARAMETEPTS ITERM = 15
 LTFRM = 78
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = -0.217
 LTERM = -0.648

CALCULATED RESULTS MEAN = -0.01
 ABS. DEV. = 0.28
 VAR. = 0.13
 STD. DEV. = 0.36

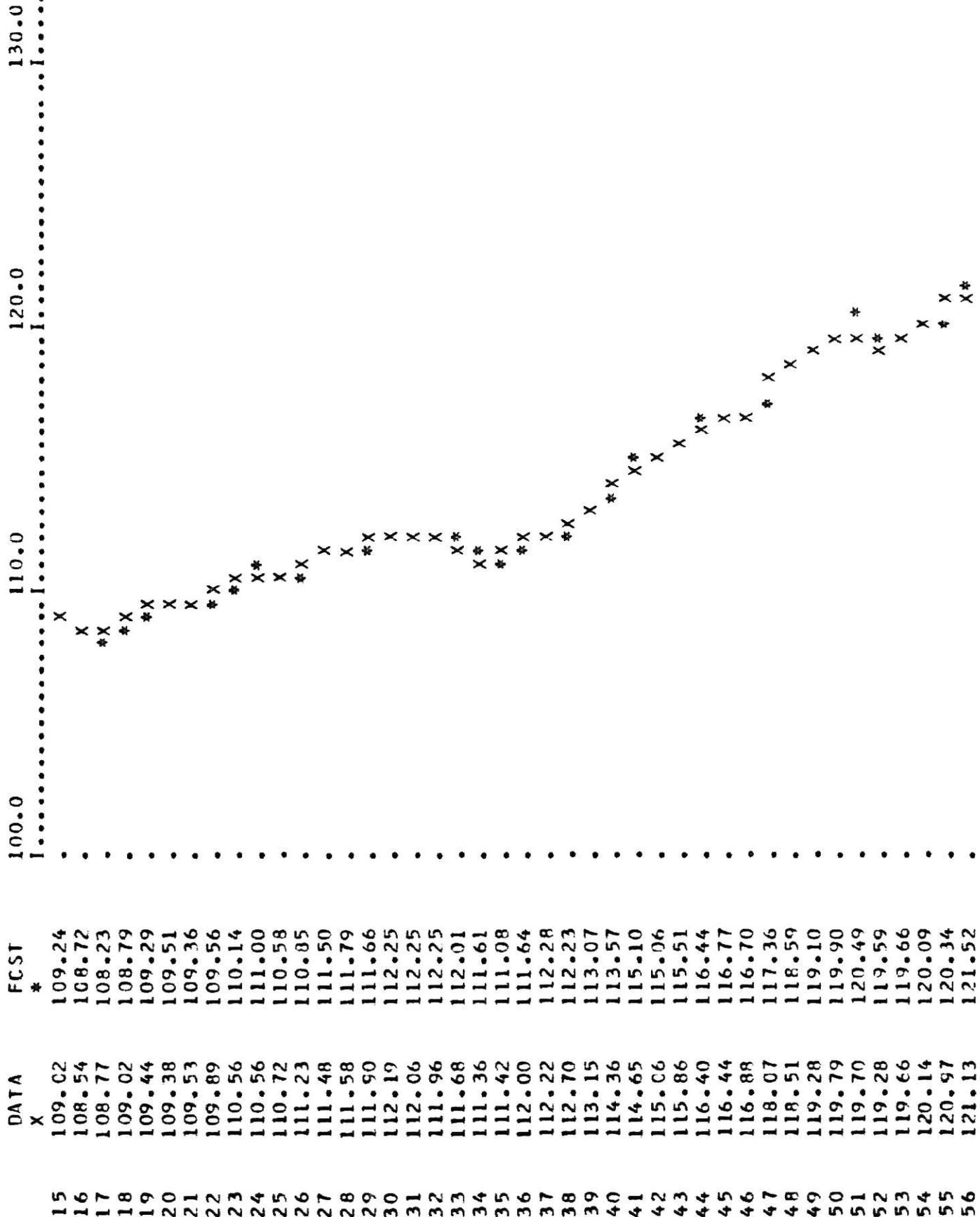
LAG AUTOCORRELATION PARTIAL ACORREL
 X

1	-0.01		-0.01
2	-0.01		-0.01
3	-0.00		-0.00
4	-0.12		-0.12
5	0.05		0.05
6	-0.07		-0.07
7	0.02		0.03
8	-0.02		-0.04
9	0.01		0.02
10	0.03		0.01
11	0.08		0.10
12	0.08		0.07



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	-0.00	2	-0.03	3	0.01	4	-0.13	5	0.05	6	-0.07	7	0.03
8	-0.03	9	0.03	10	0.01	11	0.09	12	0.07				



57	121.55	121.64		*X
58	121.96	121.91		X
59	122.26	122.29		X
60	123.79	122.91		* X
61	124.11	124.81		X*
62	124.14	124.47		X
63	123.37	124.22		X #
64	123.02	123.47		X
65	122.86	122.68		X
66	123.02	123.06		X
67	123.11	123.00		X
68	123.05	123.15		X
69	123.05	122.94		*X
70	122.83	123.02		X*
71	123.18	123.15		X
72	122.67	123.17		X*
73	122.73	122.74		X
74	122.86	122.58		X
75	122.67	123.09		X*
76	122.09	122.31		X
77	122.00	121.84		*X
78	121.23	121.88		X*

Estimation phase: (p. 85 - 93) Since ARIMA(1,1,0) has less parameters than ARIMA(1,1,1), we will try ARIMA(1,1,0) first. If there is doubt in ARIMA(1,1,0), then we will try ARIMA(1,1,1) again.

Step 1: Parameters initialization

NDIFF1 = 1

NDIFF2 = 0

IAVG = 0

KMAX = 12

LMAX = 1

NSRT1 = NSRT2 = 1

T1MIN = T1MAX = T2MIN = T2MAX = 0

Lsea = 0

Step 2: Since $N = 78$, the approximate upper bound for the standard error of a single autocorrelation is $1/\sqrt{78} = .12$. Compare with this standard error bound, the values $\gamma_4(\hat{a}) = .18$, $\gamma_7(\hat{a}) = .15$, $\gamma_{11}(\hat{a}) = .23$, and $\gamma_{12}(\hat{a}) = -.18$ (see p. 91) are all rather large. Of course, occasional large deviations occur even in random series, but taking these result as a whole, there is some doubt as to the adequacy of this model.

Step 3: Since model ARIMA(1,1,1) is also suggested in the identification phase, the iterative step will skip the identification phase, and directly get into estimation phase.

BOXJEN PRELIMINARY ESTIMATION
INPUT DATA

1	110.54	2	110.69	3	110.43	4	110.56	5	110.75	6	110.84
7	110.46	8	110.56	9	110.46	10	110.65	11	109.60	12	109.31
13	109.31	14	109.25	15	109.02	16	108.54	17	108.77	18	109.02
19	109.44	20	109.38	21	109.53	22	109.95	23	110.56	24	110.56
25	110.72	26	111.23	27	111.48	28	111.58	29	111.90	30	112.15
31	112.06	32	111.96	33	111.68	34	111.36	35	111.42	36	112.00
37	112.22	38	112.70	39	113.15	40	114.36	41	114.65	42	115.06
43	115.86	44	116.40	45	116.44	46	116.88	47	118.07	48	118.51
49	119.28	50	119.79	51	119.70	52	119.28	53	119.66	54	120.14
55	120.57	56	121.13	57	121.55	58	121.96	59	122.26	60	123.75
61	124.11	62	124.14	63	123.37	64	123.02	65	122.86	66	123.02
67	123.11	68	123.05	69	123.05	70	122.83	71	123.18	72	122.67
73	122.73	74	122.86	75	122.67	76	122.09	77	122.00	78	121.23

BJDIFF

ITRM = 1
LTERM = 78
NDIFF1 = 1
NDIFF2 = 0
IAVG = 0
AVG = 0.13

JMASR

NSRT1 = 1 T1MIN = 0.0 T1MAX = 0.0
NSRT2 = 1 T2MIN = 0.0 T2MAX = 0.0

BJARMA

T1 = 0.0
T2 = 0.0
LSEA = 0
INTERM= 2

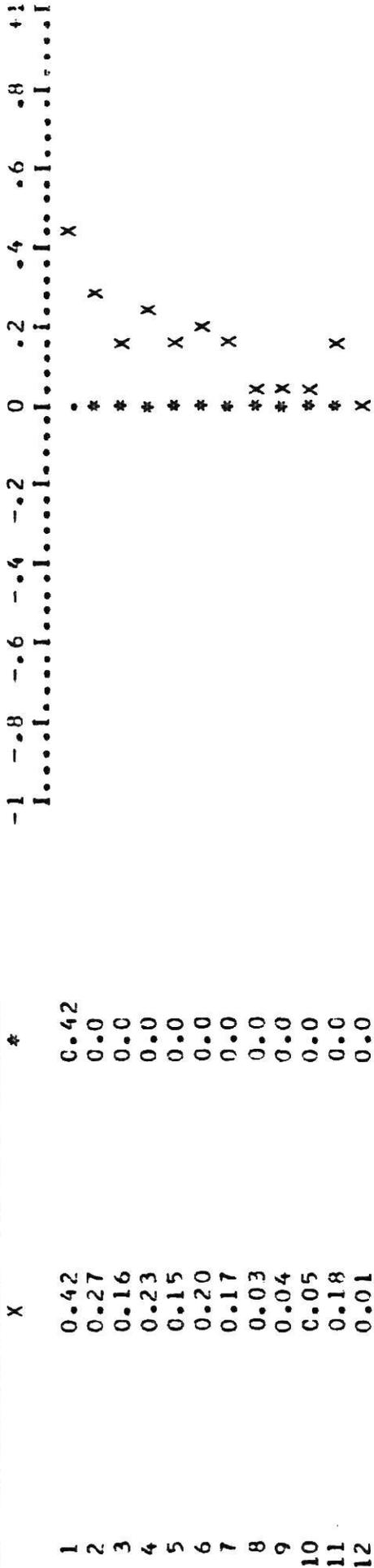
BJSTAT

INPUT PARAMETERS ITERM = 2
LTERM = 78
KMAX = 12
LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
LTERM = -0.770

CALCULATED RESULTS MEAN = 0.13
ABS. DEV. = 0.33
VAR. = 0.18
STD. DEV. = 0.42

LAG AUTOCORRELATION PARTIAL ACCORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.42

BJRES

BJSTAT

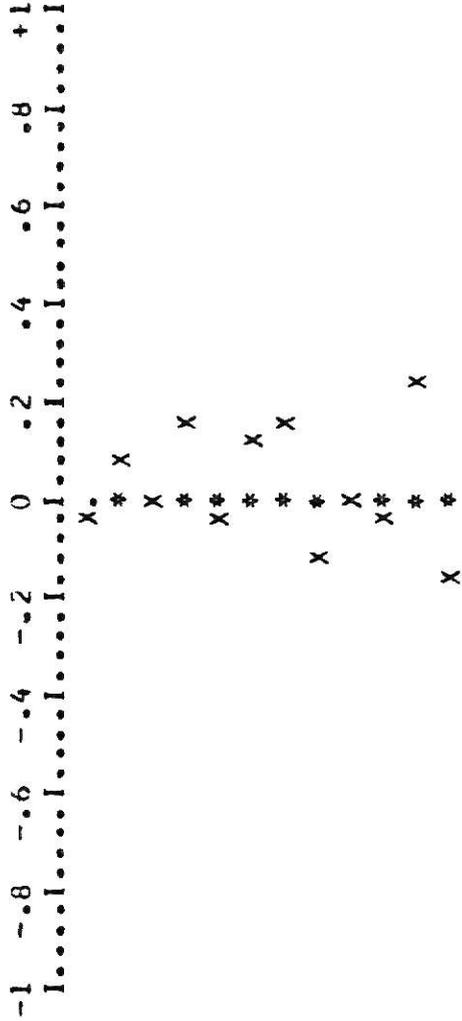
INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.155
 LTERM = -0.732

CALCULATED RESULTS MEAN = 0.08
 ABS. DEV. = 0.30
 VAR. = 0.15
 STD. DEV. = 0.38

LAG AUTOCORRELATION PARTIAL ACORREL
 X

1	-0.03		-0.03
2	0.09		0.0
3	-0.00		0.0
4	0.18		0.0
5	-0.02		0.0
6	0.13		0.0
7	0.15		0.0
8	-0.10		0.0
9	0.01		0.0
10	-0.04		0.0
11	0.23		0.0
12	-0.18		0.0



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.03

SEARCH NUMBER 1
 RT1 = 0.0 BT2 = 0.0 DRMS = 0.38 T1 = 0.0 T2 = 0.0 SIGMA = 0.38

SEARCH GRID

TABLE OF SIGMA VALUES

T2 PARAMETER VALUES

* 0.0

T1 PARAMETER

VALUES

0.0 * 0.38

BEST PARAMETERS

T1 = C.0
T2 = C.0
BRMS = C.38

BJARMA

I1 = 0.0
 I2 = 0.0
 LSEA = 0
 IINTERM = 2

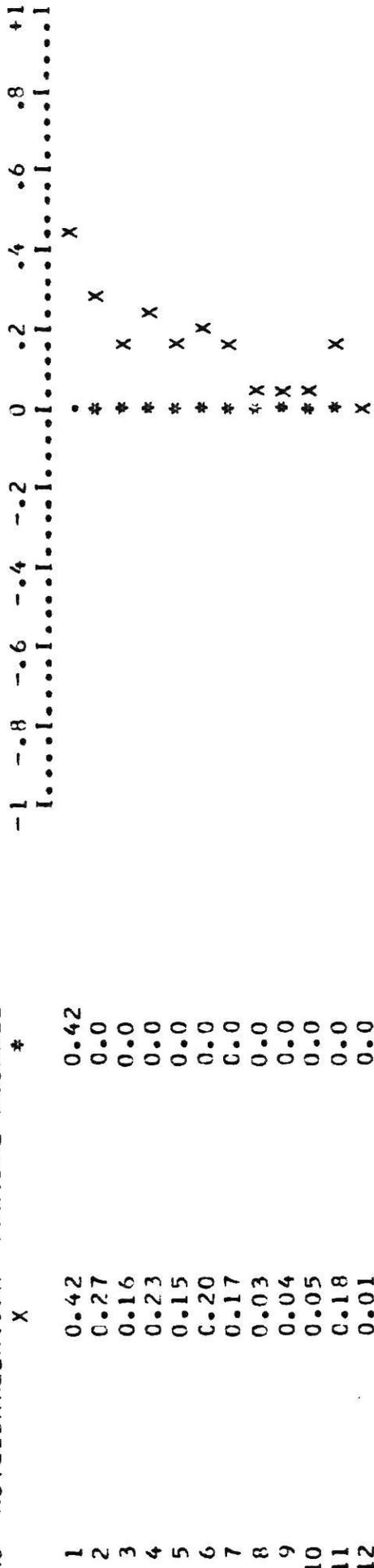
BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -0.770

CALCULATED RESULTS MEAN = 0.13
 ABS. DEV. = 0.33
 VAR. = 0.18
 STD. DEV. = 0.42

LAG AUTOCORRELATION PARTIAL ACCORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.42

BJRES

BJSTAT

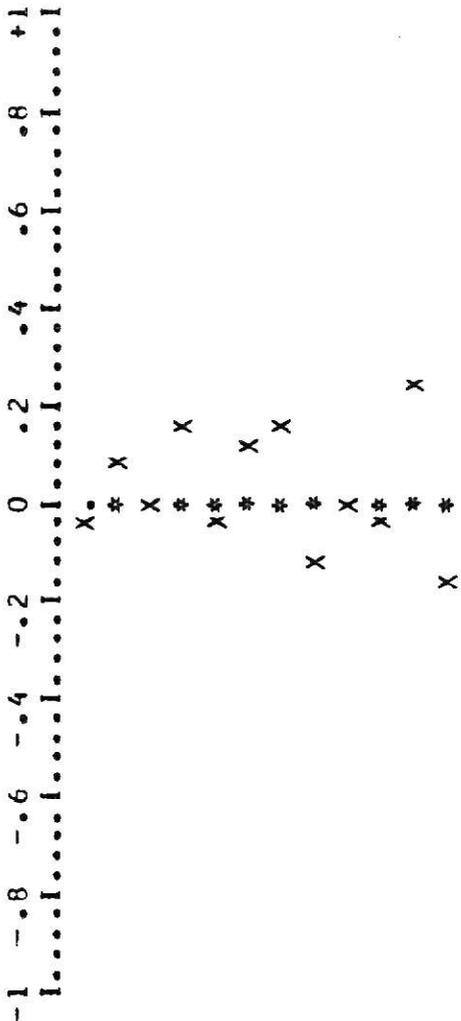
INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.155
 LTERM = -0.732

CALCULATED RESULTS MEAN = 0.08
 ABS. DEV. = 0.30
 VAR. = 0.15
 STD. DEV. = 0.38

LAG AUTOCORRELATION PARTIAL ACORREL
 X

1	-0.03		-0.03
2	0.09		0.0
3	-0.00		0.0
4	0.18		0.0
5	-0.02		0.0
6	0.13		0.0
7	0.15		0.0
8	-0.10		0.0
9	0.01		0.0
10	-0.04		0.0
11	0.23		0.0
12	-0.18		0.0



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.03

	DATA	FCST	100.0	110.0	120.0	130.0
3	X 110.43	* 110.58				
4	110.56	110.32		X		
5	110.75	110.61		*X		
6	110.84	110.83		X		
7	110.46	110.88		X*		
8	110.56	110.30		*X		
9	110.46	110.60		X		
10	110.05	110.42		X*		
11	109.60	109.88		X*		
12	109.31	109.41		X*		
13	109.21	109.19		X		
14	109.25	109.31		X		
15	109.02	109.22		X		
16	108.54	108.92		X*		
17	108.77	108.34		X		
18	109.02	108.87		X		
19	109.44	109.13		*X		
20	109.38	109.62		X		
21	109.53	109.35		X		
22	109.89	109.59		*X		
23	110.56	110.04		*X		
24	110.56	110.84		X		
25	110.72	110.56		X		
26	111.23	110.79		*X		
27	111.48	111.45		X		
28	111.58	111.59		X		
29	111.50	111.62		*X		
30	112.19	112.03		X		
31	112.06	112.31		X		
32	111.56	112.01		X		
33	111.68	111.92		X*		
34	111.36	111.56		X*		
35	111.42	111.22		*X		
36	112.00	111.45		*X		
37	112.22	112.24		X		
38	112.70	112.31		*X		
39	113.15	112.90		X		
40	114.36	113.34		*		
41	114.65	114.87		X		
42	115.06	114.77		X		
43	115.86	115.23		*X		
44	116.40	116.20		*X		

Estimation phase: (p. 95 - 119) Because of the doubt of adequacy to the model ARIMA(1,1,0), we then, try the model ARIMA(1,1,1).

Step 1: Parameters initialization

NDIFF1 = 1
NDIFF2 = 0
IAVG = 0
KMAX = 12
LMAX = 1
NSRT1 = 9
NSRT2 = 1
T1MIN = .1
T1MAX = .9
T2MIN = T2MAX = 0
LSEA = 0

Step 2: Since $N = 78$, the approximate upper bound for the standard error is $1/\sqrt{78} \approx .12$. Compare with standard error, the values $\gamma_{11}(\hat{a}) = .19$ (see p. 117) is the only rather larger autocorrelation exists.

Step 3: The autocorrelations of residuals are rather smaller than model ARIMA(1,1,0), therefore, we conclude that model ARIMA(1,1,1) is better than model ARIMA(1,1,0).

BOXJEN PRELIMINARY ESTIMATION
INPUT DATA

1	110.94	2	110.69	3	110.43	4	110.56	5	110.75	6	110.84
7	110.46	8	110.56	9	110.46	10	110.05	11	109.60	12	109.31
13	109.31	14	109.25	15	109.02	16	108.54	17	108.77	18	109.02
19	109.44	20	109.38	21	109.53	22	109.89	23	110.56	24	110.56
25	110.72	26	111.23	27	111.48	28	111.58	29	111.90	30	112.19
31	112.06	32	111.96	33	111.68	34	111.36	35	111.42	36	112.00
37	112.22	38	112.70	39	113.15	40	114.26	41	114.65	42	115.06
43	115.66	44	116.40	45	116.44	46	116.88	47	118.07	48	118.51
49	119.28	50	119.79	51	119.70	52	119.28	53	119.66	54	120.14
55	120.97	56	121.13	57	121.55	58	121.56	59	122.26	60	123.75
61	124.11	62	124.14	63	123.37	64	123.02	65	122.86	66	123.02
67	123.11	68	123.05	69	123.05	70	122.83	71	123.18	72	122.67
73	122.73	74	122.86	75	122.67	76	122.09	77	122.00	78	121.23

BJDIFF

I	TERM	=	1
L	TERM	=	78
N	DIFF1	=	1
N	DIFF2	=	0
I	AVG	=	0
A	AVG	=	0.13

JMASR

N	SRT1	=	9	T1	MIN	=	-0.90	T1	MAX	=	0.90
N	SRT2	=	1	T2	MIN	=	0.0	T2	MAX	=	0.0

BJARMA

T1 = -0.90
 T2 = 0.0
 LSFA = 0
 INTERM = 2

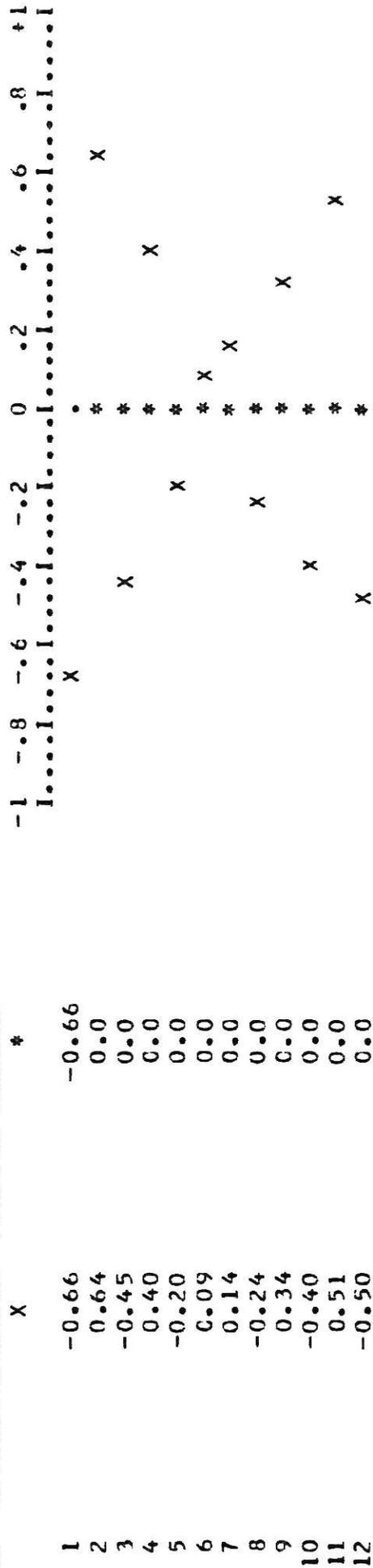
BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -1.313

CALCULATED RESULTS MEAN = 0.06
 ABS. DEV. = 0.47
 VAR. = 0.32
 STD. DEV. = 0.57

LAG AUTOCORRELATION PARTIAL ACORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.66

BJRES

BJSTAT

INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SFRIFS VALUES ITERM = -0.201
 LTERM = -0.913

CALCULATED RESULTS MEAN = 0.12
 ABS. DEV. = 0.32
 VAR. = 0.17
 STD. DEV. = 0.41

LAG AUTOCORRELATION PARTIAL ACCORFL *



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.18

SEARCH NUMBER 1
 BT1 = -0.90 BT2 = 0.0 BRMS = 0.41 T1 = -0.90 T2 = 0.0 SIGMA = 0.41

BJARMA

T1 = -0.68
 T2 = 0.0
 LSEA = 0
 INTERM = 2

BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -0.975

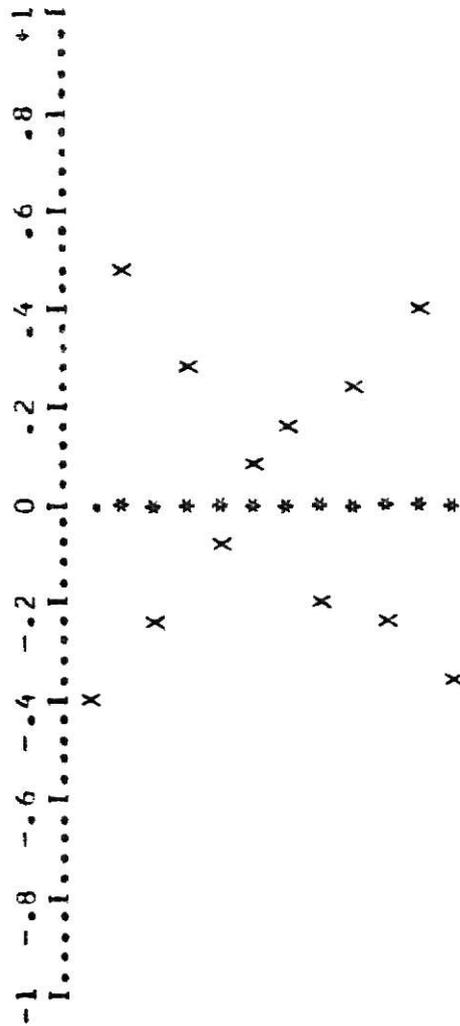
CALCULATED RESULTS MEAN = 0.07
 ABS. DEV. = 0.35
 VAR. = 0.20
 STD. DEV. = 0.45

LAG AUTOCORRELATION PARTIAL ACCORREL *

1	-0.39	X	-0.39
2	0.46		0.0
3	-0.23		0.0
4	0.30		0.0
5	-0.09		0.0
6	0.09		0.0
7	0.17		0.0
8	-0.18		0.0
9	0.24		0.0
10	-0.26		0.0
11	0.40		0.0
12	-0.36		0.0

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.39



BJRES

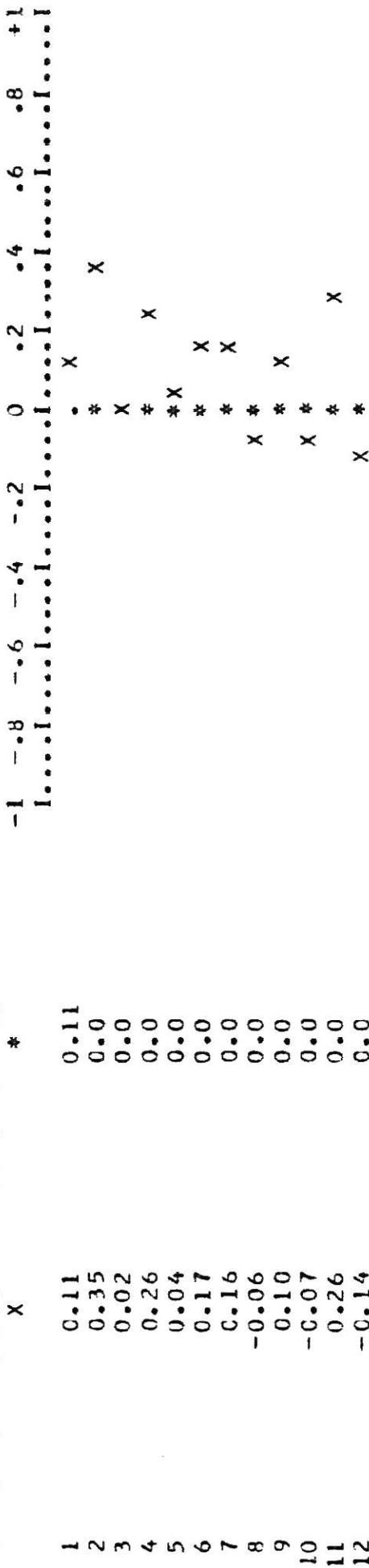
BJSTAT

INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.189
 LTERM = -0.857

CALCULATED RESULTS MEAN = 0.11
 ABS. DEV. = 0.31
 VAR. = 0.17
 STD. DEV. = 0.41

LAG AUTOCORRELATION PARTIAL ACCORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.11

SEARCH NUMBER 2
 BT1 = -0.68 BT2 = 0.0 BRMS = 0.41 T1 = -0.68 T2 = 0.0 SIGMA = 0.41

BJARMA

T1 = -0.45

T2 = 0.0

LSEA = 0

IWTERM= 2

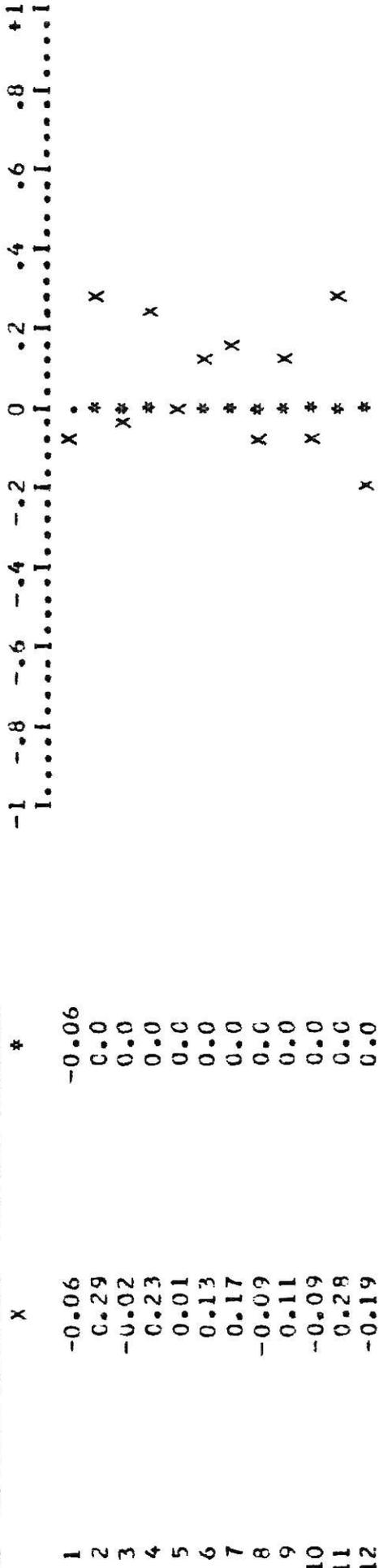
BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -0.831

CALCULATED RESULTS MEAN = 0.09
 ABS. DEV. = 0.31
 VAR. = 0.16
 STD. DEV. = 0.40

LAG AUTCCORRELATION PARTIAL ACCORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

BJRES

BJSTAT

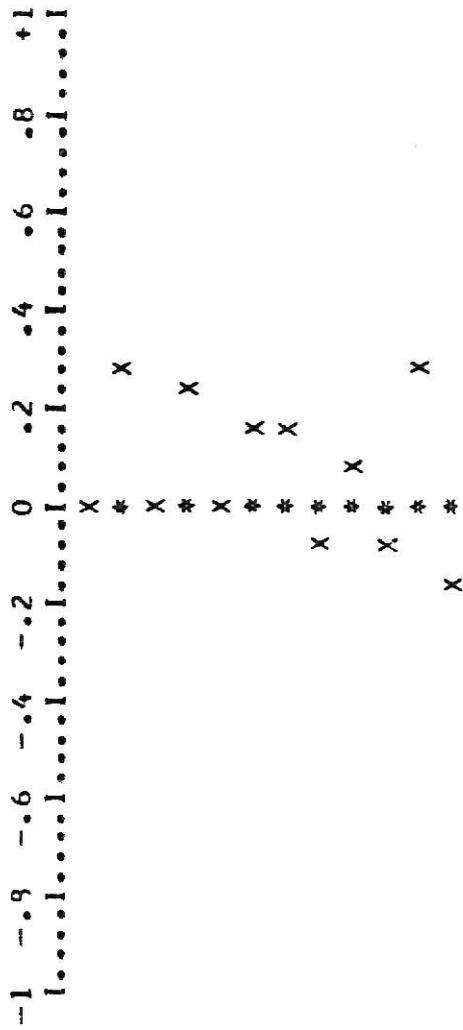
INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.163
 LTERM = -0.923

CALCULATED RESULTS MEAN = 0.10
 ABS. DEV. = 0.30
 VAR. = 0.16
 STD. DEV. = 0.40

LAG AUTOCORRELATION PARTIAL ACCORREL
 X

1	0.01		0.01
2	0.30		0.0
3	0.01		0.0
4	0.23		0.0
5	0.01		0.0
6	0.15		0.0
7	0.16		0.0
8	-0.08		0.0
9	0.09		0.0
10	-0.08		0.0
11	0.26		0.0
12	-0.17		0.0



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.01

SEARCH NUMBER 3 T1 = -0.45 T2 = 0.0 SIGMA = 0.40
 RT1 = -0.45 RT2 = 0.0 BRMS = 0.40

BJARMA

T1 = -0.23
T2 = 0.0
LSEA = 0
IMTERM= 2

HJSTAT

INPUT PARAMETERS ITERM = 2
LTERM = 78
KMAX = 12
LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
LTERM = -0.777

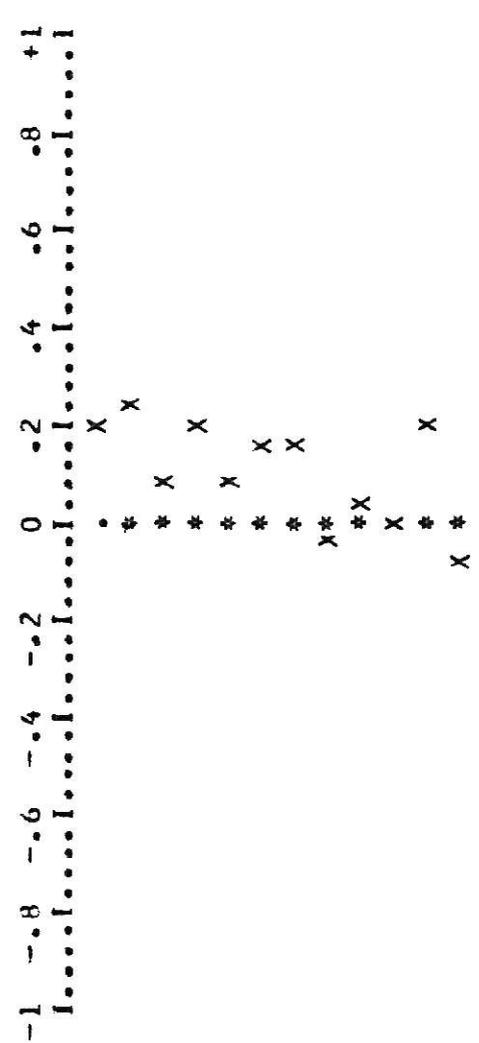
CALCULATED RESULTS MEAN = 0.11
ABS. DEV. = 0.30
VAR. = 0.16
STU. DEV. = 0.40

LAG AUTOCORRELATION PARTIAL ACORREL *

LAG	AUTOCORRELATION	PARTIAL ACORREL	*
1	0.21		0.21
2	0.24		0.0
3	0.09		0.0
4	0.22		0.0
5	0.08		0.0
6	0.17		0.0
7	0.17		0.0
8	-0.02		0.0
9	0.05		0.0
10	0.01		0.0
11	0.21		0.0
12	-0.08		0.0

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.21



BJRES

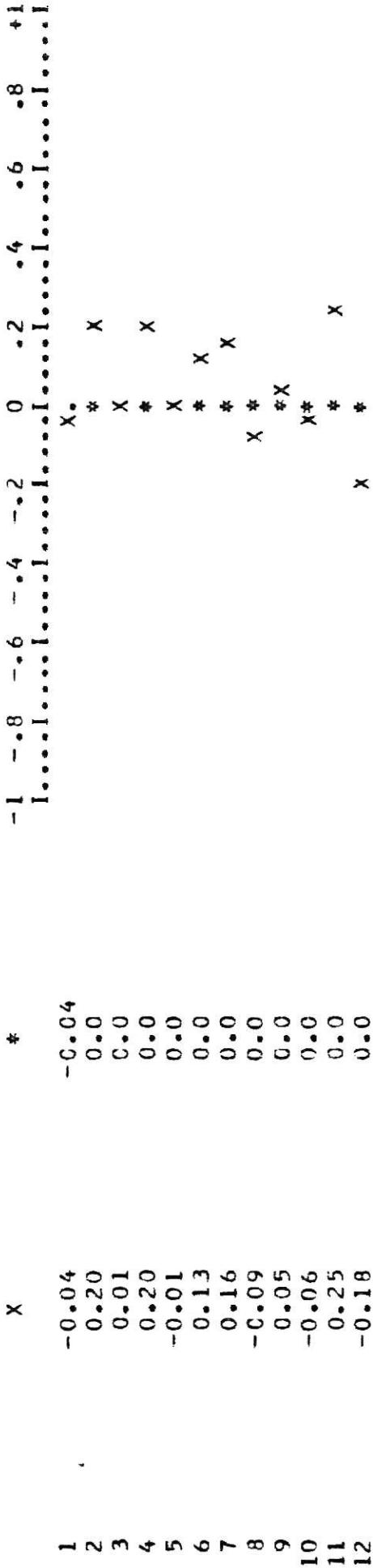
BJSTAT

INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.152
 LTERM = -0.783

CALCULATED RESULTS MEAN = 0.05
 ABS. DEV. = 0.30
 VAR. = 0.15
 STD. DEV. = 0.39

LAG AUTOCORRELATION PARTIAL ACORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.04

SEARCH NUMBER 4
 BT1 = -0.23 BT2 = 0.0 ERMS = 0.39 T1 = -0.23 T2 = 0.0 SIGMA = 0.39

BJARMA

T1 = -0.00
T2 = 0.0
LSEA = 0
INTERM = 2

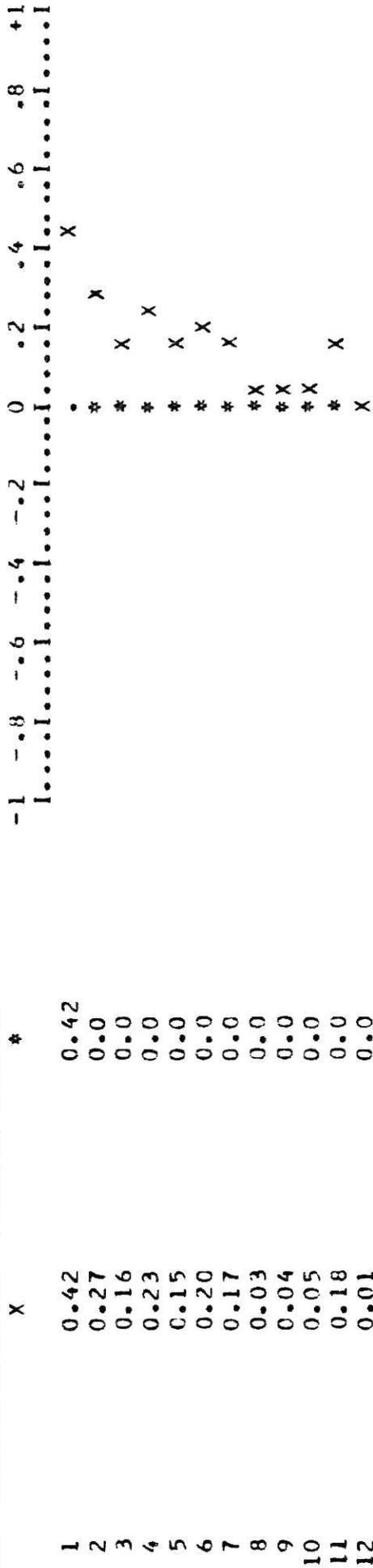
BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -0.770

CALCULATED RESULTS MEAN = 0.13
 ABS. DEV. = 0.33
 VAR. = 0.18
 STD. DEV. = 0.42

LAG AUTOCORRELATION PARTIAL ACORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.42

BJRES

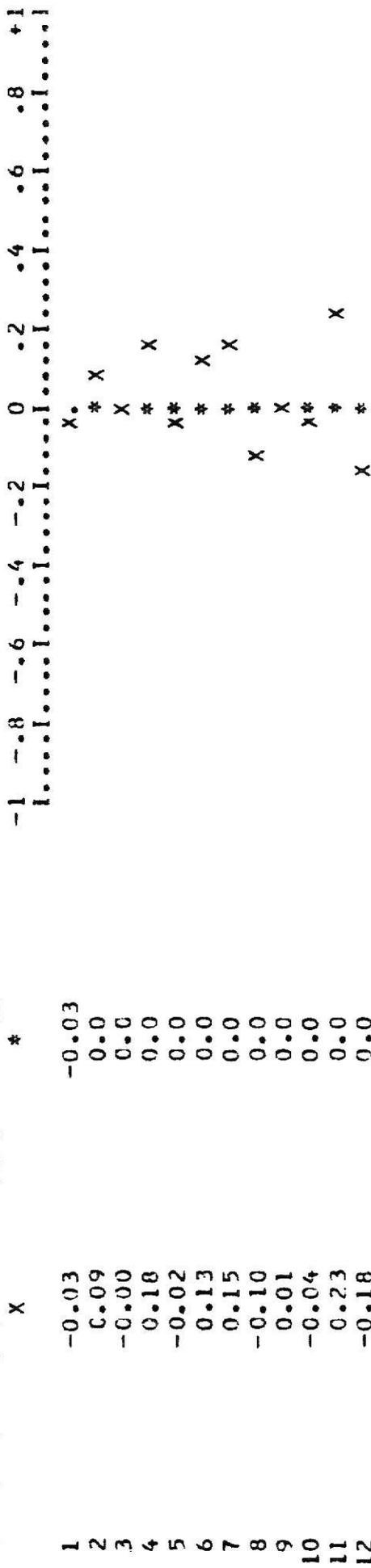
BJSTAT

INPUT PARAMETERS ITERM = 3
 LTERM = 79
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.155
 LTERM = -0.732

CALCULATED RESULTS MEAN = 0.08
 ABS. DEV. = 0.30
 VAR. = 0.15
 STD. DEV. = 0.38

LAG AUTOCORRELATION PARTIAL ACCORREL *



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 -0.03

SEARCH NUMBER 5
 BT1 = -0.00 BT2 = 0.0 BRMS = 0.38 T1 = -0.00 T2 = 0.0 SIGMA = 0.38

BJARMA

T1 = 0.22
 T2 = 0.0
 LSEA = 0
 ITERM= 2

BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -0.821

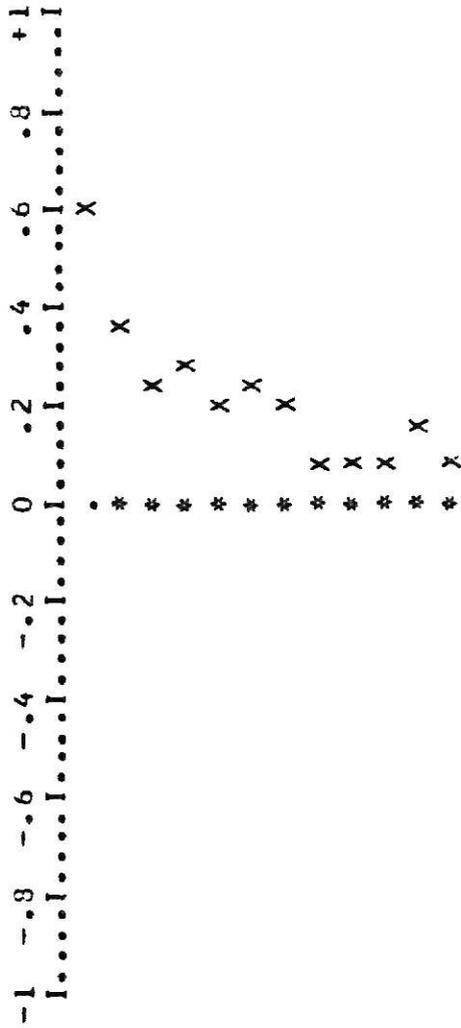
CALCULATED RESULTS MEAN = 0.18
 ABS. DEV. = 0.38
 VAR. = 0.23
 STD. DEV. = 0.48

LAG AUTOCORRELATION PARTIAL ACORREL *

LAG	AUTOCORRELATION	PARTIAL ACORREL *
1	0.60	0.60
2	0.38	0.0
3	0.25	0.0
4	0.26	0.0
5	0.22	0.0
6	0.23	0.0
7	0.19	0.0
8	0.08	0.0
9	0.06	0.0
10	0.09	0.0
11	0.16	0.0
12	0.08	0.0

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.60



BJARMA

T1 = 0.45
 T2 = 0.0
 LSEA = 0
 ITERM= 2

BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -0.942

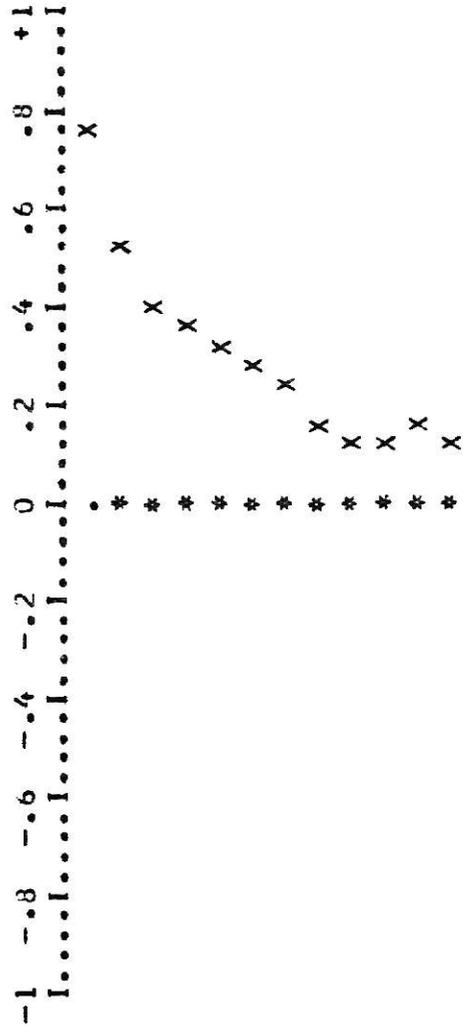
CALCULATED RESULTS MEAN = 0.25
 ABS. DEV. = 0.49
 VAR. = 0.35
 SID. DEV. = 0.59

LAG AUTCCORRELATION PARTIAL ACORREL #

LAG	AUTCCORRELATION	PARTIAL ACORREL	#
1	0.74		0.74
2	0.54		0.0
3	0.39		0.0
4	0.35		0.0
5	0.30		0.0
6	0.28		0.0
7	0.22		0.0
8	0.14		0.0
9	0.11		0.0
10	0.13		0.0
11	0.17		0.0
12	0.12		0.0

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.74



BJARMA

T1 = 0.67
T2 = 0.0
LSEA = 0
IINTERM= 2

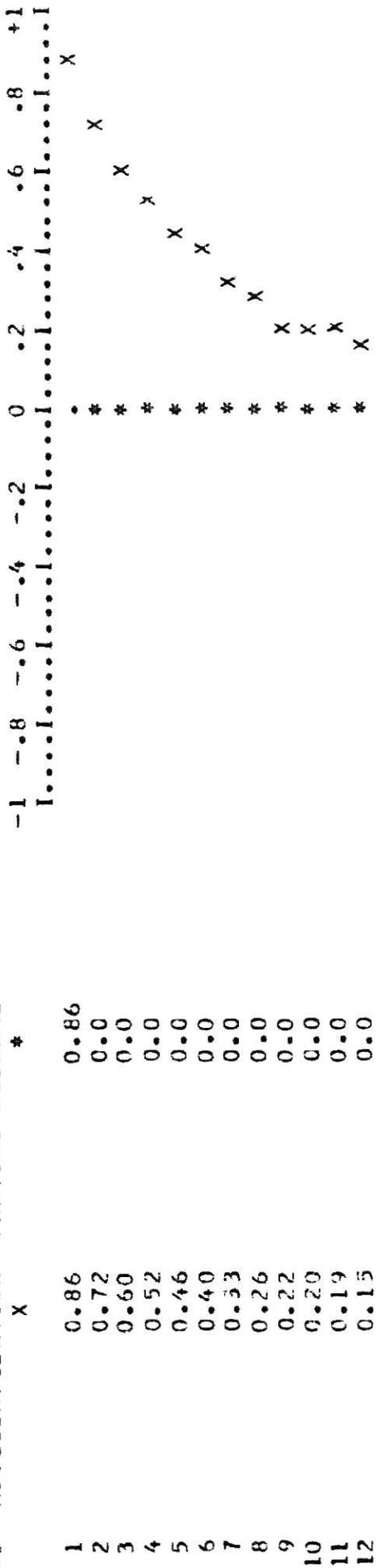
BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -1.154

CALCULATED RESULTS MEAN = 0.44
 ABS. DEV. = 0.71
 VAR. = 0.69
 STD. DEV. = 0.83

LAG AUTOCORRELATION PARTIAL ACORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.86

BJRES

BJSTAT

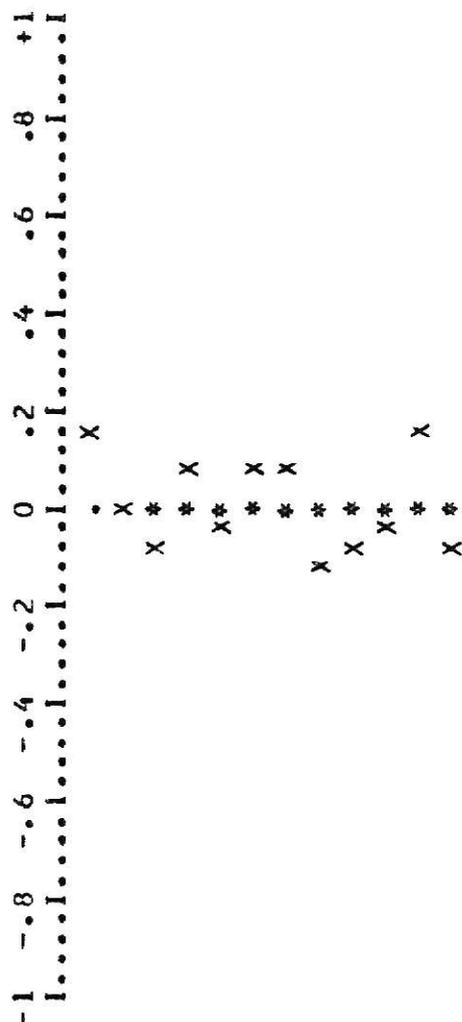
INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.213
 LTERM = -0.663

CALCULATED RESULTS MEAN = 0.05
 ABS. DEV. = 0.29
 VAR. = 0.15
 STD. DEV. = 0.38

LAG AUTOCORRELATION PARTIAL ACCORREL
 X *

1	0.17		0.17
2	0.01		0.0
3	-0.08		0.0
4	0.06		0.0
5	-0.04		0.0
6	0.08		0.0
7	0.03		0.0
8	-0.13		0.0
9	-0.08		0.0
10	-0.04		0.0
11	0.17		0.0
12	-0.07		0.0



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.17

SEARCH NUMBER 8
 BT1 = 0.45 BT2 = 0.0 BRMS = 0.38 T1 = 0.67 T2 = 0.0 SIGMA = 0.20

BJARMA

T1 = 0.90
T2 = 0.0
LSEA = 0
IINTERM= 2

BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT STRIPS VALUES ITERM = -0.250
 LTERM = -0.972

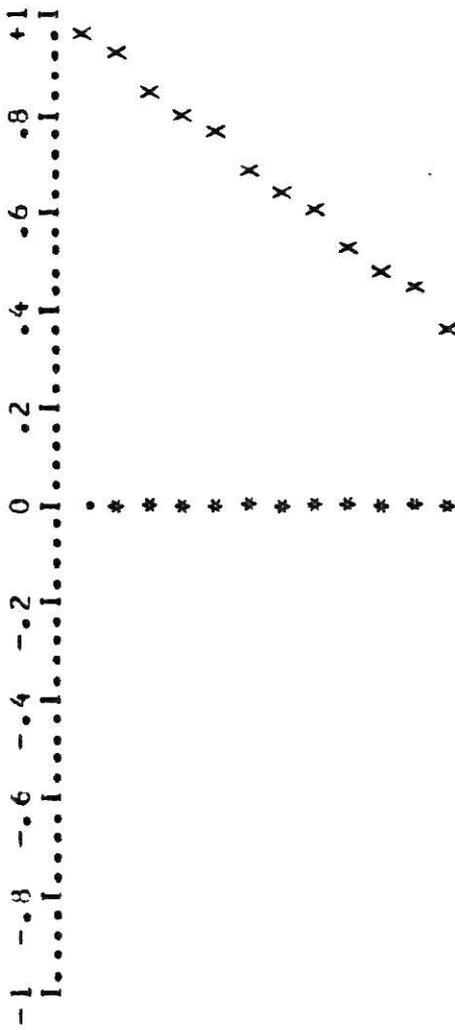
CALCULATED RESULTS MEAN = 1.45
 ABS. DEV. = 1.52
 VAR. = 3.24
 STD. DEV. = 1.80

LAG AUTOCORRELATION PARTIAL ACORREL

LAG	AUTOCORRELATION	PARTIAL ACORREL
1	0.96	0.56
2	0.91	0.0
3	0.85	0.0
4	0.80	0.0
5	0.75	0.0
6	0.70	0.0
7	0.64	0.0
8	0.59	0.0
9	0.53	0.0
10	0.48	0.0
11	0.42	0.0
12	0.37	0.0

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.96



SEARCH GRID

TABLE OF SIGMA VALUES

T2 PARAMETER VALUES

* 0.0

T1 PARAMETER

VALUES

-0.90	*	0.41
-0.68	*	0.41
-0.45	*	0.40
-0.23	*	0.39
-0.00	*	0.38
0.22	*	0.38
0.45	*	0.38
0.67	*	0.38
0.90	*	0.40

BEST PARAMETERS

T1 =	0.45
T2 =	0.0
BRMS =	0.38

BJARMA

T1 = 0.45
 T2 = 0.0
 LSEA = 0
 ITERM= 2

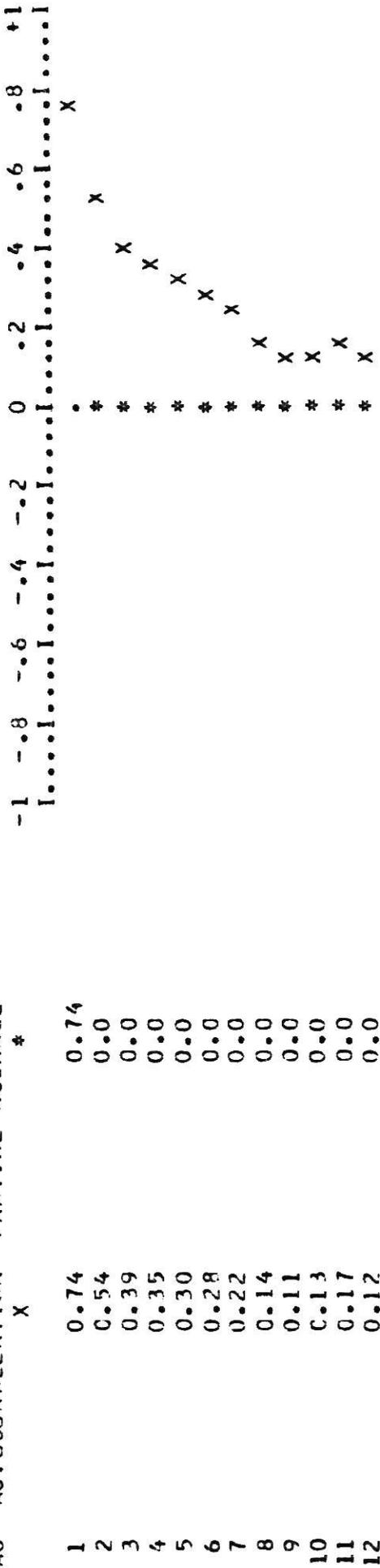
BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.250
 LTERM = -0.942

CALCULATED RESULTS MEAN = 0.25
 ABS. DEV. = 0.49
 VAR. = 0.35
 STD. DEV. = 0.59

LAG AUTOCORRELATION PARTIAL ACCORREL *



PARAMETERS OF AN AUTOREGRESSIVE MODEL

BJRES

BJSTAT

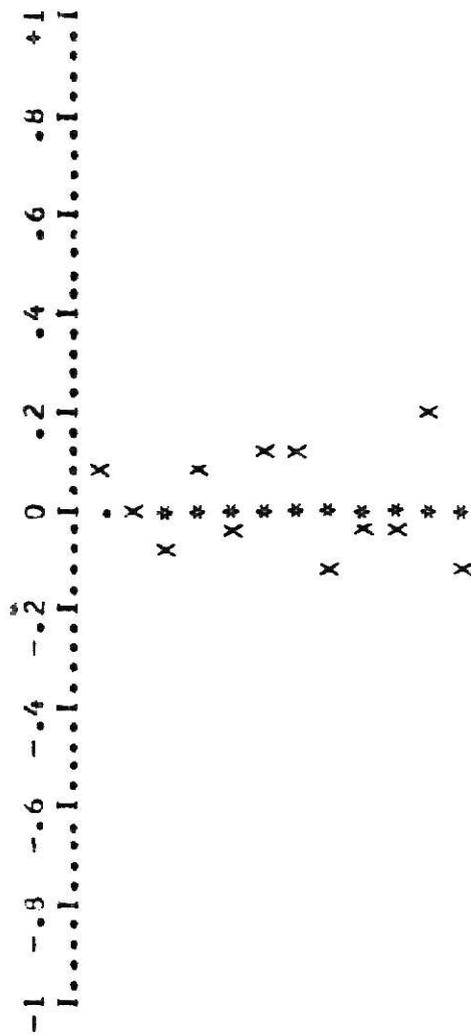
INPUT PARAMETERS ITERM = 3
 LTERM = 78
 KMAX = 12
 LMAX = 1

INPUT SERIES VALUES ITERM = -0.187
 LTERM = -0.658

CALCULATED RESULTS MEAN = 0.06
 ABS. DEV. = 0.29
 VAR. = 0.14
 STD. DEV. = 0.38

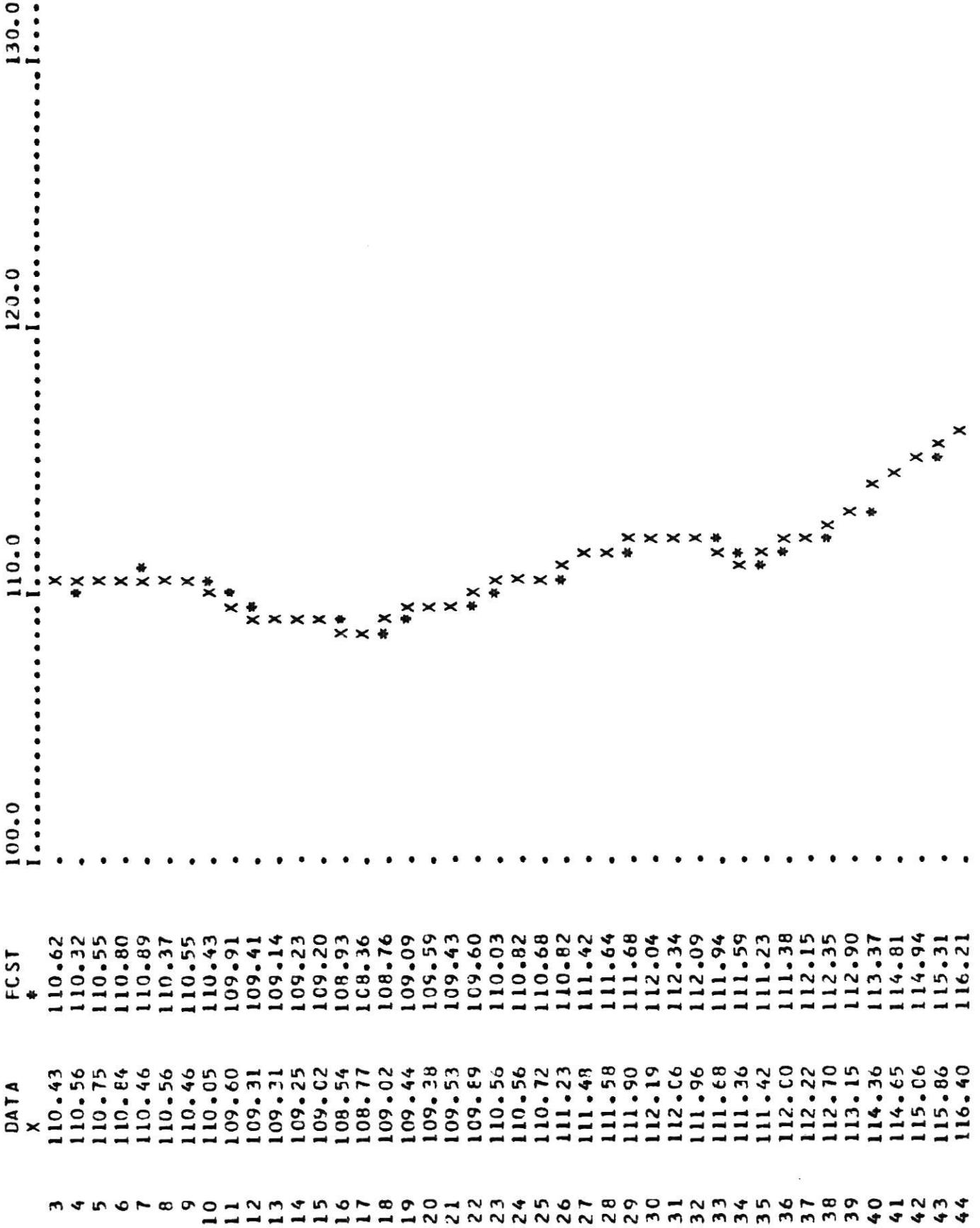
LAG AUTOCCORRELATION PARTIAL ACORREL *
 X

1	0.08		0.08
2	-0.01		0.0
3	-0.08		0.0
4	0.10		0.0
5	-0.03		0.0
6	0.11		0.0
7	0.12		0.0
8	-0.12		0.0
9	-0.05		0.0
10	-0.03		0.0
11	0.19		0.0
12	-0.11		0.0



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.08



4.3 Wolfer Sunspot Number (1770 - 1860)

Identification phase: (p. 121 - 129)

Step 1: The autocorrelations of the original series (p. 122) is in damped sine wave indicate the original series is stationary. Go through 3 models indicates model 1 (p. 122 - 123) has the smallest standard deviation. Therefore, no differencing operation is needed.

Step 2: Since the autocorrelation of the original series is in damped sine wave, it indicates that an autoregressive (AR) model is suggested.

Step 3: The partial autocorrelations (p. 122) after second lag are insignificant indicate that a second order autoregressive (AR(2)) model is suggested.

Step 4: From step 1 - 3, we have a tentative ARIMA(2,0,0) model on hand.

BOXJEN PRELIMINARY IDENTIFICATION
INPUT DATA

1	101.00	2	82.00	3	66.00	4	35.00	5	31.00	6	7.00
7	20.00	8	92.00	9	154.00	10	125.00	11	85.00	12	68.00
13	38.00	14	23.00	15	10.00	16	24.00	17	83.00	18	132.00
19	131.00	20	118.00	21	90.00	22	67.00	23	60.00	24	47.00
25	41.00	26	21.00	27	16.00	28	6.00	29	4.00	30	7.00
31	14.00	32	34.00	33	45.00	34	43.00	35	48.00	36	42.00
37	28.00	38	10.00	39	8.00	40	2.00	41	0.0	42	1.00
43	5.00	44	12.00	45	14.00	46	35.00	47	45.00	48	41.00
49	30.00	50	24.00	51	16.00	52	7.00	53	4.00	54	2.00
55	8.00	56	17.00	57	36.00	58	50.00	59	62.00	60	67.00
61	71.00	62	48.00	63	23.00	64	8.00	65	13.00	66	57.00
67	122.00	68	138.00	69	103.00	70	86.00	71	63.00	72	37.00
73	24.00	74	11.00	75	15.00	76	40.00	77	62.00	78	58.00
79	124.00	80	96.00	81	66.00	82	64.00	83	54.00	84	39.00
85	21.00	86	7.00	87	4.00	88	23.00	89	55.00	90	54.00
91	96.00	92	77.00	93	59.00	94	44.00	95	47.00	96	30.00
97	16.00	98	7.00	99	37.00	100	74.00				

BJRES

BJSTAT

INPUT PARAMETERS ITERM = 13
 LTERM = 100
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = -15.114
 LTERM = 12.626

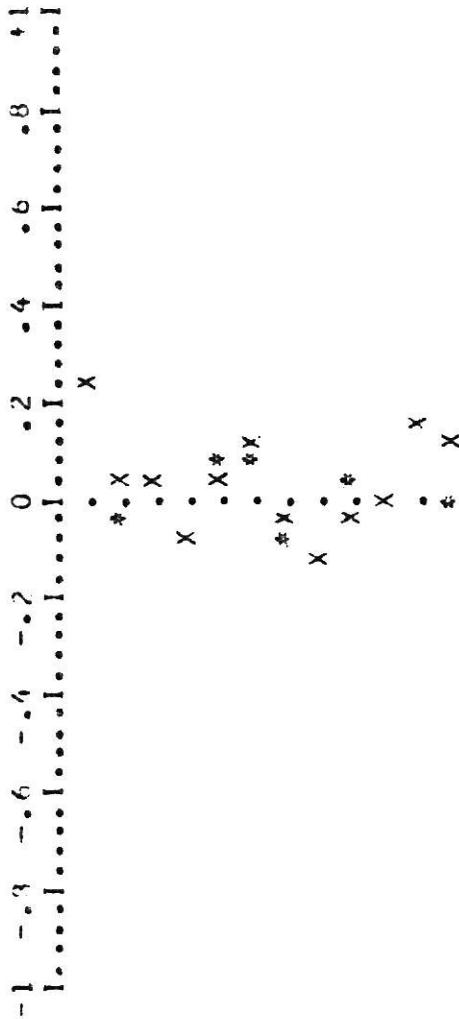
CALCULATED RESULTS MEAN = -0.59
 ABS. DEV. = 9.90
 VAR. = 161.34
 STD. DEV. = 12.70

LAG AUTOCORRELATION PARTIAL ACORREL *

1	0.25		0.25
2	0.04		-0.02
3	0.03		0.03
4	-0.07		-0.09
5	0.05		0.10
6	0.10		0.07
7	-0.03		-0.07
8	-0.11		-0.11
9	-0.03		0.04
10	-0.01		0.00
11	0.16		0.16
12	0.12		0.02

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	0.26	2	-0.02	3	0.07	4	-0.12	5	0.07	6	0.07	7	-0.02
8	-0.13	9	0.04	10	-0.04	11	0.16	12	0.02				



BJDIFF

ITERM = 1
LTERM = 100
NDIFF1 = 1
NDIFF2 = 0
I AVG = 0
AVG = -0.27

Model 2

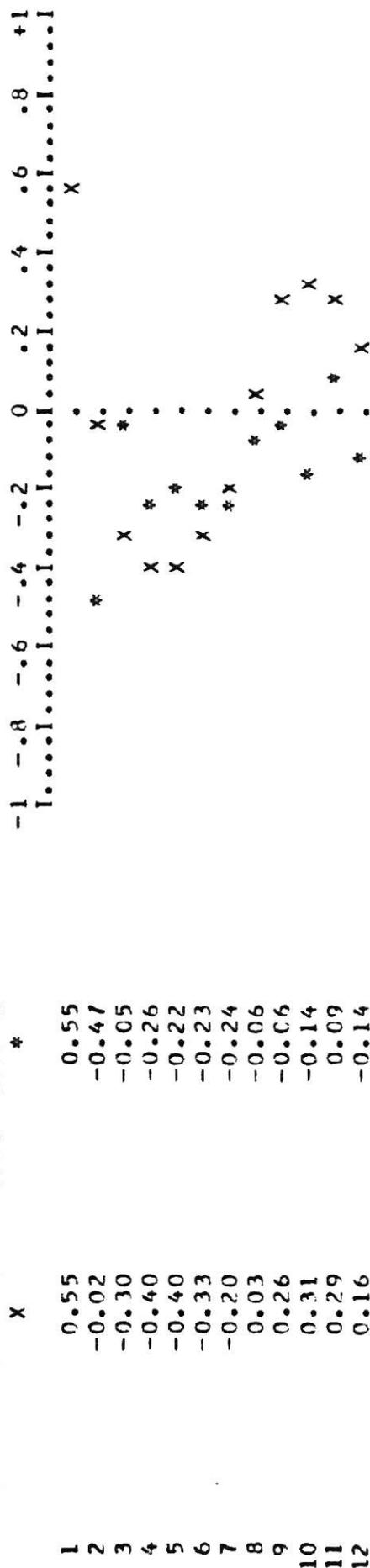
BJSTAT

INPUT PARAMETERS ITERM = 2
 LTERM = 100
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = -19.000
 LTERM = 37.000

CALCULATED RESULTS MEAN = -0.27
 ABS. DEV. = 17.09
 VAR. = 504.01
 STD. DEV. = 22.45

LAG AUTOCORRELATION PARTIAL ACCORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL
1 0.61 2 -0.57 3 -0.04
8 -0.13 9 0.07 10 -0.27

4 -0.23 5 -0.23 6 -0.13 7 -0.24
11 0.17 12 -0.14

BJRES

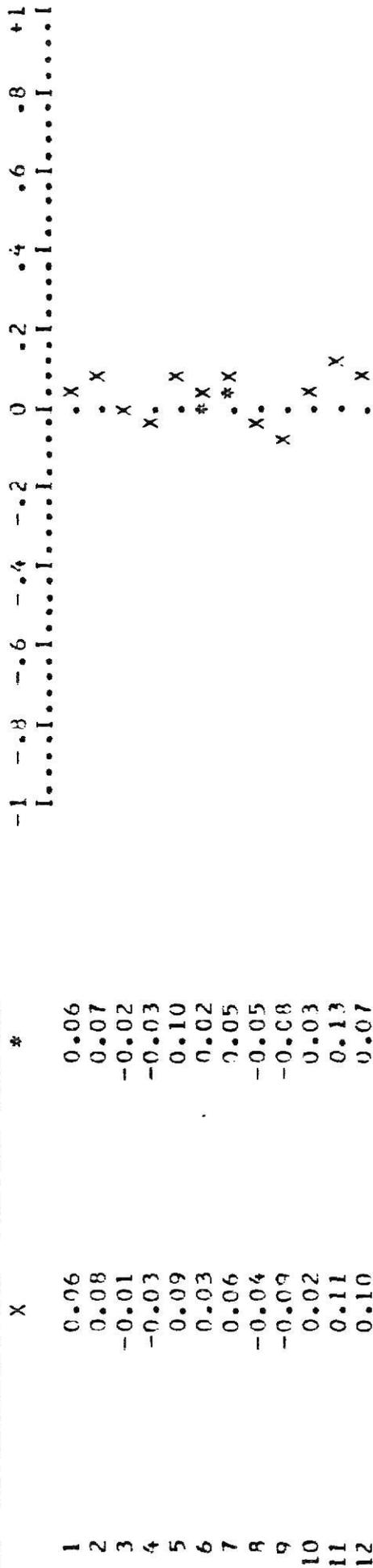
BJSTAT

INPUT PAPAMETERS ITERM = 14
 LTERM = 100
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = 1.545
 LTERM = 8.563

CALCULATED RESULTS MEAN = -0.16
 ABS. DEV. = 8.97
 VAR. = 141.40
 STD. DEV. = 11.89

LAG AUTOCORRELATION PARTIAL ACORREL *



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	0.04	2	0.09	3	-0.00	4	-0.04	5	0.08	6	0.07	7	0.05
8	-0.04	9	-0.10	10	0.02	11	0.12	12	0.07				

BJDIFF

ITRM = 1
LTERM = 100
NDIFF1 = 1
NDIFF2 = 1
IAVG = 0
AVG = 0.57

Model 3

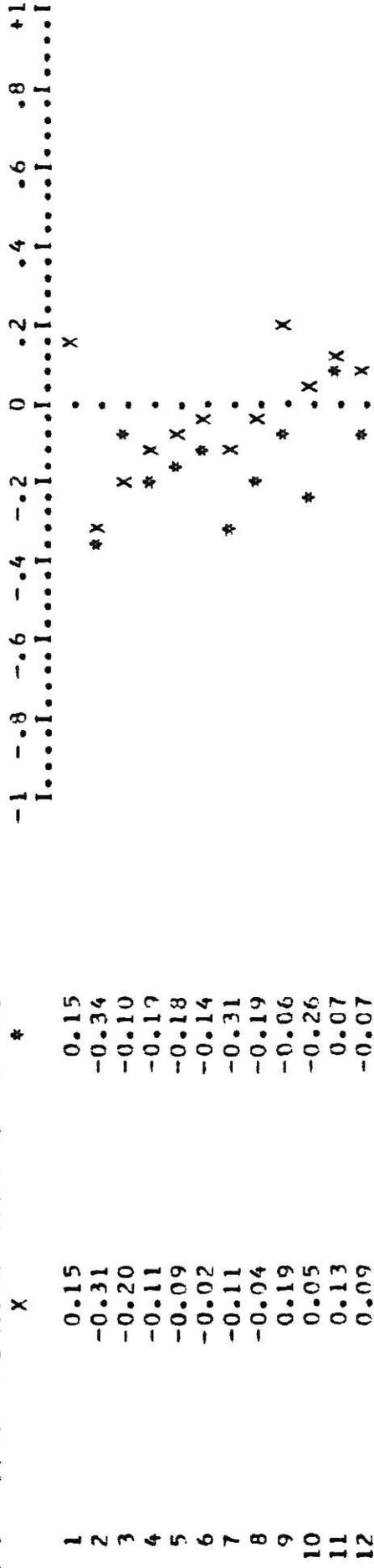
BJSTAT

INPUT PARAMETERS ITERM = 3
LTERM = 100
KMAX = 12
LMAX = 12

INPUT SERIES VALUES ITERM = 3.000
LTERM = 7.000

CALCULATED RESULTS MEAN = 0.57
ABS. DEV. = 14.31
VAR. = 438.98
STD. DEV. = 20.95

LAG AUTOCORRELATION PARTIAL ACORREL *



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	-0.02	2	-0.57	3	-0.35	4	-0.39	5	-0.45	6	-0.29	7	-0.41
8	-0.33	9	-0.06	10	-0.29	11	0.17	12	-0.07				

BJRES

BJSTAT

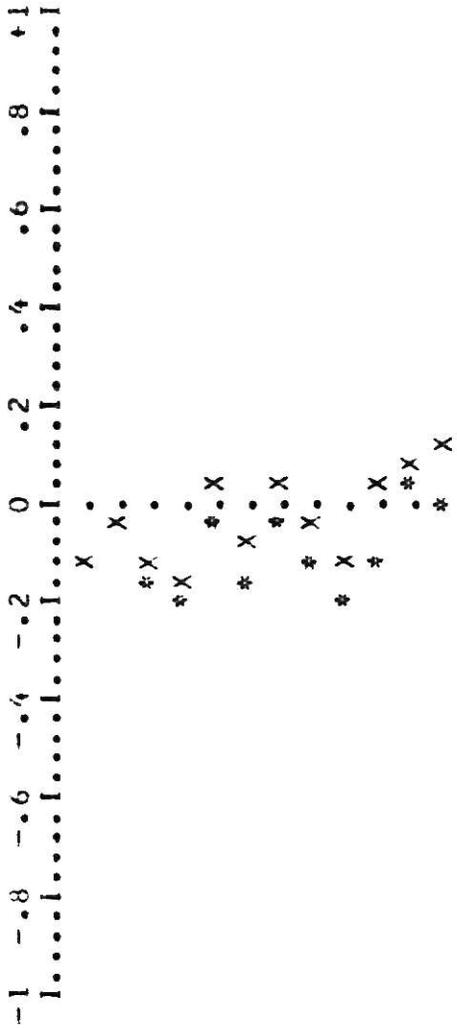
INPUT PARAMETERS ITERM = 15
 LTERM = 100
 KMAX = 12
 LMAX = 12

INPUT SERIES VALUES ITERM = -0.842
 LTERM = 6.557

CALCULATED RESULTS MEAN = 0.34
 ARS. DEV. = 10.27
 VAR. = 179.45
 STD. DEV. = 13.40

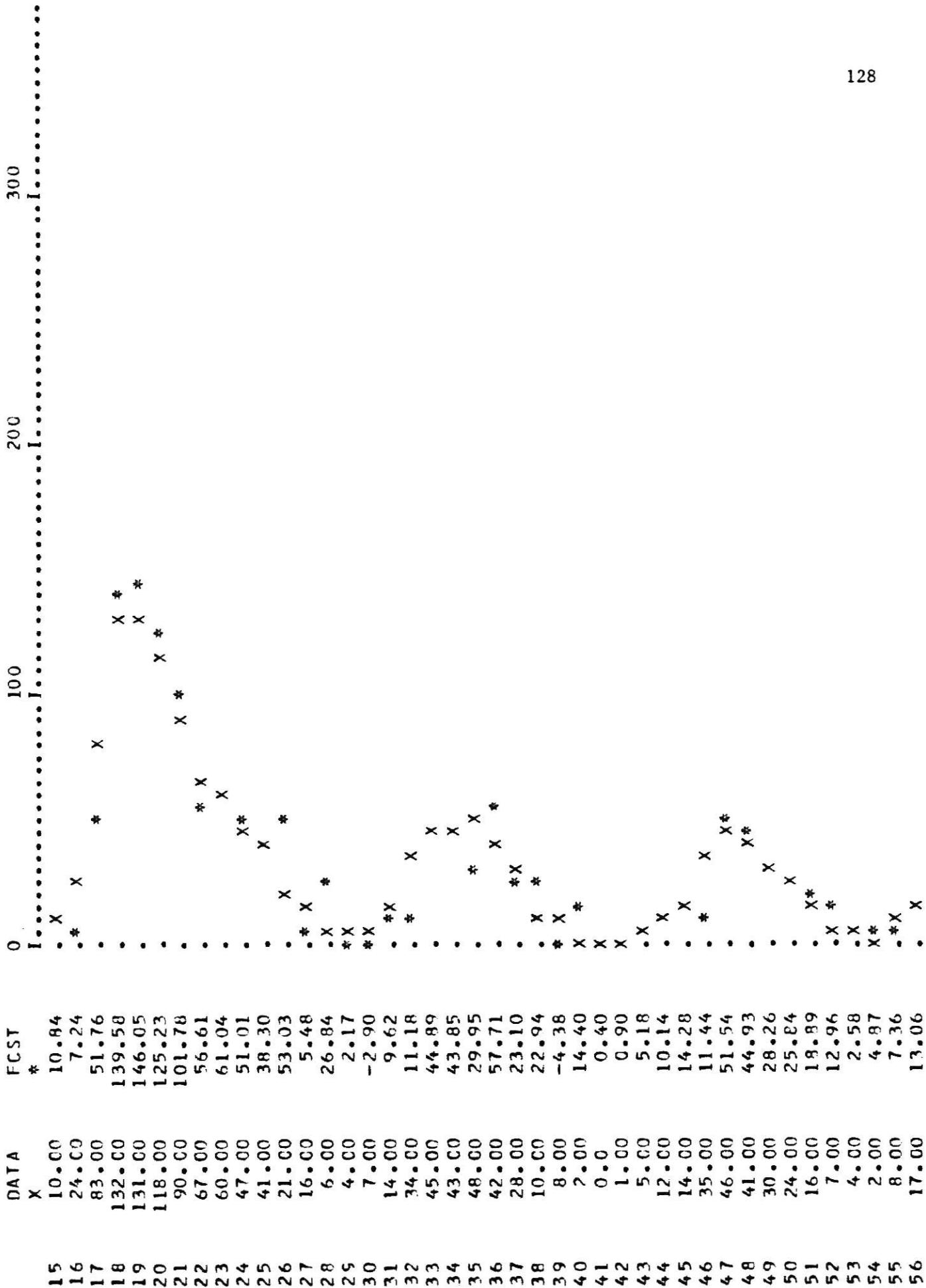
LAG AUTOCORRELATION PARTIAL ACCORREL
 X

1	-0.10		-0.10
2	-0.03		-0.04
3	-0.14		-0.15
4	-0.17		-0.21
5	0.02		-0.04
6	-0.08		-0.14
7	0.05		-0.05
8	-0.04		-0.11
9	-0.12		-0.21
10	0.02		-0.11
11	0.10		0.02
12	0.10		0.02



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1	-0.22	2	-0.15	3	-0.26	4	-0.30	5	-0.16	6	-0.23	7	-0.12
8	-0.15	9	-0.22	10	-0.10	11	0.03	12	0.02				



Estimation phase: (p. 131 - 140)

Step 1: Parameters initialization

NDIFF1 = NDIFF2 = 0

IAVG = 1

KMAX = 20

LMAX = 2

NSRT1 = NSRT2 = 1

T1MIN = T1MAX = T2MIN = T2MAX = 0

LSEA = 1

Step 2: $N = 100$, the approximate upper bound for the standard error is $1/\sqrt{100} = .1$. Compare with standard error bound, the value $\gamma_1(\hat{a})$ and $\gamma_9(\hat{a})$ are rather large, also $\gamma_{11}(\hat{a})$ and $\gamma_{12}(\hat{a})$ (see p. 137) are considerable larger than standard error limit. In random series occasional large deviation may occurred, however, this will not affect the whole series. Also, the autocorrelation will considerable smaller as K grows bigger.

BOXJEN PRFLIMINARY ESTIMATION
INPUT DATA

1	101.00	2	82.00	3	66.00	4	35.00	5	31.00	6	7.00
7	20.00	8	92.00	9	154.00	10	125.00	11	85.00	12	68.00
13	38.00	14	23.00	15	10.00	16	24.00	17	83.00	18	132.00
19	131.00	20	118.00	21	90.00	22	67.00	23	60.00	24	47.00
25	41.00	26	21.00	27	16.00	28	6.00	29	4.00	30	7.00
31	14.00	32	34.00	33	45.00	34	43.00	35	48.00	36	42.00
37	28.00	38	10.00	39	8.00	40	2.00	41	0.0	42	1.00
43	5.00	44	12.00	45	14.00	46	35.00	47	46.00	48	41.00
49	30.00	50	24.00	51	16.00	52	7.00	53	4.00	54	2.00
55	8.00	56	17.00	57	36.00	58	50.00	59	62.00	60	67.00
61	71.00	62	48.00	63	23.00	64	3.00	65	13.00	66	57.00
67	122.00	68	138.00	69	103.00	70	86.00	71	63.00	72	37.00
73	24.00	74	11.00	75	15.00	76	40.00	77	62.00	78	98.00
79	124.00	80	96.00	81	66.00	82	64.00	83	54.00	84	39.00
85	21.00	86	7.00	87	4.00	88	23.00	89	55.00	90	54.00
91	96.00	92	77.00	93	59.00	94	44.00	95	47.00	96	30.00
97	16.00	98	7.00	99	37.00	100	74.00				

BJDIFF
 ITERM = 1
 LTERM = 100
 NDIFF1 = 0
 NDIFF2 = 0
 IAVG = 1
 AVG = 46.93

JMASR
 NSRT1 = 1 T1MIN = 0.0 T1MAX = 0.0
 NSRT2 = 1 T2MIN = 0.0 T2MAX = 0.0

BJARMA

T1 = 0.0
T2 = 0.0
LSEA = 1
LWTERM = 1

BJSTAT

INPUT PARAMETERS ITERM = 3
LTERM = 100
KMAX = 12
LMAX = 2

INPUT SERIES VALUES ITERM = 19.070
LTERM = 27.070

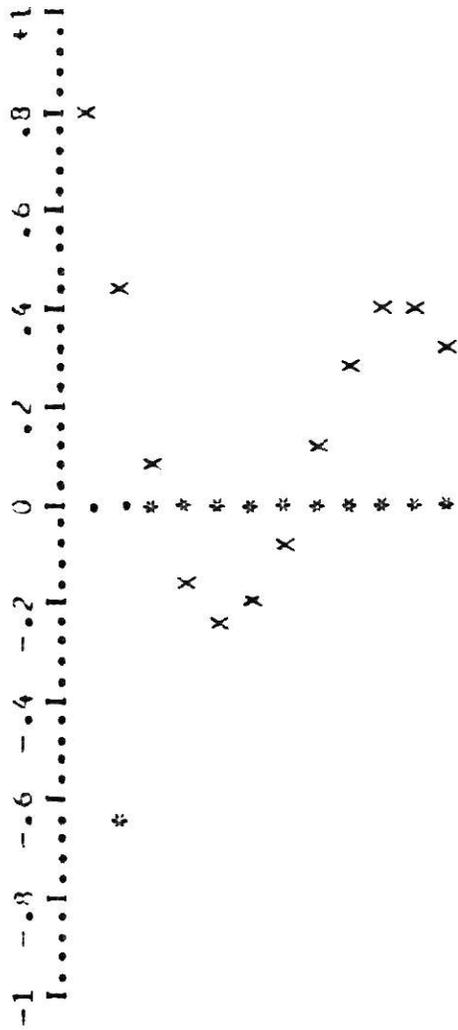
CALCULATED RESULTS MEAN = -0.91
ABS. DEV. = 29.76
VAR. = 1367.16
STD. DEV. = 36.98

LAG AUTOCORRELATION PARTIAL ACORREL *

LAG	AUTOCORRELATION	PARTIAL	ACORREL
1	0.81	X	0.81
2	0.44		-0.65
3	0.08		0.0
4	-0.16		0.0
5	-0.25		0.0
6	-0.22		0.0
7	-0.09		0.0
8	0.10		0.0
9	0.30		0.0
10	0.40		0.0
11	0.40		0.0
12	0.30		0.0

PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 1.34 2 -0.65



BJRES

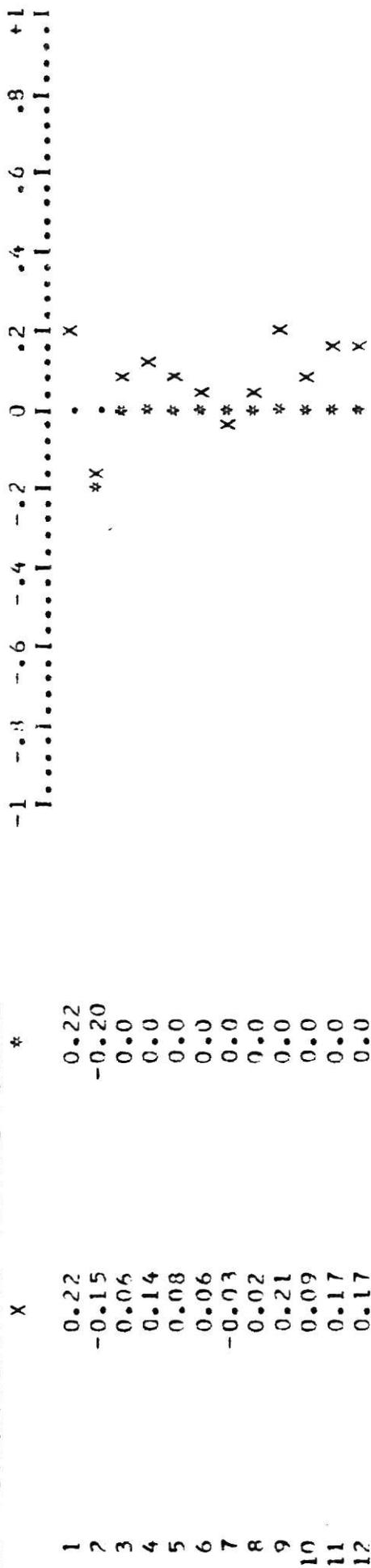
BJSTAT

INPUT PARAMETERS ITERM = 5
 LTERM = 100
 KMAX = 12
 LMAX = 2

INPUT SERIES VALUES ITERM = 12.503
 LTERM = 14.347

CALCULATED RESULTS MEAN = 0.16
 ABS. DEV. = 11.85
 VAR. = 231.37
 STD. DEV. = 15.21

LAG AUTOCORRELATION PARTIAL ACCORREL



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.26 2 -0.20

SFAPCH NUMBER 1
 BT1 = 0.0 BT2 = 0.0 PRMS = 15.21 T1 = 0.0 T2 = 0.0 SIGMA = 15.21

SEARCH GRID

TABLE OF SIGMA VALUES

T2 PARAMETER VALUES

* 0.0

T1 PARAMETER

VALUES

0.0 * 15.21

BEST PARAMETERS

T1 = 0.0
T2 = 0.0
BRMS = 15.21

BJARMA

T1 = 0.0
 T2 = 0.0
 LSEA = 1
 ITERM = 3

BJSTAT

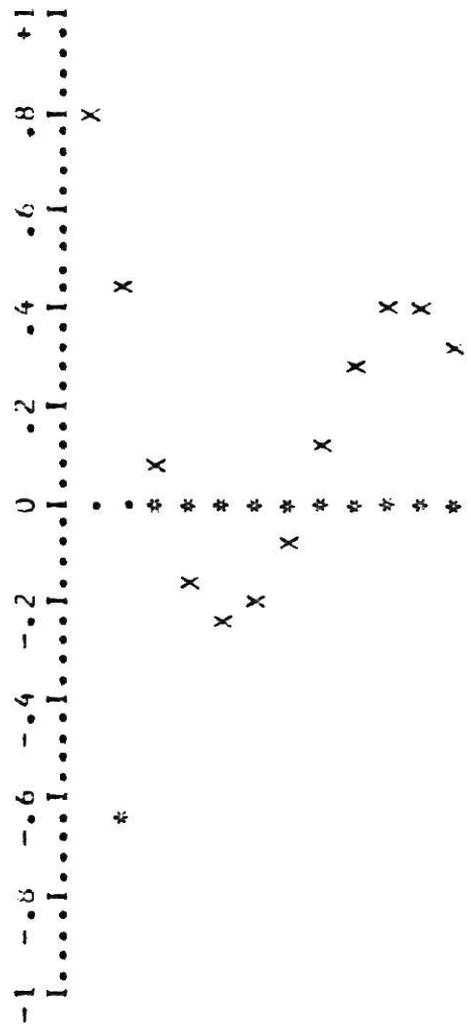
INPUT PARAMETERS ITERM = 3
 LTERM = 100
 KMAX = 12
 LMAX = 2

INPUT SERIES VALUES ITERM = 19.070
 LTERM = 27.070

CALCULATED RESULTS MEAN = -0.91
 ABS. DEV. = 29.76
 VAR. = 1367.16
 STD. DEV. = 36.98

LAG AUTOCORRELATION PARTIAL ACORREL

LAG	AUTOCORRELATION	PARTIAL	ACORREL
1	0.81	X	0.81
2	0.44		-0.65
3	0.08		0.0
4	-0.16		0.0
5	-0.25		0.0
6	-0.22		0.0
7	-0.09		0.0
8	0.10		0.0
9	0.30		0.0
10	0.40		0.0
11	0.40		0.0
12	0.30		0.0



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 1.34 2 -0.65

BJRES

BJSTAT

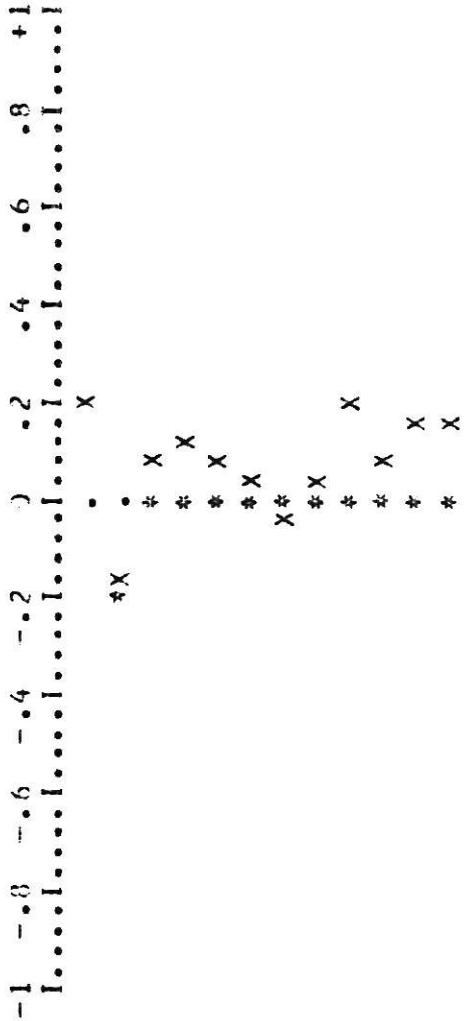
INPUT PARAMETERS ITERM = 5
 LTERM = 100
 KMAX = 12
 LMAX = 2

INPUT SERIES VALUES ITERM = 12.503
 LTERM = 14.347

CALCULATED RESULTS MEAN = 0.16
 ARS. DEV. = 11.85
 VAR. = 231.37
 STD. DEV. = 15.21

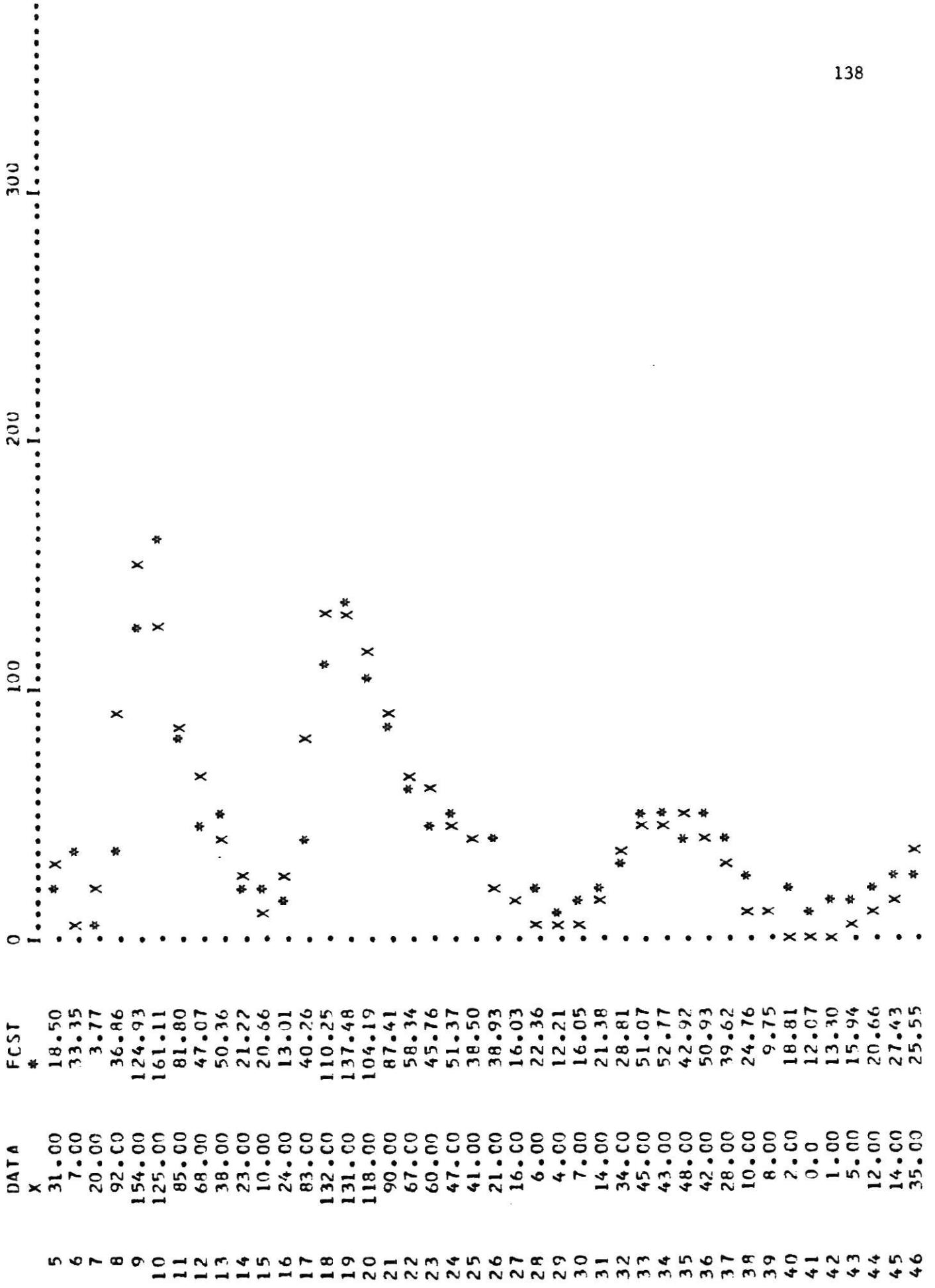
LAG AUTOCORRELATION PARTIAL ACORREL *

LAG	AUTOCORRELATION	PARTIAL	ACORREL	*
1	0.22		0.22	
2	-0.15		-0.20	
3	0.06		0.0	
4	0.14		0.0	
5	0.08		0.0	
6	0.06		0.0	
7	-0.03		0.0	
8	0.02		0.0	
9	0.21		0.0	
10	0.09		0.0	
11	0.17		0.0	
12	0.17		0.0	



PARAMETERS OF AN AUTOREGRESSIVE MODEL

1 0.26 2 -0.20



47	46.00	52.41	.	X*
48	41.00	53.46	.	X *
49	30.00	39.58	.	X *
50	24.00	28.09	.	X*
51	16.00	27.22	.	X *
52	7.00	20.41	.	X *
53	4.00	13.56	.	X *
54	2.00	15.40	X	*
55	8.00	14.68	.	X*
56	17.00	24.03	.	X *
57	36.00	32.18	.	*X
58	50.00	51.79	.	X
59	62.00	58.17	.	X
60	67.00	65.14	.	X
61	71.00	64.02	.	X
62	48.00	66.12	.	X *
63	28.00	32.67	.	X
64	8.00	20.85	.	X *
65	13.00	7.07	.	* X
66	57.00	26.82	.	*
67	122.00	82.56	.	X
68	138.00	141.03	.	*
69	103.00	120.10	.	X
70	86.00	62.73	.	*
71	63.00	62.76	.	X
72	37.00	43.00	.	X *
73	24.00	23.14	.	X
74	11.00	22.66	.	X *
75	15.00	13.70	.	X
76	40.00	27.54	.	*
77	62.00	58.46	.	X
78	98.00	71.66	.	*
79	124.00	105.58	.	X
80	96.00	116.97	.	X
81	66.00	62.47	.	*X
82	64.00	40.50	.	X
83	54.00	57.38	.	*
84	39.00	45.28	.	X*
85	21.00	31.69	.	X *
86	7.00	17.33	.	X *
87	4.00	10.30	.	X*
88	23.00	15.40	.	*
89	55.00	42.84	.	* X
90	94.00	73.36	.	* X

4.4 Summary of Result

To make more formal diagnostic checks, rather consider the $\gamma_k(a)$'s individually, taking first k autocorrelations of the a 's as a whole. We refer

$$Q = n \sum_{k=1}^K \gamma_k^2(a)$$

where

$$n = N - d$$

$$K = k - p - q$$

Table 4 shows the values of the criterion Q , based on 12 residual autocorrelations for the models that fitted to the corresponding series.

Inspection of Table 4 shows the significant large values of Q doesn't occur, which indicate these models are acceptable. However, there is a little doubt about the Wölfer Sunspot Numbers model, but the value of Q will relatively smaller as the K becomes larger. Compare the Dow Jones Utility Index model (1,1,0) and (1,1,1), there is relatively smaller Q in model (1,1,1) to show the improvement.

Table 4. Summary of results of Portmanteau test applied to residuals of various model fitted to the series in the report.

Series	n = N = d	Fitted model	Q	DF
Gross National Product (2,2,0)	37	$(1 - .27B + .3B^2)\nabla^2 Z_t = a_t$	5.98	10
Dow Jones Utility Index (1,1,0)	77	$(1 - .42B)\nabla Z_t = a_t$	7.38	11
Dow Jones Utility Index (1,1,1)	77	$(1 - .71B)\nabla Z_t = (1 - .4B)a_t$	5.24	10
Wolfers Sunspot Numbers (2,0,0)	100	$Z_t - 1.34Z_{t-1} + .65Z_{t-2} = 14.47 + a_t$	15.51	10

CHAPTER V

CONCLUSION

Some basic tools of time series analysis have discussed in Chapter 1, an important class of linear stochastic models is introduced in Chapter 2, and their properties discussed. Chapter 3 describes, in term, the process of model identification, model estimation, and model diagnostic checking, and forecasting (an interative modeling-building methodology) whereby the stochastic models introduced in Chapter 2, Chapter 4 illustrates the whole model building process by using 3 actual time series data (G.M.P., Dow Jones Utility Index, and Wolfer sunspot numbers), and analyze result.

Although the Box-Jenkins approach is of undoubted value, there are other ways of analyzing time series and making forecasts [23], [24], and in certain cases the Box-Jenkins method will be either unacceptable expensive to perform, or indeed inferior to some other approach. Thus of large numbers of series need to be quickly and cheaply forecast, say for stock-control or production-planning purposes, a fully automatic adaptive forecasting method, based on exponential smoothing, such as those associated with Holt and Winters [25], or Harrison [26], will be preferred. While, again, for short series (with strong seasonal component), one of these methods is likely to do better than Box-Jenkins method, especially at longer lead [27]. Evidently, like other statistical procedures, the Box-Jenkins approach should not be used indiscriminately.

We conclude by indicating further areas of application for the Box-Jenkins method. The approach can be used in "process control" [1]. For, if the deviation from the target can be forecast, compensatory action may be feasible. It

can be used in transfer function modeling [28], which connects inputs and output series. The whole approach promises to give rewarding results when extended to deal with multiserries situations.

REFERENCES (BIBLIOGRAPHY)

1. Box, G. E. P. and Jenkins, G. M.; Times Series Analysis Forecasting and Control, Holden-Day, San Francisco, 1970.
2. Nelson, C. R.; Applied Time Series Analysis for Managerial Forecasting, Holden-Day, San Francisco, 1973.
3. Anderson, O. D.; Time Series Analysis and Forecasting, Butterworth, London, 1976.
4. Bhat, U. N.; Elements of Applied Stochastic Processes, John Wiley and Sons, New York, 1972.
5. Kashyap, R. L. and Rao, A. R.; Dynamic Stochastic Models from Empirical Data, Academic Press, New York, 1976.
6. Graupe, D., Krause, D. J. and Moore J. B.; Identification of Autoregressive Moving Average Parameters of Time Series, IEEE Trans. Automat. Contr., Vol. AC-20, pp. 104 - 106, Feb. 1975.
7. Gersch, W.; Estimation of the Autoregressive Parameters of a Mixed Autoregressive Moving Average Time Series, IEEE Trans. Automat. Contr., Vol. AC-15, pp. 583 - 588, Oct. 1970.
8. Shinnars, S. M.; Modeling of Human Operator Performance Time Series Analysis, IEEE Trans. System, Man, and Cybernetics, Vol. SMC-4, pp. 446 - 453, Sept. 1974.
9. Wichern, D. W.; The Behavior of the Sample Autocorrelation Function for an Integrated Moving Average Process, Biometrika, Vol. 60, pp. 235 - 239, Aug. 1973.
10. Anderson, R. L.; Distribution of the Serial Correlation Coefficient, The Annual of Mathematical Statistics, Vol. XIII, No. 1, pp. 1 - 13, March 1942.
11. Bartlett, M. S.; On the Theoretical Specification and Sampling Properties of Autocorrelated Time Series, Supplement to the Journal of the Royal Statistical Society, Vol. VIII, No. 1, pp. 27 - 41, 1946.
12. Chatfield, C. and Prothero, D. L.; Box-Jenkins Seasonal Forecasting: Problem in Case-Study, Journal of the Royal Statistical Society, Series A, Vol. 136, Part 3, pp. 295 - 336, 1973.
13. Emanuel, P.; An Approach to Time Series Analysis, The Annuals of Mathematical Statistics, Vol. 32, No. 4, pp. 951 - 989, 1961.
14. Durbin, J.; The Fitting of Time-Series Models, Revue De L'Institut International De Statistique, Vol. 28, pp. 233 - 243, 1960.

15. Quenouille, M. H.; Approx. Test of Correlation in Time-Series, Journal of the Royal Statistical Society, Vol. B11, pp. 68 - 84, 1949.
16. Box, G. E. P. and Pierce, D. A.; Distribution of Residual Autocorrelation in Autoregressive Integrated Moving Average Time Series Models, Journal of the American Statistical Association, Vol. 65, pp. 1509 - 1526, Dec. 1970.
17. Walker, A. M.; Large-Sample Estimation of Parameters for Autoregressive Process with Moving-Average Residuals, Biometrika, Vol. 49, pp. 117 - 131, June 1962.
18. Quenouille, M. H.; Rapid Statistical Calculations, Hafner, New York, 1966.
19. Bartlett, M. S.; An Introduction to Stochastic Processes, Cambridge, London, 1966.
20. Yule, G. U.; On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer Sunspot Numbers, Phil. Trans., A226-227, 1967.
21. Walker, G.; On Periodicity in Series of Related Terms, Proc. Royal Society, A131-518, 1931.
22. U. S. Government (Council of Economic Advisor); Economic Report of the President, Table I, pp. 187, 1977.
23. Granger, C. W. J.; Statistical Forecasting of Economic Series: A Review of Techniques, Survey Papers in Economics, No. 8, 1973.
24. Anderson, O. D.; Comparative Forecasting and Time Series Analysis, Submitted for Publication, 1974.
25. Winters, P. R.; Forecasting Sales by Exponentially Weighted Moving Averages, Man. Sci., 6, pp. 324 - 342, 1960.
26. Harrison, P. J.; Short Term Sales Forecasting, J. R. Statist. Soc. C14, pp. 102 - 139, 1965.
27. Reid, D. J.; A Comparison of Forecasting Techniques on Economic Time Series, Forecast in Action. Ed. M. J. Bramson Et Al. Op. Res. Society, 1972.
28. Anderson, O. D.; Statistical Transfer Function Modeling, Submitted for Publication, 1974.
29. Wiener, N.; Extrapolation Interpolation and Smoothing of Stationary Time Series, John Wiley, N.Y., 1949.
30. Whittle, P.; Predication and Regulation by Linear Least-Squares Methods, English Universities Press, London, 1963.
31. I.C.L.; 1900 Series, Technical Publication 4824. Nov. 1972.

•

APPENDIX A

GENERAL DESCRIPTION OF THE BOXJEN PROGRAMS

The BOXJEN programs are based on theoretical work discussed in G. E. P. Box and G. M. Jenkins, "Time Series Analysis Forecasting and Control" [1]. The programs are written in ANS FORTRAN IV G, primarily for IBM 360/370 system.

There are two principal programs:

- 1) Univariate stochastic model identification
- 2) Univariate stochastic model preliminary estimation

A.1 Univariate Stochastic Model Identification - MAINI

The functions provided by this program are important in many stages of the development of stochastic models. The program provides valuable insights into the underlying structure of a given time series and it can rapidly determine the usefulness of particular models and forecasting systems.

The program identifies parsimonuous models for a given time series by comparing basic properties of the given data series with corresponding properties of the model. The program calculates the mean, standard deviations, absolute deviation, autocorrelation function, partial autocorrelation function, and solves the Yule-Walker equations ($\rho_j = \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \dots \quad j > 2$) to provide the parameter values of an AR model fitted to the data series. The program uses the printer to graph the autocorrelation and partial autocorrelation functions. The autocorrelation function indicates whether the given time series is non-stationary and the appropriate degree of differencing can be tested by the program. The autocorrelation function also indicates the order of the MA portion of the ARIMA model. The partial autocorrelation function indicates the order of the AR portion of the model.

This program is also used in the diagnostic checking stage of stochastic model development. The residual errors, obtained after the model is fitted to the original data, need to be checked for any remaining structure or correlations

which may indicate the need for model improvements. The functions of MAIN1 are performed by the following subroutine:

- BJDIFF - for seasonal and nonseasonal differencing operations
- BJSTAT - for calculation and display of the statistical properties of various data series
- BJRES - for calculation of residual errors and the one step ahead forecasts from the stochastic models
- BJPLOT - for plotting on the printer graphs of the autocorrelation and partial autocorrelation functions, and the given data and corresponding forecasts.

A.2 Univariate Stochastic Model Preliminary Estimation - MAIN2

This program computes estimates of the parameters of the stochastic models identified by MAIN1. Given the order of differencing required for the data series and AR portions of the model, estimates are obtained for each AR and MA parameters.

Once the parameter estimates are obtained, the residual errors are calculated by the program. This residual data series is tested to verify the goodness of fit of the model to the data series.

The functions of MAIN2 are performed using the same subroutines as needed in MAIN1 i.e. BJDIFF, BJSTAT, BJRES, and BJPLOT, and as well:

- BJMASR - for performing a search for the best values of one or two seasonal or non-seasonal MA parameters
- BJARMA - for calculation of the AR parameter residuals and forecasts of a combined ARIMA model.

A.3 Program Descriptions

A.3.1 Routine MAIN1 - Stochastic Model Identification.

Purpose: The purpose of this routine is to perform preliminary identification of an appropriate ARIMA model for the given data series. The program tests the original data series and several differenced series for stationarity. It fits a high order autoregressive model to the series to provide indications of simpler ARIMA models. It checks the statistical properties of the residuals to verify model needness of fit. It graphs the original data and forecasts using the printer.

Initialization: Input of the data series can be accomplished in many ways but for simplicity here the data is read in from a card deck. Besides needing an array DAT for the original data, arrays of the same size are needed for the differenced series w, the forecasts values FCST, and the residuals errors RES. The following parameters also need to be initialized -

DFORM - the format of the data series on the
data check

ITERM - the index of the initial term of the
DAT vector (normally equal to 1)

LTERM - the index of the last term of the DAT vector

NMODS - the number of models to be tested

NDIFF1 - a control parameter for the first differ-
encing operation

NDIFF2 - a control parameter for the second differencing operation (NDIFF1 and NDIFF2 can be any integer greater or equal to zero).
 IAVG - if IAVG = 1, the mean is subtracted from the differenced series w
 LMAX - the max. order of AR portion of the model to be fitted
 KMAX - the max. lag to calculated in the correlation function
 MAXV - the max. value of DAT and FCST to be plotted
 MINV - the min. value of DAT and FCST to be plotted
 PHFORM - the format to used for the header on the output plot of the data and forecast
 PDFORM - the format to be used for each line of the plot of I, DAT(I), and FCST(I).

Subroutines used - BJDIF
 BJSTAT
 BJRES
 BJPLOT

Data deck - The following is a description of the data deck required for execution of BOXJEN MAIN1 - identification

- A. The first card specifies the format to be used in the reading in of the data series. This card is read using 18A4 format and contains a format specification such as (8F10.2).
- B. One card containing values for ITERM and LTERM in format z110.

- C. Multiple cards containing the values of the data series the format given in card A.
- D. One card containing a value for NMODS, the number of models to be tested, in format I10.
- E. NMODS cards, each containing values of NDIFF1, NDIFF2, IAVG, KMAX, LMAX for one of the NMODS tests, in format 5110.
- F. One card containing values for MINV, MAXV, and NPTS in format 3I10.
- G. Three cards containing the header to be used for the plot header format PHFORM is of dimension 54 and each of the three cards is read using 18A4 format.
- H. One card containing the format PDFORM to be used for each line of the plot containing I, DAT(I), and FCST(I). e.g. (1X, IZ, 2F10.1, 5X, 101A1).

Capabilities - This program can handle input card data of any format. Series with more than 200 data points will require expansion of the dimensions in the main program only. The program carries out tests on NMODS different models as specified by the user. For each model, the program calls BJDIFF to calculate the required differences, BJSTAT to calculate the statistical properties of the differences, BJRES to calculate the residuals and forecasts, and BJSTAT again to calculate the statistics of the residual series. For the last model, the original DAT vector and the one step ahead forecasts FCST are plotted using the BJPLOT subroutine. The differencing

operations can be seasonal or non-seasonal. The residuals are derived for an AR model of order LMAX fitted to the differenced series.

A.3.2 Subroutine BJDIFF.

Purpose - The differencing technique carried out by this subroutine is a very powerful aid in the modeling of non-stationary series. This routine can carry out one or two seasonal or non-seasonal differencing operations and this capability is sufficient to transform most practical non-stationary time series into stationary series.

Calling sequence -

```
CALL BJDIFF (Y, W, ITERM, LTERM, NDIFF1, NDIFF2, IAVG)
```

where Y is the input data series,

W is the output data series,

ITERM is the index of the initial term in the Y series,

LTERM is the index of the last term in the Y series,

NDIFF1 is the control parameter for the first differencing operation,

NDIFF2 is the control parameter for the second differencing operation,

IAVG is a control parameter which allows the mean to be subtracted from the output W series.

Capabilities -

If the control parameter $NDIFFx = 0$ (where $x = 1$ or 2),

then no differencing operation is performed in step number x .

If $NDIFF1 = M$ where $M > 0$, the differencing operation

$W(I) = Y(I) - Y(I-M)$ is performed.

If $NDIFF1 = 0$, then $W(I) = Y(I)$.

If $NDIFF2 = N$ where $N > 0$, the differencing operation

$W(I) = W(I) - W(I - N)$ is performed.

If $NDIFF1 = M$ and $NDIFF2 = N$ are > 0 , the final result is

$$W(I) = W(I) - W(I - N) = Y(I) - Y(I - M) - Y(I - N) \\ + Y(I - M - N).$$

If $I_{AVG} = 1$, then the mean AVG of the W series is subtracted from each element of W .

Note the number of terms in the W series is fewer than in the Y series. For W , the first non-zero term, has an index $(I_{TERM} + NDIFF1 + NDIFF2)$.

A.3.3 Subroutine BJSTAT

Purpose - The purpose of this routine is to calculate and display the basic statistical properties of the input data series which are useful in the identification of

- 1) stationarity in the input series,
- 2) the orders of the AR and MA portions of the stochastic model. The routine also calculate the parameters of an AR model fitted to the input data series.

Calling sequence -

CALL BJSTAT (W, ITERM, KMAX, LMAX)

where W is the input data series

ITERM is the index of the initial term in the W series,

LTERM is the index of the last term in the W series.
 KMAX is the maximum correlation lag to be calculated,
 LMAX is the maximum order of the AR model to be fitted
 to the W series, and the order of the partial
 autocorrelation function.

Capabilities - The routine calculates the mean, variance, standard deviation, autocorrelation, and partial autocorrelation of the W series. KMAX can be set equal to any integer in the range 0 to 400. LMAX can be set equal to any integer in the range 0 to 25. KMAX should be greater than or equal to LMAX to allow calculation of the PHI array. The elements of the PHI array are the solution of the Yule-Walker equations. The partial autocorrelation function PACF(I) is equal to PHI(I, I). All the calculated results from this routine are stored in the common block STATCB. The autocorrelation, and the partial autocorrelation function are plotted on the printer using the BJPLOT subroutine.

Subroutines used - BJPLOT

A.3.4 Subroutine BJRES

Purpose - The purpose of this routine is to calculate the residuals (one step ahead forecast errors) from fitting an AR model to the input W series. One step ahead forecasts are also calculated by subtracting the residual from the original data.

Calling sequence -

CALL BJRES (DAT, W, FCST, RES, IWTERM, LTERM, LMAX)

where DAT is the original data series,

W is the series to be fitted with the AR model,

FCST is the output forecast series,
 RES is the output residual error series,
 IWTERM is the initial term of the W series,
 LTERM is the last term of the W series,
 LMAX is the order of the AR model to be fitted.

The common block STATCB contains previously calculated AR parameters for the series.

Capabilities - Forecasts and residuals are generated by this routine for terms with indices (IWTERM + LMAX) to LTERM. For the first LMAX term, the forecast is equated to the data, and the residual to zero. If LMAX = 0, the total model fitted is assumed to have only MA components, and $RES(I) = W(I)$, $FCST(I) = DAT(I) - W(I)$. The AR parameters $PHI(LMAX, L)$ are calculated by the immediately preceding call to subroutine BJSTAT (with W as input series).

Subroutines used - None.

A.3.5 Subroutine BJPLOT

Purpose - The purpose of this routine is to allow graphical display of pairs of functions such as the autocorrelation and partial autocorrelation functions, or the original data series and the corresponding forecasts.

Calling sequence -

CALL BJPLOT (X, Y, PFORM)

Where X and Y are the two input vectors,

PFORM is the format to be used in plotting each line on that graph,

the common block PLOTCB contains other control parameters,

MINV is the minimum value to be plotted,
 MAXV is the maximum value to be plotted,
 ILINE is the index of the first element in the vector
 X and Y which is to be plotted,
 LLINE is the index of the last element in the vector
 X and Y which is to be plotted,
 NEWPAG is a control parameter which can start the
 plotted output on a new page,
 NPTS is the number of possible points which can be
 used in plotting each element of X and Y between
 MINV and MAXV.

Capabilities - The variable format PFORM is initialized in the
 calling routine.

e.g. DIMENSION PFORM(7)

DATA PFORM/'(1X, I5, 2F10.2, 5X, 101A1)'/

Each output line consists of the index (or line number) I,
 the value of the element X(I), the value of the element Y(I),
 and an X and * representing X and Y placed at the appropri-
 ate point between MINV (point 1) and MAXV (point NPTS).

Subroutine used - None.

A.3.6 Routine MAIN2 - Preliminary Estimation

Purpose - The purpose of this routine is to calculate preliminary
 estimates of the parameters of an ARIMA model.

$$GPHI(B)W(I) = T(B)A(I)$$

$$\phi(B)W_t = \theta(B)a_t$$

where B is the backward shift operator

$$GPHI(B) = (1 - PHI(LMAX, 1)B - PHI(LMAX, 2)B^2 \dots)$$

$$T(B) = (1 - T1B - T2B^2) \text{ for non-seasonal models}$$

$$= (1 - T1B)(1 - T2B^{LSEA}) \text{ for seasonal models}$$

W(I) is the appropriately differenced data series.

A(I) is the residual one step ahead forecast errors series.

Initialization - The model structure used by this routine is assumed to be previously identified using program MAIN1 - Preliminary identification. Hence required for this program are the original data series DAT, the differenced data series W, the degree of differencing NDIFF1 and NDIFF2 and the average control IAVG. The MA portion of the model is assumed to belong to the very broad class of two parameter seasonal or nonseasonal models. In a typical seasonal model, NDIFF1 = 1 and LSEA = NDIFF2 = M, where M is the length of the seasonal cycle. T1 is the parameter associated with the MA term at lag 1. T2 is the parameter associated with the MA term at lag 2 for non-seasonal models, and at lag LSEA for seasonal models. LSEA is the control parameter which can specify the use of the seasonal or non-seasonal models, and also the length of the seasonal cycle.

If LSEA > 2, the seasonal model with cycle length LSEA is used. In finding the best estimates for the two MA parameters T1 and T2, a grid search is performed. The following parameters specify the range and number of searches -

NSRT1 - the number of searches on T1 between the values
T1MIN and T1MAX,

T1MIN - the minimum range of T1 used in the search,

T1MAX - the maximum range of T1 used in the search,
NSRT2 - the number of searches on T2 between the values
T2MIN and T2MAX.
T2MIN - the minimum value of T2 used in the search,
T2MAX - the maximum value of T2 used in the search,
KMAX - the maximum lag to be calculated in autocorrela-
tion function,
LMAX - the order of the AR portion of the model to
fitted if LMAX = 0 only the MA portion is fitted.
Control parameters for use in plotting the DAT and FCST
vectors also need to be specified as in MAIN1.

Subroutines used - BJDIFP

BJSTAT

BJMASR

BJARMA

BJRES

BJPLOT

Data deck - Same as used in MAIN1 except D and E

D - One card containing values for NDIFF1, NDIFF2, LAVG,
kmax, lmax in format 5I10.

Capabilities - In a similar fashion to the MAIN1 program, this
routine can handle input card data of any format. The
program search over a grid of find the best values for
the (one or two) MA parameters of an ARIMA model. The
standard deviation of the residuals is used as the criterion
for selecting the best parameters. The degree of seasonal
or non-seasonal differencing is specified by NDIFF1 and

NDIFF2. The order of the AR portion of the ARIMA model is specified by LMAX. Once the best value for the MA parameters have been found, the program re-calculates the residuals and forecasts for this best model and plots the results using the BJPLOT routine.

A.3.7 Subroutine BJARMA

Purpose - When given the suitably differenced input series, and values for the MA parameters, this routine calculates the parameters of the AR portion of a ARIMA model. The routine also calculates the residuals after both portions of the model are fitted, and calculates the one step ahead forecasts by subtracting the residuals from the original data.

Calling sequence -

```
CALL BJARMA(DAT, W, FCST, RES, E, IWTERM, LTERM, KMAX, LMAX, T1,
            T2, LSEA)
```

where DAT is the original data series,

W is the differenced series,

FCST is the calculated forecast series,

RES is the calculated residual series,

E is the series of residual errors from fitting only
the MA portion,

IWTERM is the last term of the W series,

KMAX is the maximum autocorrelation function lag,

LMAX is the maximum order of the AR portion of the model,

T1 is the MA parameter for lag 1,

T2 is the MA parameter for lag 2, or lag LSEA, LSEA is
the length of the seasonal cycle.

Capabilities-

If LSEA > 2, the seasonal model with a seasonal cycle length of LSEA is used, otherwise a MA model (order 2) is used.

If T1 = 0, the first order MA portion of the model is not used.

If T2 = 0, the second order MA portion of the model is not used.

If LMAX = 0, the model is pure MA with no AR portion.

Note the forecasts and residuals are calculated only for term from (LMAX + IWTERM) to LTERM or from (LSEA + 2 + LMAX) to LTERM (if IWTERM - LSEA - 1 less than or equal to 0).

Subroutines used - BJSTAT

BJRES

A.3.8 Subroutine BJMASR.

Purpose - The purpose of this routine is to carry out a grid search for the best estimates of the two MA parameters T1 and T2 for seasonal or non-seasonal model.

Calling sequence-

```
CALL BJMASR(DAT, W, FCST, RES, E, IWTERM, LTERM, KMAX, LMAX,
            T1, T2, LSEA)
```

where these calling parameters are the same as those described for BJARMA. The common block MASR also contains the following input parameters specifying the number and range of the searches on the parameters T1 and T2 - NSRT1, T1MIN, T1MAX, NSRT2, T2MIN, T2MAX. For a description of these parameters and their initialization see MAIN2.

Capabilities - The search on each parameter starts at the minimum value and proceeds to the maximum value in the grid. The criterion used for selection of the best parameters is the standard deviation SIGMA of the residual errors. The best values of T1 and T2 are returned in the calling parameters of the routine.

Subroutines used - EJARMA
BJSTAT

A.4 Program Listing

MAIN1 - Identification

BJDIFF

BJSTAT

BJRES

BJPLOT

MAIN2 - Preliminary Estimation

BJMASR

EJARMA

C MAIN1 -- PRELIMINARY IDENTIFICATION

```

0001 DIMENSION DFORM(18), PDFORM(13), PHFORM(54)
0002 DIMENSION DAT(200), W(200), FCST(200), RES(200)
0003 COMMON/PLOT/ MINV, MAXV, ILINE, LLINE, NEWPAG, NPTS
0004 COMMON/STATCB/ C(400), PHI(25,25), PACF(25), FMEAN, ABDEV, SIGMA
0005 WRITE(6,5)
0006 READ(5,100) DFORM
0007 READ(5,130) ITERM, LTERM
0008 READ(5, DFORM)(DAT(I), I=1:ITERM, LTERM)
0009 WRITE(6,15)(I, DAT(I), I=1:ITERM, LTERM)
0010 READ(5,130) NMCD5
0011 DO 120 NM = 1, NMCD5
0012 READ(5,130) NDIFF1, NDIFF2, IAVG, KMAX, LMAX
0013 CALL BJDIFF(DAT, W, ITERM, LTERM, NDIFF1, NDIFF2, IAVG)
0014 ITERM = ITERM + NDIFF1 + NDIFF2
0015 CALL BJSTAT(W, ITERM, LTERM, KMAX, LMAX)
0016 CALL BJRES(DAT, W, FCST, RES, ITERM, LTERM, LMAX)
0017 ITERM = ITERM + LMAX
0018 CALL BJSTAT(RES, ITERM, LTERM, KMAX, LMAX)
0019 120 CONTINUE

C
0020 READ(5,130) MINV, MAXV, NPTS
0021 READ(5,100) PHFORM
0022 WRITE(6, PHFORM)
0023 READ(5,100) PDFORM
0024 ILINE = ITERM
0025 LLINE = LTERM
0026 NEWPAG = 0
0027 CALL BJPLOT(DAT, FCST, PDFORM)

C
0028 100 FORMAT(18A4)
0029 5 FORMAT(1X, 'BOXJEN PRELIMINARY IDENTIFICATION')
0030 15 FORMAT(1X, 'INPUT DATA', //, 6(14, F12.2))
0031 130 FORMAT(5I10)
0032 STOP
0033 END

```

```

0001 SUBROUTINE BJDIFF( Y,W,ITERM,LTERM,NDIFF1,NDIFF2,IAVG)
0002 DIMENSION W(LTERM),Y(LTERM)
0003 THIS ROUTINE TRANSFORMS THE SERIES Y BY ONE OR TWO, SEASONAL
0004 NON-SEASONAL DIFFERENCING OPERATIONS AND IT STORES THE RESULT
0005 IN W. IF NDIFFX (WHERE X EQUALS 1 OR 2) = 0, NO DIFFERENCING
0006 PERFORMED IN OPERATION NUMBER X. IF NDIFFX = M WHERE M > 0,
0007 THE DIFFERENCING W(I) = Y(I) - Y(I-M) IS PERFORMED. IF IAVG = 1,
0008 OR IF X = 2, W(I) = W(I) - W(I-M) IS PERFORMED. IF IAVG = 1,
0009 MEAN OF W(I) IS SUBTRACTED FROM EACH W(I).
0010 WRITE(6,60) ITERM,LTERM,NDIFF1,NDIFF2,IAVG
0011 60 FORMAT(1H1,'BJDIFF',//,1X,'ITERM = ',I10,
0012 *,1X,'LTERM = ',I10,
0013 *,1X,'NDIFF1 = ',I10,
0014 *,1X,'NDIFF2 = ',I10,
0015 *,1X,'IAVG = ',I10)
0016 N1 = ITERM + NDIFF1
0017 N2 = N1 + NDIFF2
0018 DO 5 I = ITERM,N2
0019 W(I) = 0
0020 5 CCNTINUE
0021 IF( NDIFF1.GE.1) GO TO 10
0022 DO 20 I = N1,LTERM
0023 W(I) = Y(I)
0024 20 CCNTINUE
0025 GO TO 30
0026 10 CCNTINUE
0027 DO 25 I = N1,LTERM
0028 W(I) = Y(I) - Y(I-NDIFF1)
0029 25 CCNTINUE
0030 30 CCNTINUE
0031 IF ( NDIFF2.LE.0) GC TO 40
0032 DO 35 I = N2,LTERM
0033 J = LTERM + N2 - I
0034 W(J) = W(J) - W(J-NDIFF2)
0035 35 CCNTINUE
0036 40 CCNTINUE
0037 C
0038 AVG = 0
0039 DO 65 I = N2,LTERM
0040 AVG = AVG + W(I)
0041 65 CCNTINUE
0042 C

```

```
0029  
0030  
0031  
0032  
0033  
0034  
0035  
0036  
0037  
0038  
  
65 CCNTINUE  
   AVG = AVG / (LTERM - N2 + 1)  
   WRITE(6,75) AVG  
75  FORMAT(IX,'AVG = ',F10.2)  
   IF(IAVG.NE.1) RETURN  
   DO 70 I = N2,LTERM  
   W(I) = W(I) - AVG  
70 CCNTINUE  
   RETURN  
   END
```

```

0001 SUBROUTINE BJSTAT(W, ITERM, LTERM, KMAX, LMAX)
0002 C
0003 COMMON/STATCB/ C(400), PHI(25,25), PACF(25), FMEAN, ABDEV, SIGMA
0004 COMMON/PLOTCB/MINV, MAXV, ILINE, LLINE, NEWPAG, NPTS
0005 DIMENSION PFORM(9)
0006 DIMENSION W(LTERM)
0007 DATA PFORM/'1X', '14,2', 'X,F1', '5.2', '2X,F', '15.2', '12X', '70A',
0008 *, '1) ' /
0009 C THIS PROGRAM CALCULATES STATISTICAL PROPERTIES OF THE INPUT
0010 C SERIES W. ITERM AND LTERM ARE THE INDICES OF THE INITIAL AND
0011 C LAST TERMS OF THE W SERIES. KMAX IS THE MAXIMUM ORDER OF THE
0012 C AUTOCORRELATION FUNCTION TO BE CALCULATED. LMAX <= KMAX IS THE
0013 C ORDER OF THE AUTOREGRESSIVE MODEL WHICH IS FITTED TO THE W SERIES.
0014 NTERM = LTERM - ITERM + 1
0015 IF( NTERM.LE.0) WRITE(6,15) ITERM, LTERM
0016 15 FORMAT(1X, ' NTERM, THE NUMBER OF TERMS IN THE SERIES, IS LE 0',
0017 *, ' ITERM =', 15, ' LTERM =', 15)
0018 IF( NTERM.LE.0) NTERM= 1
0019 FMEAN = 0
0020 VAR = 0
0021 ABDEV = 0
0022 DO 440 I= ITERM, LTERM
0023 FMEAN=FMEAN+W(I)
0024 FMEAN = FMEAN/NTERM
0025 DO 435 I = ITERM, LTERM
0026 ABDEV = ABDEV + ABS(W(I) - FMEAN)
0027 ABDEV = ABDEV/NTERM
0028 DO 450 I= ITERM, LTERM
0029 VAR = VAR+(W(I)-FMEAN)**2
0030 VAR = VAR/NTERM
0031 SIGMA = SQRT (VAR)
0032 IF(KMAX.GT.400) KMAX = 400
0033 IF(KMAX.LE.LMAX) KMAX = LMAX
0034 IF(KMAX.LE.0) GO TO 715
0035 DO 600 KS = 1, KMAX
0036 C(KS) = 0
0037 J = LTERM-KS
0038 DO 605 I=ITERM, J
0039 C(KS) = C(KS)+(W(I)-FMEAN)*(W(I+KS)-FMEAN)
0040 C(KS) = C(KS)/(VAR*NTERM)
0041 600 CONTINUE
0042 PHI(1,1) = C(1)

```

```

0035 IF(LMAX.LE.1) GO TO 662
0036 DO 660 L=2,LMAX
0037 S1 = 0
0038 S2 = 0
0039 JMAX = L-1
0040 DO 650 J=1,JMAX
0041 S1 = S1+PHI(L-1,J)*C(L-J)
0042 S2 = S2+PHI(L-1,J)*C(J)
0043 650 CONTINUE
0044 PHI(L,L) = (C(L)-S1)/(1-S2)
0045 DO 640 J=1,JMAX
0046 PHI(L,J) = PHI(L-1,J)-PHI(L,L)*PHI(L-1,L-J)
0047 660 CONTINUE
0048 662 CONTINUE
0049 IF(LMAX.LE.0) GO TO 720
0050 DO 1005 I = 1,LMAX

```

```

0051 1005 PACF(I) = PHI(I,I)
0052 IF(KMAX.LE.LMAX) GO TO 705
0053 720 CONTINUE
0054 K1= LMAX+1
0055 DC 710 I = K1, KMAX
0056 710 PACF(I) = 0
0057 705 CCNTINUE
0058 715 CONTINUE
0059 WRITE(6,1004) ITERM,LTERM,KMAX,LMAX,W(LTERM),FMEAN,
* ABDEV,VAR,SIGMA
0060 1004 FORMAT(/,1X,'BJSTAT ',
//,1X,'INPUT PARAMETERS ITERM =',I15,
//,21X,'LTERM =',I15,
//,21X,' KMAX =',I15,
//,21X,' LMAX =',I15,
//,1X,'INPUT SERIES VALUES ITERM =',F16.3,
//,21X,'LTERM =',F16.3//,
*1X,'CALCULATED RESULTS MEAN =',F15.2,
//,17X,'ABS. DEV. =',F15.2,
//,22X,'VAR. =',F15.2,
//,17X,'STD. DEV. =',F15.2,
//,2X,'LAG',2X,'AUTOCORRELATION',2X,'PARTIAL ACCRREL',
//,19X,'X',16X,'*',13X,'-1',2X,'-',.8',2X,'-.6',2X,'-.4',2X,'-.2',4X,
*0',3X,'.2',3X,'.4',3X,'.6',3X,'.8',3X,'+1',
//,51X,1HI,10(4(' '),1HI))
0061 NPTS = 51
0062 ILINE = 1
0063 LLINE = KMAX
0064 NEWPAG = 0
0065 MINV = -1
0066 MAXV = 1
0067 IF( KMAX.LE.0) RETURN
0068 CALL BJPLOT(C,PACF,PFORM)
0069 IF(LMAX.GT.0) WRITE(6,20) (I,PHI(LMAX,I),I = 1,LMAX )
0070 20 FCRMAT(/,1X,'PARAMETERS OF AN AUTOREGRESSIVE MDEL',
//,7(1X,I5,F7.2))
RETURN
0071 END
0072

```



```

0001      SUBROUTINE BJPLOT(X,Y,PFORM)
C
0002      DIMENSION PFORM(1)
0003      DIMENSION X(LLINE),Y(LLINE)
0004      DIMENSION IPLOT(135)
0005      DIMENSION ICHAR(5)
0006      COMMON/PLCTCB/MINV,MAXV,ILINE,LLINE,NEWPAG,NPTS
0007      DATA ICHAR/ ' ', 'G', 'X', '*', '.,', '.,' /
C
C      NPTS IS THE NUMBER OF PCINTS WHICH MAY BE PLOTTED ON EACH
C      LINE. MINV AND MAXV ARE THE VALUES OF THE FIRST AND LAST POINT
C      ON EACH LINE.
0008      IF(NEWPAG.EQ.1) WRITE(6,1000)
0009      FCFORMAT(IHL)
0010      NS = NPTS - 1
0011      DO 30 I= ILINE,LLINE
0012          DO 10 IP = 1,NPTS
0013              10 IPLOT(IP)=ICCHAR(1)
0014              IZERO = -MINV* NS / (MAXV-MINV) + 1.5
0015              IF(IZERO.LE.1) IZERO = 1
0016              IPLOT(IZERO) = ICHAR(5)
0017              IP = (Y(I) - MINV)* NS / (MAXV-MINV) + 1.5
0018              IF(IP.GE.NPTS) IP = NPTS
0019              IF(IP.LE.1) IP = 1
0020              IPLOT(IP)=ICCHAR(4)
0021              IP = (X(I) - MINV)* NS / (MAXV-MINV) + 1.5
0022              IF(IP.GE.NPTS) IP = NPTS
0023              IF(IP.LE.1) IP = 1
0024              IPLOT(IP)=ICCHAR(3)
0025              WRITE (6,PFCRM ) I,X(I),Y(I),(IPLOT(IP),IP= 1,NPTS)
0026          30 CCNTINUE
0027      RETURN
0028      END

```

```

C
C
C
0001      MAIN2 - PRELIMINARY ESTIMATION
0002      DIMENSION DFCRM(18),PDFCRM(18),PHFCRM(54)
0003      DIMENSION DAT(200),W(200),FCST(200),RES(200),E(200)
0004      COMMON/PLCTCB/MINV,MAXV,ILINE,LLINE,LLINE,NEWPAG,NPTS
0005      COMMON/STATCB/ C(400),PHI(25,25),PACF(25),FMEAN,ABDEV,SIGMA
0006      COMMON/MASR/NSRT1,I1MIN,I1MAX,NSRT2,I2MIN,I2MAX
0007      WRITE(6,805)
0008      READ(5,100) DFCRM
0009      READ(5,130) ITERM,LTERM
0010      READ(5,DFCRM)( CAT(I),I = ITERM,LTERM)
0011      WRITE(6,15)(I,DAT(I),I = ITERM,LTERM)
0012      READ(5,130) NDIFF1,NDIFF2,IAVG,KMAX,LMAX
0013      CALL BJDIFF(DAT,h,ITERM,LTERM,NDIFF1,NDIFF2,IAVG)
0014      ITERM = ITERM + NDIFF1 + NDIFF2
0015      READ(5,160) NSRT1,I1MIN,I1MAX,NSRT2,I2MIN,I2MAX,LSEA
0016      CALL BJMASR(DAT,W,FCST,RES,E,IWTERM,LTERM,KMAX,L1,T2,LSEA)
0017      CALL BJARMA(DAT,W,FCST,RES,E,IWTERM,LTERM,KMAX,L1,T2,LSEA)
0018      ITERM = IWTERM + LMAX
0019      IF(LMAX.GT.0) CALL BJSTAT(RES,ITERM,LTERM,KMAX,LMAX)
C
0020      READ(5,130) MINV,MAXV,NPTS
0021      READ(5,100) PHFORM
0022      WRITE(6,PHFCRM)
0023      READ(5,100) PDFCRM
0024      ILINE = ITERM
0025      LLINE = LTERM
0026      NEWPAG = 0
0027      CALL BJPLOT(DAT,FCST,PDFCRM)
C
0028      100 FORMAT(18A4)
0029      15  FORMAT(1X,'INPUT DATA',/,6(14,F12.2))
0030      805  FORMAT(1X,'BOXJEN PRELIMINARY ESTIMATION')
0031      130  FORMAT(5I10)
0032      160  FORMAT(10,2F10.2,110,2F10.2,110)
0033      STOP
0034      END

```

```

0001 SUBROUTINE BJMASR(CAT,W,FCST,RES,E,IWTERM,LTERM,KMAX,LMAX,T1,T2,
      *LSEA)
C
C THIS ROUTINE SEARCHES OVER A GRID TO FIND THE BEST VALUES FOR TWO
C (SEASONAL OR NON-SEASONAL) MOVING AVERAGE PARAMETERS T1 AND T2
0002 CCMCMC/MASR/NSRT1,T1MIN,T1MAX,NSRT2,T2MIN,T2MAX
0003 COMMON/STATCB/ C(400),PHI(25,25),PACF(25),FMEAN,ABDEV,SIGMA
0004 DIMENSION ST1(10),ST2(10),SIG(10,10)
0005 DIMENSION DAT(LTERM),W(LTERM),FCST(LTERM),RES(LTERM),E(LTERM)
0006 WRITE(6,805)NSRT1,T1MIN,T1MAX,NSRT2,T2MIN,T2MAX
0007 FORMAT(/,'BJMASR',/,1X,'NSRT1 =',I10,' T1MIN =',F10.2,
      *' T1MAX =',F10.2,/,1X,'NSRT2 =',I10,' T2MIN =',F10.2,
      *' T2MAX =',F10.2)
      BRMS = 100000
      NONT = 0
      DO 5 NT2 = 1,NSRT2
      DO 10 NT1 = 1,NSRT1
      NONT = NONT + 1
      T1 = T1MIN
      T2 = T2MIN
      IF(NT1.GT.1) T1= T1MIN + (T1MAX-T1MIN)*(NT1-1)/(NSRT1-1)
      IF(NT2.GT.1) T2= T2MIN + (T2MAX-T2MIN)*(NT2-1)/(NSRT2-1)
      ST1(NT1) = T1
      ST2(NT2) = T2
      CALL BJARMA(CAT,W,FCST,RES,E,IWTERM,LTERM,KMAX,LMAX,T1,T2,LSEA)
      IRTERM = IWTERM + LMAX
      IF(LMAX.GT.0)CALL BJSTAT(RES,IRTERM,LTERM,KMAX,LMAX)
      SIG(NT1,NT2) = SIGMA
      IF( SIGMA.GE.BRMS) GO TO 20
      BRMS = SIGMA
      BT1 = T1
      BT2 = T2
20 CONTINUE
      WRITE(6,30) NONT,BT1,BT2,BRMS,T1,T2,SIGMA
30 FORMAT(/,1X,'SEARCH NUMBER ',I5,/,
      *1X,'BT1 =',F7.2,3X,'BT2 =',F7.2,3X,'BRMS =',F7.2,3X,
      *'T1 =',F7.2,3X,'T2 =',F7.2,3X,'SIGMA =',F7.2)
10 CONTINUE
5 CONTINUE
      T1 = BT1
      T2 = BT2
      WRITE(6,90) (ST2(J2),J2 = 1,NSRT2)
0030
0031
0032
0033
0034

```

```

0035 90 FORMAT(1H1,'SEARCH GRID',//,IX,'TABLE OF SIGMA VALUES',//
0036 *,20X,'T2 PARAMETER VALUES',//,IX,15X,'*',4X,10F10.2)
0037 WRITE(6,95)
0038 95 FCRMAT(//,IX,132('*'),//,IX,'T1 PARAMETER',//,IX,'VALUES')
0039 DO 100 J1 = 1,NSRT1
0040 WRITE(6,115) ST1(J1),(SIG(J1,J2),J2 = 1,NSKT2)
0041 100 CONTINUE
0042 115 FCRMAT( /,IX,F10.2,5X,'*',4X,10F10.2)
0043 WRITE(6,130) BT1,BT2,BRMS
0044 130 FCRMAT(1X,//,IX,'BEST PARAMETERS',//,IX,'T1 = ',F10.2,/,IX,
0045 *,T2 = ',F10.2,/,IX,'BRMS = ',F10.2)
0046 RETURN
0047 END

```

0001 SUBROUTINE BJARMA(DAT,W,FCST,RES,E,IWTERM,LTERM,KMAX,LMAX,T1,T2,
*LSEA)

C
C GIVEN THE SUITABLY DIFFERENCED SERIES W(I), AND VALUES FOR THE
C MOVING AVERAGE PARAMETERS T1,T2 AND LSEA, THIS ROUTINE DERIVES
C THE PARAMETERS OF THE AUTOREGRESSIVE PORTION OF A MIXED ARMA
C MODEL - T(B)*RES(I)=GPHI(B)*W(I)
C WHEKE B IS THE BACKWARD SHIFT OPERATER,
C WHEKE T(B) = 1 - T1*B - T2*B**2 FOR NON-SEASONAL MODELS,
C WHEKE T(B) = (1-T1*B)(1-T2*B**LSEA) FOR SEASONAL MODELS,
C WHEKE RES(I) ARE THE RESIDUALS,AND
C WHEKE GPHI(B) = (1-PHI(LMAX,1)*B - PHI(LMAX,2)*B**2-...).
C IF LSEA.GE.2 , THE SEASONAL MODEL IS USED, CITHERWISE THE
C NON- SEASONAL MODEL IS USED. THE RESIDUALS AND CNE STEP AHEAD
C FORECASTS ARE ALSO CALCULATED.

0002 DIMENSION DAT(LTERM),W(LTERM),FCST(LTERM),RES(LTERM),E(LTERM)
0003 CCMON/STATCB/ C(400),PHI(25,25),PACF(25),FMEAN,ABDEV,SIGMA
0004 WRITE(6,805)T1,T2,LSEA,IWTERM
0005 805 FORMAT(1H1 , 'BJARMA', //, 1X, 'T1 =', F10.2, /, 1X, 'T2 =', F10.2,

*, 1X, 'LSEA =', I10, /, 1X, 'IWTERM=', I10)
IF((IWTERM - LSEA - 1).LE.0) IWTERM = LSEA + 2
DO 10 I = 1, IWTERM
E(I) = 0

10 CONTINUE
DO 20 I = IWTERM, LTERM
IF(LSEA.GE.2) E(I)=W(I)+T1*E(I-1)+T2*E(I-LSEA)-T1*T2*E(I-LSEA-1)
IF(LSEA.LT.2) E(I)=W(I)+T1*E(I-1)+T2*E(I-2)
20 CONTINUE

C E(I) ARE THE RESIDUALS FRM FITTING ONLY THE MCV.AVG. PORTION.
C IF LMAX.LE.0 , THESE ARE THE FINAL RESIDUALS AND RES(I) = E(I).
C IF LMAX > 0, AN AUTOREGRESSIVE MODEL IS FITTED TO THE E(I).
CALL BJSTAT(F,IWTERM,LTERM,KMAX,LMAX)
CALL BJRES(DAT,E,FCST,RES,IWTERM,LTERM,LMAX)
RETURN
END

0014
0015
0016
0017

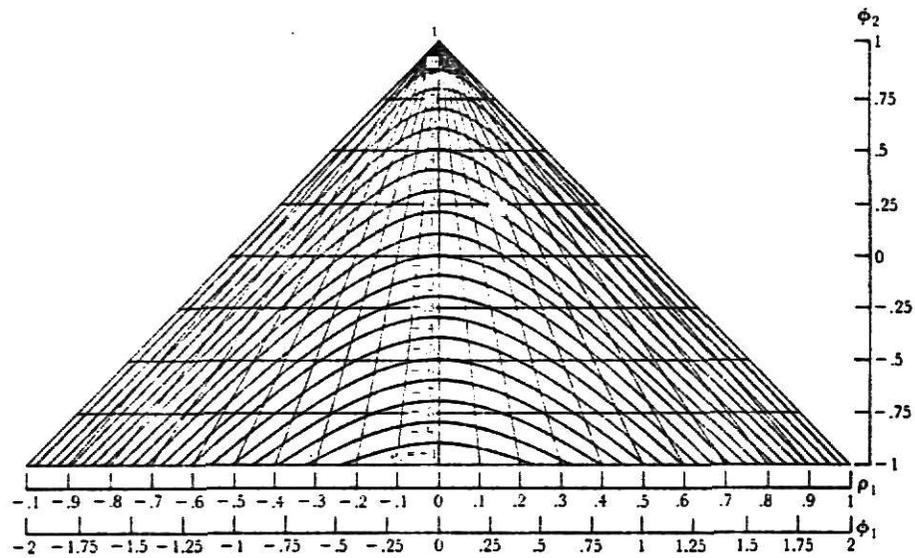
APPENDIX B
COLLECTION OF TABLES AND CHARTS

Table B.1. Table relating ρ_1 to θ for MA(1) process

θ	ρ_1	θ	ρ_1
0.00	0.000	0.00	0.000
0.05	-0.050	-0.05	0.050
0.10	-0.099	-0.10	0.099
0.15	-0.147	-0.15	0.147
0.20	-0.192	-0.20	0.192
0.25	-0.235	-0.25	0.235
0.30	-0.275	-0.30	0.275
0.35	-0.315	-0.35	0.315
0.40	-0.349	-0.40	0.349
0.45	-0.374	-0.45	0.374
0.50	-0.400	-0.50	0.400
0.55	-0.422	-0.55	0.422
0.60	-0.441	-0.60	0.441
0.65	-0.457	-0.65	0.457
0.70	-0.468	-0.70	0.468
0.75	-0.480	-0.75	0.480
0.80	-0.488	-0.80	0.488
0.85	-0.493	-0.85	0.493
0.90	-0.497	-0.90	0.497
0.95	-0.499	-0.95	0.499
1.00	-0.500	-1.00	0.500

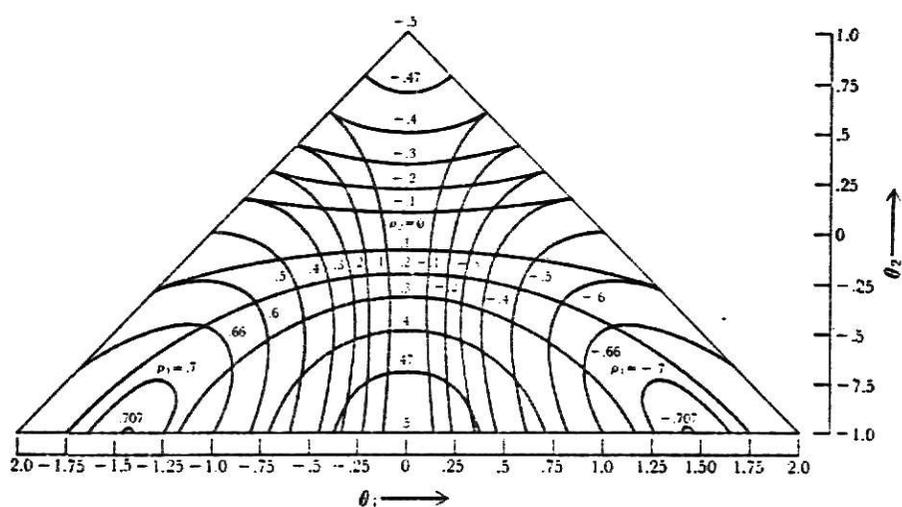
The table may be used to obtain first estimates of the parameters in the $(0, d, 1)$ process $w_t = (1 - \theta B)a_t$, where $w_t = \nabla^d Z_t$, by substituting $\gamma_1(w)$ for ρ_1 .

Chart B.2. Chart relating ρ_1 and ρ_2 to ϕ_1 and ϕ_2 for AR(2) process



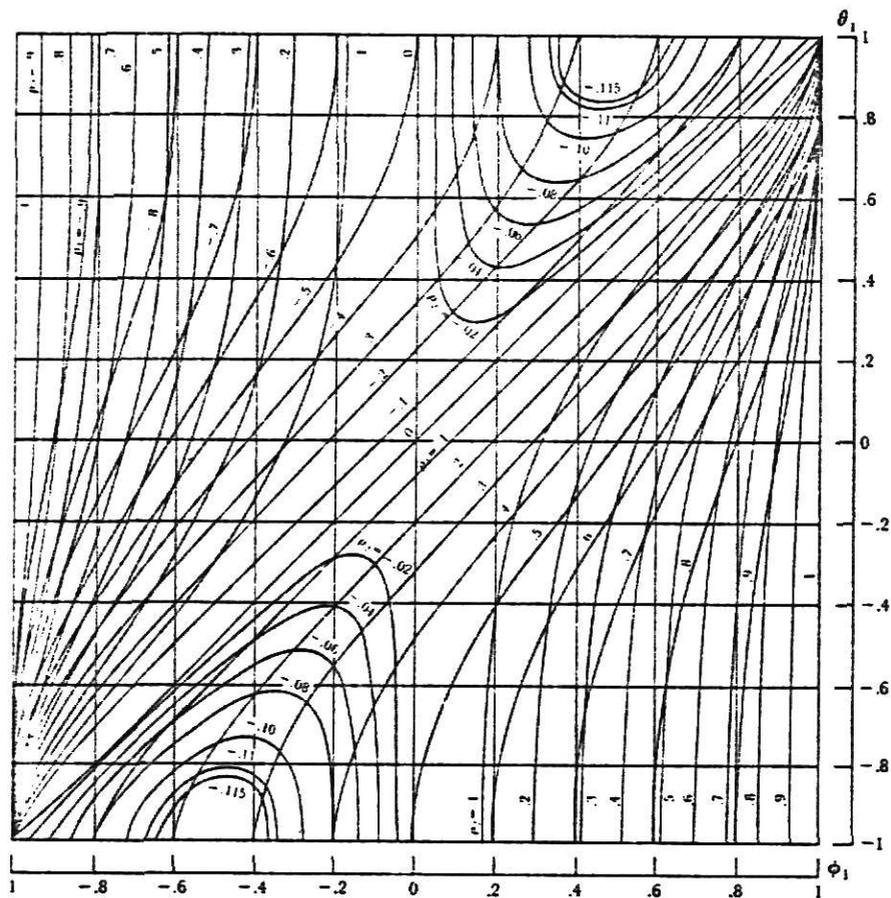
The chart may be used to obtain estimates of the parameters in the (2, d, 0) process: $(1 - \phi_1 B - \phi_2 B^2)w_t$, where $w_t = \nabla^d Z_t$, by substituting $\gamma_1(w)$ and $\gamma_2(w)$ for ρ_1 and ρ_2 .

Chart B.3. Chart relating ρ_1 and ρ_2 to θ_1 and θ_2 for a MA(2) process



The chart may be used to obtain first estimates of the parameters in the $(0, d, 2)$ process $w_t = (1 - \theta_1 B - \theta_2 B^2)a_t$, where $w_t = \nabla^d z_t$, by substituting $\gamma_1(w)$ and $\gamma_2(w)$ for ρ_1 and ρ_2 .

Chart B.4. Chart relating ρ_1 and ρ_2 to ϕ and θ for a ARMA(1,1) process



The chart may be used to obtain first estimates of the parameters in the (1, d, 1) process $(1 - \phi B)w_t = (1 - \theta B)a_t$, where $w_t = \nabla^d Z_t$, by substituting $\gamma_1(w)$ and $\gamma_2(w)$ for ρ_1 and ρ_2 .

TIME SERIES STOCHASTIC PROCESS
AND FORECASTING

by

TONY LEE-CHUIN CHIEN

Diploma, Tamkang College of Arts and Sciences, 1965

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Computer Science

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1979

Time series analysis is concerned with data which are not independent, but serially correlated, and where the relation between consecutive observations are of interest. It is a rapid growth area in statistical practice.

This report is concerned with the building of stochastic models for discrete time series in the time-domain and the use of such models in forecasting. The Box-Jenkins approach, developed in the 60's, is attracting more and more attraction. There are now very few disciplines in the sciences, business and technology, which are not investigating its possibilities.

The objective of this report is to build the models possessing maximum simplicity and the minimum number of parameters constant with representational adequacy by using Box-Jenkins approach.