APPLICATION OF GENERAL AUTOREGRESSIVE MOVING-AVERAGE STOCHASTIC MODELS TO TIME SERIES AND SIMULATION PROBLEM

by

Steven Shyan-Ming Ko

B.S., Taiwan Cheng-Kung University

Taiwan, R.O.C., 1966

5248

A MASTER'S THESIS

submitted in partial fulfillment of the

requirement for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1971

Approved by:

L. E. Grosh

Major Professor

THIS BOOK CONTAINS NUMEROUS PAGES WITH THE ORIGINAL PRINTING BEING SKEWED DIFFERENTLY FROM THE TOP OF THE PAGE TO THE BOTTOM.

THIS IS AS RECEIVED FROM THE CUSTOMER.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my major professor, Dr. L. E. Grosh, for his guidance in the preparation of this thesis. He not only suggested this research work but, more importantly, aroused my great interest in statistics.

Finally, I want to express thanks to Mrs. Marie Jirak for her invaluable assistance in typing this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

DESCRIPTION OF ARMA MODEL

Box and Jenkins define a time series as a set of observations generated sequentially in time [1]. If the set is continuous, the time series is said to be continuous. If the set is discrete, the time series is said to be discrete. In this thesis, we consider only the discrete time series where observations are made at some fixed interval h. However, the value of the time interval h is often unimportant in the appropriate model for the given time series.

As indicated by Casimer Micheal Stralkowski [2], a desirable mathematical analysis of a time series should be general enough to accommodate all types of the time series and should be embody the following qualities:

1. Parsimony of the model parameter, i.e. the models should contain as few parameters as possible.
2. The model should be simple to interpret and apply.
3. The model should accommodate theoretical as well as empirical information, i.e. should be empirical-mechanistic in nature.

Box and Jenkins introduce the autoregressive model as a mathematical model which is extremely useful in the representation of certain practically occuring series. In this model, the current value of the process is expressed as a finite, linear aggregate of previous values of the process and a random shock a_t. [3].

Another kind of model, of great practical importance in the representation of the observed time series, is the so-called finite moving average process. Box also introduces the moving average model as making $\tilde{z}_t$, i.e., is the time series observation z_i minus its mean $\bar{z}$, linearly dependent on a finite number of previous random shocks. [4]

To achieve greater flexibility in the fitting of actual time series, it is sometimes advantageous to include both autoregressive and moving-average terms in the model. This model is the so-called autoregressive moving-average model.

When the general autoregressive moving-average model is mentioned later on in this thesis, it includes all possible models, either the autoregressive process, moving-average process or autoregressive moving-average process. However, if the autoregressive moving-average model is mentioned, it includes only an autoregressive moving-average process.

The general autoregressive moving-average model is capable of representing any type of time series problem and is empirical-mechanistic in nature. The parameters in the model are as parsimonious as possible and are simple to interpret. In practice, it is frequently true that an adequate representation of an actual time series can be obtained with a low order model. The order of the model is usually not greater than two and often less than two. Hence, it possesses the characteristics of being a good mathematical model.

Recently, the general autoregressive moving-average model has been developed to represent many practical time series occuring in nature.

Examples are: scientific phenomena, such as the movement of tide, the vibrations of violin strings, the motion of the pendulum, etc [5]; in a business situation, such as the common stock market, gasoline sales by all oil company, international airline passenger fluctuation [6]; in an industrial production process, such as temperature variation, gas furnance process; in a simulation process, such as an inventory control process [7]. Not only may this model be used to represent the on going process, but it may also be used to forecast future situations.

An iterative cycle of identification, fitting, diagnostic checking and its forecasting are developed in this thesis to arrive at the appropriate function-stochastic model for the time series. This technique is to be applied to three sets of data obtained from chemical process, international airline passenger situations and simulated inventory process respectively. Both the nature of the system of the process and the optimal forecasts of future values can be acquired from this methodology.

1.1. Linear Filter Model

The mathematical models we employ are based on the idea that a time series in which successive values are highly dependent can be generated from a series of independent "shocks' a_t. [8]. These shocks are random drawings from a fixed distribution, usually assumed normal with mean zero and variance σ_a^2, such a sequence of random variable a_t, a_{t-1}, a_{t-2}, is called a white noise process.

The white noise process a_t is supposed to transform the process z_t by what is called a "linear filter", as shown in Fig. 1.1. [8]. The

THIS BOOK CONTAINS NUMEROUS PAGES WITH DIAGRAMS THAT ARE CROOKED COMPARED TO THE REST OF THE INFORMATION ON THE PAGE.

THIS IS AS RECEIVED FROM CUSTOMER.

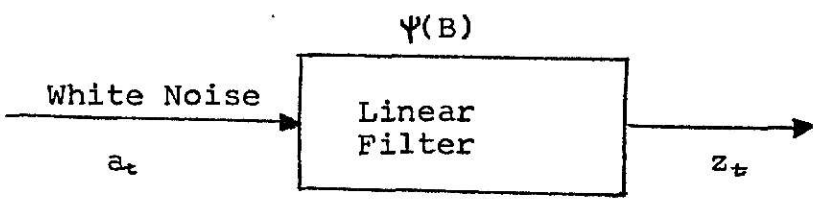


Figure 1.1 Representation of a time series as the output from a linear filter

linear filterning operation simply takes a weighed sum of previous observations, so that

$$z = \mu + \psi(B)\ a_t$$

$$= \mu + a_t + \psi_1 a_{t-1} + \psi_2\ a_{t-2} + \dots \qquad (1.1.1)$$

where μ is a parameter that determines the "level" of the process; B is the backward shift operator, i.e. $B\ z_t = z_{t-1}$; and

$$\psi(B) = 1 + \psi_1\ B + \psi_2\ B + \dots \qquad (1.1.2)$$

is the linear function that transforms a_t to z_t and is called the transfer function of the filter. The sequence ψ_1, ψ_2, ... formed by the weights may be finite or infinite. If this sequence is finite, or infinite and convergent, the filter is said to be stable and the process z_1, ... z_t to be stationary. [9]. The parameter μ in (1.1.1) is then the mean about which the process varies. Otherwise, z_t is non-stationary and μ has no specific meaning except as a reference point for the level of the process.

1.2. Stationary Process

Stationary processes play a very important role in the time series problem. Most of the time series phenomena occurring in nature, which are non-stationary or seasonal process, have to be transformed to a stationary process so that the appropriate model can be identified and the forecast values obtained.

The stationary process is based on the assumption that the process remains in equilibrum about a constant mean level. The time series is said to be strictly stationary if it is independent of time differences; that is, if the joint distribution associated with m observations $z_{t1}, z_{t2}, \ldots, z_{tm}$, made at any set of times $t_1, t_2, \ldots, t_m$, is the same as that associate with m observations $z_{t1+k}, z_{t2+k}, \ldots, z_{tm+k}$, made at times $t_{1+k}, t_{2+k}, \ldots, t_{m+k}$; thus for a discrete process to be strictly stationary, the joint distribution of any set of observations must be unaffected by shifting all of the observation times forward or backward by any integer amount k.

In Fig. 1.2, the observations of series C and D appear to fluctuate about a fixed mean with similiar pattern of irregularity. Series of this type are said to be "stationary in mean and variance". Series E, appears to fluctuate about a fixed mean but with a changing pattern of irregularity. Series of this type are said to be "stationary in the mean but nonstationary in variance"; Series A and B appear to drift with time, but appear to exhibit constant patterns of irregularity if allowance is made for the changing level and direction about which the observations are fluctuating. Series of this type are said to be "non-stationary in the mean". A more complete discussion of non-stationary process is presented in Section 1.6.

The stationary process implies that the probability distribution $P(z_t)$ is the same for all times t and may be written P(z). Hence its process has a constant mean where

$$\mu = E[z_t] = \int_{-\infty}^{\infty} z\, P(z)\, dz. \qquad (1.2.1)$$

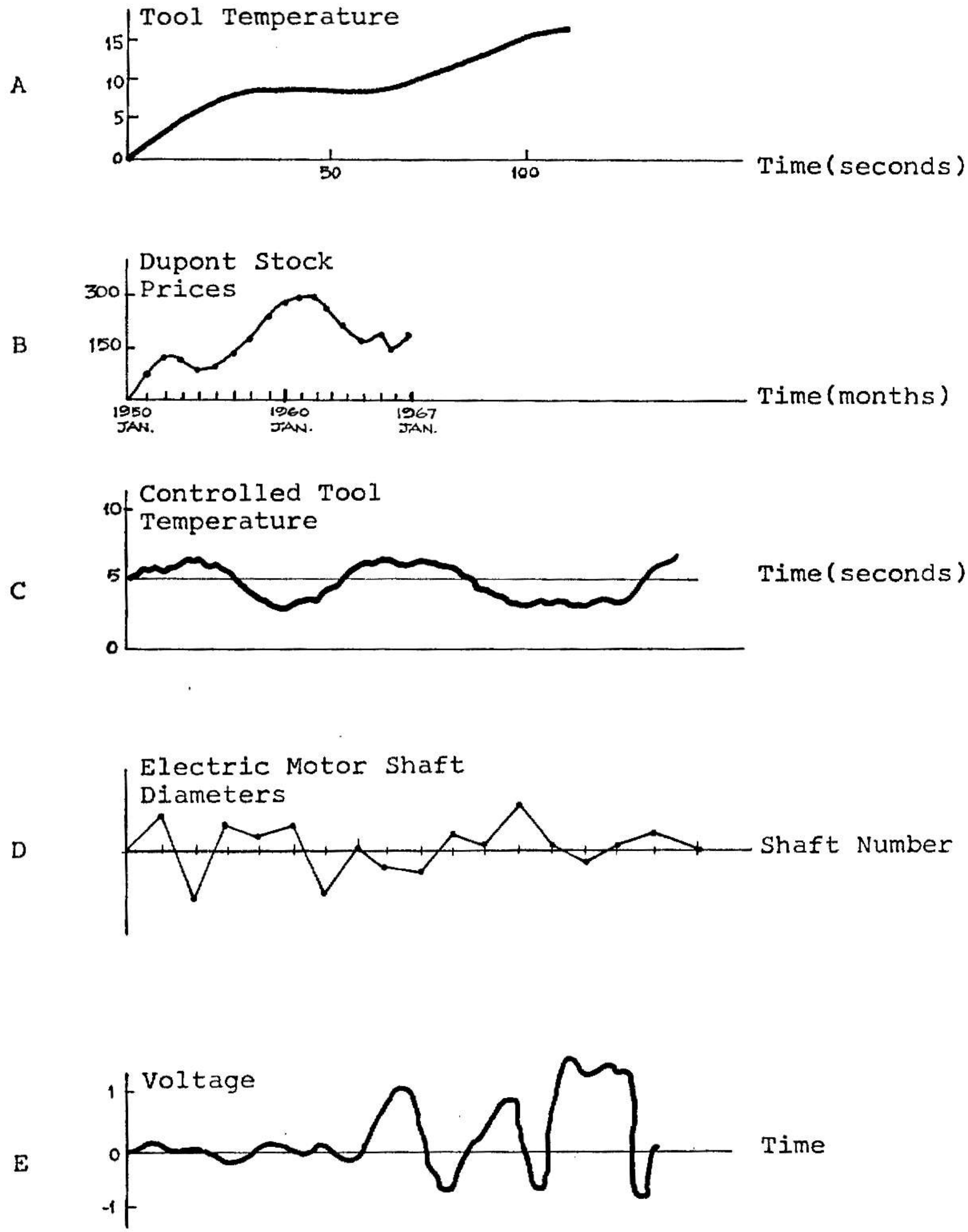


Figure 1.2 Typical Time Series

Which defines the level about which the process fluctuate; and with a constant variance

$$\sigma_z^2 = E\,[(z - \mu)^2] = \int_{-\infty}^{\infty} (z - \mu)^2 \quad P(z)\; dz \tag{1.2.2}$$

which measures its spread about the level μ.

The mean μ of the time series process can be estimated by

$$\hat{z} = \frac{1}{N} \sum_{t=1}^{N} z_t$$

and the variance σ_z^2, can be approximated by

$$\hat{\sigma}_z^2 = \frac{1}{N} \sum_{t=1}^{N} (z_t - \bar{z})^2 \tag{1.2.4}$$

1.3. Autoregressive Model

The autoregressive model of order p, or abbreviated as AR(p), can represent the given time series z_t, z_{t-1}, z_{t-2}, ..., observed at a constant time interval, as

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t \tag{1.3.1}$$

where $\tilde{z}_i = z_i - \mu$, μ is the mean of time series observation. a_t is assumed random normal and independent.

(1.3.1) can be related as a "dependent" variable $\tilde{z}_t$ regressed on a set of "independent" variable $\tilde{z}_{t-1}$, $\tilde{z}_{t-2}$, ..., $\tilde{z}_{t-p}$, plus an error term a_t. The autoregressive model can also be written as

$$\tilde{z}_t - \phi_1 \tilde{z}_{t-1} - \phi_2 \tilde{z}_{t-2} - \ldots - \phi_p \tilde{z}_{t-p} = a_t$$

or

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) \tilde{z}_t = a_t$$

where $\quad B^m \tilde{z}_t = \tilde{z}_{t-m}$

or

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

where $\quad \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$

The model contains p+2 unknown parameters μ, ϕ_1, $\phi_2, \ldots$, ϕ_p, σ_a^2, which are estimated from the time series data. The additional parameter σ_a^2 is the variance of the white noise process a_t.

The finite parameter in the autoregressive model $\phi(B) \tilde{z}_t = a_t$ can be inverted to an infinite number of random shock a_t, as

$$\tilde{z}_t = \phi^{-1}(B) a_t = \psi(B) a_t$$

where

$$\phi^{-1} (B) = \psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots \qquad (1.3.2)$$

Comparing (1.3.2) with (1.1.1) in section 1.1, for the autoregressive process to be stationary, the ϕ's must be so chosen that the weights ψ_1, ψ_2, ..., in $\psi(B)$ form a convergent series. The autoregressive process can be thought of as the output $\tilde{z}_t$ from a linear filter with transfer function $\phi^{-1}(B)$, where the input is white noise a_t.

1.4. Moving Average Model

The autoregressive model which expresses the deviation $\tilde{z}_t$ of the process as a finite weighted sum of p previous deviation $\tilde{z}_{t-1}$, $\tilde{z}_{t-2}$, ..., $\tilde{z}_{t-p}$, plus a random shock a_t, can be inverted as an infinite weighted sum of a's. The moving average model may be defined as a linear function of a number of previous shocks a's, which can be finite or infinite. [10].

For the moving average model of order q, or abbreviated as MA(q), the model form is

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_g a_{t-g} \tag{1.4.1}$$

(1.4.1) is also called a moving average process of order q. It may be written as,

$$\tilde{z}_t = \theta(B)\, a_t \tag{1.4.2}$$

where

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_g B^g \tag{1.4.3}$$

The model contains q+2 unknown parameters μ, θ, ... θ_q, σ_a^2, which in practice are determined from the data. The moving average process can be thought of as the output $\tilde{z}_t$ from a linear filter with transfer function $\theta(B)$, when the input is white noise a_t.

1.5. Autoregressive Moving-Average Model

The first order moving-average model,

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} = (1 - \theta_1 B) a_t \tag{1.5.1}$$

can be inverted to an infinite autoregressive process, as

$$a_t = (1 - \theta_1 B)^{-1} \tilde{z}_t$$

or

$$a_t = \tilde{z}_t + \theta_1 \tilde{z}_{t-1} + \theta_1^2 \tilde{z}_{t-2} + \ldots$$

or

$$\tilde{z}_t = -\theta_1 \tilde{z}_{t-1} - \theta_1^2 \tilde{z}_{t-2} - \ldots + a_t \tag{1.5.2}$$

The higher order moving average model can be inverted to an infinite autoregressive process by the same derivation. Hence, if the process can be represent by MA(1), it is impractical to obtain a non-parsimonious representation in terms of an autoregressive model; Conversely, an autoregressive model of first order can not be parsimoniously represented using a moving average process. To achieve greater flexibility in fitting actual time series, it is sometimes advantageous to include both autoregressive and moving-average terms in the model. This leads to the so-called autoregressive moving-average model, as,

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \ldots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \ldots - \theta_g a_{t-q} \tag{1.5.3}$$

or

$$\phi(B) \tilde{z}_t = \theta(B) a_t \tag{1.5.4}$$

which is also called the autoregressive moving-average process of order (p,q), or abbreviated as ARMA(p,q). The ARMA(p,q) employs p+q+2 unknown parameters μ; $\phi_1, \dots \phi_p$; $\theta_1, \dots \theta_q$; σ_a^2, that can be obtained from the data.

(1.5.4) may be inverted to

$$\tilde{z}_t = \phi^{-1}(B)\ \theta(B)\ a_t = \frac{\theta(B)}{\phi(B)}\ a_t = \frac{1 - \theta_1 B - \dots - \theta_q B^q}{1 - \phi_1 B - \dots - \phi_p B^p}\ a_t$$

The autoregressive moving-average process can then be thought of as the output $\tilde{z}_t$ from a linear filter, whose transfer function is the ratio of two polynomials $\theta(B)$ and $\phi(B)$, when the input is white noise a_t.

1.6. Non-stationary Process

Non-stationary time series have the property that their mean or variance or both may be changed with time. In other words, a non-stationary time series is depend on the time difference. When there is doubt about the choice of a nonstationary model or a stationary model to represent a time series, it is advantageous to employ the nonstationary model rather than the stationary alternative. [11]. Because the trend of a non-stationary time series can be transformed to a stationary process by differencing the data. Thus the nonstationary time series operator $\varphi(B)$ can be defined as

$$\varphi(B) = \phi(B)\ (1-B)^d \tag{1.6.1}$$

where $\phi(B)$ is a stationary operator. Thus a general model, which can represent nonstationary behavior, is of the form

$$\varphi(B)\ z_t = \phi(B)\ (1-B^d)\ z_t = \theta(B)\ a_t$$

or

$$\phi(B)\ w_t = \theta(B)\ a_t \tag{1.6.2}$$

where

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

Nonstationary behavior can therefore be represented by a model which calls for the d'th difference of the process to be stationary. In practice, d is usually one or at most two.

The process defined by (1.6.2) provides a good way for describing non-stationary time series and is called an autoregressive moving-average process of order (p,d,q), or abbreviated as ARMA(p,d,q).

1.7. Seasonal Model

A seasonal time series is defined as a series which exhibits periodic behavior with period S; i.e. when similarities in the series occur after S basic time intervals.

The monthly international airline passenger in Fig. 1.3, for example, is highly correlated twelve months apart. Sales of a particular product, like baseball equipment, will also be expected to have the same highly correlated situation. The highest sales occuring during the summer months and the months of December. Series of this type are called seasonal time series.

The fundamental part about seasonal time series with period S is that the observations which are S intervals apart are similiar. Therefore,

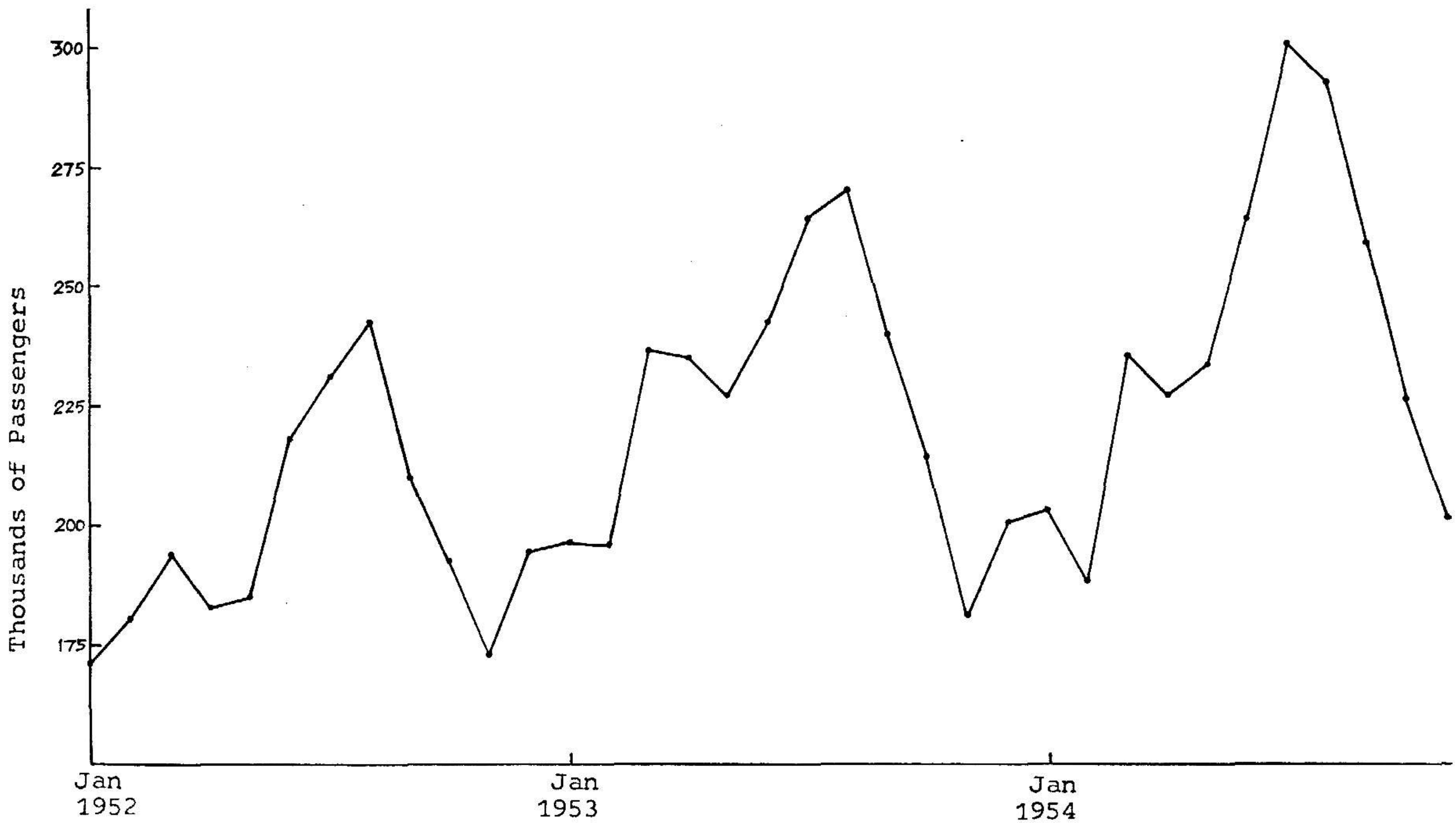


Figure 1.3 Totals of international airline passengers in thousands

$B^s z_t = z_{t-s}$ will be a powerful tool to analyze seasonal time series. The linking of the observation z_t to an observation in the previous period with period S by a general ARMA model is defined as [12],

$$\Phi(B^s)\ \nabla_s^D\ z_t = \Theta\ (B^s)\ a_t \tag{1.7.1}$$

where

$\nabla_s = 1-B^s$ and D is the number of seasonal difference.

$\Phi(B^s)$, $\Theta\ (B^s)$ are polynomials in B^s of degrees P and Q, respectively.

(1.7.1) is the autoregressive moving-average process which represents the seasonal time series, or abbreviated symbolically as ARMA$(P,D,Q)_s$.

1.8. Multiplicative Model

Suppose that a time series has shown a tendency to increase over a particular period and also to follow a seasonal pattern. Then the time series may be represented by the form

$$\phi(B)\ \Phi(B^s)\ \nabla^d\ \nabla_s^D\ z_t = \theta_q(B)\ \Theta_Q(B^s)\ a_t \tag{1.8.1}$$

which is the multiplication of

$$\Phi(B^s)\ \nabla_s^D\ z_t = \Theta\ (B^s)\ \alpha_t \tag{1.8.2}$$

and

$$\phi(B)\ \nabla^d\ \alpha_t = \theta(B)\ a_t \tag{1.8.3}$$

(1.8.2) and (1.8.3) are used to take care of seasonal fluctuations and non-stationary trend respectively. [13]. α_t and a_t are defined as a white noise process; $\phi(B)$ and $\theta(B)$ are polynominals in B of degrees p and q, respectively, and $\nabla = \nabla_1 = 1-B$.

This most general autoregressive moving-average process is said to be of order $(p,d,q) \times (P,D,Q)_s$. It represents the time series process having a non-stationary trend and cyclic pattern and can also be denoted symbolically as ARMA$(p,d,q) \times (P,D,Q)_s$.

1.9. The Selection of An Appropriate Model

The purpose of this thesis is to find an appropriate model to represent a time series process and also forecast its future value. The method to select an appropriate model can be explained briefly in Fig. 1.4. The function at different stages can be illustrated as follows.

(1) The theory and practice are to be interacted to entertain the appropriate model. The autocorrelation and the partial autocorrelation function and the knowledge of the system are employed to suggest an appropriate parsimonious model. In addition, a rough estimate of the model parameters can be achieved in the process of model identification.

(2) The efficient estimate of parameters in the tentatively entertained model is the heart of this stage. The rough estimates of the parameters obtained during the identification stage can now be used as the starting points for the least square estimation of the parameters.

(3) The entertained model is subjected to a diagnostic check to test the goodness of fit. If no inadequacy of fit is indicated, the model is ready to use.

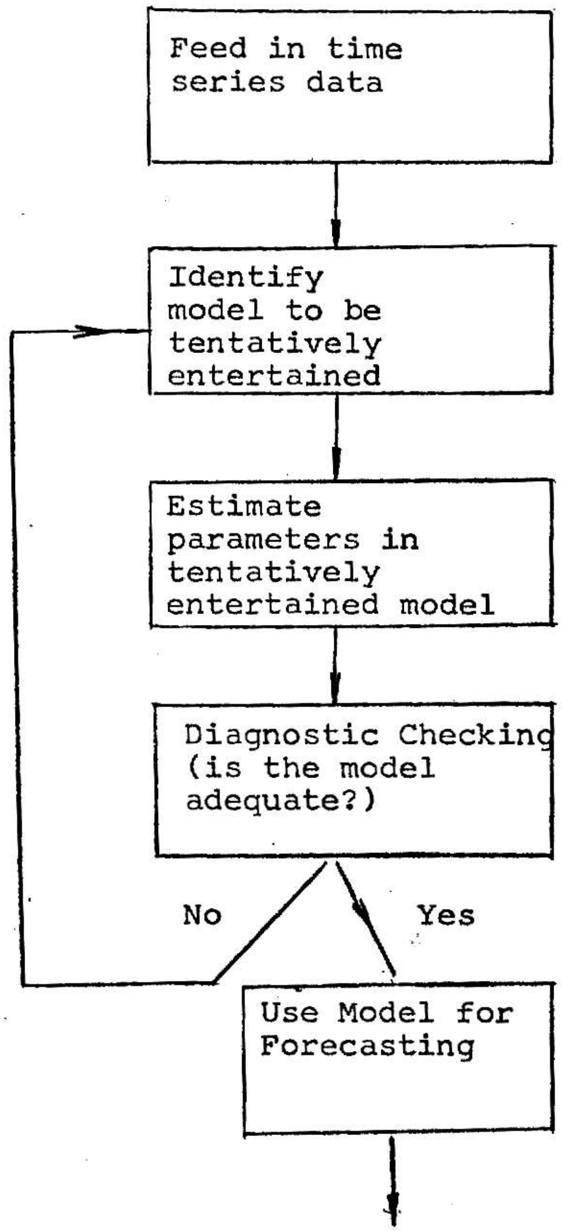


Figure 1.4 Stages in the iterative approach to model building

(4) If the model is adequate enough to represent the given time series process, the future situation can be forecasted and its confidence interval computed.

More details of model building and its application will be described in the following chapters.

CHAPTER II

MODEL IDENTIFICATION

Identification methods are the rough procedures applied to a set of time series data to indicate the kind of representational model which is worthy of further investigation. The specific aim in the identification stage is to obtain some idea about the number of parameters and the degrees of differences needed in the appropriate model and also to obtain initial estimates for the parameters. The tentative models so obtained provide a starting point for the application of the more formal and efficient estimation methods in the stimation stage.

Our approach to identify an appropriate model from the general autoregressive moving-average model family, which is

$$\phi(B)\ \nabla^d\ z_t = \theta(B)\ a_t$$

are as follows.

(a) To identify the possibility of nonstationary and cycle trend, the original series z_t is to be differenced as many time as needed. For the nonstationary time series,

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

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

For the seasonal time series,

$$\Phi(B^s)\ w_t = \Theta\ (B^s)\ a_t$$

where $w_t = \nabla_s^D z_t = (1-B^s)^p z_t$.

(b) To identify the appropriate model form of the time series data. The distribution of the time series can be well defined by its theoretical autocorrelation function, its mean and variance. Every kind of model has its specific autocorrelations and partial autocorrelation function. In view of these facts, a powerful technique for identifying a candidate model form can be achieved by estimating the correlation and partial correlation pattern from the data and mentally comparing them with the theoretical patterns. Then select the model which has the estimated correlation and partials most similiar with the theoretical correlation and partials. Many charts of lower order autoregressive moving-average models for this purpose are constructed by Box and Jenkins [14]. The unique pattern of the autocorrelation and partial autocorrelation of the general ARMA process can be used not only to identify the model, but to obtain the appropriate estimate of the parameters.

2.1. Autocorrelation Function

Each different type of autoregressive moving-average model has its own specific autocorrelation coefficient pattern. The autocorrelation coefficient process can be plotted out as a scatter diagram using pairs of values (z_t, z_{t+k}), of the time series, seperated by k lags apart. It is easy to select the appropriate model for the given time series by the plotted form of its autocorrelation function. Theoretically, the autocorrelation coefficient at lag k is

$$\rho_k = \frac{E[(z_t-\mu)\ (z_{t+k}-\mu)]}{\sqrt{E\ [(z_t-\mu)^2]\ E\ [(z_{t+k}-\mu)^2]}} \tag{2.1.1}$$

And the covariance between z_t and z_{t+k}, which is also called the autocovariance at lag k, is

$$\gamma_k = \text{Cov}\ [z_t,\ z_{t+k}] = E\ [(z_t-\mu)\ (z_{t+k}-\mu)] \tag{2.1.2}$$

To estimate the autocorrelation coefficient and autocovariance, Box and Jenkins recommends the following method [15], for autocovariance is estimated as

$$\hat{\gamma}_k = \frac{1}{N}\sum_{t=1}^{N-k} (z_t-\bar{z})\ (z_{t+k}-\bar{z}), \qquad k = 0,\ 1,\ 2,\ \ldots,\ K \tag{2.1.3}$$

The estimated autocorrelation coefficient is

$$\rho_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$$

with its variance

$$\text{Var}\ [\hat{\rho}_k] \simeq \frac{1}{N}\left\{1 + 2\sum_{j=1}^{k-1} \hat{\rho}_j^2\right\} \tag{2.1.4}$$

The square root of (2.1.4) is called the large-lag standard error [16]. It is based on the assumption that the theoretical autocorrelation

ρ_k are all essentially zero beyond some hypothesized lag $k = q$. The large lag standard error approximated the standard deviation of ρ_k for suitably large lags ($k>q$). Hence, usually, $\pm \hat{\sigma}_{\hat{\rho}_k}$ or $\pm 2\hat{\sigma}_{\hat{\rho}}$ is plotted as "control" lines about zero. This is an rough indication of whether the autocorrelation coefficient is zero beyond some specific lag, or, in other words, the autocorrelation function is being cut off after a particular lag.

The theoretical autocorrelation matrix is usually symmetrical [17], $\rho_{-k} = \rho_k$, it is only necessary to plot the positive half of the autocorrelation matrix to analysis its process. When the autocorrelation function is mentioned later, it means only the positive half of its function.

2.2. Partial Autocorrelation Function

Every autoregressive, moving-average or autoregressive moving-average model has its own specific partial autocorrelation function. Hence the partial autocorrelation function is used as an auxiliary device to identify the appropriate model for a given time series among the general ARMA family. The correlation, ρ_k, represents the dependence between z_t and z_{t-k}. However, the partial correlation, ρ'_k, represents the dependence between z_t and z_{t-k}, given that observations $z_{t-k+1}, \ldots, z_{t-1}$ are known. Hence, for the AR(p) model,

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \ldots + \phi_p \tilde{z}_{t-p} + a_t \qquad (2.2.1)$$

It can be observed that given the observations $z_{t-1}, z_{t-2}, \ldots,$

z_{t-p}, no dependence exist between z_t and the observation occuring before time t-p. The partial correlation, ρ'_k, will therefore be zero when k>p. In other words, the partial autocorrelation process of AR(p) model will be cut off after lag p. For both the moving-average process MA(q)

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \tag{2.2.2}$$

and the autoregressive moving-average process ARMA(p,q)

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \tag{2.2.3}$$

can be inverted to the infinite autoregressive process. This implies that the partial autocorrelation function of the moving-average process or ARMA process tails off rather cuts off.

The estimation of the partial autocorrelation is developed by Dubin as [18],

$$\rho'_{k+1} = \frac{\rho_{k+1} - \sum_{j=1}^{k} \rho'_{k,j}\, \rho_{k+1-j}}{1 - \sum_{j=1}^{k} \rho'_{k,j}\, \rho_j} \tag{2.2.4}$$

$$\rho'_{k+1,j} = \rho'_{k,j} - \rho'_{k+1}\ \rho'_{k,k-j+1} \qquad (j = 1, 2, \dots, K)$$

$$\rho'_0 = 1$$

$$\rho'_1 = \rho_1$$

with variance of the partial correlation coefficient as

$$\mathrm{Var}(\hat{\rho}'_k) \simeq \frac{1}{n-k} \tag{2.2.5}$$

The standard error, which is the square root of (2.2.5), can be used as a rough indication of the lag q where the partial autocorrelation is cut off.

2.3.1. The Autocorrelation Function of the Autoregressive Process

The specific autocorrelation pattern of the autoregressive process is a powerful tool to distinguish it from MA or ARMA process. On the autoregressive process

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t \tag{2.3.1}$$

Multiplying each term in (2.3.1) by $\tilde{z}_{t-k}$

$$\tilde{z}_{t-k} \tilde{z}_t = \phi_1 \tilde{z}_{t-k} \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-k} \tilde{z}_{t-2} + \dots \phi_p \tilde{z}_{t-k} \tilde{z}_{t-p} + \tilde{z}_{t-k} a_t \tag{2.3.2}$$

Then take the expected value of (2.3.2), we obtain

$$E[\tilde{z}_{t-k} \tilde{z}_t] = \phi_1 E[\tilde{z}_{t-k} \tilde{z}_{t-1}] + \phi_2 E[\tilde{z}_{t-k} \tilde{z}_{t-2}] + \dots$$

$$+ \phi_p E[\tilde{z}_{t-k} \tilde{z}_{t-p}] + E[\tilde{z}_{t-k} a_t]$$

or

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \qquad k > 0 \qquad (2.3.4)$$

z_{t-k} can only involve a_i up to time t-k; for a_t is beyond t-k, it is uncorrelated with z_{t-k}; so the expected value of $E[\tilde{z}_{t-k} \ a_t]$ vanishes. Dividing (2.3.4) by γ_0, the autocorrelation function of the AR process is

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \qquad k > 0 \qquad (2.3.5)$$

or

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \rho_k = 0$$

$$\phi(B) \rho_k = 0 \qquad (2.3.6)$$

No matter how large a k we take in (2.3.5), ρ_k is still obtainable. This fact implies that the autocorrelation function of an AR process tails off rather than cuts off.

Box and Jenkins discuss the roots of $\phi(B)$ in (2.3.6) and conclude that the autocorrelation function of an autoregressive process is either a damped exponential or damped sine wave or a mixture of damped exponential and damped since wave [19].

2.3.2. The Partial Autocorrelation Function of AR Process and Yule-Walker Equations

To decide which order of autoregressive process to fit an observed time series is analogaous to decide the number of independent variables to be included in a regression equation. For an AR process is finite itself and MA process can be inverted to an infinite AR process, any general ARMA model can be expressed in AR form, either finite or infinite. Although the proper order of an AR model to fit the time series is unknown, its parameter can be easily calculated.

For the autoregressive process

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \qquad k > 0 \tag{2.3.1}$$

By substituting k = 1, 2, ..., p, in (2.3.1) one by one, and for $\rho_{-k} = \rho_k$, it yields,

$$\begin{aligned}
\rho_1 &= \phi_1 &&+ \phi_2 \rho_1 &&+ \dots + \phi_p \rho_{p-1} \\
\rho_2 &= \phi_1 \rho_1 &&+ \phi_2 &&+ \dots + \phi_p \rho_{p-2} \\
&\vdots && && \vdots \\
\rho_p &= \phi_1 \rho_{p-1} &&+ \phi_2 \rho_{p-2} &&+ \dots + \phi_p
\end{aligned} \tag{2.3.2}$$

(2.3.2) are the Yule-Walker equation [20]. The matrix form of Yule-Walker equations can be written as

$$\rho_p = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{pmatrix} \qquad \phi = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{pmatrix} \qquad P_p = \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & & & \rho_{p-2} \\ \vdots & & & & \\ \rho_{p-1} & \rho_{p-2} & & & 1 \end{pmatrix}$$

or

$$\rho_p = P_p \, \phi$$

hence

$$\phi = P_p^{-1} \rho_p$$

Initially, which order of an AR process to fit is unknown; suppose ϕ_{kj} is the jth coefficient in an autoregressive process of order k, so that ϕ_{kk} is the last coefficient. (2.3.1) can be written as

$$\rho_j = \phi_{k1} \rho_{j-1} + \dots + \phi_{k(k-1)} \rho_{j-k+1} + \phi_{kk} \rho_{j-k} \qquad j = 1, 2, \dots, k \tag{2.3.5}$$

Hence (2.3.2) can be extended to

$$\begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-2} \\ \vdots & & & & \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & & 1 \end{pmatrix} \begin{pmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{pmatrix} \tag{2.3.6}$$

or the matrix form

$$P_k \phi_k = \rho_k$$

$$\phi_k = P_k^{-1} \rho_k \tag{2.3.7}$$

The quantity ϕ_{kk} is regarded as a partial autocorrelation function [21]. To the autoregressive parameter ϕ_{kk}, the values between ρ_1, ... ρ_k have to be known. In other words, ϕ_{kk} is dependent on the observation z_1, ..., to z_k.

(2.3.7) can be used also as the rough estimate of the autoregressive model parameter.

2.4.1. The Autocorrelation Function of a Moving-Average Process

The moving-average process has its own specific autocorrelation function. For MA process,

$$z_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \tag{2.4.1}$$

multiplying (2.4.1) by z_{t-k}, which is

$$z_{t-k} = a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q}, \quad k = 1, 2, \dots, q \tag{2.4.2}$$

then obtain

$$z_t z_{t-k} = (a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q})(a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q}) \tag{2.4.3}$$

takes expectation value on (2.4.3),

$$E[z_t z_{t-k}] = E[(a_t - \theta_1 a_{t-1} a_{t-1} - \dots - \theta_q a_{t-q})(a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q})] \tag{2.4.4}$$

Here, the random variables a_t are assumed uncorrelated [22],

$$\gamma_k = E[a_t , a_{t-k}] = \begin{cases} \sigma_a^2 & , \quad k = 0 \\ 0 & , \quad k \neq 0 \end{cases} \tag{2.4.5}$$

Then the solution of (2.4.4) is shown as [23],

$$\gamma_k = \begin{cases} (-\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \ldots + \theta_{q-k} \theta_q) \sigma_a^2, & k = 1,2,\ldots,q \\ 0 & , \ k > q \end{cases} \qquad (2.4.6)$$

with

$$\gamma_0 = 1 + \theta_1^2 + \theta_2^2 + \ldots + \theta_q^2$$

Hence the autocorrelation coefficient, $\rho_k = \frac{\gamma_k}{\gamma_0}$, will be

$$\rho_k \begin{cases} \dfrac{-\theta_k + \theta_1 \theta_{k+1} + \ldots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \ldots + \theta_q^2}, & K = 1,2,\ldots, q \\ 0 & , \ K > q \end{cases} \qquad (2.4.7)$$

(2.4.7) reveals that the autocorrelation function of MA process is zero or cut off beyond a lag of q. This fact provides the means to identify the model of the observed time series. (2.4.7) can also be employed to obtain an approximate parameter estimate for the moving-average model.

2.4.2. The Partial Autocorrelation of MA Process

The finite moving-average model $z_t = \phi(B)\ a_t$ can be inverted to an infinite autoregressive model $\phi^{-1}(B)\ z_t = a_t$ with an infinite number of parameters. In other words, the finite moving-average model can be expressed in terms of an autoregressive model with its order to be decided. Here Box and Jenkins recommend the partial autocorrelation coefficient of the moving-average process to be expressed by its inverted autoregressive parameter as [24],

$$\phi_{kk} = -\theta_1^k \left\{1 - \theta_1^2\right\} / \left\{1 - \theta_1^{2(k+1)}\right\} \qquad (2.3.2.1)$$

with

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

and

$$\rho_0 = 0$$

to decide θ_1 value.

From (2.4.2.1), if ρ_1 is positive, then θ_1 is negative, so that ϕ_{kk} is positive. Conversely, if ρ_1 is negative, θ_1 is positive, so that ϕ_{kk} is negative. This implies that the partial autocorrelation function of MA process is damped exponential. For k values in (2.4.2.1) can be substituted by any positive integer and ϕ_{kk} is still obtainable, the process has a cut off after a lag of q, whereas the autocorrelation function of the AR process tails off, whereas the partial autocorrelation function of an AR process cuts off.

2.5. Autocorrelation and Its Partial for the Autoregressive Moving-Average Process

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

or

$$\phi(B)\ \tilde{z}_t = \theta(B)\ a_t \tag{2.5.1}$$

has its autocovariance function

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

$$\theta_q\ \gamma_{za}(k-q) \tag{2.5.2}$$

(2.5.2) is obtained by multiplying (2.5.1) with z_{t-k} and then taking the expected value. Suppose $\gamma_{za}(k) = E[\tilde{z}_{t-k}\ a_t]$; since $\tilde{z}_{t-k}$ depend only on random shock a_j which have occurred up to time t-k, it follows that [25],

$$\gamma_{za}(k) = 0 \qquad k > 0$$

$$\gamma_{za}(k) \neq 0 \qquad k \leq 0 \tag{2.5.3}$$

Hence the autocorrelation function (2.5.2) will become

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p} \qquad k \geq q + 1 \tag{2.5.4}$$

by dividing (2.5.4) with γ_0, the autocorrelation function can be expressed as,

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p} \qquad k \geq q + 1 \tag{2.5.5}$$

or

$$\phi(B) \rho_k = 0 \tag{2.5.6}$$

For any k value bigger than q+1 in (2.5.5), ρ_k is always obtainable. Thus the autocorrelation function of an ARMA process tails off rather than cuts off. Box and Jenkins also investigate the roots of $\phi(B)$ in (2.5.6), and find the autocorrelation function of an ARMA process consists of a mixture of damped exponentials and/or damped sine wave. (2.5.5) can also be used to estimate the rough parameter ϕ for the appropriate ARMA model. (2.5.1) can be inverted as

$$\tilde{z}_t = \phi^{-1}(B) \theta(B) a_t$$

For $\phi^{-1}(B)$ is an infinite series in B. Hence the partial autocorrelation function of an ARMA process will be infinite in extent. Box and Jenkins conclude that the partial autocorrelation function will behave like a mixture of damped exponentials and/or damped sine waves, depending on the order of the ARMA process and the values of the parameters. [26].

We now can conclude that the appropriate model may be an ARMA process if the autocorrelation function and its partials tail off rather than cut off.

The first order ARMA model is the most common and practical to represent the appropriate ARMA process. An approximate value of the parameters of ARMA(1,d,1) process can be computed by [27],

$$\rho_1 = \frac{(1 - \theta_1 \phi_1)(\phi_2 - \phi_1)}{1 + \theta_1^2 - 2 \phi_1 \theta_1} \qquad - 1 < \phi <$$

$$\rho_2 = \rho_1 \phi_1 \qquad - 1 < \theta_1 < 1$$

A general method for obtaining initial estimate of the parameter for any ARMA process is derived by Box and Jenkins [28].

2.6. Identification of Appropriate Model

The autocorrelation function and the partial autocorrelation function have been used as a powerful tool to identify the appropriate model for a given time series. The low order model can represent the given time series quite well. The order is usually no more than 2. Many charts and tables have been constructed to describe the autocorrelation functions, partial autocorrelation functions and to provides an estimate of parameters for low order ARMA models, see Box and Jenkins [29]. The general process to identify the appropriate model for a given time series can be summarized as follows.

(a) For the non-stationary time series, its autocorrelation function will not die out quickly, or will fall off slowly, or is very nearly linearly [30]. Therefore, a tendency for the autocorrelation function not to die out quickly is taken as an indication that a nonstationary time series may exist. Then we can treat the time series as nonstationary in z_t, but possibly as stationary in ∇z_t, or in some higher difference.

It is assumed that the degree of difference d, which is required for stationarity, has been reached when the autocorrelation function of $w_t = \nabla^d z_t$ dies out fairly quickly. In practice, d is normally either 0, 1 or 2. It is usually sufficient to inspect about the first 20 estimated autocorrelation of the time series.

(b) The AR process has a autocorrelation function which is infinite in extent, but has a partial autocorrelation function that is zero beyond a certain point. Conversely, the MA process has an autocorrelation function of zero beyond a certain point, but with a partial autocorrelation function which is infinite is extend.

Table 2.1. is the summary of the properties of AR, MA and ARMA process.

Table 2.1. Summary of Properties of Autoregressive, Moving-average and ARMA Process

	autoregressive process	moving-average process	ARMA process
model in terms of previous $\tilde{z}$'s	$\phi(B)\ \tilde{z}_t = a_t$	$\theta^{-1}(B)\ \tilde{z}_t = a_t$	$\theta^{-1}(B)\ \phi(B)\ \tilde{z}_t = a_t$
model in terms of previous a's	$\tilde{z}_t = \phi^{-1}(B)\ a_t$	$\tilde{z}_t = \theta(B)\ a_t$	$\tilde{z}_t = \phi^{-1}(B)\ \theta(B)\ a_t$
autocorrelation function	infinite (damped exponentials and/or damped sine waves), tail off	finite, cut off	infinite (damped exponentials and/or damped sine waves after the first q-p lags), tail off
partial auto-correlation function	finite, cut off	infinite (dominated by damped exponentials and/or sine waves), tail off	infinite (dominated by damped expo-nentials and/or sine waves after the first p-q lags), tail off

CHAPTER III

MODEL ESTIMATION AND DIAGNOSTIC CHECK

The tentative formation of the model for the given time series is obtained in the identification stage. The more efficient estimate of the parameters in the appropriate model will be computed so as to construct a more perfect model. The rough estimated parameters obtained in the identification stage will be used as the starting points. After the model is built, it will be subject to diagnostic checks to test the fit of the model. If the model is inadequate, the time series process will be reviewed again and another modified model tried. If no lack of fit is indicated, the model is ready to use.

3.1. Maximum Likelihood Estimation of Parameters of ARMA Model

After a candidate model has been selected, it is necessary to estimate more accurate parameters to fit the time series data. The best estimate, from many points of view, is the maximum likelihood estimate [31].

For the ARMA process

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots -$$

$$\theta_q a_{t-q} + a_t \tag{3.1.1}$$

the random shocks $a_1, a_2, \dots, a_t, \dots a_n$ are assumed normally independently distributed, so,

$$f(a_1\, a_2\, \ldots\, a_n \mid \underline{\phi}, \underline{\theta}, \sigma^2) = \frac{1}{(2\pi)^{n/2}\, \sigma^n}\, e^{-\frac{1}{2\sigma^2} \sum a_t^2}$$

$$= \frac{1}{(2\pi)^{n/2}\, \sigma^n}\, e^{-\frac{1}{2\sigma^2} S(\underline{\phi},\underline{\theta})} \qquad (3.1.2)$$

where

$$S(\underline{\phi},\underline{\theta}) = \sum a_t^2 \qquad (3.1.3)$$

and

$$a_t = \tilde{z}_t - \phi_1 \tilde{z}_{t-1} - \ldots - \phi_p \tilde{z}_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \ldots$$

$$+ \theta_q a_{t-q} \qquad (3.1.4)$$

The likelihood function of $\underline{\theta}$, $\underline{\phi}$ and σ^2, can be obtained by substituting the observed value of a into (3.1.2), as

$$L(\underline{\phi}, \underline{\theta}, \sigma^2 \mid a_1, a_2, \ldots a_n) \propto \frac{1}{\sigma^n}\, e^{-\frac{1}{2\sigma^2} S(\underline{\phi},\underline{\theta})} \qquad (3.1.5)$$

The likelihood function is maximized when $S(\underline{\phi},\underline{\theta})$ is minimized. The maximum likelihood estimates of $\underline{\phi}$ and $\underline{\theta}$, denoted by $\hat{\underline{\phi}}$ and $\hat{\underline{\theta}}$, corresponds to the minimum sum of squares, $S(\hat{\underline{\phi}}, \hat{\underline{\theta}})$.

Differentiating (3.1.5) with respect to σ^2 reveals that the maximum likelihood estimate of σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{n} S(\hat{\underline{\phi}}, \hat{\underline{\theta}})$$

The maximum likelihood estimation of $\underline{\phi}$, $\underline{\theta}$ are equal to the least square estimate, which minimize the sum of squares of the residuals. The parameters in the model can be linear or non-linear after extention. For instance, $a_t = \phi^{-1}(B)\ \theta(B)\ z_t$ or $a_t = \phi(B)\ \Phi(B^s)\ \nabla^d \nabla_s^D\ z_t$. Hence the the non-linear parameter least square estimation procedure is employed to meet every possible model, either linear or non-linear, [32]. Subroutine UWHAUS in the Appendix is used to obtain the estimated parameter by a nonlinear least square method.

In order to insure that the estimation of parameters will converge to the least square point and also save machine time, it is necessary to obtain a good initial estimate of the parameters to start the computation. The initial estimate of parameters are obtained from autocorrelation and/or partial autocorrelation functions as discussed on Chapter Two.

The general ARMA model can be transformed to an equation like (3.1.4) in order to pursue the least square estimate of the parameters in the model [33]. For starting the problem, the p values z_0, z_{-1}, z_{-2}, ... z_{-p+1} among the n=N-d w_i, which is $w_i = \nabla^d z_i$, and the q values a_0, a_{-2}, ... a_{-q+1} of a are unknown. For the practical purpose, if the sample size is moderately large, the unknown a's can be assumed zero and also sacrifice the first p observations with an effective size of n-p [34].

3.2. Diagnostic Check

After the model has been identified and the parameters estimated for a time series, the model should be subject to investigation to test the fit of the model. If there is evidences of serious inadequacy of fit, the model will be adjusted and the modified model tried again.

No model form can ever represent the true time series absolutely. However, the model should nave no indicated lack of fit under different statistical tests. Box and Jenkins suggest many statistical tests for the general autoregressive moving-average model.

3.2.1. Diagnostic Checks Applied to Residuals- Autocorrelation Check

Theoretically, the random shock in the ARMA process is assumed to be white noise. Therefore, it is reasonable to expect that the study of the $\hat{a}_t$ in the ARMA model can indicate the model inadequacy. The autocorrelation function of the residual $\hat{a}_t$ is therefore a good device to test the fit of the model. Assume a general ARMA model,

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

where

$$w_t = z_t - \mu$$

being built from the interaction of the theory and practice discussed in Chapter Two and Three. Thus the residuals of the model

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

are subject to test. It is possible to show that, if the model is adequate, then [35]

$$\hat{a}_t = a_t + 0\ (\frac{1}{\sqrt{n}})$$

As the series length increases, the $\hat{a}_t$'s become close to the white noise a_t's. Hence the estimated autocorrelation coefficient $\gamma_k(a)$, of the a's, distributed approximately about zero with variance n^{-1}, or, with a standard error of $n^{-1/2}$ [36]. We can use these fact to assess the statistical significance of apparant departures of these autocorrelations from zero. If all the estimated autocorrelation coefficients of the residuals are inside the "control" line, then no inadequacy of the model is indicated. However, if the estimated autocorrelation coefficients are out of the "control" line, the suspicion of the lack of fit is hence aroused.

3.2.2. A Portmenteau Lack of Fit Test [37]

Box and Jenkins also suggest another statistical method to test the model fit. Rather than consider the $\gamma_k(\hat{a})$'s individually, the first few autocorrelation coefficients of the a's, suppose about 20, are taken as a whole to test the fit of the model. Suppose we take the first k autocorrelation coefficients $\gamma_k(\hat{a})(k=1,2,\ldots K)$ from general ARMA(p,d,q) process, then if the model is appropriate, the value of

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

will be distributed as $\chi^2(k-p-q)$, where n=N-d is the number of transformed observations w_i, where $w_i = \nabla^d z_i$, used to fit the model. On the other hand, if the model is inappropriate, the average values of Q will be inflated. Therefore, an appropriate, general, or "portmanteau" test of the fit of the model can be achieved by obtaining the value of Q and comparing it with the percentage points on the χ^2 table. If Q is greater than the percentage points on the χ^2 table, then the inadequacy of the model is indicated. Conversely, if Q is no greater than the critical χ^2 value, then no inadequacy of the model is indicated.

CHAPTER IV

FORECASTING

The model is supposed to represent the time series data adequately as no lack of fit is indicated under the statistical investigation. Then the appropriate model can represent the stochastic process as well as be used to forecast the future situations. The approximation of the forecast value of the time series process will be presented in this Chapter. The confident limits of the forecast value will be developed.

4.1. The Forecast Function of the ARMA Model

The forecast function of ARMA model, as indicated by Box and Jenkins, has three model forms, either in terms of the difference equation, or in terms of an infinite weighted sum of previous random shock a_j, or in terms of an infinite weighted sum of previous observations plus a random shock. The simplest and the most practical form is the difference equation form, which will be discussed here [38]. For the general ARMA model

$$\varphi(B)\ z_t = \theta(B)\ a_t$$

where

$$\varphi(B) = \phi(B)\ \nabla^d$$

the forecast value is defined as $z_{t+\ell}$, $\ell \geq 1$, and its estimated value is $z_t(\ell)$. In other words, the forecast $z_{t+\ell}$ is said to be made at origin t for lead time ℓ when we are currently standing at time t.

An observation $z_{t+\ell}$ generated by the process may be expressed directly in terms of the difference equation by

$$z_{t+\ell} = \varphi_1 z_{t+\ell-1} + \dots + \varphi_{p+d} z_{t+\ell-p-d} - \theta_1 a_{t+\ell-1} - \dots$$

$$- \theta_q a_{t+\ell-q} + a_{t+\ell} \qquad (4.1.1)$$

Now, suppose, standing at time t, then the forecast function $\hat{z}_t(\ell)$ of $z_{t+\ell}$ will be a linear function of current and previous observations z_t, z_{t-1}, z_{t-2}, ... and also a linear function of current and previous shocks a_t, a_{t-1}, a_{t-2}, ...; the forecast function may be written as,

$$\hat{z}_t(\ell) = \varphi_1 \hat{z}_{t+\ell-1} + \dots + \varphi_{p+d} \hat{z}_{t+\ell-p-d} - \theta_1 a_{t+\ell-1} -$$

$$\dots - \theta_q a_{t+\ell-q} + \hat{a}_{t+\ell} \qquad (4.1.2)$$

Box and Jenkins indicate (4.1.2) is the minimum mean square error forecast function [39]. To obtain the forecast value $\hat{z}_t(\ell)$, the right hand side of the forecast function in (4.1.2) should be treated as follows:

1) The z_{t-j} (j=0,1,2,...), which have already happened at time t, are left unchanged.
2) The z_{t+j} (j=0,1,2,...), which have not yet happended, are replaced by their forecasts $\hat{z}_t$ (j).
3) The a_{t-j} (j-0,1,2,...), which have happened, are available from $z_{t-j} - \hat{z}_{t-j-1}$ (1).

4) The a_{t+j} (j=1,2,...), which have not yet happened, are replaced by zero.

4.2. The Confidence Limits of the Forecast Value

Suppose the forecasts at lead time 1,2,..., L, are required. To obtain probability limits for these forecast value, it is necessary to calculated the weights ψ_1, ψ_2, ..., ψ_{L-1}, which are the parameters of the pure moving-average model; it may be written as,

$$z_t = \psi(B)\ a_t \tag{4.2.1}$$

for the general ARMA model,

$$\varphi(B)\ z_t = \theta(B)\ a_t \tag{4.2.2}$$

Comparing (4.2.1) with (4.2.2), we can obtain

$$\varphi(B)\ (1+\psi_1 B + \psi_2 B^2 + \ldots\ ) = \theta(B) \tag{4.2.3}$$

or

$$\varphi(B)\quad \psi(B) = \theta(B) \tag{4.2.4}$$

$$(1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_{p+d} B^{p+d})(1 + \psi_1 B + \psi_2 B^2 + \ldots) =$$

$$(1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q) \tag{4.2.5}$$

As we equate the coefficients of powers of B in (4.2.5), we can obtain the pure moving-average parameters in terms of general ARMA parameters φ's and θ's, which are known. Then we obtain,

$$\psi_1 = \varphi_1 - \theta_1$$

$$\psi_2 = \varphi_1\psi_1 + \varphi_2 - \theta_2$$

$$\vdots$$

$$\psi_j = {}_1\psi_{j-1} + \dots + \varphi_{p+d}\,\psi_{j-p-d} - \theta_j \qquad (4.2.6)$$

where

$$\psi_0 = 1, \quad \psi_j = 0 \qquad \text{for} \quad j < 0$$

and

$$\theta_j = 0 \qquad \text{for} \quad j > q$$

If k is the greater of the integers p+d-1 and q, then for j>k, the ψ's satisfy the difference equation [40];

$$\psi_j = {}_1\psi_{j-1} + {}_2\psi_{j-2} + \dots + {}_{p+d}\psi_{j-p-d}$$

Thus the ψ's can be easily calculated recusively.

Box and Jenkins suggest the variance of the forecast error ℓ steps ahead for any origin t is the expected value of $\hat{e}_t^2(\ell) = \{z_{t+\ell} - \hat{z}_t(\ell)\}^2$, it can be estiamted by [41]

$$v(\ell) = \{1 + \sum_{j=1}^{\ell-1} \psi_j^2\}\,\sigma_a^2$$

Then assuming that the a's are normal, and given information up to time t, the conditional probability distribution $P(z_{t+\ell}/z_t, z_{t-1} \dots)$ of a future value $z_{t+\ell}$ of the process will be normal with mean $\hat{z}_t(\ell)$ and standard deviation $\left\{1 + \sum_{j=1}^{\ell-1} \psi_j^2\right\}^{1/2} \sigma_a$.

The variance σ_a can be estimated by S_a if the numbers of observations on which such an estimate is based is at least fifty; S_a^2 is the minimum sum of square of residual and can be acquired by $\frac{S(\hat{\phi},\hat{\theta})}{n}$, [42].

Hence the approximate $1-\varepsilon$ probability limits $z_{t+\ell}(-)$ and $z_{t+\ell}(+)$ for $z_{t+\ell}$ are given by

$$z_{t+\ell}\ (\pm) = \hat{z}_t(\ell) \pm \mu_{\varepsilon/2} \left\{1 + \sum_{j=1}^{\ell-1} \psi_j^2\right\}^{1/2} S_a \qquad (4.2.7)$$

Where $\mu_{\varepsilon/2}$ is the deviate exceeded by a proportion $\varepsilon/2$ of the unit normal distribution,

for 50% limits, $\mu_{\varepsilon/2}$ is 0.674

for 95% limits, $\mu_{\varepsilon/2}$ is 1.960

The $z_{t+\ell}$ (-) and $z_{t+\ell}$ (+) mean that, given the information available at origin t, there is a probability of $1-\varepsilon$, that the actual value $z_{t+\ell}$, when it occurs. will be within them; it can be expressed statistically as,

$$P_r\ \{z_{t+\ell}\ (-) < z_{t+\ell} < z_{t+\ell}\ (+)\} = 1 - \varepsilon \qquad (4.2.8)$$

The confidence limits obtained here is applied to individual forecasts $z_{t+\ell}$ only and not jointly to the forecast values at all the different lead times.

The Program FORCAT in the appendix will calculate the forecast values as well as its confidence limits.

CHAPTER FIVE

APPLICATION

The technique of model building for time series have been discussed in previous chapters. The computer programs in the Appendix provides the model calculations [43]. The model has to be constructed by computer calculation and human reasoning.

The process of model building is concerned with relating a class of statistical models to the data at hand and involves much more than model fitting. Thus, identification techniques, designed to suggest what particular kind of model might be worth considering, are developed first and make use of the autocorrelation and partial autocorrelation function. The fitting of the identified model to a time series using the likelihood function can then supply maximum likelihood estimate of the parameters. The initially fitted model will not, necessarily provide adequate representation. Hence diagnostic checks are developed to detect model inadequacy and thus, where necessary, to initiate a further iterative cycle of identification, estimation and diagnostic checking. When the forecast is the objective, the fitted statistical model with past data is used directly to generate optimal forecasts by simple recursive calculation.

The application of these techniques are presented by three examples of time series, which are obtained from industry process [44], business situation [45] and inventory simulation process [7] respectively.

5.1. Example One

A set of data shown on Table 5.1. about an industrial chemical process is to be analyzed here. [44]. This series represent "uncontrolled" outputs of concentration from the chemical process. And they were collected on full scale processes where it was necessary to maintain some output quality characteristics as close as possible to a fixed level. To achieve this control, another variable had been manipulated to approximately concel out variation in the output. However, the effect of these manipulation on the output was in each case accurately known, so that it was possible to compensate numerically for the control action. That is to say, it was possible to calculate very nearly, the values of the series that would have been obtained if no corrective action been taken. It is these compensated value which are recorded here and referred to as "the uncontrolled" series [46].

The obtaining of the appropriate model will be explained step by step in the following sub-sections. Not only will we understand the system from the derived model, but we will acquire optimal forecast values for the series.

5.1.1. Identification of the Model

Program IDENT calculates the autocorrelations and partial autocorrelation of the time series. Since the series represent the "uncontrolled" behavior of the process output, we might expect it possess nonstationary characteristics. So differences of data are taken to see what kind of model can properly represent the series. The output of z, ∇z and $\nabla^2 z$ are shown on Table 5.2., 5.3 and 5.4. respectively. The plotting

TABLE 5-1 CHEMICAL PROCESS CONCENTRATION READINGS:
EVERY TWO HOURS
(READ ROWWISE FROM LEFT TO RIGHT)

17.0	16.6	16.3	16.1	17.1
16.9	16.8	17.4	17.1	17.0
16.7	17.4	17.2	17.4	17.4
17.0	17.3	17.2	17.4	16.8
17.1	17.4	17.4	17.5	17.4
17.6	17.4	17.3	17.0	17.8
17.5	18.1	17.5	17.4	17.4
17.1	17.6	17.7	17.4	17.8
17.6	17.5	16.5	17.8	17.3
17.3	17.1	17.4	16.9	17.3
17.6	16.9	16.7	16.8	16.8
17.2	16.8	17.6	17.2	16.6
17.1	16.9	16.6	18.0	17.2
17.3	17.0	16.9	17.3	16.8
17.3	17.4	17.7	16.8	16.9
17.0	16.9	17.0	16.6	16.7
16.8	16.7	16.4	16.5	16.4
16.6	16.5	16.7	16.4	16.4
16.2	16.4	16.3	16.4	17.0
16.9	17.1	17.1	16.7	16.9
16.5	17.2	16.4	17.0	17.0
16.7	16.2	16.6	16.9	16.5
16.6	16.6	17.0	17.1	17.1
16.7	16.8	16.3	16.6	16.8
16.9	17.1	16.8	17.0	17.2
17.3	17.2	17.3	17.2	17.2
17.5	16.9	16.9	16.9	17.0
16.5	16.7	16.8	16.7	16.7
16.6	16.5	17.0	16.7	16.7
16.9	17.4	17.1	17.0	16.8
17.2	17.2	17.4	17.2	16.9
16.8	17.0	17.4	17.2	17.2
17.1	17.1	17.1	17.4	17.2
16.9	16.9	17.0	16.7	16.9
17.3	17.8	17.8	17.6	17.5
17.0	16.9	17.1	17.2	17.4
17.5	17.9	17.0	17.0	17.0
17.2	17.3	17.4	17.4	17.0
18.0	18.2	17.6	17.8	17.7
17.2	17.4			

Table 5.2 Estimated Autocorrelations and its Partials of Chemical Process Concentration Readings about z

Lag	Autocorrelations	Partial Autocorrelations
1	0.57	0.57
2	0.49	0.25
3	0.39	0.07
4	0.35	0.06
5	0.32	0.06
6	0.34	0.12
7	0.39	0.15
8	0.32	-0.03
9	0.30	0.01
10	0.25	-0.02
11	0.18	-0.07
12	0.16	-0.02
13	0.19	0.06
14	0.23	0.08
15	0.14	-0.12
16	0.18	0.04
17	0.19	0.09
18	0.20	0.06
19	0.14	-0.07
20	0.18	0.05
21	0.10	-0.10
22	0.12	0.05

Table 5.3 Estimated Autocorrelations and its Partials of Chemical Process Concentration Readings about ∇z

Lag	Autocorrelations	Partial Autocorrelations
1	-0.41	-0.41
2	0.02	-0.18
3	-0.06	-0.16
4	-0.01	-0.14
5	-0.07	-0.19
6	-0.02	-0.21
7	0.14	-0.00
8	-0.06	-0.04
9	0.03	-0.02
10	0.02	0.04
11	-0.04	-0.00
12	-0.06	-0.07
13	-0.01	-0.10
14	0.16	0.10
15	-0.17	-0.08
16	0.03	-0.13
17	0.01	-0.09
18	0.08	0.04
19	-0.12	-0.07
20	0.15	0.09
21	-0.12	-0.07
22	0.04	0.02
23	-0.06	-0.06
24	0.04	-0.04
25	0.00	-0.01

Table 5.4 Estimated Autocorrelations and Its Partials of Chemcial Process Concentration Readings about $\nabla^2 z$

Lag	Autocorrelations	Partial Autocorrelations
1	-0.65	-0.65
2	0.18	-0.42
3	-0.04	-0.31
4	0.03	-0.20
5	-0.04	-0.17
6	-0.04	-0.31
7	0.13	-0.17
8	-0.11	-0.14
9	0.04	-0.14
10	0.02	-0.05
11	-0.02	0.02
12	-0.02	0.02
13	-0.04	-0.16
14	0.17	0.05
15	-0.19	0.06
16	0.07	-0.00
17	-0.03	-0.12
18	0.09	0.01
19	-0.16	-0.12
20	0.19	0.07
21	-0.16	-0.02
22	0.10	0.06
23	-0.08	0.01
24	0.05	-0.01
25	-0.01	0.01

of the autocorrelation function and its partials for z, ∇z and $\nabla^2 z$ are also shown on Fig. 5.1., 5.2., 5.3., respectively.

From Fig. 5.1, the autocorrelation function decreases fairly regularly after the first lag, and the partial autocorrelation has the tendency of tailing off; this is to suggest that the process might be ARMA (1,0,1). However, the autocorrelation function of z does not fall quickly. This suggests that the series might be nonstationary. The appropriate estimate of the initial parameters can be calculated from

$$\rho_1 = \frac{(1-\theta_1\phi_1)(\phi_1-\theta_1)}{1+\theta_1^2 - 2\phi_1\theta_1} , \qquad \rho_2 = \rho_1\phi_1$$

and hence we obtain $\phi_1 = 0.86$, $\theta_1 = 0.78$; the model can thus be written as

$$(1-0.86\ B)\ z_t = (1-0.78\ B)\ a_t \qquad (5.1.1.1)$$

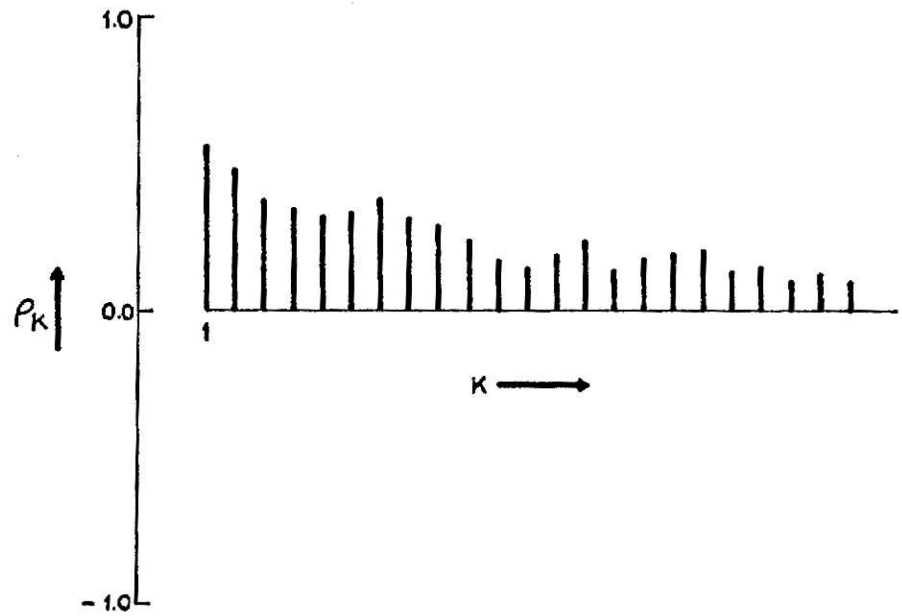
From Fig. 5.2, the autocorrelation function are small after the first lag, and the partial autocorrelation tails off. This suggests an MA(1) process; the approximate estimate of initial parameters can be calculated from

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

and hence $\theta_1 \simeq 0.5$; the model can thus be written as

$$\nabla z_t = (1-0.5\ B)\ a_t$$

or



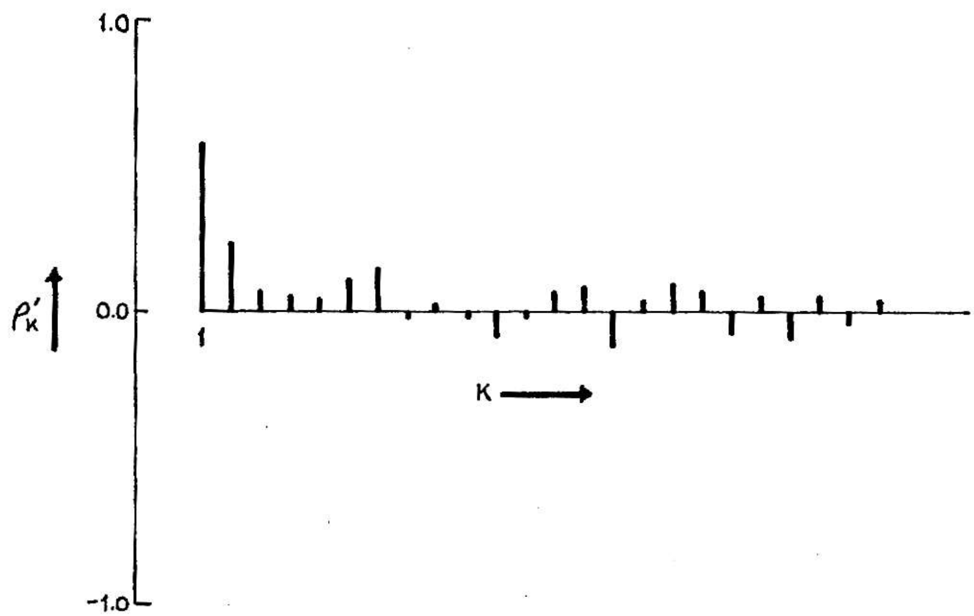


Figure 5.1 Estimated Autocorrelation and its Partials of Chemical Process Concentration Readings about z

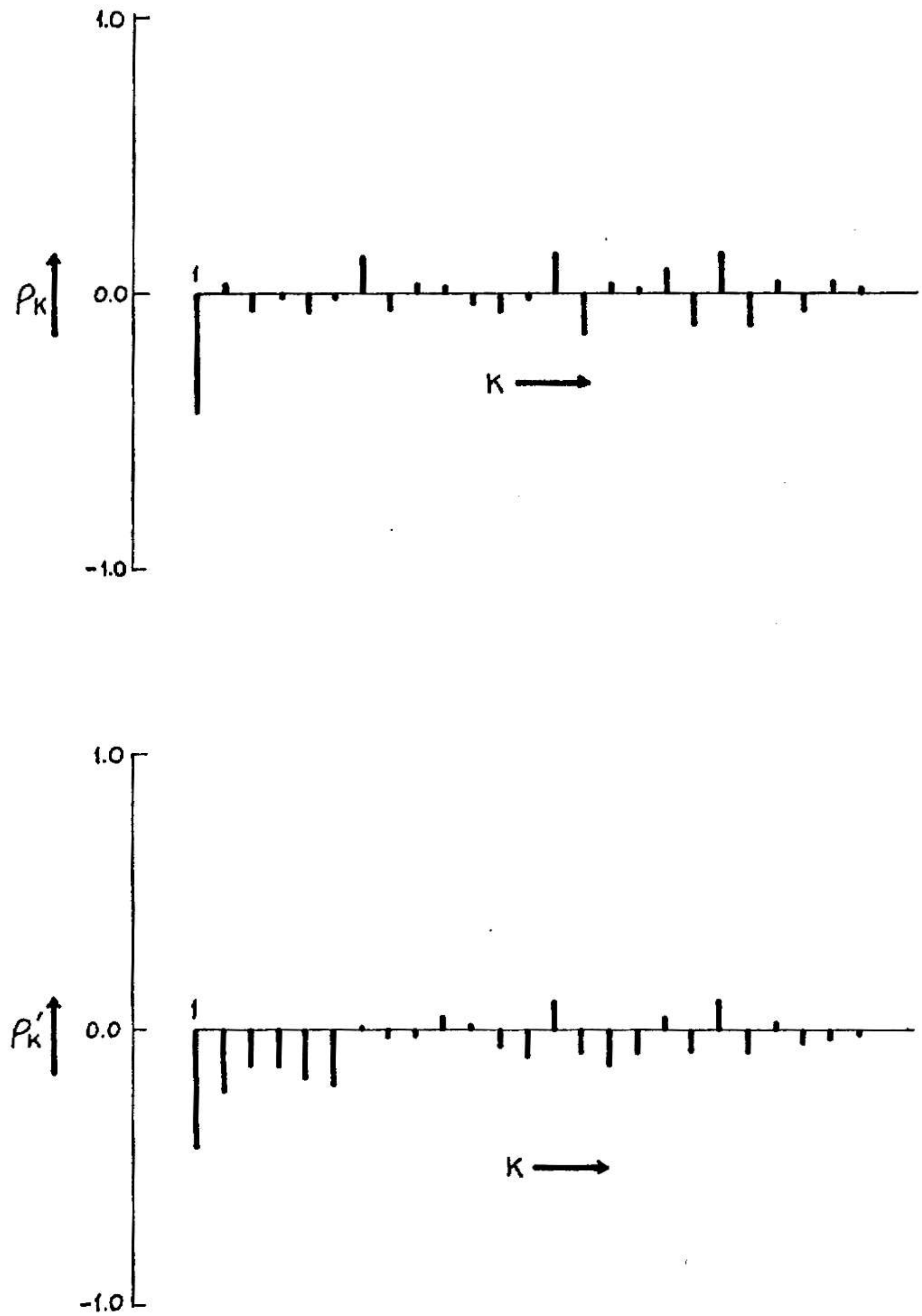


Figure 5.2 Estimated Autocorrelation and its Partials of Chemical Process Concentration Readings about ∇z

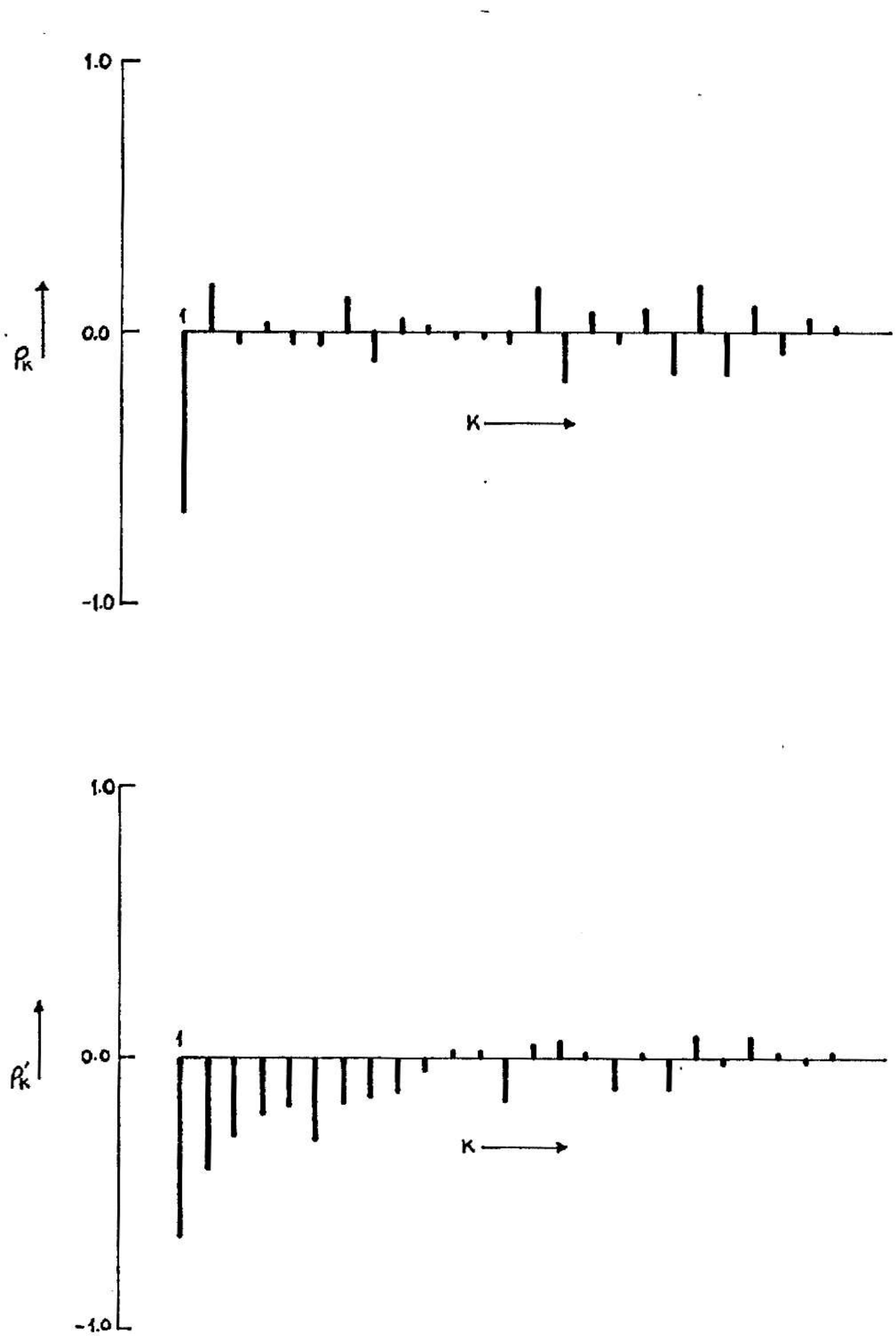


Figure 5.3 Estimated Autocorrelation and its Partials of Chemical Process Concentration Readings about $\nabla^2 z$

$$(1\text{-}B)\ z_t = (1\text{-}0.5\ B)\ a_t \qquad (5.1.1.2)$$

Comparing (5.1.1.1) with (5.1.1.2), we see two possible result in the same form. Either form might represent the time series. However, in doubtful cases, it may be advantageous in employing the nonstationary model rather than the stationary alternative. Hence, the MA(1,1) is to be adapted to represent the given time series and will be subjected to a diagnostic check.

5.1.2. Efficient Estimation of Parameters

Program ESTIM performs the maximum likelihood estimate of the parameters in (5.1.1.2). The convergent situations is shown on Table 5.5. Hence, the appropriate model to represent the time series can now be written in more perfect form as

$$\nabla z_t = (1\text{-}0.7\ B)\ a_t \qquad (5.1.1.3)$$

(5.1.1.3) will be subject to further test for the goodness of fit. The statistical methods described in Section 3.2. are applied here to investigate the model. The sample correlation coefficients of residuals is also obtained from the output of program ESTIM and are shwon in Table 5.6. By the autocorrelation check method, we compare the autocorrelation coefficients of residuals on Table 5.6. with the "control" line $\pm\ 2n^{-1/2}$. It is revealed that all the correlation coefficients of the residuals are within the "control" lines. Thus there is no suspicion of inadequacy of the model.

To test the goodness of fit by the method of a portmanteau lack of fit, the value of

Table 5.5 Iterative estimation of θ_1 for Chemical process concentration data

Iteration	θ_1
0	0.500
1	0.521
2	0.596
3	0.657
4	0.680
5	0.688
6	0.691
7	0.691

Table 5.6 The Sample Correlation Coefficients of Residuals for the Chemical Process Concentration Data

Lag	Correlation
1	0.091
2	0.010
3	-0.096
4	-0.112
5	-0.118
6	0.003
7	0.146
8	0.022
9	0.041
10	0.001
11	-0.099
12	-0.119
13	-0.036
14	0.062
15	-0.131
16	-0.010
17	0.045
18	0.073
19	-0.034
20	0.085
21	-0.091
22	-0.027
23	-0.058
24	0.037
25	0.041

$$Q = n \sum_{k=1}^{k} \gamma_k^2(\hat{a}) \tag{5.1.1.4}$$

is assumed distribute approximately as $\chi^2(k-p-q)$ if the model is adequate, when $n=N-d$ is the number of z's to fit the model. By taking the first 20 autocorrelation coefficients on Table 5.6. to substitute on (5.1.1.4), we obtain

$$Q = n \sum_{k=1}^{20} \gamma_k^2(\hat{a}) = 23.58$$

with 19 degrees of freedom. The 10% and 5% points for χ^2 with 19 degree of freedom, are 27.2 and 30.1 respectively. For 27.2 and 30.1 both far greater than 23.58, there is no significant inadequacy of the model.

5.1.3. Forecasting

Now MA(1,1) is supposed to represent the time series. The forecast values and its individual confidence limits are obtained by Program FORCAT. The forecast function can be written as

$$z_{t+\ell} = (1-0.7\ B)\ a_{t+\ell}$$

or

$$z_{t+\ell} = z_{t+\ell-1} + a_{t+\ell} - 0.7\ a_{t+\ell-1}$$

The $a_{t+\ell}$ beyond the present time is assumed as zero. Hence, for all lead time, the forecasts at origin t will follow a straight line parallel to the time axis. Table 5.7 shows the forecast values and its confidence intervals. Fig. 5.4. shows parts of the chemical process and its forecast values.

Table 5.7 Forecast Value and its 95% Confidence Limits for the Chemical Process Concentration Data

Time	Forecast Value	Upper Limit	Lower Limit
198	17.501	16.879	18.124
199	17.501	16.850	18.153
200	17.501	16.822	18.181
201	17.501	16.795	18.207
202	17.501	16.770	18.233
203	17.501	16.745	18.258
204	17.501	16.721	18.282
205	17.501	16.698	18.305
206	17.501	16.675	18.328

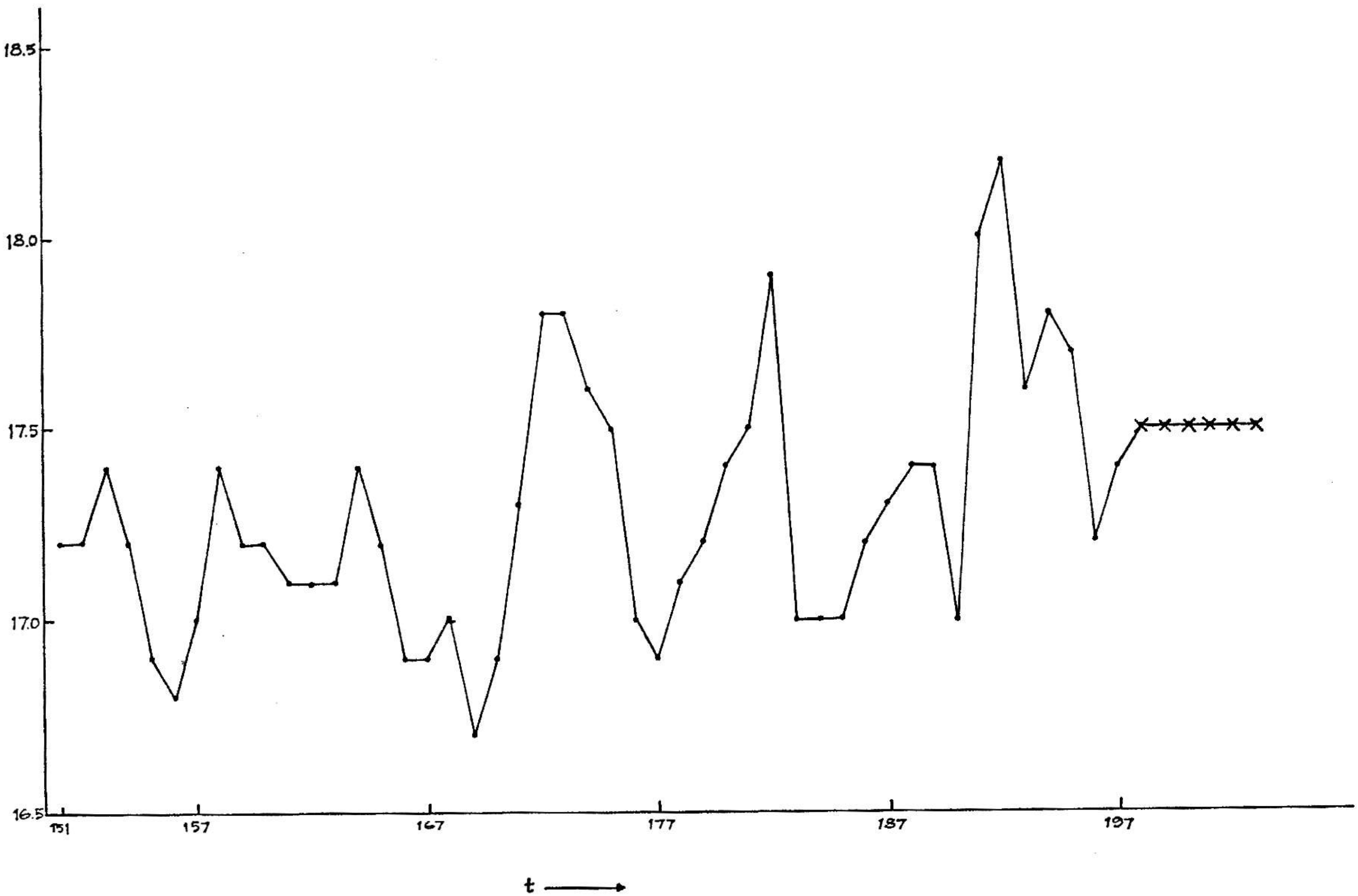


Figure 5.4 Part of Chemical Process and its Forecast vaules

5.2. Example Two

The totals of international airline passengers for 1952, 53, and 54 shown on Fig. 1.3. is to be analyzed here. It is part of a longer series (twelve years of data) quoted by Brown. [47]. The series shows a marked seasonal pattern since travel is at its highest in the late summer months.

Many other series, particularily sales data, show similiar seasonal characteristics. In general, we say that a series exhibits periodic behavior with period S, when similarities in the series occur after S basic time intervals. In this example, we can see apparantly from Fig. 1.3., the basic time interval is one month and the period is S=12 months.

When we have series exhibiting seasonal behavior with known periodicity S, it is of value to set down the data in the form of a table containing S columns. The logarithms of the airline data taken by Box and Jenkins is shown on Table 5.8. As indicated by Box and Jenkins, logarithm are often taken before analyzing sales data and other series of this kind, because it is the percentage fluctuation which might be expected to be comparable at different sales volumes [48].

5.2.1. Identification of Model

Program IDENT provides the autocorrelation and its partials of the original time series and their differences. The outputs are shown on Table 5.9., 5.10., 5.11., 5.12. and Fig. 5.5., 5.6., 5.7., 5.8. respectively.

TABLE 5.8 NATURAL LOGARITHMS OF MONTHLY PASSENGER TOTAALS(MEASURED IN THOUSANDS) IN INTERNATIONAL AIR TRAVEL

	JAN.	FEB.	MAR.	APR.	MAY	JUNE	JULY	AUG.	SEPT.	OCT.	NOV.	DEC.
949	4.718	4.771	4.883	4.860	4.796	4.905	4.997	4.997	4.913	4.779	4.664	4.771
950	4.745	4.836	4.949	4.905	4.828	5.004	5.136	5.136	5.063	4.890	4.736	4.942
951	4.977	5.011	5.182	5.094	5.147	5.182	5.293	5.293	5.215	5.088	4.984	5.112
952	5.142	5.193	5.263	5.199	5.209	5.384	5.438	5.489	5.342	5.252	5.147	5.268
953	5.278	5.278	5.464	5.460	5.434	5.493	5.576	5.606	5.468	5.352	5.193	5.303
954	5.318	5.236	5.460	5.425	5.455	5.576	5.710	5.680	5.557	5.434	5.313	5.434
955	5.489	5.451	5.587	5.595	5.598	5.753	5.897	5.849	5.743	5.613	5.468	5.628
956	5.649	5.624	5.759	5.746	5.762	5.924	6.023	6.004	5.872	5.724	5.602	5.724
957	5.753	5.707	5.875	5.852	5.872	6.045	6.146	6.146	6.001	5.849	5.720	5.817
958	5.829	5.762	5.892	5.852	5.894	6.075	6.196	6.225	6.001	5.883	5.737	5.820
959	5.886	5.835	6.006	5.981	6.040	6.157	6.306	6.326	6.138	6.009	5.892	6.004
960	6.023	5.969	6.038	6.133	6.157	6.282	6.433	6.407	6.230	6.133	5.966	6.068

Table 5.9 Estimated Autocorrelations and its Partials for the International Airline Passenger Data about z_t

Lag	Autocorrelation	Partial Autocorrelation
1	0.953	0.953
2	0.898	-0.118
3	0.851	0.055
4	0.808	0.023
5	0.779	0.115
6	0.756	0.044
7	0.737	0.040
8	0.727	0.098
9	0.734	0.203
10	0.744	0.063
11	0.758	0.114
12	0.762	-0.050
13	0.717	-0.483
14	0.663	-0.037
15	0.618	0.045
16	0.576	-0.043
17	0.544	0.027
18	0.519	0.039
19	0.501	0.039
20	0.491	0.015
21	0.498	0.075
22	0.506	-0.036
23	0.517	0.053
24	0.521	0.039
25	0.484	-0.194
26	0.437	-0.037

Table 5.10 Estimated Autocorrelations and its Partials for the International Airline Passenger Data about ∇z_t

Lag	Autocorrelation	Partial Autocorrelation
1	0.202	0.202
2	-0.122	-0.170
3	-0.150	-0.093
4	-0.320	-0.310
5	-0.083	0.010
6	0.021	-0.081
7	-0.110	-0.206
8	-0.335	-0.494
9	-0.116	-0.188
10	-0.109	-0.540
11	0.206	-0.291
12	0.842	0.585
13	0.218	0.025
14	-0.141	-0.177
15	-0.117	0.115
16	-0.277	-0.001
17	-0.053	0.020
18	0.013	-0.109
19	-0.115	0.080
20	-0.337	-0.066

Table 5.11 Estimated Autocorrelations and its Partials for the International Airline Passenger Data about $\nabla_{12}' z_t$

Lag	Autocorrelation	Partial Autocorrelation
1	0.710	0.710
2	0.616	0.224
3	0.477	-0.047
4	0.435	0.098
5	0.383	0.049
6	0.313	-0.057
7	0.241	-0.050
8	0.193	0.006
9	0.151	-0.012
10	-0.004	-0.281
11	-0.118	-0.164
12	-0.247	-0.153
13	-0.144	0.300
14	-0.142	0.057
15	-0.101	0.049
16	-0.144	-0.037
17	-0.096	0.126
18	-0.108	-0.060
19	-0.143	-0.147
20	-0.157	-0.015

Table 5.12 Estimated Autocorrelation and its Partials for the International Airline Passenger Data about $\nabla\nabla_{12}z_t$

Lag	Autocorrelation	Partial Autocorrelation
1	-0.336	-0.336
2	0.091	-0.024
3	-0.188	-0.186
4	0.009	-0.129
5	0.066	0.033
6	0.016	0.025
7	-0.046	-0.058
8	0.001	-0.013
9	0.167	0.212
10	-0.068	0.046
11	0.063	0.059
12	-0.390	-0.343
13	0.156	-0.104
14	-0.056	-0.079
15	0.139	-0.024
16	-0.127	-0.140
17	0.063	0.028
18	0.028	0.113
19	-0.024	-0.016
20	-0.106	-0.157

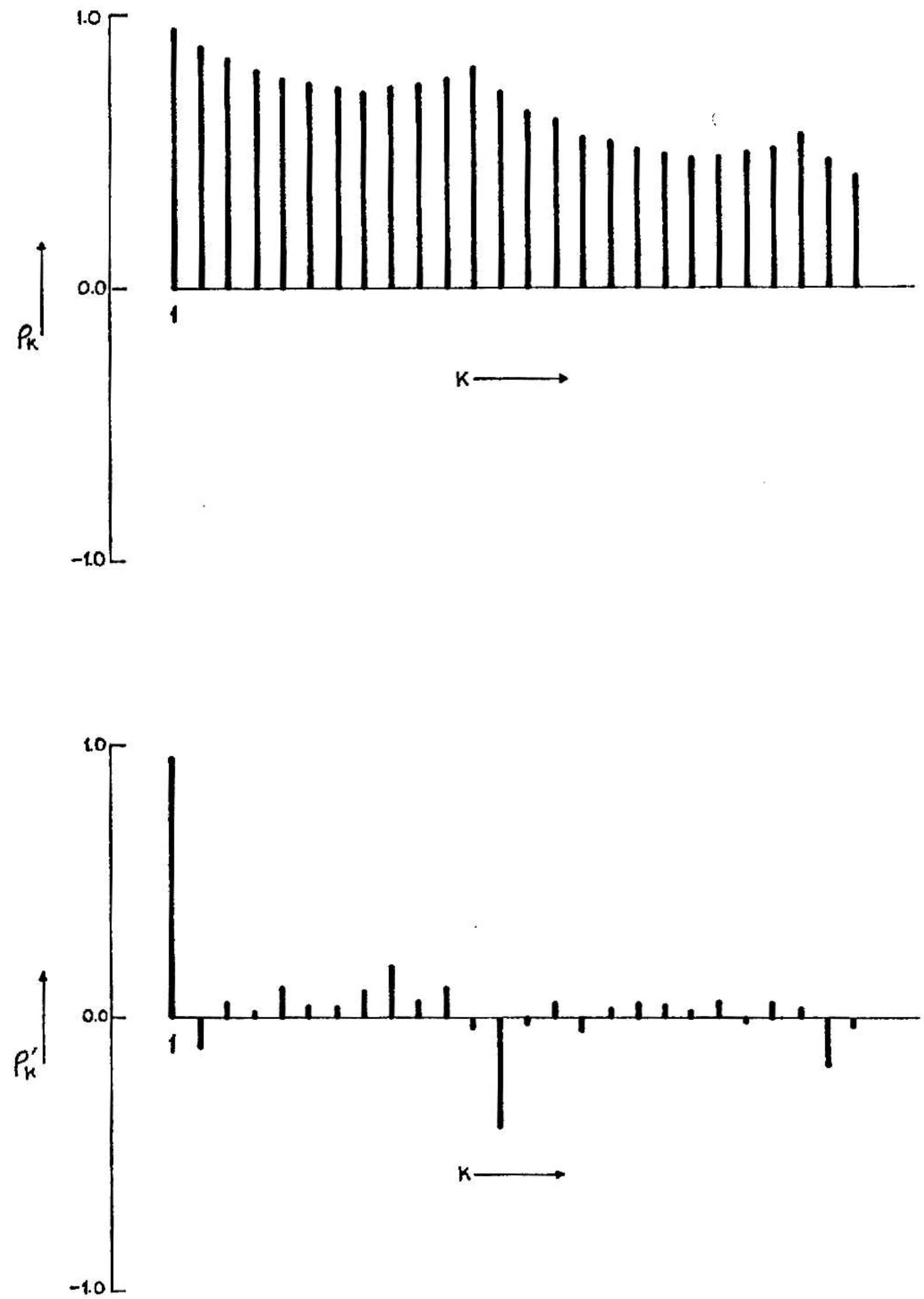


Figure 5.5 Estimated Autocorrelation and its Partials of the airline data about z_t

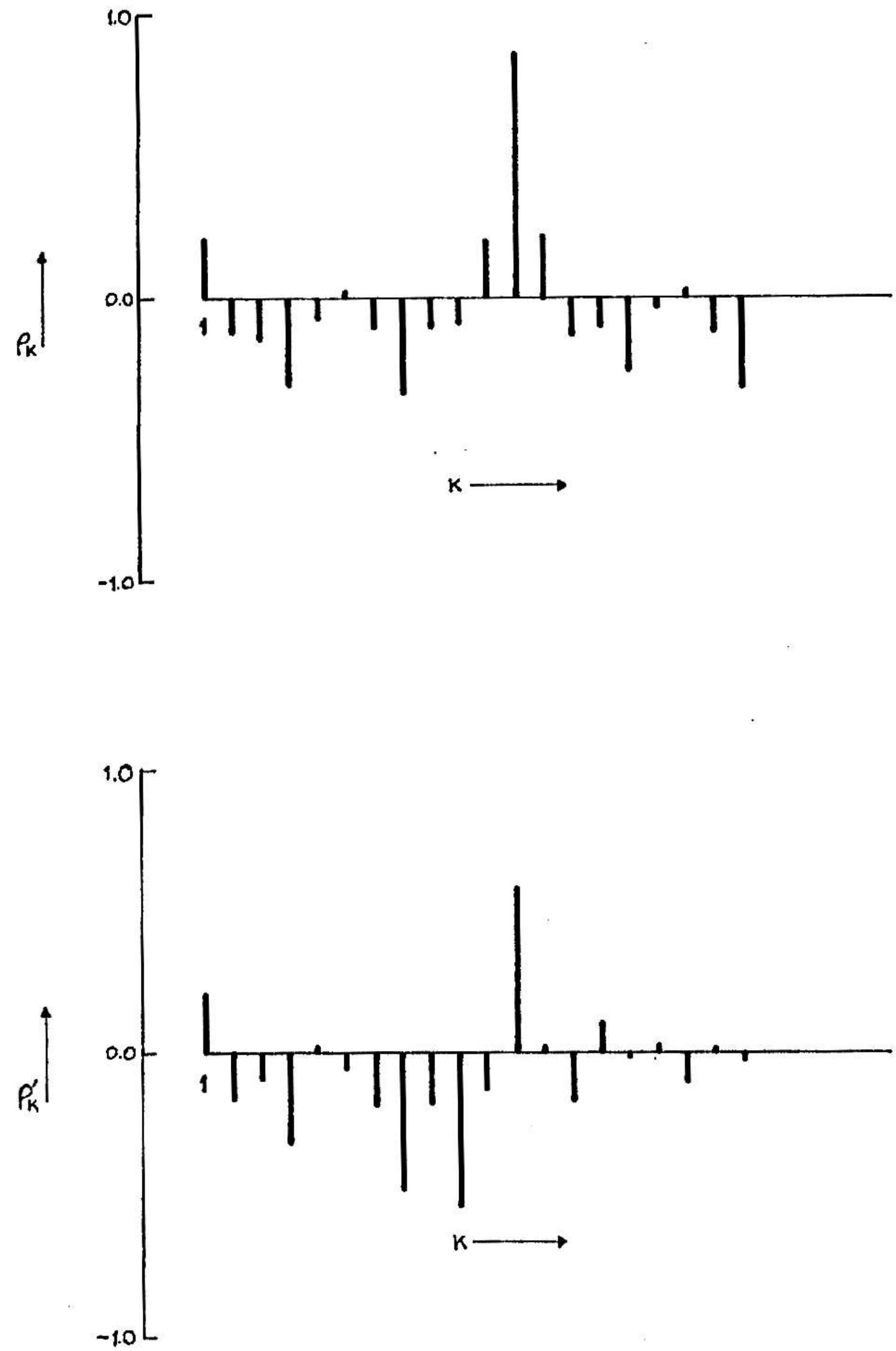


Figure 5.6 Estimated Autocorrelation and its Partials of the airline data about ∇z_t

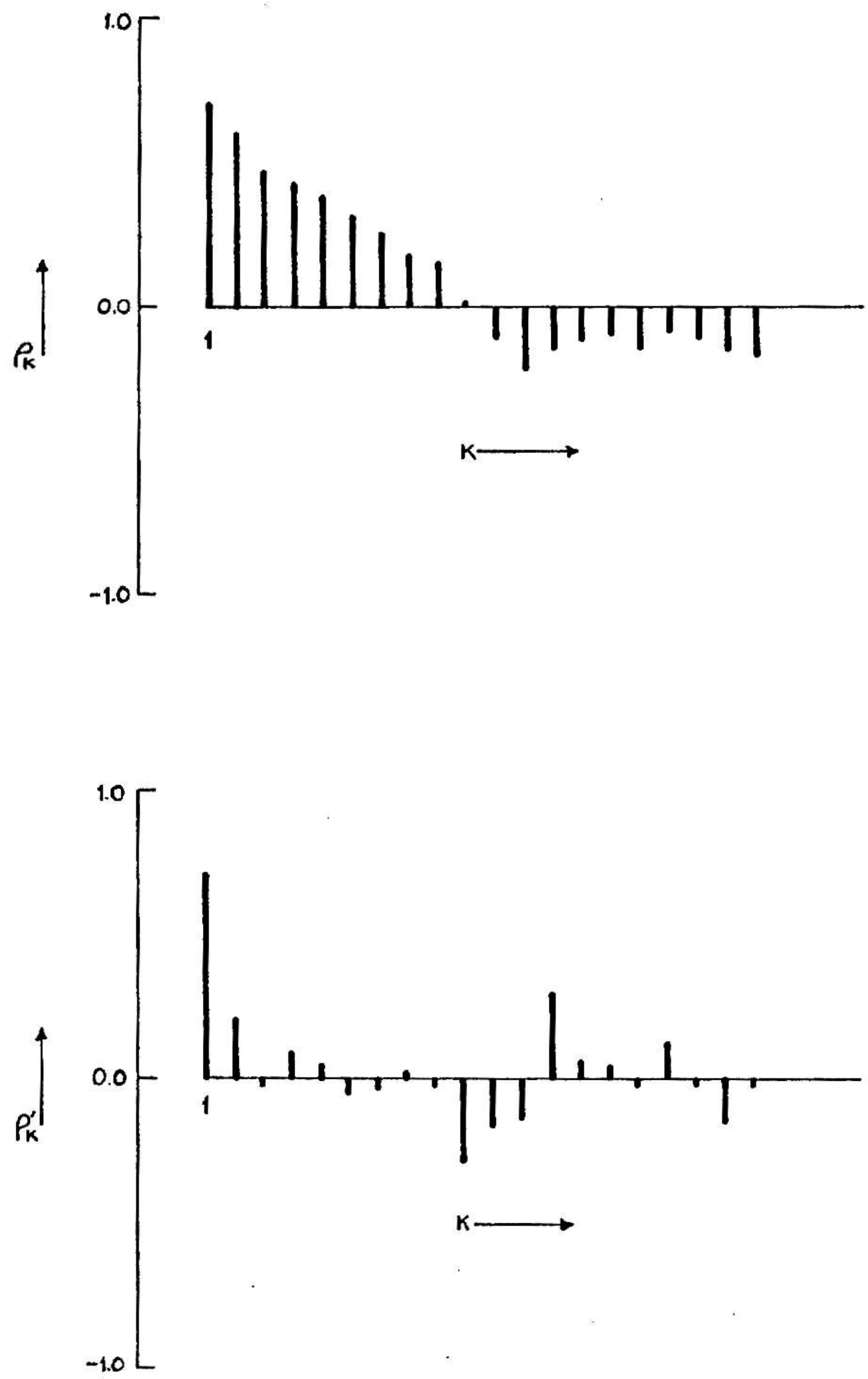


Figure 5.7 Estimated Autocorrelation and its Partials of the airline data about $\nabla'_{12} z_t$

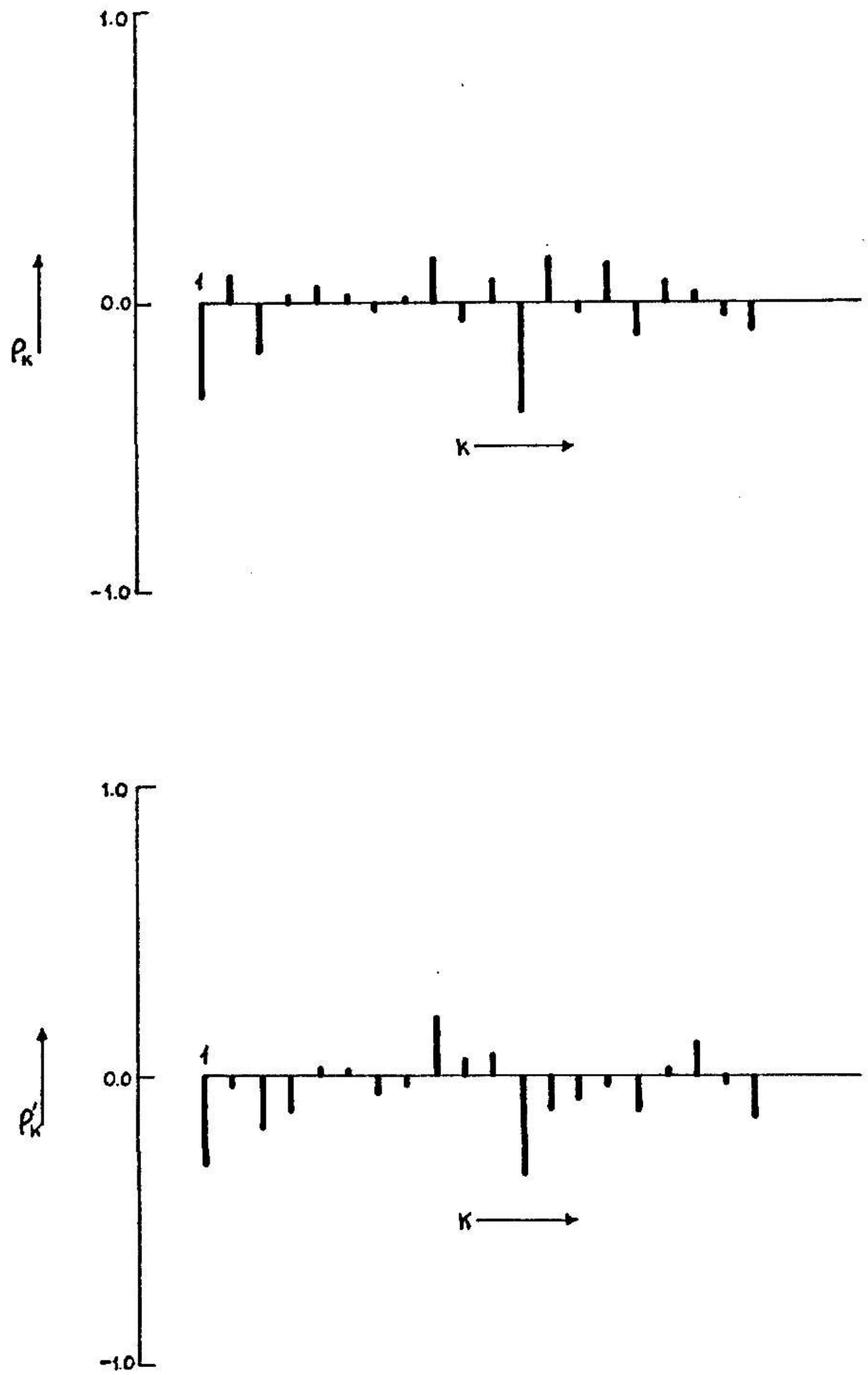


Figure 5.8 Estimated Autocorrelation and its Partials of the airline data about $\nabla'\nabla'_{12}z_t$

In Fig. 5.5., the autocorrelations for z are large and fail to die out at higher lags. This implies the possibility of nonstationarity. The highly correlated periods at lags 12, 24 suggest the seasonal period of this time series 12.

In Fig. 5.6., while the first difference reduces the correlation in general, a very heavy periodic component remains. This is inducted by the large lag of 12.

In Fig. 5.7., simple differencing with respect to a period of twelve results in correlation which are first persistently positive and then persistently negative. This implies that the cyclic component of twelve periods and nonstationarity both exist in the time series.

In Fig. 5.8., the differencing $\nabla^1\nabla^1_{12}$ markedly reduces correlation throughout. The autocorrelation beyond the first lag are compartively small. The partial autocorrelation has the tendency to tail off. Besides, as indicated by Box and Jenkins, a simple and widely applicable stochastic model for the analysis of nonstationary time series is MA(1) [49]. Hence the model to represent the time series is suggested to be

$$\nabla\nabla_{12}z_t = (1 - \theta B)\ (1 - \Theta B^{12})\ a_t \tag{5.2.1.1}$$

which will be subject to further investigation.

As with the seasonal model, by equating the observed correlation to their expected values, approximate values can be obtained for the parameters θ and Θ . On substituting the sample estimates $\rho_1 = -0.34$ and $\rho_{12} = -0.39$ in the expressions, which is obtained from (2.4.7.),

$$\rho_1 = \frac{-\theta}{1+\theta^2}, \qquad 12 = \frac{-\Theta}{1+\Theta^2}$$

The rough estimate of parameters in (5.2.1.1) is $\hat{\theta} \approx 0.39$ and $\Theta \simeq 0.48$.

5.2.2. Efficient Estimate of Parameters.

Program ESTIM provides the maximum likelihood estimates of nonlinear parameters of (5.2.1.1). Table 5.13 shows the converge situations of parameters. The entertained model of the time series can be expressed as

$$\nabla\nabla_{12} z_t = (1 - 0.436\ B)\ (1 - 0.486\ B^{12})\ a_t \qquad (5.2.2.1)$$

The program ESTIM also provides the sample correlations of residuals. Table 5.14 shows the sample correlation coefficients of residuals of the time series. The goodness of fit can be tested as follows.

(1) By autocorrelation check, comparing the autocorrelation coefficients of residuals on Table 5.14. with the "control" line $n^{-1/2}$, few individual correlations appear little large. However, among 20 random deviates one would expect some large deviation. We will further investigate the model to check the goodness of fit.

(2) By the method of a portmanteau lack of fit test, the value of

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

is approximately distributed as $\chi^2(k-p-q)$ if the model is appropriate. Hence, by taking the first 20 autocorrelation of the a's as a whole from Table 5.14., we can obtain

$$Q = n \sum_{k=1}^{20} \gamma_k^2(\hat{a}) = 20.44 \qquad (5.2.2.2)$$

Table 5.13 Iterative Estimation of θ and Θ for the logged airline data

Iteration	θ	Θ
Starting Values	0.390	0.480
1	0.396	0.482
2	0.417	0.487
3	0.432	0.488
4	0.436	0.486
5	0.436	0.486

Table 5.14 Correlation Coefficients of Residuals for the Logged Airline Data

Lag	Correlation
1	0.026
2	0.014
3	-0.138
4	-0.177
5	0.035
6	0.113
7	-0.041
8	-0.031
9	-0.091
10	-0.162
11	-0.033
12	-0.011
13	0.034
14	0.027
15	0.033
16	-0.182
17	0.015
18	0.044
19	-0.084
20	-0.077

Comparing Q with the value of $\chi^2(18)$ on χ^2 table, the 10% and 5% points for χ^2 value, with 18 degrees of freedom, are 27.2 and 30.1, respectively. The Q value in (5.2.2.2) is smaller than 27.2, no indication of lack of fit is indicated. Hence (5.2.2.1) is proposed as the appropriate model to represent the international airline passenger situation.

5.2.3. Forecasting

Program FORCAT provides the forecast values and its confidence limits of the model with given time series. The results are shown on Table 5.15. and Fig. 5.9. We can predict the future business of the international airline passengers is to be increased with the cyclic period of twelve. Travel is at its highest in the summer months, while a secondary peak occurs in the spring.

5.3. Example Three

A set of observations about an inventory simulation process shown on Table 5.16. will be analyzed [7].

5.3.1. Identification of the Model

Program IDENT computes the autocorrelation and its partials for the original time series and its differences. Table 5.17, 5.18. and Fig. 5.10., 5.11. show the output of the autocorrelation and its partials.

In Fig. 5.10., the autocorrelation function is damped exponentially and tails off. While the partial autocorrelation is cut off after the first lag. This suggests that the process might possibly be a first

Table 5.15 Forecast Value and its 95% Confidence Limits for the International Airline Passenger Data

Time	Forecast Value	Upper Limit	Lower Limit
145	6.117	6.047	6.188
146	6.053	5.972	6.134
147	6.122	6.032	6.213
148	6.217	6.118	6.316
149	6.241	6.135	6.348
150	6.366	6.253	6.480
151	6.517	6.397	6.638
152	6.491	6.364	6.618
153	6.314	6.181	6.447
154	6.217	6.078	6.356
155	6.050	5.906	6.195
156	6.173	6.023	6.322
157	6.219	6.051	6.387
158	6.155	5.977	6.334
159	6.224	6.036	6.412
160	6.319	6.121	6.517
161	6.343	6.136	6.550
162	6.468	6.253	6.683
163	6.619	6.396	6.843
164	6.593	6.362	6.825
165	6.416	6.177	6.655

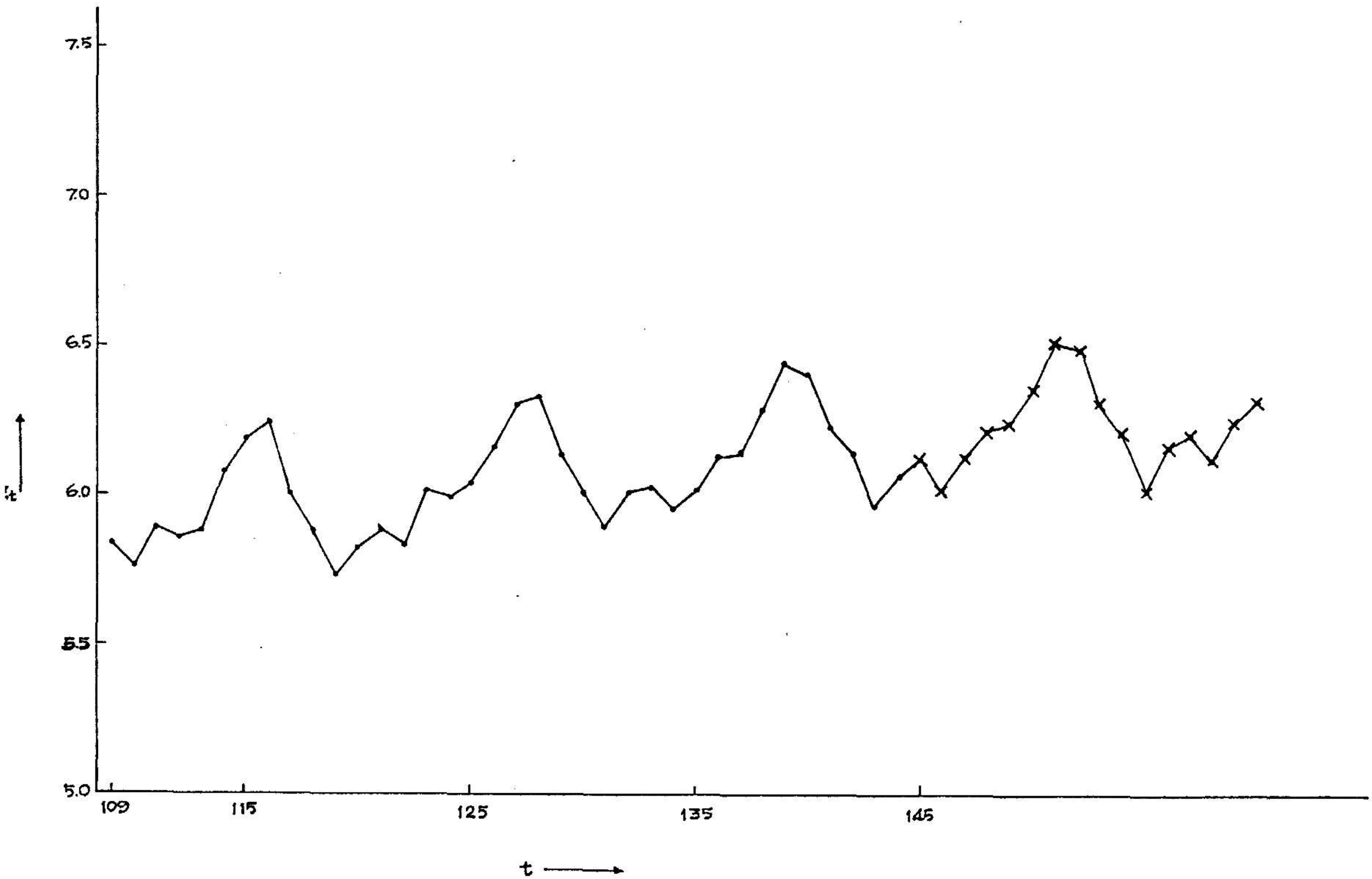


Figure 5.9 Part of the airline passenger data and its forecast

TABLE 5.16 SIMULATED INVENTORY PROCESS READINGS
(READ ROWWISE FROM LEFT TO RIGHT)

488.056	476.056	463.056	449.056	424.056
410.056	398.056	383.056	363.056	346.056
335.056	316.056	297.056	376.056	361.056
337.056	318.056	302.056	288.056	276.056
259.056	251.057	237.057	225.057	212.057
198.057	181.057	171.057	152.057	130.057
115.010	108.914	108.914	106.914	188.010
176.010	170.010	163.010	156.010	150.010
145.010	139.010	128.010	124.010	117.010
109.010	103.010	95.010	90.010	82.010
70.010	164.010	152.010	149.010	144.010
128.010	123.010	114.010	104.010	98.010
85.010	80.010	75.010	70.010	66.010
57.010	47.010	37.010	33.010	28.010
24.914	24.914	24.914	24.914	18.914
87.010	79.010	72.010	67.010	60.010
52.010	47.010	42.010	38.010	27.010
24.010	19.010	109.096	103.096	98.096
93.096	85.096	79.096	662.009	653.009
650.009	902.056	883.056	875.056	849.056
834.056	816.056	789.056	779.056	769.960
763.960	759.960	749.960	747.960	735.960
727.960	719.960	759.056	752.056	732.056
711.056	699.056	679.056	657.056	747.056
734.056	717.056	695.056	678.056	665.056
644.056	626.009	619.009	617.009	608.009
599.009	594.009	584.009	581.009	572.009
566.009	563.009	554.009	542.009	536.009
532.009	526.009	519.009	512.009	510.914
604.009	600.009	599.009	591.009	581.009
577.009	564.009	559.009	555.009	543.009
535.009	531.009	526.009	517.009	506.009
497.009	487.009	481.009	571.009	567.009
657.056	634.056	624.056	607.056	586.056
568.056	555.056	529.056	514.056	500.056
488.009	485.009	476.009	472.009	466.009
464.009	460.009	449.009	443.009	438.009
527.009	521.009	512.009	510.009	503.009
494.009	491.009	485.009	482.009	474.009
466.009	460.009	641.056	635.056	616.056
601.056	587.056	572.056	648.056	635.056
625.056	610.056	600.056	575.056	562.056
549.056	540.056	522.056	504.056	487.056
470.056	459.056	447.056	425.056	814.056
801.056	787.056	771.056	752.056	738.056

717.056	695.C56	687.056	677.056	659.056
644.056	635.C56	607.960	599.960	589.960
652.056	634.C56	616.056	598.056	580.C56
563.056	547.C56	521.056	506.056	495.056
480.056	457.C56	439.056	425.056	415.056
404.056	386.C56	464.056	456.056	844.056
836.056	816.C56	806.056	787.056	777.056
756.056	747.C56	729.056	712.056	699.056
689.056	667.C56	655.056	643.056	628.056
614.056	6C0.C56	688.056	673.056	650.056
633.056	621.C56	6C9.056	601.056	593.056
585.056	576.C56	558.056	539.056	526.056
507.056	496.C56	481.056	462.056	450.056
432.056	414.C56	396.056	380.056	367.960
450.056	433.C56	415.056	407.056	391.056
373.056	358.C56	341.056	326.056	312.056
297.056	279.C56	264.056	241.056	226.056
215.056	199.C56	273.056	254.056	240.009
233.009	227.CC9	579.056	567.056	539.056
524.056	506.C56	494.056	473.056	463.056
446.056	433.C56	423.056	410.056	397.056
381.056	364.C56	452.056	434.056	425.056
417.056	402.C56	394.056	374.056	351.056
333.056	305.C56	296.056	278.056	264.056
259.056	244.C56	235.056	224.056	212.056
194.056	184.C56	164.056	249.056	634.056
620.056	613.C56	601.056	590.056	575.056
559.056	542.C56	534.056	515.056	505.056
485.056	470.C56	454.056	438.056	420.056
409.056	390.C56	374.056	360.056	340.056
922.056	905.C56	893.960	889.960	962.056
947.056	937.C56	924.056	911.056	896.056
878.056	868.C56	842.056	827.056	813.056
801.056	792.C56	776.056	765.056	755.056
744.056	824.C56	797.056	781.056	770.056
754.056	735.C56	721.056	711.056	700.056
695.056	680.C56	665.056	656.056	647.056
634.056	626.C56	610.056	1001.056	985.056
968.056	949.960	937.960	1016.056	1003.056
992.056	980.C56	965.056	956.056	937.056
910.056	895.C56	880.056	866.056	854.056
844.056	827.C56	814.056	799.056	785.C56
870.056	856.C56	847.056	822.056	809.056
798.056	779.C56	756.056	748.056	725.056
712.056	699.C56	679.056	663.056	647.056
630.056	612.C56	600.056	579.056	571.056
554.056	640.C56	628.056	610.056	595.056
573.056	562.C56	552.056	538.056	518.056
506.056	491.C56	483.009	474.009	469.009
456.009	450.CC9	445.009	444.009	435.009

424.009	732.C56	713.056	697.056	780.056
766.056	753.C56	743.056	734.056	717.056
701.056	685.C56	665.056	649.056	637.056
631.056	6C8.C56	597.056	581.056	549.056
535.056	523.C56	506.056	495.960	483.960
475.96C	543.C56	528.056	509.056	488.056
868.056	849.C56	838.056	831.056	817.056
799.056	784.C56	773.C56	754.056	746.056
730.056	714.C56	702.056	693.960	774.056
763.056	745.C56	736.056	711.056	694.056
68C.056	665.C56	646.C56	624.056	613.056
602.056	585.C56	573.056	555.056	538.056
633.056	621.C56	608.056	598.056	584.056
580.056	564.C56	553.056	537.056	524.056
515.056	502.C56	488.056	463.056	451.056
428.056	4C4.C56	397.056	381.056	373.056
356.056	343.C56	325.056	705.960	695.960
683.960	671.960	661.960	651.960	643.960
635.960	625.960	609.960	603.960	599.960
591.960	6C9.C56	598.056	582.056	562.056
548.056	539.960	529.960	612.056	604.056
585.056	568.C56	554.056	538.056	523.056
506.056	498.C56	475.056	461.056	440.056
416.056	4C3.C56	381.C56	360.056	350.056
343.056	325.960	407.056	389.056	379.056
365.056	352.C56	341.056	330.056	318.056
301.056	288.C56	278.056	264.056	255.056
242.056	22C.C56	211.056	200.056	191.056
180.056	167.C56	151.009	143.009	132.914
130.914	218.CC9	206.009	204.009	198.009
191.009	187.CC9	500.C56	484.056	469.056
453.056	439.C56	425.056	415.056	401.056
386.056	373.C56	364.056	433.056	419.056
399.056	381.C56	373.C56	361.056	338.056
320.056	306.C56	290.056	277.056	258.056
248.056	229.C56	204.056	199.056	183.056
171.056	156.C56	150.056	126.056	212.056
197.056	184.C56	171.056	154.056	142.056
130.056	115.C56	103.056	90.009	84.009
81.009	75.CC9	666.009	659.009	651.009
976.056	955.C56	941.056	928.056	905.056
991.056	969.C56	957.056	943.056	929.056
916.056	898.C56	880.056	869.056	846.056
833.056	830.C56	813.056	796.056	784.056
772.056	756.C56	736.056	715.056	710.056
791.056	775.C56	765.C56	738.056	720.056
708.056	693.C56	684.C56	665.056	638.056
620.056	603.C56	986.056	982.056	971.056
954.056	939.C56	919.056	904.056	882.056

877.056	855.C56	836.056	829.056	807.056
796.056	786.C56	776.056	759.056	742.056
728.056	810.C56	803.056	785.056	764.056
744.056	725.C56	715.056	700.056	685.056
672.056	661.C56	646.056	624.056	598.056
565.056	554.C56	534.056	529.056	512.056
591.056	986.C56	978.056	963.056	938.056
927.056	912.C56	897.056	884.056	868.056
852.056	840.C56	822.056	807.056	791.056
780.056	763.C56	750.056	735.056	719.056
697.056	779.C56	764.056	748.056	729.056
718.056	7C3.C56	689.056	667.056	662.056
649.056	636.C56	622.056	602.056	593.056
584.056	570.C56	654.056	642.056	630.056
611.056	592.C56	575.056	568.056	554.056
527.056	5C9.C56	499.056	878.056	866.056
836.056	819.C56	797.056	774.056	757.056
738.056	724.C56	714.056	695.056	680.056
661.960	742.C56	718.056	701.056	688.056
660.056	645.C56	623.056	614.056	593.056
577.056	554.C56	528.056	508.056	495.056
585.056	572.C56	548.056	534.056	518.056
500.056	476.C56	454.056	439.056	425.056
402.056	386.C56	364.056	351.056	339.056
320.056	306.C56	301.056	273.056	361.056
755.056	734.C56	723.056	712.056	694.056
683.056	666.C56	645.056	634.056	629.056
608.056	588.C56	570.056	556.056	549.056
535.056	523.C56	509.056	497.056	490.056
473.056	458.C56	534.056	515.056	502.056
484.056	459.C56	441.056	417.056	397.056
383.056	370.C56	359.056	347.056	333.056
317.056	299.C56	282.056	266.056	353.056
336.056	320.C56	307.056	290.056	272.056
659.056	641.C56	622.056	600.056	580.056
570.056	558.C56	551.056	533.056	521.056
509.056	497.C56	490.056	478.056	471.960
457.960	443.960	433.960	425.960	487.056
470.056	452.C56	444.056	431.056	416.056
405.056	397.C56	387.056	369.056	354.056
335.056	310.C56	296.056	285.056	269.960
265.960	263.960	253.961	243.961	235.961
229.961	223.961	263.056	248.056	235.056
217.056	201.C56	189.056	170.056	260.056
236.056	225.C56	211.056	596.056	585.056
561.056	543.C56	518.056	495.056	487.056
475.056	467.C56	456.056	449.056	440.056
414.056	397.C56	383.056	364.056	346.056
328.056	912.C56	983.056	963.056	945.056
928.056	912.C56	903.056	890.056	875.056

850.056	836.C56	830.056	816.056	801.056
782.056	772.C56	752.056	736.056	724.056
704.056	787.C56	767.056	752.056	741.C56
724.056	714.C56	700.056	700.056	684.056
666.056	655.C56	640.056	615.056	600.056

Table 5.17 Sample Correlation and its Partials of the Inventory Simulation Process about z_t

Lag	Autocorrelation	Partial Autocorrelation
1	0.96	0.96
2	0.92	-0.02
3	0.88	0.01
4	0.84	-0.07
5	0.80	-0.00
6	0.76	-0.01
7	0.72	-0.00
8	0.69	0.02
9	0.66	0.01
10	0.63	0.02
11	0.60	0.02
12	0.58	0.01
13	0.56	0.16
14	0.53	-0.00
15	0.52	0.03
16	0.51	0.02
17	0.49	0.01
18	0.48	0.02
19	0.46	-0.04
20	0.45	-0.00
21	0.43	-0.03
22	0.41	-0.04
23	0.39	-0.05
24	0.37	0.01
25	0.35	0.02

Table 5.18 Sample Correlation and its Partials of the Inventory Simulation Process about ∇z_t

Lag	Autocorrelation	Partial Autocorrelation
1	-0.000	-0.000
2	-0.031	-0.030
3	0.049	0.049
4	-0.021	-0.021
5	-0.014	-0.011
6	-0.013	-0.017
7	-0.041	-0.040
8	-0.029	-0.030
9	-0.041	-0.043
10	-0.040	-0.039
11	-0.020	-0.023
12	-0.032	-0.033
13	-0.010	-0.012
14	-0.035	-0.042
15	-0.036	-0.040
16	-0.017	-0.028
17	-0.022	-0.030
18	0.039	0.031
19	0.001	-0.010
20	0.024	0.019
21	0.037	0.022
22	0.039	0.032
23	-0.015	-0.025
24	-0.028	-0.038
25	-0.022	-0.032

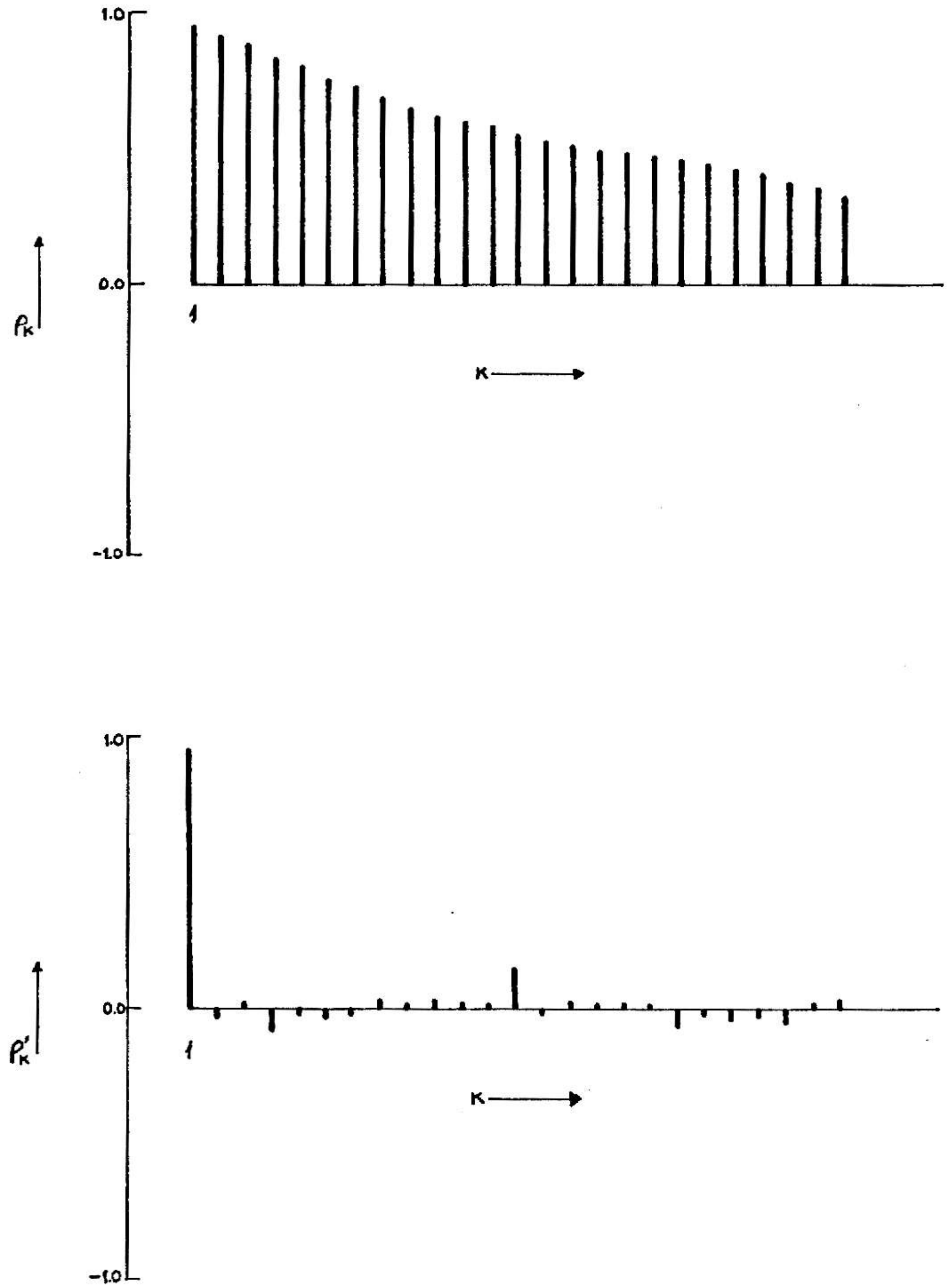


Figure 5.10 Estimated Autocorrelation and its Partials of the simulated Inventory Process about z_t

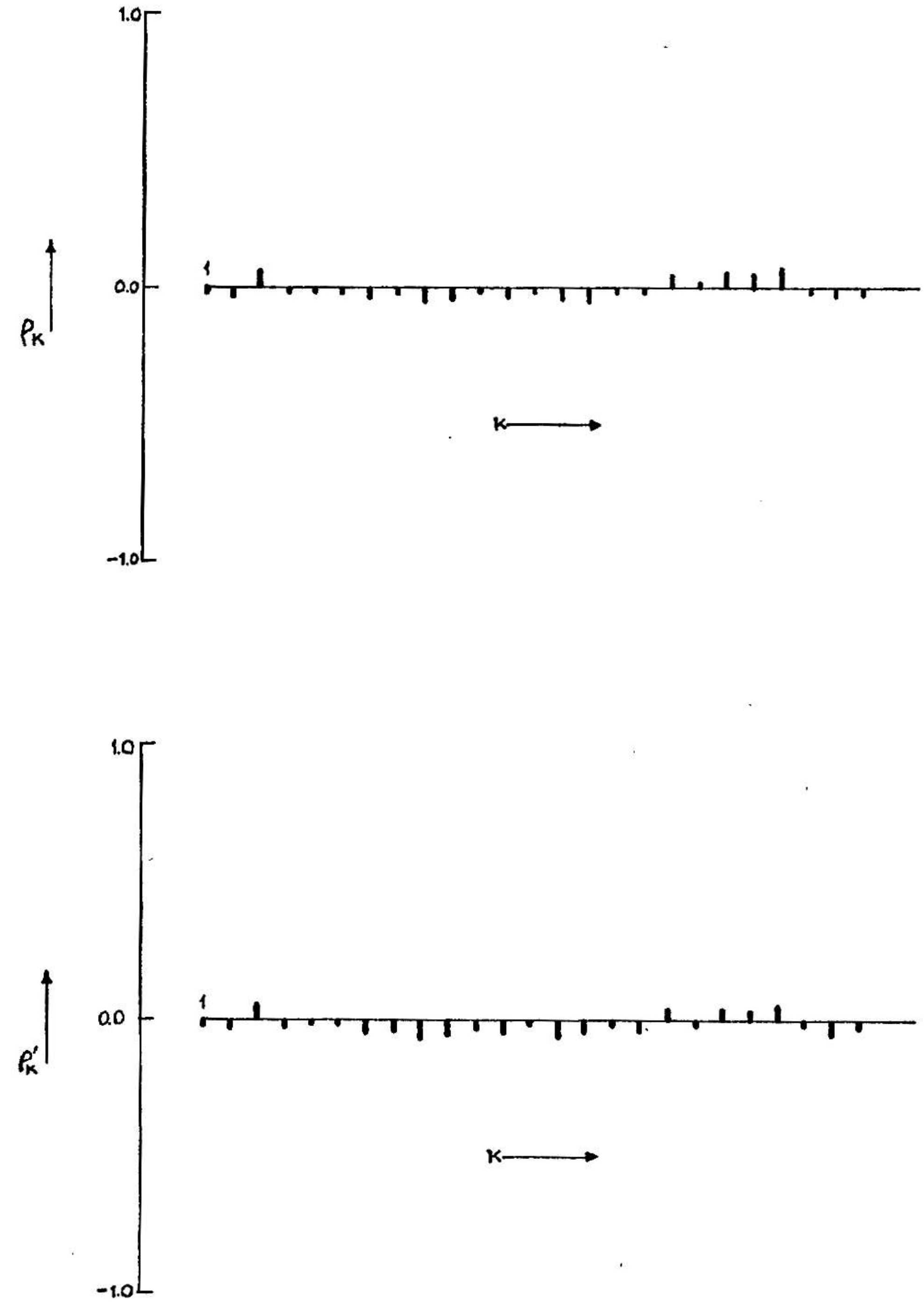


Figure 5.11 Estimated Autocorrelation and its Partials of the simulated inventory process about ∇z_t

order autoregressive process. The estimated parameter of AR(1) can be obtained by applying equation (2.3.7.)

$$\hat{\phi} = \rho_1 = 0.96$$

so, the recommended model might be

$$(1 - 0.96\ B)\ z_t = a_t \qquad (5.3.1.1)$$

In Fig. 5.11, the correlation dies out completely after the first difference. This suggests a possible model might be

$$\nabla z_t = a_t$$

or

$$(1\text{-}B)\ z_t = a_t \qquad (5.3.1.2)$$

Comparing (5.3.1.1) with (5.3.1.2), the model form is similiar except the parameter is a little different. The forecast value of (5.3.1.2) are all the same beyond the lead time ℓ=1. However, the inventory will be depleted little by little. Hence (5.3.1.1) is to be entertained to represent the simulated inventory process and will be subjected to diagnostic check.

5.3.2. Efficient Estimation of Parameter

Program ESTIM computes the maximum likelihood estimation of parameter of (5.3.1.1) with given time series. After the first iteration, the program is stopped execution with the efficient estimation parameter of θ = 0.96. Hence the tentative model is

$$(1 - 0.96\ B)\ z_t = a_t \tag{5.3.2.1}$$

The process of testing the fit the model is explained as follows.

(1) An autocorrelation check is based on the assumption that the estimated autocorrelation of residuals $\gamma_k(a)$ are uncorrelated and distributed approximately about zero with a standard error of $n^{-1/2}$, if the model can represent the time series appropriately. Hence comparing the residual autocorrelation coefficients shown on Table 5.19. with the "control" line $n^{-1/2}$, a few correlations are slightly larger than $n^{-1/2}$. Hence, the model should be subjected to more investigation.

(2) To make a more formal assesment, the portmanteau lack of fit test which is based on the assumption that if the model is appropriate, the value of $Q = n \sum_{k=1}^{k} \gamma_k^2(\hat{a})$ is approximately distributed as $\chi^2(k-p-q)$. Hence, taking 20 autocorrelation coefficients of residual as a whole, we obtain,

$$Q = n \sum_{k=1}^{k} \gamma_k^2(a) = 13.68$$

with 19 degree of freedom. From χ^2 tables, the 10% and 5% points for χ^2, with 19 degrees of freedom are 27.2 and 30.1 respectively. For Q = 13.68 is smaller than 27.2, thus no lack of fit is indicated. Hence the model of (5.3.2.1) is recommended to represent the time series.

Table 5.19 Sample correlations of Residuals of the Inventory Simulation Process Data

Lag	Correlation
1	0.0112
2	-0.0194
3	0.0595
4	-0.0104
5	-0.0052
6	-0.0045
7	-0.0325
8	-0.0213
9	-0.0333
10	-0.0327
11	-0.0134
12	-0.0248
13	-0.0032
14	-0.0291
15	-0.0294
16	-0.0109
17	-0.0161
18	0.0451
19	0.0068
20	0.0300
21	0.0426
22	0.0436
23	-0.0110
24	-0.0238
25	-0.0183

5.3.3. Forecasting

Program FORCAT provides the forecast values and its confidence intervals. The outputs are shown on Table 5.20., Fig. 5.12. respectively.

From Fig. 5.12. of the simulated inventory process, we see the inventory is being replenished when it is depleted to some extent. However, there is no definite replenishment cycle. Hence the model will not forecast replenishments. The AR(1) of $(1 - 0.96 B) z_t = a_t$ can represent this process satisfactorily. From the forecast values of Fig. 5.12., we can predict when the inventory stock will be dropped to what level, and hence, can prepare in advance to order the needed stock.

Table 5.20 Forecast Value and its 95% Confidence Limits for the Simulated Inventory Process Data

Time	Forecast Value	Upper Limit	Lower Limit
1001	576.053	443.28	708.82
1002	553.011	368.96	737.05
1003	530.890	309.88	751.89
1004	509.654	259.37	759.93
1005	489.268	214.75	763.78
1006	469.697	174.61	764.78
1007	450.909	138.05	763.76
1008	432.873	104.49	761.25
1009	415.558	73.49	757.61
1010	398.935	44.73	753.13
1011	382.978	17.94	748.01
1012	367.658	-7.08	742.39
1013	352.952	-30.51	736.42
1014	338.833	-52.50	730.71
1015	325.280	-73.17	723.73
1016	312.269	-92.63	717.17
1017	299.778	-110.97	710.53
1018	287.787	-128.289	703.86
1019	276.275	-144.64	697.19
1020	265.224	-160.11	690.56

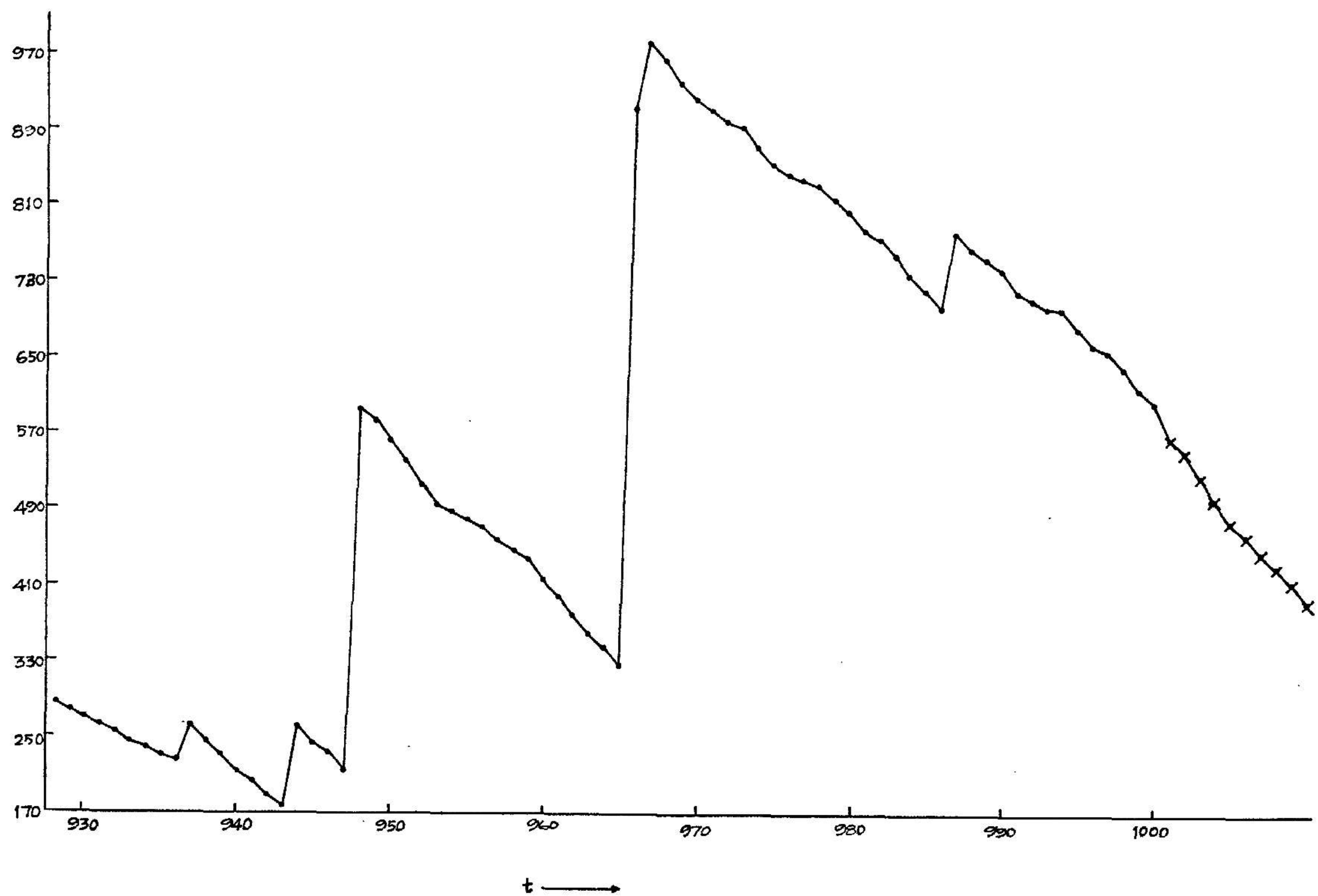


Figure 5.12 Part of Simulated Inventory Process and its Forecast Values

BIBLIOGRAPHY

[1] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970. pp. 23, Sec. 2.1.1.

[2] C.M. Stralkowski, "Lower Order Autoregressive-moving Average Stochastic Models and Their use for the Characterization of Abrasive Cutting Tools", Ph.D. Thesis, University of Wisconsin, 1968, pp 11, Sec. 2.1.

[3] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 23, Sec. 2.1.1

[4] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 10, Sec. 1.2.

[5] C. M. Stralkowski, S.M. Wu, and R.E. DeVor, "Chart for the Interpretation and Estimation of the Second Order Autoregressive Model", Technometrics, Vol. 12, No. 3, Aug. 1970, pp. 669-685.

[6] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 524.

[7] Gadad, S.V., "Numerical Optimization of a Stochastic Inventory System Under Constraint". Master Thesis, Kansas State University. 1971.

[8] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 8, Sec. 1.2.1.

[9] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 9, Sec. 1.2.1.

[10] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 10, Sec. 1.2.

[11] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 192, Sec. 6.3.5.

[12] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1 0, pp. 304, Sec. 9.1.3.

[13] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 305, Sec. 9.1.3.

[14] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., pp. 517.

[15] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., pp. 200.

[16] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 178, Sec. 6.2.2

[17] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 30, Sec. 2.1.4.

[18] Durbin, J. (1960), "The Fitting of Time Series Models", Rev. Int. Inst. Stat., 28, 233.

[19] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1 70, pp. 54, Sec. 3.2.1.

[20] G.U. Yule, "On a Method of Investigating Periodicties in Disturbed Series, with Special Reference to Wolfer's Sunspot Numbers", Phil-Trans., A226, 267, 1927.

[21] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 64, Sec. 3.2.5.

[22] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 46, Sec. 3.1.1.

[23] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 68, Sec. 3.3.2.

[24] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 70, Sec. 3.3.3.

[25] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 75, Sec. 3.4.2.

[26] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 76, Sec. 3.4.2.

[27] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 77, Sec. 3.4.3.

[28] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 201, Appendix A6.2.

[29] Box, G.E.P. and Jenkins. Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 517-520.

[30] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, p. 175, Sec. 6.2.1.

[31] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, p. 208, Sec. 7.1.1.

[32] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 210, Sec. 7.1.2.

[33] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 209, Sec. 7.1.2.

[34] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 219, Sec. 7.1.5.

[35] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 289, Sec. 8.2.1.

[36] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 290, Sec. 8.2.1.

[37] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 2902 pp. 3, Sec. 8.2.2.

[38] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 127, Sec. 5.1.

[39] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 130-131, Sec. 5.1.2.

[40] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 134, Sec. 5.2.2.

[41] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 135, Sec. 5.2.3.

[42] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 269, Appendix A7.4.

[43] C.M. Stralkowski, "Lower Order Autoregressive-moving Average Stochastic Models and Their Use for the Characterization of Abrasive Cutting Tools," Ph.D. Thesis, University of Wisconsin, 1968. Appendix A, B, C, D.

[44] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 525.

[45] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 531.

[46] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 85, Sec. 4.1.1.

[47] R. G. Brown, Smoothing, Forecasting and Preduction of Discrete Time Series, Prentice-Hall, New Jersey, 1962.

[48] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 303, Sec. 9.1.3.

[49] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 313, Sec. 9,2,3.

[50] Marquardt, D. L. (1963), "An Algorithm for Least-Square Estimation of Non-linear Parameters". J. Soc. Indust. Appl. Math., 2, pp. 431-441.

[51] Mood, A. M. Introduction to the Theory of Statistics, McGraw-Hill Book Company, Inc., New York, 1950, pp. 158-161.

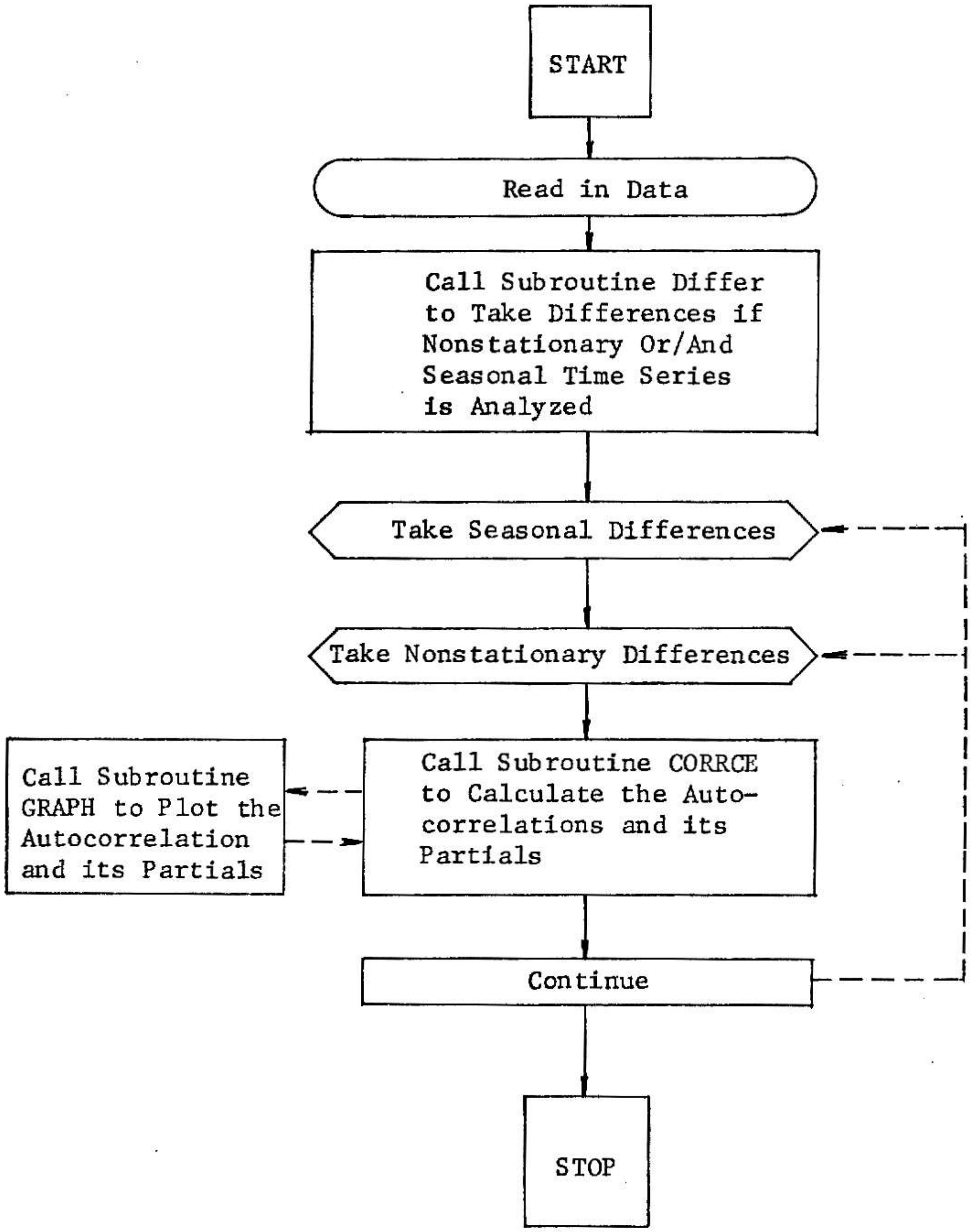
[52] Box, G.E.P. and Jenkins, Time Series Analysis Forecasting and Control, Holden Day, Co., 1970, pp. 214, Sec. 7.1.4.

APPENDIX A

Program IDENT

A-1. Description of Program

Program IDENT is developed to provide the sample correlation and partial correlation functions so as to identify the appropriate model for the given time series. The calculations performed are based on equations (2.1.3), (2.1.4), (2.2.4) and (2.2.5) in Chapter Two. Program IDENT can accommodate stationary time series, non-stationary time series and seasonal time series. The program consists of a main program and three subroutines. Subroutine DIFFER performs data differences if non-stationary or seasonal time series is analyzed. Subroutine CORRCE calculates the correlations and partial correlations of the time series. Subroutine GRAPH is modified from IBM scientific subroutine PLOT to plot out the correlations and partial correlations within $\pm$ 1 range. The flow chart of Program IDENT is constructed on next page.


START
Read in Data
Call Subroutine Differ to Take Differences if Nonstationary Or/And Seasonal Time Series is Analyzed
Take Seasonal Differences
Take Nonstationary Differences
Call Subroutine GRAPH to Plot the Autocorrelation and its Partials
Call Subroutine CORRCE to Calculate the Autocorrelations and its Partials
Continue
STOP

A.2. Description of Input Data

CARD	VARIABLE IN PROGRAM	FORMAT	DESCRIPTION
1	KK	(I10)	Number of correlations to be calculated.
2	IDW	(I10)	Number of differences required. For stationary time series, a zero is entered.
3	IDB,IS	(2I10)	Number of seasonal differences followed by seasonal lag. For stationary time series, two zeros are entered.
4	N	(I10)	Number of observations in original series.
Last N Cards	Z(I,1,1)	(2X,(F20.5))	Observations of original series.

A.3. Description of Output Data

The following output can be generated by Program IDENT.

1. Graphs of the sample correlations and partials of the original time series and of the differenced time series.
2. Sample mean of observations; sample variance of observations; Sample correlations, partials and standard errors.
3. Estimates of the autoregressive parameters for "candidate" AR(p) models, with p taken from 1 to KK, based on the sample partial correlation function.

APPENDIX A. PROGRAM IDENT

A.4 Computer Program

```
FORTRAN IV G LEVEL  18                 MAIN              DATE = 71113          03/38/05

 0001              DIMENSION Z(1000,3,2),NOB(5,4),X(1000),P(1000)
 0002              COMMON/A1/Z
 0003              COMMON/A2/NOB
 0004              COMMON/A3/X,P
             C     READ IN DATA
 0005          200 FORMAT(4I10)
             C     KK      TOTAL NO. OF CORRELATIONS AND PARTIALS TO BE CALCULATED
             C
             C     IDW     NO. OF DIFFERENCES REQUIRED.
             C             FOR STATIONARY TIME SERIES, A ZERO IS ENTERED.
             C
             C     IDB,IS  NO. OF SEASONAL DIFFERENCES FOLLOWED BY SEASONAL LAG.
             C              FOR STATIONARY TIME SERIES, TWO ZEROS ARE ENTERED.
             C
             C     N       NO. OF OBSERVATIONS IN ORIGINAL SERIES.
             C     Z     OBSERVATIONS OF ORIGINAL SERIES.
 0006              READ (1,200) KK
 0007              READ (1,200) IDW
 0008              READ (1,200) IDB,IS
 0009              READ (1,200) N
 0010          201 FORMAT(2X,(F20.5))
 0011              READ(1,201) (Z(I,1,1),I=1,N)
 0012              CALL DIFFER(IDW,IDB,IS,N,MW,MB)
 0013              DO 202 J=1,MB
 0014              DO 202 M=1,MW
 0015         202  CALL CORRCE(KK,M,J,IS)
 0016              END
```

```
FORTRAN IV G LEVEL  18                 DIFFER            DATE = 71113          03/38/05

                  SUBROUTINE DIFFER(IDW,IDB,IS,N,MW,MB)
            C     SUBROUTINE DIFFER PERFORMS DIFFERENCING OPERATIONS ON THE DATA
            C     IF NON-STATIONARY OR SEASONAL SERIES ARE ANALYZED.
            C
                  DIMENSION Z(1000,3,2),NOB(5,4)
                  COMMON/A1/Z
                  COMMON/A2/NOB
            C     CALCULATION OF NOB(M,J)
                  MW=IDW+1
                  MB=IDB+1
                  DO 151 M=1,MW
                  DO 151 J=1,MB
              151 NOB(M,J)=N+1-M+IS-J*IS
            C     DIFFERENCING WITH RESPECT TO DW, NO. OF DIFFERENCES
            C     NOBD=NOB(M,1) IS NO. OF OBSERVATIONS AFTER DIFFERENCE
            C     Z(I,M,1) IS THE OBSERVATIONS AFTER DIFFERENCE BY SUBSTRACTING
            C     THE PRECEEDING OBS. FROM THE CURRENT OBS. M IS THE DIFFERENCE NO. INDE
            C
                  IF(MW-1)152,152,153
              153 DO 154 M=2,MW
                  NOBD=NOB(M,1)
                  DO 154 I=1,NOBD
              154 Z(I,M,1)=Z(I+1,M-1,1)-Z(I,M-1,1)
            C     DIFFERENCING WITH RESPECT TO DB
            C     NOBD=NOB(M,J) IS THE NO. OF OBSERVATIONS AFTER DIFFERENCES M AND
            C     SEASONAL DIFFERENCES J.
            C     Z(I,M,J) IS THE OBSERVATIONS AFTER SEASONAL DIFFERENCE,
            C     J IS THE SEASONAL DIFFERENCE NO. INDEX HERE.
            C
              152 IF(MB-1) 155,155,156
              156 DO 157 M=1,MW
                  DO 157 J=2,MB
                  NOBD=NOB(M,J)
                  DO 157 I=1,NOBD
              157 Z(I,M,J)=Z(I+IS,M,J-1)-Z(I,M,J-1)
              155 CONTINUE
                  RETURN
                  END
```

```
FORTRAN IV G LEVEL  18                CORRCE              DATE = 71113          03/38/05

                   SUBROUTINE CORRCE(KK,M,L,IS)
             C     SUBROUTINE CORRCE CALCULATES THE SAMPLE CORRELATIONS AND PARTIALS AND
             C     THEIR STANDARD ERRORS.
             C
                   DIMENSION ZDUM(1000,3,2),NOB(5,4),Z(1000),X(1000),P(1000),C(1000)
                   DIMENSION R(100),VAR(100),S(100),T(100,100),VART(100),U(100)
                   DIMENSION E(250)
                   COMMON/A4/R
                   COMMON/A1/ZDUM
                   COMMON/A2/NOB
                   COMMON/A3/X,P
                   COMMON/A6/T
               101 FORMAT(1H12X,31HCORRELATION INFORMATION FOR DW=,I4,2X,3HDB=,I4,2X,
                  12HS=,I4//)
                   IDW=M-1
                   IDB=L-1
                   WRITE (3,101) IDW,IDB,IS
                   N=NOB(M,L)
                   DO 100 I=1,N
               100 Z(I)=ZDUM(I,M,L)
                   ZBAR=0.
                   XN=N
                   DO 102 I=1,N
               102 ZBAR=ZBAR+Z(I)
                   ZBAR=ZBAR/XN
                   CO=0.
                   DO 103 I=1,N
               103 CO=CO+(Z(I)-ZBAR)**2
                   CO=CO/XN
             C     CALCULATION OF R, ESTIMATED CORRELATION FUNCTION
                   DO 104 K=1,KK
                   C(K)=0.
                   NN=N-K
                   DO 105 J=1,NN
               105 C(K)=C(K)+(Z(J)-ZBAR)*(Z(J+K)-ZBAR)
                   C(K)=C(K)/XN
               104 R(K)=C(K)/CO
             C     CALCULATION OF T
                   IF(KK-101) 106,106,107
               107 KKK=101-1
                   GO TO 109
               106 KKK=KK-1
             C     RECURSIVE RELATIONS FOR FINDING T(K,K), PARTIAL CORRELATIONS
             C     FUNCTION, GIVEN R(K).
               109 T(1,1)=R(1)
                   T(2,2)=(R(2)-R(1)**2)/(1.-R(1)**2)
                   T(2,1)=T(1,1)-T(2,2)*T(1,1)
                   DO 203 K=2,KKK
                   B=0.
                   A=0.
                   DO 202 J=1,K
                   A=A+T(K,J)*R(K+1-J)
               202 B=B+T(K,J)*R(J)
                   A=R(K+1)-A
                   B=1.-B
                   T(K+1,K+1)=A/B
                   DO 203 J=1,K
               203 T(K+1,J)=T(K,J)-T(K+1,K+1)*T(K,K-J+1)
```

```
      C     PLOT OUT AUTOCORRELATION AND PARTIAL AUTCCORRELATION FUNCTION
      C     E IS PCLTTED FUNCTION
            DO 112 K=1,KK
            K1=KK+K
            E(K)=K
        112 E(K1)=R(K)
            E(2*KK+1)=T(1,1)
            E(2*KK+2)=T(2,2)
            DO 113 K=2,KKK
            K2=(2*KK+2)+K-1
        113 E(K2)=T(K+1,K+1)
            CALL GRAPH (1,E,KK,3,KK)
      C     CALCULATION OF VAR AND VART
      C     VAR(K) IS
      C     AN ESTIMATE OF THE VARIANCE OF THE ESTIMATE OF THE CORRELATIONS,
      C     WHICH CAN BE USED IN A ROUGH TEST FOR WHETHER CORRELATION R IS
      C     EFFECTIVELY ZERO.
      C
            VAR(1)=1./XN
            S(1)=R(1)/SQRT(VAR(1))
            A=2./XN
            DO 204 K=2,KK
            VAR(K)=VAR(K-1)+A*(R(K-1)**2)
        204 S(K)=R(K)/SQRT(VAR(K))
      C     VART(K) IS AN APPROXIMATE ESTIMATE OF THE VARIANCE OF THE SAMPLE
      C     PARTIAL CORRELATIONS, GIVEN THAT THE MODEL IS AR(K-1)
            KKK=KKK+1
            DO 205 K=1,KKK
            A=1./(N-K)
            VART(K)=A
        205 U(K)=T(K,K)/SQRT(VART(K))
      C     WRITE OUT
        601 FORMAT(/2X,5HZBAR=,F20.5,5X,7HVAR(Z)=,F20.5///)
            WRITE (3,601) ZBAR,CO
            SVAR=SQRT(VAR(KK))
            SVAT=SQRT(VART(KKK))
            WRITE(3,250) SVAR,SVAT
        250 FORMAT(/2X,12HS.D.(R(KK))=,F10.6,6X,14HS.D.(PR(KKK))=,F10.6)
        300 FORMAT(/2X,31HSAMPLE CORRELATION COEFFICIENTS//)
            WRITE(3,300)
        301 FORMAT(2X,2HR(,I3,2H)=,F10.5,6X,10HR/S.D.(R)=,F10.5)
            DO 302 I=1,KK
        302 WRITE (3,301) I,R(I),S(I)
        303 FORMAT(////2X,39HSAMPLE PARTIAL CORRELATION COEFFIEIENTS//)
            WRITE(3,303)
        304 FORMAT(2X,3HPR(,I3,2H)=,F10.5,6X,12HPR/S.D.(PR)=,F10.5)
            DO 305 I=1,KKK
        305 WRITE(3,304) I,T(I,I),U(I)
        310 FORMAT(1H12X,25HAUTORGESSIVE PARAMETERS//)
            WRITE(3,310)
        311 FORMAT(2X,2HP=,I3,5X,4HPHI(,I3,2H)=,F10.5)
            DO 312 K=1,KKK
            DO 312 I=1,K
        312 WRITE(3,311) K,I,T(K,I)
            RETURN
            END
```

```
0001               SUBROUTINE GRAPH (NO,A,N,M,NL)
             C
0002               DIMENSION OUT(101),YPR(11),ANG(9),A(1)
             C
0003             1 FORMAT(1H1,60X,7H CHART ,//)
0004             2 FORMAT(1H ,F11.4,5H+    ,101A1)
0005             4 FORMAT(10H *X015678+)
0006             5 FORMAT( 10A1)
0007             7 FORMAT(1H ,16X,101H.         .         .         .         .         .
                1 .         .         .         .         .)
0008             8 FORMAT(1H0,9X,11F10.4//)
0009           200 FORMAT(10X,'     PLOT OF AUTO-CORRELATION AND PARTIAL AUTO-CORRELA
                1TION FUNCTION')
0010           201 FORMAT(10X,'     PLOT OF AUTO-CORRELATION FUNCTION')
             C
             C     ************************************************************
             C
0011               NLL=NL
             C
             C
             C        PRINT TITLE
             C
0012            20 WRITE(3,1)
0013               GO TO (91,92), NO
0014            91 WRITE(3,200)
0015               GO TO 21
0016            92 WRITE(3,201)
0017               GO TO 21
0018            21 CONTINUE
             C
             C        DEVELOP BLANKS AND DIGITS FOR PRINTING
             C
0019               REWIND 4
0020               WRITE(4,4)
0021               REWIND 4
0022               READ(4,5)BLANK,(ANG(I),I=1,9)
0023               REWIND 4
             C
             C        FIND SCALE FOR BASE VARIABLE
             C
0024               XSCAL=(A(N)-A(1))/(FLOAT(NLL-1))
             C
             C        FIND SCALE FOR CROSS VARIABLES
             C
0025               M1=N+1
0026               M2=M*N
0027               YMIN=-1.
0028               YMAX=+1.
0029               YSCAL=(YMAX-YMIN)/100.0
             C
             C        FIND BASE VARIABLE PRINT POSITION
             C
0030               XB=A(1)
0031               MY=M-1
0032               DO 108 I=1,NLL
0033            45 F=I-1
0034               XPR=XB+F*XSCAL
             C
```

```
             C         FIND CROSS VARIABLES
             C
0035              51 DO 55 IX=1,101
0036              55 OUT(IX)=BLANK
0037              57 DO 60 J=1,MY
0038                 LL=I+J*N
0039                 JP=((A(LL)-YMIN)/YSCAL)+1.0
0040                 OUT(JP)=ANG(J)
0041                 IF(JP.EQ.51) GO TO 60
0042                 OUT(51)=ANG(9)
0043                 IF (JP.EQ.101) GO TO 60
0044                 OUT(101)=ANG(9)
0045              60 CONTINUE
             C
             C         PRINT LINE AND CLEAR, OR SKIP
             C
0046                 WRITE(3,2)XPR,(OUT(IZ),IZ=1,101)
0047             108 CONTINUE
             C
             C         PRINT CROSS VARIABLES NUMBERS
             C
0048              86 WRITE(3,7)
0049                 YPR(1)=YMIN
0050                 DO 90 KN=1,9
0051              90 YPR(KN+1)=YPR(KN)+YSCAL*10.0
0052                 YPR(11)=YMAX
0053                 WRITE(3,8)(YPR(IR),IR=1,11)
0054                 RETURN
0055                 END
```

APPENDIX B

PROGRAM ESTIM

B.1. Description of Program

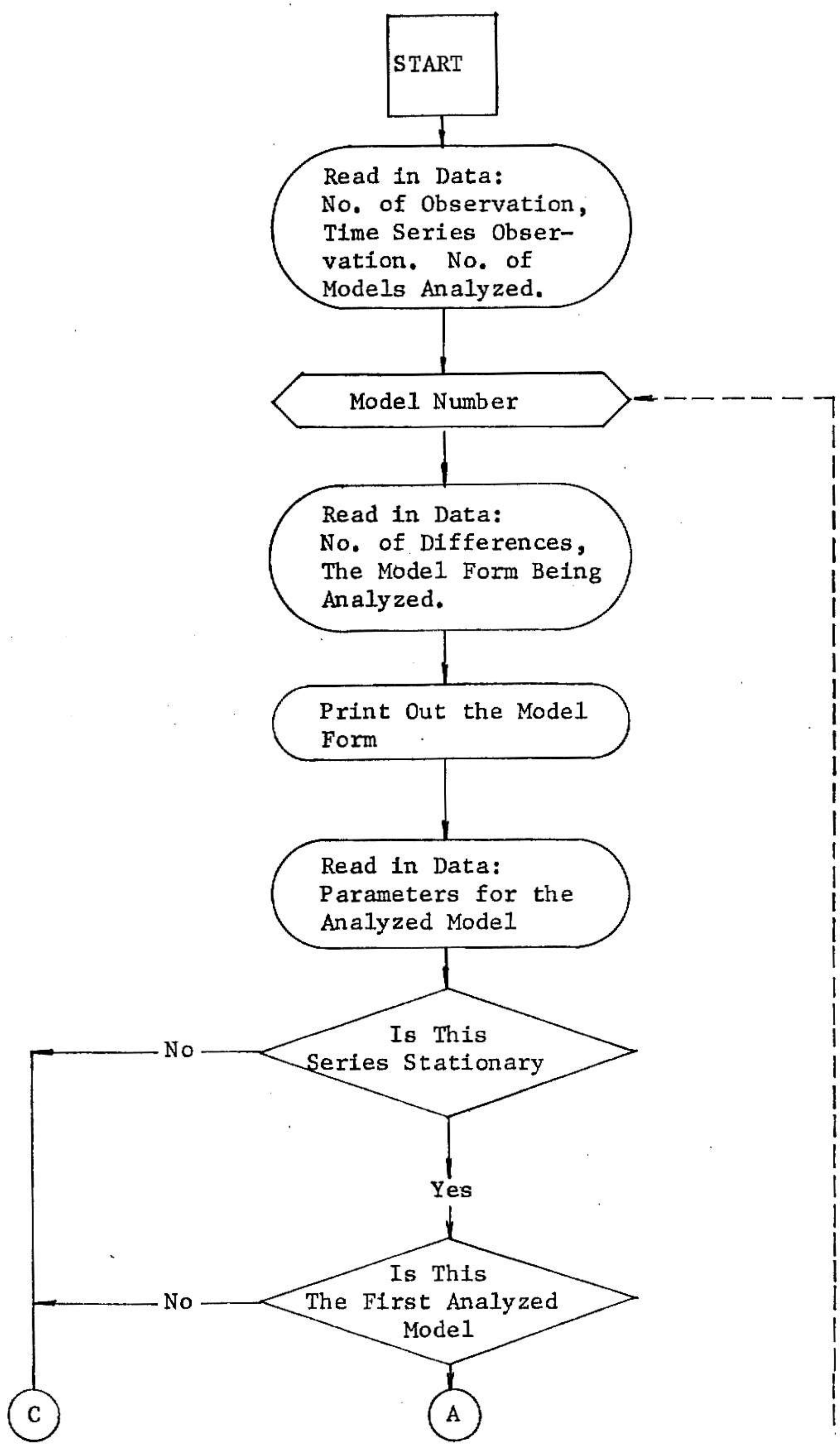
Program ESTIM is developed to determine the least square estimates of the parameters of the models entertained as candidates for acceptance. The program consists of the main program, and five subroutines; they are subroutine MODEL, subroutine CALA, subroutine DIFFER, subroutine MULT and subroutine UWHAUS. The main program provides the data needed for subroutine UWHAUS and calculates the autocorrelation function of the residuals based on the least square estimates. Program UWHAUS is used in conjunction with subroutine MODEL, CALA and MULT, and perform the operation of locating the least square estimates in an iterative manner. The complete description of subroutine UWHAUS is presented on Section B.2. Subroutine DIFFER performs the required differences on the original observations if seasonal or nonstationary time series is analyzed. Subroutine MODEL, CALA and MULT calculate the residuals of original time series required by subroutine UWHAUS for each set of parameter tested. Its calculation is based on the following equation.

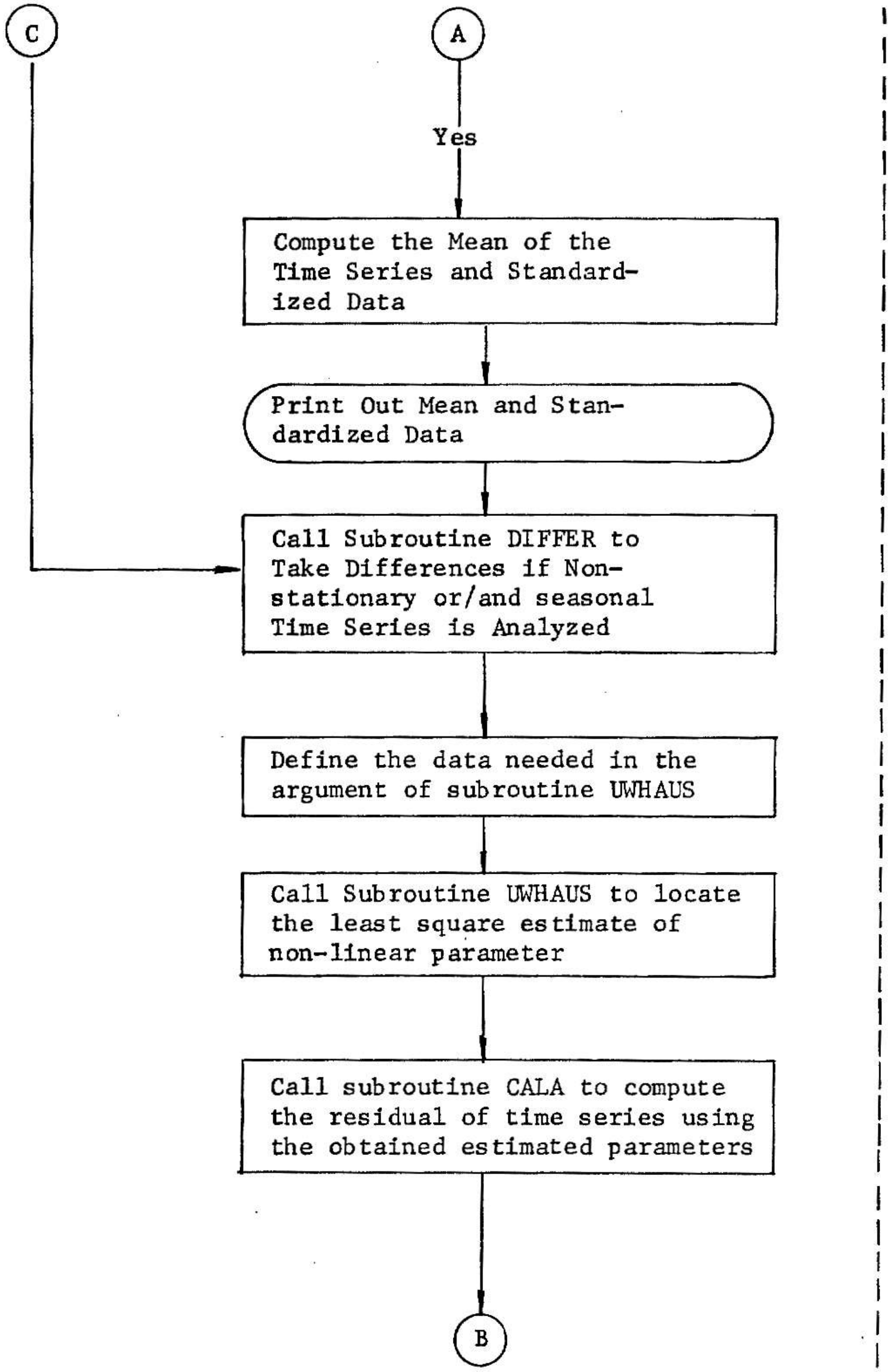
$$a_t = \tilde{w}_t - \phi_1 \tilde{w}_{t-1} - \phi_2 \tilde{w}_{t-2} - \dots - \phi_p w_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2}$$

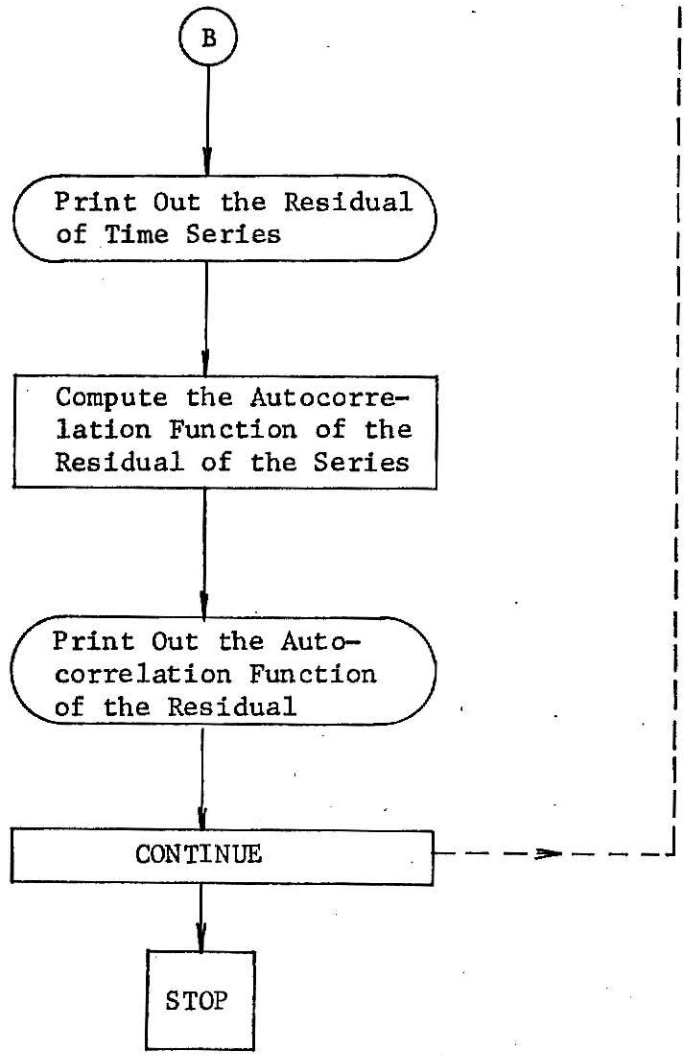
$$+ \dots + \theta_q a_{t-q} \qquad \text{(B.1.1)}$$

where $W_t = \nabla^d Z_t$ and $\tilde{W}_t = W_t - \mu$ with $E[W_t] = \mu$. a_t is the residual of general autoregressive model and is assumed as white noise process.

The flow chart of program ESTIM is as follows.




C
A
Yes
Compute the Mean of the Time Series and Standardized Data
Print Out Mean and Standardized Data
Call Subroutine DIFFER to Take Differences if Non-stationary or/and seasonal Time Series is Analyzed
Define the data needed in the argument of subroutine UWHAUS
Call Subroutine UWHAUS to locate the least square estimate of non-linear parameter
Call subroutine CALA to compute the residual of time series using the obtained estimated parameters
B


B
Print Out the Residual of Time Series
Compute the Autocorrelation Function of the Residual of the Series
Print Out the Autocorrelation Function of the Residual
CONTINUE
STOP

B.2. Description of Input Data of Program ESTIM

CARD	FORMAT	VARIABLES IN PROGRAM	DESCRIPTION
1	(I10)	N	Sample size of original series.
Next N cards	(2X,(F20.5))	ZD(I,1,1)	Observations of original series.
Next card	(I10)	INDS	Number of models to be fitted to original series.
[For each model, add the following parameter input data]			
Next card	(3I10)	IPW	Number of autoregressive parameters.
		IDW	Number of nonstationary differences.
		IQW	Number of moving-average parameters.
Next card	(4I10)	IPB	The largest power of the shift operator associated with the seasonal autoregressive parameters.
		IDB	The number of seasonal differences.
		IQB	The largest power of the shift operator, associated with the seasonal moving-average parameters.
		IS	The seasonal lag.

CARD	FORMAT	VARIABLES IN PROGRAM	DESCRIPTION
Next card	(4I10)	I1	The number of auto-regressive parameters.
		I2	The number of moving-average parameters.
		I3	The number of non-zero seasonal auto-regressive parameters.
		I4	The number of non-zero seasonal moving-average parameters.
Next card	(I10,F20.5)	J,PHIW(J)	The initial estimated value of the Jth autoregressive parameter. If I1=0, no card is entered.
	(I10,F20.5)	J,THETW(J)	The initial value of the Jth moving-average parameter. If I2=0, no card is entered.

[Add the following cards, for seasonal models, only]

CARD	FORMAT	VARIABLES IN PROGRAM	DESCRIPTION
Next card	(I10,F20.5)	J,PHIB(J)	The Jth non-zero seasonal auto-regressive parameters with index J, corresponding to the power of the adjacent shift operator in the model.
Next card	(I10,F20.5)	J,THETB(J)	The Jth non-zero seasonal moving-average parameter with index J corresponding to the power of the adjacent shift operator in the model.

B.3. Description of Output Data

The following data can be generated by Program ESTIM.

1. List of the initial estimates of the parameters.
2. For stationary time series only; a print out of the standardized observations, in other words, the observations formed by subtracting the sample mean from the original observations.
3. UWHAUS prints out summary information at each iteration which can be analyzed to determine the path in the parameter space taken by the iterations to converge on the least square estimates.
4. Tabulation of the first 25 sample correlations of the residuals based on the least square estimates.

B.4.1. Description of Subroutine UWHAUS

Subroutine UWHAUS is developed to obtain least square estimates of parameters entering non-linearly into a mathematical model. An iterative technique is used, the estimate at each iteration is obtained by a method due to Marquardt which combines the Gauss (Taylor series) method and the method of steepest descent [51]. The main program must be provided by the user to supply the input for subroutine UWHAUS. Subroutine MODEL is to specify what mathematical model is to be used. As for the general Auto-regressive Moving-average model, its model form is equation (B.1.1).

The theory behind subroutine UWHAUS can be described as follows. Suppose the mathematical model which is tentatively entertaining is

$$\eta = f(\underline{\theta},\underline{\xi}) \qquad (B.2.1)$$

where $\underline{\theta}$ is a pxl vector of unknown parameters and $\underline{\xi}$ is a vector of independent variables.

Suppose n actual observations Y are made. When the uth observation Y_u is made, the value of the independent variable is $\underline{\xi}_4$. Because of experimental error, an observation Y_u, different from the model response, η_u. Hence,

$$Y_u = \eta_u + \varepsilon_u, \qquad u = 1, \ldots, n \tag{B.2.2}$$

Marquardt also assured the theoretical optimum properties of least square estimates. Some assumption on the errors, ε_u, have to be made: [50]

(1) The errors, ε, are independent random variables with equal variance from the same probability observation (independence implies that knowledge of ε_i does not give any information about ε_j, $i \neq j$).

(2) The expected value of the errors is zero.

(3) The probability distribution of the error is the normal (Gaussian) distribution with variance σ^2.

Under the assumption (1), (2), and (3), the least square estimate, $\hat{\theta}$, is a maximum likelihood estimates and thus had certain desirable properties [51].

Now, from (B.2.2), it is desired to use the observed data to obtain estimates of the unknown parameter, $\underline{\theta}$. An estimate of $\underline{\theta}$, say $\hat{\underline{\theta}}$, obtained by minimizing

$$S(\underline{\theta}) = \sum_{u=1}^{n} [Y_u - \eta_u]^2 = \sum_{u=1}^{n} [Y_u - f(\underline{\theta}, \underline{\xi}_u)]^2 \tag{B.2.3}$$

as a function of $\underline{\theta}$, is frequently referred to as a least square estimate. Subroutine UWHAUS is intended to provide a least square estimate of $\underline{\theta}$ when the model (B.2.1) is nonlinear in the parameters, $\underline{\theta}$.

When dealing with the autoregressive moving-average problem, η_u in (B.2.2) is the random error a_t in model (B.1.1). The expected value of the errors, which are Y_u in (B.2.2), are assumed zero; this comforms with the assumption of Box and Jenkins [52] and Marquardt [50]. Hence, in the case of an autoregressive moving-average model, (B.2.3) may be written as,

$$S(\underline{\theta}) = \sum_{u=1}^{n} [Y_u - \eta_u]^2 = \sum_{u=1}^{n} a_t^2 \tag{B.2.4}$$

Now, suppose $\underline{\theta}^{(0)}$ is an initial guess, the first order Taylor series expansion about $\underline{\theta}^{(0)}$ is

$$\eta_u(\underline{\theta}) \simeq \eta_u(\underline{\theta}^{(0)}) + \sum_{i=1}^{p} (\theta_i - \theta_i^{(0)}) \left.\frac{\partial f(\theta, \xi_u)}{\partial \theta_i}\right|_{\underline{\theta}^{(0)}} \tag{B.2.5}$$

$u=1,\ldots n$

or more compactly,

$$\underline{\eta}(\underline{\theta}) \simeq \underline{\eta}^{(0)} + X\,\underline{\delta}$$

where X is the nxp matrix

$$X_{nxp} = \left\{ \left.\frac{\partial f(\underline{\theta}, \underline{\xi}_u)}{\partial \theta_i}\right|_{\underline{\theta}^{(0)}} \right\}, \qquad \begin{matrix} u=1,\ldots n \\ i=1,\ldots p \end{matrix}$$

where $\underline{\delta} = \underline{\theta} - \underline{\theta}^{(0)}$ is the pxl vector; $\eta(\theta)$ is the nxl vector $[f(\underline{\theta},\underline{\xi}_1), \ldots, f(\underline{\theta},\underline{\xi}_n)]$, and $\underline{\eta}^{(0)}$ is the nxl vector $\underline{\eta}(\theta^{(0)})$.

Now the approximation on the right hand side of (B.2.6) is linear in the parameters $\underline{\delta}$; by substituting (B.2.6) to (B.2.3), an approximation for $S(\underline{\theta})$ is,

$$S(\underline{\theta}) \simeq (\underline{y} - \underline{\eta}^{(0)} - X\,\delta_m)'\,(\underline{y} - \eta^{(0)} - X\,\underline{\delta}_m) \tag{B.2.7}$$

where

$$\underline{\delta}_m = D^{-1/2}\,(D^{-1/2}\,X'\,X\,D^{-1/2} + \lambda I)^{-1}\,D^{-1/2}\,X'\,\underline{Y} \tag{B.2.8}$$

is the correction vector, which is adapted from Marquardt's algorithm; [50]. D is a pxp diagonal matrix whose i-th diagonal element is the same as that of X'X; λ is a non-negative number.

λ should be decreased only if the progress is satisfactory, i.e., only if the sum of squares, $S(\underline{\theta})$, at the new estimate is smaller than at the old. Thus, at i-th iteration, the basic strategy as indicated by Marquardt is as follows [50]:

Denote by $S(\lambda)$ the value of $S(\underline{\theta})$ obtained by using λ in (B.2.8) to get $\underline{\theta}^{(i)}$ from $\underline{\theta}^{(i-1)}$. Let $\lambda^{(i-1)}$ be the value of λ from the previous iteration. Let $\nu > 1$.

Compute $S(\lambda^{(i-1)})$ and $S(\lambda^{(i-1)}/\nu)$.

(1) if $S(\lambda^{(i-1)}/\nu) \leq S(\underline{\theta}^{(i-1)})$, let $\lambda^{(i)} = \lambda^{(i-1)}/\nu$

(2) if $S(\lambda^{(i-1)}/\nu) > S(\underline{\theta}^{(i-1)})$, and $S(\lambda^{(i-1)}) \leq S(\underline{\theta}^{(i-1)})$, let $\lambda^{(i)}=\lambda^{(i-1)}$

(3) otherwise, increase λ by successive multiplication by ν until for smallest w , $S(\lambda^{(i-1)}\ \nu^{w}) \leq S(\underline{\theta}^{(i-1)})$. Let $\lambda^{(i)} = \lambda^{(i-1)}\ \nu^{w}$

Hence, by the definition of $\underline{\delta}$ in (B.2.8), the new guess is $\underline{\theta}^{(1)} = \hat{\underline{\delta}} + \underline{\theta}^{(0)}$, and the next iteration can be started by expanding about $\underline{\theta}^{(1)}$.

B.4.2. Description of the variable in the Argument of Subroutine UWHAUS

UWHAUS is called from the main program with a FORTRAN statement of the form:

```
CALL UWHAUS (NPROB, NOB, Y, NP, TH, DIFF, SIGNS, EPS1, EPS2, MIT, FLAM,
             FNU, SCRAT)
```

NPROB — is the problem number.

NOB — is the number of observations.

Y — is a real one-dimensional array containing the vector of observed function values; i.e., Y(I) is the Ith observed function value, I=1,..., NOB.

NP — is an integer indicating the number of unknown parameters.

TH — is a real one-dimensional array of the parameter values. i.e. TH(J) is the Jth parameter value, J = 1, ..., NP. It is very important to obtain reasonable starting guess for the parameters; not only will the computation time be decreased by a good choice of starting values, but there is also the possibility of converging to a more reasonable estimate.

DIFF — is a real one-dimensional array containing a vector of proportions in $\underline{\theta}$, for use in computing the difference quotients of the model function values. The devivatives,

$\frac{\partial f(\underline{\theta},\underline{\xi}_u)}{\partial \theta_i}$ in (B.2.5) are approximated by difference quotients within the program.

$$\frac{\partial f(\underline{\theta},\underline{\xi}_u)}{\partial \theta_i} \simeq \frac{f(\theta_1, \ldots, \theta_i + \Delta\theta_i, \ldots \theta_p, \underline{\xi}_u) - f(\underline{\theta},\underline{\xi}_u)}{(\theta_i + \Delta\theta_i 0) - \theta_i}$$

Thus at any point in the calculations, the denominator of the above difference quotient will be expressed as:

(TH(I)+DIFF(I)*TH(I))-TH(I)= DIFF(I)*TH(I)

In any case, DIFF(I) must satisfy $0 < |\text{DIFF(I)}| < 1$, (I=1,..., NP). Using a starting guess of zero for any parameter is prohibited for this method of calculation.

SIGNS is a real one-dimensional array indicating the existence of a prior sign restrictions on each of the parameters. If SIGNS(I) is set equal to any positive quantity, UWHAUS will not allow the Ith parameter to change its sign during the calculations, thus the Ith parameter, TH(I), retains the same sign as the starting guess for that parameter. If SIGNS(I) = 0, this feature is disabled for the Ith parameter.

EPS1 is a real constant indicating the sum of squares convergence criterion and is used to terminate the calculation based on the relative change in the sum of squares from one iteration to the next iteration. More precisely, if at the completion of the ith iteration, it is true that

$$\left| \frac{S(\underline{\theta}^{(i)}) - S(\underline{\theta}^{(i-1)})}{S(\underline{\theta}^{(i-1)})} \right| \leq \text{EPS1}$$

then the calculations are terminated. Roughly, this means that if EPS1=10^{-k}, the calculations will be stopped if the sum of squares for the (i-1)st and ith iteration agree to k decimal places. If EPS1 is set equal to zero, this feature is disabled.

EPS2 is a real constant which is the parameter convergence criterion and is used to terminate the calculations based on the relative change in the parameter values from one iteration to the next iteration. Suppose that after the ith iteration, the value of the jth parameter is $\theta_j^{(i)}$ (j=1,..., p). If, at the completion of the ith iteration, the following holds:

$$\left| \frac{\theta_j^{(i)} - \theta_j^{(i-1)}}{\theta_j^{(i-1)}} \right| < \text{EPS2}$$

for all j=1,..., p, then the calculations are terminated. Roughly, this means that if EPS2 = 10^{-k}, the calculations will be stopped if the value of each parameter after the ith iteration agree to k decimal place with the value of the same parameter after the (i-1)st iteration. This feature is disabled if EPS2 is set to zero.

MIT is an integer constant (where 0 < MIT < 1000) which is the maximum number of iterations to be performed. If the calculations have not been terminated for some other reasons, they will be terminated when the number of iteration equals MIT.

FLAM starting value for λ .

FNU is the value of ν .

SCRAT is an optimal parameters used to specify temporary storage for use by UWHAUS. When present in the calling sequence, SCRAT must be the name of an array containing at least the number of storage locations given by:

$5*NP+2*NP^2 + 2*NOB+NP*NOB$

The contents of these locations will be destroyed during execution of UWHAUS.

B.4.3. The restrictions of subroutine UWHAUS

At the beginning of each problem run, UWHAUS checks the input arguments to see that the following are obeyed:

(1) $1 \leq NP \leq 50$;

(2) $NOB \geq NP$

(3) $TH(I) \neq 0$, I=1,..., NP. Each starting parameter guess is non-zero.

(4) $0 < |DIFF(I)| < 1$, I=1,..., NP. Each difference proportion is between 0 and 1 in absolute value.

(5) $0 < MIT < 1000$, the maximum number of iteration is between 0 and 1000.

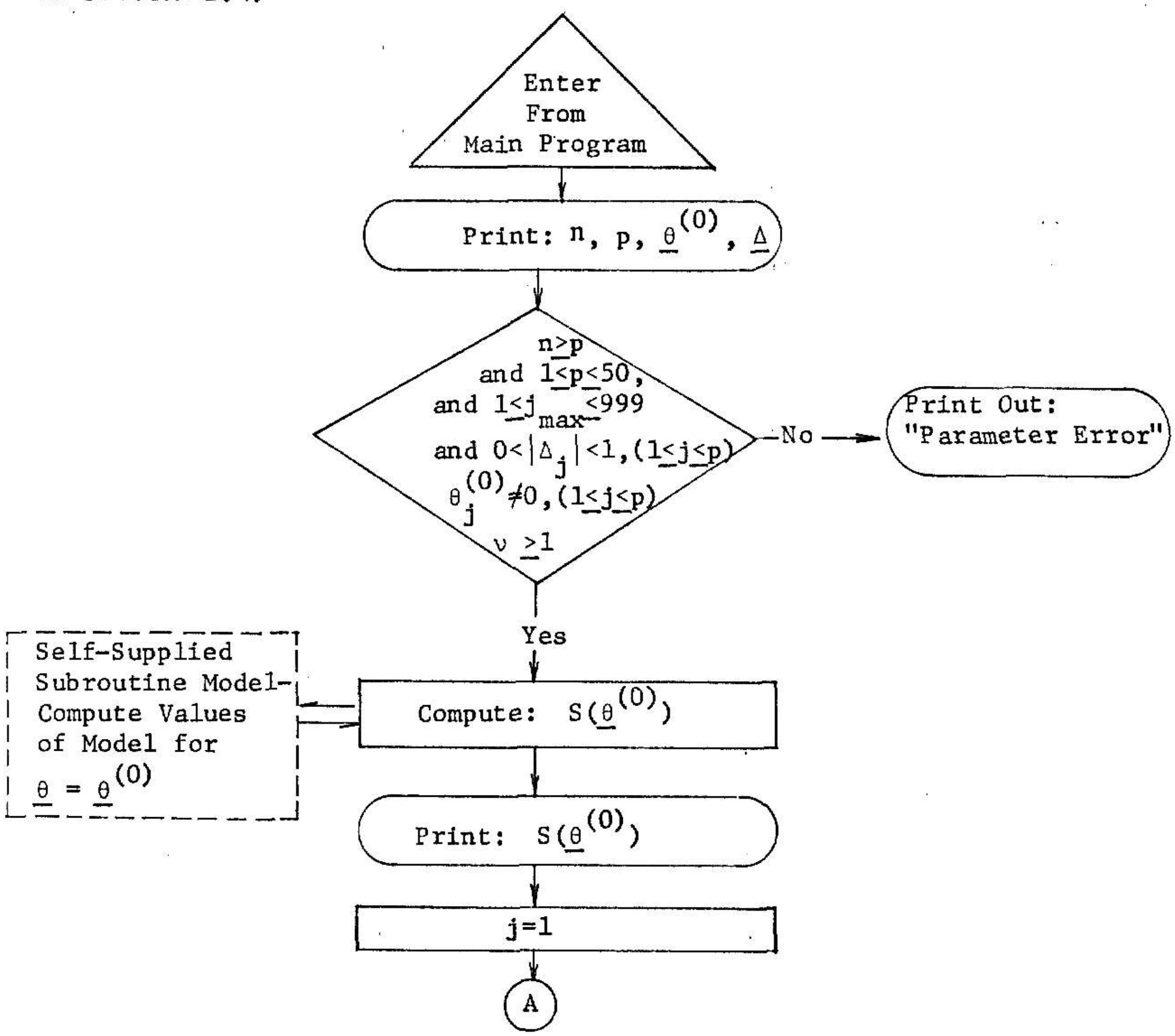
(6) $FNU \geq 1$. The starting value of ν is greater than or equal to 1.

If any of these restrictions are not obeyed, the message:

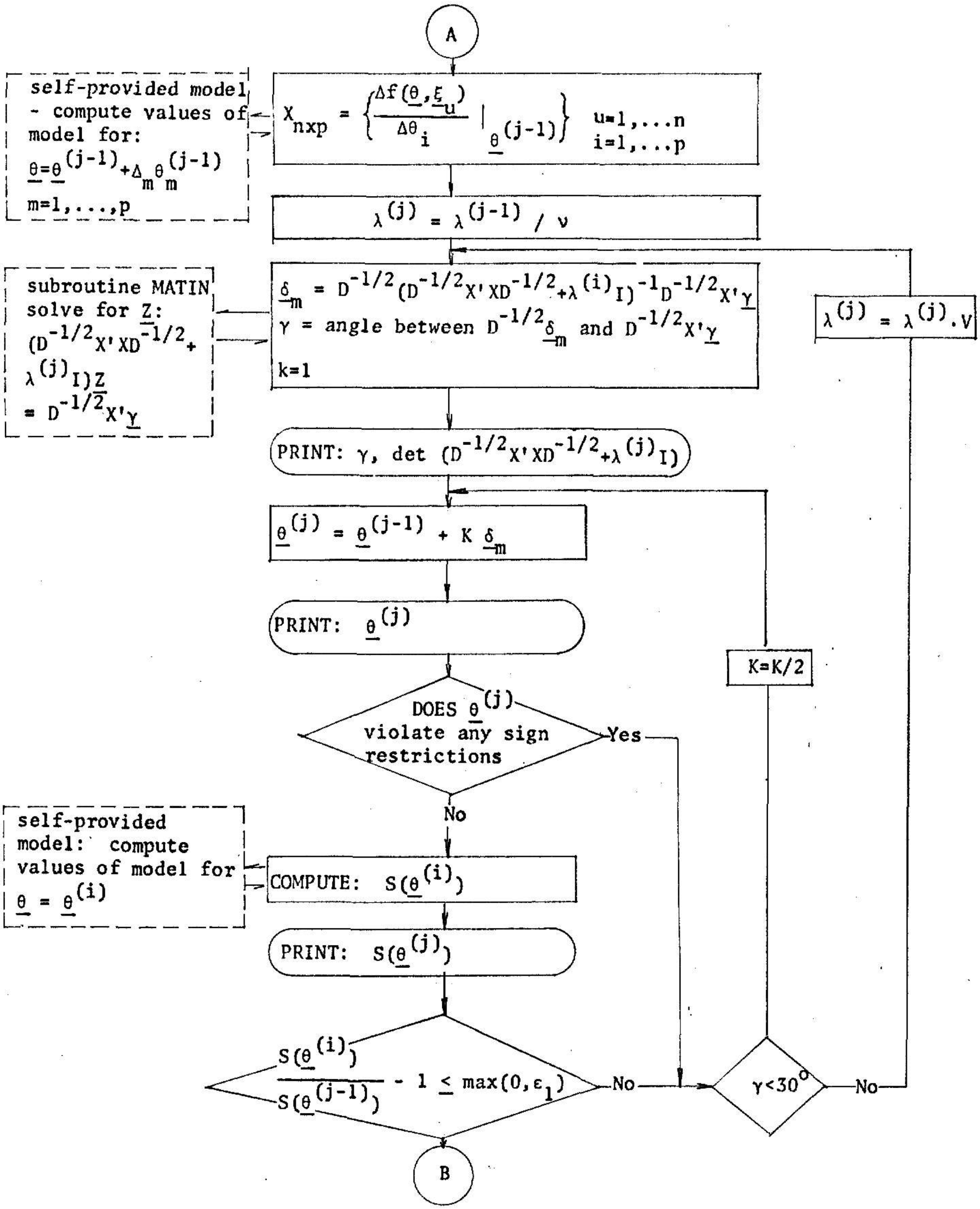
PARAMETER ERROR

will be printed on the printer output for the job, and control will be returned to the main program.

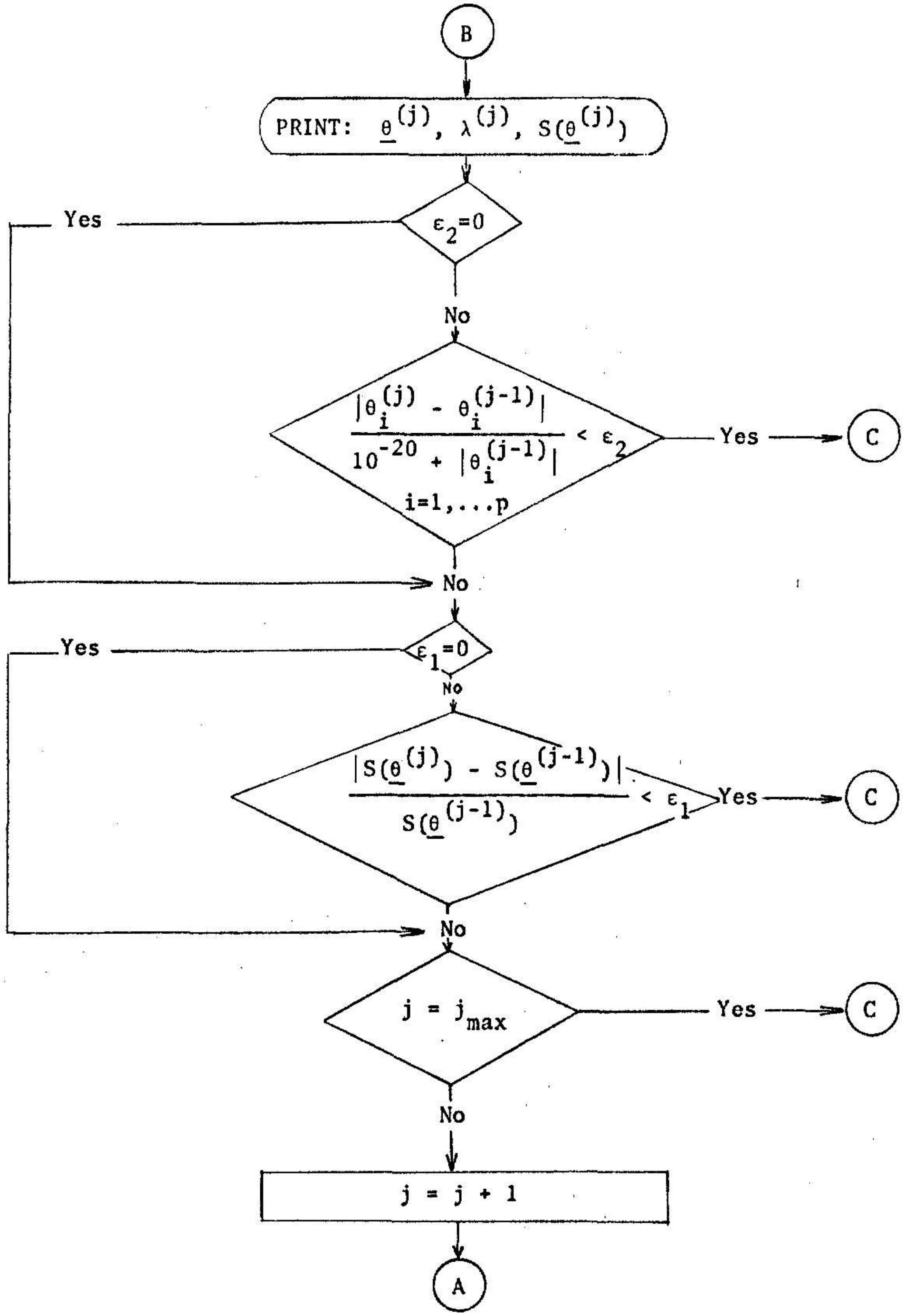
The flow chart of subroutine UWHAUS is shown as follows. The notations ε_1, ε_2, $\underline{\Delta}$ and j_{max} are the input argument EPS1, EPS2, DIFF and MIT, respectively. Other notations have the same meanings as define in Section B.4.



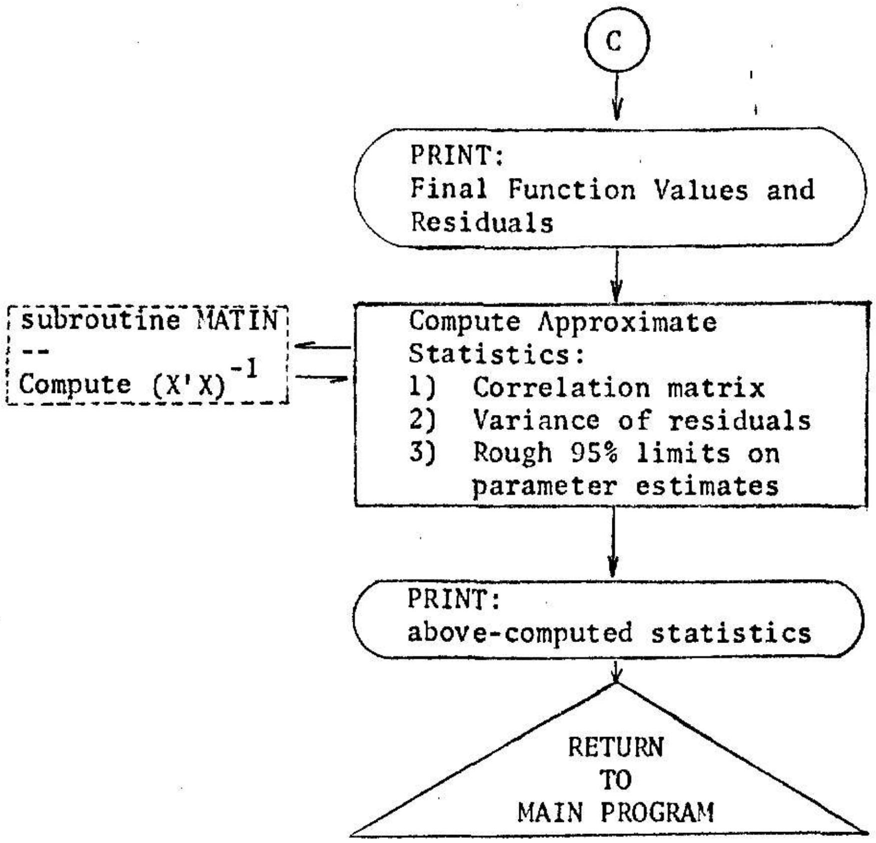
Initialization



Compute Parameter Estimation



Complete One Iteration



Iterating Complete

APPENDIX B. PROGRAM ESTIM

B.5 Computer Program

```
              C      PROGRAM ESTIM IS DEVELOPED TO DETERMINE THE LEAST SQUARES ESTIMES
              C      OF THE PARAMETERS OF MODELS ENTERMINED AS CANDIDATEA FOR ACCEPTANCE.
              C      THE PROGRAM ALSO DEVELOP THE APPROXIMATE COVARIANCE MATRIX FOR THE
              C      ESTIMATES, AND PROVIDES THE RESIDUALS BASED ON THE LEAST SQUARES ESTIMATES
              C
 0001                DIMENSION SCRAT(4018)
 0002                DIMENSION  R(100),C(100),S(100),VAR(100)
 0003                DIMENSION B(20)
 0004                DIMENSION ZERO(1000)
 0005                DIMENSION DIFF(20),SIGNS(20)
 0006                COMMON ZD(1000,4,3),NOB(4,3),Z(1000),A(1000), PHI(100),THETA(100),
                    1          PHIW(100),THETW(100),PHIB(100),THETB(100),LOC1(10),LOC2(
                    210),LOC3(10),LOC4(10),II1(1),II2(1),II3(1),II4(1),IIP(1),IIQ(1)
 0007                KCREP=0
              C       N   SAMPLE SIZE OF ORIGINAL SERIES
 0008                READ 2002, N
 0009           2021 FORMAT(2X,F20.5)
 0010                READ 2021, (ZD(I,1,1),I=1,N)
              C      READ IN DATA AND PRINT OUT
              C      ZD(I,1,1)     OBSERVATIONS OF ORIGINAL SERIES
              C      INDS     NO. OF MODELS TO BE FITTED TO ORIGINAL SERIES
              C      IPW     NO. OF AUTOREGRESSIVE PARAMETERS
              C      IDW     NO. OF NON-STATIONARY DIFFERENCES.
              C              THIS IS ZERO FOR STATIONARY SERIES.
              C      IQW     NO. OF MOVING AVERAGE PARAMETERS.
 0011                READ 2002, INDS
 0012                DO 10 IDOP=1,INDS
 0013           2002 FORMAT(4I10)
 0014                READ 2002, IPW,IDW,IQW
              C      IPB     THE LARGEST POWER OF THE SHIFT OPERATOR ASSOCIATED WITH
              C                      THE SEASONAL AUTOREGRESSIVE PARAMETERS.
              C      IDB     NO. OF SEASONAL DIFFERENCES
              C      IQB     LARGEST POWER OF THE SHIFT OPERATOR ASSOCIATED WITH THE
              C              SEASONAL MOVING AVERAGE PARAMETERS.
              C      IS     SEASONAL LAG.
              C      FOR STATIONARY SERIES, THE FOLLOWING CARD IS BLANK
 0015                READ 2002, IPB,IDB,IQB,IS
 0016           2003 FORMAT(1H12X,6HIARMA(,I3,1H,,I3,1H,,I3,3H)X(,I3,1H,,I3,1H,,I3,3H)X
                    1(,I3,1H)//)
 0017                PRINT 2003, IPW,IDW,IQW,IPB,IDB,IQB,IS
 0018                IP=IPW+IPB
 0019                IQ=IQW+IQB
 0020                IIP(1)=IP
 0021                IIQ(1)=IQ
 0022                IPIQ=IP+IQ
 0023                DO 2090 I=1,IPIQ
 0024                PHIW(I)=0.
 0025                PHIB(I)=0.
 0026                PHI(I)=0.
 0027                THETW(I)=0.
 0028                THETB(I)=0.
 0029           2090 THETA(I)=0.
 0030           100  FORMAT(///2X,20HINITIAL GUESS VALUES///)
 0031                PRINT 100
 0032           2004 FORMAT(I10,F20.5)
 0033           2005 FORMAT(2X,5HPHIW(,I3,2H)=,F11.5)
 0034           2007 FORMAT(2X,5HPHIB(,I3,2H)=,F11.5)
 0035           2006 FORMAT(2X,6HTHETW(,I3,2H)=,F11.5)
```

```
0036       2008 FORMAT(2X,6HTHETB(,I3,2H)=,F11.5)
          C      I1     THE NO. OF AUTOREGRESSIVE PARAMETERS
          C      I2     THE NO. OF MOVING AVERAGE PARAMETERS.
          C      I3     THE NO. OF NON-ZERO SEASONAL AUTOREGRESSIVE PARAMETERS.
          C      I4     THE NO. OF NON-ZERO SEASONAL MOVING AVERAGE PARAMETERS.
0037            READ 2002, I1,I2,I3,I4
0038            K=0
0039            IF(I1)2009,2009,2010
          C      PHIW(J)      THE INITIAL GUESS VALUE OF THE JTH AUTOREGRESSIVE PARAMETERS.
          C                   LOC1(I) ASSOCIATED THESE J.
          C      THTEW(J)      THE INITIAL GUESS VALUE OF THE JTH MOVING AVERAGE PARAMETERS.
          C                    LOC2(I) ASSOCIATED THESE J.
          C      PHIB(J)      THE JTH NON-ZERO SEASONAL AUTOREGRESSIVE PARAMETER,
          C                   CORRESPONDING TO THE POWER OF THE ADJACENT SHIFT OPERATOR
          C                   IN THE MODEL.
          C                   LOC3(I) ASSOCIATED THOSE J.
          C      THETB(J)     THE JTH NON-ZERO SEASONAL MOVING AVERAGE PARAMETER.
          C                   LOC4(I) ASSOCIATED THESE J.
0040       2010 DO 2011 I=1,I1
0041            READ 2004, J,PHIW(J)
0042            LOC1(I)=J
0043            K=K+1
0044            B(K)=PHIW(J)
0045       2011 PRINT 2005, J,PHIW(J)
0046       2009 IF(I2) 2012,2012,2013
0047       2013 DO 2014 I=1,I2
0048            READ 2004, J,THETW(J)
0049            LOC2(I)=J
0050            K=K+1
0051            B(K)=THETW(J)
0052       2014 PRINT 2006,J,THETW(J)
0053       2012 IF(I3) 2015,2015,2016
0054       2016 DO 2017 I=1,I3
0055            READ 2004, J,PHIB(J)
0056            LOC3(I)=J
0057            K=K+1
0058            B(K)=PHIB(J)
0059       2017 PRINT 2007, J,PHIB(J)
0060       2015 IF(I4)2018,2018,2019
0061       2019 DO 2020 I=1,I4
0062            READ 2004, J,THETB(J)
0063            LOC4(I)=J
0064            K=K+1
0065            B(K)=THETB(J)
0066       2020 PRINT 2008, J,THETB(J)
0067       2018 ISUM=I1+I2+I3+I4
0068            II1(1)=I1
0069            II2(1)=I2
0070            II3(1)=I3
0071            II4(1)=I4
0072            IF(IDW)9000,9000,9001
0073       9000 CONTINUE
          C      IF THIS IS AFTER FIRST MODEL, JUST GO TO 9001, NOT NECESSARY
          C      TO COMPUTE AGAIN.
          C
0074            KCREP=KCREP+1
0075            IF(KCREP-1)6721,6721,9001
0076       6721 CONTINUE
```

```
0077              XN=N
0078              ZBAR=0.
0079              DO 9002 I=1,N
0080        9002  ZBAR=ZBAR+ZD(I,1,1)
0081              ZBAR=ZBAR/XN
0082              DO 9003 I=1,N
0083        9003  ZD(I,1,1)=ZD(I,1,1)-ZBAR
0084        9004  FORMAT(1H12X,17HSTANDARDIZED DATA///)
0085              PRINT 9004
0086        9005  FORMAT(2X,5HZBAR=,F20.5///)
0087              PRINT 9005,ZBAR
0088              PRINT 205, (ZD(I,1,1),I=1,N)
0089        9001  CONTINUE
0090              CALL DIFFER(IDW,IDB,IS,N,MW,MB)
           C        NOBB IS NO. OF OBSERVATION AFTER DIFFERENCES AND SEASONAL DIFFERENCES.
0091              NOBB=NOB(MW,MB)
0092              DO 2022 I=1,NOBB
0093              ZERO(I)=0.
0094        2022  Z(I)=ZD(I,MW,MB)
           C      GAUSSHAUS DATA
0095              DO 11 I=1,ISUM
0096              DIFF(I)=.01
0097         11   SIGNS(I)=0.
0098              EPS1=0.
0099              EPS2=.00001
0100              MIT=18
0101              FLAM=50.
0102              FNU=10.
0103              NPROB=IDCP
0104              CALL GWHAUS(NPROB,NOBB,ZERO,ISUM,B,DIFF,SIGNS,EPS1,EPS2,MIT,FLAM,
                 1FNU,SCRAT)
           C      PUNCH OUT OF A FOR DIAGNOSTIC CHECKING
0105              CALL CALA(NOBB,IP,IQ)
0106              PRINT 2002, NOBB
0107         205  FORMAT(2X,6F20.5)
0108              PRINT 205, (A(I),I=1,NOBB)
0109        7001  FORMAT(1H12X,19HDIAGNOSTIC CHECKING///)
0110              PRINT 7001
0111              DO 7000 I=1,NOBB
0112              ZERO(I)=Z(I)-A(I)
0113        7000  CONTINUE
0114              KK=25
0115              N=NOBB
0116              XN=N
0117              ZBAR=0.
0118              DO 102 I=1,N
0119              Z(I)=A(I)
0120         102  ZBAR=ZBAR+Z(I)
0121              ZBAR=ZBAR/XN
0122              C0=0.
0123              DO 103 I=1,N
0124         103  C0=C0+(Z(I)-ZBAR)**2
0125              C0=C0/XN
           C      CALCULATION OF R
0126              DO 104 K=1,KK
0127              C(K)=0.
0128              NN=N-K
0129              DO 105 J=1,NN
```

```
FORTRAN IV G LEVEL  18                MAIN              CATE = 71113           03/39/40

           105 C(K)=C(K)+(Z(J)-ZBAR)*(Z(J+K)-ZBAR)
               C(K)=C(K)/XN
           104 R(K)=C(K)/CO
               VAR(1)=1./XN
               S(1)=R(1)/SCRT(VAR(1))
               AAA=2./XN
               CO 204 K=2,KK
               VAR(K)=VAR(K-1)+AAA*(R(K-1)**2)
           204 S(K)=R(K)/SCRT(VAR(K))
           300 FCRMAT(1H12X,44HSAMPLE CORRELATION COEFFICIENTS OF RESIDUALS//)
               PRINT 300
           305 FORMAT(//2X,5HZBAR=,F10.5//)
               PRINT 305, ZBAR
           301 FCRMAT(2X,2HR(,I3,2H)=,F10.5,6X,10HR/S.D.(R)=,F10.5)
               CO 302 I=1,KK
           302 PRINT 301, I,R(I),S(I)
           10  CCNTINUE
               STOP
               END
```

```
FORTRAN IV G LEVEL  18                  MODEL               DATE = 71113          03/39/40

 0001               SUBROUTINE MODEL(NPROB,B,F,NOBB,ISUM)
           C        SUBROUTINE EST CHANGE B(K) TO DIFFERENT PARAMETERS, WHICH IS THE SAME AS
           C        READ IN DATA VECTOR PARAMETER.      THEN, COMPUTE THE VALUE OF
           C         THE MODEL.
           C
 0002               DIMENSION B(1)
 0003               DIMENSION F(1)
 0004               COMMON ZD(1000,4,3),NOB(4,3),Z(1000),A(1000), PHI(100),THETA(100),
                   1         PHIW(100),THETW(100),PHIB(100),THETB(100),LOC1(10),LOC2(
                   210),LOC3(10),LOC4(10),II1(1),II2(1),II3(1),II4(1),IIP(1),IIQ(1)
 0005               IP=IIP(1)
 0006               IQ=IIQ(1)
 0007               IPIQ=IP+IQ
 0008               I1=II1(1)
 0009               I2=II2(1)
 0010               I3=II3(1)
 0011               I4=II4(1)
 0012               K=0
 0013               IF(I1)100,100,101
 0014           101 DO 102 I=1,I1
 0015               J=LOC1(I)
 0016               K=K+1
 0017           102 PHIW(J)=B(K)
 0018           100 IF(I2)103,103,104
 0019           104 DO 105 I=1,I2
 0020               J=LOC2(I)
 0021               K=K+1
 0022           105 THETW(J)=B(K)
 0023           103 IF(I3)106,106,107
 0024           107 DO 108 I=1,I3
 0025               J=LOC3(I)
 0026               K=K+1
 0027           108 PHIB(J)=B(K)
 0028           106 IF(I4)109,109,110
 0029           110 DO 111 I=1,I4
 0030               J=LOC4(I)
 0031               K=K+1
 0032           111 THETB(J)=B(K)
 0033           109 CALL MULTS  (PHIW,PHIB,PHI,IP)
           C         PHI(I) IS THE MINUS VALUE OF THE AUTOREGRESSIVE PARAMETERS
           C         BY MULTIPLING NON-SEASONAL AND SEASONAL PARAMETERS TOGETHER.
 0034               CALL MULTS  (THETW,THETB,THETA,IQ)
           C        THETA(I) IS THE MINUS VALUE OF THE MOVING AVERAGE PARAMETERS AFTER
           C        MULTIPLYING NON-SEASONAL AND SEASONAL PARAMETER.
 0035               CALL CALA(NOBB,IP,IQ)
 0036               DO 50 I=1,NOBB
 0037            50 F(I)=-A(I)
 0038               RETURN
 0039               END
```

```
0001                SUBROUTINE CALA(N,IPID,IQ)
           C        SUBROUTINE CALA IS TO CALCULATE THE VALUE OF MODEL BY TIME SERIES.
           C
           C      CPHI=PHI
           C
0002                COMMON ZD(1000,4,3),NOB(4,3),Z(1000),A(1000),CPHI(100),THETA(100),
                  1           PHIW(100),THETW(100),PHIB(100),THETB(100),LOC1(10),LOC2(
                  210),LOC3(10),LOC4(10),II1(1),II2(1),II3(1),II4(1),IIP(1),IIQ(1)
           C        FIX INITAL VALUES
0003                IF(IQ-IPID)201,201,200
0004            200 ISTART=IQ
0005                GO TO 203
0006            201 ISTART=IPID
0007            203 DO 204 I=1,ISTART
0008            204 A(I)=0.
0009                ISTART=ISTART+1
0010                DO 206 I=ISTART,N
0011                A(I)=Z(I)
0012                IF(IPID)207,207,208
0013            208 DO 209 J=1,IPID
0014            209 A(I)=A(I)-CPHI(J)*Z(I-J)
0015            207 IF(IQ)  206,206,211
0016            211 DO 212 J=1,IQ
0017            212 A(I)=A(I)+THETA(J)*A(I-J)
0018            206 CONTINUE
0019                RETURN
0020                END
```

```
FORTRAN IV G LEVEL  18                  DIFFER              DATE = 71113          03/39/40

 0001                  SUBROUTINE DIFFER(IDW,IDB,IS,N,MW,MB)
 0002                  COMMON  Z(1000,4,3),NOB(4,3),C(1000),A(1000), PHI(100),THETA(100),
                     1          PHIW(100),THETW(100),PHIB(100),THETB(100),LOC1(10),LOC2(
                     210),LOC3(10),LOC4(10),II1(1),II2(1),II3(1),II4(1),IIP(1),IIQ(1)
                C      CALCULATION OF NOB(M,J)
 0003                  MW=IDW+1
 0004                  MB=IDB+1
 0005                  DO 151 M=1,MW
 0006                  DO 151 J=1,MB
 0007             151  NOB(M,J)=N+1-M+IS-J*IS
                C      DIFFERENCING WITH RESPECT TO DW
 0008                  IF(MW-1)152,152,153
 0009              153 DO 154 M=2,MW
 0010                  NOBD=NOB(M,1)
 0011                  DO 154 I=1,NOBD
 0012              154 Z(I,M,1)=Z(I+1,M-1,1)-Z(I,M-1,1)
                C      DIFFERENCING WITH RESPECT TO DB
 0013              152 IF(MB-1)155,155,156
 0014              156 DO 157 M=1,MW
 0015                  DO 157 J=2,MB
 0016                  NOBD=NOB(M,J)
 0017                  DO 157 I=1,NOBD
 0018              157 Z(I,M,J)=Z(I+IS,M,J-1)-Z(I,M,J-1)
 0019              155 CONTINUE
 0020                  RETURN
 0021                  END
```

```
FORTRAN IV G LEVEL  18                MULTS              DATE = 71113          03/39/40

 0001                   SUBRCUTINE MULTS  (PHIW,PHIB,PHI,IPWPB)
              C         SUBROUTINE MULTS IS TO MULTIPLE THE PARAMETERS OF NON-SEASONAL
              C         ANC SEASCNAL MODELS.
 0002                   CIMENSICN PHIW(100),PHIB(100),PHI(100)
 0003                   IF(IPWPB)101,101,105
 0004               1C5 CO 10C I=1,IPWPB
 0005               1C0 PHI(I)=PHIW(I)+PHIB(I)
 0006                   IF(IPWPB-1)101,101,102
 0007               102 DO 104 I=2,IPWPB
 0008                   JDUM=I-1
 0009                   DO 104 J=1,JDUM
 0010               1C4 PHI(I)=PHI(I)-PHIB(J)*PHIW(I-J)
 0011               101 CONTINUE
 0012                   RETURN
 0013                   END
```

```
FORTRAN IV G LEVEL  18                  UWHAUS             DATE = 71113          03/39/40

 0001            SUBROUTINE UWHAUS(NPROB,       NOB,Y,NP,TH,DIFF,SIGNS,EPS1,EPS2,     UWHA
                1  MIT, FLAM, FNU, SCRAT)                                             UWHA
 0002            DIMENSION SCRAT(1)
 0003            DIMENSION Y(1),TH(1),DIFF(1),SIGNS(1)
 0004            IA=1
 0005            IB=IA+NP                                                             UWHA
 0006            IC=IB+NP                                                             UWHA
 0007            ID=IC+NP                                                             UWHA
 0008            IE=ID+NP                                                             UWHA
 0009            IF=IE+NP
 0010            IG=IF+NOB
 0011            IH=IG+NOB                                                            UWHA
 0012            II = IH + NP * NOB                                                   UWHA
 0013            IJ = IH                                                              UWHA
 0014                CALL  HAUS59(NPROB,       NOB,Y,NP,TH,DIFF,SIGNS,EPS1,EPS2,MITUWHA
                1  ,FLAM,FNU,SCRAT(IA), SCRAT(IB), SCRAT(IC), SCRAT(ID),             UWHA
                2  SCRAT(IE), SCRAT(IF), SCRAT(IG), SCRAT(IH), SCRAT(II),
                3  SCRAT(IJ)  )                                                       UWHA
 0015            RETURN                                                               UWHA
 0016            END                                                                  UWHA
```

```
FORTRAN IV G LEVEL  18                  HAUS59              DATE = 71113          03/39/40

0001             SUBROUTINE HAUS59(NPRBO,         NBC, Y,NQ,TH,DIFZ,SIGNS,EP1S,EP2S,UWHA  20
                1MIT,FLAM,FNU,    C,P,E,PHI,TB,F,R,A,D,DELZ)
            C     PROGRAM UWHAUS IS USED IN CONJUNCTION WITH SUBROUTINES EST, CALA
            C     AND MULTS, AND PERFORMS THE OPERATION OF LOCATING THE LEAST SQUARES
            C     ESTIMATES IN AN ITERATIVE MANNER.
            C     PROGRAM UWHAUS IS BASED ON A METHOD DUE TO MARQUARDT WHICH COMBINES
            C     THE GAUSS (TAYLOR SERIES) METHOD AND THE METHOD OF STEEPEST ASCENT.
            C     UWHAUS PRINTS OUT SUMMARY INFORMATION AT EACH ITERATION WHICH CAN BE
            C     ANALYZED TO DETERMINE THE PATH IN THE PARAMETER SPACE
            C     TAKEN BY THE ITERATION TO CONVERAGE ON THE  LEAST SQUARE ESTIMATES.
            C
            C                  FORTRAN II VERSION                                    UWHA  22
            C                  H. J. WERTZ                                           UWHA  23
            C                  ADAPTED FOR THE UNIVAC 1108  (HJW  12/68)             UWHA  24
            C                                                                        UWHA  25
            C     DIMENSION TH(NQ), DIFZ(NQ), SIGNS(NQ), Y(NBO)                      UWHA  26
            C     DIMENSION C(NQ), P(NQ), E(NQ), PHI(NQ), TB(NQ)                     UWHA  27
            C     DIMENSION F(NBO), R(NBO)                                           UWHA  28
            C     DIMENSION A(NQ,NQ), D(NQ,NQ), DELZ(NBO,NQ)                         UWHA  29
0002              DIMENSION TH(1),DIFZ(1),SIGNS(1),Y(1),C(1),P(1),E(1)                     30
0003              DIMENSION PHI(1),TB(1),F(1),R(1),A(1),D(1),DELZ(1)                       31
0004              DIMENSION CO(2)
0005              ACOS(X) = ATAN(SQRT(1.0/X**2 - 1.0))                               UWHA  32
0006              NP = NQ                                                            UWHA  33
0007              NPROB = NPRBO                                                      UWHA  34
0008              NOB = NBO                                                          UWHA  35
0009              EPS1 = EP1S                                                        UWHA  36
0010              EPS2 = EP2S                                                        UWHA  37
0011              PRINT 1000, NPROB,NOB,NP
0012              PRINT 1001                                                         UWHA  41
0013              CALL GASS60(1,NP,TH,OO)
0014              PRINT 1002                                                         UWHA  43
0015              CALL GASS60(1,NP,DIFZ,OO)
            C     TEST INPUT VALUE
0016              IF(MINO(NP-1,50-NP,NOB-NP,MIT-1,999-MIT))99,15,15                  UWHA  45
0017         15   IF(FNU-1.0)99, 99, 16                                              UWHA  46
0018         16   CONTINUE                                                           UWHA  47
0019              DO 19 I=1,NP                                                       UWHA  48
0020              TEMP = ABS(DIFZ(I))                                                UWHA  49
0021              IF(AMIN1(1.0-TEMP, ABS(TH(I))))99, 99, 19                          UWHA  50
0022         19   CONTINUE                                                           UWHA  51
0023              GA = FLAM                                                          UWHA  52
0024              NIT = 1                                                            UWHA  53
0025              MARY=1                                                                   54
0026              MAY=1                                                                    55
0027              LUCY=1                                                                   56
            C   EPS1 CAN NOT BE NEGATIVE
            C     EPS2  LESS THAN OR EQUAL TO ZERO, EPS1 GREATER THAN ZERO, THEN MAY=3
            C     EPS2 LESS THAN OR EQUAL ZERO, EPS1 LESS THAN OR EQUAL TO ZERO, THEN MAY=2
            C     EPS2 GREATER THAN ZERO, EPS1 GREATER THAN ZERO
            C     EPS2 GREATER THAN ZERO, EPS1 LESS THAN OR EQUAL TO ZERO, THEN LUCY=2
            C     SSQ=SUM OF SQUARE OF RESIDUAL
0028              IF(EPS1) 5, 10, 10                                                 UWHA  57
0029          5   EPS1 = 0                                                           UWHA  58
0030         10   IF(EPS2) 40, 40, 30                                                UWHA  59
0031         40   IF(EPS1) 60, 60, 50                                                UWHA  60
0032         60   MAY=2                                                                    61
0033              GO TO 70                                                           UWHA  62
```

```
0034     50    MAY=3
0035           GO TO 70
0036     30    IF(EPS1) 80, 80, 70
0037     80    LUCY=2
0038      70   SSQ = 0
0039           CALL MODEL(NPROB, TH, F, NOB, NP)
0040           DO 90 I = 1, NOB
0041           R(I) = Y(I) - F(I)
0042      90   SSQ=SSQ+R(I)*R(I)
0043           PRINT 1003, SSQ
         C
         C                                          BEGIN ITERATION
         C
0044     100   GA = GA / FNU
0045           INTCNT = 0
0046           PRINT 1004, NIT
0047     101   JS = 1 - NOB
0048           DO 130 J=1,NP
0049           TEMP = TH(J)
0050           P(J)=DIFZ(J)*TH(J)
0051           TH(J)= TH(J)+P(J)
0052           Q(J)=0
0053           JS = JS + NOB
0054           CALL MODEL(NPROB, TH, DELZ(JS), NOB, NP)
         C     DELZ IS THE NEW PREDICTED FUNCTION VALUE THROUGH THE MODIFIED PARAMETERS
0055           IJ = JS-1
0056           DO 120 I = 1, NOB
0057           IJ = IJ + 1
0058           DELZ(IJ) = DELZ(IJ) - F(I)
0059     120   Q(J) = Q(J) + DELZ(IJ) * R(I)
0060           Q(J)= Q(J)/P(J)
         C                                 Q=XT*R (STEEPEST DESCENT)
0061      130  TH(J) = TEMP
0062           GO TO(131,414), MARY
0063     131   DO 150 I = 1, NP
0064           DO 151 J=1,I
0065           SUM = 0
0066           KJ = NOB*(J-1)
0067           KI = NOB*(I-1)
0068           DO 160 K = 1, NOB
0069           KI = KI + 1
0070           KJ = KJ + 1
0071     160   SUM = SUM + DELZ(KI) * DELZ(KJ)
0072           TEMP= SUM/(P(I)*P(J))
0073           JI = J + NP*(I-1)
0074           D(JI) = TEMP
0075           IJ = I + NP*(J-1)
0076     151   D(IJ) = TEMP
0077     150   E(I) = SQRT(D(JI))
0078     666   CONTINUE
0079           DO 153  I = 1, NP
0080           IJ = I-NP
0081           DO 153 J=1,I
0082           IJ = IJ + NP
0083           A(IJ) = D(IJ) / (E(I)*E(J))
0084           JI = J + NP*(I-1)
0085     153   A(JI) = A(IJ)
         C                                 A= SCALED MOMENT MATRIX
```

ILLEGIBLE

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

ILLEGIBLE

```
            II = - NP
            DO 155 I=1,NP
            P(I)=C(I)/E(I)
            PHI(I)=P(I)
            II = NP + 1 + II
      155   A(II) = A(II) + GA
C
            I=1
            CALL MATIN( A,NP,P,I,DET)
C                                          P/E = CORRECTION VECTOR
             STEP=1.0
            SUM1=0.
            SUM2=0.
            SUM3=0.
            DO 231 I=1,NP
            SUM1=P(I)*PHI(I)+SUM1
            SUM2=P(I)*P(I)+SUM2
            SUM3= PHI(I) * PHI(I) + SUM3
      231   PHI(I) = P(I)
             TEMP = SUM1/SQRT(SUM2*SUM3)
            TEMP = AMIN1(TEMP, 1.0)
            TEMP = 57.295*ACOS(TEMP)
            PRINT 1041, DET, TEMP
      170   DO 220 I = 1, NP
            P(I) = PHI(I) *STEP / E(I)
            TB(I) = TH(I) + P(I)
      220   CONTINUE
            PRINT 7000
      7000  FORMAT(30HOTEST POINT PARAMETER VALUES      )
            PRINT 2006, (TB(I), I = 1, NP)
            DO 221 I = 1, NP
            IF(SIGNS(I)) 221, 221, 222
      222   IF(SIGN(1.0,TH(I))*SIGN(1.0,TB(I))) 663, 221, 221
      221   CONTINUE
            SUMB=0
            CALL MODEL(NPROB, TB, F, NOB, NP)
            DO 230 I=1,NOB
            R(I)=Y(I)-F(I)
     230    SUMB=SUMB+R(I)*R(I)
            PRINT 1043, SUMB
            IF(SUMB - (1.0+EPS1)*SSQ) 662, 662, 663
      663   IF( AMIN1(TEMP-30.0, GA)) 665, 665, 664
     665    STEP=STEP/2.0
            INTCNT = INTCNT + 1
            IF(INTCNT - 36) 170, 2700, 2700
       664  GA=GA*FNU
            INTCNT = INTCNT + 1
            IF(INTCNT - 36) 666, 2700, 2700
      662   PRINT 1007
            DO 669 I=1,NP
     669    TH(I)=TB(I)
            CALL GASS60(1, NP, TH, OO)
            PRINT 1040, GA, SUMB
            GO TO (225,270,265), MAY
      225   DO 240 I = 1, NP
            IF(ABS(P(I))/(1.E-20+ABS(TH(I)))-EPS2) 240, 240, 241
      241   GO TO (265,270) ,LUCY
       240  CONTINUE
```

```
FORTRAN IV G LEVEL  18                  HAUS59              DATE = 71113            03/39/40

0142                PRINT 1009, EPS2                                                        UWHA 178
0143                GO TO 280                                                               UWHA 179
0144           265  IF(ABS(SUMB - SSQ) - EPS1*SSQ) 266, 266, 270                            UWHA 180
0145           266  PRINT 1010, EPS1                                                        UWHA 181
0146                GO TO 280                                                               UWHA 182
0147          270   SSQ=SUMB                                                                UWHA 183
0148                NIT=NIT+1                                                               UWHA 184
0149                IF(NIT - MIT) 100, 100, 280                                             UWHA 185
0150           2700 PRINT 2710                                                              UWHA 186
0151           2710 FORMAT(//115HO**** THE SUM OF SQUARES CANNOT BE REDUCED TO THE SUMUWHA 187
                   1OF SQUARES AT THE END OF THE LAST ITERATION - ITERATING STOPS   /)UWHA 188
              C                                                                             UWHA 189
              C                                                END ITERATION                UWHA 190
              C                                                                             UWHA 191
0152           280  PRINT 1011                                                              UWHA 192
0153                PRINT 2001, (F(I), I = 1, NOB)                                          UWHA 193
0154                PRINT 1012                                                              UWHA 194
0155                PRINT 2001, (R(I), I = 1, NOB)                                          UWHA 195
0156                SSQ=SUMB                                                                UWHA 196
0157                IDF=NOB-NP                                                              UWHA 197
0158                I=0                                                                     UWHA 199
0159                CALL MATIN( D,NP,P,I,DET)
0160                DO 7692 I=1,NP                                                          UWHA 201
0161                II = I + NP*(I-1)                                                       UWHA 202
0162           7692 E(I) = SQRT(D(II))                                                      UWHA 203
0163                DO 340 I=1,NP                                                           UWHA 204
0164                JI = I + NP*(I-1) - 1                                                   UWHA 205
0165                IJ = I + NP*(I-2)                                                       UWHA 206
0166                DO 340 J = I, NP                                                        UWHA 207
0167                JI = JI + 1                                                             UWHA 208
0168                A(JI) = D(JI) / (E(I)*E(J))                                             UWHA 209
0169                IJ = IJ + NP                                                            UWHA 210
0170           340   A(IJ) = A(JI)                                                          UWHA 211
0171                PRINT 1016                                                              UWHA 213
0172                CALL GASS60(1,NP,E,OO)
0173                IF(IDF) 341, 410, 341                                                   UWHA 215
0174           341  SDEV = SSQ / IDF                                                        UWHA 216
0175                PRINT 1014, SDEV, IDF                                                   UWHA 217
0176                 SDEV = SQRT(SDEV)                                                      UWHA 218
0177                DO 391 I=1,NP                                                           UWHA 219
0178                P(I)=TH(I)+2.0*E(I)*SDEV                                                UWHA 220
0179          391   TB(I)=TH(I)-2.0*E(I)*SDEV                                               UWHA 221
0180                PRINT 1039                                                              UWHA 222
0181                CALL GASS60(2, NP, TB, P)
0182                MARY=2                                                                       224
0183                GO TO 101                                                               UWHA 225
0184           414  DO 415 K = 1, NOB                                                       UWHA 226
0185                TEMP = 0                                                                UWHA 227
0186                DO 420 I=1,NP                                                           UWHA 228
0187                DO 420 J=1,NP                                                           UWHA 229
0188                ISUB = K+NOB*(I-1)                                                      UWHA 230
0189                DEBUG1 = DELZ(ISUB)                                                     UWHA 231
              C     DEBUG1 = DELZ(K + NOB*(I-1))                                            UWHA 232
0190                ISUB = K+NOB*(J-1)                                                      UWHA 233
0191                DEBUG2 = DELZ(ISUB)                                                     UWHA 234
              C     DEBUG2 = DELZ(K + NOB*(J-1))                                            UWHA 235
0192                IJ = I + NP*(J-1)                                                       UWHA 236
0193                DEBUG3 = D(IJ)/(DIFZ(I)*TH(I)*DIFZ(J)*TH(J))                            UWHA 237
```

```
FORTRAN IV G LEVEL  18                HAUS59            DATE = 71113          03/39/40

0194           420  TEMP = TEMP + DEBUG1 * DEBUG2 * DEBUG3                              UWHA 238
0195                 TEMP = 2.0*SQRT(TEMP)*SDEV                                         UWHA 239
0196                R(K)=F(K)+TEMP                                                      UWHA 240
0197           415  F(K)=F(K)-TEMP                                                      UWHA 241
0198                PRINT 1008                                                          UWHA 242
0199                IE=0                                                                UWHA 243
0200                DO 425 I=1,NOB,10                                                   UWHA 244
0201                IE=IE+10                                                            UWHA 245
0202                IF(NOB-IE) 430,435,435                                              UWHA 246
0203            430 IE=NOB                                                              UWHA 247
0204           435  PRINT 2001, (R(J), J = I, IE)                                       UWHA 248
0205           425  PRINT 2006, (F(J), J = I, IE)                                       UWHA 249
0206           410  PRINT 1033, NPROB                                                   UWHA 250
0207                RETURN                                                              UWHA 251
0208           99   PRINT 1034                                                          UWHA 252
0209                GO TO 410                                                           UWHA 253
0210           1000FORMAT(38H1NON-LINEAR ESTIMATION, PROBLEM NUMBER   I3,// I5,         UWHA 254
                  1  14H OBSERVATIONS, I5, 11H PARAMETERS  I14, 17H SCRATCH REQUIRED)UWHA 255
0211           1001 FORMAT(/25H0INITIAL PARAMETER VALUES  )                             UWHA 256
0212           1002 FORMAT(/54H0PROPORTIONS USED IN CALCULATING DIFFERENCE QUOTIENTS )UWHA 257
0213           1003 FORMAT(/25H0INITIAL SUM OF SQUARES =    E12.4)                      UWHA 258
0214           1004 FORMAT(/////45X,13HITERATION NO.  I4)                               UWHA 259
0215           1007 FORMAT(/32H0PARAMETER VALUES VIA REGRESSION  )                      UWHA 260
0216           1008 FORMAT(////54H0APPROXIMATE CONFIDENCE LIMITS FOR EACH FUNCTION VALUWHA 261
                  1LE     )                                                             UWHA 262
0217           1009CFORMAT(/62H0ITERATION STOPS - RELATIVE CHANGE IN EACH PARAMETER LEUWHA 263
                  1SS THAN    E12.4)                                                    UWHA 264
0218           1010CFORMAT(/62H0ITERATION STOPS - RELATIVE CHANGE IN SUM OF SQUARES LEUWHA 265
                  1SS THAN    E12.4)                                                    UWHA 266
0219           1011 FORMAT(22H1FINAL FUNCTION VALUES  )                                 UWHA 267
0220           1012 FORMAT(////10H0RESIDUALS  )                                         UWHA 268
0221           1014 FORMAT(//24H0VARIANCE OF RESIDUALS =     ,E12.4,1H,I4,              UWHA 269
                  120H DEGREES OF FREEDOM   )                                           UWHA 270
0222           1016 FORMAT(////21H0NORMALIZING ELEMENTS   )                             UWHA 272
0223           1033 FORMAT(//19H0END OF PROBLEM NO.  I3)                                UWHA 273
0224           1034 FORMAT(/16H0PARAMETER ERROR  )                                      UWHA 274
0225           1039CFORMAT(/71H0INDIVIDUAL CONFIDENCE LIMITS FOR EACH PARAMETER (ON LIUWHA 275
                  1NEAR HYPOTHESIS)    )                                                UWHA 276
0226           1040CFORMAT(/9H0LAMBDA =E10.3,40X,33HSUM OF SQUARES AFTER REGRESSION = UWHA 277
                  1E15.7)                                                               UWHA 278
0227           1041 FORMAT(14H DETERMINANT =  E12.4, 6X, 25H ANGLE IN SCALED COORD. = UWHA 279
                  1  F5.2, 8HDEGREES    )                                               UWHA 280
0228           1043 FORMAT(28H0TEST POINT SUM OF SQUARES =    E12.4)                    UWHA 281
0229           2001 FORMAT(/10E12.4)                                                    UWHA 282
0230           2006 FORMAT(10E12.4)                                                     UWHA 283
0231                END                                                                 UWHA 284
```

```
      SUBROUTINE MATIN(A, NVAR, B, NB, DET)
      DIMENSION A(NVAR,1)    , B(NVAR,1)
      COMMON/GASPAR/DUMIES(7), PIVOTM
       PIVOTM = A(1,1)
       DET = 1.0
      DO 550 ICOL = 1, NVAR
       PIVOT = A(ICOL, ICOL)
      PIVOTM = AMIN1(PIVOT, PIVOTM)
       DET = PIVOT * DET
C
C     DIVIDE PIVOT ROW BY PIVOT ELEMENT
C
       A(ICOL, ICOL) = 1.0
C
      PIVOT = AMAX1(PIVOT, 1.E-20)
       PIVOT = A(ICOL, ICOL)/PIVOT
      DO 350 L=1,NVAR
 350   A(ICOL, L) = A(ICOL, L)*PIVOT
      IF(NB .EQ. 0) GO TO 371
      DO 370 L=1,NB
 370   B(ICOL, L) = B(ICOL, L)*PIVOT
C
C     REDUCE NON-PIVOT ROWS
C
 371  DO 550 L1=1,NVAR
       IF(L1 .EQ. ICOL) GO TO 550
       T = A(L1, ICOL)
       A(L1, ICOL) = 0.
      DO 450 L=1,NVAR
 450   A(L1, L) = A(L1, L) - A(ICOL, L)*T
      IF(NB .EQ. 0) GO TO 550
      DO 500 L=1,NB
 500   B(L1, L) = B(L1, L)-B(ICOL,L)*T
  550 CONTINUE
      RETURN
      END
```

```
FORTRAN IV G LEVEL  18              GASS60              DATE = 71113              03/39/40

            SUBROUTINE GASS60(ITYPE, NQ, A, B)
            DIMENSION A(NQ), B(NQ)
             NP = NQ                                                      UWHA 323
             NR = NP/10                                                   UWHA 324
             LOW = 1                                                      UWHA 325
             LUP = 10                                                     UWHA 326
     10     IF( NR )15,20,30                                              UWHA 327
     15     RETURN                                                        UWHA 328
     20     LUP=NP                                                        UWHA 329
            IF(LOW .GT. LUP) RETURN                                       UWHA 330
     30     PRINT 500, (J,J=LOW,LUP)                                      UWHA 331
            GO TO (40,60), ITYPE
     40     PRINT 600,(A(J),J=LOW,LUP)                                    UWHA 333
            GO TO 100                                                     UWHA 334
     60     PRINT 600, (B(J),J=LOW,LUP)                                   UWHA 335
            GO TO 40                                                      UWHA 336
    100      LOW = LOW + 10                                               UWHA 343
             LUP = LUP + 10                                               UWHA 344
             NR = NR - 1                                                  UWHA 345
             GO TO 10                                                     UWHA 346
    500     FORMAT(/I8,9I12)                                              UWHA 347
    600     FORMAT(10E12.4)                                               UWHA 348
            END                                                           UWHA 352
```

APPENDIX C

PROGRAM FORCAT

C.1. Description of program

Program FORCAT is developed to provide the forecast value and its confidence interval for the appropriate model of the time series, which may be stationary, nonstationary or/and seasonal. The program consists of a main program and seven subroutines. The functions of subroutines are as follows.

MULTS and EXPAND: convert the general seasonal nonstationary model into regular ARMA(p,q) forms. The form is used to calculate the π weights (pure AR weights), ψ weights (pure MA weights), confidence intervals of forecast values for the original series and the reduced stationary model.

CALPSI : calculate the ψ weights for the original model and for the reduced stationary model. The calculations are based on equation (4.2.6)

CALPIE : calculates the pure AR weights for the original model and for the reduced stationary model. The calculation is based on equation (C.1.4). For pure moving-average model, its form may be

$$\tilde{z} = \psi(B)\ a_t \tag{C.1.1}$$

where

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \ldots$$

For pure autoregressive model, its form may be

$$\pi(B)\ \tilde{z}_t = a_t \tag{C.1.2}$$

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

comparing (C.1.1) with (C.1.2), we obtain

$$\pi(B)\ \psi(B) = 1 \tag{C.1.3}$$

equating the power of B in (C.1.3), getting

$$\pi_1 = \psi_1$$

$$\pi_2 = \psi_2 - \psi_1 \pi_1$$

$$\pi_3 = \psi_3 - \psi_2 \pi_1 - \psi_1 \pi_2$$

$$\vdots$$

or in more general equation

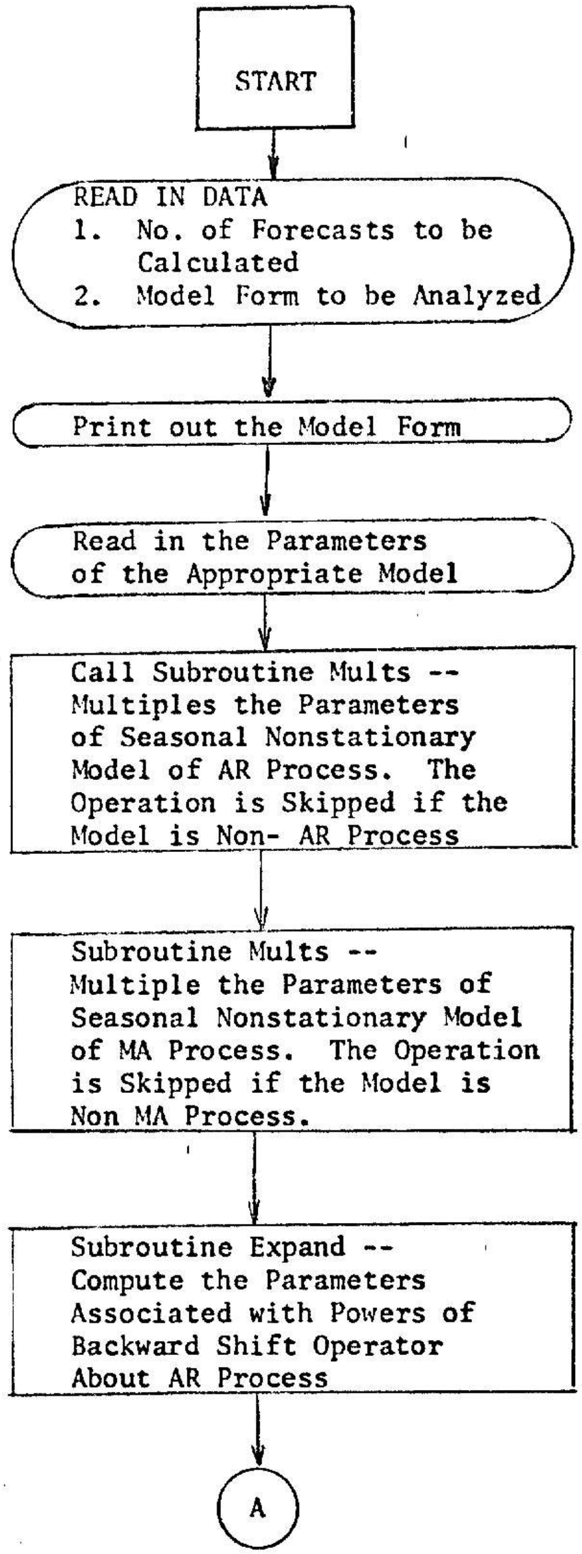
$$\pi_i = \psi_i - \sum_{j=1}^{i-1} \psi_j\ \pi_{i-j} \tag{C.1.4}$$

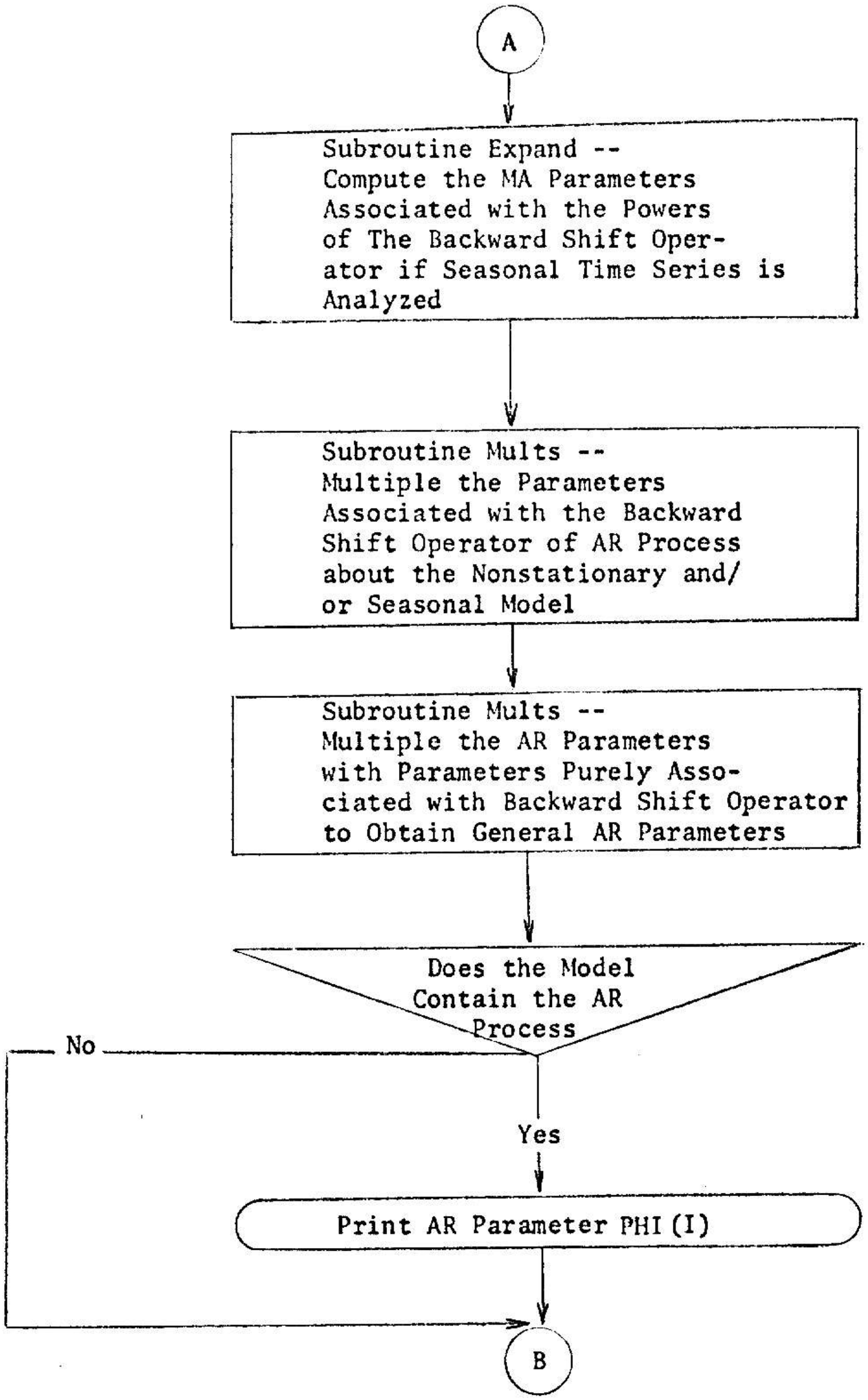
hence, the π weights of pure AR model can be calculated recursively.

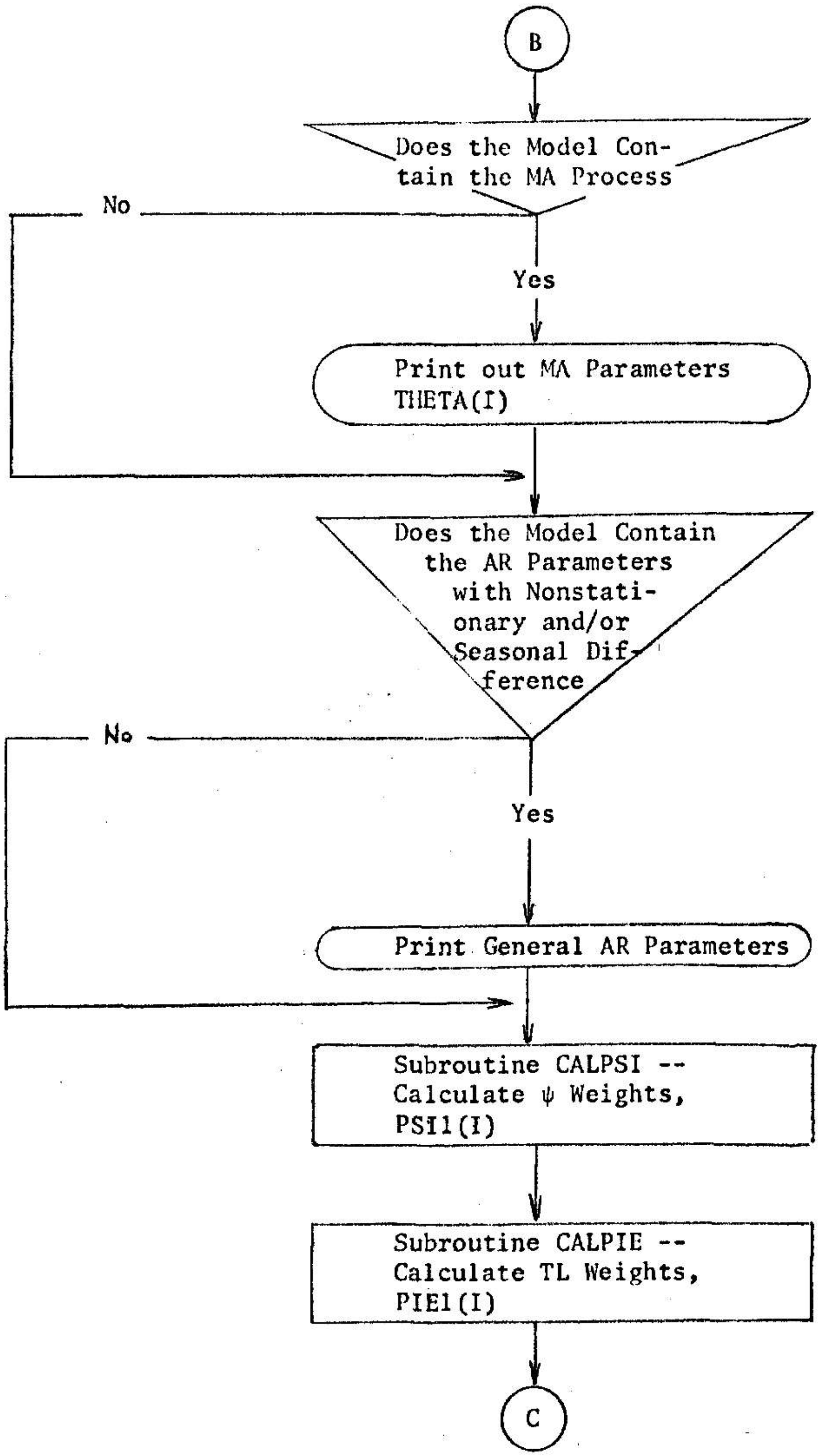
FORCAT : calculate the forecasts and confidence interval for the original time series. Its calculation is based on equations (4.1.2) and (4.2.7).

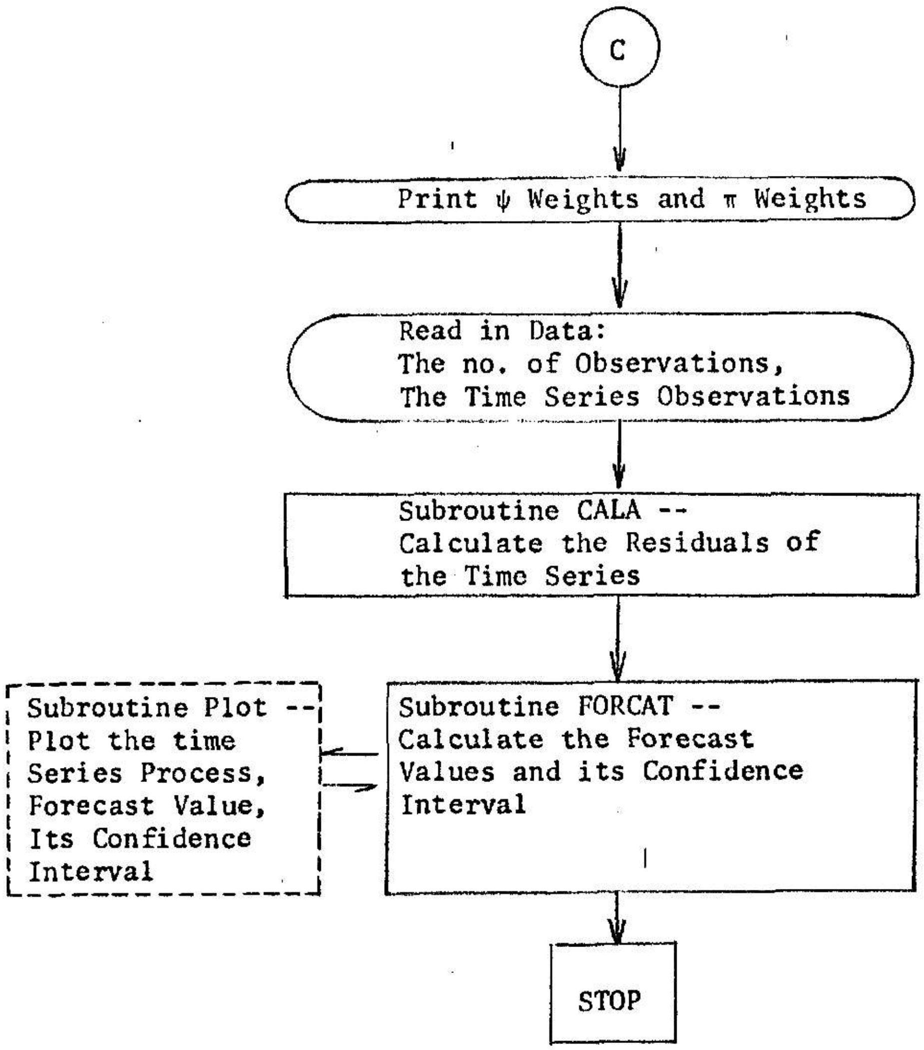
CALA : calculate the residuals of the original time series.

PLOT : plot the forecast values and its confidence intervals,


START
READ IN DATA
1. No. of Forecasts to be Calculated
2. Model Form to be Analyzed
Print out the Model Form
Read in the Parameters of the Appropriate Model
Call Subroutine Mults -- Multiples the Parameters of Seasonal Nonstationary Model of AR Process. The Operation is Skipped if the Model is Non- AR Process
Subroutine Mults -- Multiple the Parameters of Seasonal Nonstationary Model of MA Process. The Operation is Skipped if the Model is Non MA Process.
Subroutine Expand -- Compute the Parameters Associated with Powers of Backward Shift Operator About AR Process
A


A
Subroutine Expand --
Compute the MA Parameters
Associated with the Powers
of The Backward Shift Oper-
ator if Seasonal Time Series is
Analyzed
Subroutine Mults --
Multiple the Parameters
Associated with the Backward
Shift Operator of AR Process
about the Nonstationary and/
or Seasonal Model
Subroutine Mults --
Multiple the AR Parameters
with Parameters Purely Asso-
ciated with Backward Shift Operator
to Obtain General AR Parameters
Does the Model
Contain the AR
Process
No
Yes
Print AR Parameter PHI (I)
B


B
Does the Model Contain the MA Process
No
Yes
Print out MA Parameters THETA(I)
Does the Model Contain the AR Parameters with Nonstationary and/or Seasonal Difference
No
Yes
Print General AR Parameters
Subroutine CALPSI -- Calculate ψ Weights, PSI1(I)
Subroutine CALPIE -- Calculate TL Weights, PIE1(I)
C


C
Print ψ Weights and π Weights
Read in Data:
The no. of Observations,
The Time Series Observations
Subroutine CALA --
Calculate the Residuals of
the Time Series
Subroutine Plot --
Plot the time
Series Process,
Forecast Value,
Its Confidence
Interval
Subroutine FORCAT --
Calculate the Forecast
Values and its Confidence
Interval
STOP

C.2. Description of input data

card	variable in program	FORMAT	Description
first card	IFOR	(I10)	number of forecasts to be calculated
next block of cards			same as the parameter input block given in Appendix B for program ESTIM
next card	MAX	(I10)	number of correlation to be calculated
next card	N	(I10)	the number of observations to be read in.
next cards	Z(I,1)	(2X,(F20.5))	the observations of the time series.

C.3. Description of output data

The following data can be generated by program FORCAT;

1. Print out parameters and model form.
2. Pure autoregressive and pure moving-average weights of the original and reduced stationary form of the model.
3. Forecast values and confidence limits.
4. Plot out the forecast values and confidence limits.

APPENDIX C. PROGRAM FORCAT

C. 4 Computer Program

```
      DIMENSION PHIW(200),THETW(200),PHIB(200),THETB(200)
      DIMENSION PHI(200),THETA(200),CPHI(200)
      DIMENSION Z(1500,1),A(1500)
      DIMENSION PIE(200),PIE1(200),PSI(200),PSI1(200)
      DIMENSION C1(200),C2(200),C(200)
      COMMON/A1/PHIW/A2/THETW/A3/PHIB/A4/THETB/A5/PHI/A6/THETA/A7/CPHI
      COMMON/A8/Z,A/A9/X,P/A10/PIE/A11/PIE1/A12/PSI/A13/PSI1/A14/RHO
      COMMON/A15/PRHO
      COMMON/A16/TREND
C     INSERT TO READ IN DATA AND PRINT OUT
  250 FORMAT(I10)
C     MAX  NUMBER OF CORRELATIONS AND PARTIAL TO BE CALCULATED
      READ 250,MAX
C     PROGRAM PARAMETERS
C     END PROGRAM PARAMETERS
C     READ IN MODEL FROM PARAMETERS
 2002 FORMAT(4I10)
      READ 2002,IPW,IDW,IQW
      READ 2002,IPB,IDB,IQB,IS
C     END READ IN MODEL FORM PARAMETERS
C     READ IN PARAMETERS
 2003 FORMAT(1H12X,6HIARMA(,I3,1H,,I3,1H,,I3,3H)X(,I3,1H,,I3,1H,,I3,3H)X
     1(,I3,1H)//)
      PRINT 2003,IPW,IDW,IQW,IPB,IDB,IQB,IS
      IP=IPW+IPB
      IQ=IQW+IQB
      IDS=IDB*IS
      IPID=IP+IDW+IDS
      ID=IDW+IDB
      DO 2090 I=1,100
      PHIW(I)=0.
      PHIB(I)=0.
      PHI(I)=0.
      THETW(I)=0.
      THETB(I)=0.
      THETA(I)=0.
      CPHI(I)=0.
      C1(I)=0.
      C2(I)=0.
 2090 C(I)=0.
 2004 FORMAT(I10,F20.5)
 2005 FORMAT(2X,5HPHIW(,I3,2H)=,F11.5)
 2007 FORMAT(2X,5HPHIB(,I3,2H)=,F11.5)
 2006 FORMAT(2X,6HTHETW(,I3,2H)=,F11.5)
 2008 FORMAT(2X,6HTHETB(,I3,2H)=,F11.5)
 2027 FORMAT(2X,6HTREND(,I3,2H)=,F11.5)
      READ 2002,I1,I2,I3,I4
      IF(I1)2009,2009,2010
 2010 DO 2011 I=1,I1
      READ 2004,J,PHIW(J)
 2011 PRINT 2005,J,PHIW(J)
 2009 IF(I2)2012,2012,2013
 2013 DO 2014 I=1,I2
      READ 2004,J,THETW(J)
 2014 PRINT 2006,J,THETW(J)
 2012 IF(I3)2015,2015,2016
 2016 DO 2017 I=1,I3
      READ 2004,J,PHIB(J)
```

```
      2017 PRINT 2007,J,PHIB(J)
      2015 IF(I4) 2025,2025,2019
      2019 DO 2020 I=1,I4
           READ 2004,J,THETB(J)
      2020 PRINT 2008,J,THETB(J)
      2025 CONTINUE
C          END READ IN PARAMETERS
C          OPERATIONS TO REDUCE SEASONAL TO NON SEASONAL MODEL
           CALL MULTS (PHIW,PHIB,PHI,IP)
           CALL MULTS(THETW,THETB,THETA,IQ)
           CALL EXPAND(C1,IDW,1)
           CALL EXPAND(C2,IDB,IS)
           IDC=IDS+IDW
           CALL MULTS(C1,C2,C,IDC)
           CALL MULTS (PHI,C,CPHI,IPID)
C          END OPERATIONS
C          PRINT OUT OF PERTINANT DATA
      2021 FORMAT(////2X,27HPHI, THETA, AND CPHI VALUEW//)
           PRINT 2021
       107 FORMAT(2X,6H   PHI(,I3,2H)=,F10.5)
       108 FORMAT(2X,6HTHETA(,I3,2H)=,F10.5)
           IF(IP)109,109,110
       110 DO 111 I=1,IP
       111 PRINT 107,I,PHI(I)
       109 IF(IQ)112,112,113
       113 DO 114 I=1,IQ
       114 PRINT 108,I,THETA(I)
       112 CONTINUE
       115 FORMAT(/////2X,33HGENERAL AUTOREGGRESIVE PARAMETERS///)
           PRINT 115
       116 FORMAT(2X,6H CPHI(,I3,2H)=,F10.5)
           IF(IPID)1120,1120,1121
      1121 CONTINUE
           DO 117 I=1,IPID
       117 PRINT 116,I,CPHI(I)
      1120 CONTINUE
C          CALCULATION OF PSI WEIGHTS
           IF(IP-4)7000,7000,7001
      7000 IF(IQ-2)7002,7002,7001
      7002 KDOZ=MAX
           GO TO 7003
      7001 KDOZ=200
      7003 CONTINUE
C          TO GET PURE MOVING AVERAGE PARAMETERS PSI FOR STATIONARY TIME SERIES
           CALL CALPSI(PHI,THETA,PSI1,IP,IQ,KDOZ)
C          TO GET PURE AUTOREGRESSIVE PARAMETERS PIE FOR STATIONARY TIME SERIES
           CALL CALPIE(PSI1,PIE1,MAX)
       300 FORMAT(/////2X,41HPSI AND PIE WEIGHTS FOR STATIONARY SERIES//)
           PRINT 300
       302 FORMAT(2X,4HPSI(,I3,2H)=,F10.6,7X,4HPIE(,I3,2H)=,F10.6)
           DO 303 I=1,MAX
       303 PRINT 302,I,PSI1(I),I,PIE1(I)
           IF(IPID-IP)5000,5000,5001
C          IF IPID=IP, MEANS NO ANY DIFFERENCES AND SEASONAL DIFFERENCE, SO WE CAN
C          LOOK ORIGINAL SERIES AS STATIONARY TIME SERIES.
C          TO GET PURE MOVING AVERAGE PARAMETERS PSI FOR ORIGINAL SERIES
      5001 CALL CALPSI(CPHI,THETA,PSI,IPID,IQ,MAX)
C          TO GET PURE AUTOREGRESSIVE PARAMETERS FOR ORIGINAL SERIES
```

```
FORTRAN IV G LEVEL  18                MAIN                CATE = 71113          03/42/28

0098              CALL CALPIE(PSI,PIE,MAX)
0C99        5CC2 FORMAT(/////2X,39HPSI AND PIE WEIGHTS FOR ORIGINAL SERIES//)
0100              PRINT 5C02
0101              CC 5003 I=1,MAX
0102        5C03 PRINT 302,I,PSI(I),I,PIE(I)
0103              GO TO 5C06
0104        5OCC CO 5005 I=1,MAX
0105        5005 PSI(I)=PSI1(I)
0106        5C06 CCNTINUE
0107        251  FORMAT(2X,F20.5)
           C      IFCR  NC OF FORECASTS TO BE CALCULATED, MAXIMUM FORECAST LAG.
0108              READ 250,IFCR
0109              READ  250,N
0110              REAC 251,(Z(I,1),I=1,N)
           C      CALL TC GENERATE A(I)
0111              SIGMA=0.
0112              CALL CALA(N,IPID,IC,SIGMA)
           C      FCRECASTING
           C      CALL TC GENERATE FCRECASTS ANC HPD INTERVAL
0113              CALL FCRCAT(N,IPID,IQ,IFOR,SIGMA)
           C      ENC FCRCASTING
0114              STOP
0115              END
```

FORTRAN IV G LEVEL 18 CALPSI DATE = 71113 03/42/28

```
      SUBROUTINE CALPSI(CPHI,THETA,PSI,IPP,IQ,MAX)
C     SUBROUTINE CALPSI  IS TO CALCULATE THE PURE MOVING AVERAGE WEIGHTS
C     FOR THE MODEL
      DIMENSION CPHI(200),THETA(200),PSI(200)
C     FIX INITIAL VALUES SO THAT IPP=IQ=IPSI
      IF(IPP-IQ)100,101,102
  100 IDUM=IPP+1
      DO 103 I=IDUM,IQ
  103 CPHI(I)=0.
      IPSI=IQ
      GO TO 110
  101 IPSI=IQ
      GO TO 110
  102 IDUM=IQ+1
      DO 104 I=IDUM,IPP
      THETA(I)=0.
  104 CONTINUE
      IPSI=IPP
  110 CONTINUE
C     END FIX
C     CALCULATION OF PSI(1)...PSI(IPSI)
      DO 11  I=1,IPSI
   11 PSI(I)=CPHI(I)-THETA(I)
      DO 112 I=2,IPSI
      JDUM=I-1
      DO 112 J=1,JDUM
      PSI(I)=PSI(I)+CPHI(J)*PSI(I-J)
  112 CONTINUE
C     END CALCULATION
C     CALCULATION OF PSI(IPSI),.........,PSI(MAX)
      IDUM=IPSI+1
      DO 113 I=IDUM,MAX
      PSI(I)=0.
      DO 113 J=1,IPSI
      PSI(I)=PSI(I)+CPHI(J)*PSI(I-J)
  113 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE CALPIE(PSI,PIE,MAX)
      DIMENSION PSI(200),PIE(200)
      DO 100 I=1,MAX
  100 PIE(I)=PSI(I)
      DO 101 I=2,MAX
      JDUM=I-1
      DO 101 J=1,JDUM
      PIE(I)=PIE(I)-PSI(J)*PIE(I-J)
  101 CONTINUE
      RETURN
      END
```

```
FORTRAN IV G LEVEL  18                 CALA                 DATE = 71113            03/42/28

             SUBROUTINE CALA(N,IPID,IQ,SIGMA)
             DIMENSION THETA(200),CPHI(200),Z(1500,1),A(1500)
             COMMON/A8/Z,A/A6/THETA/A7/CPHI
      C      FIX INITIAL VALUES
             IF(IQ-IPID)201,201,200
        200  ISTART=IQ
             GO TO 203
        201  ISTART=IPID
        203  DO 204 I=1,ISTART
        204  A(I)=0.
             ISTART=ISTART+1
             DO 206 I=ISTART,N
             A(I)=Z(I,1)
             IF(IPID)207,207,208
        208  DO 209 J=1,IPID
        209  A(I)=A(I)-CPHI(J)*Z(I-J,1)
         207 IF(IQ) 206,206,211
        211  DO 212 J=1,IQ
        212  A(I)=A(I)+THETA(J)*A(I-J)
        206  CONTINUE
        220  FORMAT(1H12X,27HPRINT OUT OF DATA AND ERROR//)
             PRINT 220
        217  FORMAT(2X,5HTIME=,I4,2X,2HZ(,I4,2H)=,F20.5,5X,2HA(,I4,2H)=,F20.5)
             SIGMA=0.
             DO 218 I=1,N
             PRINT 217,I,I,Z(I,1),I,A(I)
             SIGMA=SIGMA+A(I)*A(I)
        218  CONTINUE
             SIGMA=SQRT(SIGMA/N)
             RETURN
             END
```

```
FORTRAN IV G LEVEL  18                  FORCAT              DATE = 71113           03/42/28

            SUBROUTINE FORCAT(IT,IPID,IQ,IFOR,SIGMA)
            DIMENSION H(1000)
            DIMENSION G(3000)
            DIMENSION CPHI(200),THETA(200),PSI(200),Z(1500,1),A(1500)
            DIMENSION ZHAT(1500),U(200),X1(200),X2(200)
            COMMON/A7/CPHI/A6/THETA/A12/PSI/A8/Z,A/A9/X,P
      100   FORMAT(1H12X,48HFORECASTS AND 95 PER CENT LIMITS FOR BASE TIME =,I
           14///)
            PRINT 100 ,IT
            DO 101 I=1,IT
      101   ZHAT(I)=Z(I,1)
      5000  CONTINUE
            IT1=IT+1
            ITFOR=IT+IFOR
            DO 102 I=IT1,ITFOR
      102   ZHAT(I)=0.
            K=1
            DO 103 I=IT1,ITFOR
            IF(IPID)104,104,105
      105   DO 106 J=1,IPID
      106   ZHAT(I)=ZHAT(I)+CPHI(J)*ZHAT(I-J)
       104  IF(I-IT-IQ) 107,107,3000
      107   DO 108 J=K,IQ
      108   ZHAT(I)=ZHAT(I)-THETA(K)*A(IT-J+K)
            K=K+1
      3000  CONTINUE
      103   CONTINUE
     C      CALCULATION OF UPPER AND LOWER 95 PER CENT POINTS
            DO 200 I=1,IFOR
      200   U(I)=1.
            DO 201 I=2,IFOR
            L=I-1
            DO 202 J=1,L
       202  U(I)=U(I)+PSI(J)**2
            U(I)=1.96*SQRT(U(I))*SIGMA
      201   CONTINUE
            U(1)=1.96*SIGMA
            DO 203 I=1,IFOR
            X1(I)=ZHAT(IT+I)-U(I)
      203   X2(I)=ZHAT(IT+I)+U(I)
      300   FORMAT(2X,5HTIME=,I4,5X,5HZHAT(,I4,2H)=,F20.5,5X,13HHPD INTERVAL=,
           1F20.5,2X,F20.5)
      400   FORMAT(2X,6HSIGMA=,F20.5///)
            PRINT 400,SIGMA
            DO 301 I=1,IFOR
     C      THIS DO LOOP IS TO SUPPLY DATA TO SUBROUTINE PLOT TO PRINT OUT
     C      THE FORECAST VALUE ZHAT(I1), THEIR CORRESPONDING UPPER AND LOWER LIMITS
            I1=I+IT
            K1=IFOR+I
            K2=2*IFOR+I
            K3=3*IFOR+I
            H(I)=I
            H(K1)=ZHAT(I1)
            H(K2)=X1(I)
            H(K3)=X2(I)
            PRINT 300,I1,I,ZHAT(I1),X1(I),X2(I)
      301   CONTINUE
            CALL PLOT(1,H,IFOR,4,IFOR,0)
```

```
              DO 120 I=1,ITFOR
              I1=ITFOR+I
              G(I)=I
          120 G(I1)=ZHAT(I)
              CALL PLOT(1,G,ITFOR,2,ITFOR,0)
          351 FORMAT(1H12X,18H1.96*S.D.(ZHAT(L))///)
              PRINT 351
          352 FORMAT(2X,2HL=,I5,5X,5HU(L)=,F20.5)
              DO 353 I=1,IFOR
        C     U(I) IS THE CONTROL LIMITS FOR CORRESPONDING FORECAST VALUE
          353 PRINT 352,I,U(I)
              RETURN
              END
```

```
FORTRAN IV G LEVEL  18                  EXPAND               DATE = 71113            03/42/28

            SUBROUTINE EXPAND(C,ID,IS)
            DIMENSION C(200)
            IF(ID)250,250,251
      251   CONTINUE
            DO 120 I=1,ID
            JA1=1
            JA2=1
            JA3=1
            DO 140 J=1,ID
      140   JA1=JA1*J
            IF(ID-I)132,132,133
      133   JDUM=ID-I
            DO 141 J=1,JDUM
      141   JA2=JA2*J
      132   DO 142 J=1,I
      142   JA3=JA3*J
            JJJ=(-1)**I
            JJJ=JJJ*JA1/(JA2*JA3)
            C(I*IS)=-JJJ
      120   CONTINUE
      250   CONTINUE
            RETURN
            END
```

```
FORTRAN IV G LEVEL  18                  MULTS               DATE = 71113          03/42/28

                SUBROUTINE MULTS(PHIW,PHIB,PHI,IPWPB)
                DIMENSION PHIW(200),PHIB(200),PHI(200)
                IF(IPWPB)101,101,105
          105   DO 100 I=1,IPWPB
          100   PHI(I)=PHIW(I)+PHIB(I)
                IF(IPWPB-1)101,101,102
          102   DO 104 I=2,IPWPB
                JDUM=I-1
                DO 104 J=1,JDUM
          104   PHI(I)=PHI(I)-PHIB(J)*PHIW(I-J)
          101   CONTINUE
                RETURN
                END
```

```
0001              SUBROUTINE PLOT(NO,A,N,M,NL,NS)
            C
0002              DIMENSION OUT(101),YPR(11),ANG(9),A(1)
            C
0003            1 FORMAT(1H1,60X,7H CHART ,I3,//)
0004            2 FORMAT(1H ,F11.4,5H+     ,101A1)
0005            4 FORMAT(10H 123456789)
0006            5 FORMAT( 10A1)
0007            7 FORMAT(1H ,16X,101H.         .         .         .         .
               1     .         .         .         .         .         .)
0008            8 FORMAT(1H0,9X,11F10.4//)
0009          200 FORMAT( 10X,'      PLOT OF FORCAST VALUE')
0010          201 FORMAT( 10X,'     PLOT OF AUTO CORR. FUNCTION')
0011          202 FORMAT( 10X,'      PLOT OF SPECTRUM')
            C
            C     .....................................................................
            C
0012              NLL=NL
            C
0013              IF(NS)16,16,10
            C
            C        SORT BASE VARIABLE IN ASCENDING ORDER
            C
0014           10 DO 15 I=1,N
0015              DO 14 J=1,N
0016              IF(A(I)-A(J))14,14,11
0017           11 L=I-N
0018              LL=J-N
0019              DO 12 K=1,M
0020              L=L+N
0021              LL=LL+N
0022              F=A(L)
0023              A(L)=A(LL)
0024           12 A(LL)=F
0025           14 CONTINUE
0026           15 CONTINUE
            C
            C        TEST NLL
            C
0027           16 IF(NLL)20,18,20
0028           18 NLL=50
            C
            C        PRINT TITLE
            C
0029           20 WRITE(3,1)NO
0030              GO TO (91,92,93),NO
0031           91 WRITE(3,200)
0032              GO TO 21
0033           92 WRITE(3,201)
0034              GO TO 21
0035           93 WRITE(3,202)
0036           21 CONTINUE
            C
            C        DEVELOP BLANKS AND DIGITS FOR PRINTING
            C
0037              REWIND 4
0038              WRITE(4,4)
0039              REWIND 4
```

```
                READ(4,5)BLANK,(ANG(I),I=1,9)
                REWIND 4
          C
          C        FIND SCALE FOR BASE VARIABLE
          C
                XSCAL=(A(N)-A(1))/(FLOAT(NLL-1))
          C
          C        FIND SCALE FOR CROSS VARIABLES
          C
                M1=N+1
                YMIN=A(M1)
                YMAX=YMIN
                M2=M*N
                DO 40 J=M1,M2
                IF(A(J)-YMIN)28,28,26
             26 IF(A(J)-YMAX)40,40,30
             28 YMIN=A(J)
                GO TO 40
             30 YMAX=A(J)
             40 CONTINUE
                YSCAL=(YMAX-YMIN)/100.0
          C
          C        FIND BASE VARIABLE PRINT POSITION
          C
                XB=A(1)
                L=1
                MY=M-1
                I=1
             45 F=I-1
                XPR=XB+F*XSCAL
          C     IF(A(L)-XPR)51,51,70         THIS CARD HAS BEEN REMOVED
          C
          C        FIND CROSS VARIABLES
          C
             51 DO 55 IX=1,101
             55 OUT(IX)=BLANK
             57 DO 60 J=1,MY
                LL=L+J*N
                JP=((A(LL)-YMIN)/YSCAL)+1.0
                OUT(JP)=ANG(J)
             60 CONTINUE
          C
          C        PRINT LINE AND CLEAR, OR SKIP
          C
                WRITE(3,2)XPR,(OUT(IZ),IZ=1,101)
                L=L+1
                GO TO 80
             80 I=I+1
                IF(I-NLL)45,84,86
             84 XPR=A(N)
                GO TO 51
          C
          C        PRINT CROSS VARIABLES NUMBERS
          C
             86 WRITE(3,7)
                YPR(1)=YMIN
                DO 90 KN=1,9
             90 YPR(KN+1)=YPR(KN)+YSCAL*10.0
```

APPLICATION OF GENERAL AUTOREGRESSIVE MOVING-AVERAGE STOCHASTIC MODELS TO TIME SERIES AND SIMULATION PROBLEM

by

Steven Shyan-Ming Ko

B.S., Taiwan Cheng-Kung University, 1966

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirement for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1971

This thesis is concerned with the development of a stochastic model (general autoregressive moving-average model) to represent a time series and forecast its future values. An iterative model-building methodology, including model identification, model estimation, model diagnostic checking and employment of the model to forecast the time series are explored and illustrated in this thesis.

Application of the general autoregressive moving-average model is illustrated by identifying the appropriate model and forecasting for an industrial chemical process, a simulated inventory system and international airline passenger fluctuation. The computer programming and human judgement both contribute to these experiments.

From the computational results, it is found that the general autoregressive moving-average model not only represents the discrete time series in the time domain, but also possess the characteristics of maximum simplicity and minimum number of parameters with representational adequacy.

Finally, further research is suggested to put the entertained model under more strictly diagnostic checks in order that it can represent the time series process adequately.