SELECTION OF SMOOTHING CONSTANTS FOR AN EXPONENTIALLY WEIGHTED TIME SERIES MODEL

By

GLENDON DANNA SELF

Bachelor of Science
University of Arkansas
Fayetteville, Arkansas
1958

Master of Science
University of Arkansas
Fayetteville, Arkansas
1959

Submitted to the Faculty of the Graduate School
of the Oklahoma State University
in partial fulfillment of
the requirements for
the degree of
DOCTOR OF PHILOSOPHY
August, 1963

SELECTION OF SMOOTHING CONSTANTS FOR AN
EXPONENTIALLY WEIGHTED TIME SERIES
MODEL

Thesis Approved:

W. J. Bentley
Thesis Adviser

Paul E. Torgersen

David L. Weeks

J. Leroy Folks

Carl E. Marshall

Robert MacVicar
Dean of the Graduate School

PREFACE

The study of methods of applied time series is usually accompanied by the formulation of a model that is used to forecast the future of the series. Since perfect information of the future is rare, these methods will have some error associated with the forecasts. The discussion of these methods in the literature is not uniform in the definition of the error, the distribution of the error or the interpretation of the error.

In this thesis, the error is defined without any assumptions of the distribution of errors in order to provide generality in the methods considered here. These methods were developed for the real world process and the examples presented are of this type.

The computational algorithms used to provide the input for the forecasting model decompose the series into its linear trend, seasonal variation, and random variation about some mean level by the use of exponentially weighted moving averages. The exponential weighting of the historical data is a function of the smoothing constants which are selected. Therefore, the selection of these constants determine the fit of forecast values to the actual data. The selection of these constants are based upon the response of the series to the application of the smoothing constants in terms of the mean forecast error squared under some defined concepts of optimality of the fit. The methods of selection of an optimum starting point within the historical data are also derived and supported by examples.

I would like to gratefully acknowledge the contribution of Dr. Leroy J. Folks for his helpful suggestions on the content and arrangement of this thesis material; of Professor Wilson J. Bentley, Head of the School of Industrial Engineering and Management, for his advice and counsel during my graduate program; of Dr. Paul E. Torgersen for his interest and suggestions during the total program; of Dr. David L. Weeks, Dr. Carl E. Marshall and Dr. Robert A. Hultquist for serving on the graduate committee and providing quality instruction during my graduate course work, and Dr. Robert D. Morrison and Dr. George F. Schrader who added to my general educational level during the course of this work. Also, a debt of gratitude is acknowledged to Sandia Corporation for providing the atmosphere and facilities for conducting this study, and to Mr. L. E. Snodgrass, both friend and supervisor, for his continued encouragement and making time and facilities available for these investigations; to Mr. R. W. Devore and Mr. C. C. Fornero for introducing me to this problem and their helpful suggestions in the improvements that were needed in the forecasting scheme; Mr. D. D. Sheldon for making the time of Mitzi Fortenbury and Sue Bell Propst available for experiments in programming and data preparation.

The final contributions to this thesis were performed by Miss Velda Davis who did an excellent job of preparing the final form and deserves a special note of thanks.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The arrangement of events by their order of occurrence in time describes a time series. Time series analysis is the planned study of the data associated with this sequence of events. In general, the purpose of a study is to provide information for the interpolation or extrapolation of some function of the observed series in order to obtain forecasts of intermediate or future events.

The generality of these definitions indicates the range of subject areas within time series analysis. Therefore, it is proposed to consider only that segment which is related to exponentially weighted moving averages and record the results of investigations in that area.

Contrasts in the History of Time Series Analysis

The astronomers of the Sixteenth Century are credited with originating the study of time series (1). These studies were conducted on the movement of the known planetary bodies within the solar system. It is fortunate for science in general and for the study of time series analysis in particular that the major influence upon the activity of the planets is the sun. The reason this is considered fortunate is that the behavior of an observed process is a function of all things which influence the process; therefore, if a dominating influence is present it usually reduces the complexity of the behavior pattern and facilitates a description of that process as a function of time. Astronomers have continued to perform time series analysis upon the activity of bodies within the universe and after 225 years of study they have a 175 page equation describing the movement of the Earth's moon. This example is extreme in the amount of emphasis it places upon patience, time, and precision, but it is illustrative of the procedural technique used throughout the development of this subject area in that data are collected and an analysis is performed to determine if a sufficient degree of consistency is demonstrated in the activity to permit the formulation of a description of the observed activity within some acceptable limits. Generally, the maturity of an area of study can be measured by the range of these limits. Since the range of the limits is quite small for astronomy, it is considered to be a mature science.

Another classical study in time series analysis is a study of sunspot data made by Wolfer about the beginning of this century. This study was particularly significant to the geophysicists of that time because Wolfer not only interpreted the pattern within the historical data, but also forecast their occurrence within reasonable limits. The isolation of these periodicities in the sunspot data created an interest in time series analysis not only in the study of sunspots but in many seemingly unrelated fields. It is ironic that similar studies of sunspot activity are presently being conducted by space scientists specifically for the purpose of forecasting such activity. This is considered a near necessity for scheduling manned space travel in the future since sunspots release large amounts of high-level radiation into outer space.

History does not relate a classical study in the field of meteorology, but is readily recognizable as a field that evolved from a mysterious curiosity into a mature science through time series analysis. Probably the first meteorologist was a mathematically inclined agriculturist that observed certain sequences in the weather and found that by recognizing the start of a particular pattern a certain type of weather activity would usually follow. The study of meteorology has advanced to the state that analyses of meteorological data are performed on an automated basis. An example of this is the computer complex operated by the Navy at San Diego, California, which routinely analyzes the data collected by a world-wide collection system and makes long range weather forecasts.

Another classical problem which is not immediately associated with time series analysis is that of the vibrating string. This is a case of a physical system that receives a known input and the response or output is observed and studied until a mathematical expression is arrived at that describes the response adequately. A solution is known to exist in terms of differential equations or harmonic analysis. This type of analysis is involved in communication engineering and statistics where the input to a system is known, and the output is predictable, but due to external sources of noise the problem becomes one of recognizing the exact points of correspondence between the input and output. This class of problems is undergoing intensive study at the present time due to the need to establish accurately the position and velocity of missiles in outer space.

Time series do not lend themselves to an exact classification system. If the term stability is used as a measure of the probability that the future of the process can be forecast in time, then an exact delineation between stable and unstable processes is not possible. Therefore, only qualitative definitions can be given for stable and unstable processes. In this thesis the unstable processes are those which have not been quantified to the extent that their description is commensurate with the reasonable use of that description. The classification of processes is one of judgment based upon the mathematical description of the process and the use of the extrapolated series.

The foremost example of unstable processes is that large class of processes associated with economic time series. Economic time series have been studied since the early Nineteenth Century and at the present time the amount of effort expended in this area continues to increase. Progress has been made, but the standard economic time series are still judged to be unstable processes. Many attempts were made to correlate economic time series with Wolfer's sunspot data, and suprisingly enough many of the trials showed a significantly high degree of correlation. This led to a frenzy of activity in finding correlations among various economic time series or any other time series that would provide a significant correlation. Published results of this nature attracted the attention of many mathematicians and scientists since these correlations could not be justified by any practical means on an a priori basis. This precipitated a number of mathematical papers on the subject of correlation and extrapolation of time series. The classics are those by Yule (2) in 1926 on why nonsense correlations exist and by Slutzky (3) in 1927 on random series. These provided the base for serious study of the unstable processes and the development of statistical techniques to provide better forecasts of the future behavior of these processes.

A majority of the studies of unstable processes are related to economic time series, but there are two other areas that are sufficiently large and should be mentioned. They are sociological and biological time series. The former is the study of people and their activity in terms of births, marriages, divorces, suicides, thefts, and any number of other such categories. The latter is the study of changes in the population of insects, fish, game animals, birds, plant and animal growth, and related topics. There are undoubtedly other logical divisions of the study within the stable and unstable processes that could be made, but those listed are considered to be the major ones.

A more detailed discussion of the above may be found by Davis (1), as well as in the introductory material in many of the texts listed in the Bibliography.

Methods of Time Series Analysis

The mathematics used in the analysis of time series vary as much as the areas of investigation. During the last 200 years a number of mathematical techniques have been developed as a direct result of time series investigations. Other standard mathematical methods and specialized methods of analysis from other fields of study have been applied to the study of time series. The following discussion will list the more commonplace techniques that have been used in developing time series study and those that are being used at the present time.

In general these methods are curve fitting techniques used as a base for extrapolation. The process of fitting a curve to the observed data is one of smoothing the observed series and interpolating for values within the series. The simple forms of curve fitting are those used for trend analyses. One procedure for this is to take the average value of all the observed results which may be done either arithemetically or graphically by freehand methods. Similarly, determination of linear, exponential, logarithmic, moving average or moving polynomial trends can by accomplished by least squares techniques or graphical methods. Special cases of these methods are the linear regression and curvilinear regression methods of curve fitting and row experiment type of analysis of variance models. Generally, the above methods tend to smooth the data more than some time series models, but also tend to provide better estimates of the mean, which may or may not be a proper mode for comparisons of methods.

One curve fitting method that can provide an exact fit of the observed data is the interpolating polynomial. The degree of the polynomial will determine generally the fit of the curve with the exact fit obtained if the degree of the curve is one less than the number of points being considered. However, this exactness is not indicative of the exactness of extrapolated values of such a polynomial. In fact, n^{th} degree polynomials of this type generally cannot predict beyond the next few values in a series as well as some of the more general smoothing techniques. Good estimates of the original series may also be obtained by using techniques of harmonic analysis. Two of the more important techniques of this type are the use of a Fourier series to fit the data within an arbitrarily small amount of error and a method of analysis known as periodgram analysis which has received considerable attention as a tool for economic time series analysis. There are a number of variations in the details of the methods of periodgram analysis and some consideration should be given to the selection of the method to be used.[2]

Correlation in observed data is a rather important consideration in time series analysis. Autocorrelation is a special type of correlation which quantifies the relationship of values within a series to other observed values in that same series. Also, correlation among different time series may be determined by direct, lag or inverse correlation. The lack of autocorrelation is important in many of the statistical treatments of time series analysis, not only within the observed data but also within the time series of errors that are generated by a lack of fit of the smoothed data.

The work of Slutzky (3) provided the fundamental concept of the moving average technique which was based upon the observed statistical properties that cumulative sums or moving averages demonstrated. There have been a number of investigations regarding the length of the interval to use for a moving average and various methods of weighting the observed data to obtain a "best" estimate of the future mean value of a process. A notable example is a weighting formula developed by Macaulay (4) which contains 43 terms. Less detailed methods will be discussed later in the thesis.

---

[2]Refer to Davis (1), Chapter 7, pp. 276-326.

There are numerous other types of studies or techniques that are associated with time series analysis and are found frequently in the literature on this subject. These include cofluence analysis, factor analysis, the variate difference method, stochastic difference equations, orthogonal functions, operational calculus methods, methods derived from the calculus of variations and others which appear less frequent. Some of the related areas of study are spectral theory, ergodic theory, stochastic processes, Brownian motion, communication theory, filtering processes, random series and servomechanisms.

Interest in Time Series Analysis

An indication of the interest in a subject may be determined by a review of publications in that area. In time series analysis, the individual contributions to the literature number in tens of thousands. The Selected Bibliography compiled by Deming (5) for the period 1930 to 1957 from the mathematical and statistical journals contains over 240 entries.[3] In addition to this, there were a number of publications on this subject in the 130 years prior to this time period, in the five years subsequent to it, and in the works of economists, sociologists, biologists, electrical and communication engineers, and applied

---

[3]These include Annals of Mathematical Statistics, Journal of American Statistical Association, Biometrika, Journal of the Royal Statistical Society, and a number of foreign publications.

statisticians.

For example, the economists have a number of journals which contain papers on the subject of economic time series. The leading journal from the standpoint of statistical treatment of this subject is Econometrica, but among the others which publish time series studies are Economica, Economic Journal, The Economist, Review of Economics and Statistics, and Harvard Business Review. In addition to these periodical publications in the field of economics, there are at least three major foundations that study the general field of economics with frequent studies in the area of economic time series. They are Harvard Economic Studies, the Netherlands Economic Institute and Cowles Commission of Research in Economics, who sponsor the preparation of monographs for publication and distribution through their individual monograph series.

Each of the other disciplines of study have similar groups of periodicals that publish their works in the study of time series, with two periodicals that are more interdisciplinary in that they encourage the contribution of any person interested in the study of time series and more specifically isolation of periodicities within historical data. These two journals are Cycles and The Journal of Cycle Research. Those journals which contained articles more closely related to this thesis are Management Science and Journal of Operations Research Society of America.

Assuming that publication rate is a suitable index of interest, then it is obvious that there is a widespread and continuing interest in time series analysis.

Scope of This Study

Within this complex of methods of analysis and various disciplines there are gaps that should be filled between the mathematical-statistical approach to the theoretical time series and the less sophisticated methods of approximation used to study real world data. Figure 1 is a graphical display of that part of time series analysis that will be considered in this thesis.

This particular area is generally associated with inventory, sales, personnel action, maintenance and market value. The purpose of this thesis is to present methods for improving the forecasts of exponentially weighted time series models by selection of smoothing constants. The selection procedure is based upon measures of optimality that are formulated in terms of observed forecast error. The thesis is directed toward the application of the methods derived to real world processes. The connotations associated with the use of real world to describe processes include: their exact future behavior cannot be known under any conditions, they are generally of the unstable type, and are assumed to be influenced by a number of variable external factors. The external influences are assumed to combine to form the basic driving mechanism of the series. Also, any one single factor does not have a significant

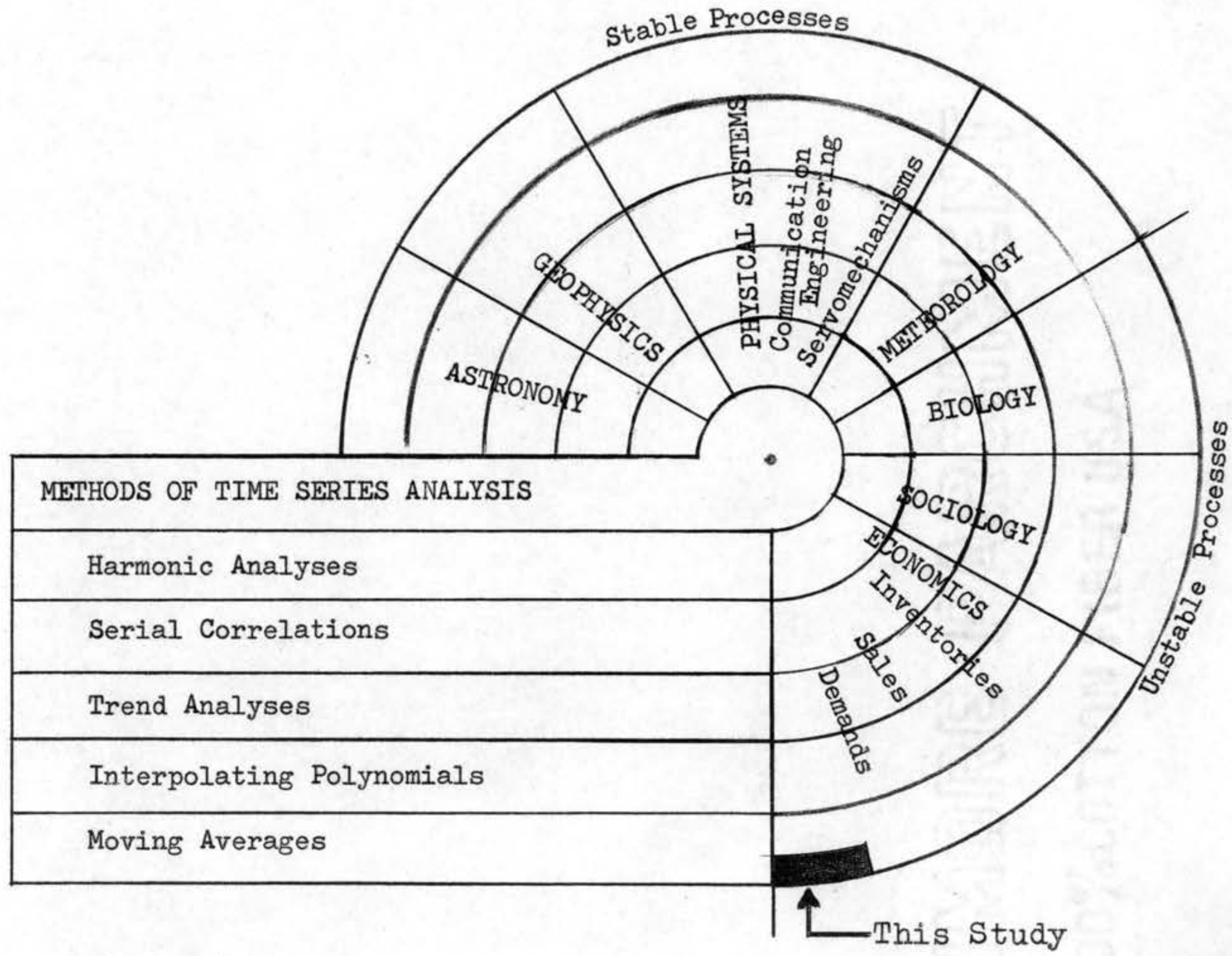

Figure 1. The Time Series Spectra of Subjects and Analyses

influence upon that series. This is one of the few assumptions that will be made in this study and it does not appear to detract from the generality of this thesis.

The method of analysis to be used is a trial and error procedure involving the use of a partially self-generated mathematical model which uses the basic technique of exponentially weighted moving averages for smoothing and interpolation of the observed time series, and extrapolation of the smoothed time series.

This basic approach has been used more recently by Magee (6), Brown (7), and Holt, Modigliani, Muth, and Simon (8) with Brown presenting the basic philosophy of exponential smoothing under varying degrees of instability. Holt, et al., modified the formulation of the mathematical model, but did not study the properties of the model under varying conditions of the unstable processes. This thesis will attempt to extend some parts of the referenced works and present original ideas in the model formulation and discussion of results.

Almost without exception, every author that discusses the real world processes and the mathematical model approach for forecasting points out that better methods are needed and that the best test of these methods is the results that they produce. Therefore, this thesis is not unlike the original studies conducted by astronomers and the goal is similar, to forecast the future.

CHAPTER II

THE EXPONENTIALLY WEIGHTED MOVING AVERAGE AS A TIME SERIES MODEL

Before describing a specific model, perhaps something should be said about models in general. A mathematical model is defined here as a formulation of relationships that is believed to describe some part of a real world process sufficiently well for use in the study of that process. A basic reason for developing a model is the belief that the process has some pattern of behavior. The model is used to extrapolate the pattern. The model is often mathematically simple, but the computations necessary for reducing the data may become complex. The computations that provide the input data for those models developed will be referred to as supporting algorithms.

Some of the general philosophies on model construction that are followed throughout this thesis are that few assumptions are made, that the economy of the model is considered by keeping the terms in the model to a minimum, and that a simple model is better than no model at all.

The supporting algorithms play an important part in the models that are developed below. The basic function of the supporting algorithms is one of smoothing various components of the historical data for use in the model. Smoothing is defined as a process of algebraic curve fitting which minimizes some function of the deviations about a local process level.

In general, it is the purpose of this thesis to develop the theory and the supporting algorithms of a general time series model for unstable processes. The objective is to extrapolate the time series to concur with the true time series as nearly as possible based upon measures used to evaluate the degree of deviation.

The true time series model is not known, but is assumed to exist for each process considered and if known it would describe the process exactly for all historical data and future occurrences.

As indicated in Figure 1 (page 12), the emphasis of this thesis will be directed toward building a model that uses exponentially weighted moving averages as estimates of the model's parameters. Due to the generality of the area of consideration, the basic mathematical theory that will be applied is that of smoothing. Some of the mathematics used will necessarily be supported by heuristic arguments.

There are some general philosophies in the study of time series that will be adhered to in this thesis. One of these is in specific reference to the model building approach and it is basically that the worth of the model is in how well it forecasts future activity. This will be the basic premise for the development of the complete model. This is necessary due to the generality of the time series that are to be considered and the fact that the conditions for application of rigorous statistical techniques cannot be satisfied. For the unstable type of time series that will be studied, it is assumed that all available information about the process is contained within the historical data. This does not require that the series be stationary, the random fluctuations be independent or the forecast errors be independent. One or more of these common restrictions in time series studies often remove many of the real world processes from consideration. However, whatever the influences are in the observed series, they are assumed to continue throughout the period for which the forecast is made.

Another general philosophy that is prevalent in the study of economic time series is that the series is composed of a trend, a systematic cyclical variation and a random variation. A combination of the first two components of variation may be referred to as the base series, driving mechanism, or signal, whereas, the latter may be termed noise. Noise is that part of the series which is not accounted for by the assumed model and may include both the inherent process error and the error due to lack of fit. However, these three sources of variation in an observed time series should be understood to be the culmination of numerous known and unknown influences. The purpose is not to isolate the three components for study on an individual basis, but is to combine them in order to extrapolate a given time series.

The Moving Average as a Model

The exponentially weighted moving average model was originated by studies of the cyclical phenomena demonstrated by moving averages or moving sums. The first notable study of this type was by Slutzky (3) in which he observed that for a series of random events or non-random events, that a moving sum containing some fixed number of observations would produce a smoothed series that was non-random as determined by the autocorrelation coefficient. In this work, he also stated and proved the law of the sinusoidal limit which states simply that if a series is smoothed a sufficient number of times by moving sums that it approaches a sinusoid within some arbitrary amount of error. This sinusoid does not necessarily represent the base series of the process. Dodd (9) has shown that misleading artificial cyclical variations may be created. Studies of the moving average and variations of it have been published by a number of persons. The work by Brown (7) provides the background and a base for the summary that will be presented here.

The moving average, as its name implies, is an average of observations of a time series which uses only the most recent n terms for purposes of computing an average. This average is considered to be the most recent estimate of the mean of the process being studied. Mathematically this idea may be expressed as

$$\bar{X}_{mov} = \sum_{i=t-n+1}^{t} X_i/n$$

where t is the latest point in time. The major problem associated with the moving average type of time series model is that of deciding how many observations are to be used in the moving average. The number that will provide the best results in terms of having the least amount of error associated with its forecasts of the mean of the process is largely dependent upon the process being studied. There are two extreme cases of the moving average that are of interest, one case includes all historical data in the estimate of the mean and the other takes only the last observed value as an estimate of the present level of activity. Either of the cases might work under certain circumstances. The former would be suitable when the process is in fact operating about some mean with only random fluctuations about that mean. Under this assumption the inclusion of all historical data in the average will provide the best estimate of the mean value of the process since the average is an unbiased estimator of the process mean. The latter case would be more accurate in a time series that has a high autocorrelation for a lag of one time interval. Within the range spanned by these two cases, there are a number of processes and values of n.

If a process has only random variations about a slowly changing mean or if the magnitude of the fluctuations or variations is large compared to the mean, then a large value of n is desirable for the moving average. This is because large values of n will tend to smooth the random fluctuations and not be influenced to any extent by large deviations in one direction. If it is desired to smooth some of the non-random variations in a process when they are known to exist, a large n equal to the period of the variations that are to be removed may be used. For example, if it is desired to remove the cyclical variation within the year from the smoothed statistics of the time series, then a twelve month moving average will provide this type of smoothing. To demonstrate this, Figure 2a indicates the effect of a twelve month moving average upon a sine wave with a period of twelve months. This holds for any systematic pattern of variation as long as the pattern is repeated in each period.

The size of n is not only dependent upon the type of smoothing desired, but it is also dependent upon the associated problem of the type of response that is desired for some corresponding change in the data. The interest in the response may arise for a number of reasons. One reason is that the mean of the process may undergo some change and shift to a new level, and if n is large, only 1/n of the shift will be added to the estimate of the level each time and it will require n periods before the moving

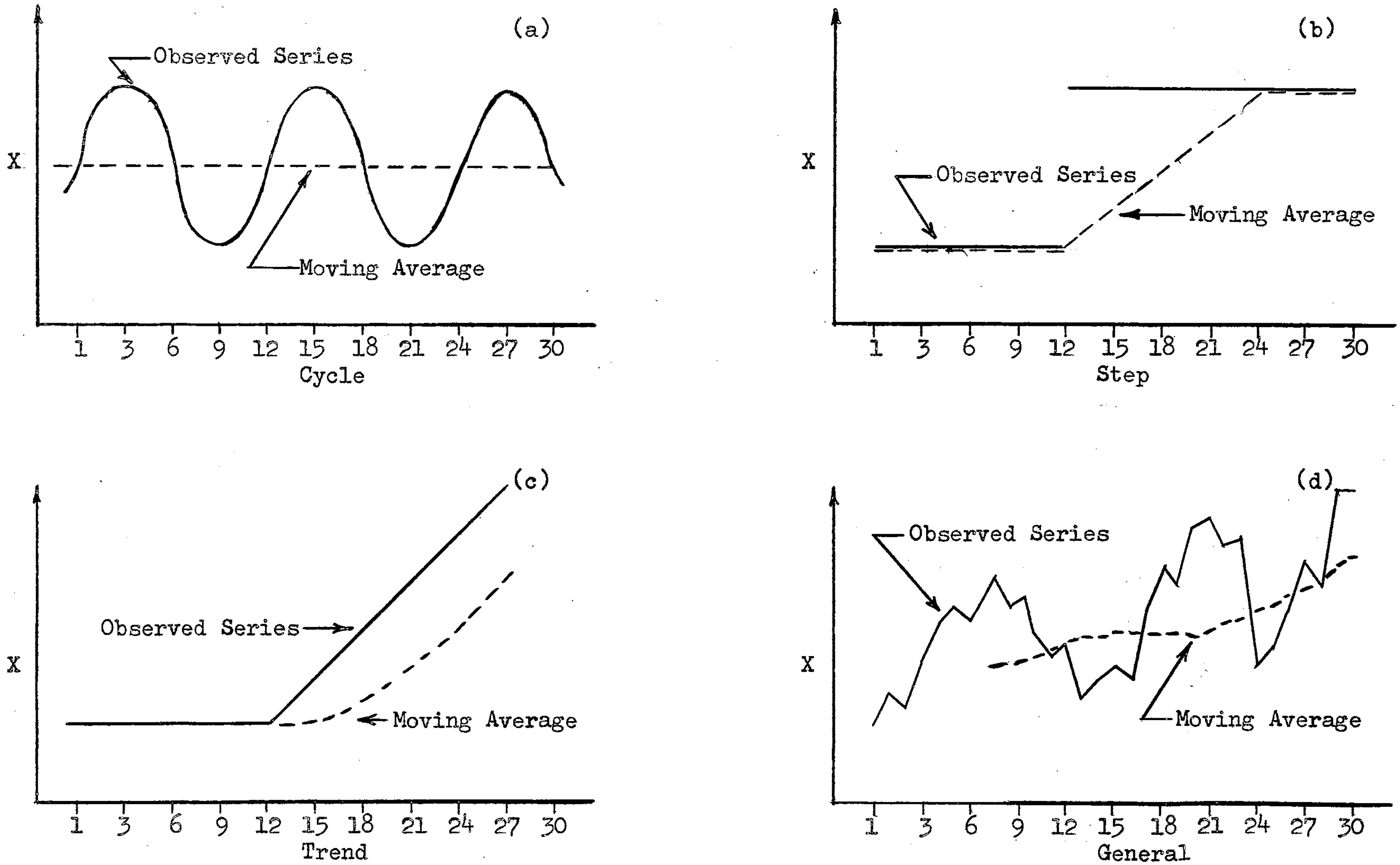

Figure 2. Response of the Twelve Month Moving Average

average has reached the new process level. This type of lag for the moving average is illustrated in Figure 2b. If the time series should experience a trend in the observations, then the moving average will always lag the observations by n/2 times the increase per time interval of the trend. This is based upon the assumption that the conventional procedure is followed in using the average to represent the point in the series at t-n/2. A typical lag is shown in Figure 2c. If the time series is composed of a trend, a cyclical variation and a random variation as illustrated by Figure 2d, then the selection of the n for the moving average is usually by some trial and error procedure. This is simply due to the mathematics involved, in that it is difficult for a simple moving average to adapt to complex situations. If an exceptionally good fit were obtained, it would be attributable to chance as opposed to the tracking capability of the model since the moving average is not a self-correcting type of mathematical procedure. This means that the magnitude or the direction of the error observed in the fit of the smoothed series to the actual series does not influence the next value computed for the smoothed series.

There are a number of improvements that can be made in the moving average that will correct for the various situations outlined above such as weighting the values included in the average by the distance they are removed from the central value and adding an increment to allow for the trend. It should also be noted that the moving average is more of a smoothing technique than an extrapolation technique, since the value of the moving average is generally associated with the mid-value of the averaged values, which means the last estimate of the mean of the process is still n/2 time units removed from the present time or the last observed value. There have been corrections devised for this deficiency also. There are a number of published articles on methods of compensating for cycles and trends in the use of the moving average. Brown (7) provides a treatment of some of these methods and the application of the moving average. However, the exponential weighting of historical data has replaced the use of the moving average to a large extent. This is due to the ease of computation, less data storage required, and the self-correction feature of exponential weighting.

The Simple Exponential Model

The moving average will be considered as a basis for comparison in terms of the error for the forecasting models to be developed. Simple smoothing or the simple exponential model is only a slightly advanced form of the moving average. The main difference is that it is self-correcting or has the ability to adjust based upon its observed errors. It is convenient at this point to establish some notation that will be followed throughout this thesis. Let

X_t be an observation at the t^{th} point in time where the points are equally spaced,

M be the total number of observations available for a given time series,

$\bar{X}_t$ be the exponentially smoothed value of the observed time series at time t,

$\bar{X}_o$ be a computed value that is used as a starting point in recurrence type computations,

A be the smoothing constant for random variations in the observed series.

T be some number of time intervals into the future,

$FX_{t,T}$ be the forecast at time t of the expected level of process activity, T intervals hence.

The constant A is sometimes referred to as a weighting factor or an attenuation factor as well as a smoothing constant, the reasons for this triple identity will become more obvious in the discussion that follows.

The smoothing performed at time t in order to obtain $\bar{X}_t$ is accomplished by adding to the previous estimate of the process mean, $\bar{X}_{t-1}$, some fraction, A, of the forecast error, $(X_t - \bar{X}_{t-1})$. This becomes the latest estimate of the mean and the forecast $FX_{t,T}$ until some new information is added to the system. Algebraically, the model may be written:

$$FX_{t,T} = \bar{X}_t \tag{1}$$

with the supporting algorithm as

$$\overline{X}_t = \overline{X}_{t-1} + A(X_t - \overline{X}_{t-1}). \qquad 0 \leq A \leq 1 \qquad (2)$$

Equation (2) may be rewritten in a more convenient form as

$$\overline{X}_t = AX_t + (1 - A)\overline{X}_{t-1}, \qquad 0 \leq A \leq 1 \qquad (2a)$$

and

$$\overline{X}_{t-1} = AX_{t-1} + (1 - A)\overline{X}_{t-2}, \qquad 0 \leq A \leq 1 \qquad (3)$$

which establishes the form of the recurrence relation. If $X_o = \overline{X}_o$ and X_o is assumed to be of the same form as X_i, then by successive substitution into Equation (2)

$$\overline{X}_t = A \sum_{i=0}^{t} (1 - A)^{t-i} X_i, \qquad 0 \leq A \leq 1 \qquad (4)$$

is obtained as the most recent estimate of the mean of the process. If $\overline{X}_o$ is not considered to be of the same form as X_i, then Equation (4) may be written

$$\overline{X}_t = A \sum_{i=1}^{t} (1-A)^{t-i} X_i + (1-A)^t \overline{X}_o, \quad 0 \leq A \leq 1 \qquad (4a)$$

where t = M. As M becomes large, the last term of Equation (4a) approaches zero. Examination of the coefficients of X_i in either Equation (4) or (4a) will show that their sum is a geometric progression. If it is assumed for a finite value of t that this geometric progression is equal to 1/A, then the sum of the weights applied to the historical data equals one. If the sum of the weights is equal to unity, it is proper to call $\overline{X}_t$ an average. The term $(1 - A)^{t-i}$ describes the system of weights and adds the exponentially weighted portion to the descriptive title of the model. The term moving is attributed to the fact that $\bar{X}_t$ is computed each time a new observation in the series is available, thus completing the basis for the title, exponentially weighted moving average.

Since the sum of the expected values is equal to the expected value of the sum and under the assumptions that the mean of the X_i series is stationary and that the series $\sum_{i=1}^{t} (1 - A)^{t-i}$ converges to $1/A$, the

$$E(\bar{X}_t) = E(X)A \sum_{i=0}^{t} (1 - A)^{t-i} \doteq E(X)A(1/A) = E(X), \text{ for } 0 < A < 1. \tag{5}$$

Therefore, $\bar{X}_t$ is shown to be an approximation to the unbiased estimate of the mean where the degree of approximation is dependent upon A and t.

From the discussion of the simple moving average, it is noted that the response of the model to changes in the process is dependent upon n, the sample size. Similarly, in the simple exponential model, the response of the model to changes in the process is dependent upon the selection of A. One extreme case is encountered when $A = 0$, which implies that $\bar{X}_0$ is the best estimate of the future behavior of the process and that all observed data are random variations about this level. This would not be unusual provided that the process was quite stable and a sufficient amount of information had been used in determining the value of $\bar{X}_0$. The other extreme case, $A = 1$, would be applicable in the same type situation as n = 1 was in the simple moving average model. This occurs when there is a high degree of autocorrelation between successive values of X_i. For those time series in the range that is spanned by these two cases, it is necessary to be able to choose some value of A that will minimize some function of the forecast error. The magnitude of A determines the influence of the historical data upon the most recent estimate of the process level. The relationship between A and the weight given to historical data is shown in Figure 3. The ordinate values in the uppermost logarithmic cycle are (1 - A).

This graph may be used to show comparisons between A and n. However, it should be realized that the simple moving average uses equal weights for each X_i contained in the sample n, where those same data in the simple exponential model are exponentially weighted. The ordinate values in Figure 3 represent that fraction of the weight unassigned. Thus, the ordinate value of the intersection of an integral abscissa value and one of the lines for A equal a constant is the total weight given to the historical data whose age is equal to or greater than the abscissa value. Therefore, in order to determine the total weight given to the last five months of historical data, it is necessary to compute one minus the value of the ordinate corresponding to (t - 6).

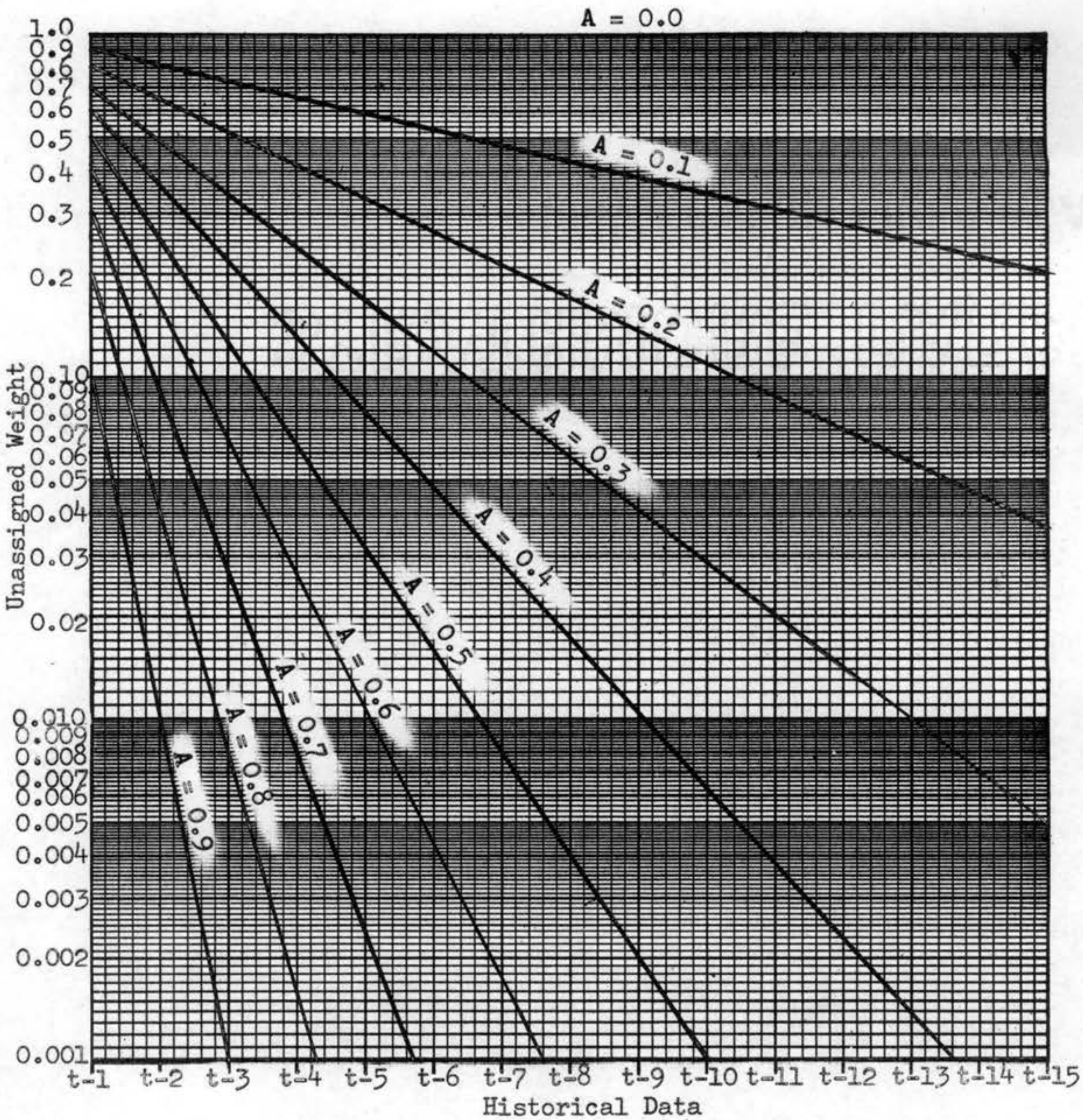

Figure 3. Attenuation of Historical Data Based Upon Choice of Smoothing Constant

As an example, if it is desired to determine the relation between a five month simple moving average and a smoothing constant of 0.5, it may be accomplished by moving vertically from an abscissa value of t-6 until the line for A = 0.5 is intersected. The ordinate value corresponding to this point is 0.016; therefore, 98.4% of the total weight has been assigned to the first five data points under the exponential smoothing scheme. If it is desired to make a number of comparisons, then some general level of unassigned weight should be specified in order that the comparisons rank properly. Since the time scale is discrete, an exact value for the unassigned weight cannot be specified.

Even though these comparisons between the moving average and simple exponential smoothing are possible, the response of the simple exponential model to change is more sensitive if the unassigned weight for the comparison is less than 0.5. Both the simple moving average and simple exponential model are more suited for smoothing the process that is relatively stable about some level with the variation attributable to random variation as opposed to those processes generated by a base series with systematic variations.

Any reasonable estimate of the mean of the process level may be used as $\bar{\bar{X}}_0$; however, if a computational form is desired the first P data points of the observed series may be averaged to obtain $\bar{X}_0$. The numerical value of P is arbitrary, but approximately one-tenth of the historical data should provide a reasonable value. Due to the exponential property discussed above, any reasonable value will be suitable since the effect that it can have on the more recent estimates of the process level will be extremely small. A detailed discussion of the selection of A will be presented in Chapter III.

The Model for Trend and Noise

If the observed time series is in fact composed of some random variations superimposed on a linear trend, then the simple exponential model will generally be less capable of smoothing the series than a model formulated specifically for this type of series.

The simple smoothing of a time series with linear trend tends to produce a lag of the type that is shown in Figure 2c (page 20). If the series does have a linear trend, a method of correction would be to determine the increment attributable to trend per time unit and add that to the estimate of the mean of the process in order to obtain the estimate of the next value in the observed series. Additional notation will be needed to describe this model. Let

Y_i be the time series of differences, $\overline{X}_i - \overline{X}_{i-1}$, which will be referred to as the trend increment series,

$\overline{Y}_t$ be the smoothed estimate of the trend increment at time t,

B be the smoothing constant associated with the variations due to the trend.

Using these definitions and those of the preceding section, the mathematical model for a time series composed of a linear trend component and a mean may be written:

$$FX_{t,T} = \bar{X}_t + T\bar{Y}_t \ . \tag{8}$$

The definition of $\bar{X}_t$ remains the same, but its function relative to the observed series, X_i, is changed. This is due to the fact that $\bar{X}_t$ does not include the estimate of the trend. However, the basic logic used to develop the smoothed estimates in the simple exponential model is applied in this model development. The smoothing is accomplished by making a forecast of the next value in the series, taking some fraction of the observed error between the actual and forecast values, and adding it to the last estimate of the mean in order to obtain the new estimate of the mean. From Equation (8), the forecast of X_t at time (t - 1) is

$$FX_{t-1,1} = \bar{X}_{t-1} + \bar{Y}_{t-1} \ . \tag{9}$$

Thus,

$$\bar{X}_t = A(X_t - \bar{X}_{t-1} - \bar{Y}_{t-1}) + \bar{X}_{t-1} + \bar{Y}_{t-1}, \quad 0 \leq A \leq 1 \tag{10}$$

$$= AX_t + (1 - A)(\bar{X}_{t-1} + \bar{Y}_{t-1}), \quad 0 \leq A \leq 1. \tag{11}$$

Applying the same smoothing techniques to the artificial variable Y_i at time t yields

$$\bar{Y}_t = B(Y_t - \bar{Y}_{t-1}) + \bar{Y}_{t-1}, \qquad 0 \leq B \leq 1 \qquad (12)$$

$$= B(\bar{X}_t - \bar{X}_{t-1} - \bar{Y}_{t-1}) + \bar{Y}_{t-1}, \qquad 0 \leq B \leq 1 \qquad (13)$$

by definition of Y_t. The Y_i variable was created as a notational convenience. The variation in the Y_i series is assumed to be a base series of the linear trend increment with some noise superimposed. The $\bar{Y}_t$ is essentially a double smoothing of the original time series. To more clearly show this, it is necessary to break Equation (13) down by substituting Equation (11) for $\bar{X}_t$ to obtain

$$\bar{Y}_t = B\left[AX_t + (1-A)(\bar{X}_{t-1} + \bar{Y}_{t-1}) - \bar{X}_{t-1} - \bar{Y}_{t-1}\right] + \bar{Y}_{t-1},$$

$$\begin{cases} 0 \leq A \leq 1 \\ 0 \leq B \leq 1 \end{cases} \qquad (14)$$

$$= B(AX_t - A\bar{X}_{t-1} - A\bar{Y}_{t-1}) + \bar{Y}_{t-1}$$

$$= BA(X_t - \bar{X}_{t-1} - \bar{Y}_{t-1}) + \bar{Y}_{t-1}$$

$$= BA(X_t - \bar{X}_{t-1}) + (1 - BA)\bar{Y}_{t-1}. \qquad (15)$$

By using the recurrence relation of Equation (11), $\bar{X}_t$ is shown to be a function of all historical data and the smoothing constants.

$$\bar{X}_t = A\sum_{i=1}^{t}(1-A)^{t-i}X_i + \sum_{i=1}^{t-1}(1-A)^{t-i}\bar{Y}_i + (1-A)^t(\bar{X}_o + \bar{Y}_o) \qquad (16)$$

This is of the same form as Equation (4a) and may be reduced to the form of Equation (4) by making simplifying assumptions about $\bar{X}_o$ and $\bar{Y}_o$.

The $\bar{Y}_t$ is an exponentially weighted moving average also and may be expressed as a function of historical data and the smoothing constant by rewriting Equation (12) as

$$\bar{Y}_t = BY_t + (1 - B)\bar{Y}_{t-1}, \qquad 0 \leq B \leq 1 \qquad (17)$$

which is of the same form as the recurrence relation for simple exponential smoothing where

$$\bar{Y}_t = B\sum_{i=1}^{t}(1-B)^{t-i}Y_i + (1-B)^t\bar{Y}_o, \quad 0 \leq B < 1 . \qquad (18)$$

The $\bar{Y}_t$ may be written in terms of the original time series by application of the recurrence relation of Equation (15).

$$\bar{Y}_t = BA\sum_{i=1}^{t}(1-BA)^{t-i}(X_i - \bar{X}_{i-1}) + (1-BA)^t\bar{Y}_o ,$$

$$0 \leq A < 1, \ 0 \leq B < 1 . \ (19)$$

Then $\bar{X}_t$ may be written in terms of the observed time series by substituting Equation (19) in Equation (16) to obtain

$$\bar{X}_t = A\sum_{i=1}^{t}(1-A)^{t-i}X_i + \sum_{i=1}^{t-1}(1-A)^{t-i}BA\sum_{j=1}^{i}(1-BA)^{i-j} \cdot$$

$$\cdot (X_j - \bar{X}_{j-1}) + (1-BA)^t\bar{Y}_o + (1-A)^t(\bar{X}_o + \bar{Y}_o),$$

$$0 \leq A < 1, \ 0 \leq B < 1 . \ (20)$$

The relationship of the smoothing constant A to the historical data from the standpoint of attenuation rate remains relatively unchanged, and from Equation (17) it would appear that the relationship between B and Y_i is similar. However, from Equation (19) it is obvious that the historical data of the original series has less overall influence, but also has a slower attenuation rate than a corresponding simple exponential smoothing process. This is a result of the product, AB, being less than A or B, since it is unlikely that for a time series composed of a trend and random variations that both A and B would be large simultaneously. However, $\overline{\overline{X}}_t$ and $\overline{Y}_t$ each possesses those properties of the exponentially weighted moving average provided A, B, and t are in the proper perspective. The mathematical model for a time series composed of a trend and some random variation is the sum of two exponentially weighted moving averages which use double smoothing for interpolating the observed time series.

Figure 3 (page 27) is applicable for A, B, or AB, and it is not necessary to repeat the discussion of the response of the model based upon the choice of these constants. The influence and selection of the smoothing constants will be discussed at length in Chapter III.

The value of $\overline{\overline{X}}_o$ is determined by the method outlined in the preceding section, and

$$\overline{\overline{Y}}_o = (\overline{\overline{X}}_o - \overline{\overline{X}}_e)/(t - P) \qquad (21)$$

where $\overline{\overline{X}}_e$ is the average of the last P observations. This is one simple method for computing the initial values, and is not to be considered a rule. Any reasonable estimate of the average trend per interval of time between observations over the range of the historical data will be suitable as a value for $\overline{Y}_o$. Like $\overline{X}_o$, $\overline{Y}_o$ is attenuated in the computational scheme.

The Complete Model for Trend, Cycles and Noise

One of the assumptions stated at the beginning of this chapter was that the general class of economic time series is composed of a trend, a cyclical variation and a random variation about some mean level. Thus, the formulation of a general model of the exponential smoothing type for the general time series completes the model building phase.

If the series includes a cyclical variation, the model developed in the previous section will attempt to track the signal by the trend factor compensating for the cyclical variation to some extent. As would be expected, the fit of the model to the observed series would not be as accurate as a model which includes direct consideration of the cyclical component.

The problem of formulating a general model is complicated at this point by the fact that cyclical variations have two common forms. The simpler cycle is assumed to be independent of the local process mean and its amplitude is not a function of the level at which the process is operating. The other general type of cyclical variation is assumed to be dependent upon the local mean and as it increases, the cyclical variations also increase as a ratio.

These two types of cyclical variations are referred to as the additive and multiplicative or ratio cycles. Both types will be considered in this section, but in general the remainder of this thesis will deal with only the ratio type, the reason being that the unstable processes are felt to be more inclined to this type of behavior. The additional definitions given in this section will also reflect the ratio type cycles. Let

Z_i be the artificial variable formed by $X_i/\overline{X}_i$,

$\overline{Z}_t$ be the smoothed estimate of the cyclical ratio at time t,

P be the period of the cyclical variation,

C be the smoothing constant for the cyclical variations.

If the cyclical variation was of the additive type, the supporting algorithms could be written as shown below where Z_i would be the artificial variable $X_i - \overline{X}_i$.

$$\overline{X}_t = A(X_t - \overline{X}_{t-1} - \overline{Y}_{t-1} - \overline{Z}_{t-P}) + \overline{X}_{t-1} + \overline{Y}_{t-1} + \overline{Z}_{t-P}, \quad 0 \leq A \leq 1 \qquad (22)$$

$$\overline{Y}_t = B(\overline{X}_t - \overline{X}_{t-1}) + (1 - B)\overline{Y}_{t-1} \qquad 0 \leq B \leq 1 \qquad (23)$$

$$\overline{Z}_t = CZ_t + (1 - C)\overline{Z}_{t-P} \qquad 0 \leq C \leq 1 . \qquad (24)$$

Since $\overline{Z}_t$ is of the simple exponential form, the model would still be of the exponentially weighted moving average type and would be written as

$$FX_{t,T} = \overline{X}_t + T\overline{Y}_t + \overline{Z}_{t-P+T} . \qquad (25)$$

Before discussing the ratio model, the definition of $\overline{X}_t$ should be reviewed in order to establish the proper perspective. The X_i in this model may be thought of as the quotient of the observed series and the level of cyclical variation at that point in time, or more simply as a ratio. Thus, $\overline{X}_t$ is the smoothed estimate of the mean ratio after the cyclical ratio and the trend have been removed. The trend, however, is determined by the difference of the successive values of $\overline{X}_t$, and it too becomes an estimate of the trend in terms of a ratio. The model for these assumptions is written as

$$FX_{t,T} = (\overline{X}_t + T\overline{Y}_t)\overline{Z}_{t-P+T}. \tag{26}$$

The forecast at time, t-1, for time t, is

$$FX_{t-1,1} = (\overline{X}_{t-1} + \overline{Y}_{t-1})\overline{Z}_{t-P} \tag{27}$$

and is a forecast of X_t and not the ratio $\overline{X}_t$. Therefore, the smoothing algorithm is written

$$\overline{X}_t\overline{Z}_{t-P} = A(X_t - \overline{X}_{t-1} + \overline{Y}_{t-1}\,\overline{Z}_{t-p}) + (\overline{X}_{t-1} + \overline{Y}_{t-1})\overline{Z}_{t-P}\ .$$

$$0 \leq A \leq 1 \tag{28}$$

By dividing both sides of the equation by $\overline{Z}_{t-P}$, and rearranging it gives:

$$\overline{X}_t = AX_t/\overline{Z}_{t-P} + (1-A)(\overline{X}_{t-1} + \overline{Y}_{t-1}). \quad 0 \leq A \leq 1 \tag{29}$$

Notationally, $\overline{Y}_t$ is unchanged from the previous form.

$$\overline{Y}_t = B(\overline{X}_t - \overline{X}_{t-1}) + (1-B)\overline{Y}_{t-1} \qquad 0 \leq B \leq 1 \tag{30}$$

$$\overline{Z}_t = C(Z_t - \overline{Z}_{t-P}) + \overline{Z}_{t-P} \qquad 0 \leq C \leq 1 \tag{31}$$

$$= CZ_t + (1-C)\overline{Z}_{t-P} \qquad 0 \leq C \leq 1 \tag{32}$$

$$\overline{Z}_t = CX_t/\overline{X}_t + (1 - C)\,\overline{Z}_{t-P}\,. \qquad 0 \leq C \leq 1 \qquad (33)$$

The index of $\overline{Z}_i$ is not the same as $\overline{X}_i$ due to the fact that each time interval within the cyclic period is assumed to be different, and the smoothed estimate of the ratio is adjusted only once each period, therefore, there is a lag of P time units in the smoothing process. This causes the attenuation of the historical values of Z_i to be at a slower rate than X_i and Y_i, but the estimate of $\overline{Z}_t$ is still a function of the historical data and the smoothing constant, as shown by the series

$$\overline{Z}_t = C\sum_{i=1}^{[t/P]} (1 - C)^{[t/P]-1}\,\frac{X_{t-[t/P-i]P}}{\overline{X}_{t-[t/P-i]P}}$$

$$+ \sum_{i=t-[t/P]P}^{0} (1 - C)^{[t/P]}\,\overline{Z}_i \qquad 0 \leq C < 1 \qquad (34)$$

where the [ ] designates the use of the next largest integer for the expression contained inside the brackets. The last term in Equation (34) represents the initial values of the cyclic ratio, and will be shown to be a function of the observed series also.

The series expansion of $\overline{Y}_t$ remains as given in Equation (19).

By applying the recurrence relation in Equation (29), and substituting Equation (34) for $\overline{Z}_{t-P}$, $\overline{X}_t$ may be written as

$$\overline{X}_t = A\sum_{i=1}^{t} \frac{(1-A)^{t-i}\, X_i}{C\sum_{j=1}^{[i/P]}(1-C)^{[i/P]-1}\frac{X_{i-[i/P-j]P}}{\overline{X}_{i-[i/P-j]P}} + \sum_{j=i-[i/P]P}^{0}(1-C)^{[i/P]}\,\overline{Z}_j}$$

$$+ \sum_{i=1}^{t-1}(1-A)^{t-i}\, BA \sum_{j=1}^{i}(1-BA)^{i-j}(X_j - \overline{X}_{j-1}) + (1-BA)^t\overline{Y}_o$$

$$+ (1-A)^t(\overline{X}_o + \overline{Y}_o)\,. \quad 0 \leq A < 1,\ 0 \leq B < 1,\ 0 \leq C < 1 \qquad (35)$$

The values of $\overline{X}_o$ and $\overline{Y}_o$ are computed as given in the previous section. The reason it is not necessary to determine their value as a ratio is due to their being determined by taking averages of P values which was shown earlier to mask the effect of cyclical variations. For an average of the observed series over some interval P to equal the average of some ratio of the observed series to the cyclical component, the sum of the cyclic ratios must equal P, at least in the initial estimates. Therefore,

$$\sum_{i=1}^{P} \overline{Z}_i = P \qquad (36)$$

will represent the system of weights applied to the data within each cycle. The computational form for $\overline{Z}_i$ is quite simple for a given P, but the general form is a more complex looking group of indices and summations that follow. Let

$$Z'_i = \frac{X_i}{\sum\limits_{j=P([i/P]-1)+1}^{[i/P]P} X_{j/P} - \left\{\frac{P+1}{2} - i\left(\left[\frac{i}{P}\right] - \left[\frac{i-P}{P}\right]\right)\right\}\overline{Y}_0} \quad i = 1, 2, \ldots M$$

(37)

$$Z''_j = \sum_{i=0}^{\left[\frac{M}{P}\right]-2} \frac{Z'_{iP+j}}{\left(\left[\frac{M}{P}\right]-1\right)} \qquad j = 1,\ 2,\ \ldots,\ P \tag{38}$$

$$\overline{Z}_{j-12} = \frac{\sum_{j=1}^{P} Z''_j}{P} Z''_j \ . \qquad j = 1,\ 2,\ \ldots,\ P \tag{39}$$

Equation (39) satisfies the constraint of Equation (36).

With the initial values of each of the smoothed series computed and the supporting algorithms for the model, the model formulation is complete with the notable exception of the smoothing constants, A, B, and C. The considerations for proper selection of these constants will be the subject of Chapter III.

CHAPTER III

THE MODEL AND FORECAST ERROR

In some of the referenced works, the authors distinguished between predicting and forecasting future events. The difference is basically that forecasting is some statistical technique that extrapolates the historical data based upon some stated procedure. Prediction is an estimate of future events that is based upon the knowledge of the person doing the predicting. He evaluates all relative processes that may influence the future of the process under consideration and may or may not analyze the historical data. The number of people capable of accurate predictions of the future is exceedingly small, and it is becoming more routine to use forecasting for the unstable type of time series. The objective of forecasting is to provide more suitable descriptions of the future activity involving various unstable processes than could be obtained by some educated guess. Thus, the forecast error that may be associated with any forecasting process becomes an important factor in the continued use of a forecasting scheme. To reiterate some earlier statements in the thesis, the proof of the mathematical model is given by the results that it produces. Magee (6) points out that to forecast the future for an unstable type process is similar to using a crystal ball, and that a forecast without an associated statement of the possible magnitude of the error is incomplete. Also, there is the possibility that the forecast by itself will be accepted as fact which would be misleading.

Another objective is the consideration of the error associated with the forecast as a means of improving the parameters of the model used for forecasting. This is essentially an additional consideration of the error, since the exponentially weighted moving average uses the consideration of observed error as the basis for adjusting the forecast values. The techniques presented in this chapter as improvements of the complete forecasting model presented in Chapter II are based upon empirical studies of both real and simulated or artificial time series that were used for investigation of the model's properties. The formulation of the concepts as presented in this chapter will be supported by numerical examples in Chapter IV.

Assumptions

A basic philosophy in the study of time series is that it is always possible to define the generating process for the observed time series. This definition may include a rather large random component from some special distribution, but nevertheless it is not considered as being mathematically undefined. If the form of the generating process is known, then the nature of the analysis of the observed series is that of estimating the constant terms within the generating process for that particular series being studied.

The particular approach used for the time series model in this thesis is as stated previously in that no specific type of model for the generating process is assumed. However, a forecasting model is used which consists of trend, cyclical variation, and random variation components. If the generating series is not of this form, then the model is used to determine the best approximation of the series that is possible within the limitations of the model and the exponential smoothing method. If the use of a particular model in forecasting provides consistently good results, then it would be difficult to attribute this success to chance alone. However, if the generating process is of the form of the model and certain statistical assumptions are satisfied, then the results obtained by the methods used here agree with those of the curve fitting types. It has been shown by Brown and Meyer (10) that exponential smoothing provides the least squares estimate of the true polynomial signal provided that the data are of a true polynomial signal and an independent noise source with the noise distributed about a mean of zero.

Testing the Model

Discussion would probably be aided by taking a specific case of the general model formulated in Chapter II. Since the applications are in the economic time series area, a model using monthly data and $P = 12$ would seem appropriate for the purposes of discussion. The value of P could be 52 or 365, but the computations would become lengthy. In future discussion, the cyclical variations become seasonal trends, averages become yearly averages and random variations may be visualized as those random effects in the process such as those caused by local weather conditions.

The basis for judging the relative merit of the constants will be the error sum of squares. The definition of the error in this case is the difference between $FX_{t,T}$ and X_{t+T} for $T = 1, 2, \ldots, N$. This is not the only method of evaluating the smoothing constants and they could be evaluated on the basis of some penalty scale associated with the accuracy of the forecast. If the relation between small forecast errors and large forecast errors is linear in terms of the penalty associated with the error, then a logical method of evaluating the constants would be the mean absolute deviation observed over the test series. The test series is that part of the observed series used for making comparisons between the forecasts of the model and the actual data. This comparison is then used as a measure of the forecast error that may be expected from the use of a particular model for a given time series.

Testing the model brings up an important point that is often ignored in some of the published works in this area. For a valid comparison of techniques, it is essential that the forecasting method has no a priori knowledge of the test series. If the test series is used in the selection of the coefficients for a model, then there is a high degree of bias introduced in favor of that particular model upon the basis of error comparisons. If the test series is included in the final form of the model, then the associated error is a smoothing error not a forecast error and should be evaluated as such. The method used here is to separate the historical data into two groups. The earlier group of historical data is used in the supporting algorithms for the preparation of input information for the model and the latter group is used as a test series for the model. The first group is that historical data which was referred to in Chapter II. The older data is used for estimating the smoothed process via the supporting algorithms for each set of constants used and then the forecasts are made T periods into the future. The error is computed as the difference between this forecast and the actual value. After each forecast, the next value in test series is absorbed into the smoothed estimates in order to update the estimates. This is the same way the model is to be used in practice. After the last observed value has been forecast with the specified T the sum of the errors squared are averaged and used as a relative measure of the forecasting ability of a particular set of constants in the model. When making real forecasts, all the historical data is assimilated by the model in order to provide the latest estimates of the model parameters.

The use of part of the historical data for the test series points out another advantage of the exponentially smoothed model approach over the curve fitting techniques which assume that all the data available are used. It would become computationally involved if a polynomial curve fit is used for repeated forecasting and the degree of the polynomial is increased each time an observation is added to the series.

The division of the historical data into two parts creates a question as to how the data should be divided between the smoothed series and test series. Unfortunately, in the economic time series, there is some difficulty in obtaining enough data with the same base series for appropriate smoothing. It then becomes a question of using the data for obtaining better estimates of the process or for evaluating the model. The test series should be less than one-half of the data and should be long enough that the estimates of the variance are reasonable from the users viewpoint. If these two conditions cannot be met simultaneously, some other form of estimation should be used since the length of the series does not provide the self-generating type of model with a fair test. The programming of this model, as presented in the

Appendix, has this decision included as a part of the computational procedure.

Hierarchy of Models

Before discussing the elements of the selection process for the smoothing constants, it should be shown that the optimal selection of the smoothing constants automatically includes the consideration of the random and linear trend models. If a series is composed of random fluctuations about some mean value, the computational form for $\overline{Y}_o$ given in Equation (21) should become $\overline{Y}_o = 0.0$, and similarly Equation (39) yields $\overline{Z}_i = 1.0$, $i = -11, \ldots, 0$. For the set of smoothing constants (A 0 0), the supporting algorithms reduce to the form

$$\overline{X}_t = AX_t + (1 - A)\overline{X}_{t-1} \qquad 0 \leq A \leq 1 \qquad (40)$$

$$\overline{Y}_t = \overline{Y}_o = 0.0 \qquad (41)$$

$$\overline{Z}_{t+1} = \overline{Z}_i = 1.0 \qquad (42)$$

and the model will be reduced to the same form as shown in Equation (1)

$$FX_{t,T} = \overline{X}_t. \qquad T = 1, 2, \ldots, N \qquad (1)$$

The $\overline{Z}_i$ are dependent upon the amount of historical data used and can be expected to approach unity as the amount of historical data becomes large and the generating process continues as a random variation about the mean.

If these same assumptions are made for $\overline{Z}_i$ and the smoothing constants (A B O) are used in the complete model, then the computational algorithms will reduce to the form

$$\overline{X}_t = AX_t + (1 - A)(\overline{X}_{t-1} + \overline{Y}_{t-1}) \qquad 0 \leq A \leq 1 \qquad (43)$$

$$\overline{Y}_t = B(\overline{X}_t - \overline{X}_{t-1}) + (1 - B)\overline{Y}_{t-1} \qquad 0 \leq B \leq 1 \qquad (44)$$

$$\overline{Z}_{t-P+1} = \overline{Z}_i = 1.0 \qquad (45)$$

and the model will become

$$FX_{t,T} = \overline{X}_t + T\overline{Y}_t. \qquad T = 1, 2, 3 \ldots, N \qquad (46)$$

Thus, by searching for the optimal smoothing constants of the complete model, the two simpler models will be considered, provided that the process being studied is of the simpler form. If the generated process is of the simpler form, this will allow the complete model to assume the forms discussed above. It should be pointed out that these search techniques and model formulations are compatible downward only; that is, they cannot be reversed in direction where the optimum value of the smoothing constant A is chosen based upon single smoothing of the observed series and the value of B and C optimized in turn for the optimal values of the preceding constant. The reason is that the optimal value of A for a series composed of trend, seasonal and random effects would attempt to compensate for all elements in the series; whereas, with the trend and seasonal effects removed, the optimal value of A would generally be expected to decrease. A similar compensating action could be expected by the other two smoothing constants. The seasonal factors, $\bar{Z}_i$'s will increase in the absence of trend factors under conditions of a steadily increasing trend in the series and will perform adequately, but become subject to oscillations under conditions of changing trend. Therefore, there are a number of interactions within the computational scheme that prevent the step-by-step optimization in the selection of smoothing constants. This is a case of the model degenerating to a simpler form but the reverse optimization procedure not being possible unless the other parameters are considered concurrently.

If the historical data file is short, it is possible that the supporting algorithms will not converge to those values that reflect the simpler process. For example, if there is some small degree of autocorrelation from one period to the next for lag P, the algorithm will tend to show this as a cyclical effect when a random series could have produced the same degree of autocorrelation.

Search Methods

Since the approach used in this thesis has been to find the best fit of the exponentially smoothed moving averages, it has now been reduced to the problem of finding the optimum set of smoothing constants. The use of the term optimum as opposed to minimum will become evident later in the discussion. The judgment of optimum will be based on some function of the mean error sum of squares over the test series. The possible values of the smoothing coefficients each range over the interval zero to one. Within this range of values, it is possible to visualize the error mean square as the response of the process to a specified set of smoothing constants. This response may be discussed in terms of a hypersurface or an analogy may be drawn to a unit cube filled with a heterogeneous substance. The axes of the cube would correspond to the smoothing constants and the density of the substance at the intersection of the coordinates of those constants as the response. The object of the search procedure is to locate that set of smoothing constants which meet the prescribed requirements of optimality.

The search is complicated by the compensation of one type of smoothing for another as was pointed out in the discussion above. This will result in the creation of local minimums in the error mean square. This type of response limits the use of some of the mathematical search techniques. The most important method eliminated is the gradient method or the method of steepest descent. The application of this technique here experiences the same difficulty as when applied in the area of experimental design. This is, a local minimum may be found instead of the actual minimum which is the same basic problem for locating the region of minimum response. With the high-speed, large-memory computer, the gradient method is marginal in some respects except for the academic fascination of reducing the number of trials required to arrive at the minimum value.

An acceptable trial and error procedure is a systematic search of the axes. This in turn reduces the search problem to one of grid size to be used in three dimensional space. For lack of any better method at this point, it is suggested that the interval zero to one be divided into equal increments and applied to all three smoothing constants. The number of values chosen for the search will be cubed in the final enumeration of points investigated if zero and one are not included. For example, if the increments were taken as 0.09 about the mid-value 0.5, this would result in 11^3 or a total of 1331 points to evaluate; whereas, if an increment of 0.15 is used, this reduces the number of possible permutations to 216. If the end values zero and one are used, the number of points to be investigated will be for (m-1) increments or m points including the end points zero and one, $m^3-2(m^2-m)$. If either, but not both zero and one, is included as an end point in the grid the number of points to be evaluated is m^3-m^2+m. This reduction in the number of points to be investigated is accomplished without a corresponding reduction in the grid size. The reason for this is when A = 0.0, B may take on any value between zero and one. Therefore, this reduces the number of combinations including A = 0.0 to the number of different values used for C. The mathematics in terms of the model may be shown by demonstrating the effect of A = 0.0 in Equations (29) and (30) by

$$\overline{X}_t = (\overline{X}_{t-1} + \overline{Y}_{t-1}) \tag{47}$$

$$\overline{Y}_t = B(\overline{X}_t - \overline{X}_{t-1}) + (1 - B)\overline{Y}_{t-1} \qquad 0 \leq B \leq 1 \tag{48}$$

$$= B(\overline{X}_{t-1} + \overline{Y}_{t-1} - \overline{X}_{t-1}) + (1 - B)\overline{Y}_{t-1}$$

$$= B(\overline{Y}_{t-1}) + (1 - B)\overline{Y}_{t-1} = \overline{Y}_{t-1}$$

$$= \overline{Y}_0 \ . \tag{49}$$

There is a similar reduction in the number of points to be investigated in the case of A = 1.0, since C may take on any value between zero and one. The effect of A = 1.0 may be demonstrated by the reduced forms of Equations (29) and (33).

$$\overline{X}_t = X_t/\overline{Z}_{t-P} \tag{50}$$

$$\overline{Z}_t = C\ X_t/\overline{X}_t + (1 - C)\ \overline{Z}_{t-P} \qquad 0 \leq C \leq 1 \tag{51}$$

$$= C\ X_t\overline{Z}_{t-P}/X_t + (1 - C)\overline{Z}_{t-P}$$

$$= C\ \overline{Z}_{t-P} + (1 - C)\overline{Z}_{t-P} = \overline{Z}_{t-P}$$

$$= \overline{Z}_i \qquad i = -11, -10, \ldots, 0\ . \tag{52}$$

If it is desired to complete the grid of constants and error variances, this may be done by using the appropriate error term as determined by A and B, and A and C in the respective cases. Error variance is synonymous with the previously defined mean error squared.

The major criticism of the coarse grid search technique is that it may not locate all local minimums. Due to the general behavior of these error values, this criticism of the coarse grid technique is not applicable in this case because the local minimums are not of prime interest. The reason the minimum is not of practical interest is due to having observed these error values and reflecting upon the computational procedures, it was found that the error variance response is flat in the region of the optimum constants. The size of the optimum region is dependent upon the form of the observed process. The more closely the process agrees with the model, the less critical the constants become. This is due to the initial determinations of the trend and the seasonal variations agreeing closely with that which is observed in the series. Then the weighting of one estimate of the process relative to an other estimate produces little change in the forecast. This is particularly true in the case of the trend and seasonal components of the series.

After the coarse grid is used in the primary search and the region of the minimum error located, a finer grid may be used to search within the region. However, due to the low sensitivity of the model in the area of minimum error, it is seldom worth the additional effort. Since the sensitivity is low in this area, it lends support to the necessity of finding the region rather than a local minimum. This is because there is some optimal set of constants and even though they are supposed to be in the region, the ones selected may be near the optimal set, but due to the flatness of the response this deviation does not become serious.

Selection Methods

Now that the search technique has been established as one of enumeration over a coarse grid, the next step is the proper selection of the constants from among the possible sets that have been explored.

The most obvious selection method has already been discounted in the discussion above; that is, the use of the set of constants associated with the minimum error mean square. However, the minimum value may be in the optimum region. This set of constants associated with the minimum value usually arouses some curosity and they are one of the program options of the program presented in the Appendix. The next most obvious method is to list the permuted constants and their associated error variances and use a manual selection procedure for determining the optimum set of constants. Even with the computer this method has some merit in that it provides the forecaster with a subjective evaluation of the sensitivity and after some education in the function of the supporting algorithms will give him a better feel for the form of the process that is being observed by the model and how closely it conforms with the assumed form.

If each smoothing constant was to perform the specific tasks assigned without the tendency to compensate for the other variations in the series, then the selection of smoothing constants would be simpler. The minimum error method of selection would work, but as a point for discussion consider the idealized case where the minimum is at one point and the entire response surface is uniformly monotone increasing in all directions about that point. Under these conditions, there are a number of methods that will work. The smoothing constants for the minimum error are obtained if, for each value of A that is in the search grid being used, the error variances for all combinations of B and C were summed, the minimum sum would provide the A coordinate of the minimum error. Likewise, if the error variance for all combinations of A and C were summed for each value of B, it would provide the B coordinate of the minimum error. The C coordinate could be determined in a similar manner. This particular method has been found successful in a number of the real and artificial series and, based upon these observations, the general class of processes for which it is most suitable as a method of selection of the constants has been determined. This occurs when the observed series is of the form of the general model or one of the simpler forms which, in essence, provides a relatively large region of minimum error. In order to support this inductively, the supporting algorithms and model will be repeated below for purposes of discussion:

$$\bar{\bar{X}}_t = AX_t/\bar{\bar{Z}}_{t-P} + (1 - A)(\bar{\bar{X}}_{t-1} + \bar{\bar{Y}}_{t-1}) \tag{29}$$

$$\bar{\bar{Y}}_t = B(\bar{\bar{X}}_t - \bar{\bar{X}}_{t-1}) + (1 - B)\bar{\bar{Y}}_{t-1} \tag{30}$$

$$\bar{\bar{Z}}_t = CX_t/\bar{\bar{X}}_t + (1 - C)\bar{\bar{Z}}_{t-P} \tag{33}$$

$$FX_{t,T} = (\bar{\bar{X}}_t + T\bar{\bar{Y}}_t)\bar{\bar{Z}}_{t-P+T} \ . \tag{26}$$

If the process is of the simple form with some random deviations about a mean, then the initial determinations of trend and seasonal effects will approach the values of zero and one respectively. Thus, for optimum fit of the data and a minimum sum of error squared, the value of A will be small, since the smaller the value of A the more historical data are included in the estimate of the mean. So, assume that the A is small and observe the effect of changing the values of B and C. In the case of B, the initial value $\bar{\bar{Y}}_o$ is assumed to be near zero due to the type of process being observed. As a result of the small value of A, the difference $(\bar{\bar{X}}_t - \bar{\bar{X}}_{t-1})$ is small since the smoothed values of X_t differ only by some small fraction of the latest observation. Therefore, regardless of the value of B, the value of $\bar{\bar{Y}}_t$ remains near zero for a small A. Looking at $\bar{\bar{Z}}_t$ for which its initial values were all near unity the error would become dependent upon the value of C since C must be small for the values of $\overline{Z}_t$ to remain near one and keep from causing the predicted values to oscillate which would increase the error values. Thus, for the fixed value of A it can be summarized that by summing over the values of C, the error terms become a function of the value of C itself and, therefore, the selection based upon this summation provides the optimum value of C for the other values, but since the range of B has essentially no effect on the forecasts, the error variance depends upon the value of C and the summation is simply a constant times the error variance associated with that particular value of C.

If the value of A is increased in the absence of autocorrelation in the observed series for a small value of B and any values of C, the forecasts tend to reflect the last observed value and, thereby, increase the variation. As the value of B is increased, the addition of a false trend component in the model tends to increase the error variance. In general, the increase of B for the larger values of A causes a monotonic increase in the error variance. In those random series and those with small autocorrelation coefficients that were studied, there is some compensation for large values of both A and C, but the compensation is not sufficient to discredit this summation method for smoothing constant determination. Similar analyses can be made for the processes that have linear trends or both linear and seasonal trends which have small values for the autocorrelation of the adjusted random series.

The definition of the term autocorrelation in the preceding discussion was on an intuitive basis. This particular statistical definition is most important in the discussion of time series. Let r_k be the autocorrelation coefficient with lag k.

$$r_k = \frac{s_{X_i, X_{i-k}}}{s_{X_i} \, s_{X_{i-k}}} \, . \qquad -1 \leq r_k \leq 1 \qquad (53)$$

This is the same form as the correlation coefficient for pairs of observations except they are from the same series with a lag of k between the observations. If there are M observations in the series for which it is desired to compute this statistic, then Equation (53) may be written in summation notation as

$$r_k = \frac{\sum_{i=k+1}^{M} X_i X_{i-k} - \left(\sum_{i=k+1}^{M} X_i \sum_{i=1}^{M-k} X_i \right) \Big/ M-k}{\left[\sum_{i=k+1}^{M} X_i^2 - \left(\sum_{i=k+1}^{M} X_i \right)^2 \Big/ M-k \right]^{1/2} \left[\sum_{i=1}^{M-k} X_i^2 - \left(\sum_{i=1}^{M-k} X_i \right)^2 \Big/ M-k \right]^{1/2}} \, . \qquad (54)$$

Often, in the more mathematical treatments of the subject of time series, r_k is assumed equal to zero. In the relatively short series of observations that are used as historical data for the processes discussed here, it is difficult to reject this hypothesis statistically. This is similar to providing statistical evidence of non-randomness in a series. Regardless of the level of significance that can be shown statistically, the statistic r_k is obviously related to the smoothing constants that are used in the exponential smoothing of the series. This relationship becomes more apparent if the numerator of Equation (53) is written in the form,

$$r_k = \frac{\sum_{i=k+1}^{M} (X_i - \bar{X}_{(i)})(X_{i-k} - \bar{X}_{(i-k)})}{s_{X_i} \; s_{X_{i-k}}} \; . \qquad (55)$$

The value of r_k is seen to be dependent upon the values of M and k for a given series. Since either of the quantities within the parentheses may be positive or negative and the summation of the cross products is for all values of i, the numerator may become small or approach the value of the denominator either as a positive or negative quantity. If the values of the series have a tendency to be on the same side of their respective means for the specified k, then this is indicative of systematic variation in the series for lag k. For a large M and relatively small k, the values of $\bar{X}_{(i)}$ and $\bar{X}_{(i-k)}$ approach the same value where

$$\bar{X}_{(i)} = \sum_{i=k+1}^{M} X_i / M-k \qquad (56)$$

and

$$\bar{X}_{(i-k)} = \sum_{i=1}^{M-k} X_i/_{M-k} \; . \tag{57}$$

There are some relationships between the value of r_k and the model that assumes the form of a trend, cyclical and random variation, that may be stated without rigorous proof or empirical data.

One point that should be made about a significant value of r_k is that all points which are integral values of k in either direction about every point in the series contributes to the relationship that is measured by r_k not just each k^{th} value beginning with the first data point. Some properties of the observed series may be determined by looking at r_k for consecutive values of k. If $r_1 > r_2 > r_3 > \ldots > r_n$, then the series is seen to be dependent upon the most recent information as a forecast of future activity. If $r_1 \doteq r_2 \doteq \ldots r_n \doteq 1$, then the series has an established trend. If the values of r_k oscillate near zero for all k, then the series is of a random nature, but if r_k oscillates about zero with a large amplitude in either direction, then it has a cyclical variation. The proof of these observations is by inspection of Equation (55). Combinations of these basic patterns may be formed for more complex forms of the time series.

The display of the autocorrelation coefficients for consecutive values of k is formally known as a correlogram. Correlograms will be presented in conjunction with the numerical examples in Chapter IV.

Since the unstable type of time series may exhibit patterns within the data that are a function of the correlation for various lag values within the data, a method for selecting the smoothing constants that evaluates this relationship should be considered.

If a single set of smoothing constants is to be used, then it is proposed that the determination of error be more nearly representative of the actual conditions of the real forecasts. One method would be to determine the error variance for lags of one through P and select the set of constants that has the minimum sum of P error variances. This procedure should provide a more realistic estimate of the error that may be expected, particularly for the stable type of series.

However, since the coarse grid search is being used, it requires only a little additional effort to arrive at an optimum set of constants for each lag value from one to P. Since the unstable series is likely to exhibit similar properties in terms of the autocorrelation for different values of the lag a more direct correspondence between the smoothing constants and the autocorrelation for a given lag could be used to improve the computed forecast error.

The basic assumption that the generating process for the series does not change over the range of the observed data becomes important in the selection of smoothing constants by this method because the error is based upon the forecast over the test series and the selection of the smoothing constants becomes some function of the autocorrelation over the test series. If the base series is unchanged in the test series as in the total historical series, then the forecast may be expected to reflect the total amount of information after smoothing; but, if the test series is not of the same general pattern as the other historical data, then the forecast will reflect the test series alone. This in effect has compounded the level of stability in the observed series either positively or negatively in the forecast of the future activity of the process. Therefore, selection of a set of smoothing constants for each interval of the forecast may result in forecasts corresponding closely to the last period of historical data.

The relative merits of these methods of selection will be demonstrated in the next chapter by use of numerical examples. However, the development of exponential smoothing as a time series model has grown from the selection of a single smoothing constant by trial and error to a systematic procedure for selection of 36 smoothing constants in order to forecast 12 months into the future.

In summary, there have been four separate methods proposed for selection of the smoothing constants for the exponentially weighted moving averages and each more suited for particular types of series. However, the general computational form provides for evaluation of all four methods simultaneously.

Selection of Starting Point

It is possible to significantly reduce the error variance by changing the starting point in the historical data by one unit in time. This particular conclusion was formed by observing the errors associated with the model as historical data were added to and deleted from the programmed computation. Generally, it is necessary to change the starting point for the historical data more than one interval of time. For example, assume that the historical data available on a process took the form shown in Figure 4. Under the type of model that is being used here for extrapolation of the series, an intuitive starting point becomes obvious for a model that has a linear trend component. This would be sometime after the first year of data.

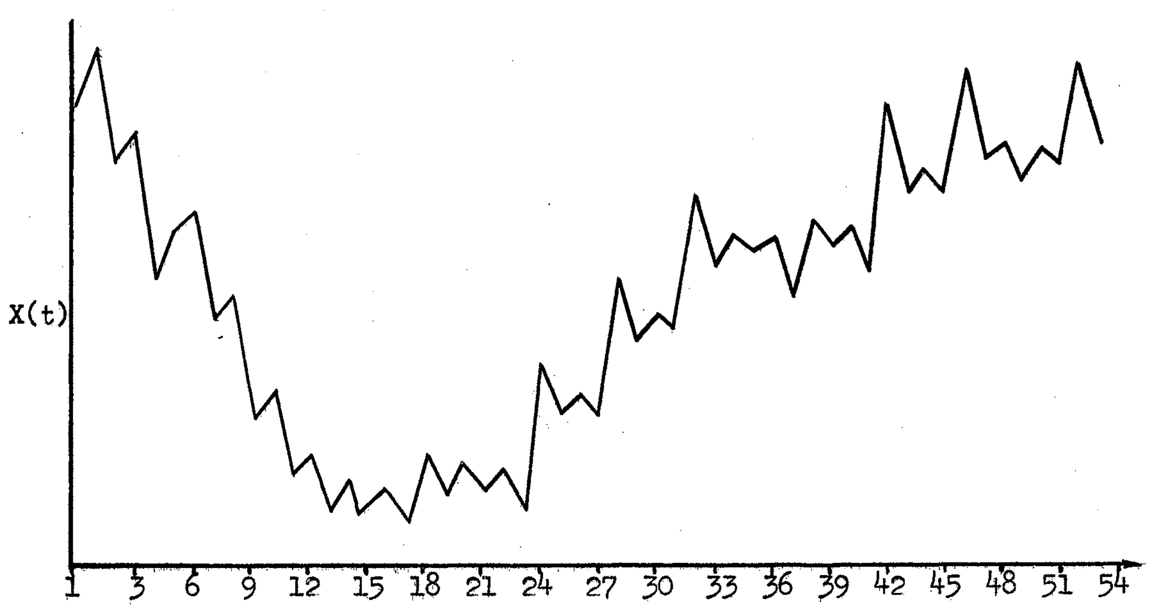

Figure 4. A Changing Process

The reduction in the error variance for a process of the type shown could be quite large. The main cause of increased error in considering the whole series is the computation of $\overline{Y}_o$ on the basis of the data given would indicate a very small trend component compared to that which actually exists. Since the model was not originally provided with this ability to determine the optimal starting point, the supporting algorithms would attempt various methods of smoothing the series, each attempt yielding a larger error variance than could be obtained with a more judicious choice of the historical origin of the series. If the value of B is large in order to discount the initial determination of the trend, then it will also fail to recognize the real trend in the later data. This will cause the trend factor to attempt to compensate for most of the variation which will distort the actual seasonal variation and increase the error variance. If the trend changes slowly, both the random factor and the seasonal factor will attempt to compensate for the lack of response in the trend. The random factor would assume that a high autocorrelation exists between successive observations and would cause sympathetic oscillations in the trend factor or the seasonal factor would interpret the data as having large seasonal fluctuations. In either case, the forecast made over the test series would result in a relatively large error in forecasting and these artificial oscillations would be carried over into the actual forecast of the process. Also in case of an observed series of this type, the seasonal factors would become inverted during the year of decline and would take a much longer time to reverse the negative trend in these factors than the linear trend. Therefore, it is desirable to determine the optimum starting point to begin the smoothing of the series in order to provide the latest information as input to the model, and to provide the most representative data of the present generating process. This, then, satisfies the requirements of the assumptions that are made in the use of the model in that the process that generated the series to which the algorithms were applied is assumed to continue into the future.

In the example given above, if it is desired to include this large variation as part of the historical data, then the period of the seasonal variation should be changed to a length such that the pattern as displayed in Figure 4 would only represent part of the period which would repeat with similar major fluctuations in the future. This would only be fair to the forecaster and to the model as presented here.

In order to solve this particular problem associated with the use of the model and to obtain less error in the forecast, the difference operator, Δ, is used to determine the best fit of a first degree equation, linear trend, to the observed data for a specified minimum interval of the historical data. This is accomplished by establishing the minimum number of historical data points that will be used in smoothing the series and then changing the starting point for the minimum sum of the absolute value of the third order differences for this interval as it is passed over the historical data available. The starting point in the series that results in the smallest sum of the absolute third order differences is then chosen as the origin of the observed series for purposes of estimating the parameters of the model. Generally, this interval would be expected to be at least seventy per cent of the historical data available, but not less than 45 months in length unless decisions are made external to the automated computational process.

The third order differences are used instead of the second order differences in order to facilitate the automatic computation and decision processes. From a theorem of finite differences, the n^{th} divided differences of a polynomial of degree n are constant. If the second order differences were all equal to zero, then an exact fit of the linear trend would have been accomplished. However, if something less than an exact fit is to be accepted, then the magnitude of the second order differences can indicate the best of the choices. For the type of data that would ordinarily be encountered in the unstable type of time series, oscillations in successive values of the early order differences could be expected. Therefore, second order differences could result in two common mistakes in judgment. The first of these would be for large alternating signed values to sum to a small value thereby indicating a best fit condition. The second would be for a best fit condition as demonstrated by constant or near constant second order differences to be rejected in favor of a set of differences that had begun oscillating. By taking the third order differences, both of these mistakes are avoided since the oscillations in the second order differences will show up as even larger values in the third order differences and any near fit will have smaller third order differences. Also, if the second order differences were in fact zero indicating a first degree fit of the data, then the third order differences would remain zero. The chances for a first degree fit and a second degree fit within the same series under the general rules given for selection of the interval are remote. A secondary benefit is derived from this procedure if the seasonal variation is of some low order polynomial, second degree or less, in that this procedure tends to provide a better fit for the complete model.

There are some modifications of this general technique discussed in the Appendix along with the presentation of this as a program option. A discussion of empirical results and examples will be given in Chapter IV.

CHAPTER IV

EMPIRICAL RESULTS

From the standpoint of rigor, the analytical approach is usually preferred over the empirical approach to problem definition and solution. However, the empirical approach cannot be discounted in its usefulness and the type of problems that it approaches. Often, the use of empirical techniques will precipitate some analytical formulation of the same result since the empirical approach provides some intuitive insight into the action or interaction of factors that provided a result which is general in form. The analytical treatment is then used to derive the result that has demonstrated ability to provide a solution. One of the more significant developments of this type was the development of the t distribution by Student (Gossett) and the subsequent analytical proof or derivation of that same distribution by Fisher. There are a number of people presently working on time series from both the analytical and empirical approaches. Of the contemporary investigators, Brown would probably be a leader from the empirical studies approach and Parzen from the analytical approach to the subject.

The empirical approach to these studies has been advanced in recent years by the use of high speed computers which make it feasible to consider enough different types of series to be able to generalize the results of the study. It was through the use of computer results that most of the material already presented was developed. The purpose of this chapter is to support those hypotheses and arguments that were advanced in the previous chapters.

Selection of Data for Study

The original data that were used in the development of the results presented in this work were actual data from personnel attrition studies. This type of time series would undoubtedly qualify as an unstable process as described in previous discussions. The influences in this type data are many. Not only are the economic conditions factors to be considered in the movement of people to other jobs, but the day of the week ending the month and other such coincidences have an influence upon the monthly attrition rate. Much of the exploratory work in reducing forecast error came about through the studies conducted on these data. These early studies also provided a real test for this type of analysis. The test of the model was a comparison between the model and its forecasts of the future activity of the process with those predictions made by an experienced personnel man that had studied the problem and used economic indicators along with some limited statistical analysis of the historical data to formulate his predictions. The model compared favorably during these early tests. Subsequent use of a model which used the minimum error variance for forecast lag value of one as the basis of smoothing constant selection pointed out some of the types of errors that could be expected under a limited computational procedure. This led to a trial and error procedure for correcting the demonstrated deficiencies in the original methods. The improvements were directed at replacing the personal judgment methods that were needed to obtain better results with an automated decision process.

In the early stages of the personnel attrition study, a number of the curve fitting techniques were used as time series models. The variation in the data and computational complexity involved in arriving at the mathematical form to be used for extrapolation, caused those methods to be rejected in favor of the exponentially weighted moving averages. Due to some proprietary considerations, the data of these early studies are not available for presentation, but it is not unlike data that will be used in the presentations in this chapter.

In addition to using real world data to study the model and the forecasting procedure, artificially created time series data were used for investigating the reaction of the model to various types of known input variations. One of the basic experiments used selected combinations of a factorial arrangement of trend, random variation and additive cyclical variations. The experiment was to determine if the model could track and extrapolate a true signal without error. The signal combinations used included: different values of slope for a linear trend, these same trends with three different levels of amplitude for the sine wave superimposed on the trend. The results of this particular phase of the experiment indicated that the multiplicative model could track a process with an additive type of cyclical variation and that signals of the simpler form could be extrapolated without error. The second phase of this experiment used the same base signals with three different levels of random variations superimposed upon the base signal. The levels of randomness were determined by the width of the interval for the values used from a table of random numbers. The basic purpose of the second phase was to observe the change in the error variance associated with the increase in the random variations for the combinations of the base signals used. This part of the study was generally to increase the confidence in the model.

The group of data that are presented in the illustrations and presentation of numerical results in this chapter were selected in order to provide a comparison with other studies that have been conducted along these lines. The data that are used in the following tables and graphs are similar to those presented by Brown (7).

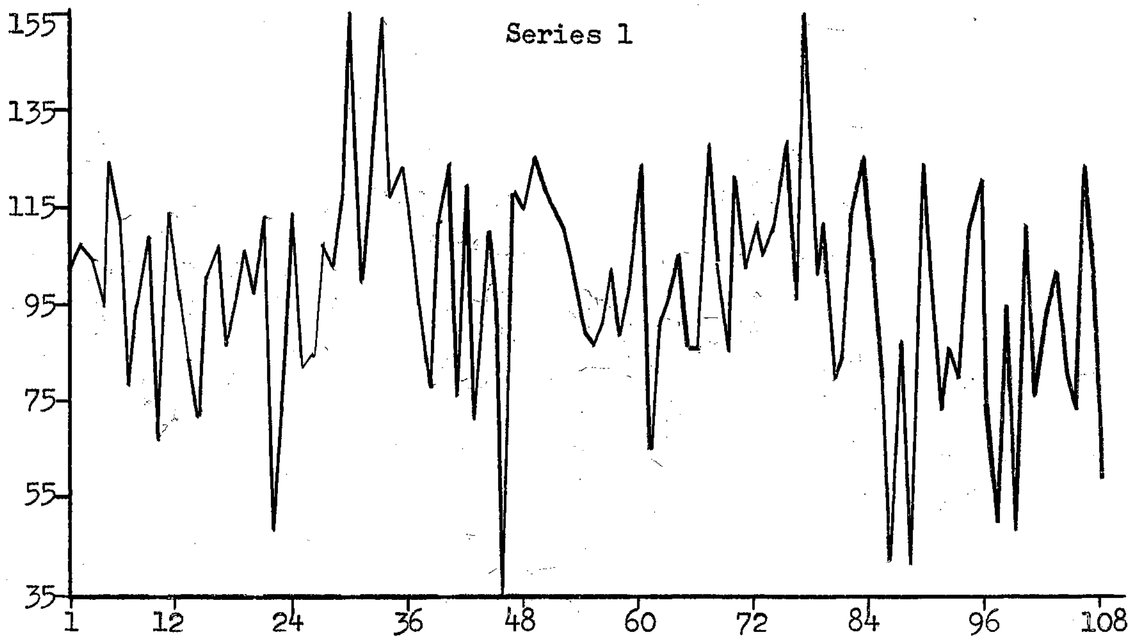

Figure 5. A Series With Noise About a Process Level

TABLE I

SERIES 1 - RANDOM TYPE SERIES

Time \ Period	1	2	3	4	5	6	7	8	9
1	102	083	083	095	126	065	104	080	051
2	107	072	085	079	119	091	111	043	095
3	104	111	108	113	116	097	128	088	049
4	094	117	103	123	110	105	096	042	111
5	124	086	118	076	103	086	165	124	077
6	112	095	168	120	089	086	103	102	093
7	079	106	100	072	087	128	111	074	102
8	095	098	114	110	091	104	080	086	081
9	109	113	154	096	102	086	084	081	074
10	067	048	118	036	089	122	115	111	124
11	114	074	123	118	098	104	126	121	102
12	095	114	116	116	124	111	104	075	060

Figure 6. A Series With Linear Autocorrelation Among Observations

TABLE II

SERIES 2 - AUTOCORRELATED TIME SERIES

Time \ Period	1	2	3	4	5	6	7	8	9
1	106	113	074	117	155	135	107	091	079
2	105	101	093	113	138	143	111	093	069
3	108	102	099	103	146	126	124	099	048
4	097	100	094	125	136	124	117	106	041
5	096	086	112	125	119	137	101	107	054
6	106	082	115	111	138	125	098	110	065
7	100	064	127	110	151	122	084	107	080
8	111	046	145	102	162	105	094	107	064
9	107	048	135	113	155	101	103	106	052
10	116	048	134	125	144	102	096	105	056
11	098	063	119	141	134	105	107	090	065
12	109	057	118	158	132	107	106	073	063

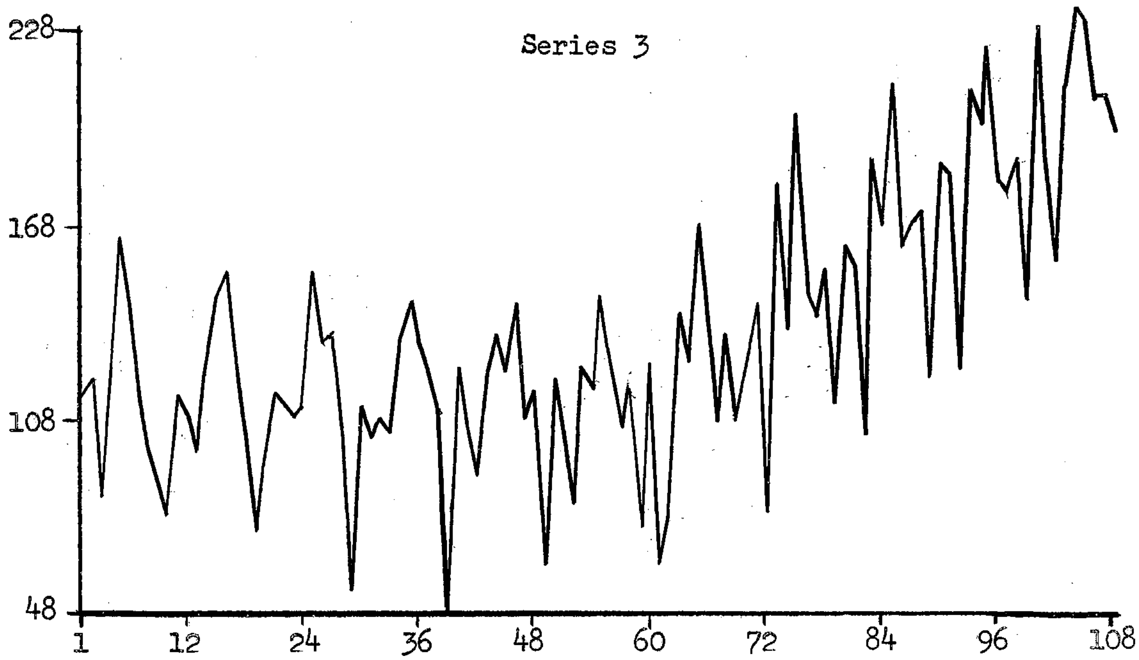

Figure 7. A Series With Noise and Trend

TABLE III

SERIES 3 - TREND AND RANDOM TYPE SERIES

Time \ Period	1	2	3	4	5	6	7	8	9
1	106	091	146	117	056	104	169	203	170
2	113	113	124	104	112	070	128	153	179
3	076	138	126	041	104	133	192	159	136
4	115	145	098	115	074	118	137	162	212
5	155	111	048	096	115	159	130	112	180
6	137	095	104	084	111	136	145	178	149
7	107	067	095	114	138	098	104	177	203
8	089	087	100	125	125	125	153	114	227
9	079	108	096	117	098	098	147	200	224
10	071	104	125	135	108	118	094	191	199
11	107	101	136	101	065	133	181	214	199
12	101	103	124	108	116	072	161	172	188

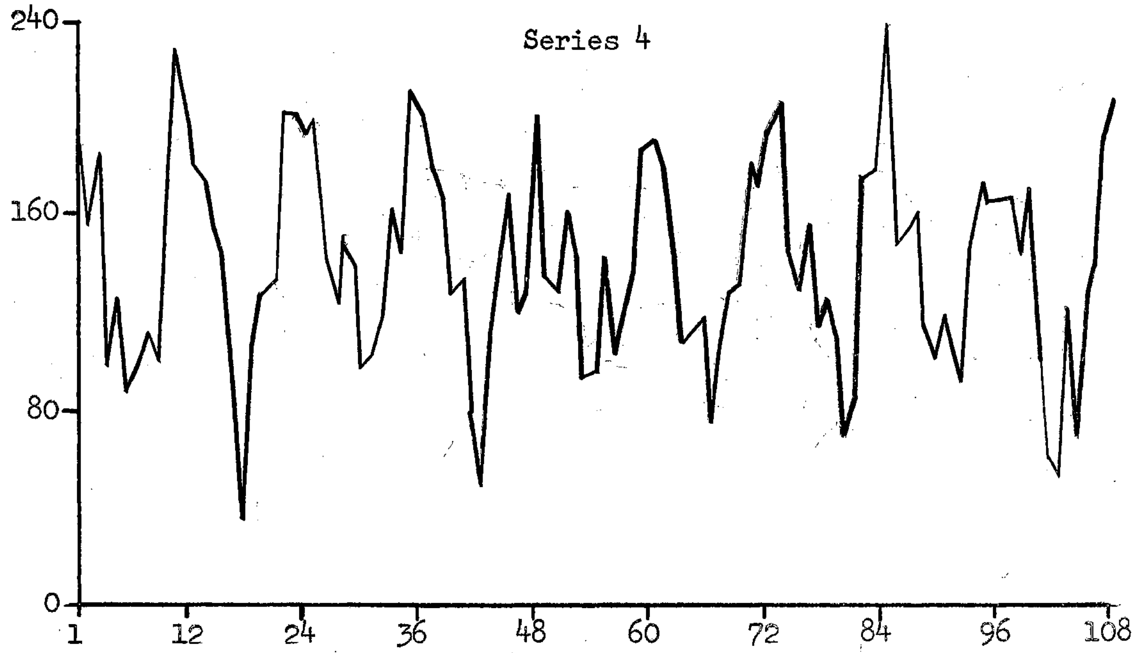

Figure 8. A Series With Cyclical and Random Variations

TABLE IV

SERIES 4 - CYCLICAL TYPE SERIES

Time \ Period	1	2	3	4	5	6	7	8	9
1	190	183	198	181	138	183	208	151	168
2	157	174	144	172	130	148	146	156	146
3	185	156	126	131	163	109	129	161	171
4	098	145	149	135	144	115	158	119	103
5	125	085	141	082	094	118	118	103	063
6	091	037	100	050	096	076	126	117	054
7	098	128	104	112	141	105	112	106	123
8	112	108	119	140	105	127	070	094	069
9	101	134	161	170	123	132	085	147	128
10	151	202	148	122	138	183	176	174	141
11	227	201	213	130	188	175	179	167	195
12	198	194	202	203	189	195	240	168	211

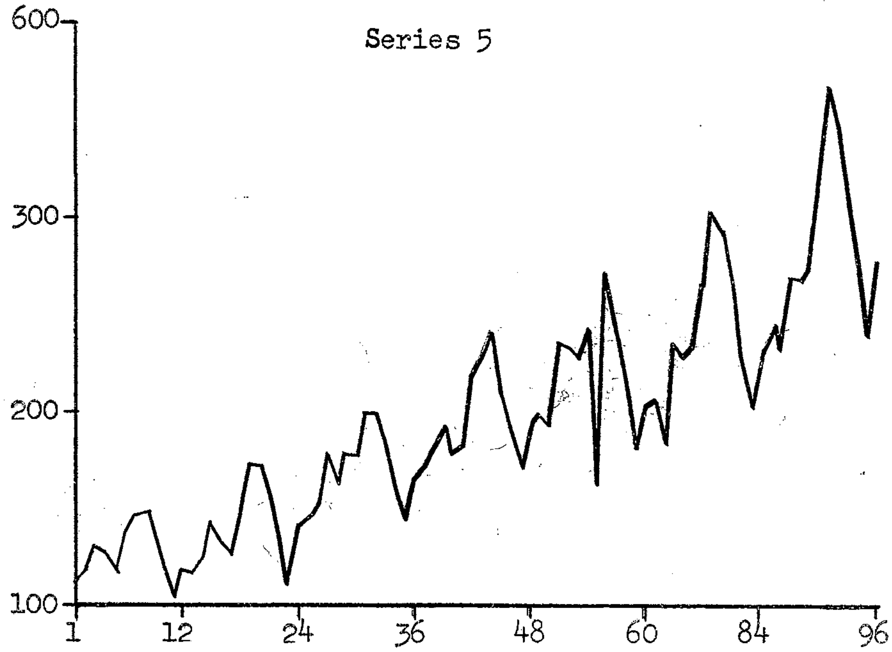

Figure 9. A Series With Trend, Cycles and Noise

TABLE V

SERIES 5 - CYCLICAL AND TREND TYPE SERIES

Time \ Period	1	2	3	4	5	6	7
1	113	116	146	172	197	205	243
2	119	125	151	179	195	189	232
3	131	142	177	192	235	234	268
4	130	134	162	180	234	228	268
5	120	126	173	182	230	233	271
6	136	148	179	217	242	263	316
7	149	171	200	229	163	301	365
8	149	171	200	241	273	294	348
9	135	157	185	210	238	260	313
10	120	132	161	192	210	228	275
11	105	113	145	171	181	202	238
12	117	141	165	193	202	228	279

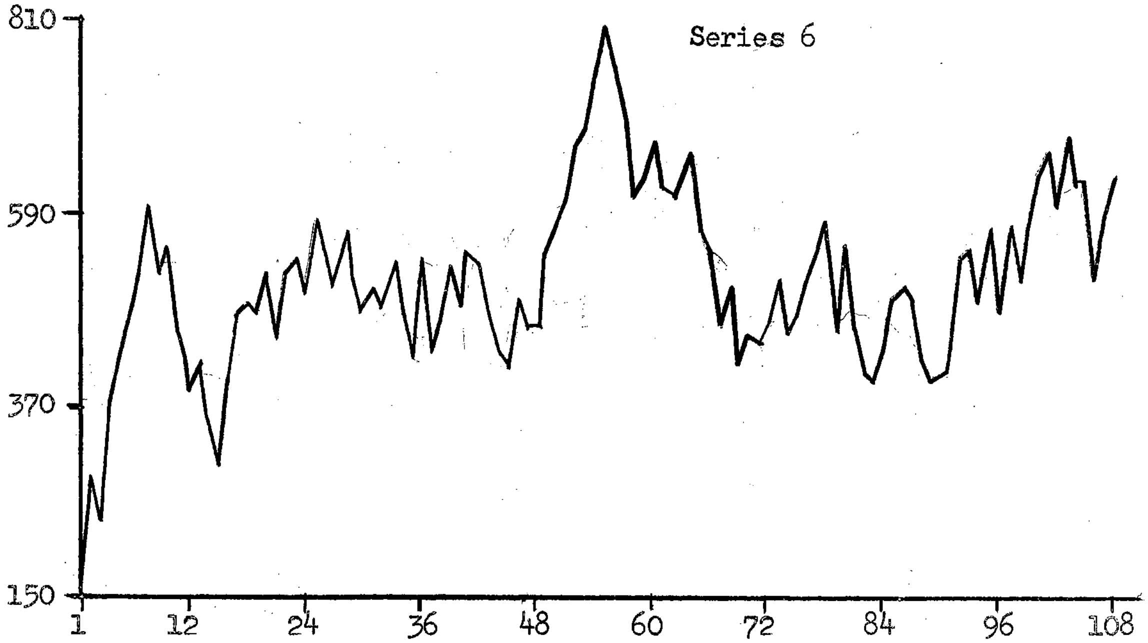

Figure 10. A Series With Autocorrelation of a Different Type

TABLE VI

SERIES 6 - NON-RANDOM TYPE SERIES

Time \ Period	1	2	3	4	5	6	7	8	9
1	168	418	582	433	547	629	511	492	570
2	294	360	550	473	582	610	456	504	510
3	244	302	514	529	619	614	471	494	568
4	373	401	575	493	671	659	514	428	634
5	433	476	518	549	689	571	549	404	658
6	481	487	482	537	748	552	581	403	596
7	521	484	507	476	793	467	466	436	661
8	595	525	486	437	765	501	554	535	628
9	522	450	535	414	698	420	460	547	631
10	552	526	479	481	612	446	407	488	518
11	462	532	426	464	628	442	401	568	589
12	391	501	539	463	673	461	437	473	635

Enumeration as a Search Method

The discussion in this chapter will parallel that of the preceding chapter when possible. The first point that was made in Chapter III was that the enumeration of the coarse grid values was the proper approach to the search for the appropriate set of smoothing constants to be used in the computational algorithms. Table VII is an enumeration of the error variances as computed for the last 60 values of the data of series 6. This grid of values demonstrates the existence of local minimums, which is the type of occurrence that was discussed relative to the use of the gradient method of searching for the minimum value. This may be determined by comparing the response at point (0.4 0.2 0.2) with its adjoining values.

The adjoining values of this point are all points that may be defined by combinations of A = 0.2, 0.4, 0.6, B = 0.0, 0.2, 0.4, and C = 0.0, 0.2, 0.4. These points may be identified directly by inspection of Table VII. A convenient means of identification is the construction of lightweight lines through the table for each value of the constants given above. This is done in an independent manner considering only one smoothing constant at a time. Upon completion of this construction of lines, there will be 27 intersections of these lines in three columns and nine values per column. The center value will be the point being investigated and the comparisons may be made

TABLE VII

ERROR VARIANCE TABLE FOR SERIES 6

				A 0.0	0.2	0.4	0.6	0.8	1.0
C	0.0	B	0.0	48815.76	6930.61	3572.14	3421.04	4011.73	5212.46
			0.2	48815.76	3242.18	2796.59	3251.35	4272.08	6143.48
			0.4	48815.76	3637.17	2820.37	3543.55	4954.73	7731.45
			0.6	48815.76	2967.26	2982.52	3974.09	5863.37	10033.60
			0.8	48815.76	2552.56	3361.08	4497.48	7011.43	13462.06
			1.0	48815.76	2559.54	3922.36	5051.27	8472.51	18884.26
C	0.2	B	0.0	48323.25	6443.78	3513.67	3411.86	4025.36	5212.46
			0.2	48323.25	2916.57	2776.83	3257.76	4293.35	6143.48
			0.4	48323.25	3151.49	3011.52	3615.87	4994.83	7731.45
			0.6	48323.25	2882.06	3517.08	4080.93	5914.49	10033.60
			0.8	48323.25	3135.07	4217.40	4588.04	7073.57	13462.06
			1.0	48323.25	3566.99	4889.84	5119.99	8560.76	18884.26
C	0.4	B	0.0	47131.48	6265.17	3624.54	3430.31	4029.90	5212.46
			0.2	47131.48	3254.42	3061.34	3297.43	4297.24	6143.48
			0.4	47131.48	3653.42	3634.53	3722.43	5009.35	7731.45
			0.6	47131.48	4311.74	4543.85	4248.46	5929.80	10033.60
			0.8	47131.48	5891.81	5801.97	4799.54	7079.64	13462.06
			1.0	47131.48	6558.93	7083.93	5367.28	8553.76	18884.26
C	0.6	B	0.0	45147.99	6411.96	3955.16	3502.23	4027.96	5212.46
			0.2	45147.99	4229.66	3761.34	3424.98	4289.09	6143.48
			0.4	45147.99	5280.75	4932.46	3953.10	5006.71	7731.45
			0.6	45147.99	7901.92	6415.52	4625.57	5921.95	10033.60
			0.8	45147.99	13094.72	8514.95	5394.91	7048.29	13462.06
			1.0	45147.99	15219.25	11118.25	6207.71	8477.13	18884.26

TABLE VII (Continued)

				A 0.0	0.2	0.4	0.6	0.8	1.0
C	0.8	B	0.0	42706.70	6890.31	4542.68	3656.03	4023.33	5212.46
			0.2	42706.70	5954.92	4952.08	3701.59	4276.00	6143.48
			0.4	42706.70	9326.05	7152.33	4413.48	4998.10	7731.45
			0.6	42706.70	15362.72	9540.63	5401.78	5908.78	10033.60
			0.8	42706.70	29304.27	12671.24	6756.24	7007.80	13462.06
			1.0	42706.70	39686.97	17286.26	8314.91	8376.25	18884.26
C	1.0	B	0.0	40476.09	7748.24	5401.78	3917.94	4020.43	5212.46
			0.2	40476.09	8962.62	6638.62	4189.16	4266.20	6143.48
			0.4	40476.09	20218.20	10493.85	5225.75	4996.92	7731.45
			0.6	40476.09	30057.83	14326.36	6818.80	5912.46	10033.60
			0.8	40476.09	60354.64	18350.76	9438.58	6994.76	13462.06
			1.0	40476.09	103716.98	25359.64	12811.68	8316.15	18884.26

directly. Since it is not convenient to use this method in this presentation, these values are displayed in Table VIII. If any part of the smoothing constants for the point being investigated as a minimum are 0.0 or 1.0, this grid size will be reduced by a corresponding amount.

TABLE VIII

ILLUSTRATION OF A LOCAL MINIMUM IN THE ERROR VARIANCES FROM TABLE VII

				A 0.2	A 0.4	A 0.6
			0.0	6931	3572	3421
	0.0	B	0.2	3242	2797	3251
			0.4	3637	3820	3544
			0.0	6444	3514	3412
C	0.2		0.2	2917	2777	3258
			0.4	3151	3012	3616
			0.0	6265	3625	3430
	0.4	B	0.2	3254	3061	3297
			0.4	3653	3635	3722

The occurrence of local minimums is not uncommon, and three or more may be found in the 216 point grid for some of the unstable series. With this one example, it is possible to discount the use of the method of steepest descent as a general technique for location of the minimum value in the search grid. The method of steepest descent would locate the position of this local minimum provided that the initial points for investigation were selected in this general area. If the method of steepest descent were used for a smaller grid size than used above, it would be reasonable to expect the number of local minimums in the hypersurface to increase.

By inspection of Table VII, the actual minimum value for the search grid that was used occurs at the points (0.2, 0.8, 0.0). It is irrelevant at this point as to which of the above is the optimum set of constants. The fact that local minimums can exist in the response of series to the smoothing constants justifies the selection of an enumeration method of search over the method of steepest descent.

Selection of the Search Grid

The search grid is intended to serve two purposes, the first is locate the region of optimum error variance and the second is to do this as economically as possible. The economy is measured in terms of the degree of convergence that is desired for the error variance and the computational effort involved.

The size of the region of optimum error variance is dependent upon the observed series. The grid should be selected fine enough to assure that at least one point in this optimum region will be obtained. Generally, this particular value for grid size is unknown, but the 0.2 increment used in this study has provided the general type of balance that is discussed above. When a local minimum is found, the grid size can be decreased for a search in the general region of that point. The finer grid search may be used upon a judgment basis; for example, if the values of Table VIII were of the same magnitude, this would indicate an optimum region and further search within that region for a smaller value of error variance could not be justified from an economic standpoint.

For the type of search grid that yields more than one local minimum, the finer grid may be used to determine which of the regions has less variation in the error variance values of that region. If there is only one local minimum, then the use of the finer grid should also be used on a judgment basis. In order to demonstrate the type of region that is desired and the method of choosing between local minimums the data from series 6 will be used in conjunction with a grid of values that are one-half the size of the original search grid values. These grids will be used for the two local minimums discussed above. The idea of a local minimum was illustrated above and is used here, with reference to the cube analogy, as a value that is less than its nearest neighbor in each plane and on the diagonals. The points (0.4 0.2 0.2) and (0.2 0.8 0.0) are the only local minimums in the error variance grid for series 6. Therefore, a grid of size 0.1 is used to search the region around each of these points. The results of these searches are displayed in Tables IX and X.

TABLE IX

ERROR VARIANCE TABLE FOR GRID SIZE 0.1
ABOUT THE LOCAL MINIMUM
(0.2, 0.8, 0.0) OF
SERIES 6

		A								
		0.1			0.2			0.3		
		B			B			B		
		0.7	0.8	0.9	0.7	0.8	0.9	0.7	0.8	0.9
C	0.0	8237	9027	8607	2678	2553	2536	2714	2794	2937
	0.1	7687	8390	7965	2591	2602	2693	2994	3241	3558

TABLE X

ERROR VARIANCE TABLE FOR GRID SIZE 0.1
ABOUT THE LOCAL MINIMUM
(0.4, 0.2, 0.2) OF
SERIES 6

		A								
		0.3			0.4			0.5		
		B			B			B		
		0.1	0.2	0.3	0.1	0.2	0.3	0.1	0.2	0.3
	0.1	2747	2737	2745	2714	2754	2800	2866	2943	3040
C	0.2	2718	2725	2779	2712	2777	2870	2959	2959	3085
	0.3	2782	2836	2969	2768	2874	3033	2891	3008	3169

The grid size may be made smaller until the values in the search area have converged to some specified fraction of the minimum value within the table. In the examples above, the smaller grid was taken as one-half the original grid. This is not a rule and any iterative technique that yields the desired level of convergence may be used.

The general computational scheme for this thesis used grid values equally spaced for each of the smoothing constants. This is not necessarily the most efficient method since a number of observations of the error variance tables for various observed series indicates that the response of the error variance is less sensitive to changes in the seasonal smoothing constant. This would indicate that a more efficient search might be obtained for the same amount of computational effort by making the increments of A and B smaller and increasing the increments of C.

Manual Selection of Smoothing Constants

One of the methods of selection of the optimum set of smoothing constants is a manual manipulation of the error variance table. This is probably the most reliable method of selection for the general case and should not be overshadowed by the refinements that are presented later.

This procedure follows that which was used to arrive at the values presented in Tables IX and X. After the local minimums have been compared, it is desirable to choose the smoothing constants associated with that local minimum which has the smallest amount of variation among those values in that region while remaining relatively near the minimum value observed. This selection could be aided by using a relative measure such as the coefficient of variation, but the additional computations cannot be justified on the basis of past experience.

If the selection of the smoothing constants is made on the basis of the information given in Tables IX and X, they will be the set (0.4 0.1 0.2). This choice is justified on the basis that the values of error variance in Table X have less variation than those of Table IX. This is particularly desirable since variations in the process would be less likely to influence the error variance significantly. This is one property of the forecast that is most important in that the "best" estimate of the error variance, as determined by the procedure outlined above, is needed since the process may vary with time. Thus, the term optimum error is used to describe the basis for smoothing constant selection rather than minimum error. The coefficient of variation for the value selected is 0.0942 which is only 0.0031 greater than the minimum value of Table IX. Therefore, the conditions of smoothing constant selection have been satisfied and arguments presented for their justification.

One other point that should be noted from this example is the efficiency of the coarse grid as a search method for locating the minimum error variance values in each of the tables. In both cases, the change in the coefficient of variation between those values that are common to Table VII (page 78) and Tables IX and X (page 83) is less than 0.0010.

The composition of the series in terms of non-random variations determines to a major degree the observed forecast error for this time series model. At this point in the discussion in this chapter, the observations have been based upon those results obtained for a forecast lag of one unit of time. The definition of lag is the number of time units between the smoothed values of the series and the forecast value. For a lag of one unit, the forecast is made one unit into the future and error observed before the smoothed estimates are recomputed and the next forecast is made. In order to summarize the numerical results that have been considered up to this point in the presentation and to provide a basis for comparison with the future developments, Table XI compares the error variance for the manually selected smoothing constants for forecast lag of one time interval with the variance of observed series.

These smoothing constants were selected by use of a coarse grid with an increment of 0.2 and the secondary grid of 0.1. The secondary grid was applied to all local minimums that were obtained with the coarse grids. The CV is a common statistical term for coefficient of variation which is the standard deviation divided by the mean. This provides a relative measure of the variation.

TABLE XI

COMPARISON FORECAST SUMMARY OF ERROR VARIANCE
AND OBSERVED SERIES VARIANCE

Series	Series Variance	Series Mean	Smoothing Constants	Error Variance	CV x 100 Series	CV x 100 Error
1	504.8	98.7	.0 .0 .2	625.1	22.77	25.34
2	544.4	104.6	1. .0 .0	153.2	22.32	11.84
3	1159.1	129.9	.0 .0 .2	1229.6	26.21	27.00
4	1589.2	140.4	.0 .0 .0	703.5	28.40	18.89
5	2145.9	207.8	.1 .3 .7	233.3	22.30	7.35
6	10050.0	526.8	.4 .1 .2	2712.2	19.02	9.88

The explanation of the differences observed in the pairs of variances are due to the form of the observed series and the fit of the time series model that is assumed for these general types of series. In the next section, arguments will be presented in explanation of Table XI.

The Observed Series and Autocorrelation

The form of the observed series can be hypothesized from inspection of the autocorrelation coefficients for consecutive lag values provided that the forms are restricted to combinations of random, trend, and cyclical variations.

The autocorrelation coefficients for lag values 1 through 12 for each of the six series presented on pages 71 through 76 are given in Table XII.

TABLE XII

AUTOCORRELATION COEFFICIENTS FOR LAG VALUES 1 THROUGH 12

Series / Lag	1	2	3	4	5	6
0	1.0	1.0	1.0	1.0	1.0	1.0
1	0.092	0.901	0.455	0.567	0.858	0.822
2	0.042	0.795	0.463	0.239	0.770	0.720
3	0.162	0.693	0.370	-0.017	0.698	0.551
4	-0.034	0.583	0.231	-0.359	0.662	0.364
5	0.132	0.502	0.431	-0.569	0.664	0.262
6	0.061	0.433	0.262	-0.636	0.656	0.176
7	-0.026	0.364	0.426	-0.564	0.656	0.112
8	-0.064	0.319	0.481	-0.263	0.648	0.155
9	-0.039	0.266	0.535	0.028	0.699	0.106
10	-0.027	0.206	0.933	0.322	0.776	0.105
11	-0.085	0.154	0.505	0.579	0.851	0.071
12	0.093	0.105	0.501	0.651	0.912	0.052

The data presented in Tables XI and XII will be discussed for each of the series in turn. The measure of the significance of the autocorrelation coefficients is usually compared to those obtained from a normally distributed random series. Anderson (11) provides an approximate value of r_k that can be used to test for a statistically significant non-random series. The value of r_k changes with the number of observations included in the computation and the lag value used. Therefore, for the purpose of discussing those example series given, a conservative value of $r_k \geq |0.2|$ will be considered indicative of a non-random series with significant autocorrelation.

Some insight may be gained by referring to the graphs of these series on pages 71 through 76 during the discussion of the inferences that are drawn from the autocorrelation statistics. Visual comparisons may be made by taking a fixed interval equal to one of the lag values and moving it along the graph of the series to determine a qualitative estimate of these relationships.

Series 1 is a series of a random nature since the r_k are all less than 0.2 in absolute magnitude. This could have been assumed by looking at the graph of the series. However, some periodicities may not be obvious from a cursory examination. Therefore, the best estimate of the future activity of this process is the mean or exponential smoothing with a very small smoothing constant. This is confirmed by the value of the variance of the series being less than the error variance for series 1 in Table XI. The optimum smoothing constants reflect this in that the random and trend constants are both zero which implies that the $\overline{X}_o$ and $\overline{Y}_o$ estimates of the process provide minimum error for this model. However, the smoothing constant for the seasonal variation was not zero since the series is not lengthy enough for the values of $\overline{Z}_i$, $i = -11, \ldots, 0$ to approach one. This made it necessary to use some smoothing in an attempt to smooth these estimates. By using the assumed model, there are forecast that are even more erroneous due to the negative values of some r_k in this series. This causes the error variance to become greater than the series variance. Based upon these observations, this series is of a random form.

Series 2 has very high autocorrelation coefficient for small lag values that diminish at a near linear rate for the larger lag values. This would imply that the more recent observations are the best estimates of what may be expected of the process in the future. Since r_{12} is small enough to be from a random series, this would indicate a lack of seasonal variation. The $r_2, \ldots, r_{11}$ decrease monotonically which indicates a lack of trend. If a trend were present, the r_k would expect to approach some common value whose magnitude would be dependent upon the evidence of a trend.

The smoothing constants reflect the analysis presented for the r_k since there is such a high degree of autocorrelation for lag one. There is little reason to improve upon the estimates of the future over the last observed value. The smoothing constants (1.0 0.0 0.0) provide this type of forecasts since the estimates of the trend, $\overline{Y}_0$, and seasonal effects, $\overline{Z}_i$, i = -11, ..., 0, are not changed in time by the smoothing and the changes in the $FX_{t,1}$ are directly dependent upon the most recently observed value, X_t. Since C can take on any value without affecting the results, the problem of smoothing the initial estimates of the seasonal factors is not a point of consideration in this series as it was in series 1.

The use of A = 1.0 allows the model to assume either of the simpler forms of variation about some mean level or about a linear trend since the seasonal variation is removed from consideration by this selection of A. This type of correlation along with A = 1.0 accounts for the error variance being less than one-third of the series variance.

Series 3 has significant autocorrelation that is relatively constant with the exception of r_4, r_6, and r_{10}. As indicated in the discussion of series 2, the relatively constant values of r_k around 0.45 indicates a trend in the series. Essentially, this states that for any lag value, the length of runs of data on one side of the mean considered is relatively long for a trend type series. This may be shown by computing r_k for a linear trend without any noise. The frequency of occurrence of identical algebraic signs in the numerator of Equation (55) will determine the magnitude of r_k. The value of r_k is an indication of the significance of the trend effects compared to other influences in the series.

Aside from the evidence of a trend with some random variation, there is a major point demonstrated in this particular series by the large value of r_{10}. This is, these data are not analyzed properly if a period of length 12 is used. The series is periodic for a period of 10 time intervals which is well established by the magnitude of r_{10} and by looking at the graph of the series. These data should be analyzed with the same basic time series model, but for P = 10.

The smoothing constants selected upon the basis of P = 12 illustrate the attempt of the model to compensate for this pattern but does not provide an estimate of error variance that could be expected if P were equal to 10. Since the values of A and B are zero as would be expected under the normal yearly data analysis and the initial determination of trend and mean value were appropriate for the series. However, this would change for P = 10. The value of C is small since the seasonal variation shifts time units each year and the model attempts to adjust each year, but to the model these are unstable seasonal variations and are weighted accordingly. This series has a trend with random variations and a seasonal variation for a period of 10 units in length. Two particular points that are worthy of note are that constant values of r_k indicate a trend and a singly large value of r_k is indicative of a cyclical variation for period of length, k.

If this series is judged on the basis of P = 12, then its form would be analyzed as a series with linear trend and random variation superimposed. If a P of 10 is used, the series is of the form of a linear trend and seasonal variation with noise superimposed.

Series 4 is another general type of non-random series. The magnitude of the r_k is important for determining the form of any series, but again the pattern of the r_k is of major importance. There are oscillations of the r_k from a significant positive value to a significant negative value and back to the positive value. This pattern is not difficult to decipher since the r_k is indicative of the relative likeness of the values of the series at intervals of the given lag. Therefore, with the fluctuations in a general pattern, the oscillations or the seasonal variation is illustrated in the r_k themself. The magnitude of r_k does provide some qualitative measure of the amount of randomness that is present in the process. Generally, the larger the amplitude of this oscillation in the r_k the more pronounced the seasonal variation. Smoothing constants selected indicate that the series is rather stable and initial determination of the pattern of the series provides the least error variance for forecast lag of one unit of time. This fit of the data by the basic arithmetic procedures provides a reduction in the error variance over the series variance by more than one-half.

Series 5 is a general class of series for which the model presented in this thesis provides the best results. By inspection of the graph on page 75, this series could be described as one with a trend and multiplicative seasonal variation. The r_k are similarly indicative of this type of series. This series combines the features of series 3 and 4, as shown by the r_k remaining nearly constant with the pattern of the seasonal variation superimposed upon the trend. The magnitude of r_{12} is indicative of the continuing seasonal influence in the series. Other than pointing out the obvious values in Table XII which provide bases of comparison for r_k and the graphical display of the series, nothing can be added to prior discussions about how the r_k reflect the series proper.

A more interesting point provided by this particular series is the selection of the smoothing constants for Table XI. The set of constants shown in the table are not that set which provides the minimum value of the error variance. This particular series has three local minimums in its error variance table. After applying the 0.1 increment grid in these local regions, the minimum points in each region were (0.2 0.2 0.1), (0.1 0.3 0.7), and (0.1 0.1 1.0) with error variances of 422.46, 233.28, and 189.70, respectively. This made it necessary to choose the constants that provided the least change in the error as a result of their being in a more optimum region or less sensitive region. The constants were selected after consideration of the local minimums created by the seasonal factors compensating for the trend. It was primarily for this reason that the second set of constants were selected. These appeared to provide more balance in the weighting of trend and seasonal effects. The interest is in providing the best estimates for forecasting the future.

The first set of constants relies heavily upon the initial determination of the series pattern and the historical data. These will not allow the model input to change quickly enough to keep pace with the process. The third set includes C = 1.0, which emphasizes the importance of the seasonal variation and appears to account for some part of the trend in its estimates of the smoothed seasonal variation. This will provide good forecast results as long as the process continues to be of this form; however, a slight deviation or change and this set of constants will not respond properly.

Series 6 is similar to series 2 in that it does not have a trend or seasonal variation that is discernible by by inspection of $r_1, \ldots r_{12}$, but it does have large values of r_k for $k < 5$. The values of r_k do not decrease in a linear fashion, but are more of an exponentially decreasing form. This indicates more randomness associated with the variation about the process level.

From the analysis above of the form of the series as established by the appearance of the r_k, the smoothing constants would be expected to be similar to those of series 2. However, they are not similar and this would appear to be one of those cases where the selection of smoothing constants provides a better fit of the data than would normally be expected. The explanation of those constants listed in Table XI is that they provide a best fit of the data and consequently are able to reduce the error variance to a value less than one-third of the series variance.

These six series are representative of the unstable type series that are found in economic processes. The foregoing discussion was presented in order that the following discussion may be more coherent. As indicated by earlier discussion and supported by the above examples, the more closely that a series conforms to the type of seasonal and trend effects assumed by the model, the smaller the forecast errors. The basic concept is to reduce the error variance through proper selection of the smoothing constants. A major part of the discussion pertaining to the relationship between smoothing constants and autocorrelation has been derived from observation of the test series and experimentation with the model.

Selection of Constants by Sums of Error Variances

For the computational procedure that uses a forecast lag of one time unit for purposes of error variance computation, there have been two methods of smoothing constant selection discussed above. One of these methods is based upon considerations of minimum values without considering the sensitivity in that region. The other method involves manual manipulation or decisions external to the computer. The prime reason for providing an additional means of constant selection is the chance for the interaction of the constants and the series to produce a minimum value of the error variance in a highly sensitive region of the error variance response. This selection procedure may be performed entirely by a computer.

The method that was proposed in Chapter III to aleviate this condition was the summation of the error variances for each value of the smoothing constant over the values of the other constants. The arguments presented were on the basis that the optimum region would influence the sums of the error variances more than isolated points could be expected to reduce the sum.

For series 6, the results of these summations are displayed in Table XIII.

By inspection, the set of constants which has the minimum sum in each column is (0.6 0.2 0.0). The sum for C = 0.2 is approximately equal to the minimum sum at C = 0.0 and this is worthy of mentioning at this point. The minimum value from Table VII occurs for smoothing constants of (0.2 0.0 0.8) and was shown in Tables VIII and IX to be less desirable than the local minimum which has smoothing constants of (0.4 0.2 0.1). The values provided by this example do not coincide exactly with this local minimum; however, they are in the same local region as defined earlier and are not in the local region of the minimum value. The difference in the coefficient of variation for selected constants in Table XI and those chosen from Table XIII is 1.21% or increases from 9.88 to 11.087. If a secondary grid is used about the set of constants for minimum sums, the local minimum resulting from that search reduces this difference to near zero.

TABLE XIII

SUMS OF ERROR VARIANCES FOR SERIES 6

Constant	Sum on A	Sum on B	Sum on C
0.0	1635607	414654	454031
0.2	463643	408825	454366
0.4	244545	450737	461707
0.6	174433	502213	485070
0.8	208216	598314	547212
1.0	368804	720505	692862
Total	3095248	3095248	3095248

This method of constant selection for the other five example series located the optimum set exactly for series 2, 3, and 4 and was an adjacent point in the coarse search grid to the selected optimum values for series 1 and 5.

For the basic computational procedure of constant selection based upon errors for forecasts with a lag of one this procedure of error summation has been demonstrated as superior to selection of constants based upon minimum error alone. If there is only one minimum in the error variance table, this method will locate that region and if there is more than one local minimum it will select the optimum region as determined by the relative sensitivity of the error variance to changes in the smoothing constants. This hypothesis is supported by the additional test series that were used in the empirical derivation of these smoothing constant selection methods.

Selection of Smoothing Constants Based Upon the Sum of Error Variances for Consecutive Lag Values

The methods of the preceding section were developed for the selection of the optimum set of smoothing constants based upon error variances computed for forecasts of lag one. Since the forecast for the series is usually made a number of periods into the future, it follows that a more conservative type of testing the forecasts for the series would be to use consecutive lag values for forecasts over the test series. This provides an estimate of the error variance for each lag value. The tests of this procedure used lag values of one through 12.

The smoothing constants selected after computation of the error variances associated with the consecutive lag values are those associated with the minimum sum of these error variances. If the immediate forecast values are considered more important than the more distant future values of the forecast, these values may be weighted to provide this as an inherent consideration of the computation process. The improvement in the over-all optimality of the smoothing constants for the total length of forecast may be seen by comparisons provided in Table XIV. The comparisons shown are between those error variances of the minimum sum for consecutive lag values and the optimum constants from the error variance tables for forecasts of lag one. The values for series 3 and 4 are not shown since both sets of constants are identical and the error variances are as shown in Table XI.

In addition to illustrating an improvement of the method of selecting a set of smoothing constants, Table XIV demonstrates the applicability of evaluating the lag of the forecast and integrating the consideration of future time by appropriate weighting of the error variances. Table XIV also provides information on the form of the process and the noise associated with the process. If the series is of the form assumed by the model, the error variances are less than those for a more complex process.

Economic time series generally have noise superimposed upon the base series. If the base series can be defined,

TABLE XIV

COMPARISON OF ERROR VARIANCES BETWEEN METHODS OF SMOOTHING CONSTANT SELECTION

Series	1	1	2	2	5	5	6	6
A	0.2	0.0	6.0	1.0	0.0	0.0	0.2	0.6
B	0.0	0.0	0.0	0.0	0.0	0.2	0.2	0.2
C	0.0	0.2	0.0	0.0	0.0	0.2	0.0	0.0
Selection by	C L*	M S**	C L	M S	C L	M S	C L	M S
Lag								
1	686	625	194	153	521	427	3242	2777
2	558	625	383	394	521	503	3490	2526
3	620	625	496	538	521	555	4895	4261
4	628	625	562	574	521	594	5725	4307
5	629	625	663	652	521	583	7600	5864
6	605	625	794	778	521	676	9120	7014
7	637	625	909	947	521	778	12159	12783
8	573	625	931	1012	521	833	13654	14905
9	588	625	888	941	521	751	16922	22364
10	594	625	879	920	521	706	18491	25864
11	529	625	857	927	521	705	20961	31862
12	666	625	787	853	521	819	22625	35654

*C L represents the selection of constants by the minimum value of error variance for values 1 through 12.

**M S represents the selection of constants by the minimum sums of error variance for lag of one.

then the errors in the extrapolation of that base series is dependent upon the distribution of the noise in the series. If, in fact, the base series as determined extends through the period of the forecast, the accuracy of the forecast cannot be expected to be better than the noise level present in the process. However, if the base series is defined by the model, the inherent process variance is approximated by the error variance. Under this condition, the error variance should be relatively constant about the level of the inherent process variance for all lag values.

Therefore, by inspection of the results produced by this type of smoothing constant selection, information on the form of the observed series is provided as a secondary benefit of the forecasting method. Series 1, 3, 4, and 5 for this method of smoothing constant selection have almost constant error variances for all lag values in each case. From the above discussion, this would indicate that the process is of the form of the model or one of its simpler forms. This contention is also supported by the discussion of the form of these series in the section on autocorrelation. The error variances for these series are estimates of the inherent process variance and provide measures of the forecast error that can be expected in the forecast of future values.

Based upon the methods above, series 2 and 6 are not of the form of the model or its simple forms. This is also supported by the discussion of these series in the section on autocorrelation. In these two cases, this method of smoothing constant selection provides the information that the series is not of the form of the model, and indicates the magnitude of the error variance that may be expected from the best fit of the model and the data. For series 2, the error variance for forecast more than two intervals into the future would be expected to be approximately four times that for forecasts one period into the future if the same smoothing constants are used. A similar analysis holds for series 6.

Two important points are illustrated in this section. One point is that the error variance and the forecast lag should be considered concurrently when selecting the smoothing constants. The other point is that the agreement of the model and the generating process is a major influence upon the magnitude of the observed error.

Selection of Smoothing Constants for Individual Lag Values

In the preceding section, the method of selecting smoothing constants provides a single set of constants for all values of the forecast lag. The use of a single set of smoothing constants is not a requirement of the model and the reduction in the error variance for forecast values may be significant if individual sets of constants are selected for each lag value that is to be used in actual forecasting practice.

The major part of the computations necessary for this smoothing constant selection process is performed by the computations associated with the methods of the previous section. In order to use this method of selection it is necessary to determine the minimum value of the error variance for each lag value used and its associated set of smoothing constants. The values for the six series used in this chapter are displayed in Table XV.

The improvements provided by this method of smoothing constant selection over the one outlined in the previous section are that a given set of constants which provide relatively small changes in the error variance for some lags may cause the relative error for other values of the lag to be multiples of their minimum value. This factor of consideration is illustrated by comparing the results for series 6 in Tables XIV and XV. This method will also provide a minimum sum of consecutive lag error variances equal to or less than that method of the previous section. The error variances for different lag values may be weighted before they are chosen, but it will not affect the selection under this method. Therefore, if the computations for the method given in the previous section uses weighted error variances, it will not affect the methods of this section.

The method of selection for each lag value is on a minimum value basis, but the sum on factors method described earlier may be used for each of the lag values

TABLE XV

MINIMUM ERROR VARIANCE AND SMOOTHING CONSTANTS FOR INDIVIDUAL LAG VALUES

Lag	Smoothing Constants	Error Variance	Smoothing Constants	Error Variance	Smoothing Constants	Error Variance
Series	1		2		3	
1	0 0 2*	625	10 0 0	153	0 0 2	1230
2	4 0 0	548	6 0 2	382	0 0 2	1230
3	2 0 0	620	0 0 0	422	0 0 2	1230
4	0 0 2	625	0 0 0	422	0 0 2	1230
5	0 0 2	625	0 0 0	422	0 0 2	1230
6	2 0 0	605	0 0 0	422	0 0 2	1230
7	0 0 2	625	0 0 0	422	2 0 2	1084
8	2 0 0	573	0 0 0	422	2 0 2	1156
9	2 0 0	588	0 0 0	422	2 2 0	918
10	2 0 0	594	0 0 0	422	10 0 0	260
11	6 0 0	439	0 0 0	422	0 0 2	1230
12	0 0 2	625	0 0 0	422	0 0 2	1230
Series	4		5		6	
1	0 0 0	703	2 10 0	418	2 8 0	2553
2	0 0 0	703	2 2 2	485	2 10 0	2344
3	0 0 0	703	2 2 2	517	2 8 0	4014
4	0 0 0	703	0 0 0	521	4 2 2	4303
5	0 0 0	703	2 2 2	504	4 2 4	5468
6	0 0 0	703	2 2 2	500	4 2 8	5368
7	0 0 0	703	0 0 0	521	4 2 10	8904
8	0 0 0	703	0 0 0	521	2 8 4	9951
9	0 0 0	703	0 0 0	521	2 2 0	16921
10	0 0 0	703	0 0 0	521	2 2 0	18491
11	0 0 0	703	0 0 0	521	2 2 0	20961
12	0 0 0	703	0 0 0	521	2 2 0	22625

*These values are 10 times the smoothing constants.

or the manual method of selection is possible but less practical since the amount of data increases directly proportional to the number of lag values used.

There is one interesting point that is displayed in Table XV which was emphasized earlier. This is the length of the period for series 3. For a lag of 10, the error variance is approximately one-fourth of that for the other lag values. The constants (1.0 0.0 0.0) essentially remove the consideration of the length of the period and the forecast value 10 intervals into the future is the present observed value plus 10 times the original trend value, $\overline{Y}_o$. The value, 260, given in the table is not necessarily the minimum value of the error variance that would be obtained if P = 10 had been used in the model for the analysis of series 3 data.

Selection of the Starting Point

The philosophy behind the deletion of historical data in order to improve the results of the forecasts was given in Chapter III. Three of the series did not require new starting points in order to minimize the absolute sum of the third order differences. Those series for which the starting point within the historical data was changed are given in Table XVI along with some comparative values to demonstrate the changes due to this procedure.

TABLE XVI

DELETION OF HISTORICAL DATA AND THE AFFECT ON ERROR VARIANCES

Series	Sum of Third Differences Orig.	Min.	Points Deleted	Smoothing Constants Old	New	Sum of Error Variances	Per Cent Reduction
2	588	368	5	6 0 0*	0 0 4	5298	36.5
5	1674	1101	7	0 0 0	0 0 0	5291	15.4
6	3055	1972	7	2 0 0	2 2 0	125881	9.4

*These values are 10 times the smoothing constants.

The emphasis of this thesis has been on the reduction of the forecast error through improved methods of selecting the smoothing constants. This particular modification in the analysis of the historical data is for the same purpose. The idea that all historical data is of value in the forecasting process is not entirely true as indicated by the discussion in Chapter III. The purpose of using the data is to obtain estimates of the future activity of the process. If some procedure provides accurate estimates of the future without using any of the historical data, then that procedure should be used in lieu of time series analysis. Therefore, the historical data is simply a means of providing information for use in forecasting the future activity of a given time series. With reduced error variance as a prime objective, it is not unreasonable to justify the methods of this section on that basis alone.

The data dsiplayed in Table XVI is indicative of the results that have been obtained for the various test series and actual problems that have been considered during the course of this investigation. There have been some isolated cases where the methods of this section did not improve the forecasts in terms of the error variance, but upon further investigation they were found to generally be of a more extreme deviation from the assumed form of the model. It is not desired to limit the generality of the procedure at this point, but simply to point out another qualitative test on the form of the series and the results that may be expected for that series.

Whether this point is basic and ignored, or whether the experimentation with adding and deleting historical data from the consideration of the model has not been investigated, this particular point seems to be missing from the published articles in this area. Even in the more obvious cases of curve fitting as a method of interpolating and extrapolating the time series, the selection of the origin within the data was not given as a prime consideration of the method. However, due to the reductions in the length of the data file along with the improvement in the computed error variance, this method cannot be ignored due to the frequency of the occurrence of the improvements in forecast error. It is also of interest that the entire grid of error variance values in the investigation of the combinations of the factors and the various lag values change significantly which essentially indicates that the model and the computational algorithms recognize the segment of the historical data as new series and treat it accordingly. This particular reduction in the forecast error for this model and its supporting algorithms are felt to be an appropriate way to end this chapter on empirical results. This development was primarily an empirically derived idea, which was later supported by the model formulation and curve fitting techniques which provided a heuristic proof of this contribution to the methods of time series analysis.

CHAPTER V

OBSERVATIONS AND CONCLUSIONS

Selection of Forecasting Methods

The selection of one forecasting method over another is primarily a matter of personal opinion. The model and the supporting algorithms presented in this thesis were developed to serve a general type of problem in the time series area. It is that of forecasting the economic type of time series which is assumed to be composed of a trend, seasonal variations and random variations superimposed upon some process level. The arguments and examples presented in support of this procedure for forecasting call attention to those points that should be considered in evaluating this particular procedure for application to a specific type of time series data. It is on the basis of demonstrated results that the methods are presented and justified.

The techniques for time series analysis that have been presented in this thesis start with the simple smoothing model which has only a single smoothing constant and the forecast is a constant value. The last model and supporting algorithms presented require the scanning of the data for selection of the optimum starting point within the data; the computation of the search grid of values for each of the lag values; and the selection of the 12 optimum sets of three smoothing constants each or 36 values from the enumeration of a 13 by 156 grid of smoothing constant combinations and the associated error variances.

Within this range of techniques, there have been a number of procedures presented to aid in forecasting the future activity of a process. The methods presented have been arranged according to the computational effort required. Generally, the increase in computational effort provides a corresponding increase in the information provided the forecaster as to the type of process, the fit of the model and the possible magnitude of the forecast errors. The selection of the particular model for analyzing a given time series should be evaluated in terms of the economics of computational effort and information provided. The increase in the information provided the forecaster is sufficient to make this consideration a necessary one. One point that should be re-emphasized is the consideration of the forecast error for a fixed number of points in the future may significantly influence the choice of forecasting methods; therefore, the weighting of error variances should be considered carefully.

The attempt has been to present a general type of computational form for the applied type of forecasting that is becoming more routine in the various fields. One point of interest is the significant influence of the autocorrelation upon the results that may be expected from the forecasting procedure. This may affect which method provides the least amount of forecast error.

The methods presented here are felt to serve the needs of generality and computational form which will permit their application on a routine basis to a broad class of problems in the field of economic time series.

Forecasting and Forecast Error

As stated earlier, a forecast without an estimate of the possible error is incomplete since the forecast is an estimate of the process without the noise superimposed. A major point to consider is whether the noise is greater than the process level or the forecast value. This has been the purpose of determining the optimal estimate of the inherent process variance in order that the estimates of the error variance are as nearly correct as is conveniently possible and not mislead the forecaster. Therefore, from the standpoint of the error variance being a good estimate of the process variance, the methods presented in this study are felt to be correct under the conditions being considered and the mathematical techniques that are employed. In general the errors associated with the forecasts of this model are not independently normally distributed, but may be from some non-normal type of distribution which has dependence among the observations since the errors of the forecast are autocorrelated. This makes the use of the two-sided tolerance limits for the normal distribution statistically unacceptable for the purposes of placing formal statistical limits upon these forecasts. However, it is suggested that the square root of the error variance for a given lag value be used as an estimate of the variation about that point of the forecast value. The effectiveness of this particular set of limits is dependent upon the process, the statistical properties of the error distribution and the goodness of the approximation of the noise by the error variance. But, for those series that were studied in the preparation of this thesis, this performed as a reasonable rule of thumb. These limits may serve in the same manner as the control chart philosophy used in quality control in that observed values beyond the limits placed on the forecast provide sufficient reason for investigating the data or the process to account for this deviation from the forecast value. In order to improve confidence in the forecasting process, all old limits should be displayed in order to measure the performance of past forecasts and the correctness of the limits on those forecasts. These limits can be adjusted to some suitable multiple of the computed error variance for the individual process under consideration.

For some operations, it may not be desirable to revise the entire forecast at the end of each period and the same smoothing constants can be used for each subsequent interval of the process without recomputing the smoothing constants and the forecast. In this case, the error variances could be used over again as limits until some out-of-control condition indicated that the process had changed and the data needed to be recomputed by regular procedures for the revision of the smoothing constants, model parameters, and estimate of the expected error.

Summary

The objective of this thesis was to improve the forecast error of the exponentially weighted moving averages time series model by the optimum selection of smoothing constants. The condition of optimality is that which provides the best estimate of the inherent process error associated with the process under consideration and the smallest error variance possible for consideration of the relationship between the forecast error and the lead time of the forecast.

In order to provide a base for the optimization of the error variance, the following models were developed: the simple exponential smoothing model; the model for trend and random noise; the model for trend, cyclical, and random variations; and the preceding model with provisions for selection of the starting point. All of these models are of the exponentially weighted moving average(s) type. During the course of these investigations, there were some results that proved of more interest than others, but all contribute to the general success of this method of forecasting.

Some of the major points include:

(a) the illustration that the minimum error and the optimum error are not necessarily the same value,

(b) the error sensitivity is low in the region of optimum error variance,

(c) the location of the optimum region may be determined by the use of manual or automated means,

(d) the method of independent sums of error variances over smoothing constants may be used as the basis for selection of the optimum region,

(e) the use of a multistage search procedure may be used in the interest of saving computational effort,

(f) the lag value is a prime consideration in the selection of the smoothing constants and that more than one lag value should be investigated to improve the estimates of the error and to improve the confidence in the model. The use of the method of consecutive lags pointed out the necessity of the consideration of the lead time of the forecast.

The ultimate use of the reduction in the error variance is to improve the confidence in the model and to place limits that will monitor the forecasting process for detectable changes in the process that may be of significant value to the user of the forecast model. The three most important points are the selection of a starting point, the optimality concept and the individual sets of smoothing constants for each lag value of the forecast model.

Extension of This Study

Probably the most challenging area of this thesis that needs further study is the statistical development of methods to evaluate the form of the distribution of the error and the appropriate development of limits for the forecast in order to estimate the error that may be expected. The other major area of possible improvement is in the basic problem of the decomposition of the observed series into other components, thereby reducing the amount of noise associated with the process. One additional endeavor that could be considered as an extension would be the proof of this particular model formulation by the use of the model over a wide range of problems. At present there are two companies using this particular model, one in the study of personnel action, the other to study an inventory problem. The proof will be completed with the use of the concepts developed here and, therefore, must be delayed until some later date.

BIBLIOGRAPHY

(1) Davis, H. T. *The Analysis of Economic Time Series*, Monograph, No. 6., Cowles Commission. Bloomington, Indiana: Principia Press, 1941.

(2) Yule, G. U. "Why Do We Sometimes Get Nonsense Correlations Between Time Series," *Journal of the Royal Statistical Society*, Vol. 89 (new series), 1926, pp. 1-64.

(3) Slutzky, E. "The Summation of Random Causes as the Source of Cyclic Processes," *Econometrica*, 5, April, 1937, pp. 105-146.

(4) Macaulay, F. R. "The Smoothing of Time Series," Publication of the National Bureau of Economic Research, Incorporated, No. 19, 1931, pp. 73-74.

(5) Deming, L. S. "Selected Bibliography of Statistical Literature, 1930 to 1957: II. Time Series," *Journal of Research* of the National Bureau of Standards - B. Mathematics and Mathematical Physics, Vol. 64B, No. 1, January-March, 1960.

(6) Magee, J. F. *Production Planning and Inventory Control*. New York: McGraw-Hill, 1958.

(7) Brown, R. G. *Statistical Forecasting for Inventory Control*. New York: McGraw-Hill, 1959.

(8) Holt, C. C., F. Modigliani, J. F. Muth, and H. A. Simon. *Planning Production, Inventories, and Work Force*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1960.

(9) Dodd, E. L. "The Problem of Assigning a Length to the Cycle to be Found in a Simple Moving Average and in a Double Moving Average of Chance Data," *Biometrica*, Vol. 9, 1941, pp. 25-37.

(10) Brown, R. G., and R. F. Meyer. "The Fundamental Theorem of Exponential Smoothing," *Operations Research*, 9, October, 1961, pp. 673-685.

(11) Anderson, R. L. "The Distribution of Circular Serial Coefficients," Annals of Mathematical Statistics XIII, 1942, pp. 1-13.

(12) Forrester, J. W. Industrial Dynamics. Cambridge: The Massachusetts Institute of Technology Press, 1961.

(13) Grenander, U., and M. Rosenblatt. Statistical Analysis of Stationary Time Series. New York: John Wiley and Sons, 1957.

(14) Hald, A. The Decomposition of a Series of Observations Composed of a Trend, a Periodic Movement, and a Stochastic Variable. Copenhagen: Gads Forlag, 1948.

(15) Hannan, E. J. Time Series Analysis. New York: John Wiley and Sons, Inc., 1960.

(16) Kendall, M. G. "The Effect of the Elimination of Trent on Oscillations in Time Series," Journal of the Royal Statistical Society, Vol. 104, 1941, pp. 43-52.

(17) Muth, J. F. "Optimal Properties of Exponential Weighted Forecasts of Time Series With Permanent and Transitory Components," Journal of the American Statistical Association, 55, June, 1960, pp. 299-306.

(18) Parzen, E. "An Approach to Time Series Analysis," Annals of Mathematical Statistics, 32, December, 1961, pp. 951-989.

(19) Wiener, N. Extrapolation, Interpolation, and Smoothing of Stationary Time Series With Engineering Applications. New York: John Wiley and Sons, Inc., 1949.

(20) Wylie, C. R., Jr. Advanced Engineering Mathematics. 2nd ed. New York: McGraw-Hill Book Company, Inc., 1960.

APPENDIX

DISCUSSION OF COMPUTER PROGRAMS

INTRODUCTION

The IBM 7090 and IBM 1620 FORTRAN programs presented in this appendix have been run on the IBM 704, IBM 650 and CDC 1604 with certain modifications. The computer program is simply the automated computation scheme for the supporting algorithms and forecasting model that has been presented in the text of this thesis. Since there are a number of options that may be exercised through the use of the computer, this appendix is continuous in that all the parts are related but for purposes of clarity these options will be discussed in turn. The major divisions of the material to be presented will be as follows:

I. The Complete FORTRAN Program

II. Basic Program for Lag 1 and Minimum Error Variance

III. Option - Sense Switch 1 for Enumeration of Grid and Error Variances

IV. Option - Sense Switch 2 for Graphic Display

V. Option - Sense Switch 3 for Selection of Constants Based Upon Independent Sums

VI. Option - Sense Switch 4 for Selection of Constants Based Upon Consecutive Lags

VII. Option - Sense Switch 5 for Selection of Optimum Starting Point

VIII. Option - Sense Switch 6 for External Control of Smoothing Constants

IX. Option - Sense Light 4 for Selection of Individual Sets of Constants for Each Lag Value

X. Table of Option Combinations

XI. Preparation of Data

XII. IBM 1620 Programs.

Through a more detailed explanation of the programs written for the time series analysis and extrapolation model, it is hoped to provide a working knowledge of the various features of the program. The data that are used in the displays on the following pages are those presented in Tables I through VI. The format of this Appendix is to discuss verbally the provisions of each option and present the related display.

SECTION I

THE COMPLETE FORTRAN PROGRAM

The listing of the program that follows this discussion on the next eight pages is the source program which contains all the programming for the options that will be explained in the other sections. This program was written with other users in mind, and a number of aids are incorporated in the program for easy modification by those who wish to do so. After changes have been made, however, do not attribute errors to the original program as listed here. Those aids that will help in the reading of the program are in the form of "Comment" cards placed in the program at the entry point of each of the options. In addition, for each of the options that have been inserted into the basic program, a new numbering sequence is used. The basic program uses statement numbers less than 100, the sense switch 3 option uses numbers between 100 and 199, sense switch 6 option uses numbers between 200 and 299, the option of sense switch 5 uses those numbers between 300 and 399, and similarly the option of sense switch 4 uses those numbers between 400 and 499. In addition to these tracing aids in the form of the statement numbers, the numenoic code that is used should aid in the understanding of the program. Some examples of the code are MOTOL for month total, DEL for the difference operator, COMP for compare, MYR for years, YRAV for year average, SEA for seasonal, EST for estimate, TREND for trend, OBE for observed, VAR for variance, and SIGS for smallest sigma or square root of the error variance. The complete program was rewritten in order to place all statement numbers in consecutive and ascending order within each numerical sequence used. This aids in discussion of the program and in locating points that will be referenced in this Appendix.

```
*      DATE
*      PROGRAM TIMES, RETURN TO GLEN SELF
*      XEQ
*      LIST8
CTIMES
C      TIME SERIES ANALYSIS AND FORECAST USING OPTIMAL
C      SMOOTHING CONSTANTS FOR THE EXPONENTIALLY WEIGHTED
C      MOVING AVERAGES MODEL
       DIMENSION OBE(600), OBS(12,50), YRAV(50), TREND(600),
      1UNAD(12), SEA(600), EST(600), PRED(620), PRI(612),
      2SAGE(217), SUMA(6), SUMB(6), SUMC(6), FA(6), FB(6),
      3FC(6), SENA(12), SMOTH(600), SSS(12), SA(12), SB(12),
      4SC(12), SIGIX(12), SEANL(12,50), ARRAY(5), DEL(600),
      5PRIS(12), SEAS(12), XLABEL(28), X(50)
       CALL RELICF(30000)
  800  READ INPUT TAPE 2,1
    1  FORMAT(1H6,51H                                              )
       WRITE OUTPUT TAPE 3,1
       MOTOL = 0
       I = 1
    2  READ INPUT TAPE 2,3,OBE(I)
    3  FORMAT(F10.6)
       IF(OBE(I) - 999.999999) 4,5,99
    4  MOTOL = MOTOL + 1
       I = I + 1
       GO TO 2
    5  MYR = MOTOL/12
C      IF SENSE SWITCH 5 IS ON, THE PROGRAM WILL SELECT
C      THE OPTIMAL STARTING POINT WITHIN THE HISTORICAL DATA
C      AND MAKE THE DELETIONS BEFORE CONTINUING THE PROGRAM
       IF(SENSE SWITCH 5)300,311
  300  MTA1=12*(MYR-1)
       MTA2=MTA1-1
       MTA3=MTA2-1
       COMP=0.9E+25
       DO 301 I=2,MTA1
  301  DEL(I-1)=OBE(I)-OBE(I-1)
       DO 302 I=2,MTA2
  302  DEL(I-1)=DEL(I)-DEL(I-1)
       DO 303 I=2,MTA3
  303  DEL(I-1)=DEL(I)-DEL(I-1)
       WRITE OUTPUT TAPE 3,304
  304  FORMAT(1H0,9X,23HFIRST DATA    SUM  OF  /
      110X,24HPOINT USED   3RD DIF    )
       DO 308 IN=1,12
       SUMDL=0.0
       MTA4=IN+12*(MYR-2)-4
       DO 305 I=IN,MTA4,12
       DO 305 J=1,9
       LL=(I+J-1)
       ADEL=DEL(LL)
  305  SUMDL=SUMDL+ABSF(ADEL)
       WRITE OUTPUT TAPE 3,306,IN,SUMDL
```

```
306 FORMAT(1H ,I15,F16.5)
    IF(COMP-SUMDL)308,308,307
307 COMP=SUMDL
    INIT=IN
308 CONTINUE
    WRITE OUTPUT TAPE 3,309,INIT
309 FORMAT(1H0,9X,31HTHOSE DATA POINTS PRECEDING THE I3
   1,26HTH UNIT HAVE BEEN DELETED )
    DO 310 I=INIT,MOTOL
    L=(I-INIT+1)
310 OBE(L)=OBE(I)
    MOTOL=MOTOL-INIT+1
    MYR=MOTOL/12
311 IF(MOTOL-12*MYR-9)6,7,7
  6 MOT = 12*(MYR-1)
    MYR = MYR - 1
    GO TO 8
  7 MOT = 12*MYR
  8 FMYR = MYR
    L=1
    DO 10 J = 1, MYR
    YRAV(J) = 0.0
    DO 9 I = 1, 12
    UNAD(I) = 0.0
    SSS(I)=(0.9E+25)*(0.9E+25)
    OBS(I,J) = OBE(L)
    YRAV(J) = YRAV(J) + OBE(L)/12.0
  9 L=L+1
 10 CONTINUE
    AMOT = MOT
    TREN = (YRAV(MYR)-YRAV(1))/(AMOT-12.0)
    DO 11 J = 1, MYR
    SUSEA = 0.0
    DO 11 I = 1, 12
    FI = I
    SEANL(I,J) = OBS(I,J)/(YRAV(J)-(6.5-FI)*TREN)
    UNAD(I) = UNAD(I) + (SEANL(I,J)/FMYR)
 11 SUSEA = SUSEA + ABSF(UNAD(I))
    PLY = 12.0/SUSEA
    DO 12 I = 1,12
    SEA(I)=UNAD(I)*PLY
 12 SENA(I)=SEA(I)
    A = 0.0
    B = 0.0
    C = 0.0
C   IF KALL4 IS A 9 THE PROGRAM WILL SELECT INDIVIDUAL
C   SMOOTHING CONSTANTS FOR EACH VALUE OF THE LAG
    READ INPUT TAPE 2,509,KALL4
509 FORMAT(I1)
    IF(KALL4-9)511,510,99
510 SENSE LIGHT 4
511 CONTINUE
```

```
      KONA=0
      SIGS = (0.9E+25)*(0.9E+25)*(0.9E+25)
      IF(SENSE SWITCH 1) 13,15
   13 WRITE OUTPUT TAPE 3,14
   14 FORMAT(1H0,9X,6HRANDOM,8X,8HSEASONAL,8X,5HTREND,6X,
     111HSQUARE ROOT,3X,3HLAG/9X,8HCONSTANT,7X,8HCONSTANT,
     27X,8HCONSTANT,5X,9HOF E.M.S./)
   15 M=1
      IF(SENSE SWITCH 6)200,16
  200 READ INPUT TAPE 2,201,A,B,C
  201 FORMAT(3F10.6)
      SENSE LIGHT 1
      SENSE LIGHT 3
      GO TO 202
   16 EST(1)=A*(OBE(1)/SENA(1))+(1.0-A)*(YRAV(1)+TREN)
      SEA(1)=B*(OBE(1)/EST(1))+(1.0-B)*SENA(1)
      TREND(1)=C*(EST(1)-YRAV(1))+(1.0-C)*TREN
      SMOTH(1) = EST(1)*SENA(1)
      DO 20 I= 2,MOTOL
      IF(I-12)17,17,18
   17 J=I
      SEA(J)=SENA(J)
      GO TO 19
   18 J=I-12
   19 EST(I) = A*(OBE(I)/SEA(J)) + (1.0-A)*(EST(I-1)+TREND
     1(I - 1))
      SEA(I) = B*(OBE(I)/EST(I))+(1.0-B)*SEA(J)
      TREND(I)=C*(EST(I)-EST(I-1))+(1.0-C)*TREND(I-1)
   20 SMOTH(I)=EST(I)*SEA(I)
      LT = 0
      T = 1.0
      NOT = MOT + 1
      LOW = MOT + 1
      MAX = MOTOL
      SISUM = 0.0
      SUMER = 0.0
      GO TO 401
  400 LT = LT + 1
      LOW = NOT - LT
      MAX = MOTOL - LT
      T = T + 1.0
  401 DO 26 K = LOW,MAX
      MAD=K+LT
      PRED(MAD)=(EST(K-1)+T*TREND(K-1))*SEA(MAD-12)
      ERROR=(OBE(MAD)-PRED(MAD))
      ERSQ = ERROR*ERROR
      SUMER = SUMER + ERSQ
      IF(SENSE LIGHT 1) 23,26
   23 IF(SENSE LIGHT 3) 24,26
   24 SENSE LIGHT 1
      SENSE LIGHT 3
      WRITE OUTPUT TAPE 3,25, OBE(MAD),PRED(MAD),ERROR
```

```
   25 FORMAT(1H ,3F19.8)
   26 CONTINUE
      DENOM = MOTOL - 12*MYR
      VAR = SUMER/DENOM
      SIGE = SQRTF(VAR)
      IF(SENSE LIGHT 4)501,503
  501 SENSE LIGHT 4
      SIGIX(LT+1)=SIGE
      IF(SIGIX(LT+1)-SSS(LT+1))502,502,503
  502 SSS(LT+1)=SIGIX(LT+1)
      MI = (MOTOL-11+LT)
      SA(LT+1)=A
      SB(LT+1)=B
      SC(LT+1)=C
      PRIS(LT+1)=(EST(MOTOL)+T*TREND(MOTOL))*SEA(MI)
      SEAS(LT+1)=SEA(MI)
  503 SAGE(M)=SIGE
      M=M+1
      SISUM = SISUM + SIGE
      IF(SENSE LIGHT 4)504,505
  504 SENSE LIGHT 4
      GO TO 36
  505 IF(SENSE LIGHT 1)27,402
   27 IF(SENSE LIGHT 3) 28,402
   28 SENSE LIGHT 1
      SENSE LIGHT 3
      WRITE OUTPUT TAPE 3,29
   29 FORMAT(1HO,8X,12HPERIODS INTO,7X,10HFORECASTED,10X,
     18HSEASONAL  /1H ,9X,10HTHE FUTURE,11X,5HVALUE,13X,
     26HFACTOR  )
      DO 31 N = 1,12
      T = N
      JK = (N-12+MOTOL)
      MOTON = MOTOL+N
      PRI(MOTON) = (EST(MOTOL)+T*TREND(MOTOL))*SEA(JK)
      WRITE OUTPUT TAPE 3,30,N,PRI(MOTON),SEA(JK)
   30 FORMAT(1HO,I14,F24.8,F19.8)
   31 CONTINUE
      MOTON = MOTOL+12
C     IF SENSE SWITCH 4 IS ON, THE SMOOTHING CONSTANTS WILL
C     BE SELECTED ON THE BASIS OF MINIMUM CUMULATIVE SUM FOR
C     LAGS 1 THROUGH 12
  402 IF(SENSE SWITCH 4)36,32
   32 IF(SIGE-SIGS)33,34,36
   33 SIGS = SIGE
      KOUNT = 0
      GO TO 35
   34 KOUNT = KOUNT + 1
   35 AS = A
      BS = B
      CS = C
   36 IF(SENSE SWITCH 1) 37,39
   37 WRITE OUTPUT TAPE 3,38,A,B,C,SIGE,T
```

```
   38 FORMAT(1H ,3F15.4,F15.6,F7.1)
   39 IF(SENSE SWITCH 4)403,410
  403 IF(LT-11)404,405,99
  404 SUMER=0.0
      IF(SENSE LIGHT 1) 40,401
  405 IF(SISUM-SIGS)406,407,407
  406 SIGS=SISUM
      AS=A
      BS=B
      CS=C
  407 IF(SENSE SWITCH 1)408,410
  408 WRITE OUTPUT TAPE 3,409,A,B,C,SISUM
  409 FORMAT(1HO,3F15.4,F15.6,5X,5HTOTAL  / )
  410 IF(SENSE LIGHT 1)40,43
   40 SENSE LIGHT 1
      WRITE OUTPUT TAPE 3,41
   41 FORMAT(1HO, 31H         CHECK ON VALUES USED    /)
      WRITE OUTPUT TAPE 3,38, A,B,C,SIGE
   42 IF(SENSE LIGHT 3) 53,43
   43 IF(0.98-C)44,44,46
   44 IF(0.98-B)45,45,47
   45 IF(0.98-A)49,49,48
   46 C = C + 0.2
      IF(KONA-0)99,44,16
   47 B = B + 0.2
      C = 0.0
      IF(KONA-5)16,49,99
   48 A = A + 0.2
      KONA=KONA+1
      C = 0.0
      B = 0.0
      GO TO 16
   49 IF(SENSE LIGHT 4)506,508
  506 SENSE LIGHT 4
      WRITE OUTPUT TAPE 3,14
      WRITE OUTPUT TAPE 3,507,(SA(M),SB(M),SC(M),SSS(M),M,M=1,12)
  507 FORMAT(1H ,3F15.4,F15.6,I6)
      WRITE OUTPUT TAPE 3,29
      WRITE OUTPUT TAPE 3,30,(M,PRIS(M),SEAS(M),M=1,12)
      GO TO 99
  508 WRITE OUTPUT TAPE 3,50,AS,BS,CS,SIGS
   50 FORMAT(1HO,9X,6HRANDOM,8X,8HSEASONAL8X,5HTREND4X,
     111HSQUARE ROOT /9X,8HCONSTANT,7X,8HCONSTANT,7X,
     28HCONSTANT,3X,9HOF E.M.S.   //3F15.4,E15.6)
      WRITE OUTPUT TAPE 3,51, KOUNT
   51 FORMAT(1HO,9X,34HNUMBER OF POINTS WITH EQUAL ERROR  I6
     1/10X,40HFOR RANDOM CONSTANT = 0 OR 1 THERE ARE 5)
      A = AS
      B = BS
      C = CS
C     IF SENSE SWITCH 3 IS ON, THE CONSTANTS WILL BE
C     SELECTED UPON THE BASIS OF THE MINIMUM SUM FOR EACH
```

```
C     SMOOTHING CONSTANT SUMMED OVER THE OTHER TWO CONSTANTS
      IF(SENSE SWITCH 3) 100,112
  100 KAA=2
      SUMA(1)=6.0*(SAGE(1)+SAGE(2)+SAGE(3)+SAGE(4)+SAGE(5)+
     1SAGE(6))
  101 DO 103 L=7,115,36
      SUMA(KAA) = 0.0
      N = L + 35
      DO 102 M=L,N
  102 SUMA(KAA) = SUMA(KAA) + SAGE(M)
  103 KAA = KAA + 1
      SUMA(6)=6.0*(SAGE(151)+SAGE(152)+SAGE(153)+SAGE(154)+
     1SAGE(155)+SAGE(156))
      KAB=1
      DO 105 M=7,37,6
      SUMB(KAB)=SUMA(6)/6.0
      DO 104 L=M,145,36
  104 SUMB(KAB)=SAGE(L)+SAGE(L+1)+SAGE(L+2)+SAGE(L+3)+
     1SAGE(L+5)+SUMB(KAB)+1.5*SAGE(KAB)
  105 KAB=KAB+1
      DO 107 M=7,12
      SUMC(M-6)=SUMA(1)/6.0
      DO106L=M,216,6
  106 SUMC(M-6)=SUMC(M-6)+SAGE(L)+6.0*SAGE(M+144)
  107 CONTINUE
      FA(1)=0.0
      FA(2)=0.2
      FA(3)=0.4
      FA(4)=0.6
      FA(5)=0.8
      FA(6)=1.0
      WRITE OUTPUT TAPE 3,108
  108 FORMAT(1H0,7X,9HCONSTANTS12X,8HSUM ON A ,9X,9HSUM ON B
     1,8X,9HSUM ON C  )
  109 WRITE OUTPUT TAPE 3,110,(FA(K),SUMA(K),SUMC(K),SUMB(K)
     1,K=1,6)
  110 FORMAT(1H ,11X,F4.2,10X,E14.8,3X,E14.8,3X,E14.8)
      FB(1)=0.0
      FB(2)=0.2
      FB(3)=0.4
      FB(4)=0.6
      FB(5)=0.8
      FB(6)=1.0
      FC(1)=0.0
      FC(2)=0.2
      FC(3)=0.4
      FC(4)=0.6
      FC(5)=0.8
      FC(6)=1.0
      CALL SORT(SUMA,6,FA)
      A=FA(1)
      CALL SORT(SUMB,6,FB)
```

```
      B=FB(1)
      CALL SORT(SUMC,6,FC)
      C=FC(1)
      WRITE OUTPUT TAPE 3,111,A,B,C
  111 FORMAT(1H0, 9X,26HCONSTANTS FOR MINIMUM SUMS //
     1F16.4,F20.4,F18.4)
  112 CONTINUE
      SENSE LIGHT 1
      SENSE LIGHT 3
      BIG = 0.0
      SMALL = 0.0
  202 WRITE OUTPUT TAPE 3,52
   52 FORMAT(1H0,10X,8HOBSERVED,9X,10HFORECASTED,12X,
     15HERROR/12X,5HVALUE,14X,5HVALUE )
      GO TO 16
C     IF SENSE SWITCH 2 IS ON, THE PLOT ROUTINE WILL BE
C     EXECUTED FOR ANY OF THE OTHER PROGRAM OPTIONS
   53 IF(SENSE SWITCH 2) 54,99
   54 DO 62 I=1,MOTOL
      IF(BIG-SMOTH(I))55,56,56
   55 BIG = SMOTH(I)
   56 IF(SMALL-SMOTH(I))58,58,57
   57 SMALL = SMOTH(I)
   58 IF(BIG-OBE(I))59,60,60
   59 BIG = OBE(I)
   60 IF(SMALL-OBE(I))62,62,61
   61 SMALL = OBE(I)
   62 CONTINUE
      IF(BIG-PRED(I))63,64,64
   63 BIG = PRED(I)
      DO 66 I =NOT,MOTOL
   64 IF(SMALL-PRED(I))66,66,65
   65 SMALL = PRED(I)
   66 CONTINUE
      MOTO = MOTOL + 1
      DO 70 I = MOTO,MOTON
      IF(BIG-PRI(I))67,68,68
   67 BIG = PRI(I)
   68 IF(SMALL-PRI(I))70,70,69
   69 SMALL = PRI(I)
   70 CONTINUE
      WRITE OUTPUT TAPE 3, 71
   71 FORMAT(1H1,40X,27HPLOT OF INPUT AND ANALYSIS ,
     111HINFORMATION  )
      DO 72 I=1,35
B  72 XLABEL(I)=606060606060
      CALL MDFBCD( SEE SUBROUTINE MANUAL FOR INFORMATION  )
      MTPO=MOTON+1
      CALL PLOT1A(3,SMALL,BIG,XLABEL(1),4,1)
      DO 77 I=1,MTPO
      ARRAY(1)=I
```

```
   IF(I-MOT)73,73,74
73 ARRAY(2)=OBE(I)
   ARRAY(4)=OBE(I)
   ARRAY(3)=SMOTH(I)
   ARRAY(5)=SMOTH(I)
   CALL PLOTA(ARRAY)
   GO TO 77
74 IF(I-MOTOL)75,75,76
75 ARRAY(4)=OBE(I)
   ARRAY(5)=SMOTH(I)
   ARRAY(3)=PRED(I)
   ARRAY(2)=PRED(I)
   CALL PLOTA(ARRAY)
   GO TO 77
76 ARRAY(5)=PRI(I)
   ARRAY(3)=PRI(I)
   ARRAY(4)=PRI(I)
   ARRAY(2)=PRI(I)
   CALL PLOTA(ARRAY)
77 CONTINUE
99 GO TO 800
   END
```

SECTION II

BASIC PROGRAM

The basic program, with no sense switches turned on, provides for the selection of the smoothing constants upon the basis of the set that has the smallest error variance over the test series. This basic program also makes the decision as to what part of the data shall be used as the test series. Since this is based upon the programmer's judgment, the program as written will not use less than nine nor more than twenty observations for the test series. The program logic determines if the number of observations left after the largest integral number of years of data is removed is equal to or greater than nine: if so, it uses those; if not, the latest full year of data is added to compose the test series. This minimum may be raised by changing statement number 311.

After execution of the program and the constants are selected upon the basis of the minimum value computed for the error variance, the display on the following page is generated for purposes of making the forecast and providing appropriate supporting evidence for the forecast.

Since this is the standard display, it will be described in detail here and only the additions made by the other options will be described as they are considered.

The first line is a label of the study that is prepared in accordance with the instructions in Section XI. The next group of titles are self-explanatory with the numerical values that appear immediately below them being the constants used for the extrapolation of the series along with the square root of the error variance obtained for the test series using these values. The next line of the display is a protective measure for the forecaster by the fact that it alerts him to the existence of other sets of constants that provide the same error variance. The set used in forecasting will be the most recently computed set. The titles and numerical values that follow are the test series observations, forecasts and errors that were used for the error variance determination. The test values are computed only one period into the future. The next group of titles and numerical values is the forecast of the future activity of the process being studied. The seasonal factor column was added as programmer's information, but it also provides information for the forecaster in that it describes the behavior of the seasonal variations and it indicates whether compensation is taking place in the computation of the supporting algorithms. If the sum of these values is not approximately twelve, then some compensation is taking place within the computations. Two of the more likely conditions are compensation for trend by the seasonal factor, and compensating for overestimates of the random influence upon the mean level of the process. The last line of title and numerical values is a built-in check upon the program and computer. The first values were stored in memory and displayed upon request; this latter set of values are recomputed with the displays and are provided as a check set of values.

If it is desired to make forecasts further into the future, the seasonal estimates may be used over as many times as desired since they represent the latest estimate of this variation in the process. This may be accomplished by changing the "DO 31" loop.

This is the standard display that will be obtained if none of the sense switch options are exercised.

SERIES 1 DATA

RANDOM CONSTANT	SEASONAL CONSTANT	TREND CONSTANT	SQUARE ROOT OF E.M.S.
0.0000	0.2000	0.0000	0.259249E 02

NUMBER OF POINTS WITH EQUAL ERROR 0
FOR RANDOM CONSTANTS = 0 OR 1 THERE ARE 5

OBSERVED VALUE	FORECASTED VALUE	ERROR
51.00000000	82.56811619	-31.56811619
95.00000000	76.80335236	18.19664764
49.00000000	96.86866283	-47.86866283
111.00000000	82.41220284	28.58779716
77.00000000	104.69667244	-27.69667244
93.00000000	95.52171230	-2.52171230
102.00000000	86.07085133	15.92914867
81.00000000	85.37516403	-4.37516403
74.00000000	86.97857189	-12.97857189
124.00000000	87.80468559	36.19531441
102.00000000	102.18780994	-0.18780994
60.00000000	93.22591114	-33.22591114

PERIODS INTO THE FUTURE	FORECASTED VALUE	SEASONAL FACTOR
1	74.34796333	0.91513077
2	78.42724037	0.96740887
3	85.10323334	1.05201089
4	85.91246796	1.06429842
5	96.65735435	1.19998877
6	92.61672211	1.15230846
7	86.99682140	1.08473043
8	82.35617828	1.02909569
9	82.23734379	1.02984491
10	92.62204075	1.16241801
11	99.54193020	1.25199135
12	84.36525154	1.06342836

CHECK ON VALUES USED

0.0000	0.2000	0.0000	25.924923

SECTION III

ENUMERATION OF GRID AND ERROR VARIANCES

The display immediately following is obtained by turning sense swith 1 ON while executing the program. This will list all 156 of the combinations of the smoothing constants and square root of the error mean square obtained by their use over the test series. The standard display discussed in the preceding section is also part of the computer output for this option.

This particular option is in line with some suggestions made in the text as a means of smoothing constant selection. The display of the entire grid provides an indication of the sensitivity of the model to the observed series and at the same time will provide an experienced person with a better understanding of the composition of the series. If it is desired to search the region of minimum error variance with a finer grid, this display is a near necessity in order to establish the region originally and indicate the direction of search which would most likely provide smaller values of the error variance. The latter technique could be based upon the steepest descent methods since the nature of the response is assumed to be unimodal in this region.

If a finer grid is desired in the original search, the numerical values in statements 46, 47 and 48 may be changed to the length of interval desired. It is not necessary to change all the intervals by the same amount; for example, the constant A may be considered more critical for optimum fit of the data and it alone could be changed to obtain a finer search grid along that axis. Care should be taken to not reduce the interval to extremely small values since this would require excessive computer time. If the output is desired on punched cards in order to sort and display the values on the basis of one of the other smoothing constants or by ordered values of the error variance, statement 37 may be changed to write on output tape 5. The tape number may vary depending upon the particular machine installation. This particular procedure of ordering the various columns is recommended for a person desiring to understand more of the relationship of the constants and the response of the series as measured by the error.

Since this program, as written, uses both of the end points, zero and one, some additional programming was used to reduce the amount of computation and length of the display. This is the reason for only six values for A = 0.0 and A = 1.0, since in the former case B may take on any value between zero and one, and in the latter C may take on any value between zero and one. For a full 216 point grid, remove all IF statements between statements 46 and 48 and replace with "GO TO 16."

SERIES 1 DATA

RANDOM CONSTANT	SEASONAL CONSTANT	TREND CONSTANT	SQUARE ROOT OF E.M.S.
0.0000	0.0000	0.0000	26.051385
0.0000	0.2000	0.0000	25.924923
0.0000	0.4000	0.0000	27.016693
0.0000	0.6000	0.0000	29.003139
0.0000	0.8000	0.0000	31.254974
0.0000	1.0000	0.0000	33.392497
0.2000	0.0000	0.0000	28.952703
0.2000	0.0000	0.2000	29.451722
0.2000	0.0000	0.4000	31.225420
0.2000	0.0000	0.6000	33.417658
0.2000	0.0000	0.8000	35.656991
0.2000	0.0000	1.0000	37.560613
0.2000	0.2000	0.0000	28.320306
0.2000	0.2000	0.2000	28.621742
SECTION OF THIS TABLE OMITTED FOR DISPLAY PURPOSES			
0.4000	0.0000	0.0000	31.857025
0.4000	0.0000	0.2000	33.538881
0.4000	0.0000	0.4000	35.413026
0.4000	0.0000	0.6000	36.804882
0.4000	0.0000	0.8000	37.631910
SECTION OF THIS TABLE OMITTED FOR DISPLAY PURPOSES			
0.6000	0.4000	0.0000	52.890210
0.6000	0.4000	0.2000	468.995960
0.6000	0.4000	0.4000	141.760754
0.6000	0.4000	0.6000	65.011344
0.6000	0.4000	0.8000	120.221645
SECTION OF THIS TABLE OMITTED FOR DISPLAY PURPOSES			
0.8000	1.0000	0.0000	55.310760
0.8000	1.0000	0.2000	298.433895
0.8000	1.0000	0.4000	511.670395
0.8000	1.0000	0.6000	174.524853
0.8000	1.0000	0.8000	230.225592
0.8000	1.0000	1.0000	581.910812
1.0000	0.0000	0.0000	234.908638
1.0000	0.0000	0.2000	256.184334
1.0000	0.0000	0.4000	269.630707
1.0000	0.0000	0.6000	289.528286
1.0000	0.0000	0.8000	305.879368
1.0000	0.0000	1.0000	313.895615

THE STANDARD DISPLAY OF SECTION II NORMALLY FOLLOWS

SECTION IV

GRAPHICAL DISPLAY

A plot of the original data along with some of the computed statistical information may be obtained by turning sense switch 2 ON during execution of the program. The legend describing the symbols used in the plot of values appears in the upper left corner of the display. The title and the labels used may be changed by altering the information in statement 71 and the statement immediately following statement 72. To add other sequences of points to the plot, the plot subroutine description should be consulted.

For purposes of generality, the scales of the axes are of the floating point format. The exponents as shown are powers of 10 for left and top justified fractions, respectively; for example, 120 2 = 12. The abscissa scale is determined by a scanning routine in the program and will not truncate or discard any of the values to be plotted as it is now written. The plots of the values are positioned in the appropriate cell and are not on a continuous scale as an analog output would be.

The values that are plotted are generally self-explanatory, but in terms of the text discussions, the observed series is X_t, the smoothed data are $FX_{t,0}$ or in terms of the model $\overline{X}_t\overline{Z}_{t-P}$, the trial values are $FX_{t,1}$ except when the sense switch option 4 is used this becomes $FX_{t,12}$, and the forecasted value is $FX_{t,T}$, $T = 1, 2, \ldots, 12$.

This particular feature of the program provides an aid for the interpretation of numerical data that is often difficult to comprehend from tabular presentations.

PLOT OF INPUT AND ANALYSIS INFORMATION

```
* IS THE FORECASTED VALUE
T IS THE TRIAL VALUE FOR COMPUTING ERROR
O IS THE OBSERVED OR INPUT DATA
S IS THE EXPONENTIALLY SMOOTHED DATA

            0 3 6 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9 9 9 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
            0 3 7 0 3 6 0 3 6 0 3 7 0 3 7 0 3 7 0 3 7 0 3 7 0 4 7 0 4 7 0 0 0 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6
            0 6 2 1 4 8 2 5 9 2 6 0 3 7 0 4 8 1 5 8 2 6 9 3 6 0 4 7 1 4 1 4 8 1 4 8 1 4 8 1 4 8 1 4 8 1 5 8 1 5 8

            0 1 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3

      100 1                                                       S      O
      200 1                                                       S         O
      300 1                                                               OS
      400 1                                                         OS
      500 1                                                                   S       O
      600 1                                                                  S O
      700 1                                                S      S
      800 1                                                          S
      900 1                                                             S   O
      100 2                                         O        S
      110 2                                                                  S  O
      120 2                                                          O   S
      130 2                                                  O   S
      140 2                                            O       S
      150 2                                                                S  O
      160 2                                                            S          O
      170 2                                                    O          S
      180 2                                                          O    S
      190 2                                                        S       O
      200 2                                                         S O
      210 2                                                              S     O
      220 2                              O              S
      230 2                                             O               S
      240 2                                                               S     O
      250 2                                                  O  S
      260 2                                                   SO
      270 2                                                               S O
      280 2                                                           S  O
      290 2                                                                S      O
```

```
300 2                                                                   S                              O
310 2                                                      S    O
320 2                                                         S         O
330 2                                                                 S                        O
340 2                                                S                    O
350 2                                                             S          O
360 2                                                             S      O
370 2                                                   S    O
380 2                                               OS
390 2                                                             S    O
400 2                                                           S            O
410 2                                             O           S
420 2                                                                   S  O
430 2                                           O      S
440 2                                                         S      O
450 2                                                        O     S
460 2                     O                    S
470 2                                                             S       O
480 2                                                             S      O
490 2                                                      S                   O
500 2                                                   S                  O
510 2                                                             S      O
520 2                                                           S    O
530 2                                                         S  O
540 2                                                    O          S
550 2                                                  SO
560 2                                                     O S
570 2                                                            OS
580 2                                            S       O
590 2                                                         O S
600 2                                                               S         O
610 2                                       O          S
620 2                                                  S  O
630 2                                                         O S
640 2                                                           S  O
650 2                                                  O   S
660 2                                                  O        S
670 2                                                      S                    O
680 2                                                       S     O
690 2                                                  O      S
700 2                                                S                       O
710 2                                                          S  O
720 2                                                              S  O
730 2                                                   S         O
740 2                                                    S           O
```

```
750 2                                                               S             O
760 2                                                          OS
770 2                                                                S                                  O
780 2                                                             S O
790 2                                                         S         O
800 2                                                  O     S
810 2                                                    O     S
820 2                                                      S               O
830 2                                                               S             O
840 2                                                               SO
850 2                                                  O S
860 2                            O                    S
870 2                                                      O      S
880 2                           O                        S
890 2                                                                  S         O
900 2                                                            S  O
910 2                                              O        S
920 2                                                         OS
930 2                                                  O    S
940 2                                                       S           O
950 2                                                               S         O
960 2                                               O           S
970 2                                    O              S   T
980 2                                                    T S        O
990 2                                   O                      S     T
100 3                                                       T  S             O
101 3                                                    O            S  T
102 3                                                             O S
103 3                                                         T S       O
104 3                                                  O ST
105 3                                              O     S T
106 3                                                          T    S                O
107 3                                                                   S
108 3                                          O               S  T
109 3                                              *
110 3                                                 *
111 3                                                     *
112 3                                                     *
113 3                                                            *
114 3                                                         *
115 3                                                      *
116 3                                                   *
117 3                                                   *
118 3                                                         *
119 3                                                             *
120 3                                                    *
```

SECTION V

SELECTION OF CONSTANTS BY SUMMATION

The sense switch 3 option provides for the selection of the constants based upon the minimum sum of error variances for each constant over all values of the other constants. Each of the sums exhibited in the display is composed of 36 values and each column sums to the total for all 216 error variance values. The theory or basis for this method of smoothing constant selection is given in Chapter III.

The particular display associated with this option is a composite of the standard display and the special features of the summations. The first numerical values are the constants and associated error based upon the minimum value of the error variance per the standard display. The second group of numerical values are the respective sums for each value of the specified constant, summed over all other values of the other constants that appear in combination with the specified constant. From these sums, the set of constants which have the minimum sum are selected and displayed below the tableau of sums. These factors are then used to extrapolate the observed series in accordance with the computational forms and the model presented in Chapter II. This particular model uses the check at the end of the display as the only display of the error variance associated with the particular set of smoothing constants. It could be computed from the test series also. This provides comparative values of the error variance for the selection of the constants upon the basis of independent sums or a strictly minimum value.

If a finer grid is used, corresponding changes must be made in the 100 series of statement numbers.

SERIES 1 DATA

RANDOM CONSTANT	SEASONAL CONSTANT	TREND CONSTANT	SQUARE ROOT OF E.M.S.
0.0000	0.2000	0.0000	0.259249E 02

NUMBER OF POINTS WITH EQUAL ERROR 0
FOR RANDOM CONSTANTS = 0 OR 1 THERE ARE 5

CONSTANTS	SUM ON A	SUM ON B	SUM ON C
0.00	0.10358616E 04	0.50681942E 04	0.14613655E 04
0.20	0.20867509E 05	0.56946499E 04	0.14617125E 04
0.40	0.70068999E 04	0.58260104E 04	0.14884080E 04
0.60	0.16331500E 04	0.59647417E 04	0.38282454E 04
0.80	0.18327073E 04	0.58131324E 04	0.28935964E 04
1.00	0.21647408E 04	0.61741400E 04	0.23407540E 05

CONSTANTS FOR MINIMUM SUMS

0.0000	0.0000	0.0000

OBSERVED VALUE	FORECASTED VALUE	ERROR
51.00000000	82.56811619	-31.56811619
95.00000000	76.80335236	18.19664764
49.00000000	96.86866283	-47.86866283

THIS PART OF THE STANDARD DISPLAY OMITTED FOR BREVITY

10	92.62204075	1.16241801
11	99.54193020	1.25199135
12	84.36525154	1.06342836

CHECK ON VALUES USED

0.0000	0.0000	0.0000	26.051385

SECTION VI

SELECTION OF CONSTANTS FOR CONSECUTIVE LAGS

The program will select the set of smoothing constants which have a minimum sum of error variances for lag values of one through 12 if sense switch 4 is ON during the program execution. The computation is basically the same as that of the standard program of Section II except it is repeated for the additional values of the lag between the time of the forecast and the point in time for which the forecast is made. For each value of the lag, there is a change in the point within the process from which the forecast is made in order to provide the same number of observations in the error variance. This is accomplished by making the forecasts at an earlier point in the smoothing of the series each time. The forecasts in each case are over the same test series as for a lag of one. The set of smoothing constants that provide the minimum sum of the standard deviations of error is used for the actual forecasting.

The basic reasons for developing this particular form of smoothing constant selection were to locate the optimum region of error variance and to provide more confidence in the model. The use of this procedure without further investigation for local minimums is justified upon the basis of those empirical results that have been studied and the highly improbable case of smoothing constant compensation for all values of the lag. The forecaster's confidence in the model should be increased if for no other reason than the increase in the number of tests that are run on each set of smoothing constants. If the autocorrelation coefficients for the series and this display are available, it is to the forecaster's benefit to justify the smoothing constant selection for this method.

The display of the grid values and the associated standard deviation of the error for the lag values and its total is provided by this option if sense switch 1 is ON at the time of the computation. A sample of the display of this type is shown following this discussion. If sense switch 1 is OFF then the computation is performed, the selection is made as discussed above, and the display is limited to the standard form as given in Section II. These methods presented in this section are for the case of a single set of smoothing constants to be used for forecasting after selection of smoothing constants based upon consideration of error variances for consecutive lag values.

If the more distant forecasts are not as important a consideration as the immediate values of the forecast, a method of weighting the error variances may be used to automatically consider this in the smoothing constant selection. This is possible with only a minor modification of the program. The modification of SISUM located between statement numbers 503 and 504 will provide the type of weighting desired provided that it can be expressed as some function of the lag value. For example, if the weights assigned to the forecast are to be in the form of a simple harmonic progression, it could be arranged by dividing SIGE by T, where T is the lag value for the particular standard deviation of the error that is being added to the sum for a given set of smoothing constants.

An additional feature of this option is the computation and display of the standard deviation of the error for each lag value within each set of smoothing constants. Even though the statistical soundness precludes the direct use of these values as limits on the forecast, they provide an estimate of that which may be expected in terms of forecast error and can be used to judge the value of this method of forecasting against other methods that may be available. Repeated use of this procedure on a given process for a period of time would probably provide some degree of assurance to the forecaster as the magnitude of the error that may be expected relative to the values of the standard deviation of the error that are presented in the display. This would provide for establishing control limits on the forecasting process to be used in the sense of a quality control chart.

SERIES 1 DATA

(A SAMPLE OF THE DISPLAY FROM THE SENSE SWITCH 4 OPTION)

RANDOM CONSTANT	SEASONAL CONSTANT	TREND CONSTANT	SQUARE ROOT OF E.M.S.	LAG
0.2000	0.0000	0.0000	28.952703	1.0
0.2000	0.0000	0.0000	27.175207	2.0
0.2000	0.0000	0.0000	26.581068	3.0
0.2000	0.0000	0.0000	27.391627	4.0
0.2000	0.0000	0.0000	27.316981	5.0
0.2000	0.0000	0.0000	25.753752	6.0
0.2000	0.0000	0.0000	27.286433	7.0
0.2000	0.0000	0.0000	25.205850	8.0
0.2000	0.0000	0.0000	25.562231	9.0
0.2000	0.0000	0.0000	27.056058	10.0
0.2000	0.0000	0.0000	24.111557	11.0
0.2000	0.0000	0.0000	27.701136	12.0
0.2000	0.0000	0.0000	320.094593	TOTAL
0.2000	0.0000	0.2000	29.451722	1.0
0.2000	0.0000	0.2000	27.190519	2.0
0.2000	0.0000	0.2000	26.003510	3.0
0.2000	0.0000	0.2000	27.638329	4.0
0.2000	0.0000	0.2000	27.843888	5.0
0.2000	0.0000	0.2000	25.727711	6.0
0.2000	0.0000	0.2000	30.096730	7.0
0.2000	0.0000	0.2000	28.276462	8.0
0.2000	0.0000	0.2000	34.180559	9.0
0.2000	0.0000	0.2000	40.202959	10.0
0.2000	0.0000	0.2000	39.081363	11.0
0.2000	0.0000	0.2000	47.930486	12.0
0.2000	0.0000	0.2000	383.624226	TOTAL
0.2000	0.0000	0.4000	31.225420	1.0
0.2000	0.0000	0.4000	29.597482	2.0
0.2000	0.0000	0.4000	29.481549	3.0
0.2000	0.0000	0.4000	32.422558	4.0
0.2000	0.0000	0.4000	30.865085	6.0
0.2000	0.0000	0.4000	36.538964	7.0
0.2000	0.0000	0.4000	32.857842	8.0
0.2000	0.0000	0.4000	44.160725	9.0
0.2000	0.0000	0.4000	54.561029	10.0
0.2000	0.0000	0.4000	57.119162	11.0
0.2000	0.0000	0.4000	71.090055	12.0
0.2000	0.0000	0.4000	483.151886	TOTAL

THE STANDARD DISPLAY OF SECTION II NORMALLY FOLLOWS

SECTION VII

SELECTION OF THE STARTING POINT

The method of finite differences as explained in Chapter III for the selection of the starting point for application of the supporting algorithms may be included in the analysis by use of sense switch 5. The program provides some additional optimization aids over those presented in the discussion in the body of the thesis while providing the same basic technique. The interval of historical data that is used as the minimum length of data to be used in the computation is passed over the absolute values of the third order differences, but these differences are summed in groups that correspond to year intervals of the original series. This still provides for the location of the best linear fit of the data by including 80% of the second order differences in the third order differences, but the additional benefit to be gained is that any systematic changes which appear in the data will be taken into consideration. For example, if a particular process had its major activity during one month and then the activity tapered off until the same month the following year and the cycle is repeated, the conventional use of finite differences would indicate that this is a major oscillation in the series and a relatively high degree polynomial would be required to fit the observed data. However, through the proper indexing, the sum is taken in terms of those third order differences determined by each year alone and not the entire sequence. This, then, provides the method whereby the best starting point is selected for the periodic behavior of the series. The importance of this can only be demonstrated by taking the third order differences of a sawtooth function and observing the difference in the absolute sum of the third order differences for each year in the series as opposed to the absolute sum of the third order differences for the whole series. This technique is considered to be relatively important for those series that experience a particularly sharp change at a given interval within each period.

In the program listed in Section I of the Appendix, the interval for the historical data is taken as one year less than the largest integral number of years of data available, therefore, it necessarily limits the selection of the starting point to the first year of data. However, in a number of the trials made during the course of this investigation, this was sufficient to change the error variance significantly.

The additional output, as a result of this addition to the standard display, is a listing of the months within the first year as starting points and the sum as described earlier of the third order differences, along with the number of data points that have been deleted from further consideration as part of the series.

To change the length of the minimum interval for this part of the program, all that is necessary is change the right side of statement number 300, make a corresponding change in the second term in the definition of MTA4 and change the upper limit of the "DO 308" loop to consider the number of data points that are in the historical data, but not in the minimum interval specified.

This addition to the algorithms used in support of the time series model is felt to be an original contribution arising from this study.

SERIES 1 DATA

FIRST DATA POINT USED	SUM OF 3RD DIF
1	4760.00000
2	4893.00000
3	4947.00000
4	4672.00000
5	4501.00000
6	4294.00000
7	4616.00000
8	4979.00000
9	4825.00000
10	4657.00000
11	4579.00000
12	4944.00000

THOSE DATA POINTS PRECEDING THE 6TH UNIT HAVE BEEN DELETED

THE STANDARD DISPLAY OF SECTION II NORMALLY FOLLOWS

SECTION VIII

EXTERNAL CONTROL OF SMOOTHING CONSTANTS

The program user may exercise external control over the selection of the constants to be used in the model by use of sense switch 6 and the addition of a card containing the constants to the data deck. The addition of this card and its format will be discussed in Section XI of the Appendix. The primary purpose of this addition to the program is to provide for the use of a finer search grid in a local area. It also provides for additional displays of various types once the smoothing constants are arrived at through the use of one of the other options of the program. For example, the plot may be rather lengthy to obtain as a routine output for each of the combinations of analyses that are used in preliminary investigations. After the set of smoothing constants is chosen, the program may be executed in a small fraction of the time required to carry out the iterations and provide the output. By making a minor program change, this option may be used to start the search grid at some arbitrary point with the programmed increment used over the remainder of the grid. This alternative is available by removing the three cards in the source program between statements 201 and 213.

In general this option should not be used for forecasting in the absence of the use of the error analysis information that is provided by the other options. If the person doing the forecasting uses the external control option exclusively for forecasting, he is defeating one of the prime purposes in the study that has been conducted in improving the search and selection methods associated with this model. The forecast of the model becomes a prediction under these conditions of external control.

However, if coefficients have been established through the use of the error method of selection in past periods and the process appears to be stable, the use of the same constants for updating the estimates would be a reasonable approach for the more stable time series. The primary intent of this particular option is as an exploratory convenience for the forecaster.

The display for this option is identical to the display of Section II and is not repeated in this section.

SECTION IX

INDIVIDUAL SETS OF CONSTANTS FOR EACH LAG VALUE

Since the IBM 7090 has only six sense switches, it is necessary to use a control card and a sense light for this particular programming option. The provisions of this program are essentially the same as those explained for the sense switch option for selection of constants based upon consecutive lags. This particular option was explained in the text of the thesis and it is used under the assumption that each data point in the cyclical period possesses certain relationships to those points that precede and follow it and this relationship is often a direct function of the lag; that is, the relationships between a data point and two other points that are different becomes a certain function to be considered in the forecasting of events. Thus, the recognition of a relationship between the autocorrelation and lag within the series will contribute to a reduction in the error associated with the forecast error that is observed in the series. In order to provide for this, the option provides for the selection of the minimum error for each of the consecutive lags that are applied to the test series. Therefore, the autocorrelation or the relationship for the smoothing or the forecasting of events that extends one period into the future have different forms than those that extend some other number of periods into the future. An example of this is the use of a polynomial fit of the data. The fit for the prediction of one period into the future could be a much higher order polynomial than one that is to predict a greater number of periods into the future. This may be related more directly to this model since the seasonal factor empirically determines the degree of fit that is established in the model for the series. The amount of smoothing that takes place in the seasonal data could well have an effect upon the accuracy of the forecast of the future for various lags in the data. Consideration of this, at least over the test series, reduces the sum of the error for the consecutive lags. This is illustrated by a comparison of Tables XIV and XV for test series 6.

The display for this option provides the minimum standard deviation of the error and its associated smoothing constants for each lag value. The forecasts and seasonal factors are computed individually for each lag value and displayed as shown on the following page.

RANDOM CONSTANT	SEASONAL CONSTANT	TREND CONSTANT	SQUARE ROOT OF E.M.S.	LAG
0.8000	1.0000	0.0000	44.127215	1
0.8000	1.0000	0.0000	50.507354	2
0.8000	0.8000	0.0000	59.934388	3
0.8000	0.8000	0.0000	55.577856	4
0.8000	0.8000	0.0000	41.282528	5
0.8000	0.8000	0.0000	41.980234	6
0.8000	0.8000	0.0000	51.553693	7
0.8000	0.8000	0.0000	50.768961	8
0.8000	0.8000	0.0000	59.581101	9
0.8000	1.0000	0.0000	55.555433	10
0.8000	0.8000	0.0000	39.487572	11
0.8000	1.0000	0.0000	29.471549	12

PERIODS INTO THE FUTURE	FORECASTED VALUE	SEASONAL FACTOR
1	60.7010298	0.890472
2	65.2863646	0.956042
3	56.3915958	1.013998
4	57.0791225	0.985599
5	55.9890518	1.241133
6	59.8567791	1.069360
7	72.9479608	1.112799
8	92.9617882	0.990582
9	95.5730228	0.954830
10	48.6098137	1.302878
11	54.9222951	1.285091
12	45.3881621	1.046367

SECTION X

TABLE OF OPTIONS

The foregoing discussions of the available options were presented as though each one was an entity in itself. This is true, but certain combinations may be used to the advantage of the forecaster and the multiple provisions that the program is capable of providing. Table XVII presented below summarizes the available combinations.

TABLE XVII

TABLE OF PROGRAM OPTIONS

Standard Display	0	-	*	*	*	*	*	*	*
Error Variance Grid	1	*	-	*	*	*	*		*
Plot Routine	2	*	*	-	*	*	*	*	*
Independent Sums	3	*	*	*	-		*		
Consecutive Lags	4	*	*	*		-	*		
Starting Point	5	*	*	*	*	*	-	*	*
External Control	6	*		*			*	-	
Correlated Lags	SL4	*	*	*			*		-(4)
Sense Switches	↳→	0	1	2	3	4	5	6	SL4

NOTE: (4) Sense switch 4 must be ON.

All displays common to the points of intersection in the table for the options exercised will be provided in the computer output.

In order to use the sense light 4 option, sense switch 4 must be ON and the last card in data deck must have a nine in column 1.

As an example, if switches 2, 3, and 5 are on, the program will select the optimum starting point, select the constants on the basis of minimum sums, provide the display of Section V, and plot the observed series and its analysis according to Section IV.

SECTION XI

PREPARATION OF INPUT DATA

The arrangement of the data deck is shown below. This is to be accompanied by the compiled program listed in Section I of this Appendix. The illustration is self-descriptive, but will be supplemented by verbal descriptions. The first card of the input is any title the user desires provided that it is only 51 characters long and the spacing of the output will correspond to the spacing on the input card. The card may be blank if desired, but must be included, otherise the first data card will appear as a title and will not be included in the analysis. The data follows the title with one observation per card. This may be changed easily by changing statement 3, but was programmed in this manner to facilitate changes in the historical data file. The present program has a limit of 600 data points: this may be increased by changing the DIMENSION statement. The next card is an end of data card: this notifies the computer that all the data has been read in. The format is the same as the data cards with nines in each numeric column. The next card is optional, depending upon whether sense switch 6 is to be used during the execution of the program. If it is not to be turned ON during the execution, this card should be omitted from the data deck. If sense switch 6 is to be ON during execution, then this card should contain the smoothing constants in a 3F10.6 format. The last card must be on the deck with a nine in the first column if sense light 4 option is to be used and blank if not. This completes the description of the data deck for the program input.

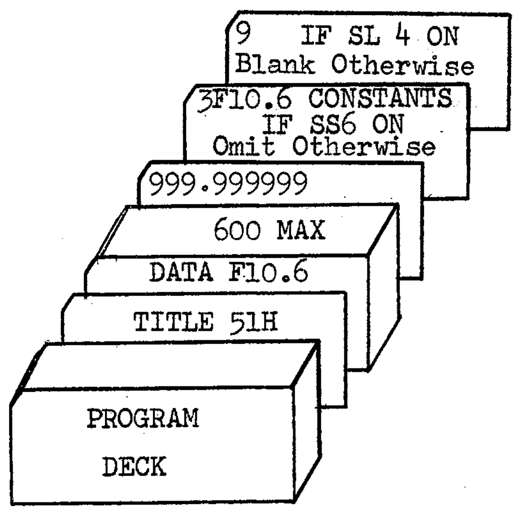

Figure 11. Composition of Program and Data Deck

SECTION XII

IBM 1620 PROGRAMS

Due to the memory requirements of this program, it is necessary to divide the program discussed in the preceding sections into five parts in order to fall within the limitations of the IBM 1620. For the convenience of the discussions the displays of these programs have been labeled in the comment cards placed at the beginning of each program. They are in the normal sequence of their execution in the analysis of a time series. This also includes the cases of deletion of some of the options in the data analysis. The programs have been reduced considerably in the amount of historical data that they are capable of handling. Therefore, rather than create confusion, the capacity of the memory is reflected consistently in all the individual programs and the maximum that can be accommodated by the programs is five years of historical data. Certain other deletions were made in the output displays primarily due to the limitation of memory and due to the 1620 installations being less formal in the amount of personal contact between the computer user and the computer operator than the IBM 7090 installations. This reduces the need for explicit means of identification of the output and strict rules of computer input.

Program 1 is for the selection of the optimum starting point within the first year of historical data. The results are identical with those of the 7090 output and it is self-explanatory when displayed by the typewriter output of the 1620. This program makes the recommendation that a certain number of the data points be eliminated from the data deck before further consideration, since the program cannot exercise the control over the situation and must rely on the computer operator to remove those cards from the data deck.

The input for the program is the historical data in the F10.6 format with the 999.999999 end-of-data card.

The output of the program is a typed display of the sums of the third order differences for the data as described in Section VII.

The use of a point within the first year of the data as the starting point is particularly applicable in the use of the 1620 programs since the limit on the historical data is five years and the deletion of more than one year of these data could significantly affect the estimates of the process level and subsequently the forecasts. Any number of data decks may be processed sequentially by pushing START on the 1620 console after the completion of each display.

The approximate time for reading the data, computation, and displaying the results is three minutes.

Program 2 is the basic program for the computation of the error variances for a given grid of constants and a lag value of one. The grid used in this program is arbitrary and is controlled external to the machine by use of punched card input in the 3F10.6 format for the sets of smoothing constants to be used. Once a primary or course search grid is established for the consideration of the processes, a complete deck of these grid values in some ordered sequence should be prepared to facilitate the use of the programs. These values are computed one at a time and the results may be obtained in three ways. The smoothing constants and the error variance value can be on punched cards, they may be typed out, both of these methods may be used simultaneously, or they may be computed and the minimum error variance and its constants stored in memory to be displayed after the completion of the search grid by turning sense switch 2 ON during the last computation. However, if sense switch 3 is ON, the manual start must be used after each set of constants. This is to encourage use of punched card output and to save machine time. The program selects the minimum value of the error variance up to the point that it is requested to display that information, and continues to re-evaluate the minimum for all values as long as a given set of series data is stored in memory. Therefore, a number of investigations may be performed by using the smaller machine and having access to the display of the computations as they are made. This is particularly true when a fine grid search is being made of a local minimum. The computation time for an iteration is approximately 24 seconds or one hour for the 156 value grid used in these studies.

Program 3 is the modification of Program 2 in order to consider the consecutive lag values and their error variance. The display is limited to card output due to the length of the answers and the length of the program, except during the final computation of the series sense switch 2 may be turned on to obtain both a typed and punched display of the set of constants and the value of the minimum sum of the error variance. Once this option has been exercised the program transfers control back to the beginning of the program and reads in the next series. Sense switch 2 provides for the display of the minimum sum and the initiation of the computations for a new series, but once this transfer has been made the switch should be turned to OFF otherwise the program will transfer after the first computation and will necessitate reloading the data again to restart the computation. This program requires about 45 seconds per iteration or a total of two hours for the 156 value grid.

Program 4 provides for the summing of the error variances for each of the smoothing constants over all values of the other constants. The input to the program is the output of Program 1. The output of this program is typed out by the console typewriter and is of the form of the display for Section V. Due to the small number of comparisons involved in determining the minimum value for each of the constants it is not automatically computed and displayed. Time is not of importance at this point, since it is necessary to load Program 5 before the optimum value determined by the minimum sums may be used in forecasting. The time required for this program to read, compute, and display the results is approximately 3 minutes per deck of input data and any number of data decks may be considered by pushing start at the end of each previous computation.

Program 5 provides a display of the forecasted values of the process and the seasonal factors. The input to this program is the standard data deck and the smoothing constants to be used for computations. The options permit the use of more than one set of constants for each data deck if sense switch 1 is ON. If sense switch 2 is ON, the program will consider more than one data deck in sequence with one set of smoothing constants for each deck. These switches may be manipulated to suit the immediate needs of the computation. The display is by typewriter and is of the same form as the 7090 displays. The computation time required is less than 2 minutes per series iteration.

The displays of these programs are not presented in this section since they are identical in nature to those of the 7090 output presented in the preceding sections and it would be redundant to include them. The only difference is the display of the error variance in the 1620 displays as opposed to the display of its square root in the 7090 displays.

```
C     PROGRAM 1 - SELECTION OF STARTING POINT
C     INPUT- DATA DECK WITH 999.999999 AS THE LAST DATA CARD
C     OUTPUT - TYPEWRITER
C     OPTIONS - MORE THAN ONE DATA DECK, PUSH START
      DIMENSION OBE(60),DEL(60)
    1 MOTOL = 0
      I=1
    2 READ3,OBE(I)
    3 FORMAT (F10.6)
      IF(OBE(I) -999.999999) 4,5, 99
    4 MOTOL = MOTOL +1
      I = I + 1
      GO TO 2
    5 MYR = MOTOL/12
      MTA1=12*(MYR-1)
      MTA2=MTA1-1
      MTA3=MTA2-1
      COMP=0.9E+25
      DO 301 I=2,MTA1
  301 DEL(I-1)=OBE(I)-OBE(I-1)
      DO 302 I=2,MTA2
  302 DEL(I-1)=DEL(I)-DEL(I-1)
      DO 303 I=2,MTA3
  303 DEL(I-1)=DEL(I)-DEL(I-1)
      PRINT 304
  304 FORMAT(10X,23HFIRST DATA    SUM  OF  )
      PRINT 3042
 3042 FORMAT(10X,21HPOINT USED   3RD DIF  )
      DO 309 IN=1,12
      SUMDL=0.0
      MTA4=IN+12*(MYR-2)-4
      DO 306 I=IN,MTA4,12
      DO 306 J=1,9
      LL=(I+J-1)
      ADEL=DEL(LL)
      IF(ADEL)305,306,306
  305 ADEL = (-1.0)*ADEL
  306 SUMDL = SUMDL + ADEL
      PRINT 307,IN,SUMDL
  307 FORMAT(I15,F16.5)
      IF(COMP - SUMDL)309,309,308
  308 COMP = SUMDL
      INIT = IN
  309 CONTINUE
      PRINT 310,INIT
  310 FORMAT(10X,31HTHOSE DATA POINTS PRECEDING THE I4,2HTH)
      PRINT 311
  311 FORMAT(10X,23HVALUE SHOULD BE DELETED  )
      PAUSE
      GO TO 1
   99 END
```

```
C      PROGRAM 2 - PROGRAM FOR COMPUTATION OF ERROR
C      VARIANCES FOR LAG OF 1
C      INPUT - DATA DECK AND DECK OF SMOOTHING CONSTANTS
C      OUTPUT - CARDS AND/OR TYPEWRITER
C      OPTIONS
C      SENSE SWITCH 1 FOR PUNCHED CARD OUTPUT
C      SENSE SWITCH 2 TO OBTAIN MINIMUM VALUES OF THE ERROR
C      VARIANCE AND ITS ASSOCIATED CONSTANTS AT ANY TIME
C      DURING THE COMPUTATION
C      SENSE SWITCH 3  FOR TYPEWRITER OUTPUT, PUSH START
       DIMENSION OBE(60),OBS(12,5),YRAV(5),TREND(60),UNAD(12)
       DIMENSION SEA(60),EST(60),PRED(60),SENA(12)
       DIMENSION SEANL(12,5)
     1 MOTOL = 0
       I=1
     2 READ3,OBE(I)
     3 FORMAT (F10.6)
       IF(OBE(I) -999.999999) 4,5, 99
     4 MOTOL = MOTOL +1
       I = I + 1
       GO TO 2
     5 MYR = MOTOL/12
       IF(MOTOL - 12* MYR -9) 6,7,7
     6 MOT = 12*(MYR-1)
       MYR = MYR - 1
       GO TO 8
     7 MOT = 12* MYR
     8 FMYR = MYR
       L = 1
       DO 10 J = 1, MYR
       YRAV(J)=0.0
       DO 9 I = 1, 12
       UNAD(I)=0.0
       OBS (I,J) = OBE (L)
       YRAV(J) = YRAV(J) + OBE (L)/12.0
     9 L = L + 1
    10 CONTINUE
       DENOM = MOTOL - MOT
       NOT = MOT +1
       SMVAR = (0.9E+25)*(0.9E+25)
       AMOT = MOT
       TREN      = (YRAV(MYR)-YRAV(1))/(AMOT-12.0)
       DO 12 J = 1, MYR
       SUSEA = 0.0
       DO 12 I = 1, 12
       FI = I
       SEANL(I,J) = OBS(I,J)/(YRAV(J)-(6.5-FI)*TREN)
       UNAD(I) = UNAD(I) + (SEANL(I,J)/FMYR)
       IF(UNAD(I))11,12,12
    11 UNAD(I) = (-1.0)*UNAD(I)
    12 SUSEA = SUSEA + UNAD(I)
```

```
      PLY = 12.0 /SUSEA
      DO 13 I = 1,12
13    SENA(I) = UNAD(I) * PLY
14    READ 15,A,B,C
15    FORMAT(F10.6,F10.6,F10.6)
      EST(1)=A*(OBE(1)/SENA(1))+(1.0-A)*(YRAV(1)+TREN)
      TREND (1)=C*(EST(1)-YRAV(1))+(1.0-C)*(TREN     )
      SEA(1) = B*(OBE(1)/EST(1))+(1.0-B)*SENA(1)
      DO 19 I = 2, MOTOL
      IF(I-12)16,16,17
16    J=I
      SEA(J) = SENA(J)
      GO TO 18
17    J=I-12
18    EST(I)=A*(OBE(I)/SEA(J))+(1.0-A)*(EST(I-1)+TREND(I-1))
      SEA (I) = B*(OBE(I)/EST(I))+(1.0-B)*SEA(J)
19    TREND (I) = C*(EST(I)-EST(I-1))+(1.0-C)*TREND(I-1)
      SUMER = 0.0
      DO 20 K = NOT, MOTOL
      PRED(K) = (EST(K-1) + TREND (K-1))*SEA(K-12)
      ERROR = (OBE(K) - PRED(K))
      ERSQ = ERROR * ERROR
20    SUMER = SUMER + ERSQ
      VAR = SUMER/DENOM
      IF(SENSE SWITCH 1) 21,22
21    PUNCH 26,A,B,C,VAR
22    IF(VAR-SMVAR) 23,24,24
23    SMVAR = VAR
      SA = A
      SB = B
      SC = C
24    IF(SENSE SWITCH 3) 25,14
25    PRINT 26, A,B,C, VAR
26    FORMAT(F7.3,F7.3,F7.3,F20.6)
      PAUSE
      IF(SENSE SWITCH 2) 27,14
27    PRINT 26,SA,SB,SC,SMVAR
      PUNCH 26,SA,SB,SC,SMVAR
      GO TO 1
99    END
```

```
C      PROGRAM 3 - PROGRAM FOR COMPUTATION OF ERROR
C      VARIANCES FOR LAG VALUES 1 THROUGH 12
C      INPUT - DATA DECK AND DECK OF SMOOTHING CONSTANTS
C      OUTPUT - CARDS
C      OPTIONS AT END OF PROGRAM TURN SWITCH 2 ON DURING LAST
C      COMPUTATION TO OBTAIN MINIMUM ERROR VARIANCE AND ITS
C      ASSOCIATED SMOOTHING CONSTANTS
       DIMENSION OBE(60),OBS(12,5),YRAV(5),TREND(60),UNAD(12)
       DIMENSION SEA(60),EST(60),PRED(60),SENA(12),VAR(12)
       DIMENSION SEANL(12,5)
     1 MOTOL = 0
       I=1
     2 READ3,OBE(I)
     3 FORMAT (F10.6)
       IF(OBE(I) -999.999999) 4,5, 99
     4 MOTOL = MOTOL +1
       I = I + 1
       GO TO 2
     5 MYR = MOTOL/12
       IF(MOTOL - 12* MYR -9) 6,7,7
     6 MOT = 12*(MYR-1)
       MYR = MYR - 1
       GO TO 8
     7 MOT = 12* MYR
     8 FMYR = MYR
       L = 1
       DO 10 J = 1, MYR
       YRAV(J)=0.0
       DO 9 I = 1, 12
       UNAD(I)=0.0
       OBS (I,J) = OBE (L)
       YRAV(J) = YRAV(J) + OBE (L)/12.0
     9 L = L + 1
    10 CONTINUE
       DENOM = MOTOL - MOT
       NOT = MOT +1
       SMVAR = (0.9E+25)*(0.9E+25)
       AMOT = MOT
       TREN      = (YRAV(MYR)-YRAV(1))/(AMOT-12.0)
       DO 12 J = 1, MYR
       SUSEA = 0.0
       DO 12 I = 1, 12
       FI = I
       SEANL(I,J) = OBS(I,J)/(YRAV(J)-(6.5-FI)*TREN)
       UNAD(I) = UNAD(I) + (SEANL(I,J)/FMYR)
       IF(UNAD(I))11,12,12
    11 UNAD(I) = (-1.0)*UNAD(I)
    12 SUSEA = SUSEA + UNAD(I)
       PLY = 12.0 /SUSEA
       DO 13 I = 1,12
    13 SENA(I) = UNAD(I) * PLY
    14 READ 15,A,B,C
```

```
15 FORMAT(F10.6,F10.6,F10.6)
   EST(1)=A*(OBE(1)/SENA(1))+(1.0-A)*(YRAV(1)+TREN)
   TREND (1)=C*(EST(1)-YRAV(1))+(1.0-C)*(TREN     )
   SEA(1) = B*(OBE(1)/EST(1))+(1.0-B)*SENA(1)
   DO 19 I = 2, MOTOL
   IF(I-12)16,16,17
16 J=I
   SEA(J) = SENA(J)
   GO TO 18
17 J=I-12
18 EST(I)=A*(OBE(I)/SEA(J))+(1.0-A)*(EST(I-1)+TREND(I-1))
   SEA (I) = B*(OBE(I)/EST(I))+(1.0-B)*SEA(J)
19 TREND (I) = C*(EST(I)-EST(I-1))+(1.0-C)*TREND(I-1)
   SUMVA = 0.0
   DO 21 LT = 1,12
   SUMER = 0.0
   LOW = (NOT - LT + 1)
   MAX = (MOTOL - LT + 1)
   DO 20 K = LOW,MAX
   N = (K + LT - 1)
   T = LT
   PRED(N) = (EST(K-1) + T*TREND(K-1))*SEA(N-12)
   ERROR = (OBE(N) - PRED(N))
   ERSQ = ERROR * ERROR
20 SUMER = SUMER + ERSQ
   VAR(LT) = SUMER/DENOM
21 SUMVA = SUMVA + VAR(LT)
   PUNCH 26, A,B,C,SUMVA
   PUNCH 22,VAR(1),VAR(2),VAR(3),VAR(4),VAR(5),VAR(6)
   PUNCH 22,VAR(7),VAR(8),VAR(9),VAR(10),VAR(11),VAR(12)
22 FORMAT(F12.4,F12.4,F12.4,F12.4,F12.4,F12.4)
   IF(SUMVA - SMVAR) 23,24,24
23 SMVAR = SUMVA
   SA = A
   SB = B
   SC = C
24 IF(SENSE SWITCH 2) 27,14
26 FORMAT(F7.3,F7.3,F7.3,F20.6)
27 PRINT 26,SA,SB,SC,SMVAR
   PUNCH 26,SA,SB,SC,SMVAR
   GO TO 1
99 END
```

```
C     PROGRAM 4 - PROGRAM FOR COMPUTATION OF SUM OF ERROR
C     VARIANCES FOR EACH VALUE OF THE SMOOTHING CONSTANTS
C     INPUT - OUTPUT DATA FROM PROGRAM 2
C     OUTPUT - TYPEWRITER
C     OPTIONS - MORE THAN ONE DATA DECK, PUSH START
      DIMENSION S(156),SUMA(6),SUMB(6),SUMC(6),FA(6)
    1 DO 100 I = 1,156
  100 READ 101, S(I)
  101 FORMAT(21X,F20.6)
      KA = 2
      SUMA(1)=6.0*(S(1)+S(2)+S(3)+S(4)+S(5)+S(6))
      DO 103 L=7,115,36
      SUMA(KA) = 0.0
      N = L + 35
      DO 102 M=L,N
  102 SUMA(KA) = SUMA(KA) + S(M)
  103 KA = KA + 1
      SUMA(6)=6.0*(S(151)+S(152)+S(153)+S(154)+S(155)+S(156))
      KB = 1
      DO 105 M=7,37,6
      SUMB(KB) = SUMA(6)/6.0
      DO 104 L=M,150,36
      K = KB
      SUMB(K)=S(L)+S(L+1)+S(L+2)+S(L+3)+S(L+4)+S(L+5)+SUMB(K)
  104 SUMB(K) = SUMB(K)+1.5*S(K)
  105 KB = KB + 1
      DO 107 M=7,12
      SUMC(M-6) = SUMA(1)/6.0
      DO 106 L=M,150,6
  106 SUMC(M-6) = SUMC(M-6) + S(L) + 0.25*S(M+144)
  107 CONTINUE
      FA(1) = 0.0
      FA(2) = 0.2
      FA(3) = 0.4
      FA(4) = 0.6
      FA(5) = 0.8
      FA(6) = 1.0
      PRINT 108
  108 FORMAT(32HSUM OF ERROR VARIANCES ON A,B,C    )
      DO 109 K=1,6
  109 PRINT 110,FA(K),SUMA(K),SUMC(K),SUMB(K)
  110 FORMAT(F7.3,F20.6,F20.6,F20.6)
      PAUSE
      GO TO 1
   99 END
```

```
C      PROGRAM-5  THIS PROGRAM PROVIDES THE FORECASTS
C      INPUT REQUIRED - ORIGINAL DATA DECK AND SMOOTHING
C      CONSTANTS IN 3F10.6 FORMAT
C      OUTPUT - TYPEWRITER
C      OPTIONS
C      SENSE SWITCH 1 - MORE THAN ONE SET OF CONSTANTS
C      SENSE SWITCH 2 - MORE THAN ONE SET OF DATA
       DIMENSION OBE(60),OBS(12,5),YRAV(5),TREND(60),UNAD(12)
       DIMENSION SEA(60),EST(60),PRI(12),SENA(12),SEANL(12,5)
     1 MOTOL = 0
       I=1
     2 READ3,OBE(I)
     3 FORMAT (F10.6)
       IF(OBE(I) -999.999999) 4,5, 99
     4 MOTOL = MOTOL +1
       I = I + 1
       GO TO 2
     5 MYR = MOTOL/12
       IF(MOTOL - 12* MYR -9) 6,7,7
     6 MOT = 12*(MYR-1)
       MYR = MYR - 1
       GO TO 8
     7 MOT = 12* MYR
     8 FMYR = MYR
       L = 1
       DO 10 J = 1, MYR
       YRAV(J)=0.0
       DO 9 I = 1, 12
       UNAD(I)=0.0
       OBS (I,J) = OBE (L)
       YRAV(J) = YRAV(J) + OBE (L)/12.0
     9 L = L + 1
    10 CONTINUE
       DENOM = MOTOL - MOT
       NOT = MOT +1
       SMVAR = (0.9E+25)*(0.9E+25)
       AMOT = MOT
       TREN      = (YRAV(MYR)-YRAV(1))/(AMOT-12.0)
       DO 12 J = 1, MYR
       SUSEA = 0.0
       DO 12 I = 1, 12
       FI = I
       SEANL(I,J) = OBS(I,J)/(YRAV(J)-(6.5-FI)*TREN)
       UNAD(I) = UNAD(I) + (SEANL(I,J)/FMYR)
       IF(UNAD(I))11,12,12
    11 UNAD(I) = (-1.0)*UNAD(I)
    12 SUSEA = SUSEA + UNAD(I)
       PLY = 12.0 /SUSEA
       DO 13 I = 1,12
    13 SENA(I) = UNAD(I) * PLY
    14 READ 15,A,B,C
    15 FORMAT(F10.6,F10.6,F10.6)
```

```
      EST(1)=A*(OBE(1)/SENA(1))+(1.0-A)*(YRAV(1)+TREN)
      TREND (1)=C*(EST(1)-YRAV(1))+(1.0-C)*(TREN     )
      SEA(1) = B*(OBE(1)/EST(1))+(1.0-B)*SENA(1)
      DO 19 I = 2, MOTOL
      IF(I-12)16,16,17
   16 J=I
      SEA(J) = SENA(J)
      GO TO 18
   17 J=I-12
   18 EST(I)=A*(OBE(I)/SEA(J))+(1.0-A)*(EST(I-1)+TREND(I-1))
      SEA (I) = B*(OBE(I)/EST(I))+(1.0-B)*SEA(J)
   19 TREND (I) = C*(EST(I)-EST(I-1))+(1.0-C)*TREND(I-1)
      PRINT 500
  500 FORMAT(44HPERIODS INTO        FORECASTED        SEASONAL)
      PRINT 501
  501 FORMAT(44H   FUTURE              VALUE             FACTOR )
      DO 502 LT=1,12
      T = LT
      IND = (MOTOL + LT - 12)
      PRI(LT) = (EST(MOTOL) + T*(TREND(MOTOL)))*SEA(IND)
  502 PRINT 503, LT, PRI(LT), SEA(IND)
  503 FORMAT(I7,F22.5,F14.6)
      IF(SENSE SWITCH 1) 14,504
  504 IF(SENSE SWITCH 2) 1,99
   99 END
```

VITA

Glendon Danna Self

Candidate for the Degree of

Doctor of Philosophy

Thesis: SELECTION OF SMOOTHING CONSTANTS FOR AN EXPONENTIALLY WEIGHTED TIME SERIES MODEL

Major Field: Engineering

Biographical:

Personal Data: Born in Waveland, Arkansas, January 1, 1938, the son of Charlie W. and Alma Carrie Self.

Education: Attended grade school in Havana, Roland, and Jacksonville, Arkansas; graduated from Jacksonville High School in 1954; received the Associate in Science degree from the Arkansas Polytechnic College, with a major in Engineering, in May, 1956; received the Bachelor of Science degree from the University of Arkansas, with a major in Industrial Engineering, in June, 1958; received the Master of Science degree from the University of Arkansas, with a major in Industrial Engineering, in January, 1959; completed requirements for the Doctor of Philosophy degree in August, 1963.

Professional experience: Worked for Sandia Corporation as a Statistical Quality Control Engineer and as a Consultant since March, 1959, except for the school year of 1960-61; the work on time series analysis of personnel attrition rates began during June, 1960, and has been a continuing part-time assignment that is concluded with the preparation of this thesis.