

DEVELOPMENT OF A DEMAND FORECASTING MODEL  
FOR  
FLORAL DATA

by 544

WILLIAM FRANCIS IRVINE

B. S., University of Notre Dame, 1966

---

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

1968

Approved by:



---

Major Professor

LD  
2668  
T4  
1968  
I 185  
1.2

## TABLE OF CONTENTS

	Page
INTRODUCTION . . . . .	1
CHAPTER I. Representation of the Time Series. . . . .	5
CHAPTER II. Data Analysis . . . . .	16
CHAPTER III. Forecasting the Time Series. . . . .	45
CHAPTER IV. Choosing the Smoothing Constant and the Initial Coefficient Vector . . . . .	65
CHAPTER V. Gathering the Data . . . . .	80
CHAPTER VI. Discussion of Results . . . . .	86
CHAPTER VII. Conclusions . . . . .	106
BIBLIOGRAPHY . . . . .	109
ACKNOWLEDGEMENTS . . . . .	110
APPENDIX A. Derivation of Statistical Regression Equation . . . . .	111
APPENDIX B. Derivation of Exponential Smoothing Equation. . . . .	114
APPENDIX C. Polynomial Regression . . . . .	118
APPENDIX D. The Z-Transform . . . . .	120
APPENDIX E. Tables and Figures. . . . .	122
APPENDIX F. Computer Programs . . . . .	184

## INTRODUCTION

A common problem in business and industry today is the accurate estimation of the level of inventory necessary to meet some future demand. In some situations this task may be quite critical. This can best be explained using an example. Since this research concerns the floral industry, let us use as an example the demand for a floral product.

Consider a plant which has a growing period of two months from the time it is first planted until it is ready to be sold. In order to know how many plants to start now, we must have some idea of how many will be needed for sale two months from now. If we underestimate this future demand, we will not have sufficient plants in stock that are ready for sale. Conversely, if we overestimate future demand we will have too many plants and may be forced to sell them at a loss or even throw them away. Thus we see that it is desirable to be able to forecast this future demand with the confidence that this forecast will not be too far from the actual demand two months after planting. The importance of forecasting can be seen almost anywhere an inventory is maintained.

The logical basis for a forecast of demand for any item is the demand history of that item. Normally, such past demand data is in the form of a time series, a sequence of observations made at specific and equal time intervals. There are several ways in which future demands may be interpreted from this time series. Among these methods, subjective reasoning is probably the simplest. This is merely the opinion of

an experienced manager of what demand will be. Other methods include statistical regression and graphical curve fitting. All of these methods may be classified as static; that is, they deal in some way with fitting a curve to the data as it now exists. Perhaps a more realistic approach is to give the recent data more weight than older data on the assumption that this newer data more accurately foretells the behavior of demand in the future. Methods employing such weighting factors may be classified as dynamic. The moving average is one approach to weighting. This method takes an average of the last  $N$  data values as a forecast of demand for the next sampling interval, where  $N$  is some pre-determined number of sampling intervals. Notice that if  $N$  is large we consider much past data, and if  $N$  is small we put the emphasis on only the most recent data. Thus we may control the rate of response of this forecasting system by varying  $N$ . If we wish to include all past data, we may use the general exponential smoothing technique. This is the method used for this research. General exponential smoothing gives data taken  $k$  periods in the past the weight  $\alpha\beta^k$ , where  $\alpha$  and  $\beta$  are constants and their sum equals one. Now we may control the system response rate to changes in the data by varying the value of  $\beta$ , known as the discount factor.

General exponential smoothing essentially fits a least squares curve to the weighted time series. By least squares we mean that the curve is determined in a manner that minimizes the sum of weighted square residuals, or fitting errors. This method provides added adaptability to changes in the data by re-evaluating the model coefficients at each sampling interval. Let us illustrate the meaning of this statement. We

may describe the time series by:

$$x(t) = \xi(t) + \epsilon(t)$$

where  $x(t)$  = the past data time series at time  $t$

$\xi(t)$  = the systematic component of  $x(t)$  at time  $t$

$\epsilon(t)$  = some random variation of  $x(t)$  from  $\xi(t)$  at time  $t$ .

We are assuming, of course, that the time series contains some mathematically describable process  $\xi(t)$ . We wish to formulate a model to describe  $\xi(t)$  of the form

$$\hat{x}(t) = \sum_{i=1}^n a_i(t) f_i(t)$$

where  $\hat{x}(t) = \xi(t)$ , if the model above is the correct one

$f_i(t)$  = a deterministic function of time

= the  $i$ th fitting function

$a_i(t)$  = the least squares coefficient of  $f_i(t)$ .

Now, by a re-evaluation of model fit we mean that general exponential smoothing re-computes the least squares fitting function coefficients of the model at each sampling interval. In this updating procedure the error made on the previous forecast is taken into account. Thus, we may note several advantages of general exponential smoothing over the methods discussed earlier. First, it accounts for all past data. Secondly, exponential smoothing gives the investigator more immediate control over the forecasting system response rate. As is noted in the text to follow, this method also requires less data storage once it is

started, since all past data contained in one word of information. A fourth advantage of general exponential smoothing is the re-evaluation of the least squares coefficients in the light of each additional amount of data.

The methods used here are taken primarily from Smoothing, Forecasting and Prediction of Discrete Time Series by Robert Goodell Brown (2). All computations are performed on the IBM System 360/50 digital computer at Kansas State University. The data used consists of four discrete time series of demand for potted plants:

- (1) Wholesale demand for chrysanthemums in monthly dollar receipts.
- (2) Retail demand for chrysanthemums in monthly dollar receipts.
- (3) Wholesale demand for lilies in daily plant sales.
- (4) Wholesale demand for poinsettias in daily plant sales.

All time series were taken from the sales records of Manhattan Retail and Wholesale Floral Companies in Manhattan, Kansas. Chrysanthemums are sold throughout the entire year, while lilies and poinsettias are seasonal items, sold only during the Easter and Christmas seasons, respectively.

This particular data was used due to the interest of the author. Having been employed by Manhattan Floral Company for several years, the author feels that he has a greater understanding of the problems involved in this area than in any other. In addition, data from this source could be more easily obtained and verified.

## 1. REPRESENTATION OF THE TIME SERIES

### 1.1 Form of the Time Series Model

A time series is a set of observations taken at specific points in time, usually at equal intervals. The analyses that follow assume that the time series involved may be separated into two components, systematic and random. The systematic component can be represented as a deterministic function of time. Thus, the time series may be represented mathematically as:

$$x_t = \xi_t + \epsilon_t$$

where

$x_t$  = the actual time series observation at time  $t$ .

$\xi_t$  = the systematic component of the observation at time  $t$ .

$\epsilon_t$  = the random component of the observations at time  $t$ .

The distribution of the random component of the time series,  $\epsilon_t$ , possesses the following properties:

$$E[\epsilon_t] = \text{Expected value of } \epsilon_t = 0$$

$$E[\epsilon_t \cdot \epsilon_{t-j}] = 0 \text{ for } j \neq 0$$

$$= \sigma^2 \text{ for } j=0$$

where  $\sigma^2$  is the variance of the random noise distribution. Successive samples of the random component are assumed to have no serial correlation.

Having defined  $\epsilon_t$  as random with a mean value of zero, we make no

attempt to describe it. So, in order to represent the time series  $x_t$ , except for the random variation,  $\epsilon_t$ , we must be able to describe the systematic component,  $\xi_t$ . Once we are able to describe  $\xi_t$ , we may predict future values of  $x_t$ , since  $\xi_t$  is a function of time only and time is always known for the future.

To better visualize the time series, let us look at the plot shown below.

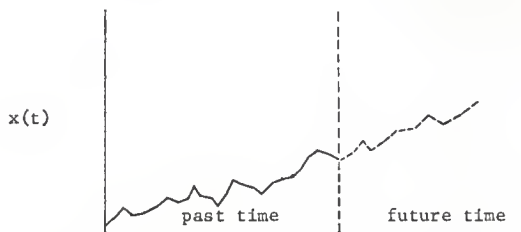


Fig. 1:1:1 Plot of hypothetical time series

This is the plot of a time series  $x_t$  on the vertical axis, versus time on the horizontal axis.  $x_t$  may be any time series for which we wish to predict future values.

The following are a few of the methods that may be employed to attack this forecasting problem:

1. Subjective estimate
2. Graphical curve fitting
3. Statistical regression analysis
4. Simulation
5. Moving averages
6. Exponential smoothing.



For reasons that will be explained in sections 1.2 and 1.3, we may classify the first three methods above as static methods, and the last three as dynamic methods.

## 1.2 Static Methods of Time Series Representation

The first of the static methods, subjective estimate, implies an opinion, guided by reason, experience, and judgment, of what  $x_t$  will be for a particular point in time,  $t$ . This prediction would normally be made by a person or persons familiar with the time series involved. It is not based on any particular analytical foundation.

Graphical curve fitting may be accomplished by merely observing the time series plot, like the one in Figure 1:1:1, and drawing a representative curve through the data points shown in the "past time" section of the plot. These points are actual data gathered from past performance of  $x_t$ .

By extending this curve into "future time" a prediction of the future values of  $x_t$  may be made. This method, though not totally subjective, may fail to adequately represent any cyclic fluctuation inherent in the time series.

Probably the major analytical method of static time series analysis is statistical regression. This method attempts to formulate a model that will describe the systematic component of the time series,  $\xi_t$ . The model may contain any independent variables that are known exactly for all future time. It is up to the investigator to choose the proper terms and their form within the model. Statistical regression techniques are

used to determine the proper coefficients of the terms in the model. These techniques will give coefficients that effect a minimization of the sum of squared residuals,  $S$ , for the time series values given (Appendix A).

$$S = \sum_{t=0}^N [x_t - \hat{x}_t]^2$$

where

$N$  = the number of data values given

$x_t$  = the actual time series value for time  $t$

$\hat{x}_t$  = the time series value for time  $t$  as determined from the model.

Although this method minimizes the sum of the squared fitting errors for the data given, it makes no provision for a change in the general form of the time series through time. As is true with all static methods of time series representation, statistical regression makes no attempt to reevaluate the model coefficients at each point in time to reflect the forecast error made in the previous prediction.

### 1.3 Dynamic Methods of Time Series Representation

Simulation techniques attempt to formulate a process model that will behave like the actual process over some specified period of time. For example, we attempt to establish probability distributions that are representative of the variables contributing to the process, such as demand, production, etc. These distributions will generally be based on

past behaviour of the actual process, if such data is available. By using a model to simulate the true process over some period of time, we may gain an insight as to what values the variables will exhibit in the future.

Normally we would expect the more recent values in the time series to be more representative of future values than the older data. The last two dynamic methods to be discussed account for this probability by paying much attention to recent data and little or no attention to older data. This important feature of dynamic methods of time series representation allows the estimates,  $\hat{x}_t$  to reflect a change in the form of the time series as time progresses.

The moving average is a simplified form of dynamic time series representation. For purposes of illustration, consider a time series for which the systematic component may be represented as a constant process.

$$x_t = a_t + \epsilon_t$$

where

$x_t$  = the observed time series value for time  $t$

$a_t = \xi(t)$  = the value of a constant process at time  $t$

$\epsilon_t$  = the random variation from the true process at time  $t$ .

It will become clear in the explanation that follows that  $a_t$  may actually change slowly with time. Since this is a dynamic representation, this change in the true process will be accounted for in the analysis.

Since we have defined the systematic component of the time series,  $\xi_t$ , as a deterministic function of time, we may construct a function, or

model, that will describe the time series, except for a random variation,  $\xi_t$ . In this case  $a_t$  is our model of the true process, with time as the only independent variable.

As an estimate of  $a_{T+1}$ , we might take an average of all the observations,  $x_t$ , up to the present time,  $T$ . So,

$$\hat{a}_{T+1} = \sum_{t=0}^T x_t / T$$

where

$$\hat{a}_{T+1} = \text{the estimated value of the true process, } a_t, \text{ at} \\ \text{time } t = T + 1.$$

A better estimate of  $a_{T+1}$  might be an average over only the most recent  $N$  observations. This is what is known as a moving average  $M_t$ . Now,

$$\begin{aligned} \hat{a}_{t+1} &= M_t = \frac{x_t + x_{t-1} + \dots + x_{t-N+1}}{N} \\ &= M_{t-1} + \frac{x_t - x_{t-N}}{N} \end{aligned}$$

$M_t$  = the moving average of  $x$  over the most recent  
 $N$  observations, or the moving average of  
period  $N$ .

$M_t$  minimizes the sum of squares of differences between the most recent  $N$  observations and our estimates  $\hat{a}_{t+1}$  of the process model. Note that when  $N$  is large we are giving more weight to older data than when  $N$  is

small.

The moving average, though it is a dynamic method of forecasting, still has two important limitations. First, it does not consider all past data, and gives equal weight to the observations that it does consider. Second, it is difficult to change the response rate of a system based on moving averages. Though the system response is based on the value of  $N$ , it would be difficult to alter a record system based on a particular  $N$  value. Also, if there are significant changes in the true process through time, the moving average will give highly unsatisfactory forecasts.

Again for purposes of illustration, suppose that we lost all our data, except the value of the previous moving average  $M_{T-1}$ . Then, by this same scheme, our best estimate for  $x_{T+1}$  would be,

$$\begin{aligned}\hat{a}_{T+1} &= \hat{M}_T = M_{T-1} + \frac{x_T - M_{T-1}}{N} \\ &= \frac{1}{N}x_T + \left(1 - \frac{1}{N}\right) M_{T-1}\end{aligned}$$

where

$$\hat{M}_T = \text{an estimate of } M_T.$$

This process is known as "smoothing." Now let,

$$\begin{aligned}\alpha &= \frac{1}{N} \\ \beta &= 1 - \frac{1}{N} \\ S_t(x) &= \hat{M}_t.\end{aligned}$$

We now have

$$S_t(x) = \alpha x_t + \beta S_{t-1}(x)$$

where

$S_t(x)$  = the smoothed value of  $x$  at time  $t$ .

$S_{t-1}(x)$  = the smoothed value of  $x$  at time  $t-1$ .

$\alpha$  = the smoothing constant

$\beta$  = the discount factor

$\alpha$  and  $\beta$  are not exactly the same as  $\frac{1}{N}$  and  $(1 - \frac{1}{N})$ , but are similar to them. We shall see their significance later in this chapter. Note here that,

$$\alpha, \beta \leq 1$$

$$\alpha + \beta = 1.$$

This operation, carried out over a sequence of observations, is called exponential smoothing. To see this effect going back in time we merely expand our original smoothing equation.

$$\begin{aligned} S_t(x) &= \alpha x_t + \beta S_{t-1}(x) \\ &= \alpha x_t + \beta [\alpha x_{t-1} + \beta S_{t-2}(x)] \\ &= \alpha x_t + \alpha \beta x_{t-1} + \alpha \beta^2 x_{t-2} \\ &\quad + \dots + \alpha \beta^n x_{t-n} + \dots \\ &\quad + \beta^n x_0. \end{aligned}$$

The exponent of  $\beta$  increases geometrically as we go back in time. So, the weight given the initial  $x$  value,  $x_0$ , after  $k$  observations is  $\beta^k$ . Hence, we call this exponential smoothing.

This procedure offers several advantages over the methods discussed earlier. First, all the past data is taken into consideration. The older the data is, the less weight it receives when determining  $S_t(x)$ . Also, we may change the response rate of the system by merely altering  $\alpha$ , the smoothing constant, or  $\beta$ , the discount factor. So, for  $\beta$  small, we are giving more weight to past data than if  $\beta$  were larger. If  $\beta$  is large, then, the system will adjust more quickly to a wide fluctuation in the observations. The third advantage of smoothing is that all past data is presented by only one value,  $S_t(x)$ , instead of  $N$  observations as in the case of the moving average. Lastly, exponential smoothing minimizes the weighted sum of squared residuals (Appendix B).

So far in our discussion, we have been dealing with the simplest of processes, the constant process. Let us now consider a general model of the true process.

$$\begin{aligned} x(t) &= a_1(t) f_1(t) + a_2(t) f_2(t) + \dots + a_n(t) f_n(t) + \varepsilon(t) \\ &= \sum_{i=1}^n a_i(t) f_i(t) + \varepsilon(t) \end{aligned}$$

where

$f_i(t)$  = the  $i^{\text{th}}$  fitting function at time  $t$ .

$a_i(t)$  = the coefficient of the  $i^{\text{th}}$  fitting function at time  $t$ .

See section 2.1 for a discussion of allowable fitting functions. In this

expression the systematic component of the time series is represented by

$$\sum_{i=1}^n a_i(t) f_i(t),$$

which then represents the model of our system. The fitting functions are deterministic functions of time. In the case of the constant process, we had only one fitting function whose value was the constant, 1. The coefficients  $a_i(t)$  are treated as constants over any local period of time. Nevertheless, the underlying process  $\xi(t)$  can be going through a slow random walk with respect to one or more of the coefficients. As will be shown later, the exponential smoothing procedure accounts for this.

For future reference, let us write the general time series representation in matrix notation.

$$x(t) = \sum_{i=1}^n a_i(t) f_i(t) + \epsilon(t)$$

$$= a'(t) f(t) + \epsilon(t)$$

where

$$a(t) = \begin{bmatrix} a_1(t) \\ a_2(t) \\ \vdots \\ a_n(t) \end{bmatrix}, \text{ and } f(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix}$$

Hence

$$\hat{x}(t) = \hat{a}'(t) f(t).$$



It should be apparent at this point that the first requirement in the formulation of a forecasting procedure for a time series is the proper choice of the fitting functions that make up the model. There are several guidelines we may follow in making this choice. These guidelines are discussed in the next chapter.

The exponential smoothing procedure will be dealt with in much greater detail in Chapter III.

## 2. DATA ANALYSIS

### 2.1 Choice of Fitting Functions for the Model

In the first chapter we described the systematic component of the time series.  $\xi_t$  is composed of fitting functions,  $f_i(t)$ , and their corresponding coefficients,  $a_i(t)$ . In matrix notation this is

$$\xi_t = \hat{x}(t) = \hat{a}'(t) f(t)$$

In this chapter we shall discuss some methods for determining the proper fitting functions to be used in the model,  $\hat{x}_t$ .

The time series data with which we are dealing dictates the form of the model fitting functions. The problem arises in determining from the data those fitting functions that will best describe the process inherent within the time series. A discussion of the possible fitting function forms will serve to clarify the alternatives available. Keep in mind that, by definition, the fitting functions must be known exactly for future time.

Brown (2,58) classifies time series models into four groups:

1. Algebraic models
2. Transcendental models
3. Composite models
4. Regression models.

The first of these classifications, algebraic models, takes in all polynomial functions of time, including the constant model with which we dealt in Chapter 1. Our forecast in the case of a constant model, or polynomial model of degree zero, is

$$\begin{aligned}
 \hat{x}(t) &= \hat{a}(t) f(t) \\
 &= \hat{a}(t) t^0 \\
 &= \hat{a}(t) (1) \\
 \hat{x}(t) &= \hat{a}(t)
 \end{aligned}$$

where  $\hat{a}(t)$  = an estimate from the smoothing process of the fitting function coefficient. As we shall show in Chapter 3, the coefficients are re-estimated by the exponential smoothing technique each time additional data has become available. Thus we allow for gradual changes in the process described by our model. This model might be used to represent a time series that changes gradually, if at all, from a constant value. A pictorial example of this type of process is shown in Figure 2:1:1. The plot is not shown as a smooth line because of the random variation contributed by  $\epsilon(t)$ , the "noise" element in  $x(t)$ .

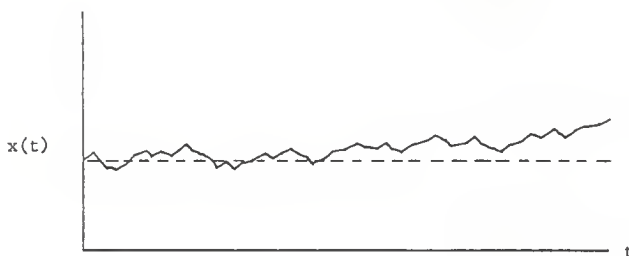


Fig. 2:1:1 Constant process time series

Going one step further, consider the case where the time series exhibits a nearly linear form, as in Figure 2:1:2. A first degree poly-

nomial model

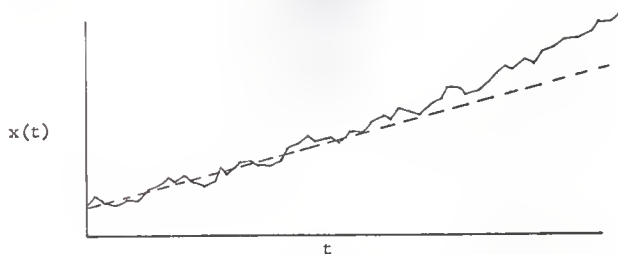


Fig. 2:1:2 Linear process time series

would probably describe this process adequately. So,

$$f_1(t) = 1$$

$$f_2(t) = t$$

and

$$\begin{aligned}\hat{x}(t) &= \sum_{i=1}^2 \hat{a}_i(t) f_i(t) \\ &= \hat{a}_1(t) (1) + \hat{a}_2(t) t \\ &= \hat{a}(t) \hat{f}(t)\end{aligned}$$

where

$$\hat{a}(t) = \begin{bmatrix} \hat{a}_1(t) \\ \hat{a}_2(t) \end{bmatrix} \quad \text{and } f(t) = \begin{bmatrix} 1 \\ t \end{bmatrix}$$

Again, the plot shows, as always, the existence of some random "noise" in the data.

We may similarly formulate a polynomial model of any desired degree. Let us describe an  $n^{\text{th}}$  degree polynomial model as one containing a constant

term and a polynomial of degree  $n-1$ . Then, an  $n^{\text{th}}$  degree polynomial model would be,

$$\begin{aligned}\hat{x}(t) &= \sum_{i=1}^n \hat{a}_i(t) f_i(t) \\ &= \hat{a}_1(t) (1) + \hat{a}_2(t)t + \frac{\hat{a}_3(t)t^2}{2!} \\ &\quad + \dots + \frac{\hat{a}_n(t)t^{n-1}}{(n-1)!} \\ &= \hat{a}(t) \hat{f}(t)\end{aligned}$$

where

$$\hat{a}(t) = \begin{bmatrix} \hat{a}_1(t) \\ \hat{a}_2(t) \\ \hat{a}_3(t) \\ \vdots \\ \hat{a}_n(t) \end{bmatrix} \quad \text{and } \hat{f}(t) = \begin{bmatrix} 1 \\ t/1! \\ t^2/2! \\ \vdots \\ t^{n-1}/(n-1)! \end{bmatrix}$$

Note that the  $(n-1)^{\text{th}}$  degree polynomial model requires the estimation of  $n$  fitting function coefficients.

Brown (2, 63) notes that the number of observations in the span covered by the model should be several times the number of degrees of freedom,  $n$ . Logically, forecasts based on only a few observations would be very unreliable.

The question now arises as to what degree of polynomial is appropriate

for a particular time series. If the observations may be exactly represented by a polynomial of degree  $n$ , i.e. no "noise", that degree may be determined by successive differences. In this case,

$$x_t = \xi(t)$$

$$\varepsilon_t = 0$$

We denote a first difference as

$$\Delta_t \xi = \xi(t) - \xi(t-1)$$

and a second difference as

$$\begin{aligned} \Delta_t^2 \xi &= \Delta_t \xi - \Delta_{t-1} \xi \\ &= \xi(t) - 2\xi(t-1) + \xi(t-2) \end{aligned}$$

Thus, a difference of order  $k$  is simply

$$\Delta_t^k = \Delta_t^{k-1} \xi - \Delta_{t-1}^{k-1} \xi$$

If the appropriate polynomial is of degree  $n$ , the difference  $\Delta^{n+1}$  will be identically zero. In the more practical case in which  $\varepsilon_t$  is not zero, the  $(n+1)$  st differences of the data  $\Delta^{n+1} x$  will average zero over time, if the process  $\xi$  is a polynomial of degree  $n$ . The variance of the differences will grow as the sum of squares of the binomial coefficients in the expansion of  $\Delta^{n+1} x$ . So, if  $\varepsilon$  has a variance  $\sigma^2$ , the values  $\Delta^{n+1} x$  will have variance

$$\sigma^2 [2(n+1)]! / [(n+1)!]^2$$

This variance is larger than  $\sigma^2$  for  $n+1 \geq 5$  (Brown, 2, 63).

In many cases the proper polynomial degree may be determined by merely inspecting a plot of the data versus time. Statistical least squares regression analysis is another very useful method of attaining this end. This last approach will be mentioned in our discussion of trend analysis.

Brown's second classification of time series models, transcendental models, is characterized by two sub-classifications:

- a. Exponential models
- b. Trigonometric models

These two will be dealt with separately.

An exponential function may be used to describe a process in which the rate of growth is proportional to the state of growth in time, like compound interest. Here, the change in value from  $\xi_{(p)}$  to  $\xi_{(t+1)}$  may be expressed as a constant percentage or  $\xi(t)$ , for any  $t$ . The form of  $\xi$  for a process which is a simple exponential function is

$$\xi_{(p)} = \hat{x}_{(p)} = ka^t$$

Note that

$$\xi_{(t-1)} = ka^{t-1}$$

so

$$\xi(t)/\xi_{(t-1)} = ka^t/ka^{t-1} = a$$

where

$a =$  a constant in the short run.

This simple model, when plotted on semilogarithmic paper, will define a nearly straight line with slope  $\log a$ , since

$$\log \xi_{(A)} = \log k + t \log a$$

The general form of the exponential model is

$$\begin{aligned} \xi(t) = \hat{x}(t) = & k_1 \binom{t}{0} a^t + k_2 \binom{t}{1} a^{t-1} b + k_3 \binom{t}{2} a^{t-2} b^2 \\ & + \dots + k_n \binom{t}{n-1} a^{t-n+1} b^{n-1} \end{aligned}$$

where

$$\binom{t}{k} = \frac{t!}{(t-k)!k!}$$

is the binomial coefficient of the  $(k+1)$ st term in the expansion of  $(a+b)^t$ .

To describe processes which are periodic, it is most convenient to use the second of the transcendental model types, trigonometric models. By periodic we mean a process whose time plot shows repetitive patterns. The length of time between similar sequence patterns is called the period. Perhaps the best graphical examples of this phenomenon are the

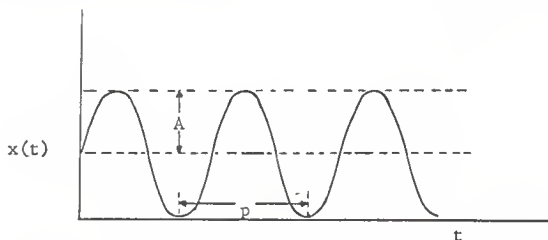


Fig. 2:1:3 Sine wave with amplitude  $A$  and period  $p$



sine and cosine wave forms.

There are several familiar causes of cyclical variations in an economic time series. Davis (3, 26) mentions some of these, a few of which are mentioned below. Perhaps most obvious cause for this phenomenon is found in the seasons of the year. A 12-month period usually results from this seasonal variation. A second common pattern in economic data is the "40-month cycle" which corresponds to that of our national economy. Davis also mentions a 15-20 year "building cycle" shown by many production series. We cannot expect these cyclical variations to be completely uniform, since in nearly every case, we are dealing with human conduct.

The theory of Fourier series provides the basis for time series representation by trigonometric models. Fourier showed that any reasonable periodic function of time can be represented by taking a sufficient number of terms in the series

$$\begin{aligned} \xi(t) = & A_0 + A_1 \cos\left[\frac{2\pi}{P}(t-\alpha_1)\right] + A_2 \cos\left[\frac{4\pi}{P}(t-\alpha_2)\right] \\ & + \dots + A_n \cos\left[\frac{2k\pi}{P}(t-\alpha_k)\right] + \dots \end{aligned}$$

where

$$A_0, A_1, A_2, \dots, A_n, \dots = \text{constants.}$$

Applying the identities

$$\begin{aligned} A_1 &= \sqrt{A_1'^2 + B_1'^2} = \text{amplitude} \\ \alpha_1 &= \arctan \frac{A_1'}{B_1'} = \text{phase angle} \end{aligned}$$

we may transform our equation for  $\xi(t)$  into a series of sines and cosines.

$$\begin{aligned}\xi(t) = & A'_0 + A'_1 \sin \frac{2\pi t}{P} + A'_2 \sin \frac{4\pi t}{P} + \dots \\ & + A'_k \sin \frac{2k\pi t}{P} + \dots \\ & + B'_1 \cos \frac{2\pi t}{P} + B'_2 \cos \frac{4\pi t}{P} + \dots \\ & + B'_k \cos \frac{2k\pi t}{P} + \dots\end{aligned}$$

where

$p/k$  = the periods of the cycles represented by each sine-cosine pair

and

$A'_0, A'_1, A'_2, \dots, B'_1, B'_2, \dots$  = constants.

As before, we could write this model in the form

$$\begin{aligned}\xi(t) = \hat{x}(t) &= \sum_{i=1}^n \hat{a}_i(t) f_i(t) \\ &= \hat{a}'(t) f(t)\end{aligned}$$

where

$$\hat{a}'(t) = \begin{bmatrix} A'_0 \\ A'_1 \\ B'_1 \\ A'_2 \\ B'_2 \\ \vdots \end{bmatrix} \quad \text{and} \quad f(t) = \begin{bmatrix} 1 \\ \sin 2\pi t/P \\ \cos 2\pi t/P \\ \sin 4\pi t/P \\ \cos 4\pi t/P \\ \vdots \end{bmatrix}$$

Again note that the coefficients are time-dependent to allow for gradual changes in the process. Thus they must be re-estimated periodically as additional data is made available. By including two terms, sine and cosine, for each cycle period, we allow the origin of time and the mean to be arbitrary; i.e., we allow for a difference in phase angle. Observing the process at discrete sampling intervals, the highest frequency that can be detected is the Nyquist frequency such that there are two observations per cycle. Brown (2, 74) warns us that trigonometric models should be used only where there is a known underlying cause for a periodic time series form.

Having described the basic mathematical time series forms in their pure states, we move now to the consideration of more practical examples. This brings us to Brown's third time series model classification, composite models. These present nothing new, except the combination of the previous models, algebraic and transcendental. The methods of exponential smoothing make it possible to estimate coefficients for any model composed of terms that are sums or products of algebraic and transcendental functions of time, in any combination whatsoever.

An example may best serve to illustrate the utility of such a composite model. Consider the hypothetical, though typical, curve shown in Figure 2:1:4.

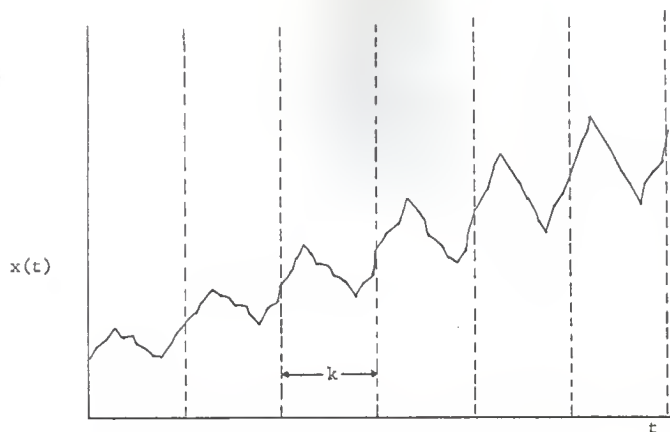


Fig. 2:1:4 Hypothetical time series plot

There are several obvious elements in the process which describes this time series. First, the model will probably contain constant and linear terms to show the starting point and apparent linear trend of the series. In addition to these terms, a pair of cyclic terms with period  $k$  will probably be necessary. Since the amplitude of the cyclic variation appears to be increasing, we may want to include the product of a polynomial term and a pair of trigonometric terms. Thus far, we have formulated the following representation for the time series,  $x(t)$ .

$$\begin{aligned}
 x(t) &= \xi(t) + \varepsilon(t) \\
 &= \hat{x}(t) + \varepsilon(t) \\
 &= \sum_{i=1}^6 a_i(t) f_i(t) \\
 &= a_1 + a_2 t + a_3 \sin \frac{2\pi t}{k} + a_4 \cos \frac{2\pi t}{k} \\
 &\quad + a_5 t \sin \frac{2\pi t}{k} + a_6 t \cos \frac{2\pi t}{k}
 \end{aligned}$$

$$= a_1 + a_2 t + (a_3 + a_5 t) \sin \frac{2\pi t}{k} \\ + (a_4 + a_6 t) \cos \frac{2\pi t}{k}.$$

where  $\xi(t) = \hat{x}(t)$ , if our model is the correct one

$a_i = \text{constant for } i = 1, 2, \dots, 6$ , for the short run.

In actual practice,  $a_i = \hat{a}_i(t)$ ; that is, we must estimate the fitting function coefficients from our smoothing process. So,  $\xi(t) = \hat{x}(t)$  only if we estimate  $a_i(t)$  accurately and our model is the correct one. If it is felt necessary, additional terms can be included in the process model, such as a higher harmonic to the cyclic component. It may seem at first glance that we should try to describe  $\xi(t)$  as accurately as possible; however, as Brown notes, for each degree of freedom added to the model the work required to estimate the coefficients will approximately double, and the computing time will increase proportionally to  $n^3$ .

The fourth and last of Brown's time series model classifications is regression models. This classification covers a wide group of linear forecast models of the form

$$\xi(t) = a_1 f_1(t) + a_2 f_2(t) + \dots + a_n f_n(t)$$

where  $f_i(t)$  can be any arbitrary function at all. Many times the fitting functions will be one or a combination of "leading" series related to  $x(t)$ . Brown offers several arguments against the use of regression models. First, there seems to be no statistical test which

can be applied to the model. Secondly, one usually doesn't know the independent series long enough in advance to be useful in a forecast. Finally, the computational techniques involved are considerably less efficient than those of exponential smoothing with algebraic and transcendental types of fitting functions. As we will show in a later section, the efficiency of computation in exponential smoothings depends on the mathematical relationships between fitting function values at different sampling intervals. We can not establish such relationships when the fitting functions are arbitrary time series.

## 2.2 Trend Analysis

By trend, or secular trend as it is more commonly called, we mean "...that characteristic of the series which tends to extend consistently throughout the entire period" (Davis, 3, 15). The period referred to in this definition is the total time covered by the available data. Trend analysis is an attempt to determine this secular trend in the time series. It is performed on the data for several reasons:

1. To determine the polynomial component of the time series model.
2. To aid in detrending the time series for the auto correlation and spectral analyses to be discussed in the next two sections of this chapter. By detrending we mean to reduce the series in such a manner that the residual fitting errors will have an expected value of zero.
3. To aid in identification of any amplitude growth that may be occurring within the cyclic components of the time series.

Davis goes on to describe four of the more commonly encountered trend forms in economic time series analysis. The first of these is the straight line trend, described by a first degree polynomial on  $t$ , time. Another common trend type is a straight line fitted to data plotted on a logarithmic scale. This type corresponds to the simple exponential model discussed earlier. He calls the third trend line type a "logistic curve" or "curve of growth". This curve is more applicable to a description of the growth of new industries and population series. It may be described as a transition line between a lower initial level of  $x(t)$ , to a stable "saturation" point on a higher level. The logistic curve is represented mathematically by

$$\xi(t) = \frac{K_1}{1 + e^{-K_2 t}}$$

where

$K_1, K_2 = \text{constants}$

$\xi(t) = \text{only the trend component of the time series in this case.}$

The form of the logistic curve is shown in Figure 2:2:1.

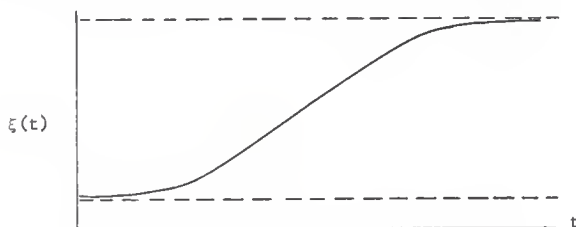


Fig. 2:2:1 Example of logistic curve

The last of the trend lines described by Davis is the familiar moving average.

We may consider  $\xi(t)$  as being composed of two components, trend and periodic.

$$\begin{aligned}\xi(t) &= \text{trend} + \text{periodic} \\ &= T + P\end{aligned}$$

where

$T$  = the trend component

$P$  = the periodic component

and both  $T$  and  $P$  are deterministic functions of time. We wish to describe  $T$  so that we may study  $P$  more closely by performing the following operation:

$$\begin{aligned}x(t) &= \xi(t) + \epsilon(t) \\ &= T + P + \epsilon(t) \\ P + \epsilon(t) &= x(t) - T\end{aligned}$$

As noted earlier, the periodic component is expressed as a series of sines and cosines. These trigonometric functions, however, may be expanded into a series of polynomial terms.

$$\begin{aligned}\sin y &= y - \frac{y^3}{3!} + \frac{y^5}{5} - \dots \\ \cos y &= 1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \dots\end{aligned}$$

where

$$y = \frac{2k\pi t}{p} \text{ as before.}$$

Here we can see the danger of mixing the trend and periodic components,



especially when high degree polynomials are used to describe trend. Unless we are cautious, we may remove part of the periodic variation when we are detrending the time series. This would probably lead to grave forecast errors later on. A plot of the trend line with the actual observations may aid the investigator in avoiding this pitfall.

In order to illustrate the detrending process, let us consider an ideal model of the form

$$x(t) = \xi(t) = a_1 + a_2 t + (a_3 + a_5 t) \sin \frac{2\pi t}{p} \\ + (a_4 + a_6 t) \cos \frac{2\pi t}{p}$$

where  $\epsilon(t)$  is always zero. This time series is shown graphically in Figure 2:2:2.

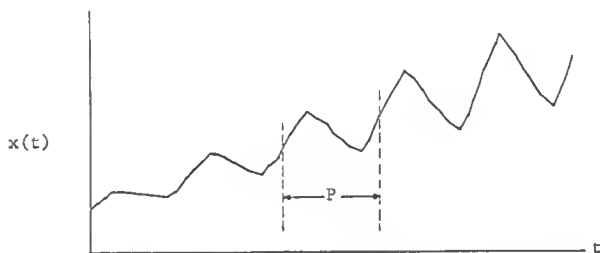


Fig. 2:2:2 Ideal time series with no random variation

In this example we may identify the process components as

$$T = a_1 + a_2 t$$

$$P = (a_3 + a_5 t) \sin \frac{2\pi t}{p} + (a_4 + a_6 t) \cos \frac{2\pi t}{p}$$

The detrended data will have the form of P alone, as shown in Figure 2:2:3.

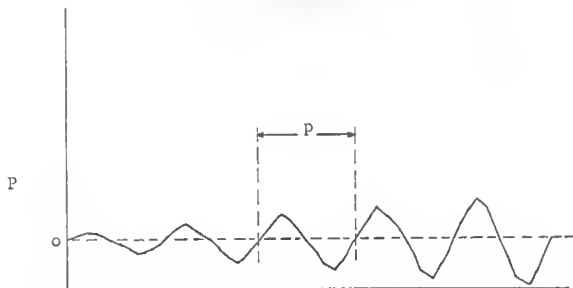


Fig. 2:2:3 Detrended form of ideal time series above

The series now has an expected value of zero. Thus, it is easier to identify the cycles present in the data and any growth in amplitude that may be occurring. The detrended data may then be subjected to autocorrelation and spectral analysis to determine the periodic components present. Polynomial regression (Appendix C) is perhaps the best method to use in detrending the data, and it is the method used in this research.

### 2.3 Autocorrelation Analysis

Though "detrending" aids in the identification of periodic components, it is usually still inadequate. Random noise and the superposition of one harmonic upon another often make it impossible to recognize all of the periodic components present in the process involved. Autocorrelation analysis is one method of manipulating the data to indicate

- (1) Whether or not a process exists in the detrended data
- (2) Whether or not the data contains periodic components
- (3) At least the largest periodic component detectable in the data
- (4) Any lag relationship between data values a certain time period apart in the time series

By applying the "detrending" techniques of the previous section, our data is now in the form

$$x_d(t) = P + \epsilon(t) = x(t) - T$$

where  $x_d(t)$  = the detrended form of the data and  $P$  and  $T$  are the periodic and trend components of the time series, respectively. Since  $x_d(t)$  now has an expected value of zero, by the principal theorem of Fourier we may write

$$x_d(t + k) \approx x_d(t)$$

where  $k$  = the period of the fundamental periodic component of the series. This equation assumes, of course, that a periodic component exists, without any amplitude growth. The fundamental periodic component here is the largest detectable period in the available time series data.

Autocorrelation analysis utilizes a function known as the average lagged product, which may be written as

$$\frac{\sum_{t=k+1}^T x_t x_{t-k}}{T - 1 - k}$$

where  $k$  is the lag, or period of time between the time series values

whose products we are summing. This average lagged product, performed on a sequence of numbers whose expected value is zero, is the autocovariance for lag  $k$ ,  $R_{xx}(k)$ .

$$R_{xx}(k) = \frac{\sum_{t=k+1}^T x_d(t)x_d(t-k)}{T - 1 - k}$$

The variance of a sequence of numbers whose expected value is zero is

$$\sigma_{x_d}^2 = \frac{\sum_{t=1}^T x_d(t)^2}{N}$$

where  $N$  = the total number of observations.

This is just  $R_{xx}(0)$ , the average lagged product for a lag of  $k = 0$ .

$$R_{xx}(0) = \sigma_{x_d}^2$$

Note here that the function  $R_{xx}(k)$  will always be symmetrical, so that

$$R_{xx}(-k) = R_{xx}(k)$$

The normalized form of the autocovariance function  $R_{xx}(k)$  is simply

$$\rho(k) = \frac{R_{xx}(k)}{R_{xx}(0)}$$

Since  $R_{xx}(0) \geq R_{xx}(k)$  for any  $k$ ,

$$-1 \leq \rho(k) \leq +1$$

and

$$\rho(0) = +1.$$

For essentially random noise  $\rho(k)$  will be nearly zero for all  $k \neq 0$ , and  $\rho(0) = 1$ . So if no periodic component exists in the data, a plot of the autocorrelation function against the lag  $k$  would appear as shown in Figure 2:3:1.

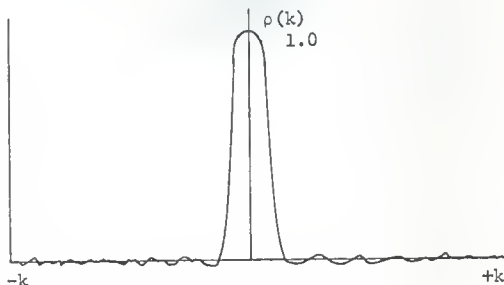


Fig. 2:3:1 Autocorrelation function for time series with no period component

$\rho(k)$  is shown in Figure 2:3:1 as  $\hat{\rho}(k)$  to indicate that this is a computed value subject to random variation. Since  $R_{xx}(k)$  is symmetrical,  $\rho(k)$  will also be symmetrical, and we need only plot  $\hat{\rho}(k)$  for  $k \geq 0$  to study the autocorrelation function. The standard deviation of the computed autocorrelation coefficients,  $\hat{\rho}(k)$ , is  $\sigma = (T-k-1)^{-1/2}$  for a sample of  $T$  observations. Brown warns us that the autocorrelation coefficients should be treated as zero unless they are clearly greater in magnitude than the standard deviation,  $\sigma$ . He also says that the maximum lag should not exceed  $k_{\max} = T/20$ . This restriction may be too tight to follow in practice, since to study a yearly cycle this means that at least twenty years of monthly data would be necessary.

It may be helpful at this point to discuss the possible interpreta-

tions that may be made from the autocorrelation function plot. Consider the plot shown in Figure 2:3:2. The pattern shows a

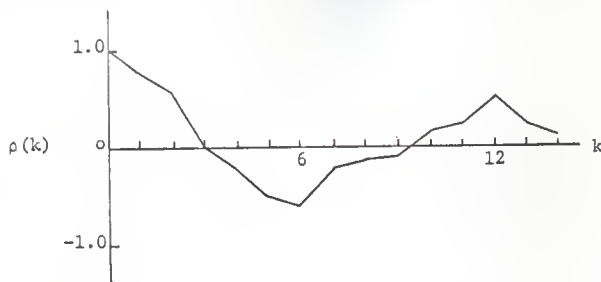


Fig. 2:3:2 Example of an autocorrelation function

trough at six months, a peak at twelve months, and inflection points at three months and nine months. This is indicative of the existence of a twelve month periodic component in the data. The squared nature of the trough and crest are indications that higher harmonics may be present also. Consider now the plot in Figure 2:3:3. Notice again the apparent twelve month periodic.

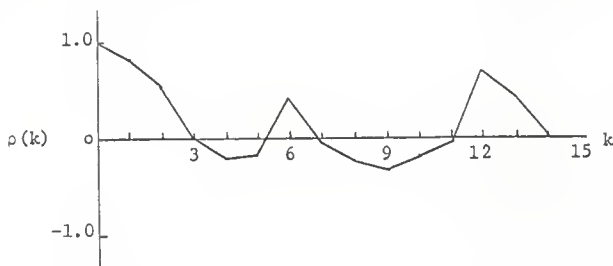


Fig. 2:3:3 Example of an autocorrelation function

However, notice also the smaller peak at six months and the cosine wave-like form of the curve between  $k = 1$  and  $k = 6$ . This indicates that a six month harmonic may also be included in the periodic component of the process.

We may summarize the autocorrelation analysis procedure in the following steps:

- (1) Detrend the time series.
- (2) Compute autocorrelation function,  $\rho(k)$ , for the desired lag values,  $k$ .
- (3) Plot  $k$  versus  $\rho(k)$ .
- (4) Analyze and interpret the plot in (3).

It may be noted here that the more completely we detrend the data, the sharper will be the peaks on the autocorrelation function plot, and hence, the more easily we may recognize the existing periodics. We must, of course, be careful not to remove any periodic variation by mistaking it for trend.

Autocorrelation analysis is especially advantageous when the time series may be described using fitting functions which are previous values of the same time series  $k$  observations in the past. This is referred to as an autoregressive model. For time dependent periodic components another analytical technique has attributes not offered by autocorrelation analysis. This latter technique, harmonic analysis, is discussed in the following section.

## 2.4 Harmonic Analysis

Recall from Section 2.1 that, according to the fundamental theorem

of Fourier series, any reasonable periodic function of time with zero mean can be represented arbitrarily closely (Brown, 2, 396) by taking terms from the infinite Fourier series

$$\xi(t) = \frac{1}{2} A_0 + \sum_{n=1}^{\infty} A_n \cos\left(\frac{n\pi t}{a}\right) + \sum_{n=1}^{\infty} B_n \sin\left(\frac{n\pi t}{a}\right),$$

$$-a \leq t \leq a$$

where

$A_n, B_n$  = constant coefficients

$A_0$  = constant (zero for the detrended data)

$a$  = one half the range of the main, or longest, period.

The coefficients  $A_n$  and  $B_n$  may be computed from the following integrations:

$$A_n = \frac{1}{a} \int_a^a x(t) \cos\left(\frac{n\pi t}{a}\right) dt$$

$$B_n = \frac{1}{a} \int_{-a}^a x(t) \sin\left(\frac{n\pi t}{a}\right) dt$$

The period,  $p$ , and amplitude  $R$ , of the  $n$ th harmonic term may be written

$$p_n = \frac{2a}{n}$$

$$R_n = \sqrt{A_n^2 + B_n^2}$$

The periods represented in this Fourier series make up the Fourier sequence, that is,



$$\frac{2a}{1}, \frac{2a}{2}, \frac{2a}{3}, \dots, \frac{2a}{n}, \dots$$

The advantage of the Fourier sequence is revealed in the following identity:

$$R_1^2 + R_2^2 + R_3^2 + \dots + R_n^2 + \dots = 2\sigma^2$$

where  $\sigma^2$  = the variance of the data series,  $\xi(t)$ .

Thus we see that the ratio  $R_n^2/2\sigma^2$  gives us the fractional variation, or energy, explained by the inclusion of the  $n$ th harmonic term in our time series model. Our goal in describing the periodic components of the data is not, however, to fit a Fourier series to the data. Rather we wish to determine the periods which the time series possesses and, if possible, determine the contribution of each of these periods in describing the time series. It is conceivable that one or more of these periods would not be included in the Fourier sequence. Hence we may define harmonic analysis as the technique of determining the principal harmonic elements of a given set of data.

For our purposes, a better sequence of periods to study is the arithmetic sequence

$$1, 2, 3, 4, \dots, a.$$

This sequence may also be used to describe the time series data, though the energies for these periods are no longer additive. We may now describe the time series in the form

$$\xi(t) = \frac{1}{2} A_0 + \sum_{p=1}^a A_p \cos\left(\frac{2\pi t}{p}\right) + \sum_{p=1}^a B_p \sin\left(\frac{2\pi t}{p}\right)$$

where  $A_p, B_p$  = the coefficients of the harmonic terms for a period of  $p$ .

It is helpful, for reasons to be discussed later in this section, to subject the detrended data to the harmonic analysis rather than the raw time series data. Since the detrended data may be written

$$x_d(t) = P + \varepsilon(t)$$

and the periodic component,  $P$ , has a zero expected value, we may now write

$$P = \sum_{p=1}^a A_p \cos\left(\frac{2\pi t}{p}\right) + \sum_{p=1}^a B_p \sin\left(\frac{2\pi t}{p}\right).$$

We may determine the amplitudes of the periodic terms from the following formulas:

$$A_p = \frac{2}{N'} \sum_{t=1}^{N'} x_d(t) \cos\left(\frac{2\pi t}{p}\right)$$

$$B_p = \frac{2}{N'} \sum_{t=1}^{N'} x_d(t) \sin\left(\frac{2\pi t}{p}\right)$$

$$R_p = \sqrt{A_p^2 + B_p^2}$$

where  $N'$  is chosen equal to the largest integral multiple of  $p$  in the total number of observed data values,  $N$ .

Having determined the values,  $R$ , of the amplitudes of the harmonics in the arithmetic sequence, we wish to compare their magnitudes. This

will give us a relative measure of the contribution of each harmonic to the description of  $x_d(t)$ . The most widely used method of comparison is the "periodogram," introduced by Sir Arthur Schuster (1851-1934) (Davis, 3, 31). A "periodogram" is the graph of  $R_p$  or  $R_p^2$  against  $p$ . Graphs of  $R_p$  versus  $p$  have been employed throughout this research.

This graph will contain peaks above the periods corresponding to harmonics present in the data,  $x_d(t)$ . Generally, the higher the peak, the more significant is that particular harmonic in describing the periodic component of the process. Hence, inclusion of that harmonic in the time series model will account for a greater decrease in error variation than would a harmonic with a lower peak.

Consider the hypothetical "periodogram" shown in Figure 2:4:1. From this graph

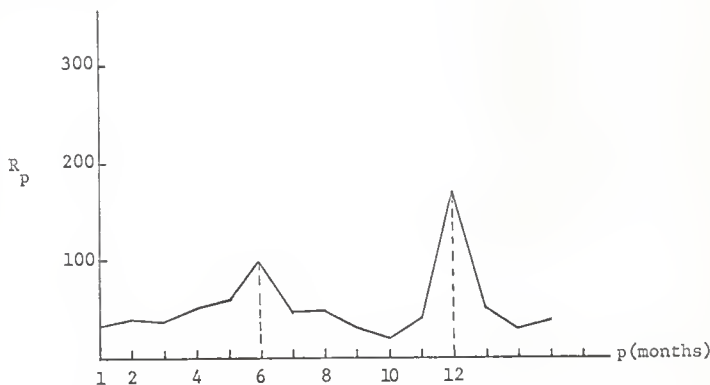


Fig. 2:4:1 Hypothetical harmonic analysis periodogram

we may infer that a periodic component with  $P = 12$  months should probably be included in our time series model. Also, a six-month harmonic may be desired in the model.

We may obtain some idea of the effect of the inclusion of certain periodic terms in the time series model by computing a per cent reduction in error variance for  $x_d(t)$ . For example, let  $\sigma_{xd}^2$  be the variance of the detrended data, and let  $\sigma_{x'd}^2$  be the variance of detrended data minus the periodic terms we wish to include. For the example in Figure 2:4:1 this would be:

$$x_d(t) = p + \varepsilon(t)$$

$$x'_d(t) = x_d(t) - p'$$

where

$$p' = A_6 \cos\left(\frac{2\pi t}{6}\right) + B_6 \sin\left(\frac{2\pi t}{6}\right)$$

$$+ A_{12} \cos\left(\frac{2\pi t}{12}\right) + B_{12} \sin\left(\frac{2\pi t}{12}\right)$$

Now we may write

$$E_{p'} = 100\left[1 - \frac{\sigma_{x'd}^2}{\sigma_{xd}^2}\right]$$

where

$$E_{p'} = \text{the per cent reduction in } \sigma_{xd}^2 \text{ due to } p'.$$

Ideally we would like our "periodograms" to exhibit sharp peaks for the harmonics that should be included in the time series model. Since our data still contains some random variations, sharp peaks are not

always possible. By working only with the detrended data, we lessen the chances of misinterpretation of the graph. Inclusion of the trend component in the harmonic analysis may give erroneous peaks on the graph, or hide harmonics that should have been included in the model. Even after trend is removed from the data, careful interpretation of the "periodogram" is essential. A very broad peak, such as that shown in Figure 2:4:2, suggests the existence of more than one harmonic.

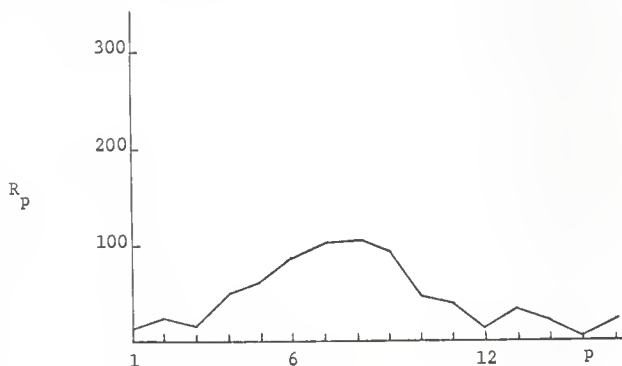


Fig. 2:4:2 Example of periodogram with broad peak

Existing periods of very low frequency may often be handled adequately by the adaptive nature of the smoothing process.

We may see from this discussion that harmonic analysis not only identifies the harmonics existing in the time series, but also indicates their relative contributions in describing the process. Nevertheless, it is still subject to error and misinterpretation. Probably a wise procedure is to perform both the autocorrelation and harmonic analyses,

so that one may serve as a check on the other.

### 3. FORECASTING THE TIME SERIES

#### 3.1 Justification for Exponential Smoothing

We have derived a simplified formulation of the exponential smoothing process in section 1.3 of Chapter 1. In deriving the expressions shown there, we were dealing with the simplest of time series models, the constant process. It was shown that the function  $S_t(x)$  was a good estimate of the value of the process at time  $(t+1)$ . For this model we had only one fitting function,  $f_1(t) = 1$ , where

$$\begin{aligned}\hat{x}(t) &= \hat{a}(t) f_1(t) \\ &= \hat{a}(t) \quad (1) \\ \hat{x}(t) &= \hat{a}(t)\end{aligned}$$

Thus we see that once we have formulated a good model of the true process, we use exponential smoothing to estimate the coefficients,  $\hat{a}_i(t)$ , of the fitting functions in the model.

To bring us up to date, let us briefly review our progress at this point:

- (1) The time series may be described as

$$x(t) = T + P + \epsilon(t)$$

where

$T$  = the trend component

$P$  = the periodic component

$\epsilon(t)$  = the random variation. Both  $T$  and  $P$  are deterministic functions of time.

- (2) By means of polynomial regression we may determine the fitting functions in T.
- (3) We may determine the fitting functions in P through autocorrelation and harmonic analyses.
- (4) Now, through exponential smoothing, we may estimate the coefficients of all the fitting functions in the time series model, and then make forecasts of future values of the time series.

From our previous development of exponential smoothing we arrived at the expressions

$$\begin{aligned}
 S_t(x) &= \alpha x_t + \beta S_{t-1}(x) \\
 &= \alpha x_t + \alpha \beta x_{t-1} + \alpha \beta^2 x_{t-2} + \dots \\
 &\quad + \alpha \beta^n x_{t-n} + \dots + \beta^t x_0
 \end{aligned}$$

or

$$S_t(x) = \alpha \sum_{n=0}^{t-1} \beta^n x_{t-n} + \beta^t x_0$$

where

$S_t(x)$  = the exponentially smoothed value of  $x(t)$  up to time  $t$ .

$\alpha$  = smoothing constant

$\beta = 1 - \alpha$  = discount factor.

Note in this expression once again the advantages offered by exponential smoothing:

- (1) Data is weighted to reflect decreasing importance with increasing age.



- (2) All past data is included in the statistic  $S_t(x)$ , where  $t$  is taken to be the present time.
- (3) Storage of past data requires only the retention of the current value of  $S_t(x)$ .

As a criterion for justification of the exponential smoothing process, we shall say that, if the expected value of  $S(x)$  equals the expected value of  $x$ , the procedure is justified. Thus,  $S_t(x)$  is a valid method for estimating  $a_1(t)$ , the constant process fitting function coefficient. So, we may say

$$E[S(x)] = E[\alpha \sum_{k=0}^{\infty} \beta^k x_{t-k}] + E[\beta \alpha x_0] \quad .$$

The last term shown,  $E[\beta \alpha x_0]$ , is zero, since  $\beta$  is a number less than one.

Now,

$$\begin{aligned} E[S(x)] &= E[\alpha \sum_{k=0}^{\infty} \beta^k x_{t-k}] \\ &= \alpha \sum_{k=0}^{\infty} \beta^k E[x_{t-k}] \\ &= E[x] \alpha \sum_{k=0}^{\infty} \beta^k \end{aligned}$$

$\sum_{k=0}^{\infty} \beta^k$  is the summation of a converging geometric series. It has the value  $\frac{1}{1-\beta}$  for an infinite number of terms. So,

$$E[S(x)] = E[x] \alpha \frac{1}{1-\beta}$$

$$= E[x] \frac{\alpha}{\alpha}$$

$$E[S(x)] = E[x]$$

and the procedure is justified. In addition to the advantages cited above, the response rate of exponential smoothing is quite flexible and the computations simple, as will be shown in later discussions. Note also another characteristic of the smoothing process. Let us rearrange the previous expression for  $S_t(x)$  as shown below.

$$\begin{aligned} S_t(x) &= \alpha x_t + \beta S_{t-1}(x) \\ &= \alpha x_t + (1-\alpha)S_{t-1}(x) \\ &= S_{t-1}(x) + \alpha(x_t - S_{t-1}(x)) \end{aligned}$$

The second expression on the right is merely the current observation,  $x_t$ , minus the estimate of  $x_t$  at time  $t-1$ ,  $S_{t-1}(x)$ , all multiplied by the smoothing constant. So, this is a measure of the error of the previous forecast. Hence, the smoothing procedure updates the coefficient estimates to adjust for the previous estimation error. The accuracy of the procedure may be inferred from the fact that the function  $S(x)$  minimizes the weighted sum of squared residuals (Appendix B).

### 3.2. Discounted Multiple Regression

Until now, we have been working with a time series that may be represented by a constant process. We now wish to consider a more practical example, where the process is composed of an arbitrary set of functions. For these fitting functions we determine values of the coefficients that

minimize the weighted sum of square residuals, as before.

Consider again the general process mentioned in section 1.3.

$$\begin{aligned}\xi(t) &= a_1(t)f_1(t) + a_2(t)f_2(t) + \dots \\ &+ a_n(t)f_n(t) \\ &= \sum_{i=1}^n a_i(t)f_i(t)\end{aligned}$$

where

$$\begin{aligned}\xi(t) &= \text{the process, or the time series minus random noise} \\ f_i(t) &= \text{the } i^{\text{th}} \text{ fitting function evaluated at time } t \\ a_i(t) &= \text{the } i^{\text{th}} \text{ fitting function coefficient estimated at} \\ &\quad \text{time } t.\end{aligned}$$

The fitting functions may be any deterministic functions of time, as long as their values are known exactly for the present and for the time in the future for which a forecast is desired. This restriction allows the inclusion of:

- (1) Simple mathematical functions of time.
- (2) Empirical functions that satisfy the restriction.
- (3) Previous observations of the dependent time series.

Having determined the fitting functions to be used, we must estimate their coefficients so that the desired  $x$ -value may be obtained from

$$\begin{aligned}\hat{x}(t) &= \sum_{i=1}^n \hat{a}_i(t)f_i(t) \\ &= \hat{a}'(t)f(t)\end{aligned}$$

where  $a(t)$  and  $f(t)$  are  $n$ -component column vectors evaluated at time  $t$ . Recall that the process  $\xi(t)$  may change slightly over a long period of time without injecting significant error into the smoothing process. Over the short run, however, we may treat the coefficients as constants and obtain a least squares estimate of their values from the exponentially discounted data.

The forecast of  $x$  for  $\tau$  periods into the future, considering the present time as  $t$ , would then be

$$\begin{aligned}\hat{x}(t+\tau) &= \hat{a}_1(t)f_1(t+\tau) + \hat{a}_2(t)f_2(t+\tau) \\ &+ \dots + \hat{a}_n(t)f_n(t+\tau) \\ &= \sum_{i=1}^n \hat{a}_i(t)f_i(t+\tau) \\ &= \hat{a}'(t)f(t+\tau)\end{aligned}$$

where

$a_i(t)$  = a least squares estimate of  $a_i(t)$  based on the weighted data,  $x$ , from  $t = 0$  up to the present time,  $t$ .

$f_i(t+\tau)$  = the  $i^{\text{th}}$  fitting function evaluated at the future time,  $t+\tau$  for which the forecast is desired.

We are assuming that the estimates,  $a_i(t)$ , based on all past data, are the best that we can obtain. The least squares estimate of  $a_i(t)$  is similar to a multiple regression solution; however, in this case, we are minimizing the discounted sum of squared residuals

$$\sum_{j=1}^T \beta^j [x(T-j) - \sum_{i=1}^n \hat{a}_i(T) f_i(T-j)]^2$$

where

$T$  = present time.

Hence, we call this procedure "discounted multiple regression."

The ensuing discussion will show the technique for finding these least square estimates of the coefficients. First, the following definitions should aid the understanding of the procedures that follow.

$$x = [x_1, x_2, x_3, \dots, x_T]$$

= all past observed time series values up to the present time  $T$ .

=  $1 \times T$  row vector.

$$f(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} = \text{an } n \times 1 \text{ column vector of the } n \text{ fitting functions evaluated at time } t.$$

$$a(t) = \begin{bmatrix} a_1(t) \\ a_2(t) \\ \vdots \\ a_n(t) \end{bmatrix} = \text{an } n \times 1 \text{ column vector of fitting coefficients for time } t.$$

$$\mathcal{F}(t) = \begin{bmatrix} f_1(1) & f_1(2) & \dots & f_1(T) \\ f_2(1) & f_2(2) & \dots & f_2(T) \\ \vdots & \vdots & & \vdots \\ f_n(1) & f_n(2) & \dots & f_n(T) \end{bmatrix}$$

= an  $n \times T$  matrix with elements  $f_i(t)$  in the  $i^{\text{th}}$  row and the  $t^{\text{th}}$  column.

$\hat{a}(t)$  is our least squares estimate of  $a(t)$  and is of the same vector form as that shown above for  $a(t)$ . Since we have defined the time series estimate as

$$\hat{x}(t) = \hat{a}'(t) f(t)$$

we may denote the corresponding residual as

$$\begin{aligned} e(t) &= x(t) - \hat{x}(t) \\ &= x(t) - \hat{a}'(t)f(t) \end{aligned}$$

Now the discounted sum of squared residuals would be

$$\sum_{t=1}^T w^2(t) e^2(t)$$

where

$w(t)$  is the  $t^{\text{th}}$  diagonal element of a  $T \times T$  matrix  $W(t)$ .

$$W(t) = \begin{bmatrix} w(1) & & & & 0 \\ & w(2) & & & \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & w(T) \end{bmatrix}$$

where

$$w^2(T-j) = \beta^j, \quad j = 0, 1, 2, \dots, T-1.$$

Now, if we define a row vector of residuals,  $e(t)$ ,

$$e(t) = x(t) - a'(t)\mathcal{F}(t),$$

then the discounted sum of squared residuals is

$$\begin{aligned} S_a &= [e(t)W(t)] [e(t)W(t)]' \\ &= [x(t)W(t) - a'(t)\mathcal{F}(t)W(t)] [x(t)W(t) - a'(t)\mathcal{F}(t)W(t)]' \end{aligned}$$

for a particular set of coefficients  $a(t)$ . To minimize this sum, we need to select  $a(t)$  such that

$$\frac{\partial S_a}{\partial a(t)} = 0.$$

This expression gives us  $n$  simultaneous linear equations which can be solved for the  $n$  coefficients  $a_1(t)$ . As shown in Appendix B, the solution to this expression is

$$a'(t) = x(t)W(t)W'(t)\mathcal{F}'(t)F^{-1}(t)$$

where

$$\begin{aligned} F(t) &= \mathcal{F}(t)W(t)W'(t)\mathcal{F}'(t) \\ &= \text{an } n \times n \text{ symmetrical matrix} \end{aligned}$$

As long as  $T \gg n$ ,  $F(t)$  will have an inverse. If we define a data vector  $g(t)$  as

$$g'(t) = x(t)W(t)W'(t)\mathcal{F}'(t)$$

then

$$a'(t) = g'(t)F^{-1}(t).$$

Note that  $F(t)$  does not depend on the data. It does, however, depend on time, and must be updated at each sampling interval. Looking at one element in the matrix  $F(t)$ ,

$$F_{ik}(T) = \sum_{j=0}^{T-1} \beta^j f_i(T-j) f_k(T-j)$$

where

$T$  = present time, we see that it may be updated by

$$F_{ik}(T+1) = f_i(T+1) f_k(T+1) +$$

$$\sum_{j=0}^{T-1} \beta^j f_i(T-j) f_k(T-j)$$

So, generally,  $F(t)f'(t) + \beta F(t-1)$

Since  $F(t) \neq F(t-1)$ , then  $F^{-1}(t) \neq F^{-1}(t-1)$  and a new  $F^{-1}$  must be computed at each sampling interval. Reference to an expression for the  $i^{\text{th}}$  data vector component at time  $T$ ,

$$g_i(T) = \sum_{j=0}^{T-1} \beta^j x(T-j) f_i(T-j),$$

indicates that  $g(t)$  may be updated at each sampling interval by

$$g(t) = x(t)f(t) + \beta g(t-1)$$

where

$t$  = the new present time value

$t-1$  = the previous time period.

So, having updated  $F(t)$  and  $g(t)$  and after a sufficient amount of data



has been gathered to allow the existence of  $F^{-1}(t)$ , we may estimate  $a$  by

$$\hat{a}'(t) = g(t)F^{-1}(t)$$

and make a forecast of future observations by

$$\hat{x}(t+\tau) = \hat{a}'(t)f(t+\tau).$$

### 3.3 Construction of the Transition Matrix

We can simplify the computations of the previous section considerably if we place a restriction on the class of fitting functions; that is, the vector of fitting function values at time  $t+1$  must be a linear combination of the values of these functions at time  $t$ . Specifically, we can define an  $n \times n$  matrix  $L$  with components  $L_{ij}$  such that

$$f_1(t+1) = L_{11}f_1(t) + L_{12}f_2(t) + \dots + L_{1n}f_n(t)$$

$$f_2(t+1) = L_{21}f_1(t) + L_{22}f_2(t) + \dots + L_{2n}f_n(t)$$

$$\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

$$f_n(t+1) = L_{n1}f_1(t) + L_{n2}f_2(t) + \dots + L_{nn}f_n(t)$$

or

$$f(t+1) = L f(t)$$

The transition matrix,  $L$ , must have an inverse, as will be shown later.

The only fitting functions which satisfy this restriction are

- (1) Polynomials
- (2) Exponentials
- (3) Sinusoids

Once we know  $f(0)$  and  $L$ , we may find  $f(t)$  for any  $t$  since,

$$\begin{aligned} f(t) &= Lf(t-1) \\ &= L[Lf(t-2)] \\ &= L^t f(0) \end{aligned}$$

The matrix  $L$  may be constructed, by inspection of the fitting functions desired, according to rules regarding each allowable class of fitting functions.

For a polynomial of degree  $n$ ,  $L$  is an  $(n+1) \times (n+1)$  matrix with ones on the diagonal, ones in the first element to the left of the diagonal and zeros everywhere else.

$$L = \begin{bmatrix} 1 & & & & 0 \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & \ddots \\ & & & & & \ddots \\ & & & & & & 1 \\ 0 & & & & & & & 1 \end{bmatrix}$$

where  $L$  has  $n+1$  rows and  $n+1$  columns. The mathematics involved in updating  $L$  requires that polynomial fitting functions now be of the form

$$f_n(t) = \frac{1}{n!} t(t-1)(t-2) \dots (t-n+1).$$

As mentioned earlier, for a trigonometric fitting function the sine and cosine term of each harmonic must both be included to account for any phase angle shift. There are two basic forms which the corresponding transition matrix may take. For example, suppose we want  $L$  for the

trigonometric pair

$$f_1(t) = \sin \omega t$$

$$f_2(t) = \cos \omega t$$

L would be the 2 x 2 matrix

$$L = \begin{bmatrix} \cos \omega & \sin \omega \\ -\sin \omega & \cos \omega \end{bmatrix}$$

If we want to include this same harmonic, but with linearly growing amplitude as well, the fitting functions would be:

$$f_1(t) = \sin \omega t$$

$$f_2(t) = \cos \omega t$$

$$f_3(t) = t \sin \omega t$$

$$f_4(t) = t \cos \omega t$$

and L would be the 4 x 4 matrix

$$L = \begin{bmatrix} \cos \omega t & \sin \omega t & 0 & 0 \\ -\sin \omega t & \cos \omega t & 0 & 0 \\ \cos \omega t & \sin \omega t & \cos \omega t & \sin \omega t \\ -\sin \omega t & \cos \omega t & -\sin \omega t & \cos \omega t \end{bmatrix}$$

The elements of L always take the same position relative to the diagonal of L, no matter what the size of L. So, for models which include many fitting functions or combinations of fitting functions, the transition matrix may be constructed by stacking the simple matrices shown along the diagonal of L. All other elements will be zeros. For the construction of exponential transition matrices, refer to Brown (2, 166).

### 3.4 Adaptive Smoothing

Observing the restriction placed on  $f(t)$  in the previous section and moving the origin of time to the most recent observation, we may now write

$$\hat{x}(t+\tau) = f'(\tau)\hat{a}(t)$$

This shift in time is reflected only in the fitting functions, so that their values depend only upon the increment of time into the future for which a forecast is desired.  $\hat{x}$  still shows the actual time to which the forecast applies, and  $\hat{a}$  shows the total data upon which it is based, all data up to time  $t$ . So, if we shall always be forecasting one time period into the future,  $\tau = 1$ , we need to compute  $f(\tau) = f(1)$  only once.

The sum of squares to be minimized now is

$$\sum_{j=0}^T \beta^j [x(T-j) - f'(-j)\hat{a}(T)]^2.$$

Differentiating the weighted sum of squared errors with respect to  $a(t)$  and setting the result equal to zero as before, we obtain

$$\hat{a}(t) = F^{-1}(t)g(t),$$

where

$$F_{ik}(t) = \sum_{j=0}^t \beta^j f_1(-j)f_k(-j)$$

and

$$g_i(t) = \sum_{j=0}^t \beta^j f_1(-j)x(t-j)$$

Again this assumes that sufficient observations have been made for  $F(t)$

to have an inverse, that is, at least  $n$  observations where  $F(t)$  is an  $n \times n$  matrix.

At each sampling interval, we must again update  $g(t)$  and  $F(t)$ . Since both are functions of  $f(t)$ , and  $f(t)$  has a different time origin than in discounted multiple regression, the updating procedure will not be the same as before. The data vector may be expressed

$$g(t) = x(t)f(0) + \sum_{j=1}^t \beta^j f(-j)x(t-j)$$

Since

$$f(-j+1) = Lf(-j), \text{ or } f(-j) = L^{-1}f(-j+1),$$

$$g(t) = x(t)f(0) + \sum_{j=1}^t \beta^j L^{-1}f(-j+1)x(t-j).$$

If we change the index of summation to  $k = j - 1$ , the data vector recursive relationship becomes

$$g(t) = x(t)f(0) + \beta L^{-1}g(t-1)$$

where the coefficients of  $x(t)$  and  $g(t-1)$  no longer depend on time and thus become program constants. The matrix of fitting functions may also be computed recursively by

$$\begin{aligned} F(t) &= \sum_{j=0}^t \beta^j f(-j)f'(-j) \\ &= F(t-1) + \beta^t f(-t)f'(-t) \end{aligned}$$

When the fitting functions are restricted to trigonometric functions or polynomials,  $\beta^t$  approaches zero faster than  $f(-t)$  grows, since  $\beta < 1$ .

Thus,  $F$  will reach a steady state value after a sufficient number of sampling intervals, and does not require any further updating. Once steady state has been reached we may write

$$F(t) = F(t-1) = F(\infty)$$

If one of the fitting functions is of the form  $f(t) = e^{-at}$ , steady state may be reached only if  $\beta e^{-2a}$ . In the case of trigonometric and polynomial models, the fastest growing function will be the highest degree polynomial term present. If this term is  $t^n$ , then

$$T = \frac{7 + 5.1n}{(1-\beta)^{0.95}}$$

= the number of sampling intervals necessary to reach steady state (Brown, 2, 170),

where the convergence criterion is that the largest quantity added to any element of  $F$  is less than  $10^{-6}$  times the previous value of that element.

Having reached steady state, we may estimate the coefficients by

$$\hat{a}(t) = F^{-1}g(t)$$

where  $F$  is no longer time dependent. The forecast would then be computed from

$$\begin{aligned}\hat{x}(t+\tau) &= \hat{a}'(t)f(\tau) \\ &= [F^{-1}g(t)]'f(\tau) \\ &= g'(t)F^{-1}f(\tau) \\ &= g'(t) \ c(\tau)\end{aligned}$$

where  $c(\tau)$  is a column vector dependent on the values of the fitting functions at time  $\tau$ , but not on absolute time. Note that only one matrix inverse will be necessary as long as  $\beta$ ,  $\tau$ , and the fitting functions are not changed, so the computations have been greatly simplified. When additional data is available, we must merely update  $g(t)$  and compute  $\hat{x}(t+\tau)$ .

### 3.5 Derivation of General Exponential Smoothing

We have said that  $F$  will converge to a steady state and may be handled as a program constant if: (1) successive fitting function values may be generated by a fixed transition matrix  $L$ , that is,  $f(t+1) = L f(t)$ ; (2) the origin of time for the fitting functions is taken as the most recent observation. Furthermore, the data vector,  $g(t)$ , may be defined recursively from the previous values of its  $n$  components,  $g(t-1)$ , and the new observation. We wish now to minimize the computational effort required in obtaining successive estimates of the coefficients,  $\hat{a}(t)$ .

If there were no random noise in the data, and we had developed a model to describe the process exactly, we could recursively evaluate  $\hat{a}(t)$  from  $\hat{a}'(t-1)$ . In this case  $\hat{x}(t) = x(t)$ , so

$$\begin{aligned}\hat{x}(t) &= \hat{a}'(t-1)f(1) = x(t) \\ x(t) &= \hat{a}'(t-1)f(1) \\ &= \hat{a}'(t-1)Lf(0)\end{aligned}$$

Note also that

$$x(t) = \hat{a}'(t) f(0)$$

Equating the last two equations,

$$a'(t)f(o) = a'(t-1)Lf(0)$$

So,

$a'(t) = a'(t-1)L$  and we may obtain the value of the next observation, that at time  $t$ , from

$$\begin{aligned}\hat{x}(t) &= \hat{a}'(t-1) f(1) \\ &= \hat{a}'(t-1)L_f(0) \\ &= a'(t)f(0)\end{aligned}$$

We cannot expect the data to conform exactly to the model of the process, however. Each observation will differ from the corresponding forecast by an amount  $e$ ,

$$e(t) = x(t) - \hat{x}(t).$$

If this fitting error can be taken into account in successive estimates of the coefficients, we may still obtain a recursive relationship for  $a$ . In order to account for these fitting errors, we shall define an  $n$  component smoothing vector,  $h$ , which depends only on  $\beta$  and the fitting functions.

$$h = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} = F^{-1}f(o) = \text{the smoothing vector.}$$

$h$  may also be treated as a program constant for a particular  $\beta$  and  $a$



particular model.

As in the previous discussions, the minimum discounted squared residual is satisfied when

$$a(t) = F^{-1}g(t)$$

where  $F^{-1}$  is used in its steady state. This may also be written

$$g(t) = Fa(t).$$

If we substitute this expression into the recursive relation for  $g(t)$  obtained in the previous section, we have

$$Fa(t) = x(t)f(o) + \beta L^{-1}Fa(t-1)$$

or

$$a(t) = x(t)F^{-1}f(o) + \beta F^{-1}L^{-1}Fa(t-1)$$

Note that the coefficients of both  $x(t)$  and  $a(t-1)$  are time independent, and that  $F^{-1}f(o)$  is  $h$ , the smoothing vector. By manipulating the equation that defines  $F$ , it can be shown that

$$F^{-1}L^{-1}F = L' - hf'(1)$$

= the coefficient of  $a(t-1)$  in the expression for  $a(t)$  above.

Now we may write

$$a(t) = hx(t) + L'a(t-1) - hf'(1)a(t-1)$$

But

$$f'(1)a(t-1) = \hat{x}(t),$$

so,

$$\begin{aligned} a(t) &= hx(t) + L'a(t-1) - \hat{hx}(t) \\ &= L'a(t-1) + h[x(t) - \hat{x}(t)] \\ a(t) &= L'a(t-1) + he(t). \end{aligned}$$

Thus we have a recursive relationship for  $a$  that accounts for previous fitting errors.

### 3.6 Summary of General Exponential Smoothing

The relationships used are valid only for trigonometric and polynomial models, and exponential models satisfying the restriction mentioned earlier (Section 3.4). The general forecast equation is

$$\hat{x}(t+\tau) = \hat{a}'(t)f(\tau).$$

The fitting function coefficients may be obtained recursively by

$$a(t) = L'a(t-1) + he(t)$$

where

$$h = F^{-1}f(o).$$

The fitting function vector may be obtained recursively if necessary by

$$f(t) = Lf(t-1).$$

Appropriate choices of  $\beta$  and the initial conditions,  $a(o)$ , are important to all the methods discussed. These problems will be discussed in the following chapter.

#### 4. CHOOSING THE SMOOTHING CONSTANT AND THE INITIAL COEFFICIENT VECTOR

##### 4.1 Effect of Data on Discount Factor Choice

There are essentially three ways by which we may control the efficiency with which our forecasts follow the actual data:

- (1) Choice of the model
- (2) Choice of the discount factor
- (3) Choice of the initial conditions.

The first control was discussed at length in an earlier chapter. At this time we wish to develop some guidelines to follow in choosing  $\beta$ , our second control over the model efficiency. The literature suggests some general rules that should be followed in making a  $\beta$  choice, however, prior research in this area appears to be inadequate. Consequently, much of the research presented in this thesis is based on the effect of  $\beta$  in describing the true process.

Initially, our choice of a discount factor should be based on the nature of the data and our expectations of future data. If the deviations from our model of the process seem to be mainly random fluctuations, then a large  $\beta$  value will probably be appropriate. This will put more emphasis on older data and, hence, tend to smooth out the effect of the fluctuations of the actual data on our forecasts. If, on the other hand, these fluctuations are large in magnitude and do not seem to imply a revision of the model, a small  $\beta$  will put greater weight on more recent data, and allow the model to react more quickly to these fluctuations. In terms of the forecasting system response to the actual data, we may say:

- (1) For large values of  $\beta$ , system response rate is low;
- (2) For small values of  $\beta$ , system response rate is high.

If the investigator expects the appearance of a new pattern in the future values of the actual data, i.e., if he expects relatively large changes in the fitting function coefficient values, again a small  $\beta$  should be used to allow quick response.

We shall defer discussion of actual numerical values of  $\beta$  until a later section. We wish first to develop some general guidelines on the range of  $\beta$ , as dictated by the data. As in numerous other problems we are faced with two conflicting criteria by which we judge the efficiency of our  $\beta$  choice, that is,

- (1) Ability to respond to fluctuations in the true process
- (2) Ability to smooth out random fluctuations in the data.

A smaller  $\beta$  gives quicker response to fluctuations in the true process, but also gives more drastic response to the random fluctuations.

#### 4.2 Average Age of the Data

Brown (2, 116) warns that if you must use a  $\beta$  smaller than 0.7, a larger than 0.3, you may actually need a different model, specifically, a constant model. He feels that with  $\beta < 0.7$  the response rate is so high that a constant model may be quite adequate. It may be that an undetected periodic in the data forces the use of a high  $\beta$ . Again we may use a more reasonable value of the discount factor if we alter the model to include the necessary periodic fitting functions. We may infer from these observations that it would be wise to know what the weighting

factors will be for specific  $\beta$  values. Brown (2, 107) has presented "average age of the data" to aid the investigator in grasping the meaning of a particular  $\beta$  value. We shall apply average age only to a constant model for now. In Section 4.4 we will see how this term applies to more complicated models. We may define the age of the data as being one for the present observation, two for the previous observation, three for the one before that, and so on. So, the observation  $N$  sampling intervals age will have age  $N$ . If we multiply the age of each data value used in forecasting by the weight received by that data, we have a figure representing the average age of all the data used. Hence, for a moving average of period  $N$ ,

$$\bar{k} = \frac{1}{N} (0) + \frac{1}{N} (1) + \frac{1}{N} (2) + \dots + \frac{1}{N} (N-1)$$

$$\bar{k} = \frac{N-1}{2}$$

where

$\bar{k}$  = average age of the  $N$  data points

$\frac{1}{N}$  = the weight received by each data point

Applying this to exponential smoothing, we see that the weight given to data  $n$  sampling intervals old is  $\alpha\beta^n$ . Hence, the average age of the data would be

$$\begin{aligned}\bar{k} &= \alpha(0) + \alpha\beta(1) + \alpha\beta^2(2) + \dots + \alpha\beta^T(T) \\ &= \alpha \sum_{n=0}^T \beta n\end{aligned}$$

where

$T$  = the total number of observations used in making the forecast.

If  $T$  is large, that is, if we have much data from which to make a forecast,  $\beta^T$  will be very small. We may then approximate  $\bar{k}$  by letting  $T$  approach infinity without introducing significant error into the calculations (Brown, 2, 135). So,

$$\bar{k} = \alpha \sum_{n=0}^{\infty} n \beta^n$$

$$= \alpha \cdot \frac{\beta}{(1-\beta)^2}$$

$$= \frac{\alpha \beta}{\alpha^2}$$

$$\bar{k} = \beta / \alpha .$$

We might want to define an exponential smoothing process as being equivalent to a moving average with the same average age of the data. That is,

$$\frac{\beta}{\alpha} = \frac{N-1}{2}$$

or

$$\beta = \frac{N-1}{N+1}$$

$$N = \frac{1+\beta}{1-\beta} \quad \frac{1+\beta}{\alpha}$$

This, of course, would only apply to exponential smoothing with a constant model of the process.

### 4.3 System Response to Standard Input Signals

We may gain some feeling for the effect of  $\beta$  on the system by analyzing the system response to various standard input signals. The inferences to be drawn are only of a general nature, since the constant process model assumed here would often be inadequate.

Before we study system response, it is necessary to establish a few definitions. Any discrete, linear, time-invariant system can be completely described by its impulse response (Brown, 2, 112). All of our discussions have been based on the assumption of discrete data, that is, data taken only at specific and equal time intervals. Linear implies the following system characteristic regarding input to and output from the system:

If input  $x_1(t)$  produces output  $y_1(t)$ , and input  $x_2(t)$  produces output  $y_2(t)$ , then input  $ax_1(t) + bx_2(t)$  will produce output  $ay_1(t) + by_2(t)$ , where  $a$  and  $b$  are constants.

The system is time-invariant if a particular input signal will cause the same system output at any time,  $t$ .

A unit impulse signal may be as shown in Figure 4:3:1.

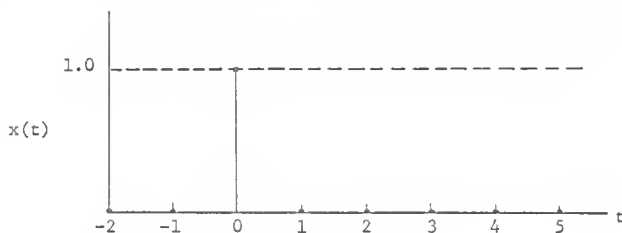


Fig. 4:3:1 Unit impulse function

It may be written as

$$\begin{aligned} x(t) &= 1 & t &= 0 \\ x(t) &= 0 & t &\neq 0 \end{aligned}$$

where

$$\begin{aligned} x(t) &= \text{the input signal} \\ t &= \text{time.} \end{aligned}$$

Since we have not defined our exponential smoothing system for negative time, we may assume a time range of  $0 \leq t \leq \infty$ . Let  $h(t)$  be the unit impulse response of the system  $t$  periods after the unit impulse input. The response of a linear system,  $y(t)$ , to an arbitrary signal  $x(t)$  may be expressed as

$$y(t) = \sum_{n=-\infty}^{\infty} x(n)h(t-n)$$

Since our system is undefined for  $t < 0$ ,

$$y(t) = \sum_{n=0}^{\infty} x(n)h(t-n)$$

$$= x(t) * h(t)$$

where the symbol  $*$  implies a convolution of the sequence  $[x(t)]$  with the sequence  $[h(t)]$ . Note from the previous equation that for a linear function of past data, as in exponential smoothing with a constant model,

$$h(t-n) = \alpha w(t-n) = \alpha \beta^n.$$

Hence,



$$y(t) = \alpha \sum_{k=0}^{\infty} x(n)w(t-n).$$

So, having found an expression for the unit impulse response for exponential smoothing, we may find the system response to an arbitrary input  $x(t)$ .

First, let us simplify the computations by transforming the convolution equation above. The Z-transform (Appendix D) is particularly convenient for this analysis, since the convolution becomes a multiplication in the Z-domain. The Z - transform,  $F(Z)$ , may be written

$$F(Z) = \sum_{n=-\infty}^{\infty} f(n)Z^n$$

where

$f(n)$  = the function to be transformed

$F(Z)$  = the Z - transform of  $f(n)$ .

Note that the previous equation for  $y(t)$  may now be expressed as a multiplication,

$$Y(Z) = X(Z) \cdot H(Z)$$

where

$Y(Z)$  = the Z- transform of  $y(t)$

$X(Z)$  = the Z- transform of  $x(t)$

$H(Z)$  = the Z- transform of  $h(t)$ .

Once we have found  $Y(Z)$ , we may apply an inverse transform to obtain  $y(t)$ . The system response,  $y(t)$ , is determined below for several standard inputs.

For a unit impulse input at time zero, as described above, we may write

$$\begin{aligned}
 X(Z) &= \sum_{n=0}^{\infty} x(n)Z^n \\
 &= 1 \cdot Z^0 + 0 + 0 + \dots \\
 &= 1 \\
 H(Z) &= \sum_{n=0}^{\infty} \alpha \beta^n Z^n = \alpha \sum_{n=0}^{\infty} \beta^n Z^n \\
 &= \alpha \left( \frac{1}{1-\beta Z} \right)
 \end{aligned}$$

Now,

$$\begin{aligned}
 Y(Z) &= X(Z) \cdot H(Z) \\
 &= 1 \cdot \alpha \left( \frac{1}{1-\beta Z} \right) \\
 Y(Z) &= 1 \cdot \sum_{n=0}^{\infty} \alpha \beta^n Z^n
 \end{aligned}$$

So,

$$y(t) = \alpha \beta^t.$$

Hence,  $y(t)$  for a unit impulse input is precisely the weight given to the data. Note that the system response is dependent upon our choice of  $\beta$ , and hence of  $\alpha$ . This effect may be seen in Figure 4:3:2.

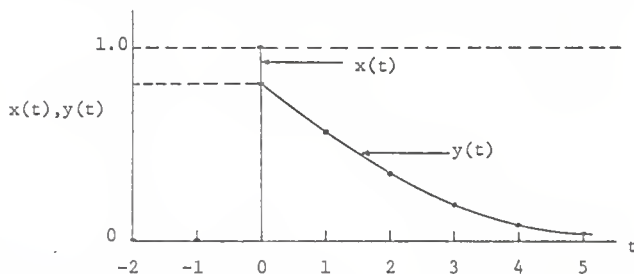


Fig. 4:3:2 System response for unit impulse input to a constant process

For lower values of  $\beta$  the system will adjust back to its original state at a faster rate, since  $w(t-n) = \beta^n$  will decrease at a faster rate than it would for a higher  $\beta$  value.

Suppose the input signal is a unit step, as shown in Figure 4:3:3.

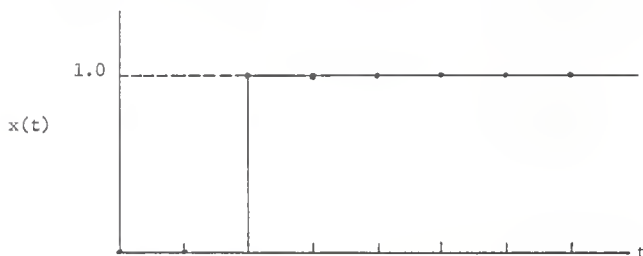


Fig. 4:3:3 Unit step function

This may be written

$$\begin{aligned} x(t) &= 0 & t < 0 \\ x(t) &= 1 & t \geq 0. \end{aligned}$$

Determining the Z -transforms

$$X(Z) = \sum_{n=0}^{\infty} (1) \cdot Z^n = \frac{1}{1-Z}$$

$$H(Z) = \frac{\alpha}{1-\beta Z},$$

We now have

$$Y(Z) = \left( \frac{1}{1-Z} \right) \cdot \left( \frac{\alpha}{1-\beta Z} \right).$$

Now,

$$\begin{aligned}
 \sum_{n=0}^{\infty} y(n)Z^n &= \left( \sum_{n=0}^{\infty} Z^n \right) \cdot \left( \sum_{n=0}^{\infty} \alpha \beta^n Z^n \right) \\
 &= \sum_{n=0}^{\infty} \alpha Z^n \sum_{k=0}^{\infty} \beta^k \\
 &= \sum_{n=0}^{\infty} \frac{\alpha(1-\beta^{n+1})}{1-\beta} Z^n
 \end{aligned}$$

and

$$y(t) = 1 - \beta^{t+1}.$$

Note again the effect of  $\beta$  on  $y(t)$ . For lower  $\beta$  values the system will adjust to the input at a much faster rate. This effect is shown in Figure 4:3:4.

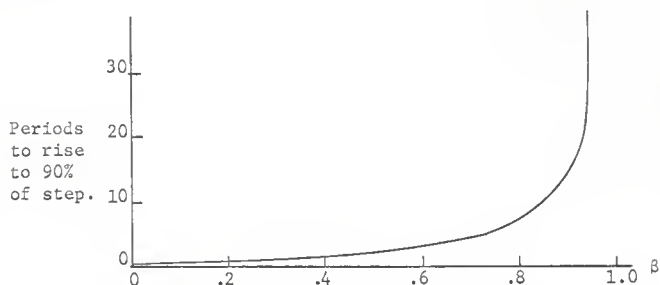


Fig. 4:3:4 Effect of discount factor on system response to a unit step input

Since exponential smoothing gives some weight to all past data, it will never actually reach the step, but approaches it asymptotically as shown in Figure 4:3:5.

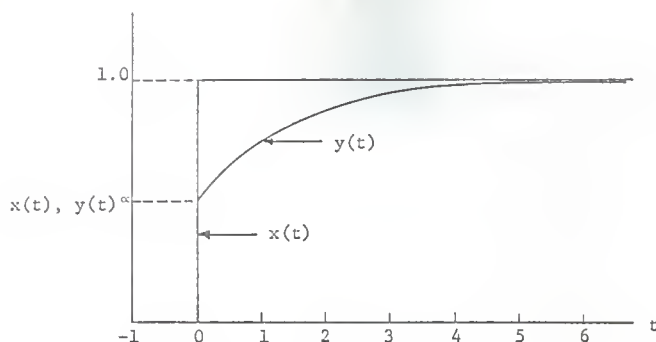


Fig. 4:3:5 System response to unit step input

Through the same procedure we will find that for a ramp input signal

$$x(t) = t, \quad t \geq 0,$$

$$y(t) = t - \frac{\beta}{\alpha} (1 - \beta^t).$$

Again we see the dependence of system response on our choice of the discount factor.

Summarizing our observations of system response characteristics, let us re-iterate once more the general rule regarding our discount factor choice:

The lower our choice of  $\beta$ , the higher will be the system response rate.

#### 4.4 Equivalent Discount Factor

We may describe an exponential smoothing process as having an order equal to the number of coefficients in the corresponding model or the number of coefficients to be estimated. The constant process model with which we have been dealing represents first order exponential smoothing.

Let us make the safe assumption that smoothing of any order  $n$  includes a constant process term in the model. There is some  $\beta_n$  for  $n^{\text{th}}$  order smoothing that will give the same estimate for the constant term as  $\beta_1$  in single smoothing with a constant model.

Consider the response of a system to a unit impulse input at time zero. The estimate of the constant term coefficient,  $\hat{a}_1(t)$ , is given by

$$\hat{a}_1(t) = 1 - \beta_n^n$$

for exponential smoothing of order  $n$  (Brown, 2, 147). We may now equate the  $\hat{a}_1(t)$  estimates for single smoothing and smoothing of order  $n$ .

$$\hat{a}_1(t) = 1 - \beta_1^1 = 1 - \beta_n^n$$

Hence, the equivalent discount factor is defined by

$$\beta_1^1 = \beta_n^n.$$

Thus, if  $\beta_n$  is the discount factor used in an  $n^{\text{th}}$  order exponential smoothing process, a  $\beta_1$  equal to  $\beta_n^n$  would give the same estimate of  $a_1(t)$  in single smoothing. Stated another way, if  $\beta_1$  is the discount factor used in single smoothing, a  $\beta_n$  equal to

$$\beta_n = \beta_1^{1/n}$$

will give the same  $\hat{a}_1(t)$  in  $n^{\text{th}}$  order exponential smoothing.

We infer from this discussion that an  $n^{\text{th}}$  order smoothing system will have response characteristics similar to a first order system with an equivalent discount factor. Let us reiterate the major inference to be drawn here:

A smaller  $\beta$  gives more weight to more recent data than does a larger  $\beta$  value.

#### 4.5 Choosing the Initial Coefficient Vector

Exponential smoothing always requires some initial condition from which to start the smoothing process. This logically entails some initial estimate of the coefficients of the model fitting functions, or  $\hat{a}(0)$ . Again we must note that very little conclusive research on the choice of  $\hat{a}(0)$  was found in the literature. The relative merit of an accurate choice is a subject for debate, however, since the need for accuracy here is greatly dependent upon our choice of the discount factor and the amount of past data available. For example, if we have very little past data at the time the smoothing process begins, we may want to estimate  $a(0)$  as accurately as possible to shorten the length of time necessary for the smoothing system to stabilize in the light of new data. On the other hand, if we have ample past data, the smoothing system will have more information from which it may stabilize before making a forecast. The dependence of  $a(0)$  on  $\beta$  may be expressed generally in this way:

If you are very confident that your estimate of  $a(0)$  is a good one, use a large  $\beta$ , since quick system response to errors in  $\hat{a}(0)$  will probably not be necessary. Conversely, if you have very

little confidence in your initial estimate, use a smaller  $\beta$ , so the system may adjust to the possible error more quickly.

There are several means of obtaining the initial coefficient estimate.

If the investigator is thoroughly familiar with the process involved, he may be satisfied with a subjective estimate of  $a(0)$ . He may employ a graph of past data against time to estimate the initial conditions. Perhaps the most logical approach is to use the least squares regression coefficients. This, however, requires some past data, the more the better. If we have no past data, we may be able to compare the process with another one for which we have ample data. This might be true in a case where  $x(t)$  represents demand for some new product. Past data on a similar product may serve as a good predictor of  $a(0)$  for the new product.

Brown (2, 194) presents a scheme for estimating the initial coefficient values for models containing a constant and some periodic terms. First, if the constant coefficient is  $a_1(t)$ , let  $\hat{a}_1(0)$  be the average of the past data over any number of full cycles. Let  $a_2(t)$  and  $a_3(t)$  be the coefficients of a pair of periodic terms for a period of  $p$  defined by the fitting function pair for which we want coefficient estimates. If there are several cycles of past data available, average corresponding data samples across them. If there are no data, estimate the cyclic pattern. By one of these methods we obtain a sequence of  $p$  numbers  $\hat{x}_1, \hat{x}_2, \dots, \hat{x}_p$ . The initial coefficient estimates are then obtained from



$$\hat{a}_2(0) = \frac{2}{p} \sum_{k=1}^p x_k \sin \frac{2\pi}{p} k.$$

$$\hat{a}_3(0) = \frac{2}{p} \sum_{k=1}^p \hat{x}_k \cos \frac{2\pi}{p} k.$$

where

$p$  = the number of sampling intervals in the period under  
study

and

$k$  ranges from 1 to  $p$ .

## 5. GATHERING THE DATA

### 5.1 Time Series Requirements

The general exponential smoothing techniques developed earlier require the formulation of some model that is representative of the time series for which we want a forecast. Generally, the forecaster cannot adequately determine the fitting functions to be included in this model without some analysis of past data. Data analysis may also aid the forecaster in making a good initial coefficient vector choice. Such analyses were treated in previous chapters.

The past time series data must conform to the analyses for which it will be used. So, it should be discrete data, taken at specific and equal sampling intervals, from the present time back to some past time. Furthermore, the data should be in a form, i.e., dollars, gallons, etc., that is easily manipulated and interpreted. Consistency of the data is vital in that a particular time series value should truly represent the time series process for some particular sampling interval.

Data exhibiting such characteristics is not always readily available. Ideally, a forecaster would want to design the record-keeping system from which he will obtain the time series. Unfortunately, this is not always possible, and we must use what data is available at the present time. There are, however, some alternatives to the manner in which the data is manipulated. It may be desirable to transform the data. By doing so we may produce a more meaningful series, or one that is more easily predicted. Transformation of data can usually be performed by some

mathematical operation, such as a multiplier, log, etc.

Suppose the time series for which we want forecasts is incomplete, or not always available. We may find that there is some related time series that may be more easily predicted. If there is some functional relationship between the two series, we can obtain the desired forecast from a forecast of the related series.

Another possibility is sub-classification of the time series. It may be easier to forecast separate parts of a time series and sum these forecasts to get a forecast of the whole. This, of course, would involve more work than a single forecast, but the need for accuracy may warrant such a procedure.

In our discussion of model types (Section 2.4), we implied a further forecasting alternative. We may be able to find one or a combination of other series that acts as a leading series to the one which we want to predict. That is, this other series may serve as an indicator of what the original series will look like a certain number of sampling intervals in the future. In this case we may wish to use an autoregressive model in the adaptive smoothing technique mentioned earlier.

## 5.2 Actual Data Used

The data used in this research consists of four time series taken from the Manhattan Wholesale and Retail Floral Companies of Manhattan, Kansas. These time series represent demand for four types of plants produced by Manhattan Wholesale:

1. Potted chrysanthemums sold only through Manhattan Wholesale.

2. Potted chrysanthemums sold only through Manhattan Retail.
3. Potted poinsettias sold through Manhattan Wholesale.
4. Potted lilies sold through Manhattan Wholesale.

Chrysanthemums are produced and sold throughout the year. The two time series related to this product group represent end-of-month dollar receipts on sales of these plants. The latter two products are sold strictly on a seasonal basis. Poinsettias are sold only as a Christmas season item, while lilies are sold only as an Easter plant. Data on seasonal items is in number of pots sold per day.

Both the chrysanthemum time series extend over a period from September, 1958, to May, 1967, inclusive. This gives a total of 105 observations, or more than eight complete yearly cycles. Figures reflect the actual receipt of payments during a month for chrysanthemum sales. So, a figure does not necessarily represent the sales that occurred during a particular month, since not all buyers pay their bills promptly each month. Nevertheless, this was the only data available on these sales at the time of this research. This fact may explain the high noise level of the data, as we shall see later. Plots of the data are shown in Appendix E.

It should be noted here that the data available at the time a forecast is desired will probably be anything but ideal. Hence, future data compilation should be altered, if necessary, to best reflect the characteristics of the data for which a forecast is desired.

The retail receipts chrysanthemum time series shows some growth. This can probably be attributed to a combination of inflation and market expansion. However, the market in the case of retail sales is probably

nearly saturated, since it is confined mainly to Manhattan alone. The wholesale chrysanthemum time series shows a much greater growth trend. Most of this is due to market expansion. The wholesale market at present includes numerous cities other than Manhattan. Though many of the cities to which deliveries now go are near saturation, the market as a whole is not.

The poinsettia time series, also plotted in Appendix E, covers the demand periods for the years 1963 through 1966. Thus, four complete sets of data were available. The market area is roughly the same as for wholesale chrysanthemum sales. The demand period per year for poinsettias was arbitrarily set as the thirty-five days immediately preceding Christmas Day. Since the earliest purchase date for the data used was thirty-three days before Christmas, all data is included within the range of the thirty-five day period, and we have a total of 140 observations. Normally, deliveries were not made on Sundays, so the data shows a fluctuation about the zero line, unlike the two series previously discussed.

The lily time series, Appendix E, has a form similar to that for the poinsettias, as might be expected. The data also covers the sales periods for the years 1963 through 1967. This time, however, a demand period of fourteen days before Easter is large enough to include all of the data. Hence, we have 70 observations.

The market growth may be seen in these seasonal items, poinsettias and lilies, in the form of an amplitude growth rather than a trend line. For these items the data available was more complete; that is, it was broken down by customer. Hence, we are able to segregate the market for

poinsettias and lilies in some fashion, if desirable. By comparing the total number of potential buyers (usually retail florists) in each city with the number to whom we now deliver, the local market saturation becomes apparent. We might interpret this to mean that unless new market areas are exploited, the growth of sales may level off somewhat.

Since all of the customer outlets are in essentially the same geographic area, segregation by geographic area was not feasible. Though most cities in the market area are saturated, there are one or two larger cities in which relatively few outlets buy from Manhattan Floral Wholesale Company. Hence, segregation of customers by city size is meaningless. Data was not available by which segregation by customer size could be performed.

Another point should be clarified before we continue. The existing productive capacity of the Manhattan Floral Wholesale Company is nearing full utilization, so much further market expansion would be impossible without an expansion of productive capacity also.

The possibility of correlating some independent variable or variables with a particular time series was also looked at briefly. The obvious problem here is finding a time series that has a high correlation with one of the afore mentioned series, and is also known for future time. In addition, assuming a related series is available we incur the added expense of keeping multiple data records. Such a series, though none could be found, would have enabled the application of adaptive smoothing with an auto-regressive model.

Due to the data limitations presented above, all time series were

used in the form in which they were gathered. The following chapter contains the results of the application of general exponential smoothing to these four time series.

## 6. DISCUSSION OF RESULTS

### 6.1 Experimental Procedure

All four time series were initially plotted (Appendix E). The forms of the resulting curves were noted in order to gain some insight into the fitting functions that should be included in the models for each time series process. In addition, detrending, auto-correlation, and harmonic analyses were performed on each time series. From the results of these analyses and from intuition, various models were proposed. These models were then used in the general exponential smoothing techniques for various initial conditions, various initial fitting function coefficient vectors, and a range of discount factor values.

In applying general exponential smoothing to a time series, the steady-state F-matrix was first determined, for the particular  $\beta$  and  $a(0)$ , from the data given. Then, using the methods discussed in Section 3.5, forecasts were made for all sampling intervals for which data was actually available. So, starting at time zero, a forecast was made for the first time series value and compared to the actual value at observation one. Based on this comparison, a forecast was then made for time series value two, and compared to actual value two, and so on. Each forecast was projected one sampling interval into the future. Forecasts were performed only for time series values for which actual data existed. In this way we could measure the relative efficiency of a particular model - initial condition - discount factor combination. Note that it would be impossible to measure the accuracy of a forecast for which no actual data existed.



A discussion of the various efficiency criteria follows.

Generally, the experimental procedure for a single time series may be summarized as follows:

- (1) Plot the time series against time.
- (2) Perform a detrending analysis.
- (3) Perform an auto-correlation analysis.
- (4) Perform a harmonic analysis.
- (5) Formulate various models to be tried.
- (6) Determine the F-matrix.
- (7) Use general exponential smoothing to forecast the original time series. Do this for various values of  $\alpha$  and  $\beta$ .
- (8) Compute the efficiency criteria for each exponential smoothing iteration.
- (9) Compare the results and draw conclusions.

## 6.2 Efficiency Criteria

In order to evaluate and compare the forecasting efficiency of the models used, it is necessary to establish certain efficiency measurement methods. The obvious approach to this problem is to measure the accuracy of fit of the general exponential smoothing process using a particular model-initial condition-discount factor combination. With this goal in mind, the following efficiency criteria were proposed:

- (1) Variance of the forecasts,  $\sigma_f^2$
- (2) Standard deviation of the forecasts,  $\sigma_f$ .
- (3) Mean error,  $\bar{e}$

- (4) Mean absolute error,  $|\bar{\epsilon}|$
- (5) Mean percent error,  $\bar{\epsilon}\%$ .
- (6) Variance of the error,  $\sigma_{\epsilon}^2$ .
- (7) Error standard deviation about the mean error,  $\sigma_{\bar{\epsilon}}$ .

The first criterion above,  $\sigma_{\bar{\epsilon}}^2$ , is not truly the variance of the forecasts, but the variance of the fitting errors about zero. The computations necessary to determine the true forecast variance are quite lengthy, so this simple computation has been used instead (Brown, 2, 393). We shall call the former quantity a forecast variance to distinguish it from the error variance about the mean error, which is defined later. So, let us compute the variance of the forecasts as:

$$\sigma_{\bar{\epsilon}}^2 = \frac{\sum_{i=1}^N \epsilon_i^2}{N}$$

where

$$\epsilon_i = x_i - \hat{x}_i$$

$x_i$  = the  $i^{\text{th}}$  actual observation

$\hat{x}_i$  = the exponential smoothing forecast of the  $i^{\text{th}}$  observation

$N$  = the total number of data points observed and forecasted.

Similarly, we may now estimate the standard deviation of the forecasts as:

$$\sigma_{\bar{\epsilon}} = \sqrt{\sigma_{\bar{\epsilon}}^2}.$$

The mean error may be computed from

$$\bar{\epsilon} = \frac{1}{N} \sum_{i=1}^N \epsilon_i.$$

Note that  $\bar{\epsilon}$  may be positive or negative. If we consider a shortage and a surplus as being equally detrimental, we may prefer to look at the mean absolute error,  $|\bar{\epsilon}|$ , or

$$|\bar{\epsilon}| = \frac{1}{N} \sum_{i=1}^N |\epsilon_i|.$$

A similar error measurement method is the mean per cent error,  $\bar{\epsilon}\%$ , or

$$\bar{\epsilon}\% = 100 \cdot \frac{1}{N} \sum_{i=1}^N \frac{|\epsilon_i|}{x_i}.$$

The difference between this efficiency criterion and the mean absolute error needs emphasis.  $|\bar{\epsilon}|$  is in the same units as the original time series, whereas  $\bar{\epsilon}\%$  is a percentage of the actual data. Thus, when  $|\epsilon_i|$  is small,  $\epsilon_i\%$  may be quite large if  $x_i$  is also small. The converse is also true.

The variance of the error is a true variance computed relative to the mean of the errors. It may be written:

$$\sigma_{\bar{\epsilon}}^2 = \frac{1}{N} \sum_{i=1}^N (\epsilon_i - \bar{\epsilon})^2.$$

Then  $\sigma_{\bar{\epsilon}}$  is merely the square root of the variance above,  $\sqrt{\sigma_{\bar{\epsilon}}^2}$ . Note that if there were no noise in the data, and if our model truly described the process, then  $\sigma_{\bar{\epsilon}}$  would be identical to  $\sigma_f$ , since  $\bar{\epsilon}$  would be zero.

Not all of the criteria above were studied. Since variance and standard deviation basically represent the same error characteristic, only standard deviation was analyzed. Thus, we are left with five efficiency criteria with which to compare various models:

- (1) Standard deviation of the forecasts.
- (2) Mean error.
- (3) Mean absolute error.
- (4) Mean percent error.
- (5) Error standard deviation about mean error.

Mean error proved to be difficult to interpret and, consequently, it was discarded as an efficiency criterion. The mean error over the entire range of available data represented such a small percentage of the corresponding time series average that  $\sigma_f$  and  $\sigma_e$  were nearly identical. Hence, the standard deviation of the forecasts was also discarded as an efficiency criterion.

### 6.3 Computer Programs

All computations were performed on the IBM System 360/50 digital computer at Kansas State University. The programs used are shown in Appendix F. The most important of these programs are:

- (1) AUTODET
- (2) POWSPEC
- (3) PLOT001
- (4) EXPOSMO

AUTODET first detrends the original time series using a first degree

polynomial. It then computes the auto-correlation coefficients of the detrended data for all lags specified by the user. POWSPEC uses the detrended data from AUTODET to perform a harmonic analysis on the time series. It computes values of a periodogram for the periods specified by the user. The autocorrelation function and periodogram may then be plotted using PLOT001.

EXPOSOMO is the computer program which performs the general exponential smoothing process. Inputs to the program include the original time series, the initial coefficient vector, the transition matrix, and the values of the discount factor the user wishes to try. The program computes and lists the transition matrix, the F and F-inverse matrices, the smoothing vector, the variances of the coefficients in the model, the forecasts, and the error criteria values. In addition, the program outputs a plot against time of the actual, forecast, and error data.

EXPOSOMO, AUTODET, and POWSPEC are taken from programs written by Mr. Raymond Miller, a former student at Kansas State University. For a more complete explanation of their operation refer to his thesis (Miller, 7, 65). Only minor changes were made to AUTODET and POWSPEC. Significant changes were made to the input-output structure of the exponential smoothing program written by Miller in developing EXPOSOMO. The basic computational procedure, however, was not altered.

#### 6.4 Effect of the Initial Coefficient Vector

Two approaches to choosing the initial coefficient vector were tried. In one approach the model coefficients determined in the least squares

detrending process were used. Recall that in detrending we used first degree polynomial regression, so coefficients were computed only for the two corresponding polynomial fitting functions. In the initial coefficients were used when applicable; all other initial coefficients were set at zero.

The other choice of initial conditions was purposely made to be erroneous; that is, initial coefficients were purposely set far from their proper values. The resulting exponential smoothing fitting errors were very large for earlier sampling intervals.

Efficiencies around the best value of the discount factor (the  $\beta$  which minimizes the efficiency criteria for that model) are similar for either initial coefficient vector choice; however, for the second choice the degree of error increases very rapidly as we move away from this best  $\beta$ . Thus, when a bad choice of initial conditions is made, the system is less stable, making the proper choice of the discount factor more critical.

## 6.5 Model Identification Notation

Throughout the remainder of this text, it will be convenient for us to develop a shorthand notation for identification of the various models used. Basically, for all models the fitting functions used fall into one of three general categories:

- (1) Cyclic terms
- (2) Polynomial terms
- (3) Combination of cyclic and polynomial terms.

Cyclic terms are sine and cosine function pairs for some specific integral period. We shall denote such terms as  $C_p$ , where  $C$  tells us that both a sine and cosine fitting function are included and  $p$  is the number of sampling intervals taken as the period for these functions. Thus, the notation  $C_{12}$  would indicate that the following terms are included in the model:

$$a_i \sin \frac{2\pi t}{12} + a_{i+1} \cos \frac{2\pi t}{12}$$

If more than one pair of cyclic terms is present, they will be written  $C_{p_1}, P_2, P_n$ , or

$$\begin{aligned} & a_i \sin \frac{2\pi t}{p_1} + a_{i+1} \cos \frac{2\pi t}{p_1} + a_{i+2} \sin \frac{2\pi t}{p_2} \\ & + a_{i+3} \cos \frac{2\pi t}{p_2} + a_{i+4} \sin \frac{2\pi t}{p_n} + a_{i+5} \cos \frac{2\pi t}{p_n} . \end{aligned}$$

Polynomial fitting functions include all terms of the form shown in Section 3.3. In all of the models containing polynomial terms, a complete polynomial of some degree is used. Note that a polynomial of degree zero designates the inclusion of a constant fitting function only. We shall denote polynomials as  $P_d$ , where  $P$  stands for polynomial and  $d$  is an integer designating the degree of the polynomial used. Thus, the notation  $P_3$  tells us that the following terms are present in the model:

$$a_0 + a_1 t + \frac{a_2}{2} t(t-1) + \frac{a_3}{6} t(t-1)(t-2).$$

By combinations of cyclic and polynomial terms, we mean the product

of such fitting functions. These combination fitting functions are encountered in only one form in this research, that is, the product of a cyclic pair and time,  $t$ . Since such terms produce linearly increasing values for the amplitude of the sine and cosine functions, we may refer to them as amplitude growth terms, or more simply, growth terms. We shall designate such terms with the notation  $G_p$ , where  $G$  indicates growth terms and  $p$  is the period of the cyclic terms to which this linear amplitude growth is applied. Thus,  $G_6$  tells us that the following terms are included in the process model:

$$a_i t \sin \frac{2\pi t}{6} + a_{i+1} t \cos \frac{2\pi t}{6} .$$

Models containing terms from more than one of the categories discussed above may be designated in the same fashion, with commas separating the fitting function notations of different categories. So, the notation  $G_6, 12, P_2, G_{12}$  stands for the following time series process model:

$$\begin{aligned} \hat{x}(t) = & a_0 + a_1 t + \frac{a_2}{2} t(t-1) + (a_3 + a_5 t) \sin \frac{2\pi t}{12} \\ & + (a_4 + a_6 t) \cos \frac{2\pi t}{12} + a_7 \sin \frac{2\pi t}{6} \\ & + a_8 \cos \frac{2\pi t}{6} . \end{aligned}$$

The notation described in this section will be used throughout the remainder of this text, as well as in Appendix E. In addition, all values discussed will be the equivalent single smoothing value, regardless of the order of the model involved (Section 4.4).



## 6.6 Wholesale Chrysanthemum Receipts Time Series

All tables and figures are contained in Appendix E. Hence, in referring to a specific table or figure, only the appropriate table or figure number will be given to avoid unnecessary repetition of the appendix reference.

The wholesale chrysanthemum receipts time series are shown in Fig. E-1. Note the growth trend of the curve. This tells us that some degree polynomial should probably be included in the model of this time series. The autocorrelation analysis, the results of which are shown in Fig. E-2, has peaks at 6 and 12 month periods, or lags. Thus we probably need to include at least the 6-month cyclic function pair, since the highest autocorrelation function value occurred at this period. Fig. E-2 also shows minor peaks at 21, 27, and 34 months. Due to the limitation in the number of observations available, we should put greater weight on the lower period peaks, if they appear significant. The squared nature of the peaks indicates that harmonics may be present.

The harmonic periodogram on this time series (Fig. E-3) exhibits peaks, in order of decreasing magnitude, at 6, 21, 8 and 10, 4, 26, and 17 months. Again we reject the higher periods due to the inadequacy of the data. Notice that the 12-month periodic coefficient is of a very low magnitude relative to the others. Since chrysanthemums are sold throughout the year, we would expect this periodic to be more significant. The 6-month periodic has a sharp peak, confirming the previous analysis. It seems that the 12-month peak in the autocorrelation function plot was merely a reflection of the 6-month periodic in the data. The harmonic

analysis becomes unstable as it nears the Nyquist frequency (Section 2.4) for this data, as we would expect. At a period of 35 months, for example, there are only 3 data points upon which to base the value of  $R$ . This is obviously not an adequate representation. Note, however, that even if a periodic does exist at 35 months, the period is long enough that it can probably be handled adequately by the adaptive nature of the general exponential smoothing process.

The various models tried on the wholesale chrysanthemum time series are listed in Table E-2. In analyzing results of the various model-discount factor combinations, let us consider each efficiency criterion in order. The model efficiency results have been plotted according to three general classifications:

- (1) Cyclic models and combination models with polynomials of degree one or zero;
- (2) Strictly polynomial models;
- (3) Combination models with polynomials of degree greater than one.

This method of classification is justified by the similarity of the plots for models within the same model group above. So, we shall first discuss each efficiency criterion as it relates to each model classification. We may refer to the model classifications above as, respectively:

- (1) Cyclic models,
- (2) Polynomial models,
- (3) Combination models.

Table E-2 shows the values of  $|\bar{\epsilon}|$ , mean absolute error, and  $\sigma_{\bar{\epsilon}}$ ,

error standard deviation, for the models and  $\beta$ 's tried. All of the cyclic models exhibit a unique minimum  $|\bar{\varepsilon}|$  for a  $\beta$  of about 0.40 or 0.45 (equivalent first order  $\beta$ ). This must be interpreted only as an approximate best  $\beta$ -value, since it was not economically feasible to cover a large number of values of the discount factor due to computer time requirements. Nevertheless, this range of best  $\beta$ 's is considerably different from that anticipated by Brown (2, 179). Brown indicates that a  $\beta$  less than 0.7 will rarely be necessary.

The order of efficiency by model and by model classification is also shown in Table E-2. The best of the cyclic models was C12, 6, P1. It seems significantly better than the other cyclic models on both the  $|\bar{\varepsilon}|$  and  $\sigma_{\bar{\varepsilon}}$  plots, perhaps due to its relative simplicity. Though adding amplitude growth to the 12-month periodic in this model did not improve it over the original model, it was still better than the same model with amplitude growth on the 6-month periodic. This fact seems to contradict the results of the harmonic analysis which shows the 12-month periodic as insignificant. The inclusion of other periodic terms in the cyclic models did not improve the  $|\bar{\varepsilon}|$  and  $\sigma_{\bar{\varepsilon}}$  values. One model was run with only a zero degree polynomial. It gave high values for the efficiency criteria, indicating that the inclusion of polynomial terms is aiding the model efficiency. Other model results are shown in Figures E-4 and E-5.

Polynomial models of the third, fourth, and fifth degrees were applied to the wholesale chrysanthemum receipts data. Though Figures E-6 and E-7 show very similar minimum efficiency criteria for all three models, we may note that the fifth degree model is more stable. The efficiency

values for this higher degree polynomial do not diverge from the minimum value as rapidly as for the other two models. This is a desirable characteristic since the consequences of choosing a bad  $\beta$  value will be less grave for the fifth degree model. As shown, the best polynomial model, P5, produced the best forecasts for  $\beta = 0.7$ . It is doubtful whether higher degree models would give sufficiently better results to warrant the expense for additional computer time. It is not surprising that a polynomial of such a high degree as the fifth degree is necessary to describe the time series when we consider that it must account for the cyclic movements within the data.

Among the many combination models tried (Table E-2) C6, 8, P5 gave the best efficiency criteria values. Models containing higher degree polynomials were generally better, as shown in Figures E-8 and E-9. The combination of trigonometric fitting functions that performed best was that with periods of 6 and 8 months. Recall that these were identified as the most significant periodics in the harmonic analysis. The associated  $\beta$  for greatest efficiency with this model was 0.6. This model, C6, 8, P5 was also the best model among all models tried. The inclusion of amplitude growth on periodic terms was detrimental to the efficiency of the models in which this was tried.

#### 6.7 Retail Chrysanthemum Receipts Time Series

The retail monthly chrysanthemum receipts time series plot (Fig. E-10) shows some growth trend, but this growth is not nearly as explicit as that in the wholesale receipts time series. This is probably due to the fact that market expansion is not as extreme, since Manhattan Floral

Retail deals mainly in Manhattan, Kansas alone. Thus, we might suspect that a polynomial of a degree less than the fifth will perform adequately for this data. Again, the data is very noisy, since we are using dollar receipts to represent sales per month. We pointed out earlier that these monthly money receipts are not necessarily for sales in that particular month.

The autocorrelation analysis on the retail receipts time series (Fig. E-11) indicates cyclic behavior for periods of 12, 24, 8 and 6 months. These peaks on the autocorrelation function plot are not clear. Their squared form suggests the presence of further harmonics in the data. The harmonic analysis on this time series (Fig. E-12) shows peaks at periods of 6, 12, 4, and 8 and 10 months respectively, in order of decreasing magnitude. Again, this plot appears to degenerate at periods greater than 35 months due to the inadequacy of the data.

The models tried in the general exponential smoothing of this data are listed in Table E-4. The best performing cyclic model was C12, 6, 4, P1, as shown in Figures E-13 and E-14. It appears from these cyclic models that all three periods (12, 6, 4) are significant enough to warrant inclusion in the model. Models containing periodic terms with growing amplitudes did not perform as well as the same models without amplitude growth. Most cyclic models performed best for  $\beta = 0.8$ . The best cyclic model, however, performed better for  $\beta = 0.6$ .

Again, third, fourth, and fifth degree polynomial models were tried (Figures E-15 and E-16). Just as for the wholesale receipts time series, the fifth degree polynomial model gave better forecasts than the other two.

The efficiency curves, however, are very close for the fourth and fifth degree polynomials. For these models the best  $\beta$  was 0.8, just as for the cyclic models.

Only one combination model (C12, 6, P3) was tried on the retail receipts time series. This model performed better than all others minimizing  $|\bar{e}|$  and  $\sigma_{\bar{e}}$  for  $\beta = 0.8$  (Figures E-17 and E-18). The periodics contained in this model are those indicated by the harmonic analysis as being the most significant.

#### 6.8 Lily Time Series

Figure E-19 is a plot of the potted lily sales time series for yearly demand periods of 14 days each. This time series is considerably different from the previous two. Instead of an obvious trend line, we see a deviation upward from the zero demand line. The data appears to be strongly periodic in nature, exhibiting amplitude growth on the periodics present.

The autocorrelation analysis of the lily data shows peaks at 14 and 5 days. The peak at 14 days is squared and unclear, indicating harmonics. The harmonic analysis shows periodics for 7 and 14 days. From this result we may suspect that the 5-day periodic in the autocorrelation analysis actually occurs for a 7-day lag. The 14-day peak in the harmonic analysis is also squared. The 7-day peak, on the other hand, is very sharp. Thus, in this analysis, as well as in the autocorrelation analysis, the 14-day peak may actually be a reflection of the 7-day peak, and not a separate periodic.

Table E-6 contains a list of the models used in attempts to forecast

the lily time series. All of the cyclic models (Figures E-22 and E-23) minimized the efficiency criteria at  $\beta = 0.8$ . This best  $\beta$  is considerably higher than for both the wholesale and retail chrysanthemum receipts time series. Perhaps this is due to the absence of excess random variation in the lily data. If this is the casual factor, the lily time series would not have to respond as quickly to changes in the actual data; that is, the lily data models would perform as well with higher values of the discount factor.

All of the cyclic models tried performed similarly. Two models, C14, 7, P1, G14, 7 and C14, 7, P1, may be identified as being slightly more efficient than the others. Note that these two models are identical except for the amplitude growth of the first one. Adding amplitude growth does not appear to affect the model performance one way or the other in the long run. This is somewhat surprising since the data plot (Fig. E-19) shows apparent amplitude growth in the actual data. It may also seem surprising that, in nearly all cases, a first degree polynomial performed more efficiently than P0. The periodic components present in the two models cited above are those indicated as significant by both the autocorrelation and harmonic analyses.

It did not seem feasible to run any models composed strictly of polynomial fitting functions, since this data appears to be mainly cyclic in nature. Hence, we shall move on to a discussion of combination models.

The combination model results are tabulated in Table E-6 and plotted in Figures E-24 and E-25. Again, the two periodic terms, 14 and 7-day periods, aid the model performance. Since both models gave similar results,

the addition of a higher degree polynomial had very little affect upon model performance. The fact that we need polynomial terms at all is somewhat surprising. The combination model, C14, 7, P2, was slightly better than the best cyclic model noted earlier. It produced troughs on the efficiency criteria plots at  $\beta = 0.9$ , again higher than the wholesale and retail receipts best  $\beta$  values.

#### 6.9 Poinsettia Time Series

The poinsettia data plot (Fig. E-26) is very similar to the lily data plot. Recall that the demand period each year chosen for this data is 35 days in length. This demand period always terminates on the same date of the year, December 25, instead of the same day of the week, Easter Sunday in the case of the lily time series. Because of this fact, week-ends will occur at different locations within the demand period from year to year. Since deliveries are rarely made on Sundays, the data will probably contain a 7-day periodic. However, this periodic, as well as any others, will be obscured due to the "week-end effect" pointed out above. Hence, we may expect some additional noise in this data, not encountered in the lily time series. Just as for the lily data, we should expect the best poinsettia time series model to contain cyclics with growing amplitudes. Recall, however, that our suspicions were not entirely verified by the experimental results of the lily time series models.

The autocorrelation function plot for the poinsettia time series (Fig. E-27) indicates cyclics at periods of 31, 33, 6, and 28 days. These periods are not easily explained unless we consider the "week-end effect"



upon the data. With this in mind, we may probably interpret the 31 and 33-day peaks as reflections of the expected 35-day periodic. Similarly, the 6-day peak should probably be taken as an indication of a 7-day, or weekly, cycle. The harmonic analysis (Fig. E-28) is also very obscure. Nevertheless, we may interpret the peaks at 32 and 8 days as periodics of 35 and 7 days respectively, since these two periodics seem to be the logical choices for the data.

Table E-8 lists all the models and their resulting efficiencies for the poinsettia data. All of the cyclic models used gave very similar minimum error results (Fig. E-29 and E-30). However, some models were more stable than others. The best performing cyclic model was C35, 7, P0, at a discount factor value of 0.8. Models containing amplitude growth terms were generally less stable than those without these terms. Still, they seem to give almost equal performances at  $\beta = 0.8$ . The periodic functions included in these models are those interpreted as significant from the autocorrelation and harmonic analyses. P0 gave better results than P1 as we hypothesized from the time series plot.

Again, no polynomial models were run. The most efficient combination model used was C35, 7, P2, as shown in Figures 3-31 and E-32. The tendency for this time series was toward the lower degree polynomials, as evidenced by both the cyclic and combination models. C35, 7, P2 gave minimum error at  $\beta = 0.9$ .

The best model overall for the poinsettia time series was the cyclic model C35, 7, P0. We must keep in mind, however, that all models, cyclic and combination, gave very similar results.

Mean percent error was computed for the best model in each model classification for both the retail and wholesale chrysanthemum time series. For each of these models  $\bar{e}\%$  covers a range of  $\beta$  values. The results of these computations are shown in Table E-9 and Figures E-33 and E-34. This efficiency criterion agrees very closely with the results shown by  $\bar{F}$  and  $\sigma_e^2$ . For both time series the combination models gave the lowest  $\bar{e}\%$  for  $\beta$  values of 0.6 (wholesale) and 0.8 (retail). The mean percent error figures are relatively high due to the very high noise level of these data.  $\bar{e}\%$  could not be computed for the lily and poinsettia data since both of these time series contain zero demand observations.

#### 6.10 Notes on Harmonic Analysis

Throughout the data analyses above it appears that harmonic analysis gives a clearer identification of the periods in the data than does autocorrelation analysis. However, for extremely noisy data, even harmonic analysis is quite obscure.

The analysis is performed on detrended data. This detrended data is merely the residuals from a first degree polynomial regression on the original time series. Since the object of detrending is to remove the trend from the data, some time series may require polynomial regression of a higher degree than the first. Brown (2, 72) warns that detrending with higher degrees is likely to remove some of the cyclic variation in the data. This is undoubtedly true for cyclics with very long periods, but detrending with high degree polynomials may actually aid in the identification of short period cyclics. Such was the case in this research.

The wholesale receipts data was detrended with second, third, fourth, and fifth degree polynomial regression. The residuals from each of these regressions were used to obtain harmonic analysis periodograms (Figures E-35, E-36, E-37, E-38). The higher degree detrending seemed to emphasize the major cyclic periods and de-emphasize the minor ones. This effect is seen as sharper peaks at the major periods and a leveling of the less significant period peaks. It is doubtful that a polynomial regression of a higher degree would have aided the investigation further, since the fifth degree polynomial fits the wholesale data trend quite well.

## 7. CONCLUSIONS

From the discussions in Chapter 6 we may conclude that for a particular time series and a particular model of this time series, there is a unique value of the discount factor that maximizes model efficiency. We may define maximum efficiency as the minimization of the three efficiency criteria proposed in Chapter 6, that is, standard deviation of the forecast, mean absolute error, and error standard deviation about the mean error. The nature of the data and form of the model are major factors contributing to the value of minimum error discount factor. If the time series is very noisy,  $\beta$  must be lower to allow the model to respond more quickly, and vice-versa. Moreover, if the model is not appropriate, that is, does not describe the time series process well,  $\beta$  must again be low to allow quick model response to changes in the actual data.

For the two chrysanthemum receipts time series studied, the most efficient value of  $\beta$  was lower than expected. For the wholesale data cyclic models it was 0.4. This fact seems to contradict the findings of Brown (2, 116), who implied that  $\beta$  should never be less than 0.7. These two chrysanthemum time series contained excessive noise, however. The lily and poinsettia time series models performed best for a discount factor value of 0.8. Since these time series contained less superimposed noise than those for chrysanthemums, the  $\beta$  value is higher, as expected.

The inclusion of polynomial terms in the models of all the time series studied tended to cause the most efficient value of  $\beta$  to be higher than for models that were mainly cyclic. This may be seen in the classifications made on the models used, that is, cyclic, polynomial,

and combination. In general, the minimum error  $\beta$  values were smaller for cyclic models than for polynomial models, and smaller for polynomial models than for combination models. It may be that combination models were more appropriate in describing the time series encountered in this research. Though the most efficient models for the retail receipts and poinsettia time series were not combination models but cyclic models, it may also be that the appropriate combination model was not tried.

The model classifications used in this research were justified by the error criteria plots. Models within the same classification produced curves belonging to the same family on the error criteria plots; furthermore, different model classifications produced different families of curves. Models within the same classifications also produced best results for a unique  $\beta$  value. These phenomena hold for a particular time series.

The efficiency criteria plots for the retail and wholesale receipts time series were quite similar, as were the same plots for the lily and poinsettia time series. The plots of these two pairs of time series, however, differed from each other. Recall that the former time series pair dealt with items sold throughout the year, while the latter pair dealt with seasonal items. The inference here is that the demand characteristics of the product studied may affect the proper choice of the discount factor. By demand characteristics we mean whether the product has constant demand, seasonal demand, whether it is a luxury, a necessity, etc.

Of the methods used to determine the initial coefficient vector, one method was slightly superior. This method used the coefficients

determined by the detrending analysis as initial coefficients for polynomial fitting functions, and zeroes as initial coefficients for all other fitting functions. The use of other initial conditions, even completely erroneous initial conditions, did not affect the best  $\beta$  choice. The general exponential smoothing process adapted well to inaccurate initial condition choices. It produced efficiencies for some of the error criteria almost equivalent to those with accurate initial condition choices for the same models. The main effect of a bad choice of the initial coefficient vector was to make the choice of  $\beta$  more critical; that is, the trough on the error criteria plots was much narrower when the initial conditions were improperly chosen.

The periodic analyses used, autocorrelation analysis and harmonic analysis, were quite reliable in detecting periodics in the data. Harmonic analysis gave a clearer indication of what these periodics were than did autocorrelation analysis. When the trend movement in the time series was apparently of a higher polynomial degree than the first, detrending with higher degree polynomial regression gave sharper and more clearly defined peaks at the major periods in the periodogram, while leveling off the less significant peaks.

In order to make an accurate forecast of the time series studied in this research, it is felt that better data will have to be acquired. The two chrysanthemum receipts time series contain excessive noise, while the lily and poinsettia time series do not cover a sufficient number of past cycles. It should also be noted here that there is still much to be learned about general exponential smoothing before this technique can be most efficiently used in forecasting.

## BIBLIOGRAPHY

- (1) Brown, Robert G. Decision Rules for Inventory Management. New York: Holt, Rinehart, and Winston, Inc., 1967.
- (2) Brown, Robert G. Smoothing Forecasting and Prediction of Discrete Time Series. New Jersey: Prentice-Hall International, Inc., 1963.
- (3) Davis, Harold T. The Analysis of Economic Time Series. Bloomington, Indiana: The Principia Press, Inc., 1941.
- (4) Davis, Harold T. The Theory of Econometrics. Bloomington, Indiana: The Principia Press, Inc., 1941.
- (5) Draper, Norman R. and Smith, Harold, Jr. Applied Regression Analysis. New York: John Wiley and Sons, Inc., 1966.
- (6) Graybill, Franklin A. An Introduction to Linear Statistical Models. New York: McGraw-Hill Book Company, Inc., Volume I, 1961.
- (7) Miller, Raymond C. The Application of Exponential Smoothing to Forecasting a Time Series. (Unpublished Master's Thesis. Kansas State University, 1967.)

## ACKNOWLEDGEMENTS

The author would like to express his gratitude to Dr. L. Eugene Grosh for his ideas and guidance throughout this research. The author is also grateful to Manhattan Floral Company for cooperating in the collection of data.



## APPENDIX A. DERIVATION OF STATISTICAL REGRESSION EQUATION

Let us suppose that we have a series of  $n$  discrete observations  $y(t)$  which we wish to describe by some model. We assume that the time series  $y(t)$  is composed of some process plus superimposed random variation,  $\epsilon(t)$ . We have at our disposal related independent variables  $x_i(t)$  which we may use to estimate  $y(t)$ . Thus we may write

$$\begin{aligned} y(t) &= \hat{y}(t) + \epsilon(t) \\ &= \beta_0 + \beta_1 x_1(t) + \beta_2 x_2(t) + \dots \\ &\quad + \beta_{k-1} x_{k-1}(t) + \epsilon(t) \end{aligned}$$

where

$y(t)$  = the value of the original time series at time  $t$

$\hat{y}(t)$  = our estimate of  $y(t)$

$x_i(t)$  = the  $i^{\text{th}}$  independent variable at time  $t$

$\beta_0$  = a constant

$\beta_i$  = the coefficient of  $x_i(t)$

$\epsilon(t)$  = the error of our estimate at time  $t$

$$= y(t) - \hat{y}(t).$$

We may write this equation in matrix notation as

$$Y = X\beta + \epsilon$$

where

$$Y = \begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(n) \end{bmatrix} \quad \begin{array}{l} \text{= an } n \times 1 \text{ column vector of time series} \\ \text{observations} \end{array}$$

$$X = \begin{bmatrix} 1 & x_1(1) & \dots & x_{k-1}(1) \\ 1 & x_1(2) & \dots & x_{k-1}(2) \\ 1 & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 1 & x_1(n) & \dots & x_{k-1}(n) \end{bmatrix}$$

= an  $n \times k$  matrix of independent variables

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \cdot \\ \cdot \\ \beta_{k-1} \end{bmatrix} = \text{a } k \times 1 \text{ column vector of coefficients}$$

$$\epsilon = \begin{bmatrix} \epsilon(1) \\ \epsilon(2) \\ \cdot \\ \cdot \\ \epsilon(n) \end{bmatrix} = \text{an } n \times 1 \text{ column vector of fitting errors}$$

We wish to determine the vector independent variable coefficients that minimize the sum of squared errors. So, let us first determine this squared residual sum,  $S$ .

$$S = \sum_{i=1}^n (e_i)^2 = e'e$$

where

$e' =$  the transpose of  $e$ .

Note that we may write  $e$  thusly:

$$e = Y - X\beta$$

So,

$$\begin{aligned} S &= (Y - X\beta)' (Y - X\beta) \\ &= (Y' - \beta'X') (Y - X\beta) \end{aligned}$$

Multiplying this out, we find that

$$\begin{aligned} S &= Y'Y - Y'X\beta - \beta'X'Y + \beta'X'X\beta \\ &= Y'Y - 2\beta'X'Y + \beta'X'X\beta \\ &= Y'Y - 2(Y'X)\beta + \beta'X'X\beta . \end{aligned}$$

If we find the derivative of  $S$  with respect to  $\beta$ , and set the result equal to zero, we may solve for  $\beta$ .

$$\frac{\partial S}{\partial \beta} = -2X'Y + 2X'X\beta = 0$$

$$X'Xb = X'Y$$

Hence

$$b = (X'X)^{-1} X'Y$$

is our least squared residuals estimate of  $\beta$ .

## APPENDIX B. DERIVATION OF EXPONENTIAL SMOOTHING EQUATION

Let us suppose that we have discrete time series  $x(t)$  which is composed of some time-model-dependent process plus random variation. This may be written:

$$x(t) = \xi(t) + \varepsilon(t)$$

where

$x(t)$  = the time series observation at time  $t$

$\xi(t)$  = the value of the time-dependent process at time  $t$

$\varepsilon(t)$  = random variation at time  $t$

We wish to approximate  $\xi(t)$  by a linear function of time,  $\hat{x}(t)$ .

$$\hat{x}(t) = \sum_{i=1}^n a_i(t) f_i(t)$$

where

$\hat{x}(t)$  = our estimate of  $\xi(t)$

$f_i(t)$  = the  $i^{\text{th}}$  fitting function evaluated for time  $t$

$a_i(t)$  = the coefficient of the  $i^{\text{th}}$  fitting function.

The fitting functions are all deterministic functions of time.

The equation above may be written in matrix notation as:

$$\hat{x}(t) = a'(t)f(t)$$

where

$$a(t) = \begin{bmatrix} a_1(t) \\ a_2(t) \\ \vdots \\ a_n(t) \end{bmatrix} = \text{an } n \times 1 \text{ column vector of fitting function coefficients}$$

$$f(t) = \begin{bmatrix} f_1(t) \\ f_2(t) \\ \vdots \\ f_n(t) \end{bmatrix} = \text{an } n \times 1 \text{ column vector of fitting functions}$$

Substituting this back into the equation for  $x(t)$ , and again using matrix notation, we may write:

$$\begin{aligned} x &= \hat{x} + e \\ &= a'Z + e \end{aligned}$$

where

$x = [x(1), x(2), \dots, x(T)]$  = a  $1 \times T$  row vector of data observations up to time  $T$ .

$$\hat{x} = [\hat{x}(1), \hat{x}(2), \dots, \hat{x}(T)]$$

= a  $1 \times T$  row vector of estimates of  $x(t)$  given by the model

$$Z = \begin{bmatrix} f_1(1) & f_1(2) & \dots & f_1(T) \\ f_2(1) & f_2(2) & \dots & f_2(T) \\ \vdots & \vdots & & \vdots \\ f_n(1) & f_n(2) & \dots & f_n(T) \end{bmatrix}$$

= an  $n \times T$  matrix of fitting functions

$$\begin{aligned}
e &= [e(1), e(2), \dots, e(T)] \\
&= a \text{ } 1 \times T \text{ row vector of residuals} \\
&= x - \hat{x} .
\end{aligned}$$

We wish to determine the coefficient vector,  $a$ , that minimizes the weighted sum of squared residuals. In exponential smoothing the time series observation  $k$  periods in the past is given weight  $\alpha\beta^k$ . We shall define a  $T \times T$  matrix  $W$  in which  $W_{ii}$  is the square root of the weight given the observation  $(T-i)$  sampling intervals age, and all off-diagonal elements are zero. Now we may express the weighted sum of squared residuals as

$$S = eW(eW)'$$

Since

$$\begin{aligned}
e &= x - a'F, \text{ we may write} \\
S &= eW(eW)' = (xW - a'FW)'(xW - a'FW)'.
\end{aligned}$$

The vector of fitting function coefficients that minimizes this weighted sum of squared residuals may be determined from the equation (Brown, 2, 392).

$$\frac{\partial S}{\partial a} = 0.$$

So,

$$\begin{aligned}
S &= xWW'x' - a'FWW'x' \\
&\quad - xWW'F'a + a'FWW'F'a
\end{aligned}$$

and

$$\frac{\partial S}{\partial a} = 0 = -xWW'F' + a'FWW'F'$$

or

$$xWW'Z' = a'Z'WW'Z'.$$

To simplify the equation above, let us make the following definition:

$$F = Z'W(Z'W)$$

$$= Z'WW'Z'$$

= an  $n \times n$  symmetrical matrix of weighted fitting  
functions

$$g = [xWW'Z']' = \text{an } n \times 1 \text{ data vector.}$$

Now, the previous equation reduces to

$$g' = a'F$$

or

$$a' = g'F^{-1}$$

where

$T$  must be  $\geq n$  for  $F^{-1}$  to exist.

This vector of fitting function coefficients minimizes the weighted sum of squared residuals.

## APPENDIX C. POLYNOMIAL REGRESSION

Suppose we have a discrete time series of  $T$  observations  $y(t)$  to which we wish to fit a polynomial of degree  $(k - 1)$ . We may express a polynomial of degree  $(k - 1)$  on time as:

$$p(t) = 1 + t + t^2 + \dots + t^{k-1}$$

where

$$t = \text{time.}$$

This polynomial, evaluated for  $t$  from 0 to  $T$ , may be compactly expressed in the form of a  $T \times k$  matrix  $X$ ,

$$X = \begin{bmatrix} 1 & 1 & 1 & \dots & 1^{k-1} \\ 1 & 2 & 4 & \dots & 2^{k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & T & T^2 & \dots & T^{k-1} \end{bmatrix}$$

where

$$\begin{aligned} X_{ij} &= \text{the element of } X \text{ contained in row } i \text{ and column } j \\ &= (i)^{j-1} . \end{aligned}$$

We wish to describe an observation  $y(t)$  by

$$y(t) = \sum_{i=0}^{k-1} \beta_i t^i + \epsilon(t)$$

where

$$y(t) = \text{the time series observation made at time } t.$$



$\beta_i$  = the coefficient of  $t^i$  for any  $t$ .

$\epsilon(t)$  = the fitting error at time  $t$ .

In matrix notation this equation would be

$$Y = X\beta + \epsilon$$

where

$Y$  = a  $T \times 1$  column vector of time series observations.

$X$  = the  $T \times k$  matrix defined earlier.

$\beta$  = a  $k \times 1$  column vector of polynomial term coefficients.

$\epsilon$  = a  $T \times 1$  column vector of fitting errors.

The vector  $\beta$  must be chosen such that the sum of squared residuals  $\epsilon(t)$  is minimized. The least squares estimate of  $\beta$ ,  $b$ , is given by (Appendix A),

$$\begin{aligned} b &= (X'X)^{-1} X'Y \\ &= \text{a } k \times 1 \text{ column vector of least squares estimates of } \beta. \end{aligned}$$

So, having determined  $b$ , we may write the resulting regression estimates as:

$$\hat{Y} = Xb$$

where

$\hat{Y}$  = a  $T \times 1$  column vector of  $(k - 1)$  degree polynomial regression estimates for  $y(t)$ .

## APPENDIX D. THE Z-TRANSFORM

The Z-transform is a useful technique for the analysis of a linear, discrete, time-invariant system (Brown). By discrete we mean a sampled data system. The linearity of a system may best be described by the following:

For input =  $x_1(t)$ , then output =  $y_1(t)$ .

For input =  $x_2(t)$ , then output =  $y_2(t)$ .

Now, when input =  $a x_1(t) + b x_2(t)$ , where  $a$  and  $b$  are constants, the system is linear if

$$\text{output} = a y_1(t) + b y_2(t).$$

A system is said to be time-invariant if input  $x(t)$  causes output  $y(t)$  regardless of the value of  $t$ .

If we let  $f(n)$  be any function defined for discrete intervals, we may define its Z-transform,  $F(Z)$ , as

$$F(Z) = \sum_{n=-\infty}^{\infty} f(n)Z^n.$$

This transform is especially useful in representing convolutions.

A convolution may be defined as

$$\begin{aligned} y(t) &= \sum_{n=-\infty}^{\infty} x(n) n(t-n) \\ &= x(t) * h(t) \end{aligned}$$

where

$x(t)$  and  $h(t)$  are sequences of observations.

In the case of system response analysis,  $x(t)$  may be the input signal,

and  $h(t)$  the response to a unit impulse  $t$  periods after its input.

Taking the Z-transform of both sides of the convolution equation, we may see that a convolution in the time domain is replaced by a multiplication, a much simpler mathematical operation, in the Z-domain.

$$Y(Z) = X(Z) \cdot H(Z)$$

$Y(Z)$  = the Z-transform of  $y(t)$

$X(Z)$  = the Z-transform of  $x(t)$

$H(Z)$  = the Z-transform of the impulse response  $h(t)$

= the transfer function.

## APPENDIX E - TABLES AND FIGURES

Each time series model referred to in this appendix is classified into one of the following three categories:

- (1) Cyclical Models - models which contain trigonometric fitting functions and polynomials of degree less than two.
- (2) Polynomial Models - models which contain only polynomial fitting functions.
- (3) Combination Models - models which contain trigonometric fitting functions and polynomials of degree greater than one.

Model components are indicated according to the following notation:

$$Cp_1, p_2, p_k, P_n, Gp_1, p_2, p_k$$

where  $P_n, n = \text{integers.}$

This symbolic notation is interpreted as follows:

$Cp_1, p_2, p_k$  indicates that the model contains both a sine and cosine fitting function for each of the integral periods,  $p_1, p_2,$  and  $p_k$ .

$P_n$  indicates that the model contains fitting functions corresponding to all of the terms of an  $n^{\text{th}}$  degree polynomial.

$Gp_1, p_2, p_k$  indicates the inclusion of linear amplitude growth on each of the sine-cosine pairs in the model for periods of  $p_1, p_2,$  and  $p_k$ . These terms are just trigonometric fitting functions multiplied by time,  $t$ .

If a model does not contain fitting functions of the C, P, or G type, the corresponding symbol (C, P, or G) is omitted in the notation for this model.

## TABLE IDENTIFICATION CHART

<u>TABLE NUMBER</u>	<u>DESCRIPTION</u>
E-1	Monthly Wholesale Chrysanthemum Receipts Actual and Detrended Time Series.
E-2	Wholesale Chrysanthemum Receipts Time Series Models Tried and Resulting Efficiencies.
E-3	Monthly Retail Chrysanthemum Receipts Actual and Detrended Time Series.
E-4	Retail Chrysanthemum Receipts Time Series Models Tried and Resulting Efficiencies.
E-5	Lily Daily Sales Actual and Detrended Time Series.
E-6	Lily Time Series Models Tried and Resulting Efficiencies.
E-7	Poinsettia Daily Sales Actual and Detrended Time Series.
E-8	Poinsettia Sales Time Series Models Tried and Resulting Efficiencies.
E-9	Mean Percent Error Values for Selected Models of Wholesale and Retail Chrysanthemum Time Series.

## FIGURE IDENTIFICATION CHART

The table below is provided as a reference to all of the figures which follow.

<u>FIGURE NUMBER</u>	<u>DESCRIPTION</u>
E-1	Wholesale Monthly Chrysanthemum Receipts Time Series Plot.
E-2	Autocorrelation Function for Wholesale Chrysanthemum Receipts Data.
E-3	Periodogram for Wholesale Chrysanthemum Receipts Data.
E-4	Mean Absolute Error Versus Discount Factor for Wholesale Chrysanthemum Data--Cyclic Models.
E-5	Standard Deviation of Error Versus Discount Factor for Wholesale Chrysanthemum Data--Cyclic Models.
E-6	Mean Absolute Error Versus Discount Factor for Wholesale Chrysanthemum Data--Polynomial Models.
E-7	Standard Deviation of Error Versus Discount Factor for Wholesale Chrysanthemum Data--Polynomial Models.
E-8	Mean Absolute Error Versus Discount Factor for Wholesale Chrysanthemum Data--Combination Models.
E-9	Standard Deviation of Error Versus Discount Factor for Wholesale Chrysanthemum Data--Combination Models.
E-10	Retail Monthly Chrysanthemum Receipts Time Series Plot.
E-11	Autocorrelation Function for Retail Chrysanthemum Receipts Data.
E-12	Periodogram for Retail Chrysanthemum Receipts Data.
E-13	Mean Absolute Error Versus Discount Factor for Retail Chrysanthemum Data--Cyclic Models.
E-14	Standard Deviation of Error Versus Discount Factor for Retail Chrysanthemum Data--Cyclic Models.

FIGURE NUMBERDESCRIPTION

E-15	Mean Absolute Error Versus Discount Factor for Retail Chrysanthemum Data--Polynomial Models.
E-16	Standard Deviation of Error Versus Discount Factor for Retail Chrysanthemum Data--Polynomial Models.
E-17	Mean Absolute Error Versus Discount Factor for Retail Chrysanthemum Data--Combination Model.
E-18	Standard Deviation of Error Versus Discount Factor for Retail Chrysanthemum Data--Combination Model.
E-19	Lily Daily Sales Time Series Plot.
E-20	Autocorrelation Function for Lily Data.
E-21	Periodogram for Lily Data.
E-22	Mean Absolute Error Versus Discount Factor for Retail Lily Data--Cyclic Models.
E-23	Standard Deviation of Error Versus Discount Factor for Retail Lily Data--Cyclic Models.
E-24	Mean Absolute Error Versus Discount Factor for Lily Data--Combination Models.
E-25	Standard Deviation of Error Versus Discount Factor for Lily Data--Combination Models.
E-26	Poinsettias Daily Sales Time Series Plot.
E-27	Autocorrelation Function for Poinsettia Data.
E-28	Periodogram for Poinsettia Data.
E-29	Mean Absolute Error Versus Discount Factor for Poinsettia Data--Cyclic Models.
E-30	Standard Deviation of Error Versus Discount Factor for Poinsettia Data--Cyclic Models.
E-31	Mean Absolute Error Versus Discount Factor for Poinsettia Data--Combination Models.
E-32	Standard Deviation of Error Versus Discount Factor for Poinsettia Data--Combination Models.

FIGURE NUMBERDESCRIPTION

E-33	Mean Percent Error Versus Discount Factor for Best Wholesale Chrysanthemum Time Series Models.
E-34	Mean Percent Error Versus Discount Factor for Best Retail Chrysanthemum Time Series Models.
E-35	Periodogram for Wholesale Chrysanthemum Data Using Second Degree Polynomial Regression Residuals.
E-36	Periodogram for Wholesale Chrysanthemum Data Using Third Degree Polynomial Regression Residuals.
E-37	Periodogram for Wholesale Chrysanthemum Data Using Fourth Degree Polynomial Regression Residuals.
E-38	Periodogram for Wholesale Chrysanthemum Data Using Fifth Degree Polynomial Regression Residuals.



PERIOD	ACTUAL	DETRENDED
1.	274.00000000	230.94224548
2.	275.00000000	188.88450623
3.	253.00000000	123.82678223
4.	178.00000000	5.76904297
5.	373.00000000	157.71128845
6.	364.00000000	105.65354919
7.	570.00000000	268.59594727
8.	444.00000000	95.53808564
9.	986.00000000	598.48046275
10.	473.00000000	42.42259216
11.	403.00000000	-70.63500977
12.	545.00000000	32.30712891
13.	630.00000000	70.24951172
14.	760.00000000	157.19163513
15.	1100.00000000	454.13378906
16.	924.00000000	235.07615662
17.	544.00000000	-187.98143005
18.	1039.00000000	263.96069336
19.	951.00000000	132.90306091
20.	1008.00000000	206.84519958
21.	1053.00000000	148.78755766
22.	1543.00000000	575.72973633
23.	1380.00000000	389.67187500
24.	1505.00000000	475.61401367
25.	1651.00000000	574.65664063
26.	1601.00000000	481.49877930
27.	2719.00000000	1554.44116211
28.	1554.00000000	348.38305664
29.	1772.00000000	523.32543945
30.	876.00000000	-415.73193359
31.	1142.00000000	-192.78977966
32.	1404.00000000	26.15234375
33.	1648.00000000	227.09471130
34.	1637.00000000	173.03684598
35.	1553.00000000	50.97924805
36.	1177.00000000	-373.07836914
37.	1361.00000000	-232.17623047
38.	1550.00000000	313.80566406
39.	1943.00000000	263.74829102
40.	1337.00000000	-385.30932617
41.	1143.00000000	-622.36718750
42.	1081.00000000	-727.42504883
43.	1435.00000000	-416.48266602
44.	1554.00000000	-340.54028320
45.	1432.00000000	-499.59790039
46.	1247.00000000	-733.65600586
47.	1171.00000000	-852.71362305
48.	1291.00000000	-775.77124023
49.	234.00000000	-1275.82285742
50.	575.00000000	-1173.88671875
51.	1016.00000000	-1179.94458008
52.	742.00000000	-1491.00244141

TABLE E-1

Monthly Wholesale Chrysanthemum Receipts  
Actual and Detrended Time Series

PERIOD	ACTUAL	DETRENDED
53.	1450.00000000	-832.05981445
54.	976.00000000	-1349.11767518
55.	1001.00000000	-1367.17573711
56.	1153.00000000	-1218.23335844
57.	1024.00000000	-1430.29077148
58.	2028.00000000	-469.34863281
59.	774.00000000	-1766.40625000
60.	1161.00000000	-1422.46435547
61.	1356.00000000	-1270.52197260
62.	2408.00000000	-261.57958984
63.	1624.00000000	-1088.63720703
64.	1860.00000000	-895.69531250
65.	2210.00000000	-588.75292969
66.	1701.00000000	-1140.81079102
67.	2944.00000000	59.13155180
68.	2404.00000000	-523.52602539
69.	2684.00000000	-266.98388672
70.	3568.00000000	553.95800781
71.	2826.00000000	-231.09934958
72.	1659.00000000	-1401.15698242
73.	2077.00000000	-1066.21484375
74.	2658.00000000	-528.27270508
75.	4306.00000000	1076.66943359
76.	3604.00000000	331.61181641
77.	4116.00000000	800.55419522
78.	3102.00000000	-256.50366211
79.	2878.00000000	-523.56127930
80.	4806.00000000	1361.38061523
81.	3422.00000000	-65.67674255
82.	4150.00000000	605.26538086
83.	3438.00000000	-135.79223633
84.	3663.00000000	46.14988708
85.	3970.00000000	310.09228516
86.	5526.00000000	1823.03442383
87.	5575.00000000	1828.97656250
88.	4956.00000000	1166.91870117
89.	3450.00000000	-382.13842773
90.	4620.00000000	744.80346680
91.	4256.00000000	337.74584961
92.	4369.00000000	407.68774414
93.	5469.00000000	1404.63012695
94.	6117.00000000	2009.57250977
95.	4546.00000000	455.51489258
96.	3488.00000000	-645.54296875
97.	4238.00000000	61.40234375
98.	4200.00000000	-19.65625000
99.	4031.00000000	-231.71484375
100.	5341.00000000	1035.22656250
101.	3712.00000000	-636.82812500
102.	3697.00000000	-654.88671875
103.	4673.00000000	238.05468750
104.	3557.00000000	-921.00390625
105.	4368.00000000	-153.06250000

TABLE E-1 (cont'd.)

CYCLIC MODELS  
DISCOUNT FACTOR,  $\beta$ MODEL AND  
CRITERION

	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55	.60	.65	.70	.75	.80	.85	.90	.95
C12, 6, P1																		
K1	661	593	544	510	485	469	461	459	464	473	485	497	513	529	546	565	574	586
O <sub>C</sub>	882	779	716	674	647	631	622	620	624	633	645	660	677	695	714	733	748	756
C12, 6, P1, G12																		
K1	707	598	553	526	509	497	497	501	510	518	526	536	548	562	574	581	585	591
O <sub>C</sub>	969	812	740	700	678	668	666	670	677	687	699	712	724	737	743	758	763	765
C12, 6, P1																		
K1			518				498				525				572			
O <sub>C</sub>			707				663				696				743			
C12, 6, P1, G6																		
K1			628				500				537				572			
O <sub>C</sub>			870				670				720				743			
C6, 4, P1, G6																		
K1			529				501				531				569			
O <sub>C</sub>			730				676				700				737			
C10, 8, P1, G, P0																		
K1			533				515				546				582			
O <sub>C</sub>			742				696				717				749			
C12, 6, P1, G6																		
K1			544				535				561				584			
O <sub>C</sub>			747				715				738				757			

TABLE E-2

Wholesale Chrysanthemum Receipts Time Series Models  
Tried and Resulting Efficiencies

POLYNOMIAL MODELS  
DISCOUNT FACTOR,  $\beta$

MODEL AND CRITERION	.20	.40	.60	.70	.80	.85	.90	.95
P5 $\overline{K}$ $\overline{O_c}$	5419 9334	498 725	430 618	418 588	431 589	455 612	503 673	569 746
P4 $\overline{K}$ $\overline{O_c}$	18423 42260	586 812	428 613	420 592	430 586	459 618	508 677	570 747
P3 $\overline{K}$ $\overline{O_c}$	767 1243	436 615	420 588	432 589	459 619	507 677	568 745	

TABLE E-2 (cont'd.)

COMBINATORIAL MODELS  
DISCOUNT FACTOR,  $\beta$

MODEL AND CRITERION	.20	.40	.60	.70	.80	.85	.90
C6, 8, P5  kl		420	394		453		
O <sub>6</sub> <sup>F</sup> C6, 8, P4  kl		575	524		605		
O <sub>6</sub> <sup>F</sup> C12, 6, P4  kl		433	402		469		554
O <sub>6</sub> <sup>F</sup> C12, 6, P4  kl		595	544		630		725
O <sub>6</sub> <sup>F</sup> C6, P4  kl		436	405	423	477	515	560
O <sub>6</sub> <sup>F</sup> C6, P4  kl		599	552	568	635	682	727
O <sub>6</sub> <sup>F</sup> C6, P4  kl		457	409	407	449		
O <sub>6</sub> <sup>F</sup> C12, 6, P3  kl		636	566	558	598		
O <sub>6</sub> <sup>F</sup> C12, 6, P3  kl	555	442	413		494		
O <sub>6</sub> <sup>F</sup> C12, 6, P5  kl	738	598	516		657		
O <sub>6</sub> <sup>F</sup> C12, 6, P2  kl		448	412		471		557
O <sub>6</sub> <sup>F</sup> C12, 6, P2  kl		632	573		632		728
O <sub>6</sub> <sup>F</sup> C6, 4, P2, 66  kl	554	443	436		518		
O <sub>6</sub> <sup>F</sup> C6, 4, P2, 66  kl	731	595	586		684		
O <sub>6</sub> <sup>F</sup> C6, 4, P2, 66  kl	537	448	476		542		
O <sub>6</sub> <sup>F</sup> C6, 4, P2, 66  kl	713	612	629		711		

TABLE E-2 (cont'd.)

PERIOD	ACTUAL	DETRENDED
1.	301.00000000	300.5725077
2.	398.00000000	397.14477539
3.	344.00000000	342.7177344
4.	257.00000000	255.29049483
5.	345.00000000	342.66303711
6.	305.00000000	362.43530273
7.	495.00000000	492.00805064
8.	325.00000000	321.58081055
9.	191.00000000	127.15336445
10.	245.00000000	240.77019629
11.	307.00000000	302.29858398
12.	151.00000000	145.87142944
13.	262.00000000	256.44384766
14.	456.00000000	450.01635742
15.	347.00000000	340.58911133
16.	345.00000000	235.16162109
17.	414.00000000	406.73437500
18.	444.00000000	436.30712871
19.	573.00000000	504.87939453
20.	271.00000000	262.45239256
21.	288.00000000	279.02465020
22.	192.00000000	182.59705025
23.	227.00000000	277.17016602
24.	237.00000000	220.74208940
25.	557.00000000	541.31542969
26.	414.00000000	402.82769531
27.	749.00000000	737.46069336
28.	251.00000000	239.03337067
29.	282.00000000	269.60571289
30.	373.00000000	360.17846000
31.	388.00000000	374.75037856
32.	466.00000000	451.32373047
33.	370.00000000	365.89624623
34.	336.00000000	321.46275000
35.	281.00000000	266.04174805
36.	269.00000000	253.61434137
37.	312.00000000	290.18676758
38.	355.00000000	338.75927734
39.	433.00000000	416.33203125
40.	120.00000000	102.96483093
41.	326.00000000	302.47705072
42.	304.00000000	280.05004483
43.	546.00000000	521.62231445
44.	295.00000000	276.19531250
45.	178.00000000	158.76794434
46.	223.00000000	203.34056091
47.	150.00000000	129.61319275
48.	193.00000000	172.48580933
49.	706.00000000	223.05834901
50.	287.00000000	263.03051073
51.	254.00000000	232.20327472
52.	121.00000000	98.77020089

TABLE E-3  
Monthly Retail Chrysanthemum Receipts Actual  
and Detrended Time Series

53.	141.00000000	118.34652173
54.	270.00000000	246.92153331
55.	216.00000000	152.40415030
56.	356.00000000	372.00040825
57.	294.00000000	289.03946430
58.	207.00000000	122.21200000
59.	267.00000000	341.70442383
60.	166.00000000	140.35727929
61.	263.00000000	236.92987061
62.	288.00000000	201.50244141
63.	280.00000000	323.07470703
64.	135.00000000	107.64773500
65.	309.00000000	281.21997070
66.	208.00000000	239.79299927
67.	471.00000000	442.35347852
68.	247.00000000	212.53821716
69.	235.00000000	205.51026426
70.	245.00000000	215.08346558
71.	177.00000000	140.65002215
72.	164.00000000	133.27672925
73.	294.00000000	262.60102539
74.	329.00000000	297.37377930
75.	266.00000000	333.54628900
76.	245.00000000	212.51921082
77.	762.00000000	269.09179088
78.	417.00000000	323.06406250
79.	418.00000000	384.23706055
80.	430.00000000	395.20932617
81.	209.00000000	174.36230896
82.	252.00000000	218.95449408
83.	358.00000000	322.52734375
84.	286.00000000	250.10015869
85.	503.00000000	466.67236326
86.	231.00000000	154.24540710
87.	402.00000000	370.81702095
88.	398.00000000	360.39002500
89.	286.00000000	247.96327209
90.	431.00000000	392.53564453
91.	442.00000000	403.10039244
92.	600.00000000	560.08066406
93.	457.00000000	417.25366211
94.	282.00000000	241.82638350
95.	170.00000000	135.39900208
96.	277.00000000	235.57151865
97.	412.00000000	370.54394531
98.	616.00000000	574.11669922
99.	385.00000000	342.65945313
100.	429.00000000	386.20171875
101.	437.00000000	453.83471680
102.	463.00000000	419.40092242
103.	362.00000000	347.97998047
104.	389.00000000	344.55224609
105.	479.00000000	434.12500000

TABLE E-3 (cont'd.)

CYCLIC MODELS  
DISCOUNT FACTOR,  $\beta$

MODEL AND  
CRITERION

	.10	.15	.20	.25	.30	.35	.40	.50	.60	.70	.80	.90	.95
C12, 6, 4, P1			107				90	85			83		
Kl			135				113	106			107		
C12, 6, P1													
Kl													
147	129	118	109	104	100	96	90	87	84	84	85	86	
184	162	148	138	130	124	120	113	109	107	107	108	110	
C12, 6, 4, P1, G12													
Kl			109			92		87			84		
O <sub>c</sub>			136			115		109			108		
C12, 6, 4, P1, G6													
Kl			111			92		87			84		
O <sub>c</sub>			139			117		109			108		
C12, 6, P1, G12													
Kl			126			99		88			85		
O <sub>c</sub>			157			122		111			109		
C12, 6, P1, G6													
Kl			177			96		90			86		
O <sub>c</sub>			231			122		112			109		
C12, 6, 3, P1													
Kl			112			96		89			86		
O <sub>c</sub>			144			122		112			110		

TABLE E-4

Retail Chrysanthemum Receipts Time Series  
Models Tried and Resulting Efficiencies



POLYNOMIAL MODELS  
DISCOUNT FACTOR,  $\beta$

MODEL AND CRITERION	.20	.40	.60	.70	.80	.90	.95
P5							
K1	557	110	97		91		
O1	867	140	126		115		
P4							
K1	1642	119	97		91		
O1	3612	147	126		115		
P3							
K1		503	130	101	100	107	
O1		898	160	129	125	133	

COMBINATION MODELS  
DISCOUNT FACTOR,  $\beta$

MODEL AND CRITERION	.20	.40	.60	.70	.80	.90	.95
G12, 6, P3							
K1		96	85		82	85	88
O1		122	108		104	108	111

TABLE E-4 (cont'd.)

PERIOD	ACTUAL	DETRENDED
1.	0.0	-1.26489258
2.	0.0	-2.52978611
3.	20.00000000	16.20530701
4.	190.00000000	184.54041443
5.	0.0	-6.32446766
6.	10.00000000	2.41063881
7.	0.0	-8.85425377
8.	0.0	-10.11914730
9.	0.0	-11.38404179
10.	72.00000000	59.35105896
11.	335.00000000	321.08593750
12.	0.0	-15.17872143
13.	20.00000000	3.55639553
14.	0.0	-17.70848083
15.	0.0	-18.97337341
16.	0.0	-20.23825651
17.	16.00000000	-5.50318909
18.	57.00000000	34.23190308
19.	12.00000000	-12.03297329
20.	18.00000000	-7.29786587
21.	161.00000000	134.43722534
22.	0.0	-27.82765198
23.	0.0	-29.09254456
24.	220.00000000	189.64256287
25.	0.0	-31.62231445
26.	289.00000000	256.11254883
27.	0.0	-34.15211487
28.	0.0	-35.41700745
29.	0.0	-36.68188477
30.	3.00000000	-34.54677734
31.	6.00000000	-33.21166952
32.	16.00000000	-24.47659302
33.	62.00000000	20.25849915
34.	217.00000000	173.99362183
35.	0.0	-44.27125549
36.	0.0	-45.53614807
37.	209.00000000	162.19892883
38.	195.00000000	146.93403625
39.	335.00000000	285.66870117
40.	10.00000000	-40.59571838
41.	0.0	-51.86061096
42.	0.0	-53.12551880
43.	0.0	-54.35041138
44.	34.00000000	-21.65530396
45.	23.00000000	-33.92018127
46.	17.00000000	-41.18507385
47.	177.00000000	117.55000305
48.	227.00000000	166.28509521
49.	17.00000000	-44.97976685
50.	17.00000000	-46.24465942

TABLE E-5

Lily Daily Sales Actual and Detrended Time Series

PERIOD	ACTUAL	DETRENDED
51.	172.00000000	107.45043274
52.	18.00000000	-47.77445924
53.	290.00000000	322.56044922
54.	78.00000000	9.69575405
55.	17.00000000	-52.56912231
56.	17.00000000	-53.83401489
57.	12.00000000	-60.09892273
58.	38.00000000	-35.36381531
59.	54.00000000	-20.62870789
60.	44.00000000	-31.89358521
61.	89.00000000	11.84150696
62.	213.00000000	134.57661438
63.	22.00000000	-57.08827820
64.	22.00000000	-58.95318604
65.	223.00000000	150.78192139
66.	242.00000000	152.51702881
67.	207.00000000	122.25212097
68.	266.00000000	179.98722839
69.	46.00000000	-41.27764893
70.	0.0	-82.54252625

TABLE E-5 (cont'd.)

CYCLIC MODELS  
DISCOUNT FACTOR,  $\beta$

MODEL AND CRITERION	.20	.40	.60	.80
Cl4, 7, Pl, Gl4, 7				
Kl	113	89	78	
O $\xi$	152	118	101	
Cl4, 7, Pl				
Kl	116	90	77	
O $\xi$	155	119	102	92
Cl4, 7, Pl, Gl4				
Kl	119	91	78	70
O $\xi$	157	120	103	92
Cl4, 7, PO				
Kl	113	88	77	72
O $\xi$	151	118	101	91
Cl4, 7, PO, Gl4				
Kl	116	91	79	74
O $\xi$	155	119	102	91

TABLE E-6

Lily Time Series Models Tried and Resulting Efficiencies

COMBINATION MODELS  
DISCOUNT FACTOR,  $\beta$

MODEL AND CRITERION	.40	.60	.80	.85	.90	.95
C14, 7, P2 $\overline{C_1}$	90	77	69	68	68	72
	120	102	92	89	88	84
C14, 7, P3 $\overline{C_1}$	91	77	69		68	
	121	103	92		88	

TABLE E-6 (cont'd.)

PERIOD	ACTUAL	DETREND
1.	0.0	-1.135617199
2.	0.0	-1.135617199
3.	0.0	-1.135617199
4.	0.0	-1.135617199
5.	0.0	-1.135617199
6.	0.0	-1.135617199
7.	0.0	-1.135617199
8.	0.0	-1.135617199
9.	0.0	-1.135617199
10.	0.0	-1.135617199
11.	0.0	-1.135617199
12.	0.0	-1.135617199
13.	0.0	-1.135617199
14.	0.0	-1.135617199
15.	0.0	-1.135617199
16.	0.0	-1.135617199
17.	0.0	-1.135617199
18.	0.0	-1.135617199
19.	0.0	-1.135617199
20.	0.0	-1.135617199
21.	0.0	-1.135617199
22.	0.0	-1.135617199
23.	0.0	-1.135617199
24.	0.0	-1.135617199
25.	0.0	-1.135617199
26.	0.0	-1.135617199
27.	0.0	-1.135617199
28.	0.0	-1.135617199
29.	0.0	-1.135617199
30.	0.0	-1.135617199
31.	0.0	-1.135617199
32.	0.0	-1.135617199
33.	0.0	-1.135617199
34.	0.0	-1.135617199
35.	0.0	-1.135617199
36.	0.0	-1.135617199
37.	0.0	-1.135617199
38.	0.0	-1.135617199
39.	0.0	-1.135617199
40.	0.0	-1.135617199
41.	0.0	-1.135617199
42.	0.0	-1.135617199
43.	0.0	-1.135617199
44.	0.0	-1.135617199
45.	0.0	-1.135617199
46.	0.0	-1.135617199
47.	0.0	-1.135617199
48.	0.0	-1.135617199
49.	0.0	-1.135617199
50.	0.0	-1.135617199
51.	0.0	-1.135617199
52.	0.0	-1.135617199

TABLE E-7

Poinsettia Daily Sales Actual and Detrended Time Series

PERIOD	ACTUAL	DETRENDED
53.	24.0000000	121.46435547
54.	23.0000000	10.10102110
55.	24.0000000	11.10102110
56.	125.0000000	-41.1357741
57.	54.0000000	44.1477677
58.	64.0000000	-113.15116640
59.	21.0000000	15.15116640
60.	129.0000000	-15.15116640
61.	26.0000000	17.15116640
62.	163.0000000	-27.13603210
63.	24.0000000	11.772545
64.	42.0000000	22.73852379
65.	74.0000000	44.1477677
66.	65.0000000	45.1477677
67.	111.0000000	75.15116640
68.	412.0000000	203.46348408
69.	6.0	-211.60277664
70.	6.0	-214.60970825
71.	6.0	-217.73603496
72.	6.0	-220.86313110
73.	2.0000000	-224.10102110
74.	0.0	-227.2555336
75.	6.0	-230.41004616
76.	6.0	-233.56455877
77.	61.0000000	-175.1408123
78.	0.0	-236.71911163
79.	4.0000000	-239.87362424
80.	103.0000000	-145.1477677
81.	238.0000000	-13.47024441
82.	117.0000000	117.55075559
83.	240.0000000	-57.66350037
84.	114.0000000	117.55075559
85.	6.0	-240.73100055
86.	341.0000000	14.1477677
87.	261.0000000	11.12108372
88.	13.0000000	-20.1477677
89.	712.0000000	41.1357741
90.	415.0000000	11.12108372
91.	415.0000000	22.1477677
92.	6.0	-240.73100055
93.	702.0000000	41.1357741
94.	314.0000000	22.1477677
95.	1219.0000000	724.55075559
96.	516.0000000	22.1477677
97.	927.0000000	31.1477677
98.	1119.0000000	11.12108372
99.	947.0000000	51.1357741
100.	70.0000000	11.12108372
101.	1371.0000000	135.1477677
102.	74.0000000	-236.71911163
103.	274.0000000	11.12108372
104.	6.0	-240.73100055
105.	6.0	-240.73100055
106.	6.0	-240.73100055

TABLE E-7 (cont'd.)

PERIOD	ACTUAL	DETRENDED
107.	0.0	-311.13757541
108.	0.0	-311.26434570
109.	0.0	-311.76944400
110.	0.0	-311.3313477
111.	50.00000000	-311.177516
112.	30.00000000	-313.7143555
113.	0.0	-311.3309116
114.	55.00000000	-274.14438647
115.	111.00000000	-184.71133086
116.	78.00000000	-277.72026125
117.	112.00000000	-13.19506716
118.	115.00000000	-214.17111671
119.	12.00000000	-332.43623842
120.	66.00000000	-302.11435231
121.	171.00000000	-100.7177208
122.	599.00000000	224.26131287
123.	275.00000000	112.24332227
124.	533.00000000	212.72301208
125.	377.00000000	34.33662500
126.	233.00000000	-113.40551753
127.	0.0	319.47216777
128.	1273.00000000	810.46833750
129.	558.00000000	477.30453125
130.	484.00000000	83.77737036
131.	538.00000000	146.26676633
132.	863.00000000	438.19433544
133.	167.00000000	217.7229917
134.	1387.00000000	366.18102516
135.	1657.00000000	1267.0937542
136.	549.00000000	131.92747434
137.	501.00000000	6.160765
138.	554.00000000	-67.28531055
139.	511.00000000	106.72727900
140.	0.0	-429.23935647

TABLE E-7 (cont'd.)



CYCLIC MODELS  
DISCOUNT FACTOR,  $\beta$

MODEL AND CRITERION	.20	.40	.60	.80
C35, 7, P0 Kl O <sub>f</sub>	269 391	235 328	224 300	212 281
C35, 7, P1 Kl O <sub>f</sub>	279 405	241 336	227 304	213 284
C35, 7, P1, G35 Kl O <sub>f</sub>	283 417	268 367	222 306	214 284
C35, 7, P1, G35, 7 Kl O <sub>f</sub>	534 917	246 348	228 306	214 284
C35, 7, P0, G35 Kl O <sub>f</sub>	1665 3479	261 382	221 301	216 281
C35, 7, P0, G35, 7 Kl O <sub>f</sub>	708 1274	241 332	229 303	217 281

TABLE E-8  
Poinsettia Sales Time Series Models Tried  
and  
Resulting Efficiencies

COMBINATION MODELS  
DISCOUNT FACTOR,  $\beta$ 

MODEL AND CRITERION	.40	.60	.80	.90
C35, 7, P2				
$\frac{ d }{O_{\xi}}$	241 340	228 307	212 285	205 277
C35, 7, P3				
$\frac{ d }{O_{\xi}}$	241 341	230 309	212 286	205 277

TABLE E-8 (cont'd.)

DISCOUNT FACTOR,  $\beta$

MODEL AND TIME SERIES	.10	.20	.30	.40	.50	.60	.70	.80	.90
Wholesale Chrysanthemum Receipts Time Series									
C12, 6, P1	35.85	28.90	25.67	24.59	25.22	26.84	29.24	32.53	36.63
P5		153.16		27.47		24.07	23.32	23.96	28.77
C6, 8, P5				23.37		22.31		25.41	
Retail Chrysanthemum Receipts Time Series									
C12, 6, 4, P1		35.43		29.70		28.37		23.92	
P5		164.17		33.53		34.38		32.28	
C12, 6, P3				33.60		29.60		23.30	30.10

TABLE E-9

Mean Percent Error Values,  $\bar{\epsilon}_\%$ , for Selected Models  
Of Wholesale and Retail Chrysanthemum Time Series

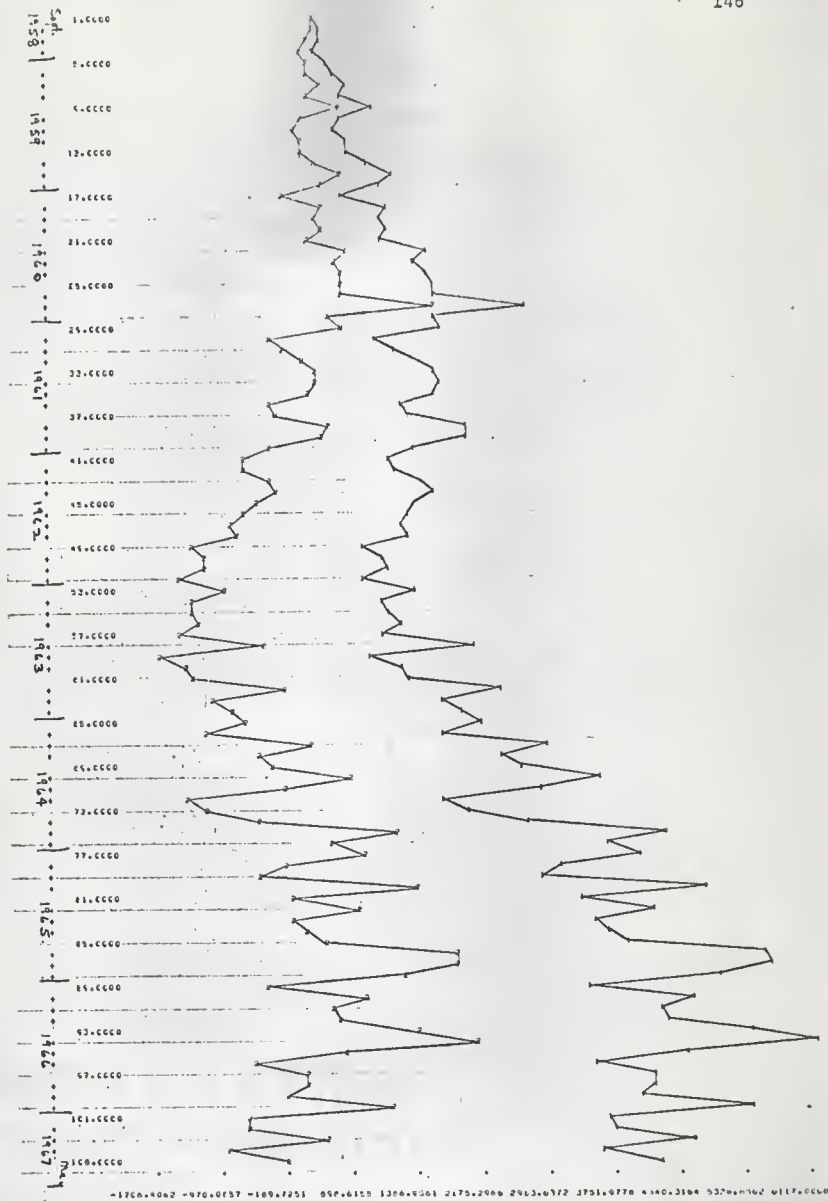


FIG. E-1

Wholesale Monthly Chrysanthemum Receipts Time Series Plot

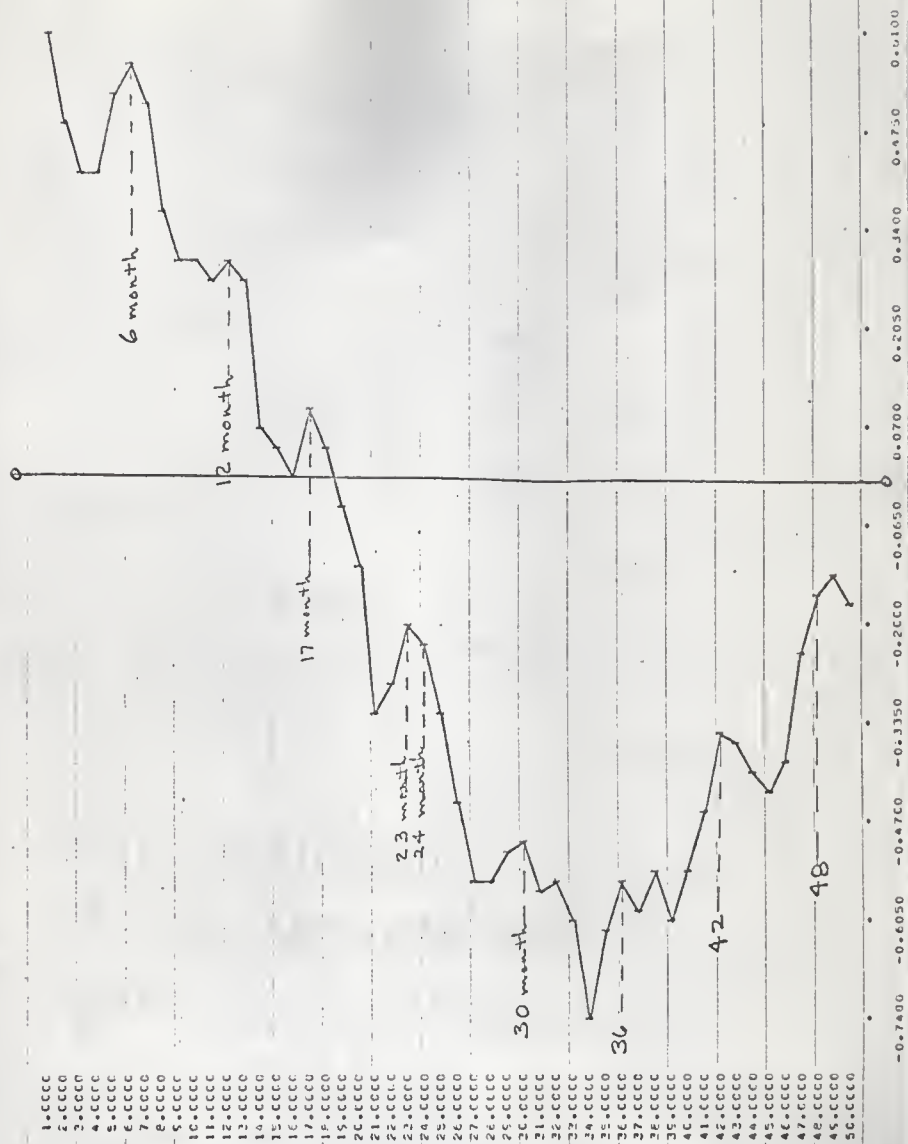
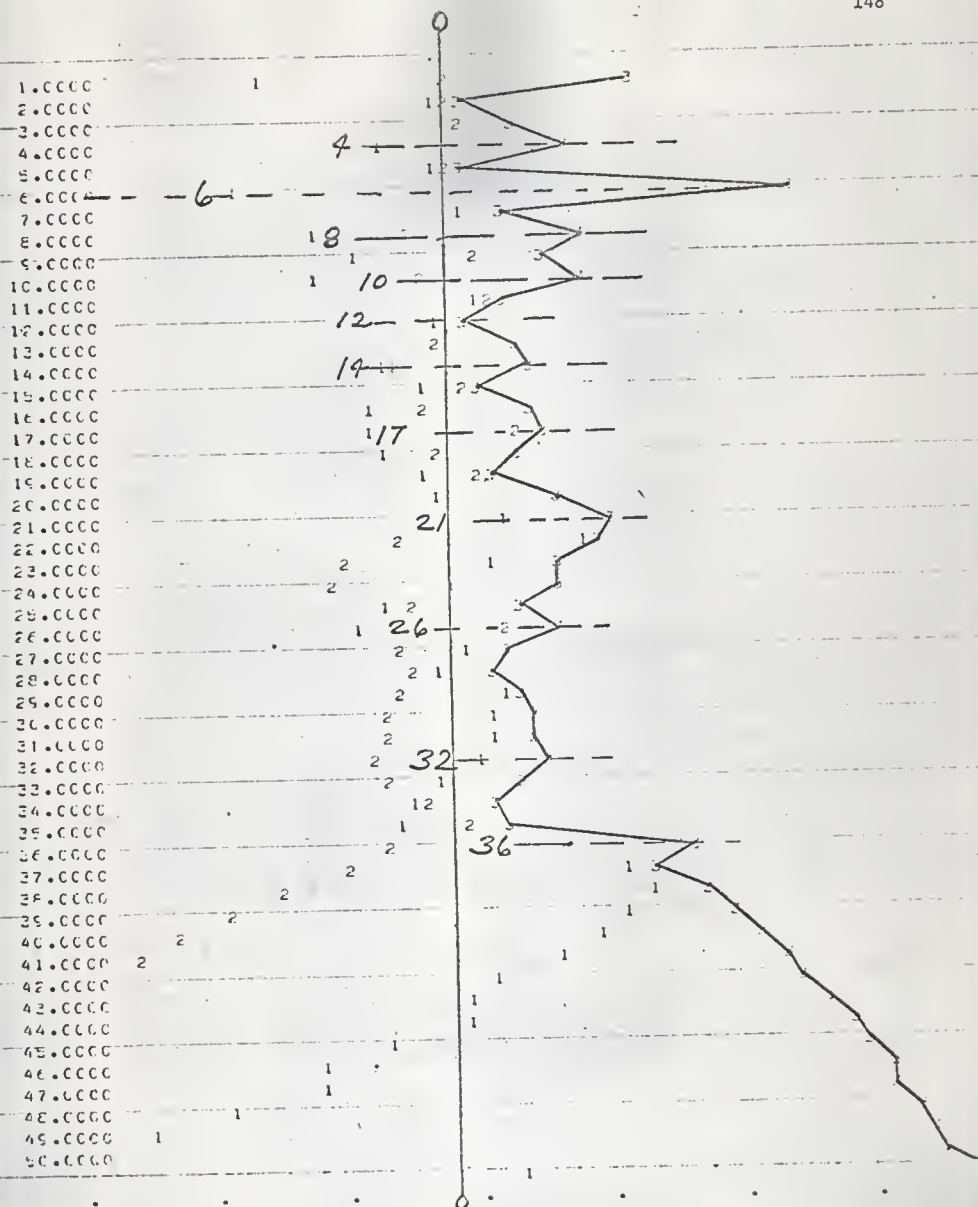


FIG. E-2

Autocorrelation Function for Wholesale Chrysanthemum Receipts Data



-303.7192 -233.4316 -167.1460 21.1436 169.4313 277.7117 406.0003

FIG. E-3

Periodogram for Wholesale Chrysanthemum Receipts Data

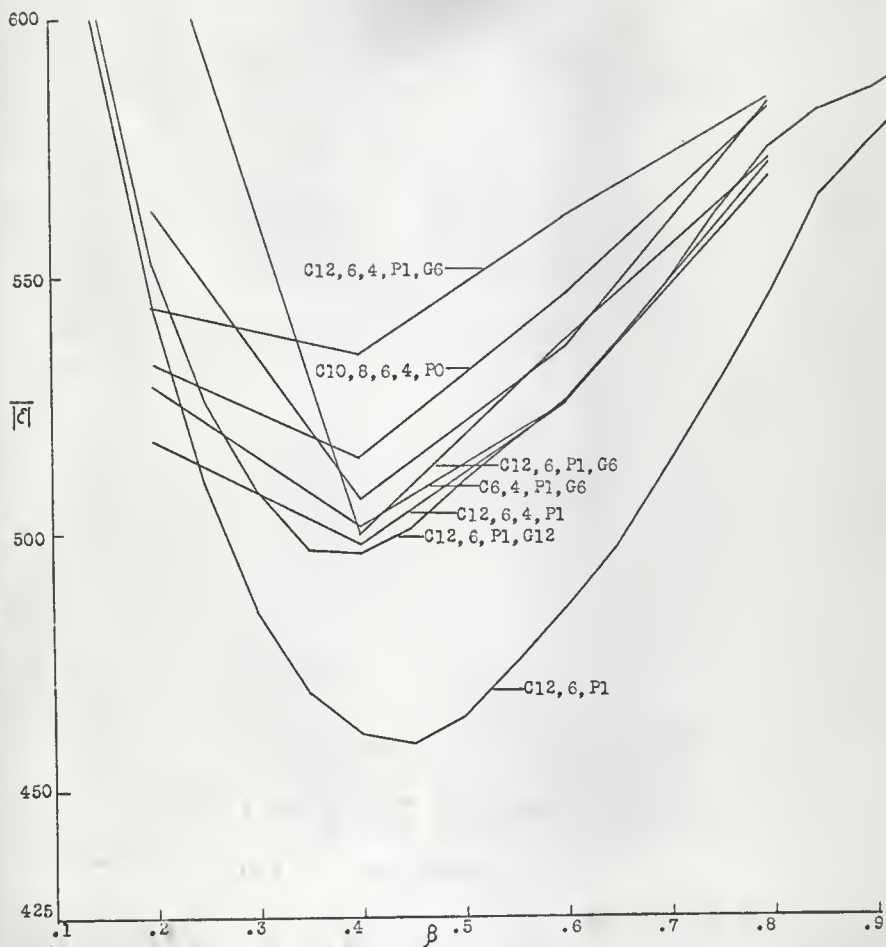


FIG. E-4

Mean Absolute Error versus Discount Factor for Wholesale  
Chrysanthemum Data--Cyclic Models

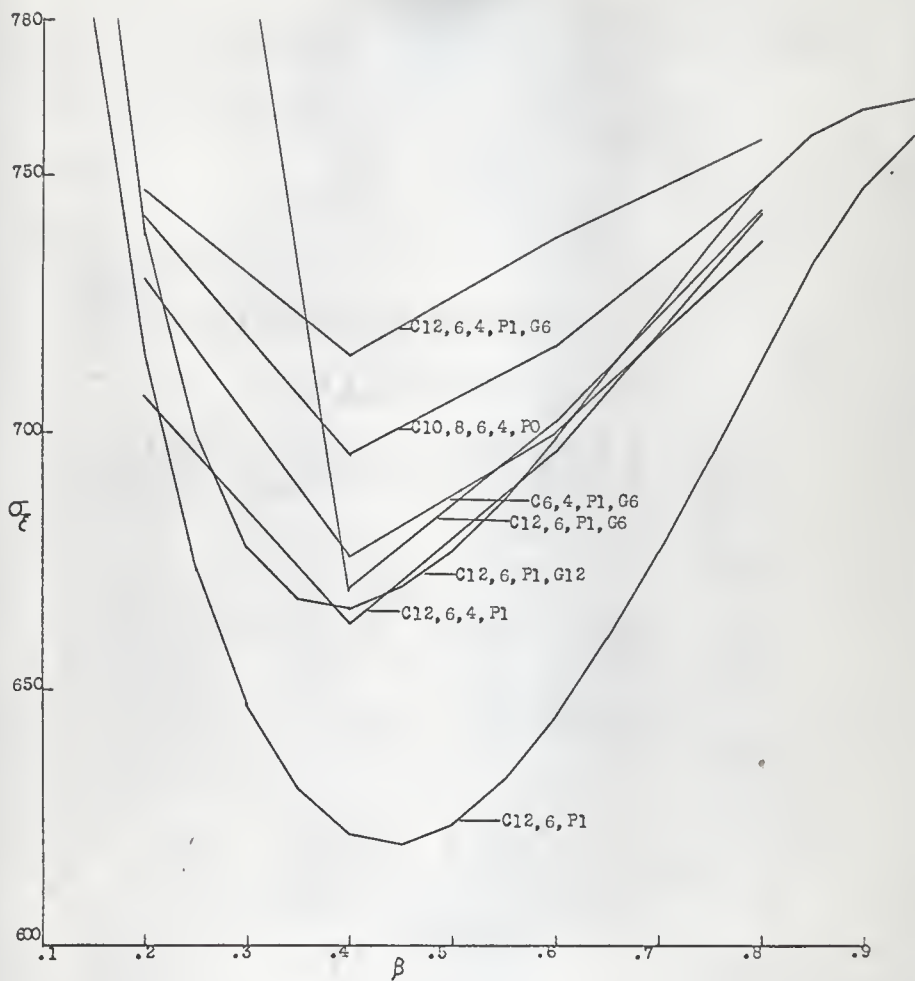


FIG. E-5

Standard Deviation of Error Versus Discount Factor for  
Wholesale Chrysanthemum Data---Cyclic Models



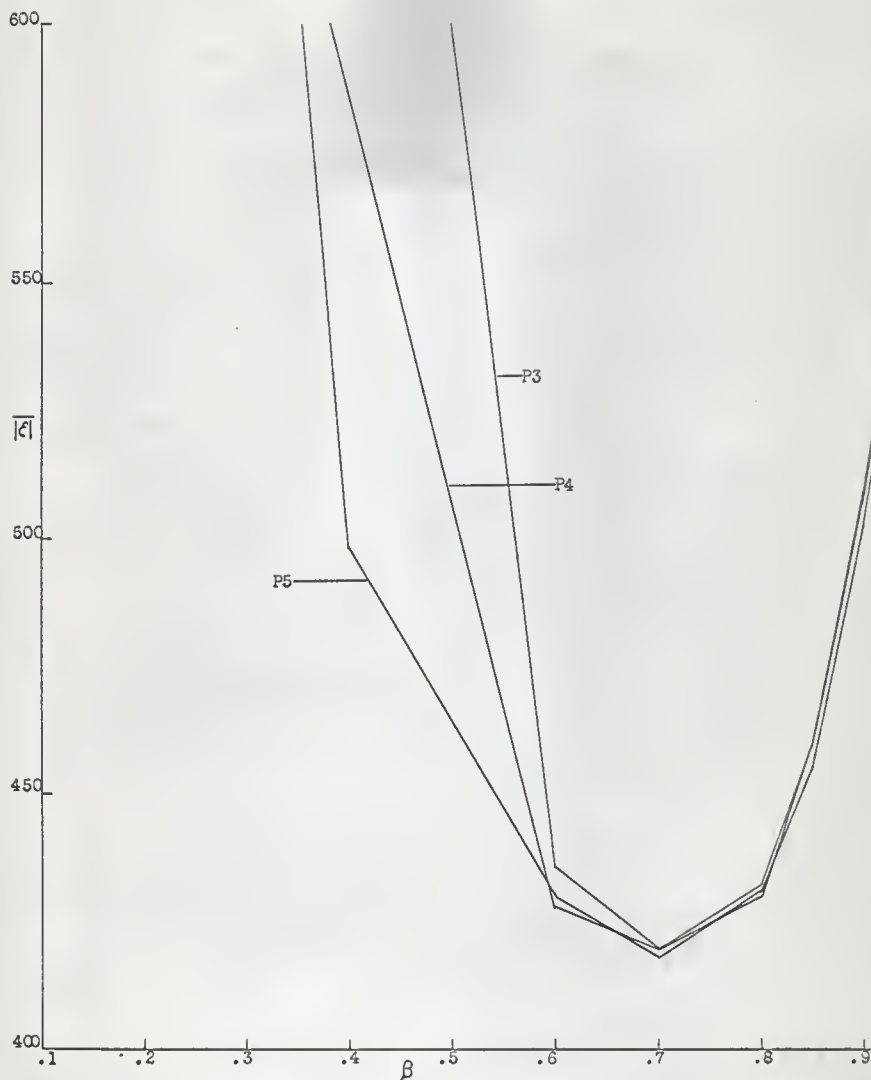


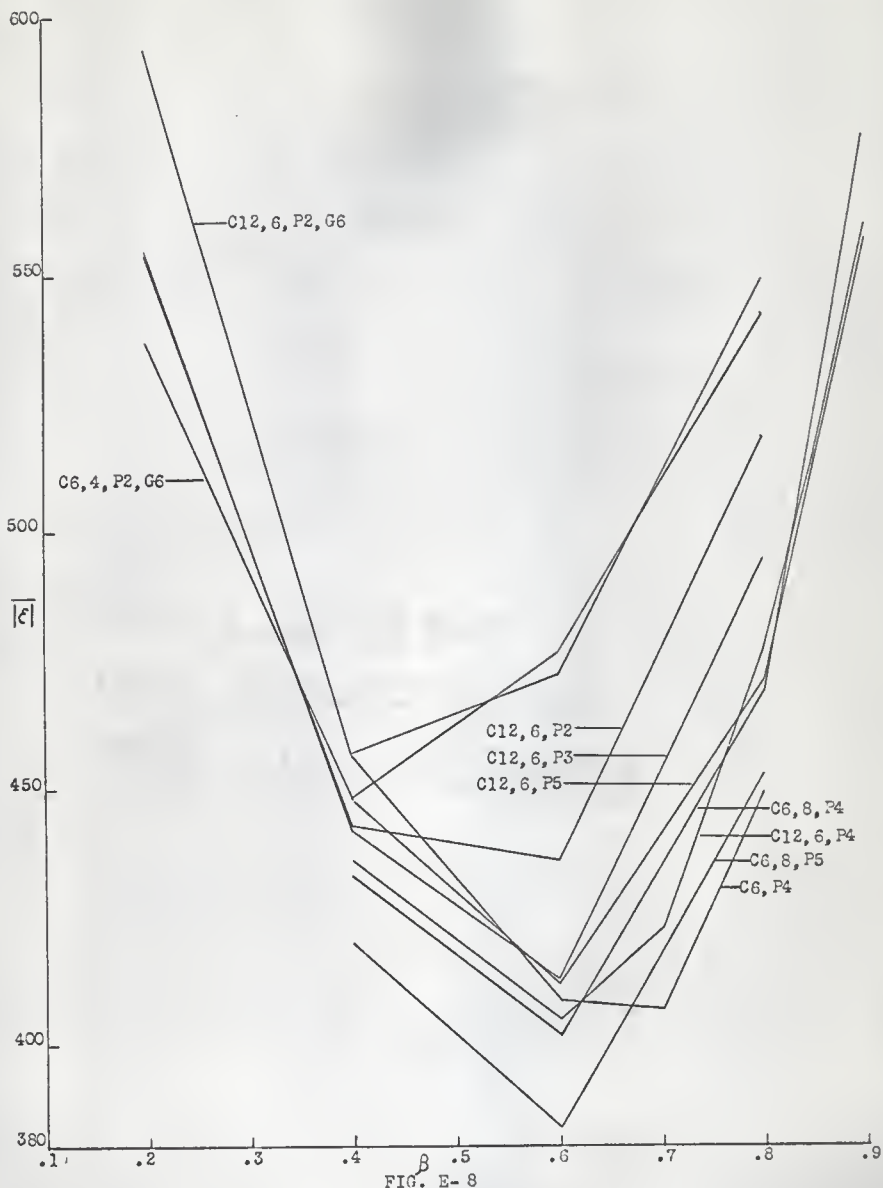
FIG. E-6

Mean Absolute Error Versus Discount Factor for Wholesale  
Chrysanthemum Data--Polynomial Models



FIG. E-7

Standard Deviation of Error Versus Discount Factor for  
Wholesale Chrysanthemum Data--Polynomial Models



Mean Absolute Error Versus Discount Factor for Wholesale Chrysanthemum Data--Combination Models

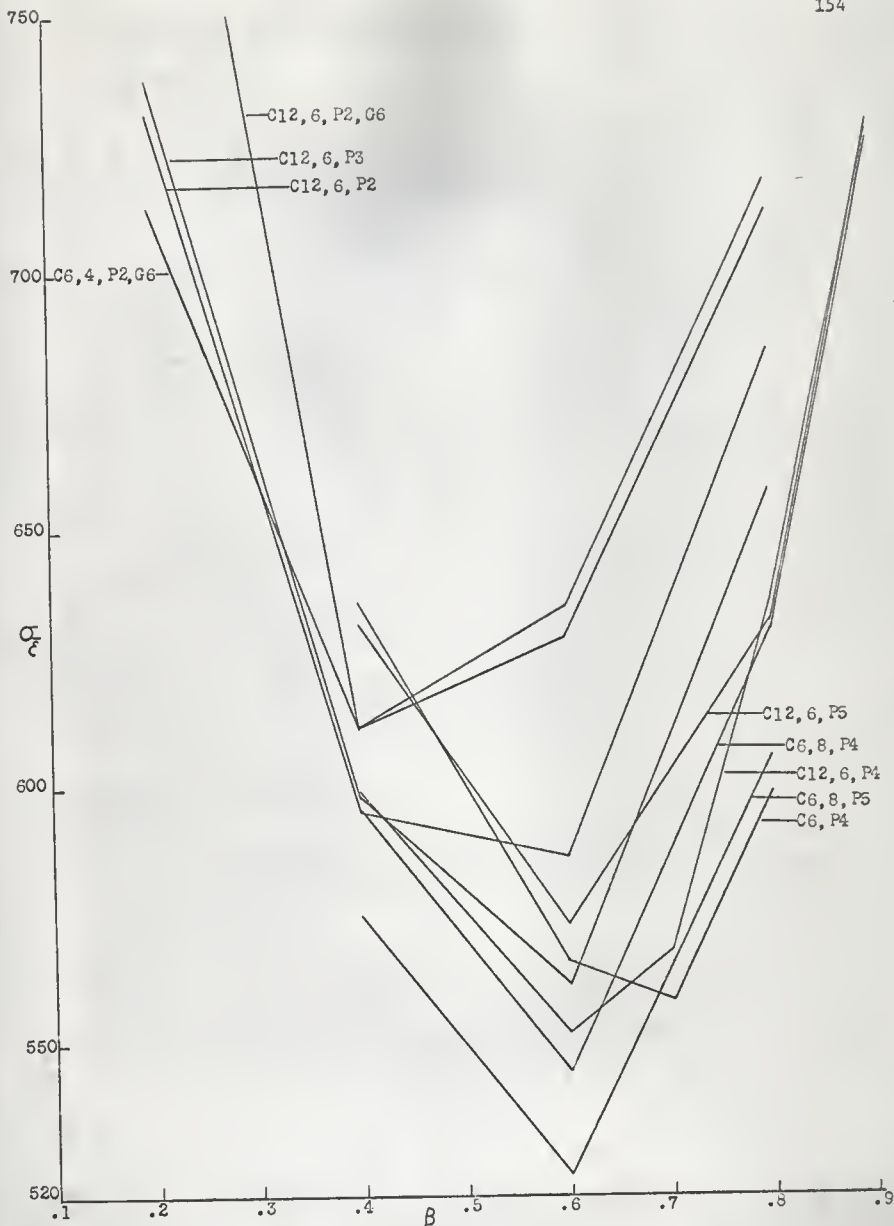


FIG. E-9  
Standard Deviation of Error Versus Discount Factor for  
Wholesale Chrysanthemum Data--Combination  
Models

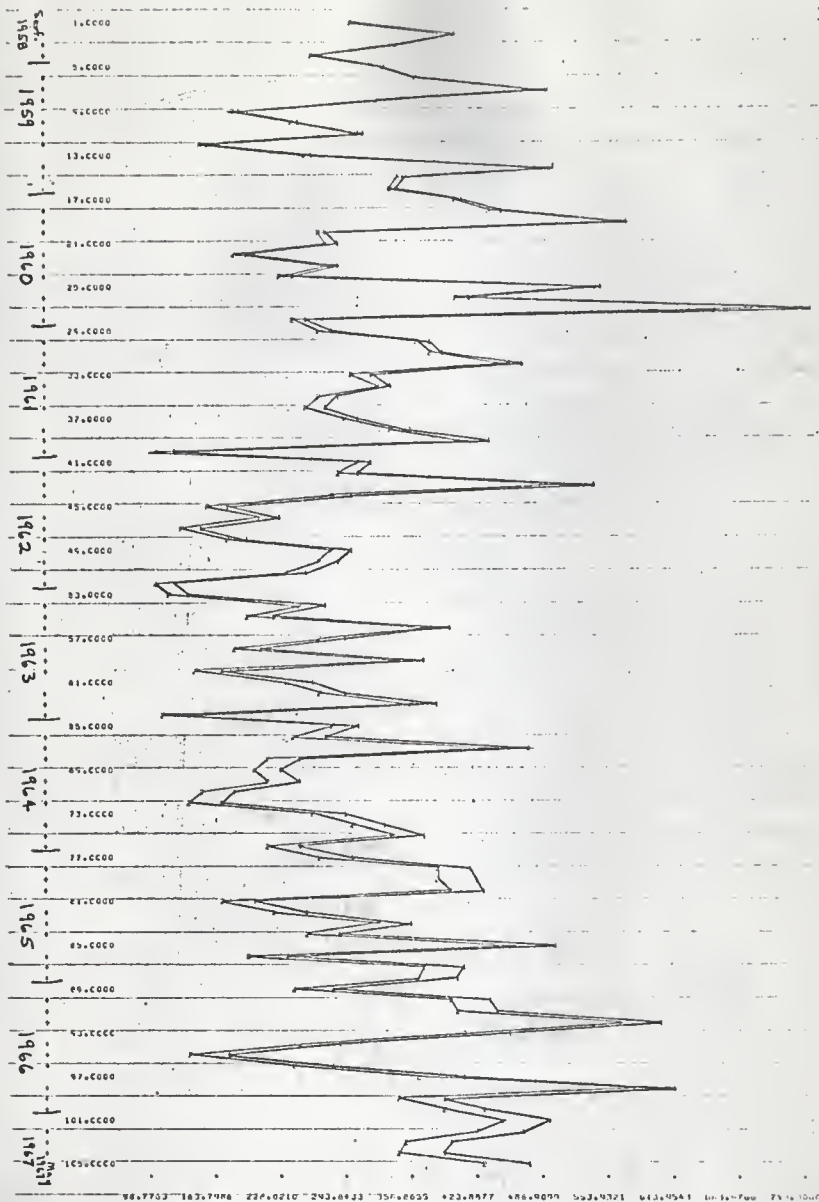


FIG. E-10

Retail Monthly Chrysanthemum Receipts Time Series Plot

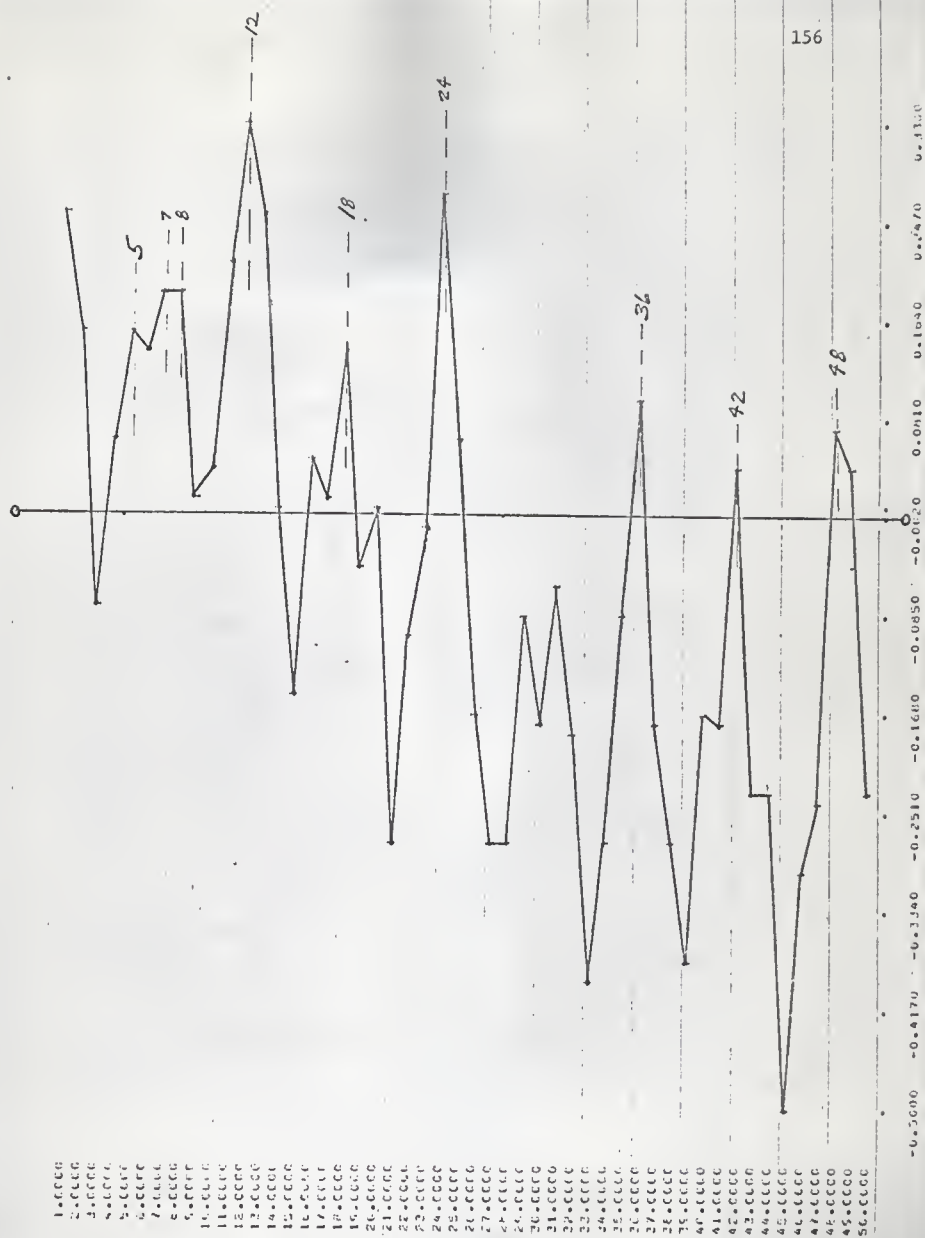


FIG. E-11

Autocorrelation Function for Retail Chrysanthemum Receipts Data

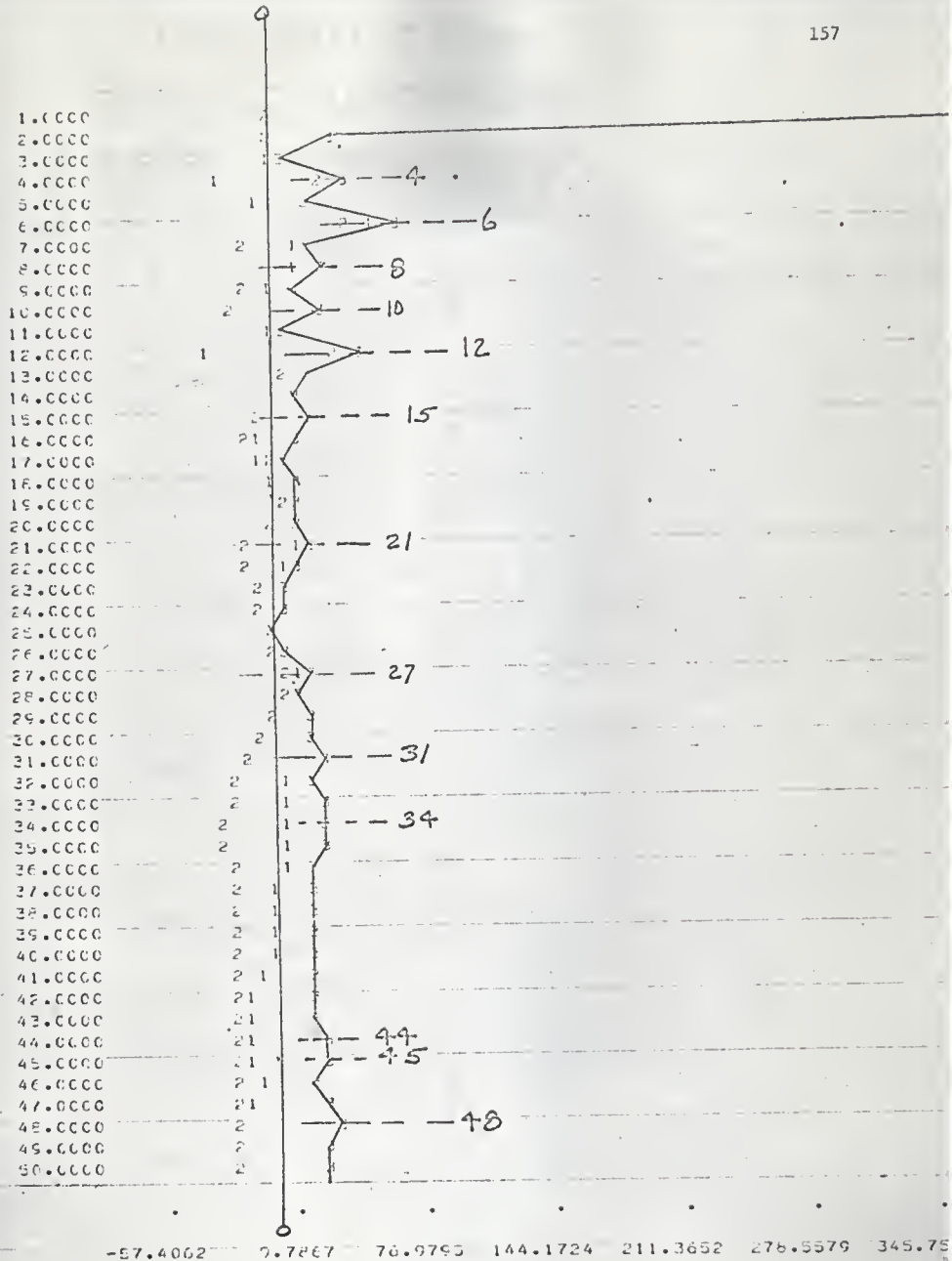


FIG. E-12

Periodogram for Retail Chrysanthemum Receipts Data

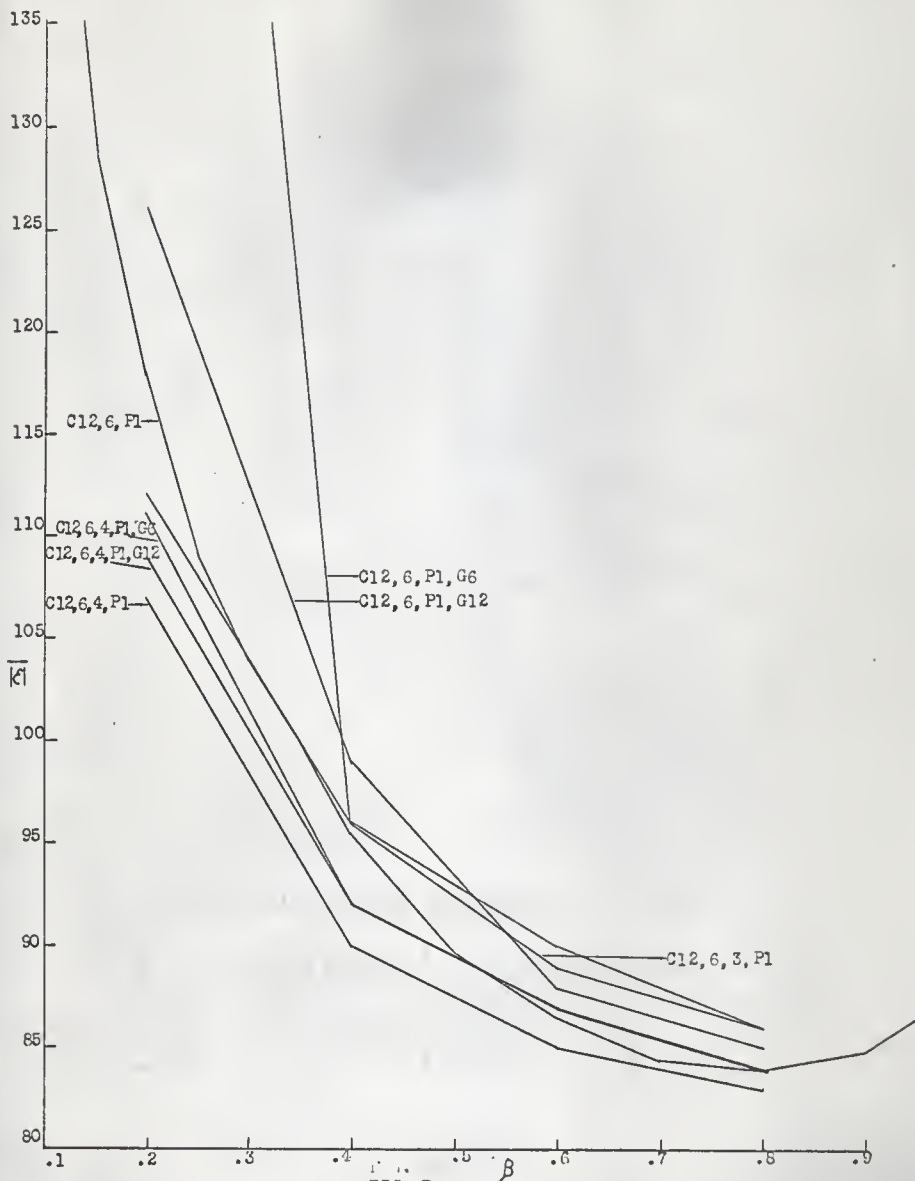


FIG. E-13.

Mean Absolute Error Versus Discount Factor for Retail  
Chrysanthemum Data--Cyclic Models



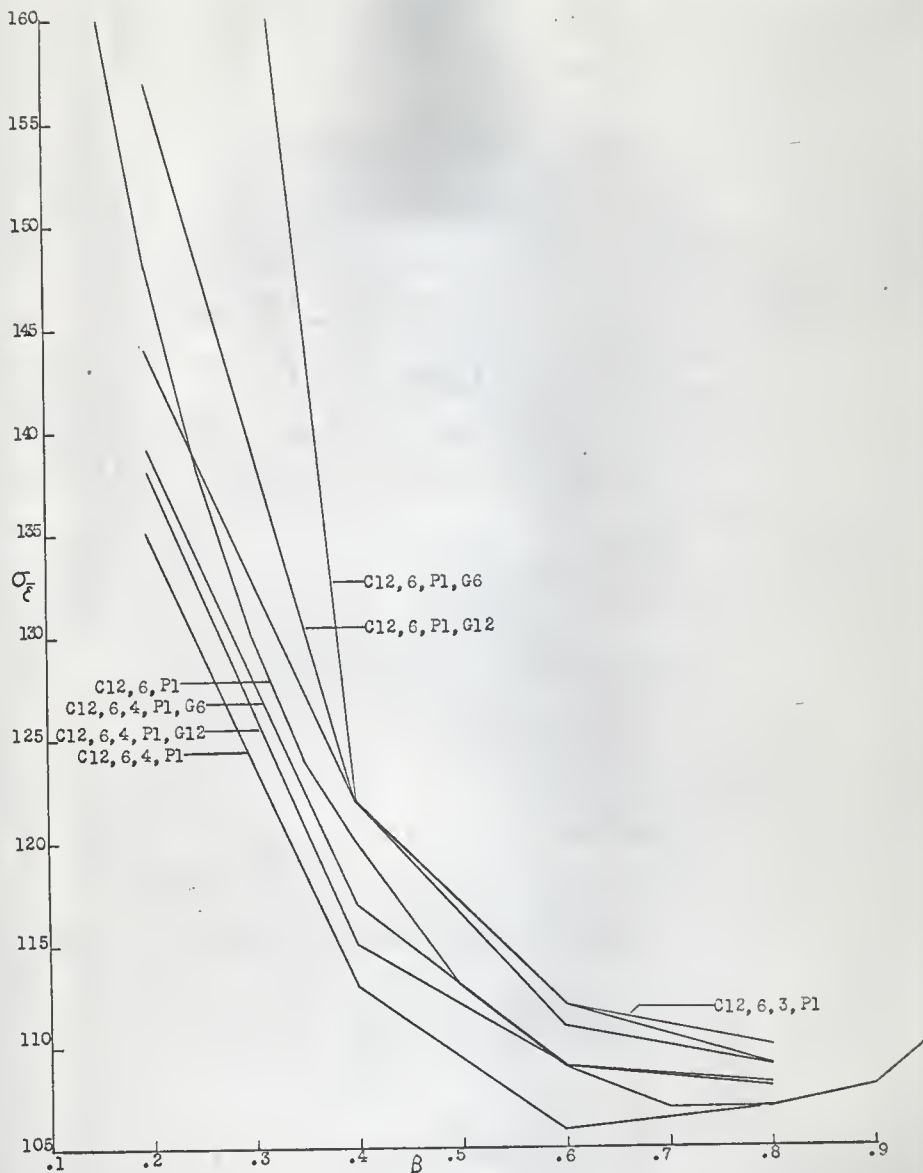


FIG. E-14

Standard Deviation of Error Versus Discount Factor for  
Retail Chrysanthemum Data--Cyclic Models



FIG. E-15

Mean Absolute Error Versus Discount Factor for Retail  
Chrysanthemum Data--Polynomial Models



FIG. E-16

Standard Deviation of Error Versus Discount Factor for  
Retail Chrysanthemum Data--Polynomial Models

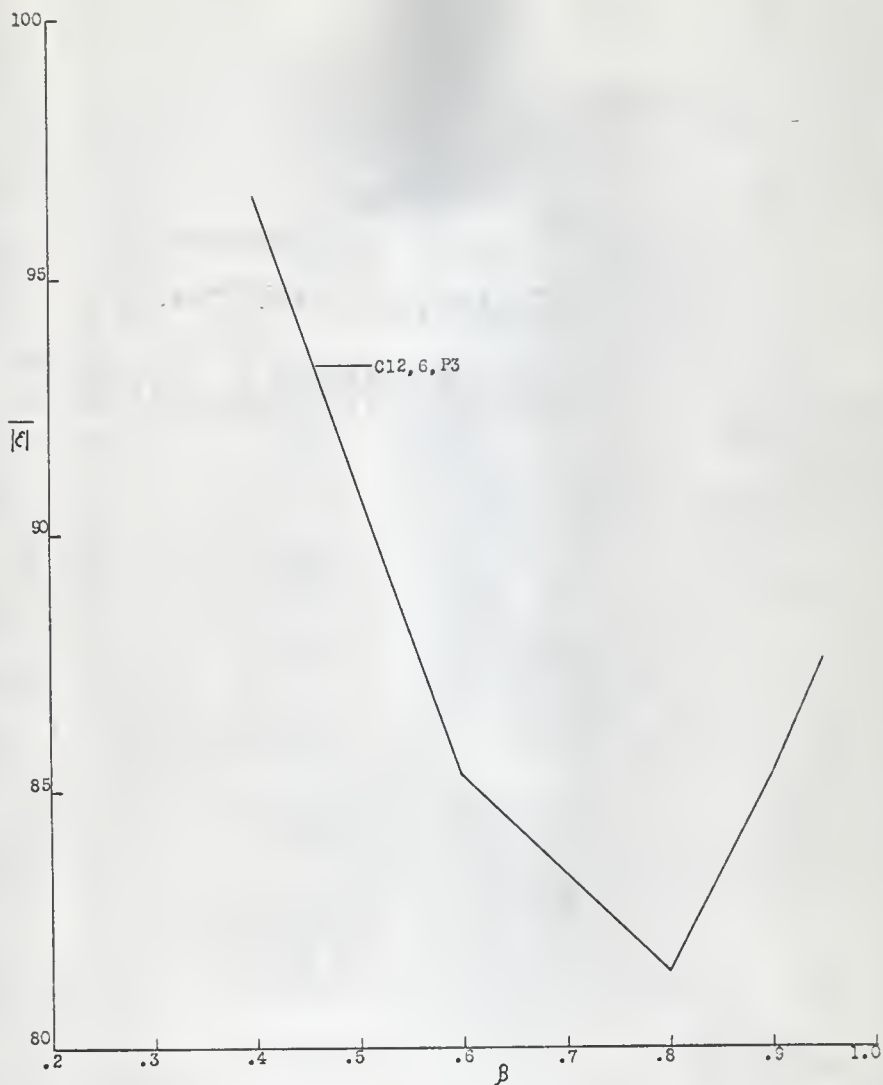


FIG. E-17

Mean Absolute Error Versus Discount Factor  
for Retail Chrysanthemum Data--Combination Model

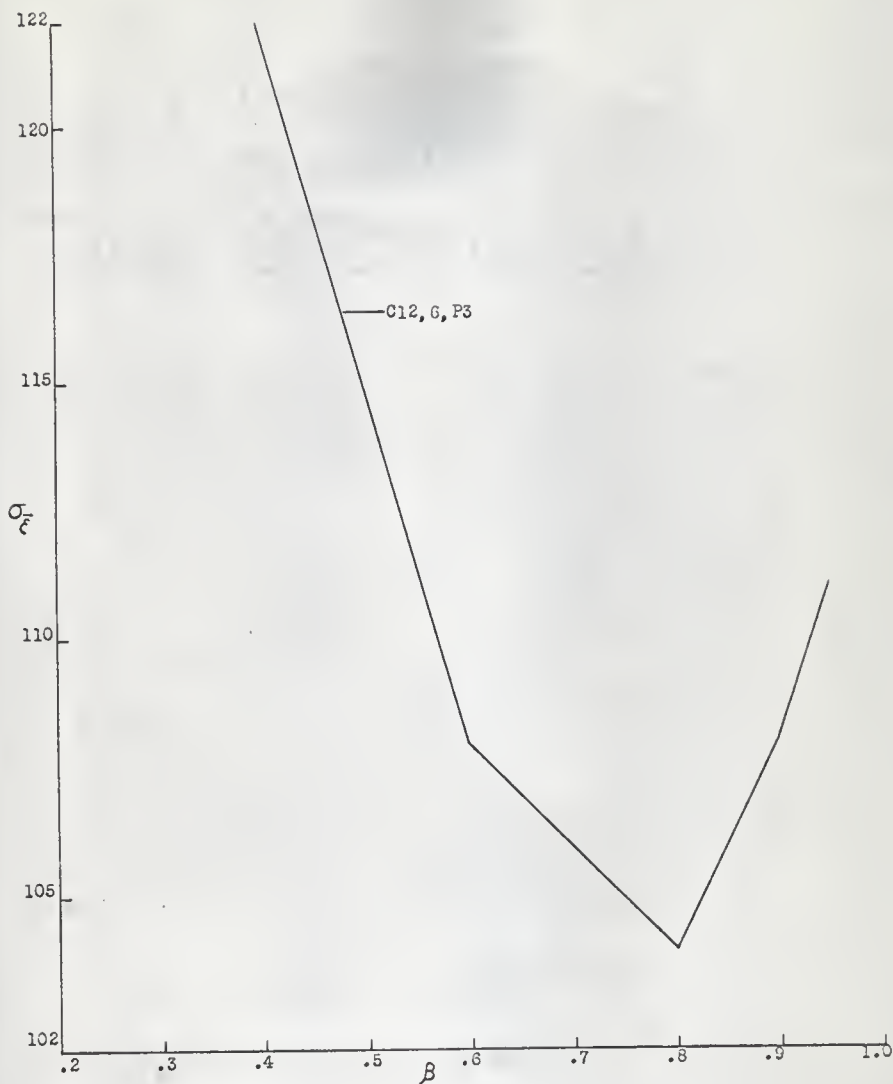


FIG. E-18

Standard Deviation of Error Versus Discount Factor  
for  
Retail Chrysanthemum Data--Combination Models

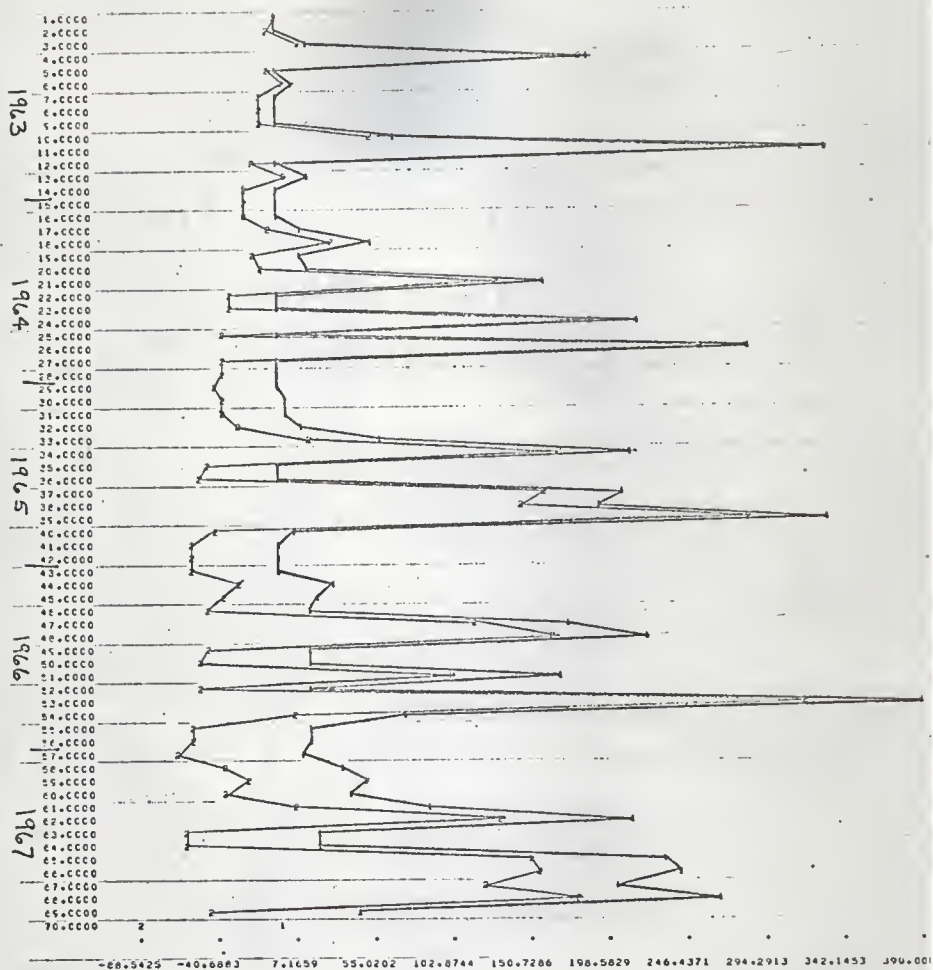


FIG. E-19

Lily Daily Sales Time Series Plot

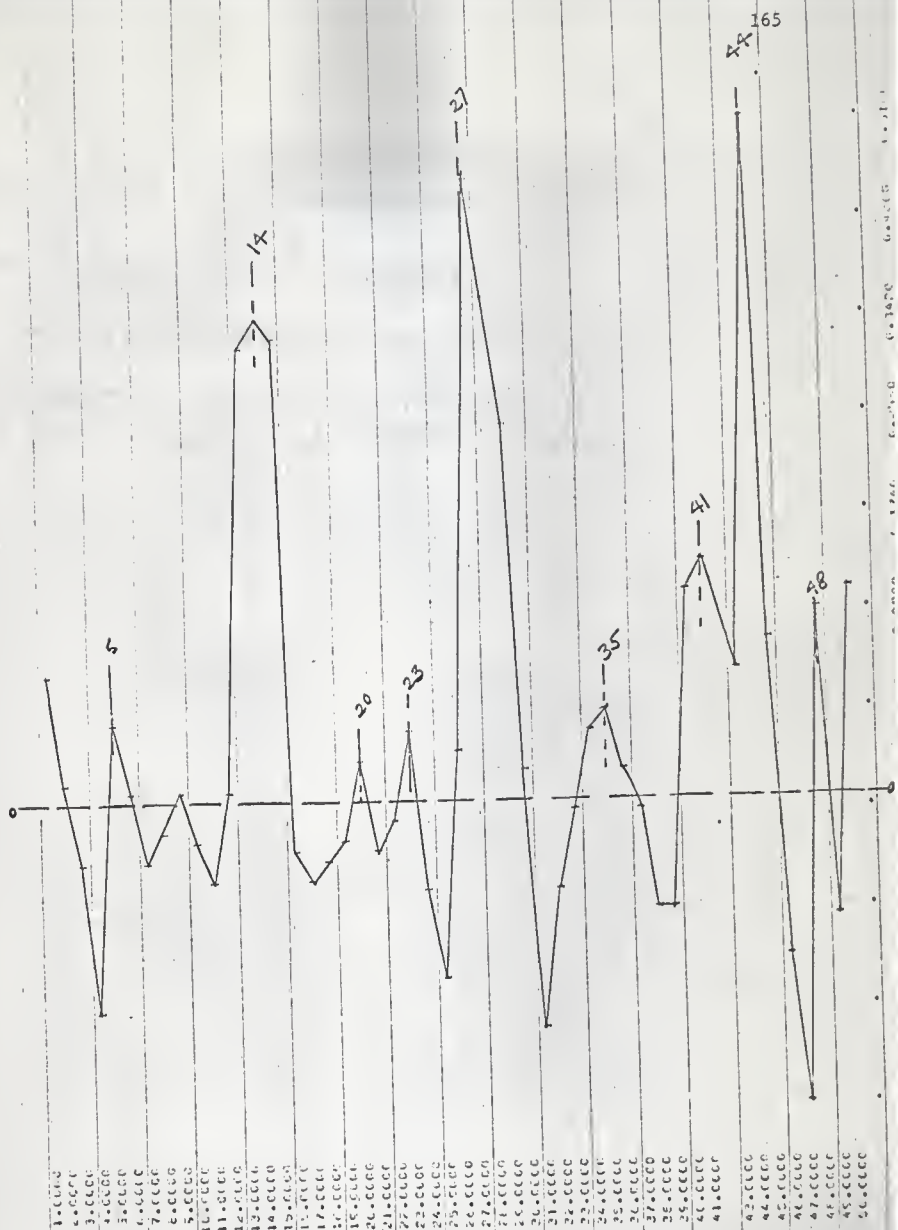


FIG. E-27

Autocorrelation Function for Lily Data

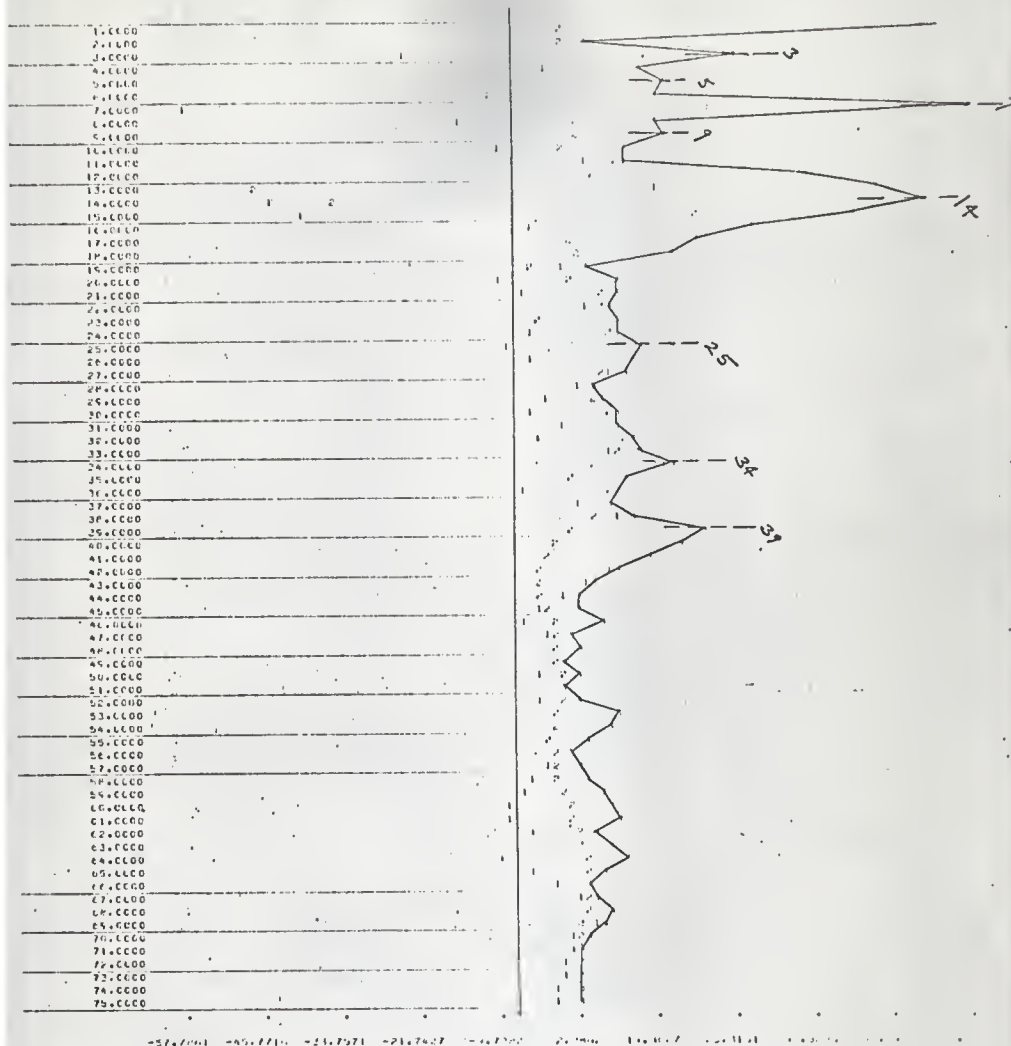
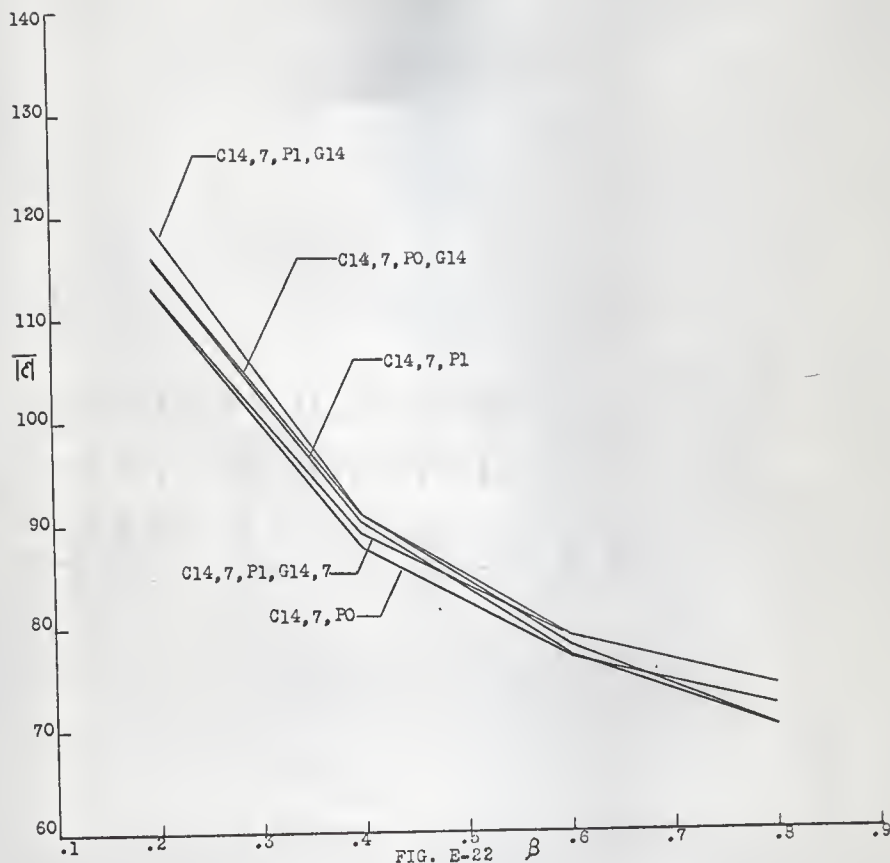


FIG. E-21

Periodogram for Lily Data





Mean Absolute Error Versus Discount Factor for Retail Lily  
Data--Cyclic Models

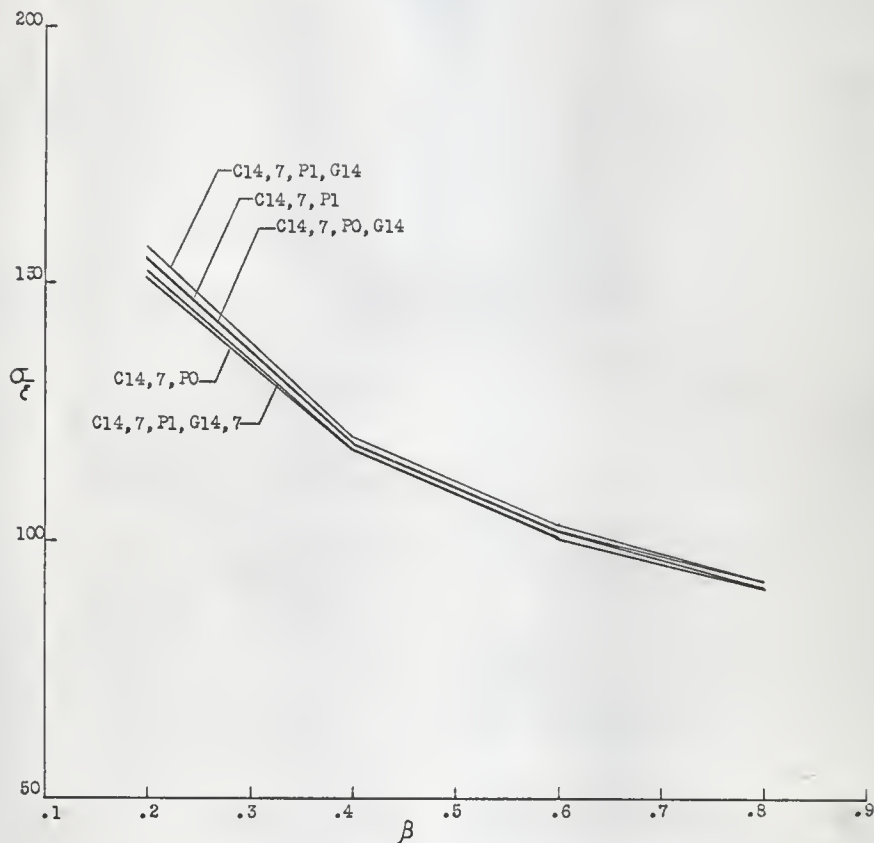


FIG. E-23

Standard Deviation of Error Versus Discount Factor for  
Retail Lily Data--Cyclic Models

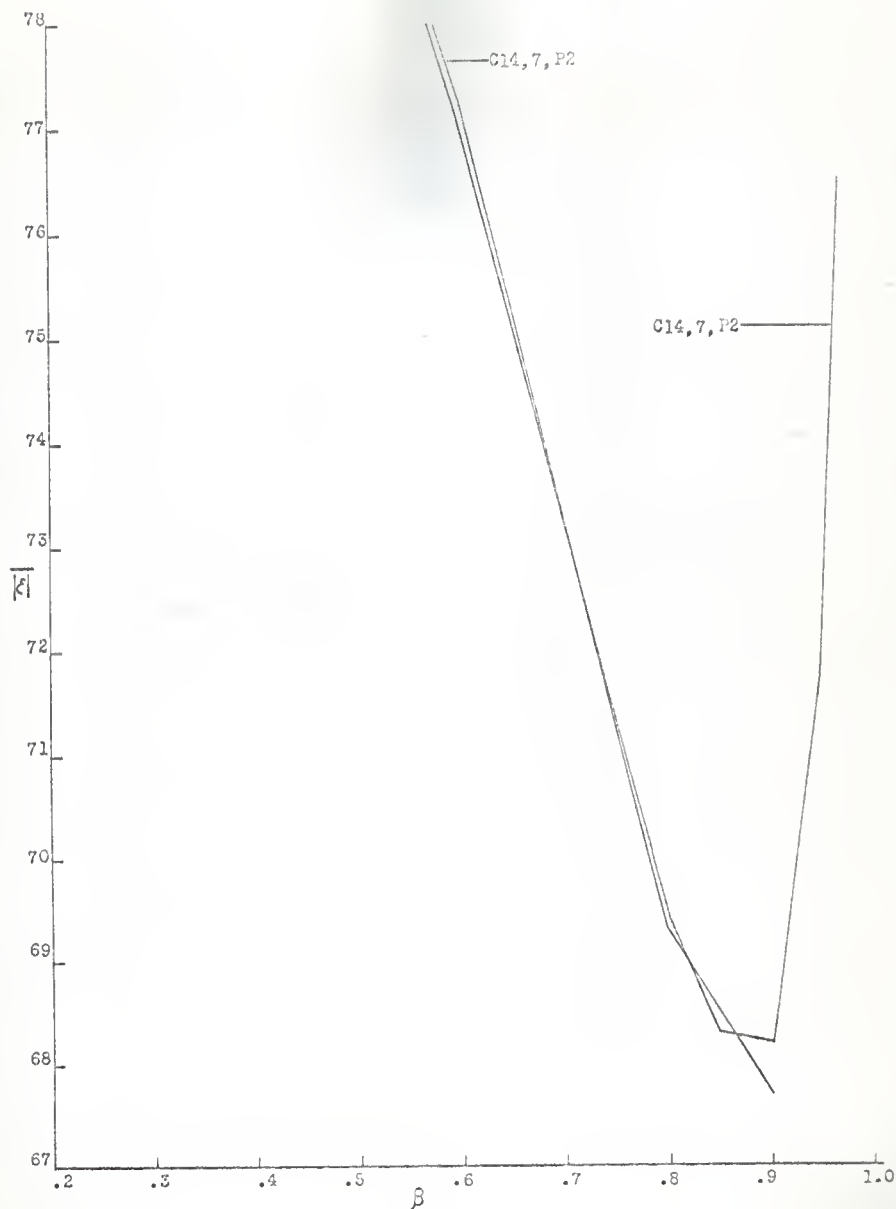


FIG. D-24

Mean Absolute Error Versus Discount Factor  
for Lily Data--Combination Models

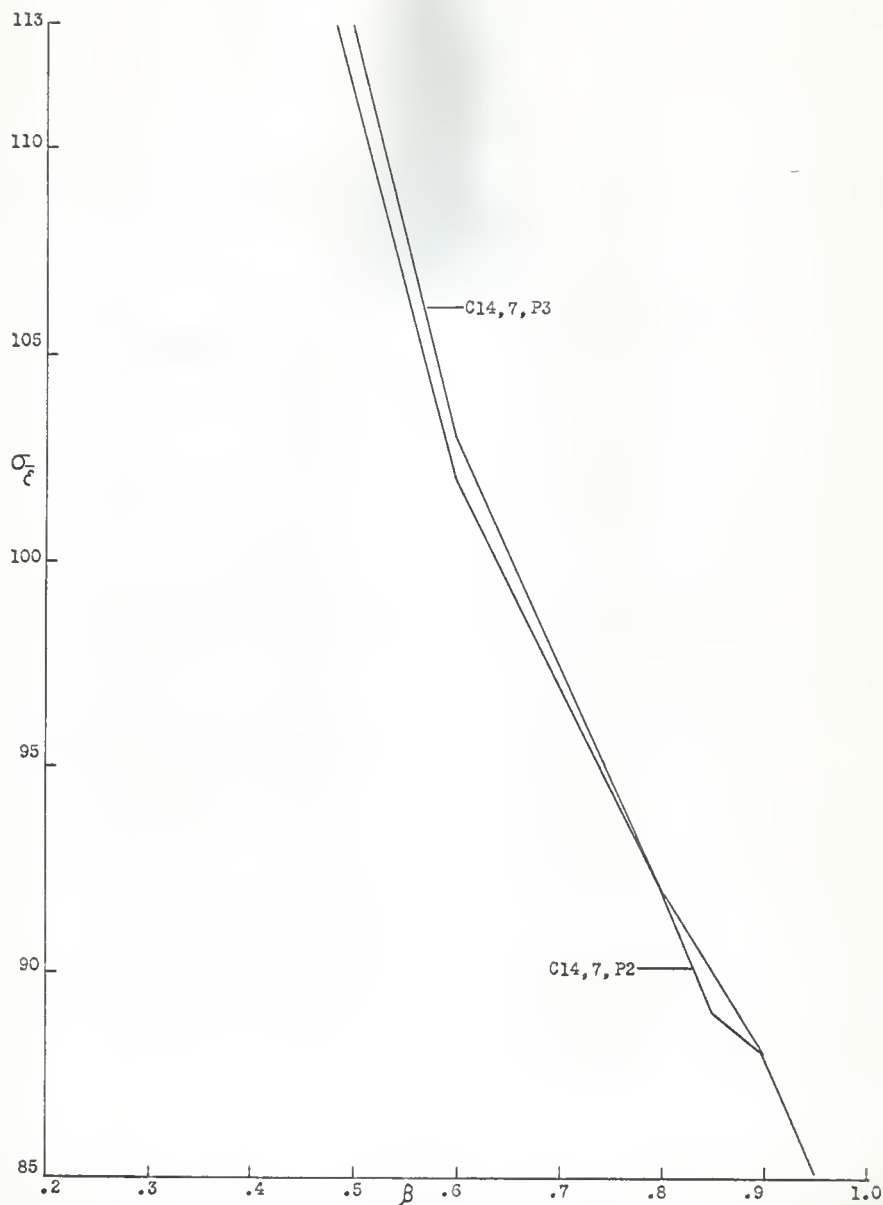


FIG. E-25  
Standard Deviation of Error Versus Discount  
Factor for Lily Data--Combination Models

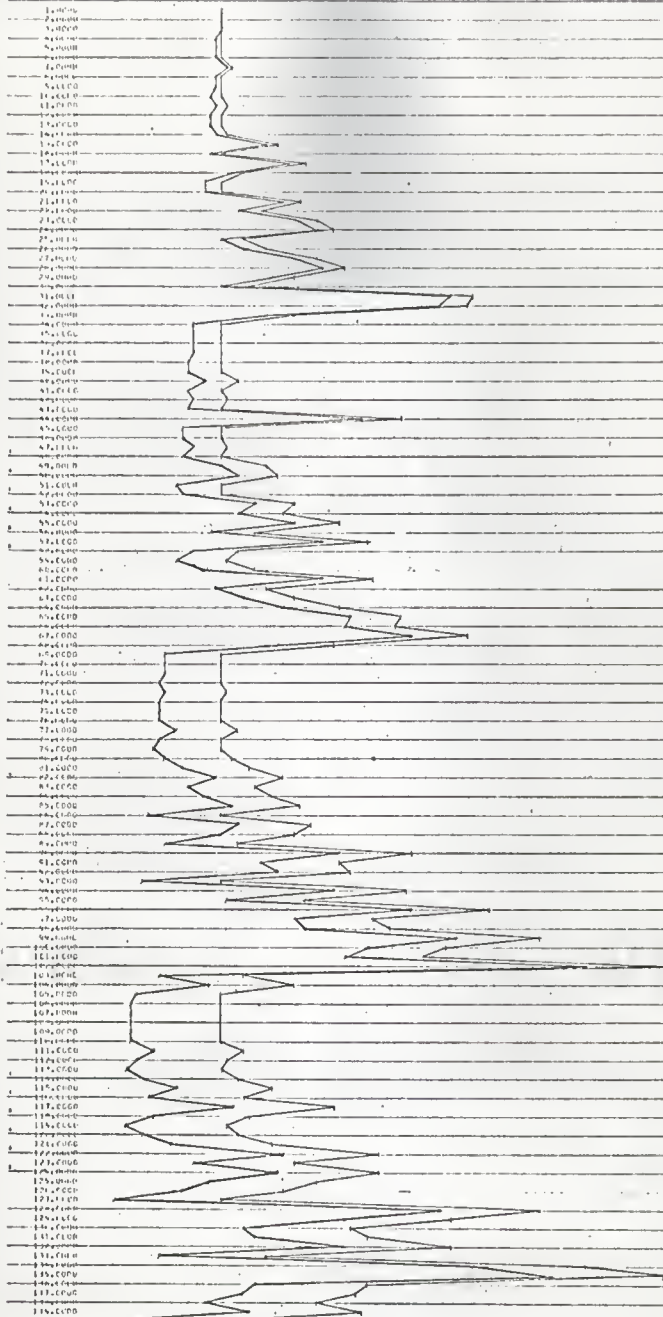


FIG. E-26  
Poinsettias Daily Sales Time Series Plot

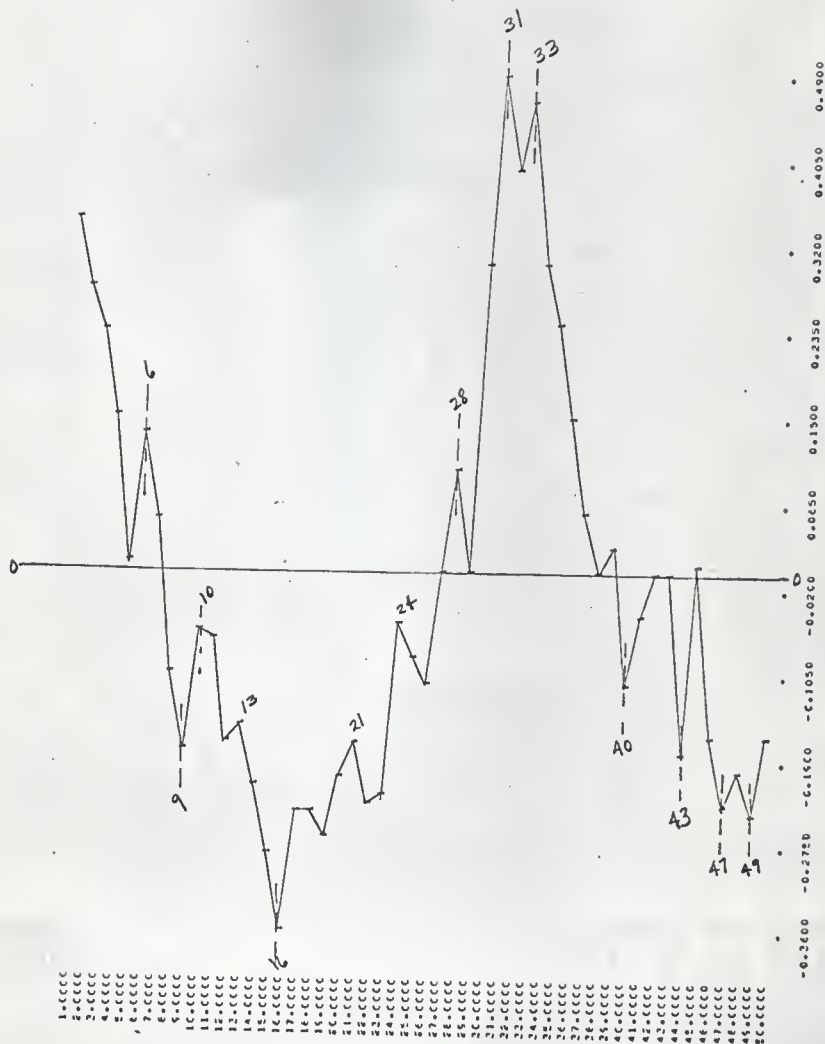


FIG. E-27

Autocorrelation Function for Peinsettia Data

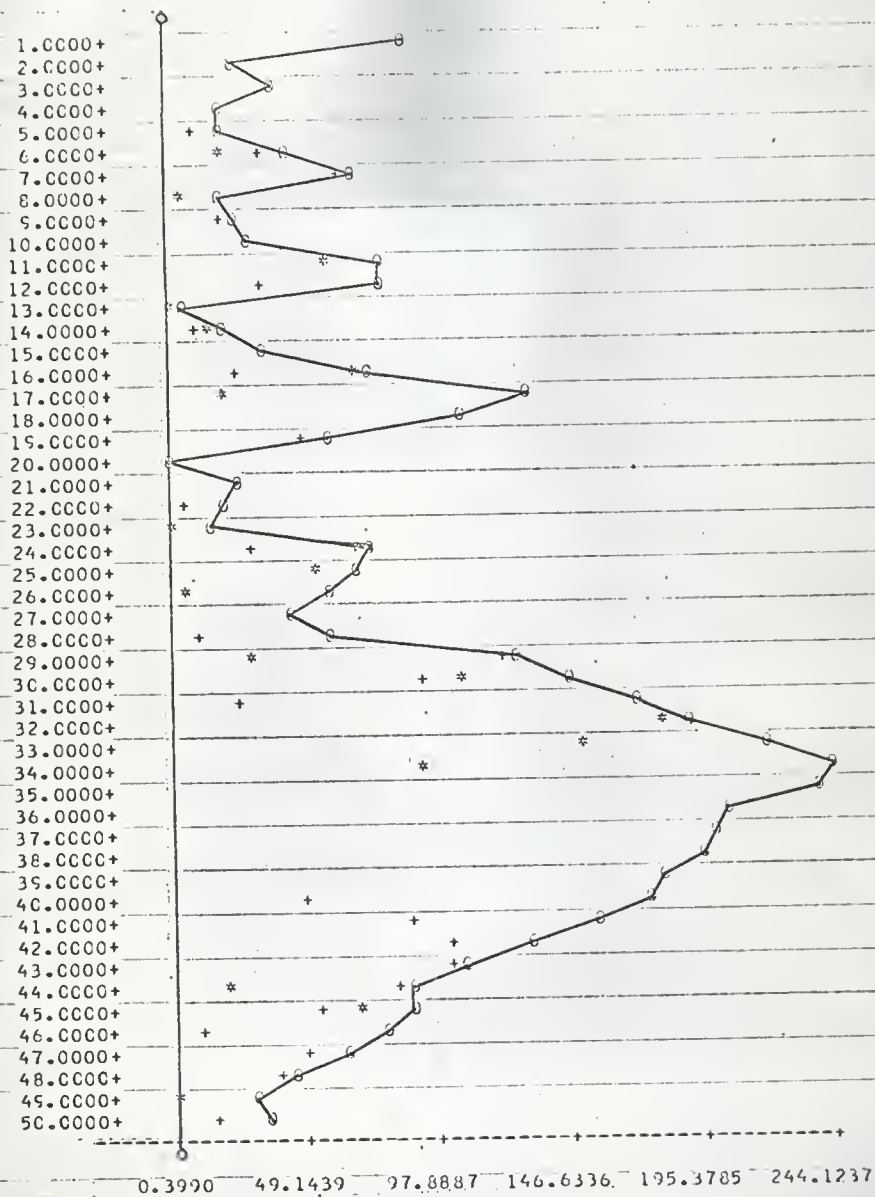


FIG. E-28

Periodogram for Poinsettia Data

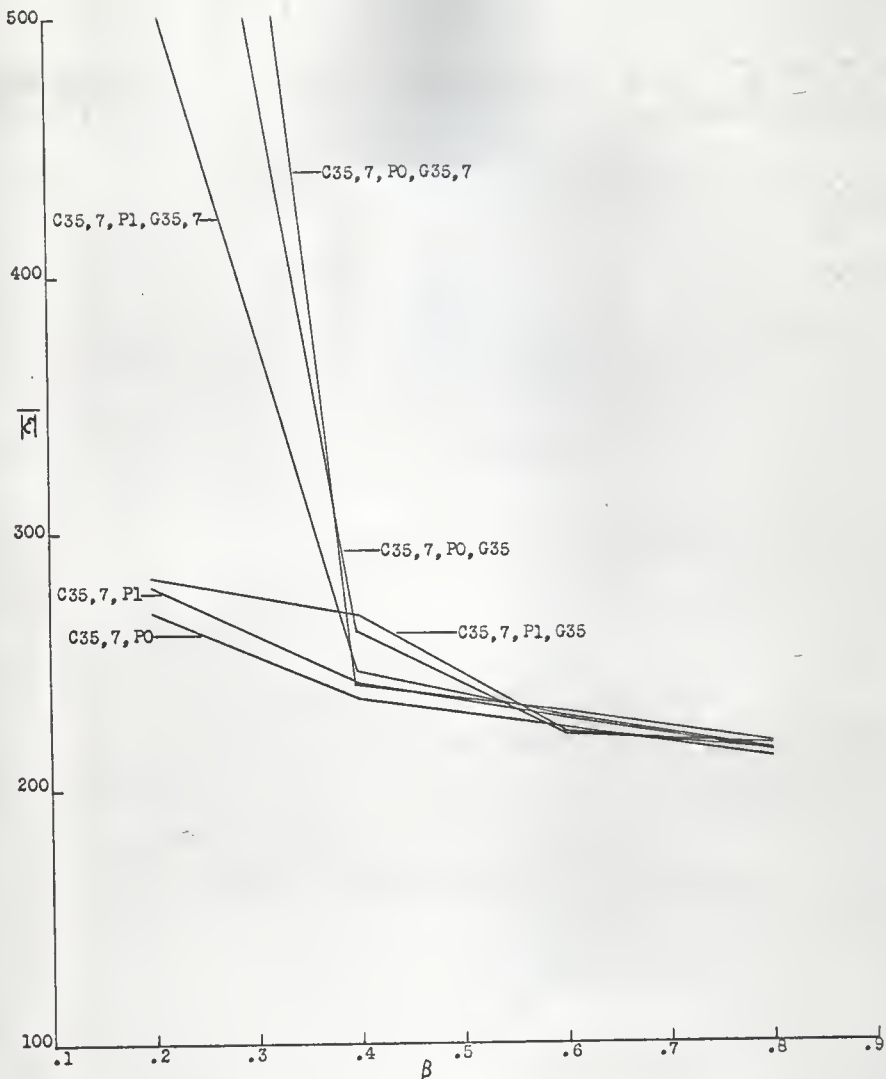


FIG. E-29

Mean Absolute Error Versus Discount Factor for Poinsettia  
Data---Cyclic Models



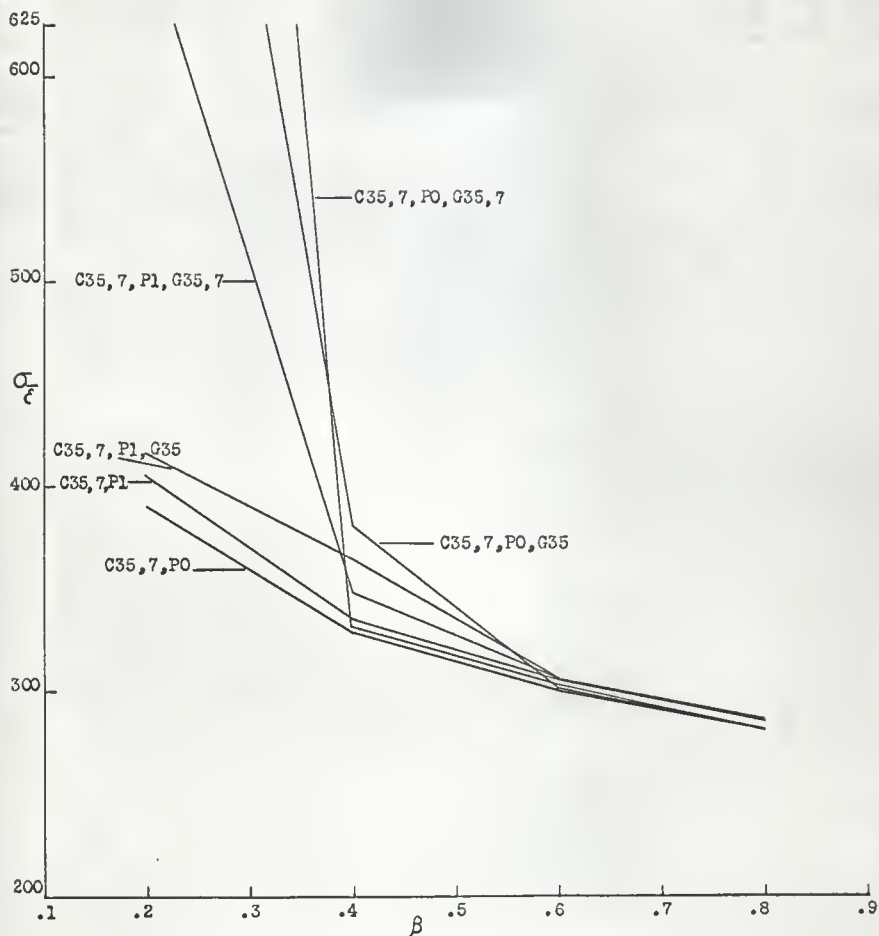


FIG. E-30

Standard Deviation of Error Versus Discount Factor for  
Poinsettia Data--Cyclic Models

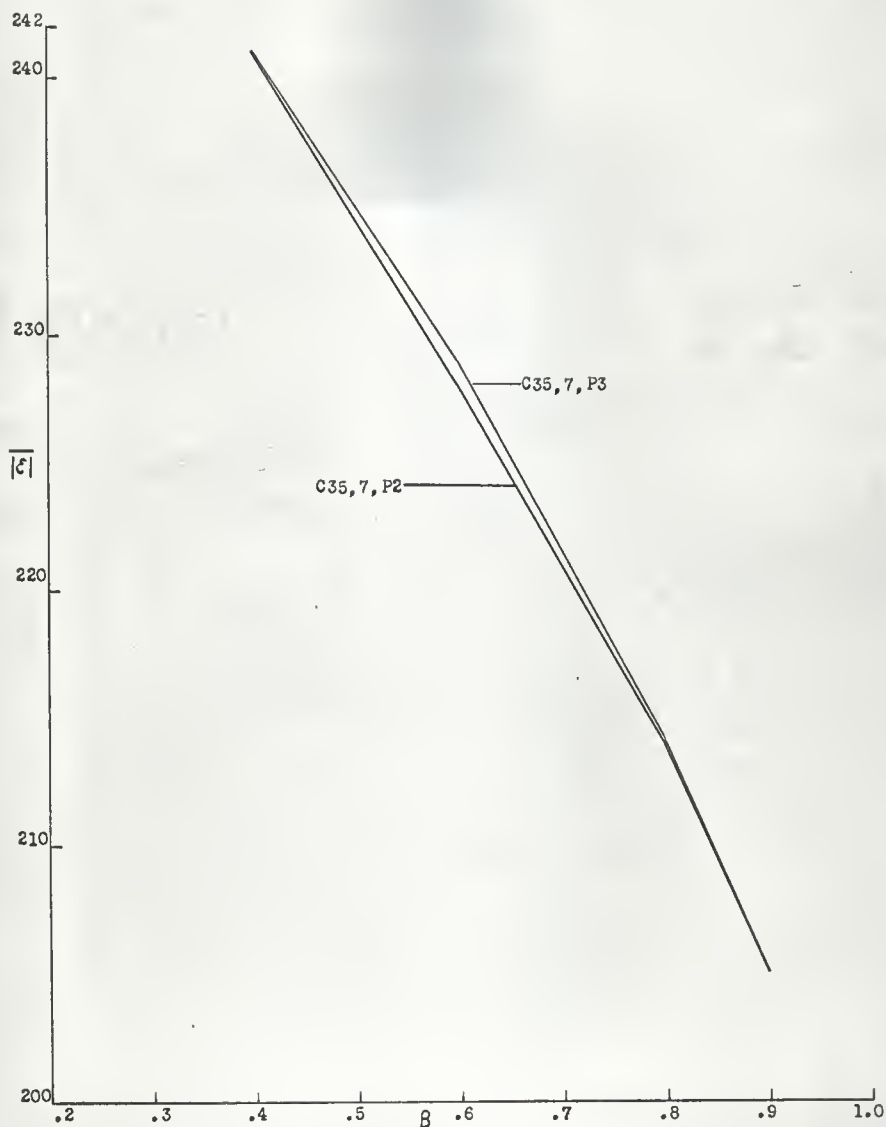


FIG. E-31

Mean Absolute Error Versus Discount Factor  
for Poinsettia Data--Combination Models

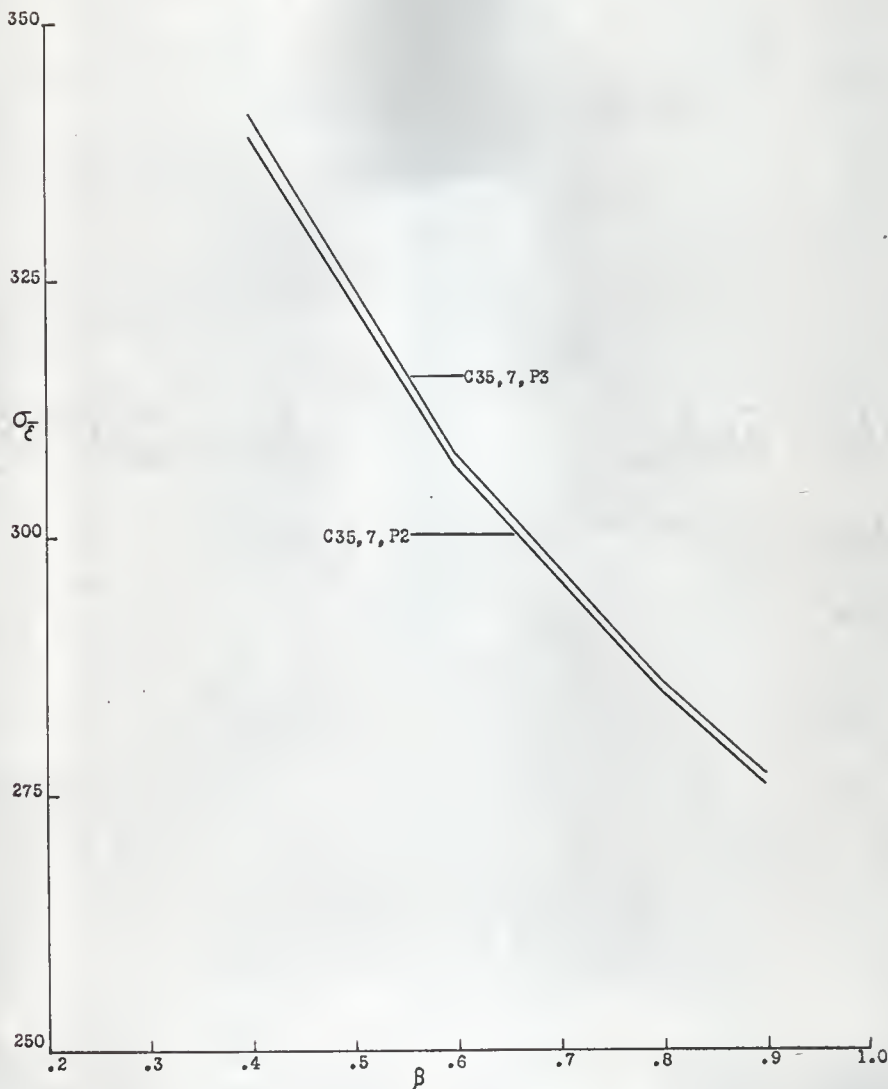


FIG. E-32

Standard Deviation of Error Versus Discount  
Factor for Poinsettia Data--Combination Models

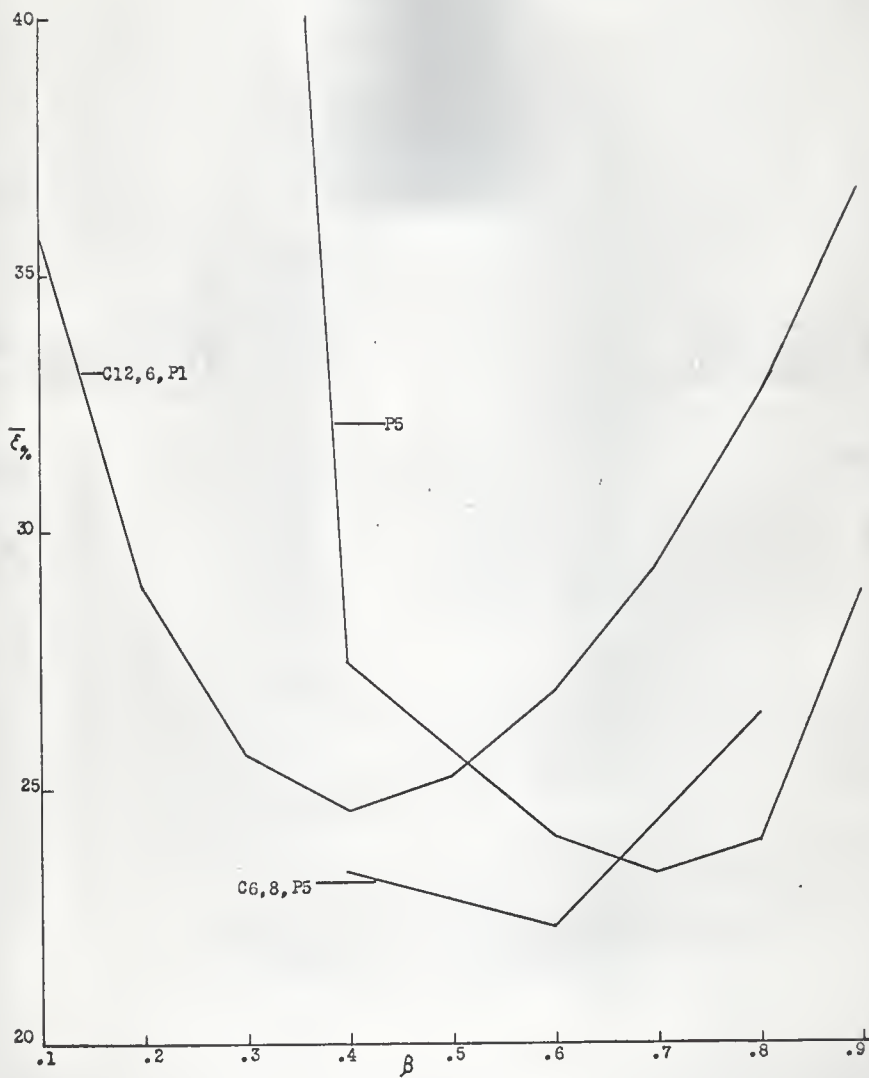


FIG. E-33

Mean Percent Error Versus Discount Factor for  
Best Wholesale Chrysanthemum Time Series Models

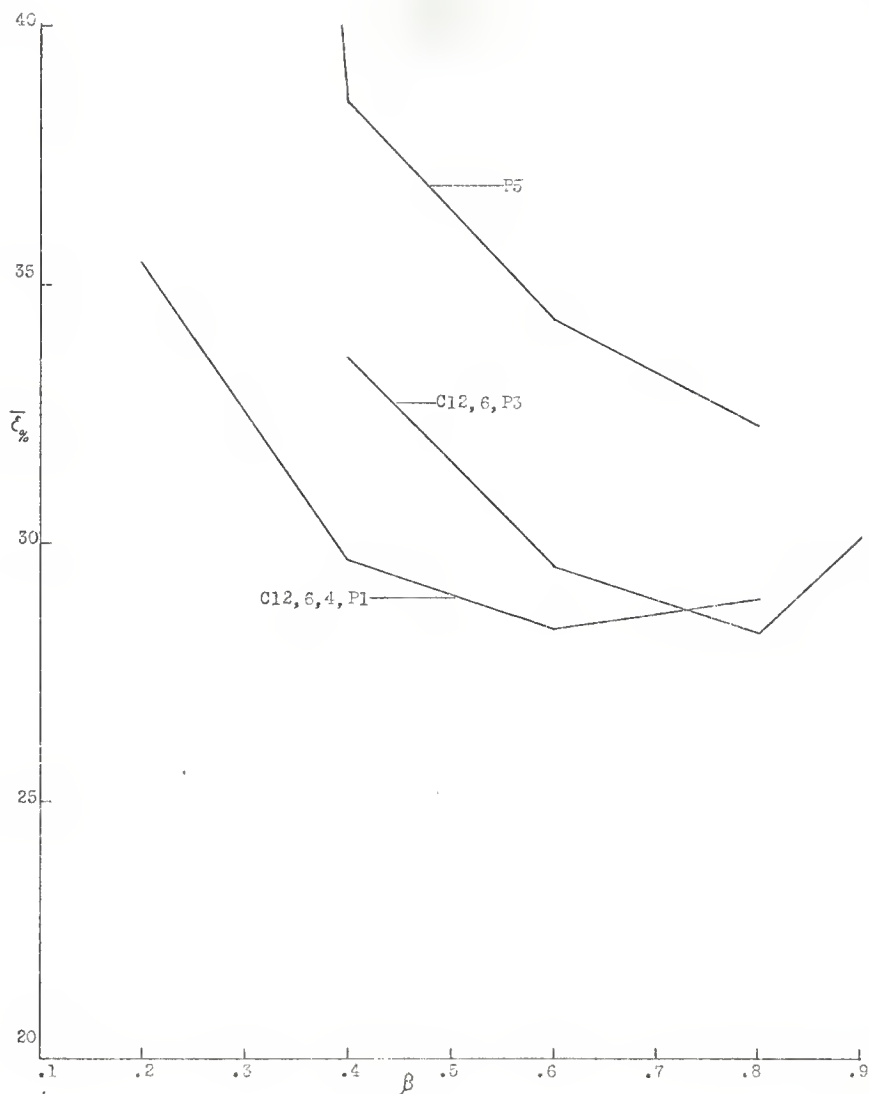


FIG. E-34

Mean Percent Error Versus Discount Factor for  
Best Retail Chrysanthemum Time Series Models

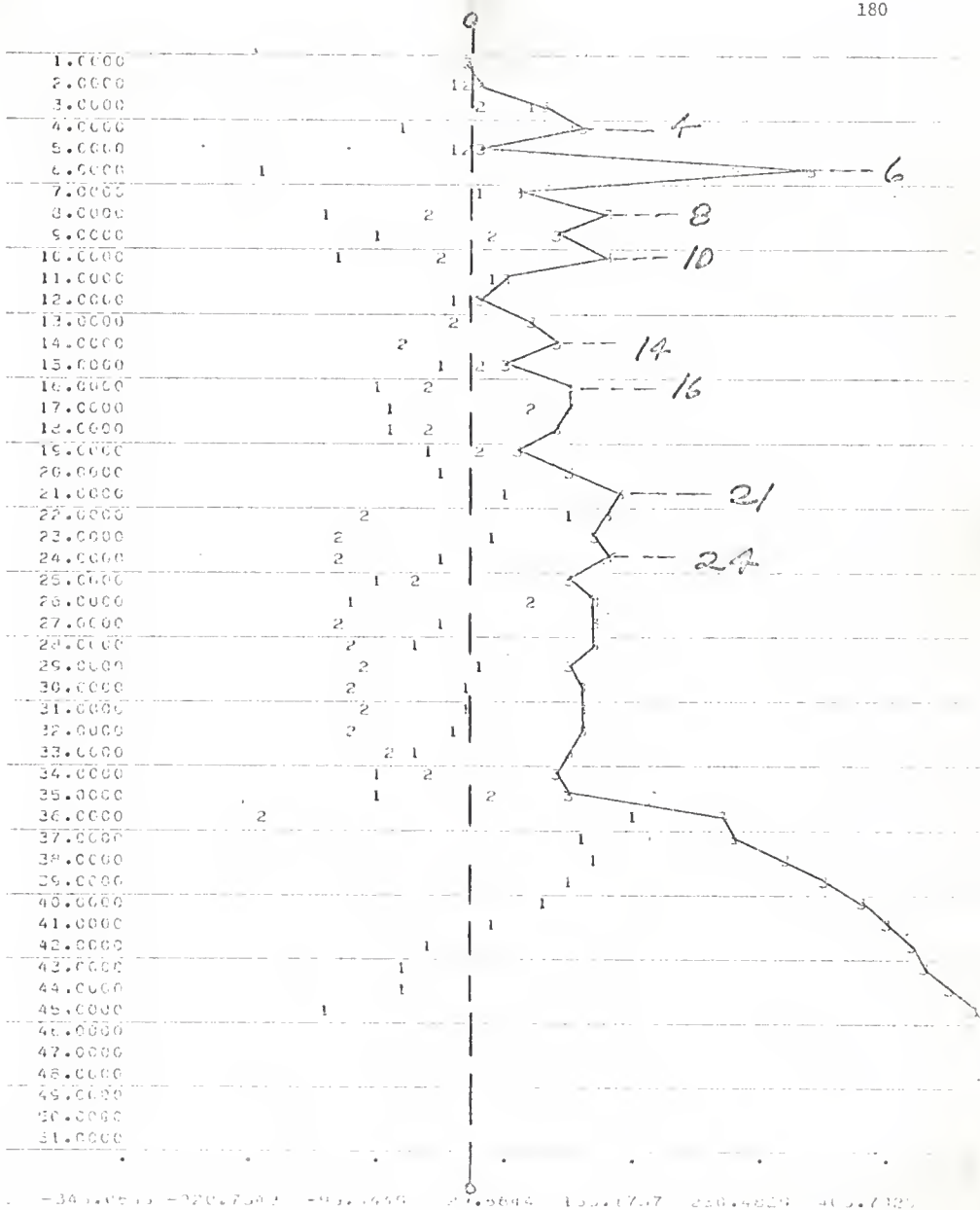
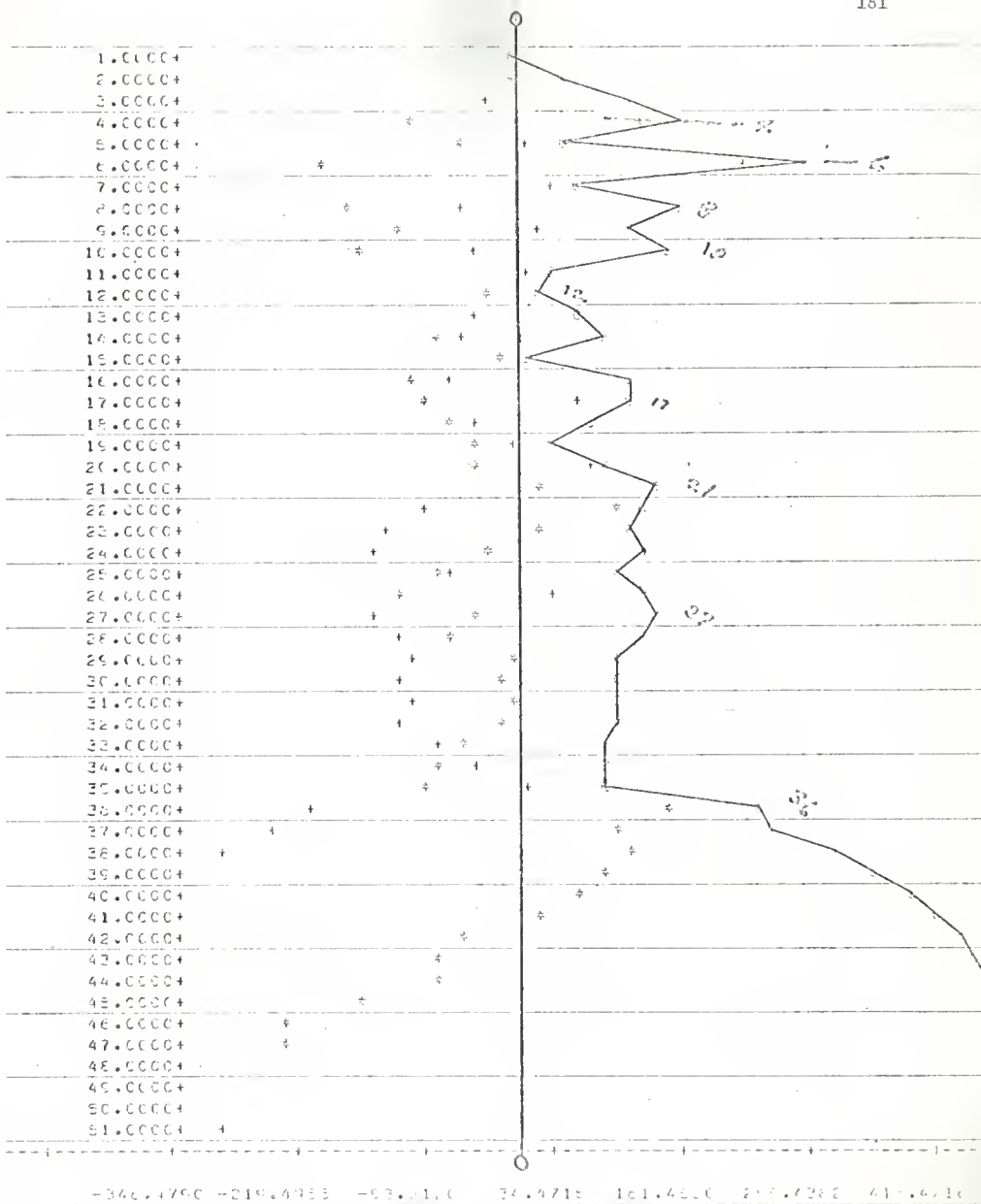


FIG. B-35

Periodogram for Wholesale Chrysanthemum Data Using  
Second Degree Polynomial Regression Residuals



Periodogram for *Chrysanthemum* Data  
Using Third Degree Polynomial Regression Residuals

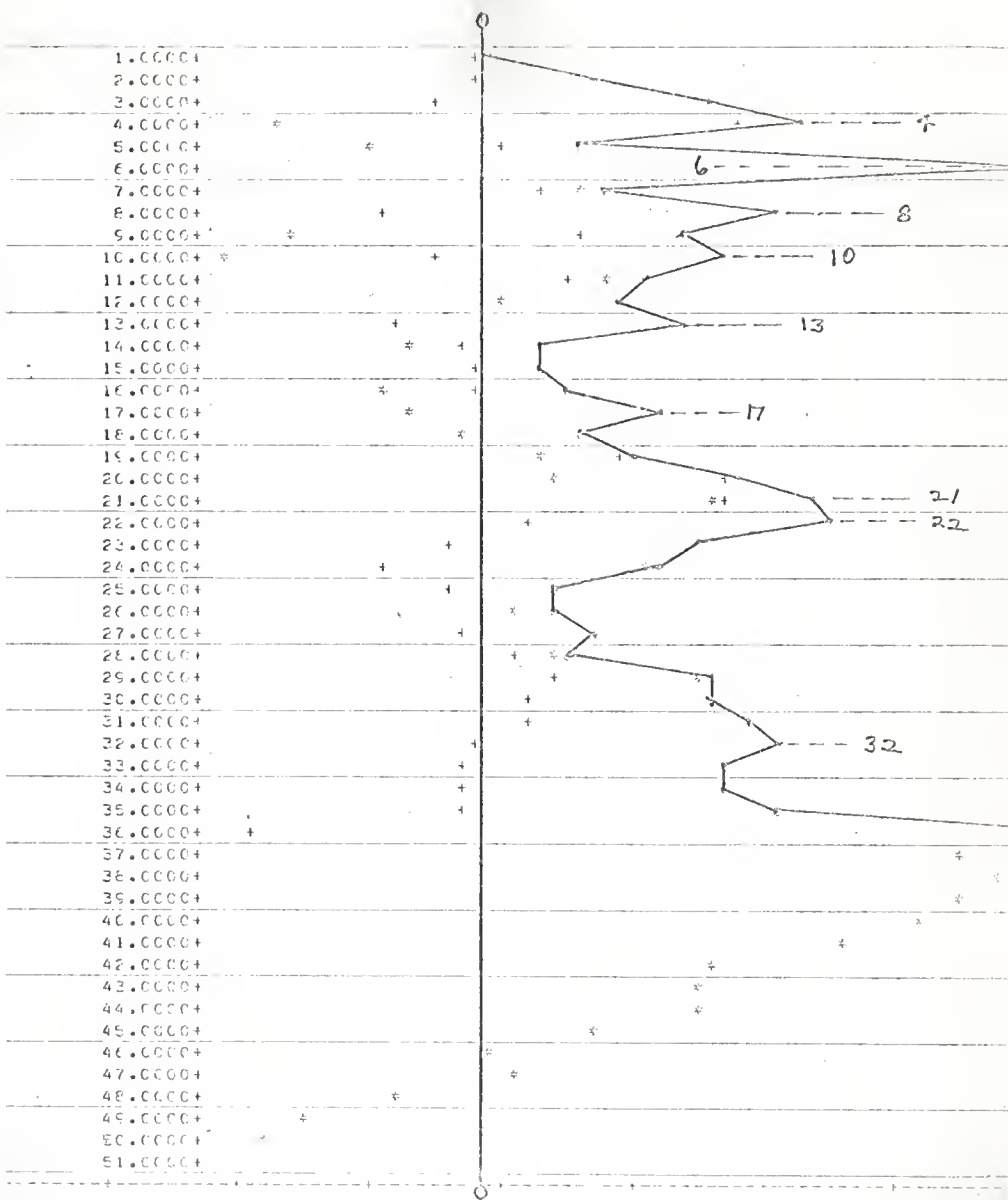


FIG. B- 37

Periodogram for Wholesale Chrysanthemum  
Data Using Fourth Degree Polynomial Regression Residuals



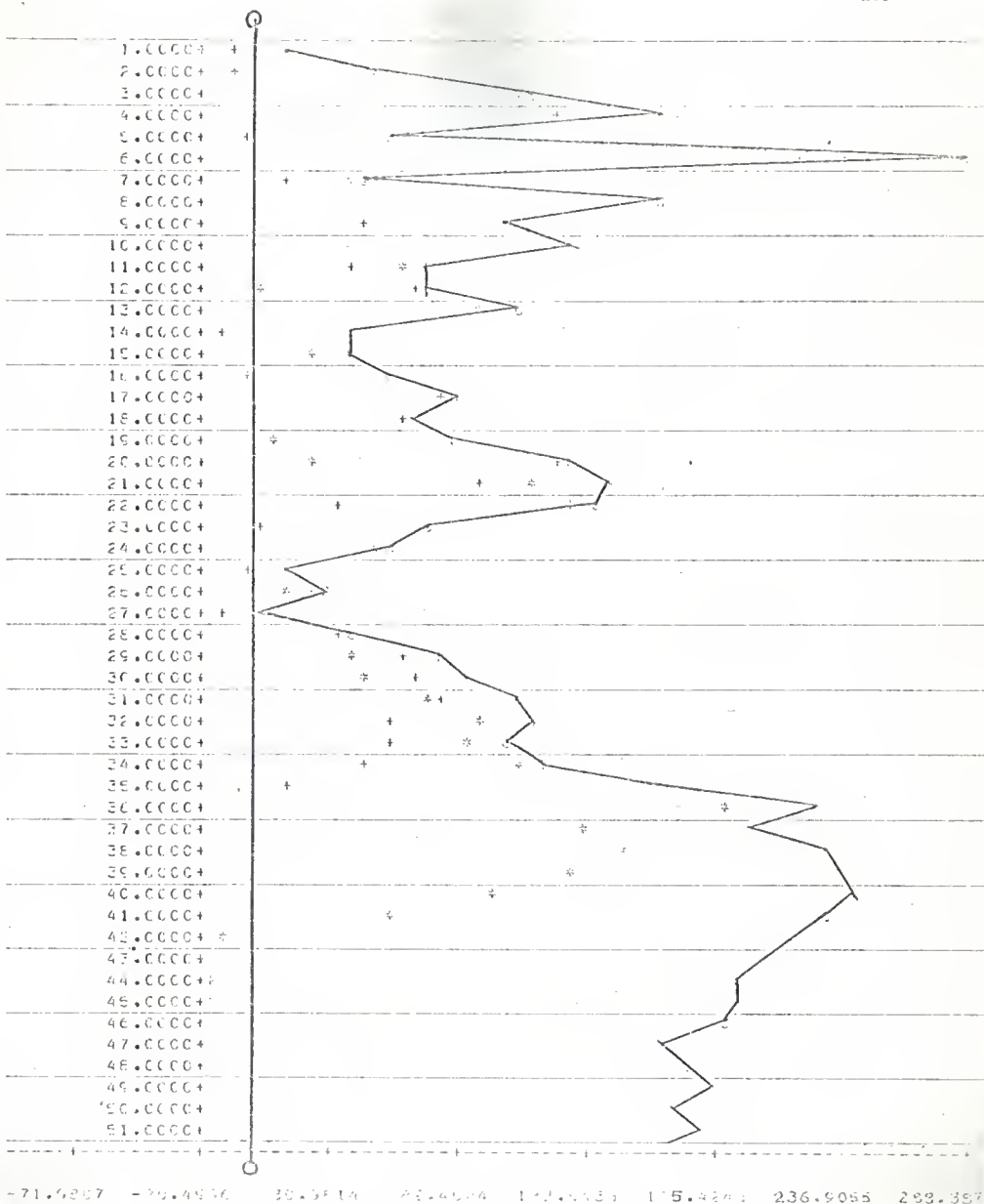


FIG. E-38

Pericogram for *Chrysanthemum* Data Using  
Fifth Degree Polynomial Regression Residuals

## APPENDIX F - COMPUTER PROGRAMS

Of the programs listed in this appendix, AUTODET, PCWSPEC, and EXPOSNO are revisions of programs originally written by Raymond Miller (7, 65). The Computer programs used were:

- (1) AUTODET
- (2) PCWSPEC
- (3) AUTOPLT
- (4) PLOT001
- (5) EXPOSNO
- (6) LATREGR

AUTODET performs a first degree polynomial regression on the original time series. The resulting residuals form the detrended time series from which the program computes an autocorrelation function for specified lag values. PCWSPEC uses the detrended time series from AUTODET to perform a harmonic analysis for specified periods. It computes the coefficients of harmonics in the arithmetic trigonometric sequence explained in Section 2.4. The resulting coefficient values may be used to construct a periodogram for the time series. AUTOPLT plots the original and detrended time series from AUTODET. PLOT001 produces autocorrelation function and periodogram values computed in AUTODET and PCWSPEC, respectively.

EXPOSNO performs the general exponential smoothing process on the original time series. At each sampling interval a forecast is made of the next time series value. Each forecast is based on all data up to that particular point in time. In addition to the time series itself, the following data must be fed into this program:

- (1) Number of observations and number of model fitting functions.

- (2) Initial coefficient vector.
- (3) Input format for transition matrix.
- (4) Transition matrix.
- (5) Input format for time series.
- (6) Values of discount factor to be tried.

For a more comprehensive explanation of AUTODET, POWSEEC, and EXPOSNO refer to the master's thesis by Miller, (7, 65). MATREGR is just a general regression program. The matrix of independent variables may be of any desirable form. This program outputs an analysis of variance table and a plot containing the actual, estimated, and residual time series.

## COMPUTER PROGRAM IDENTIFICATION CHART

The list below is provided as a reference to all of the computer programs which follow:

<u>PROGRAM NUMBER</u>	<u>NAME AND DESCRIPTION</u>
F-1	AUTODET - detrends the original time series and computes the autocorrelation function.
F-2	POWSPEC - computes periodic term coefficients for an arithmetic sequence of periods.
F-3	AUTOPLT - Plots original and detrended time series from AUTODET.
F-4	PLOT001 - plots autocorrelation function from AUTODET and periodogram from POWSPEC.
F-5	EXP0010 - Exponentially smoothes the original time series for any model and discount factor.
F-6	MATREGR - performs a multiple regression for independent variables of any desired form.



```

SUBROUTINE TREND(AX,X,N)
  DIMENSION X(500),AX(500)
11 FORMAT(20H THIS IS DETRENDED DATA)
1  FORMAT(2F20.8)
4  FORMAT(F20.8)
6  FORMAT (F5.0)
  SUMY=0
  SSQY=0
  SUMXY=0
  SUMX=0
  SSCX=0
  DO50I=1,N
  READ(1,6)(AX(I))
  S=I
  SUMY=SUMY+AX(I)
  SSQY=SSQY+AX(I)**2
  SUMXY=SUMXY+AX(I)*I
  SUMX=SUMX+AX(I)
50 SSCX=SSCX+AX(I)**2
  R=N
  A1=(R*(SUMXY)-(SUMX)*(SUMY))/(R*(SSQY)-(SUMY**2))
  WRITE(3,4)(A1)
  WRITE(2,4)(A1)
  WRITE(3,11)
  WRITE(2,11)
  A2 = SUMX/N - A1*(SUMY/N)
  DO 60I=1,N
  S=I
  X(I) = AX(I) - A1*S - A2
  WRITE(3,1)(X(I),AX(I))
60 WRITE(2,1)(X(I),AX(I))
  RETURN
END

```

```

      DIMENSIONX(500),A(100),B(100),C(100)
1  FORMAT(13)
2  FORMAT(3F20.7,2110)
32  FORMAT (F20.8)
      PI=3.1415927
      READ(1,1)ND
      INC=ND
      READ(1,1)LIM
      DC18I=1,ND
12  READ(1,32)X(1)
      CCNT=0.0
      DC3I=1,ND
3  CCNT=CCNT+X(1)
      CONT=CCNT/INC
C****CALCULATE THE FOURIER COEFFICIENTS
      DO5CN=1,LIM
      A(N)=0.0
      B(N)=0.0
      TCU=N
      NR=N
      AIS=N
66  NR=NR+NIS
      IF(ND-NR)22,44,66
22  NR=NR-NIS
44  CONTINUE
      DC40I=1,NR
      F=1-1
      B(N)=B(N)+SIN(2.*PI*F/TCU)*X(1)
      A(N)=A(N)+COS(2.*PI*F/TCU)*X(1)
46  CONTINUE
      ANR=NR
      B(N)=(2./ANR)*B(N)
      A(N)=(2./ANR)*A(N)
      R(N)=(A(N)**2+B(N)**2)**.5
      WRITE(3,2)A(N),B(N),R(N),N,NR
      WRITE(2,2)A(N),L(N),R(N),N,NR
56  CONTINUE
      STOP
      END

```

```
DIMENSION X(400),CUT(101),YPR(11),ANG(9)
1 FORMAT (2I3)
2 FORMAT (2F20.0)
3 FORMAT (25X,37FACTUAL(1), DETRENDED(2) DATA VS. TIME/)
4 FORMAT (F5.0,2F20.0)
5 FORMAT (5H TIME,10X,6FACTUAL,14X,9HDETRENDED/)
READ (1,1) N,NL
READ (1,2) (X(I+2*N),X(I+N),I=1,N)
DO 100 I=1,N
100 X(I) = I
WRITE (3,5)
WRITE (3,4) (X(I),X(I+N),X(I+2*N),I=1,N)
WRITE (3,3)
CALL PLCT (I,X,N,3,NL,0)
STOP
END
```



```

C FOR PLOT OF AUTO COEFF.
C P = BASE, A = 1 = AUTO COEF.
C FOR PLOT OF POWSPEC DATA
C LAG = BASE, A = 1, B = 2 R = 3
      DIMENSION P(130), A(200),GLT(101),YPR(11),ANG(9)
1  FORMAT (I3)
2  FORMAT (F10.0,F20.2)
3  FORMAT (3F20.7,F10.0)
4  FORMAT (F10.0,3F20.7)
5  FORMAT (3X,4HLACS,10X,10HAUTO COEFFS/)
6  FORMAT (10X,FREQUENCY,10X,2FA ,18X,2FB ,18X,2FR /)
      READ (1,1) N
      READ (1,2) (P(I),P(I+N),I=1,N)
      WRITE (3,5)
      WRITE (3,2) (P(I),P(I+N),I=1,N)
      CALL PLOT (1,P,N,2,0,0)
      READ (1,1) N
      READ (1,3) (A(I+N),A(I+2*N),A(I+3*N),A(I),I=1,N)
      WRITE (3,5)
      WRITE (3,4) (A(I),A(I+N),A(I+2*N),A(I+3*N),I=1,N)
      CALL PLOT (2,A,N,4,75,0)
      STOP
      END

```

```

      DIMENSION FCST(1000),PLT(800)
      DIMENSION AT(10,1),CGI(10)
      DIMENSION FK(10,10),F(10,10),AC(10,1),H(10,1)
      DIMENSION CHK(10)
      DIMENSION XC(200)
      DIMENSION CI(10),A(10,1),TM(10,10)
      DIMENSION FMT(10),XFMT(10)
      DIMENSION ER(200),PCTER(200)
      DOUBLE PRECISION F,FK,C1,C11
1   FORMAT (5I3)
2   FORMAT (3F10.5)
3   FORMAT (13A4)
4   FORMAT (F5.3)
5   FORMAT (' TRANSITION MATRIX')
6   FORMAT (1H,10F12.7)
      NN1 = 1
9   READ(1,1) ND,NL,N,MN,NTOU
      IF(ND) 13,13,14
14  CONTINUE
      READ(1,2) (CGI(I),AD(I,1),CHK(I), I=1,N)
      READ(1,3) FMT
      READ(1,FMT) ((TH(I,J), J=1,N), I=1,N)
      WRITE(3,5)
      DO 15 I=1,N
15  WRITE(3,6) (TM(I,J), J=1,N)
      READ(1,3) XFMT
C READ IN THE OBSERVATIONS
      READ(1,XFMT) (X(I), I=1,ND)
      TCU = NTOU
      PI = 3.1415927
10  READ(1,4) ABETA
      IF(ABETA) 12,12,11
C SAVE THE INITIAL VALUES OF THE CONSTANTS
C SAVE THE INITIAL VALUES OF THE FITTING FUNCTIONS
11  DO 16 I=1,N
      A(I,1) = AD(I,1)
16  CI(I) = CGI(I)
      CALL RAY (ND,NTOU,N,A,X,CI,MN,CHK,TM,CBETA,FK,F,AD,AI,ABETA,CO)
      CALL MINV(F,N)
      CALL HVEC(F,AC,N,H,ABETA)
C***FORM THE V MATRIX
      CALL VAR1(FK,F,N)
      CALL FPCST (N,ND,CI,AI,X,ABETA,MN,TM,F,H,FCST)
      DO 50 II = 1,ND
      PLT(II) = 11
      PLT(II+ND) = X(II)
      PLT(II+2*ND) = FCST(II)
      PLT(II+3*ND) = X(II) - FCST(II)
50  PLT(II + 4*ND) = 0.0
      CALL PLOT (NN1,PLT,ND,5,NL,0)
      NN1=NN1+1
      GO TO 10
12  CONTINUE
      GO TO 9
13  STOP
      END

```

```

SUBROUTINE RAY (ND,NTOL,N,A,Y,C1,YN,CHK,TV,CBETA,FK,F,AG,AI,ABET
1CUT)
  DIMENSION A(10,1),B(1,10),C1(10,10),F(10,10),CUI(10)
  DIMENSION X(207),C1(10),A1(10,1),H(10,1),AG(10,1),TH(10,10)
  DIMENSION AP(10,1)
  DOUBLE PRECISION I,FK,C1,C11
  1 FORMAT (' F-MATRIX')
  2 FORMAT ('IN',10D13.5)
  TN=N
  AP=(1./TN)
  C STARTING HERE THE F MATRIX IS RECALCULATED WITH A NEW BETA
  67 DBETA=(ABETA)**AP
  C REINITIALIZE THE VECTOR OF FITTING FUNCTIONS AND CONSTANTS
  DB42I=1,N
  C1(I)=CUI(I)
  42 AA(1,1)=AG(1,1)
  C COMPUTE THE F MATRIX FOR ONE MODEL AND ONE BETA
  DB3I=1,N
  DB3J=1,N
  FK(I,J)=0.0
  8 F(I,J)=0.0
  BETA=1.0
  CNT=1.
  K=1
  99 DB3OI=1,N
  30 B(1,1)=A(1,1)
  DB4OI=1,N
  X=1
  DB4OJ=X,N
  C1(1,J)=A(1,1)*B(1,J)*(BETA)
  40 CONTINUE
  DB43I=1,N
  X=1
  DB43J=X,N
  43 C11(I,J)=C1(I,J)*BETA
  IF(K-1)66,16,66
  C****CHECK THE CONVERGENCE OF THE F MATRIX AT EACH 50TH ITERATION
  C****THIS ROUTINE CHECKS UNTIL IT FINDS AN ELEMENT WHICH IS NOT STABLE
  C****IZED
  66 IF(CNT-50.)16,44,16
  44 CNT = 0.0
  DO 80 I=1,N
    F=1
    DB8CJ=X,N
    DENZF=F(1,J)
    ANUM=C1(1,J)
    RATIO=ANUM/DENZF
    IF(RATIO-.CCCCC1)80,80,16
  80 CONTINUE
  C****NORMAL EXIT FROM THIS LOOP OCCURS WHEN ALL OF THE ELEMENTS OF
  C****THE F MATRIX MEET THE CONVERGENCE CRITERIA WHEN THIS OCCURS THE
  C****F MATRIX IS CLAMPED BY SYMMETRY AND THE PROGRAM CHECKS
  C****TO SEE IF THE RANGE OF BETA HAS BEEN COVERED
  DB87I=2,N
  L=1-1
  EXPOSPO (cont'd.)

```

```

DO67J=1,L
FK(1,J)=FK(J,1)
67 F(1,J)=F(J,1)
WRITE(3,1)
DO68 I=1,N
28 WRITE(3,2) (F(1,J), J=1,N)
GO TO 45
C*****THE PROGRAM BRANCHES TO HERE WHEN THE CONVERGENCE
C*****CRITERIA IS NOT MET AND NORMAL ITERATION CONTINUES
16 DO50I=1,N
NEXT
DO50J=1,N
FK(1,J)=FK(1,J)+C11(1,J)
50 F(1,J)=F(1,J)+C11(1,J)
C*****NOW UPDATE THE VECTOR OF FITTING FUNCTIONS USING THE L MATRIX
DO31I=1,N
31 AF(1,1)=0.0
DO30I=1,N
DO30J=1,N
90 AF(I,1)=FK(1,J)*AA(J,1)+AF(I,1)
DO100I=1,N
AA(1,1)=AF(1,1)
100 A(1,1)=AF(1,1)
C*****THIS IS A ROUTINE FOR CONVERTING THE VECTOR OF FITTING FUNCTION
C*****FOR T TO THE VECTOR EVALUATED FOR -T
DO63I=1,N
CK=CHK(1)
IF(CK-1.)63,65,63
65 A(1,1)=-A(1,1)
63 CONTINUE
C SAVE F(1) FOR FORECAST
IF(K-1)62,22,62
22 DO62I=1,N
A1(1,1)=AA(1,1)
62 CONTINUE
BETA=DOBETA*BETA
CNT=CN+1.
K=K+1
GO TO 99
45 CONTINUE
RETURN
END

```

```

SUBROUTINE MATINV(C,N)
DIMENSION C(10,10),V(10),F(10,10)
DOUBLE PRECISION F,FK,C1,C11,C
NM1=N-1
DO 6050 L=1,N
ALL=C(1,1)
DO 4440 J=1,NM1
4440 V(J)=C(1,J+1)/ALL
V(N)=1./ALL
DO 9990 I=1,NM1
IP1=I+1
CIP11=C(IP1,1)
DO 5550 J=1,NM1
5550 C(I,J)=C(IP1,J+1) - CIP11*V(J)
9990 C(I,N)=(-1.0)*CIP11+V(N)
DO 6660 J=1,N
6660 C(N,J)=V(J)
RETURN
END

```

```

SUBROUTINE VARIX(FK,F,N)
  DIMENSION F(10,10),FK(10,10)
  DIMENSION V2(10,10)
  DIMENSION V1(10,10)
  DOUBLE PRECISION F,FK,C11
777 FORMAT(///10X,35H THE VARIANCE OF THE COEFFICIENTS IS )
789 FORMAT (F12.6)
  DO14I=1,N
  DO14J=1,N
14 V2(I,J)=0.0
  DO11I=1,N
  DO11K=1,N
  DO11J=1,N
11 V2(K,I)=F(K,J)*FK(J,I)+V2(K,I)
  DO13I=1,N
  DO13J=1,N
13 V1(I,J)=0.0
  DO12I=1,N
  DO12K=1,N
  DO12J=1,N
12 V1(K,I)=V2(K,J)*F(J,I)+V1(K,I)
  WRITE(3,777)
  DO122K=1,N
122 WRITE(3,789) V1(K,K)
  RETURN
END

```

EXPOSNO (cont'd.)

```

SUBROUTINE HYEC(F,AD,N,H,BETA)
  DIMENSION F(10,10),AD(10,1),H(10,1),FF(10,1)
  DOUBLE PRECISION F,FK,C1,C11
  1 FORMAT (' EFFECTIVE BETA FOR',14,' DIMENSION SMOOTHING IS')
  789 FORMAT (F14.5)
  727 FORMAT(///1X,16H THE H VECTOR IS      )
  444 FORMAT (1H ,10D13.5)
  888 FORMAT(///1X,11H F INVERSE      )
  889 FORMAT(///20X,6H BETA=,F9.6)
  DO32I=1,N
  32 H(1,1)=0.0
  DO543I=1,N
  DO543J=1,N
  543 F(1,1)=H(1,1)+ F(1,J)*AD(J,1)
  629 CONTINUE
  C*****WRITE THE EFFECTIVE BETA
  WRITE(3,1) N
  WRITE(3,889)DBETA
  C*****WRITE F INVERSE
  WRITE(3,388)
  DO333I=1,N
  333 WRITE(3,444)(F(1,J),J=1,N)
  C WRITE OUT THE H VECTOR
  WRITE(3,727)
  DO600I=1,N
  600 WRITE(3,789)(H(1,1))
  RETURN
  END

```

```

SUBROUTINE FORCST (N,ND,C1,A1,X,ABETA,MN,TM,F,H,FCST)
  DIMENSION FCST(1000),ER(200),F(10,10),C0F(10),A(10,1),A(10,1)
  DIMENSION X(200),H(10,1),C1(10),TM(10,10),CIF(10),CIL(10)
  DIMENSION ITM(10,10), PCTER(200)
  DOUBLE PRECISION F,ER,C1,CIL
  1 FORMAT (' STD DEV FORECASTS IS ',F12.6/)
  2 FORMAT (' MEAN ERROR IS ',F12.6/)
  3 FORMAT (' MEAN ABSOLUTE ERROR IS ',F12.6/)
  4 FORMAT (' VARIANCE OF THE ERROR IS ',F15.6/)
  5 FORMAT (' STD DEV ABOUT MEAN OF THE ERROR IS ',F12.6/)
  6 FORMAT (' STD DEV ABOUT ZERO OF THE ERROR IS ',F12.6/)
  7 FORMAT (' MEAN PERCENT ERROR IS ',F12.6/)
  222 FORMAT(///1X,31H THE VARIANCE OF THE FORECASTS=,F13.6/1H1)
  107 FORMAT(1X,F20.7,F20.4)
  108 FORMAT(1X,15H NEW FORECAST )
  101 FORMAT(1X,11H MODEL NO.=,I2,20X,6H BETA=,F4.3)
  193 FORMAT(1X,20H***THE ERROR IS*** )
  973 FORMAT(85H  PC10D      OBSERVATION      FORECAST      ERROR
    I      CUM. ERROR      PCT. ERROR )
  104 FORMAT(1X,F15.3,F15.2,F15.2,F15.2,F15.2)
  106 FORMAT(24H SUM OF SQUARE OF ERROR=,F15.3)
  C      PRINT OUT THE MODEL NO. AND BETA FOR THIS F MATRIX
  WRITE(3,101)MN,ABETA
  C MAKE THE TRANSPOSE OF THE TRANSITION MATRIX
  SUM=0.0
  SUMA=0.0
  SUMX=0.0
  SUMAB = 0.0
  SUMPCT = 0.0
  SEDSQ = 0.0
  DO429I=1,N
  DO429J=1,N
  429 ITM(I,J)=TM(J,I)
  C MAKE THE FORECAST
  WRITE(3,973)
  DO191K=1,ND
  Z=0.0
  DO711I=1,N
  ZN=C(1)*A1(I,1)
  711 Z=Z+ZN
  FCST(K)=Z
  C CALCULATE THE ERROR
  ER(K) = X(K) - Z
  PCTER(K)=(ABS(ER(K))/X(K))*100.0
  C CALCULATE THE CUMULATIVE ERROR
  SUM = SUM + ER(K)*ER(K)
  SUMA = SUMA + ER(K)
  SUMAB = SUMAB + ABS(ER(K))
  SUMPCT = SUMPCT + PCTER(K)
  C***CALCULATE THE SUM OF THE OBSERVATIONS
  SUMX=SUMX+X(K)
  C UPDATE THE CONSTANTS
  DO443I=1,N
  449 CIL(I)=0.0
  DO444I=1,N
  DO444J=1,N
  944 CIL(I)=CIL(I)+ITM(I,J)*C1(J)
  DO726I=1,N

```

EXPOSNO. (cont'd.)



```

920 CI(I) = CIL(I) + H(I,1)*ER(K)
WRITE(J,104) (K,X(K),Z,ER(K),SUMA,PCTER(K))
191 CONTINUE
ERM = SUMA/ND
ABERM = SUMAB/ND
PERM = SUMPCI/ND
DO 20 I=1,ND
20 SEDSQ = SEDSQ + (ER(I) - ERM)**2
VARE = SEDSQ/ND
SDEVE = SQRT(VARE)
SDEV = SQRT(SUM/ND)
WRITE(J,103)
WRITE(J,106)(SUM)
C****CALCULATE THE VARIANCE OF THE FORECASTS
VAR=SUM/ND
SDFST = SQRT(VAR)
WRITE(J,22)VAR
WRITE(J,1) SDFST
WRITE(J,2) ERM
WRITE(J,3) ABERM
WRITE(J,4) VARE
WRITE(J,5) SDEVE
WRITE(J,6) SDEV
WRITE(J,7) PERM
RETURN
END

```



C

```

N1=N+1
YMIN=A(N1)
YMAX=YMIN
M2=M+1
DO 40 J=M1,M2
IF(A(J)-YMIN) 28,26,26
26 IF(A(J)-YMAX) 40,40,30
28 YMIN=A(J)
GO TO 40
30 YMAX=A(J)
40 CONTINUE
YSCAL=(YMAX-YMIN)/100.0

```

C

C

FIND BASE VARIABLE PRINT POSITION

C

```

XB=A(I)
L=1
MY=M-1
I=1
45 F=I-1
XPR=XB+F*XSCAL
IF(A(L)-XPR) 49,49,70

```

C

C

FIND CROSS-VARIABLES

C

```

49 LSW=0
50 IF (L-N) 51,51,80
51 DO 55 IZ=I,101
55 OUT(IZ)=BLANK
57 DO 60 J=1,NY
LL=L+J*N
JP=((A(LL)-YMIN)/YSCAL)+1.0
OUT(JP)=ANG(J)
60 CONTINUE
L=L+1

```

C

C

PRINT LINE AND CLEAN, OR SKIP

C

```

IF(LSW) 62,61,62
61 WRITE(3,2)XPR,(OUT(IZ),IZ=I,101)
LSW=1
IF(A(L)-XPR-XSCAL) 50,90,80
62 WRITE(3,3) (OUT(IZ),IZ=I,101)
IF(A(L)-XPR-XSCAL) 50,90,80

```

C

```

70 WRITE(3,3)
80 I=I+1
IF(I-NLL) 45,84,85
84 XPR=A(N)
GO TO 49

```

C

C

PRINT CROSS-VARIABLES NUMBERS

C

```

80 WRITE(3,7)
YPR(I)=YMIN
DO 90 KN=1,9
90 YPR(KN+1)=YPR(KN)+YSCAL*10.0

```

EXPOS'0 (cont'd.)

```
YPR(11)=YMAX  
WRITE(3,3)(YPR(IR),IR=1,11)  
RETURN  
END
```

EXFCSHO (cont'd.)

```

DIMENSION XX(700),YY(100),XPY(3),YPY(1)
DIMENSION XPY(6),JPY1(6),JPY2(6),J(6),BPXPY(1)
DIMENSION YFAT(100),FLT(100)
DIMENSION S(100), A(100), R(100)
DIMENSION RESID(100),RX(100),RXX(100),RXXX(100)

1 FORMAT(16,I3)
2 FORMAT(F5.0)
3 FORMAT(1H0/24X,20HANALYSIS OF VARIANCE/)
4 FORMAT(1H0,3X,19HSOURCE OF VARIATION,7X,9HDEGREE OF,7Y,6HSUM OF,
1,4HMEAN,10X,1HF/33X,7HREGRSS,3X,7HSQUARES,7X,6HSUM OF,
5 FORMAT(20H0 DUE TO REGRESSION,12X,16,F17.5,F17.5,F13.5)
6 FORMAT(30H DEVIATION ABOUT REGRESSION,16,F17.5,F17.5,5///)
7 FORMAT(1H0,4HRS =,F10.5,10X,4HRS =,F10.5,10X,4HRS =,F10.5,10X,4H
1 =,F10.5,10X,4HRS =,F10.5,5HRS =,F10.5)
16 FORMAT(1H0/24X,20HDECOMPOSITION OF RESIDUAL/)
17 FORMAT(1H0,3X,19HSOURCE OF VARIATION,7X,9HDEGREE OF,7X,6HSUM OF
1X,4HMEAN,10X,1HF/33X,7HREGRSS,3X,7HSQUARES,7X,6HSUM OF,
2)
18 FORMAT(14H0 LACK OF FIT,18X,16,F17.5,F17.5,F13.5)
19 FORMAT(13H PURE ERROR,18X,16,F17.5,F17.5,5///)
20 FORMAT(3F12.5)
21 FORMAT(34H ACTUAL ESTIMATE RESIDUAL/)
22 FORMAT(F20.8)
98 READ(1,1) N,NC
IF(N) 204,204,99
99 READ(1,2) (YY(I),I=1,N)
DO 400 I=1,N
XX(I) = 1.0
XX(I+N) = 1
XI = 1
XX(I+2*N) = XI**2
XX(I+3*N) = XI**3
XX(I+4*N) = XI**4
400 XX(I+5*N) = XI**5
CALL GTPRD(XX,XX,XPY,N,6,6)
CALL GTPRD(YY,YY,YPY,N,1,1)
CALL GTPRD(XX,YY,XPY,N,6,1)
CALL MINV(XPX,6,DET,JPY1,JPY2)
CALL GTPRD(XPX,XPY,E,6,6,1)
CALL GTPRD(S,XPY,BXPY,6,1,1)
YSUM = 0.0
DO 100 K=1,N
100 YSUM = YSUM + YY(K)
TN = N
CFAC = (YSUM**2)/TN
CSSDR = BPXPY(1) - CFAC
RGMS = CSSDR
CSSRL = YPY(1) - BPXPY(1)
RLMS = CSSRL/(TN-6)
JRG = 5
JRL = N-6
FVAL = RGMS/RLMS
WRITE(3,3)
WRITE(3,4)
WRITE(3,5) JRG,CSSDR,RGMS,FVAL
WRITE(3,6) JRL,CSSRL,RLMS
WRITE(3,7) C(1),C(2),C(3),C(4),C(5),C(6)
199 DO 200 I=1,N

```

```

200 YHAT(I) = B(I) + XX(N+1)*B(P) + XX(P*N+1)*B(3)+XX(3+.1)*P(4)+Y
1*(N+1)*B(5)+XX(5*N+1)*B(6)
DO 201 J=1,N
201 PLT(J) = Y(N+J)
DO 202 K=1,N
202 PLT(Y+K) = YY(K)
DO 203 L=1,N
203 PLT(L+2*Y) = YHAT(L)
CALL PLOT (NO,PLT,N,3,53,1)
DO 300 I=1,N
300 RESID(I)=YY(I) - YHAT(I)
WRITE (3,21)
WRITE (3,20) (YY(I),YHAT(I),RESID(I),I=1,N)
WRITE(2,22) (RESID(I), I=1,N)
GO TO 98
204 STOP
END

```

# SUBROUTINE GTPRD

## PURPOSE

PREMULTIPLY A GENERAL MATRIX BY THE TRANSPOSE OF ANOTHER  
GENERAL MATRIX

## USAGE

CALL GTPRD(A,B,R,N,M,L)

## DESCRIPTION OF PARAMETERS

A - NAME OF FIRST INPUT MATRIX  
E - NAME OF SECOND INPUT MATRIX  
R - NAME OF OUTPUT MATRIX  
N - NUMBER OF ROWS IN A AND B  
M - NUMBER OF COLUMNS IN A AND ROWS IN R  
L - NUMBER OF COLUMNS IN B AND R

## REMARKS

MATRIX R CANNOT BE IN THE SAME LOCATION AS MATRIX A  
MATRIX R CANNOT BE IN THE SAME LOCATION AS MATRIX B  
ALL MATRICES MUST BE STORED AS GENERAL MATRICES

## SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED

NONE

## METHOD

MATRIX TRANSPOSE OF A IS NOT ACTUALLY CALCULATED. INSTEAD  
ELEMENTS OF MATRIX A ARE TAKEN COLUMNWISE RATHER THAN  
ROWWISE FOR POSTMULTIPLICATION BY MATRIX B.

```

SUBROUTINE GTPRD(A,B,R,N,M,L)
DIMENSION A(100),B(100),R(100)

```

MATREG (cont'd.)

```

IR=0
IK=-N
DO 10 K=1,L
  IJ=0
  IK=IK+N
  DO 10 J=1,M
    IB=IK
    IR=IR+1
    R(IR)=0
    DO 10 I=1,N
      IJ=IJ+1
      IB=IB+1
      IC R(IR)=R(IR)+A(IJ)*B(IB)
    RETURN
  END

```

# ..... SUBROUTINE MINV

## PURPOSE

INVERT A MATRIX

## USAGE

CALL MINV(A,N,D,L,M)

## DESCRIPTION OF PARAMETERS

A - INPUT MATRIX, DESTROYED IN COMPUTATION AND REPLACED BY  
RESULTANT INVERSE.

N - ORDER OF MATRIX A

D - RESULTANT DETERMINANT

L - WORK VECTOR OF LENGTH N

M - WORK VECTOR OF LENGTH N

## REMARKS

MATRIX A MUST BE A GENERAL MATRIX

## SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED

NONE

## METHOD

THE STANDARD GAUSS-JORDAN METHOD IS USED. THE DETERMINANT  
IS ALSO CALCULATED. A DETERMINANT OF ZERO INDICATES THAT  
THE MATRIX IS SINGULAR.

.....

SUBROUTINE MINV(A,N,D,L,M)

DIMENSION A(100),L(2),M(2)

.....

IF A DOUBLE PRECISION VERSION OF THIS ROUTINE IS DESIRED, THE  
C IN COLUMN 1 SHOULD BE REMOVED FROM THE DOUBLE PRECISION  
STATEMENT WHICH FOLLOWS.

KATREGR (cont'd.)

```

C
C      DOUBLE PRECISION A,D,BIGA,FOLD
C
C      THE C MUST ALSO BE REMOVED FROM DOUBLE PRECISION STATEMENTS
C      APPEARING IN OTHER ROUTINES USED IN CONJUNCTION WITH THIS
C      ROUTINE.
C
C      THE DOUBLE PRECISION VERSION OF THIS SUBROUTINE MUST ALSO
C      CONTAIN DOUBLE PRECISION FORTRAN FUNCTIONS.  ABS IN STATEMENT
C      10 MUST BE CHANGED TO DABS.
C
C      .....
C
C      SEARCH FOR LARGEST ELEMENT
C
      D=1.0
      NK=-N
      DO 80 K=1,N
        NK=NK+A
        L(K)=K
        A(K)=K
        KF=NK+K
        BIGA=A(K*K)
        DO 20 J=<,N
          IZ=N*(J-1)
          DO 20 I=K,N
            IJ=IZ+I
            10 IF( ABS(BIGA)- ABS(A(IJ))) 15,20,20
            15 BIGA=A(IJ)
            L(K)=I
            M(K)=J
          20 CONTINUE
C
C      INTERCHANGE ROWS
C
      J=L(K)
      IF(J-K) 35,35,25
      25 KI=K-N
      DO 30 I=1,N
        KI=KI+N
        FOLD=-A(KI)
        JI=KI-K+J
        A(KI)=A(JI)
      30 A(JI)=FOLD
C
C      INTERCHANGE COLUMNS
C
      35 I=M(K)
      IF(I-<) 45,45,38
      38 JP=N*(I-1)
      DO 40 J=1,N
        JK=NK+J

```

MATEGR (cont'd.)



```

      JI=JP+J
      HOLD=-A(JK)
      A(JK)=A(JI)
40  A(JI) =HOLD
C
C      DIVIDE COLUMN BY MINUS PIVOT (VALUE OF PIVOT ELEMENT IS
C      CONTAINED IN BIGA)
C
45  IF(BIGA) 48,46,48
46  D=0.0
      RETURN
48  DO 55 I=1,N
      IF(I-K) 50,55,50
50  IK=NK+I
      A(IK)=A(IK)/(-BIGA)
55  CONTINUE
C
C      REDUCE MATRIX
C
      DO 65 I=1,N
      IK=NK+I
      HOLD=A(IK)
      IJ=I-N
      DO 65 J=1,N
      IJ=IJ+N
      IF(I-K) 60,65,60
60  IF(J-K) 62,65,62
62  KJ=IJ-I+K
      A(IJ)=HOLD*A(KJ)+A(IJ)
65  CONTINUE
C
C      DIVIDE ROW BY PIVOT
C
      KJ=NK-N
      DO 75 J=1,N
      KJ=KJ+N
      IF(J-K) 70,75,70
70  A(KJ)=A(KJ)/BIGA
75  CONTINUE
C
C      PRODUCT OF PIVOTS
C
      D=D*BIGA
C
C      REPLACE PIVOT BY RECIPROCAL
C
      A(KK)=1.0/BIGA
80  CONTINUE
C
C      FINAL ROW AND COLUMN INTERCHANGE
C
      K=N

```

MATREGR (cont'd.)

```

100 K=(K-1)
    IF(K) 150,150,105
105 I=L(F)
    IF(I-K) 120,120,105
100 JQ=N*(F-1)
    JR=N*(I-1)
    DO 110 J=1,N
        JK=JQ+J
        HOLD=A(JK)
        JI=JR+J
        A(JK)=-A(JI)
110 A(JI)=HOLD
120 J=V(K)
    IF(J-K) 100,100,125
125 KI=K-N
    DO 130 I=1,N
        KI=KI+N
        HOLD=A(KI)
        JI=KI-K+J
        A(KI)=-A(JI)
130 A(JI)=HOLD
    GO TO 100
150 RETURN
    END

```

MATREGR (cont'd.)

DEVELOPMENT OF A DEMAND FORECASTING MODEL  
FOR  
FLORAL DATA

by

WILLIAM FRANCIS IRVINE

B. S., University of Notre Dame, 1966

---

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

1968

The ability to forecast with reasonable accuracy the future demand for an item of production has value anywhere that an inventory must be maintained. This research investigates one approach to the problem of forecasting, general exponential smoothing.

The logical basis for forecasting the future demand of an item is the demand behavior of this item in the past. Such past behavior may be summarized in the form of a time series, a sequence of observations of demand taken at specific and equal intervals of time. We assume that this time series is composed of a systematic component, which may be described by a linear combination of deterministic functions of time, and random variation. The deterministic functions of time are known as fitting functions, and their sum, along with the appropriate fitting function coefficients, makes up the model of the time series.

Assuming that recent data is more representative of the future time series, we wish to weight the data, giving the most recent observations more weight than older data. We do this by assigning a weight  $\alpha\beta^k$  to an observation made  $k$  sampling intervals before the present time, where  $\alpha$  and  $\beta$  are constants whose sum is unity. Now, we may control the general exponential smoothing response rate by merely varying  $\beta$ , the discount factor. When  $\beta$  is smaller, we give recent data more weight, and the forecasting system may respond more quickly to changes in the actual data.

General exponential smoothing fits a least squares curve to the time series much like statistical regression. It does this by determining the fitting function coefficients in the model that will minimize

the sum of weighted squared residuals. The main differences between these two methods are that general exponential smoothing operates on weighted data, and the fitting function coefficients are re-computed each time new data becomes available. Hence, this process is able to adapt to small changes in the true time series process.

To simplify the computations involved in making a forecast, the fitting functions must consist of: (1) trigonometric functions, (2) polynomial functions, (3) some exponential functions, (4) mathematical combinations of the first three types. In addition to this restriction, the number of time series observations available must be at least equal to the number of fitting function coefficients to be computed.

General exponential smoothing offers several advantages over other fore-casting methods: (1) Response rate is easily controlled; (2) All past data is considered and contained in one word of information; (3) Model coefficients are re-computed at each sampling interval.

The data used herein consists of four time series of demand for plants produced and sold by Manhattan Wholesale and Retail Floral Companies. For each time series various combinations of models and discount factors were tried in an attempt to determine the combination that best forecasts the time series. Forecasts were evaluated on the basis of three efficiency criteria: standard deviation of the forecasts, mean absolute error, and standard deviation of the error. The effects of model and discount factor choice on fore-casting efficiency were also studied in an attempt to gain a clearer understanding of the general exponential smoothing technique.

All computations are performed on the IBM System 360/50 digital computer at Kansas State University.