BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-82/10

FINAL REPORT TO ASA/CENSUS SPECIAL PROGRAM FOR TIME SERIES METHODS DEVELOPMENTS

by

Adi Raveh

Hebrew University, Jerusalem and Stanford
University, California

and

currently, an ASA Junior Research Fellow,
Bureau of the Census, Washington, D.C.

Recommended by: Myron J. Katzoff

Report completed: July 18, 1982

Report issued: July 18, 1982

Dear Ted: 8.82

Many thanks for your great help hope to keep in touch.

yours friendly

Adi

Final Report to ASA/CENSUS SPECIAL PROGRAM

FOR TIME SERIES METHODS DEVELOPMENTS

by

Adi Raveh

Hebrew University, Jerusalem and Stanford University, California

and

currently, an ASA Junior Research Fellow, Bureau of the Census, Washington D.C.

July 18, 1982

School of Business Administration
Hebrew University
Jerusalem 91904

D'israeli 9|B street
Talbia
Jerusalem 92222
Israel

ISRAEL

Final Report to ASA/CENSUS SPECIAL PROGRAM
FOR TIME SERIES METHODS DEVELOPMENTS

by

Adi Raveh

Hebrew University, Jerusalem and Stanford University, California

and

currently, an ASA Junior Research Fellow, Bureau of the Census,
Washington, D.C.

July 18, 1982

School of Business Administration

Hebrew University

Jerusalem 91904

ISRAEL

D'israeli 91B street
Talbia
Jerusalem 92222

Israel

TABLE OF CONTENTS

Part 2

Part 3

Part 4

Appendix

PREFACE

The ASA Fellowship presented a unique opportunity for the author to work on the topic "Data Analysis of Time Series." This report is an attempt to summarize the author's work on the ASA Special Project for Time Series methods development at the Bureau of the Census.

This report consists of four main parts. The first two deal with Seasonal Adjustment and Forecasting quantitative series, based mainly on nonmetric filters. Analysis and Prediction by various types of examples, mainly economic data, are presented as well as a brief description of X-11 and a comparison of our approach to X-11. The third part deals mainly with analysis and forecasting qualitatives series. In the final part, we deal mainly with graphical methods in order to study relationships among a given set of empirical series.

I would like to thank Louis Guttman, Ingram Olkin and Arnold Zellner for their outstanding comments and for encouraging me before and during my fellowship period. My gratitude is extended to Estella B. Dagum, Charles Tapiero, George Tiao and Joe Kruskal for their help and suggestions. Thanks must also go to T.W. Anderson, David Brilinger, W.P. Cleveland, W.S. Cleveland, Morris Hamburger, Joseph Kadene, Charles Nelson, and John Tukey for fruitful discussions and helpful comments.

Let me take this opportunity to comment on the special atmosphere which allows me to work with the talented people in the Statistical Research Division (SRD) at the Bureau. My friends and co-workers in the Division and within the ASA Special Projects Group deserve special recognition for their continued interest and support. Special thanks go to David Findley of SRD for his kindness, for sharing comments and ideas, and for offering methodological help; and to Ted Holden for his outstanding pragramming help. I would like to thank Bill Bell, Will Gersch, John Irvine, Genshiro Kitagawa, Sandy MacKenzie, Nash Monsour, and Kirk Wolter. Finally, my deepest thanks and appreciation to Lillian Wilson for her efforts to improve my English, and for typing this report. It was a real pleasure to be helped by Mrs. Wilson during my ten months with SRD.

INTRODUCTION

This report deals with data analysis of empirical series. The main concept in such data analysis is the concept of order among the observations. An important special case of these series are the Time Series in which the order is determined by Time. Indeed, most of the examples included in this work are (economic) Time Series, but not exclusively. We are concerned with relationships between the values of the observations and their order, namely, the behavior of observations over time.

Our point of view is, briefly, that Data Analysis requires a loss function to be minimized (or a measure for goodness-of-fit to be maximized). The loss function is based on definitions that are related to the research problem. Actually, we measure the amount of deviation of empirical data from a priori definitions.

Our point of view is that analysis of empirical series is a special case of the general problem of dividing the space of indices of observations into intervals. For time series the division is into both equal and unequal intervals. Equal intervals are needed for estimating fixed seasonal patterns. Unequal intervals are needed for the trend, moving seasonality, etc.

Three main goals achieved in analyzing economic time series are:

(a) Decomposing into components while knowing the periods length.

(b) Seasonal adjustment of current data, and

(c) Forecasting.

The question of model identification and estimation may (or may not) be involved in the above three topics. In order to do (a) and (b), the well-known X-11 program as well as X-11-ARIMA were developed. The latter was developed especially for goal (b). The X-11 procedure, see Shiskin, Young and Musgrave (1967) represents the culmination of a major phase of continuing research in the area of seasonal adjustment. Today, the X-11 program is also the most widely used for economic time-series.

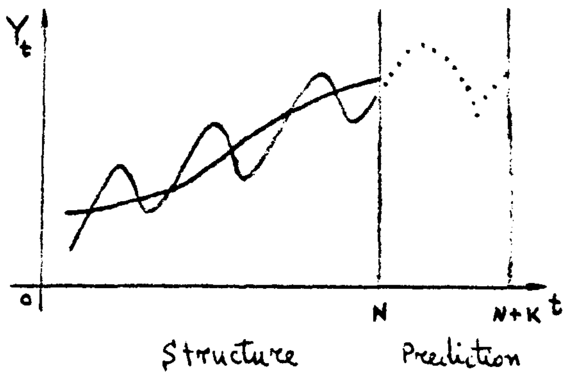
Four main parts comprise this report: The first two parts deal essentially with seasonal adjustment and forecasting quantitatives series, mostly based on economic data. The third part deals with analysis and prediction of qualitative time series. In the last part, interrelationships among various components of Time Series are treated and used for graphical methods.

We will only consider discrete time series with observations y_t made at times t=1,...,N, where N, the length of the series, is the total number of observations made. In Figure 1, a typical Time Series study is exhibited. Our goals are mainly two-fold: (a) Study the structure of a given empirical series,

and (b) Prediction for some range ahead.

Hence, for the first N data points, our goal is to reveal their structure while Prediction is the main goal for the next k points of time, presuming that the structure remains the same.

Figure 1: A typical Time Series Study: Structure Analysis of the first N data point and prediction the next k points over time.



In order to study the structure of a Quantitative Series, a Nonmetric approach is suggested in part 1. Analysis by examples of various types of series are given as well. The proposed method is designed to compose an empirical time series into its main three components: trend, periodicity and irregularity. Filters are used in two stages:

(a) remove either fixed or moving seasonality in order to produce seasonally adjusted data (S.A.D);

(b) Remove irregularity from S.A.D. in order to estimate it as well as the trend.

Both seasonality and irregularity could be either purely multiplicative or purely additive fashion. Seasonality could also be a kind of a mixed model. The filters are nonmetric since the loss function has no specific formula but a very general shape called polytonicity (or monotonicity as a special case). The method search for the smallest number of tones (monotone segments) possible for trend, or in other words, minimize the number of turning points. Thus it is called - Least Polytone Trend Analysis (LPTA).

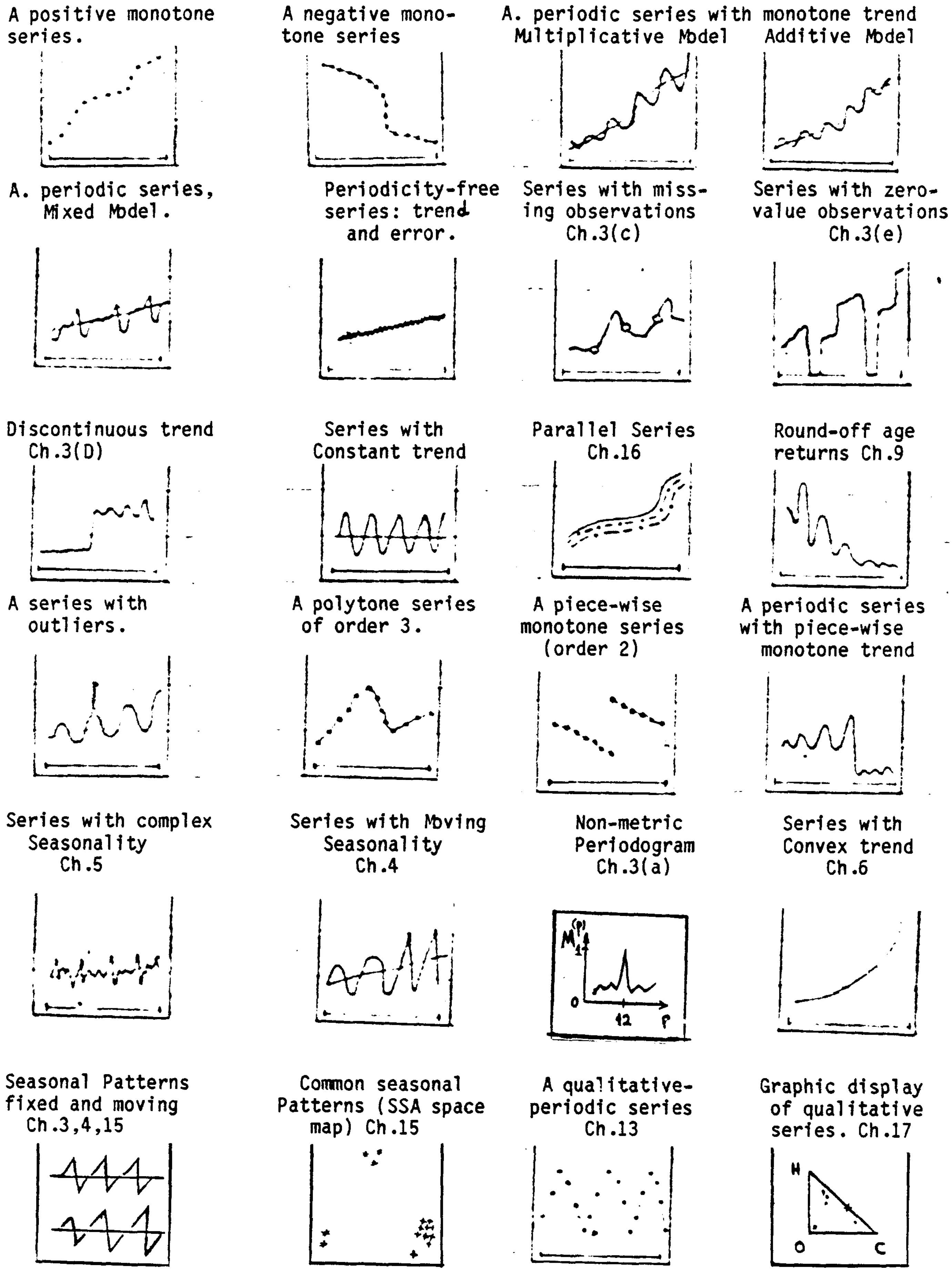
A computer program has been developed which enables analysis of arbitrary series, either by a prespecified length of period or by estimating the period's length if not known in advance. Robustness of the nonmetric approach enables analysis of very short series, series with missing values, and other series with limitations that cannot be easily handled otherwise.

In Chapters 5 and 6 the LPTA method is extended to deal with complex seasonality as well as convex (concave) series. To conclude part 1, in chapter 8 some theoretical and empirical results for economic time-series obtained by this approach are compared with those from the X-11 program. A brief description of X-11 and some notes are given in chapter 7.

In order to deal with prediction, two approaches are suggested in part 2, chapters 10 and 11. In chapter 12 a way to improve X-11 is discussed. Examples are given throughout the report; specific series are presented in Chapter 3(a)-(e), 4(a), (b) and in chapter 9.

For Qualitative series the main two goals of revealing the structure and doing forecast is presented in part 3. Various methods to compete with a multivariate time series in a special way is discussed in part 4. In Figure 2, charts of various types of series that are analyzed in this report are presented.

Figure 2: Charts of various types of series which are analyzed in this report.



The report ends in a chapter for conclusions. A family of coefficients of monotonicity and polytonicity is presented in some detail in Appendix A. A special case of this family is used very intensively in this report. Some of the original series we used are given in appendix B. The estimation of trading days effects and their appropriate adjustment would hopefully be extended in further research in the future.

PART 1

DECOMPOSITION OF QUANTATIVE SERIES

2. SEASONAL ADJUSTMENT OF QUANTITATIVE SERIES.

This chapter presents a nonmetric technique for periodic analysis of numerical empirical series, such as seasonal time-series. Illustrations will be given for the decomposition of economic time series into (polytone) trend, fixed seasonality, and irregular components.

Existing techniques for analyzing periodic time series tackle separately the problems of estimating the period-length (for example, by spectral techniques) and of decomposing the series into trend, periodicity, and irregular components. The decomposition is often carried out by first employing one of the moving averages techniques (filters) to estimate the trend, and then fitting a function (trigonometric, polynomial or any other) to estimate periodic components. For a comprehensive survey of data analytic techniques for time-series, see Makridakis (1976). Recent development in Seasonal Adjustment is given in Pierce (1980). Discussions of specific methods are contained in Burman (1965), B.L.S. (1966), Shiskin et.al. (1967), Durbin and Murphy (1975), Cleveland, et al. (1978), Raveh (1981), Akaike (1981), and others. Fase et al. (1973) and Kuiper (1978) made an instructive comparison of several decomposition methods.

We propose here an alternative technique which is not based on either moving-averages nor on regression, as are most other approaches. Data Analysis techniques usually require a figure of merit, namely criterion of fit be maximized (or loss function to be minimized), and a set of definitions, in order to measure the goodness-of-fit or the amount of deviations from 'ideal' prespecified series. Following the above point of view, let Y_t denote the value of a quantitative time series at time t. One way for presenting a decomposition of a mixed model of Y_t into its components is given in eq. (2.1) below:

$$(2.1) \qquad Y_t = T_t \cdot I_t \cdot S_t + s_t + i_t \qquad t=1,\ldots,N$$

where T_t denotes the underlying trend at time t. S_t and s_t are the multiplicative and additive seasonal components, respectively. I_t and i_t are multiplicative and additive irregular components, respectively. The purely multiplicative model is obtained by using the constraints $s_t = i_t = 0$ for all t. A purely additive model is obtained by using the constraints $S_t = I_t = 1$ for all t.

In trying to decompose empirical time-series, one is faced with the problem of estimating at least some $2 \cdot n$ parameters (for the simplest model) from 'just' n given numerical observations. Thus, some constraints are required in order to reduce the arbitrariness in the estimation process.

Obviously, there are infinite ways to express a given series by eq. (2.1). First, we limit oursevles to a simpler model of eq. (2.2).

(2.2) $$Y_t = T_t \cdot I_t \cdot S_t + s_t \qquad t=1,\ldots,N$$

$$= Z_t \cdot S_t + s_t$$

where $Z_t = T_t \cdot I_t$ is the periodicity-free series which is known in literature as Seasonally Adjusted Data (S.A.D.). The coefficients S_t, s_t present the seasonality pattern. These coefficients could be constants or any systematic function of time depending on whether the seasonality is fixed or of a moving fashion, respectively. The trend T_t is a polytone series of order m.

Most authors estimate the trend using moving-average filters. After elimination of the trend from the original data, the seasonal component is fitted by various approaches. B.L.S. (1966) and Shiskin et.al. (1967) computed moving-averages (within months for monthly series), while Durbin and Murphy (1975) fitted the seasonal component by means of a stepwise regression method applied to additive and multiplicative Fourier components.

In this report, definitions for the periodicity (seasonality) and trend components will be given simultaneously for an 'ideal' series (e.g. series without irregular components). For empirical series including irregular components, we estimate in the first stage the seasonal components; in the second (and final stage, we estimate the trend and the irregular components simultaneously.

SOME DEFINITIONS AND NOTATIONS

A numerical time-series is a sequence of numerical observations $[Y_t]$ over some real interval $a \leq t \leq b$. Such a sequence is called periodic if the interval (a,b) can be partitioned into sub-intervals of equal length, called periods, so that there may be a change in the general level of the Y_t <u>between</u> periods but no change in the pattern of observations <u>within</u> periods. The term periodic pattern will be used loosely to designate, for a given or an assumed period length, a periodically recurring shape (if such exists) of the piecewise linear graph connecting successive points (t, Y_t).

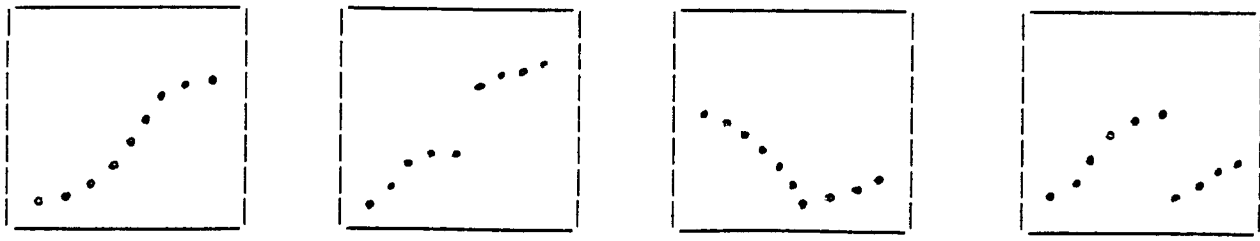
Let us restrict ourselves to the discrete case where t assumes a finite number of equally spaced values and write t=1,...,N, instead of $a \leq t \leq b$.

A series $[Y_t]$ is <u>polytone</u> of order m if there are (m-1) turning points so that the series $[Y_t]$ is monotone between successive turning points, the sign of the monotonicity on one side of a turning point being reverse of that on the opposite side of that point. The series $[|t|]$, t= -N, -N+1,...,-1,0,1,...,N, for example, is a series of polytone order m=2 with 2N + 1 elements. The single turning point is located at the (N+1)th observations.

A series $[Y_t]$ is <u>piecewise</u> <u>monotone</u> of order m if there are m segments of indexes within which $[Y_t]$ is monotone with the same direction (either positive or negative). For example, the series $t - [t]$, t=1/N, 2/N,...,i/N,...,1,...,3 (where [t] is the greatest integer which is less or equal to the quotient) is a piecewise monotone of order m=3 with 3·N observations. Clearly, if m=1 the polytone series is a monotone one.

In Figure 2.1 below, two kinds of positive monotone series, a polytone series (of order m=2) and a piece-wise monotone series are plotted.

Figure 2.1: (a) and (b) are two types of monotone series; (c) is a polytone series of order m=2. (d) is a piece-wise monotone series.



(a) (b) (c) (d)

To express $[Y_t]$ in periodic terms, it will be useful to replace the observation index t by an index of the form i + pa, where p is the proposed period length, i is the position of the observation within a period, and a is the period index in the sequence of periods, with the first indexed 0, the second 1, etc. We denote the number of complete periods by n, so that a = 0,1,...,n-1. Given this notation, a sequence $[Y_t]$ (t=1,...,N) can be written as $[Y_{i+pa}]$ (i=1,...,p; a=0,1,...,n-1).

The series $[Z_{i+pa}]$ given in eq. (2.3) is said to be linear periodic transformations of $[Y_t]$:

(2.3) $$Z_{i+pa} = (Y_{i+pa} - s_i^{(a)})/S_i^{(a)} \quad (i=1,\ldots,p;\ a=0,1,\ldots,n-1)$$

where the transformation coefficients $s_i^{(a)}$ and $S_i^{(a)}$ represent multiplicative and additive periodic coefficients, respectively. When $S_i^{(a)} \neq 1$ and $s_i^{(a)} \neq 0$, Equation (2.3) represents a mixed multiplicative-additive seasonality model. Equation (2.3) can be written in a different way:

(2.4) $$Y_{i+pa} = Z_{i+pa} \cdot S_i^{(a)} + s_i^{(a)}$$

which is similar to eq. (2.2) above and to Durbin and Murphy's model, except for the irregular component.

For fixed seasonality, the linear transformations are periodic in the strong sense, namely they depend on the period length p and on the observation position i but not on the particular period a. In other words, $S_i^{(a)} = S_i^{(a-1)}$ and $s_i^{(a)} = s_i^{(a-1)}$ for all a=1,...,n-1. Hence, for this case let us use the notation of S_i and s_i, respectively. Transformations such as (3) enable us to remove the variation in data caused by periodic effects. If the trend is monotone and the variation is proportional to it, then a multiplicative model might be appropriate. An additive model might be adequate when variation is independent on the trend level. For a constant level trend there is little difference between the above two models. Mixed models of course can capture much more complicated variations in data caused by periodic events. Models of fixed but mixed seasonality can also capture variations which look like moving seasonality where either multiplicative or additive models are adopted.

For moving seasonality, the linear transformations are periodic in a weak sense. In other words, they depend on the observation position i and on a specific function of (time) period a. Hence, $S_i^{(a)} = f_i^{(a)} S_i^{(a-1)}$ for a=1, 2,...,n-1. Fixed seasonality is a special case of moving seasonality when $f_i^{(a)} \equiv 1$ for all i=1,...,p and a=0,1,...,n-1.

If $[Y_t]$ is not a Polytone (Monotone) series, it might be possible that a period length p and coefficients $S_i^{(a)}$ and $s_i^{(a)}$ can be found for which the transformed series $[Z_t]$ is a polytone or nearly polytone. Then, such $[Z_t]$ can be regarded as an underlying (Periodicity-free) polytone trend T_t or as seasonally adjusted data namely "trend and error" respectively. The p pairs of coefficients $S_i^{(a)}$, $s_i^{(a)}$ define the periodic pattern of observations i.e., the seasonal components.

In empirical time series an irregular component usually exists, thus, the transformed series, namely, the seasonally adjusted data $[Z_t]$ is a "trend and error" curve which means that it is only approximately Polytone. In a second stage, a decomposition of $[Z_t]$ into trend and irregularity components is obtained. In order to deal with empirical deviations from ideal polytonicity a family of coefficients which designate to measure polytone association is used. (see Raveh, 1982b). Some background on a family of Monotonicity and Polytonicity coefficients is given in Appendix A. Specifically, to assess the extent to which any series, say $[Y_t]$ is polytone (of order m), the formula below is used:

$$\mu_m = \frac{\sum_{K=1}^{m} \sum_{i>j}^{I_K} (Y_i - Y_j) w_{ij} \delta_k}{\sum_{K=1}^{m} \sum_{i>j}^{I_K} |Y_i - Y_j| w_{ij}} \quad (2.5)$$

where the original series is partitioned into m consecutive sub-series I_K, K=1, ...,m and $\delta_K = (-1)^{K-1}$ within I_K. The inner summation is over all $(i,j) \varepsilon I_K$, such that i>j. The outer summation is over all m sub-series I_K, such that K=1, ...,m. The weights w_{ij} are non-negative numbers linked to each pair of observations i and j. Obviously, $-1 \leq \mu_m \leq 1$, and $|\mu_m| = 1$ only if the series is perfectly polytone, whether of positive or negative slope interchangeably.

The coefficient of Polytonicity for transformed series $[Z_t]$, t=1,...,N is given in eq. (2.6) and denoted by $\mu_m^{(p)}$. It is a function of the 2p coefficients S_t, s_t as well as the original series. The lower index m indicates the order of Polytonicity, while the upper index p indicates the period length. For the rest of this report, unless otherwise indicated, $w_{ij} \equiv$ constant and the formula will be simpler. For example,

$$\mu_m^{(p)} = \frac{\sum_{K=1}^{m} \sum_{i>j}^{I_K} (Z_i - Z_j) \delta_k}{\sum_{K=1}^{m} \sum_{i>j}^{I_K} |Z_i - Z_j|} \quad (2.6)$$

A series $[Y_t]$ is said to be a fixed periodically and polytone series if there exists a series of linear periodic transformations (2.3) in the strong sense which transforms $[Y_t]$ into a polytone series $[Z_t]$. In practice, a perfect transformation will not be insisted upon (depending on the irregular component). Instead, only a "sufficiently large" value for the figure of merit $|\mu_m^{(p)}|$ will be sought. In data analysis a critical value for goodness-of-fit does not exist as in the case of principal component analysis, as well as for the goodness-of-fit in Multidimensional Scaling methods such as Kruskal's stress or Guttman's coefficient of alienation. The researcher should have some feelings to the criteria of fit values as well as to the data. A similar unsolved problem is how to choose appropriate α level, for Statistical inference purposes. We have to keep it in mind whenever the expressions "close" or "close enough" are used.

LEAST POLYTONE TREND ANALYSIS (LPTA): WHAT IT IS, WHAT DOES IT DO?

The order of polytonicity of the unobserved trend $[Y_t]$ should be assessed. This is done by estimating the smallest order of polytonicity, m (m=1,2,...) of the original data. We keep in mind the parsimony principle, namely minimum turning points* for high fitness, and thus the procedure's name, "Least Polytone Trend Analysis." Second, if $|\mu_m|$ departs substantially from 1, say $|\mu_m| < .95$, such departure may be assumed to originate from periodic fluctuations (the seasonal components, modifying the polytonicity of the trend) or by an irregular component, or both. In such cases we search for a series of linear periodic transformations with a suitable period length p and coefficients $S_i^{(p)}$, $s_i^{(p)}$ (i=1,...,p) and function f(a) that converting the original series $[Y_t]$ into an approximately polytone series $[Z_t]$ in an optimal manner. That is, bringing $|\mu_m^{(p)}|$ as close to 1 (the theoretical maximum) as possible. The closer Max $|\mu_m^{(p)}|$ is to 1, the closer the series $[Y_t]$ is to being periodically and Polytone. This does not imply that the deseasonalized series should have as few turning points as possible. The criterion of fit (2.7), which is based on (2.6), might be 'close enough' to 1 while there may be relatively many turning points, each having small deviations from an 'ideal' basic polytone series. A minimum number of turning points is required at the step of estimating the trend component.

---

* Both turning points and outliers are inconsistent with previous, recent observations. The difference between them is that turning points are consistent with later observations while outliers are not.

The maximization of $|\mu_m^{(p)}|$ as a function of the 2p variables $(S,\ldots,S_p) = \underline{S}_p$, $(s_1,\ldots,s_p) = \underline{s}_p$ in the general mixed model (or only p variables in the simple purely multiplicative or purely additive model) for fixed seasonality, may be achieved by known Quasi-Newton or Powell, or Zangwill algorithms, see Zangwill (1967). These algorithms require an initial guess for the 2p unknown variables and by a successive procedure converge to optimal values. As an initial guess, the coefficients $\underline{S}_p = \underline{1}$, $\underline{s}_p = \underline{0}$, have been used, presumably the neutral assumption of no seasonal effects. For the usual case where $|\mu_m| < 1$ (recall that $|\mu_m^{(1)}| = |\mu_m|$), the measure $M_m^{(p)}$ is defined as the improvement in terms of Polytonicity gained by the transformations,

$$(2.7) \qquad M_m^{(p)} = \frac{\text{Max}|\mu_m^{(p)}| - |\mu_m^{(1)}|}{1 - |\mu_m^{(1)}|}$$

Clearly, $0 \leq M_m^{(p)} \leq 1$. Furthermore, $M_m^{(p)} = 1$ if and only if the series $[Y_t]$ is perfectly (without irregularities) periodically Polytone of order m where $[Z_t]$ is the polytone trend and $\underline{S}_p$, $\underline{s}_p$ are periodicity multiplicative and additive components, respectively. $M_m^{(p)} = 0$ if and only if there is no periodic component, namely the series is of "trend and error" type.

When $M_m^{(p)}$ is 'close enough' to 1, Z_t is 'only' periodicity-free series, i.e., S.A.D., and $\underline{S}_p$, $\underline{s}_p$ are the periodicity components. In the next stage, the trend T_t would be estimated simultaneously with the I_t. When $M_m^{(p)}$ is 'close' to zero it means that no periodicity component (at least with period's length p) exists and this component is negligible in further analysis. When the period length of a series is not known in advance, a previous step is to estimate it. it. This is done by selecting the smallest period length which yields a peak value "sufficiently close" to 1 for the series $M_m^{(p)}$; p=2,3..., (see chapter 3(a). The order of polytonicity is estimated as the minimal suitable parameter m. Assuming a polytone trend, the seasonal effects (the 2p coefficients) are estimated simultaneously with the periodicity-free series $[Z_t]$ and finally the trend component (T_t) is estimated simultaneously with the irregularity component (I_t).

So far we have discussed the problem of estimating both the period's length and the seasonal components for a decomposition model. The $[Z_t]$ series is a periodicity-free series representing the "trend and error". In order to separate the S.A.D. into two parts, trend and irregularities, trend T_t is estimated as the closest polytone (monotone) series to the S.A.D. In other words, the required T_t $t=1,\dots,N$ for multiplicative irregulars are those numbers that minimize (2.8a) subject to the constraint of polytonicity. For example, $T_1 \leq T_2 \leq \dots \leq T_N$ for the positive monotonicity case. I(T) mirrored the irregular magnitude. For purely additive irregular component eq.(2.8b) is used.

(2.8a) $$I(T) = \frac{1}{N} \sum_{t=1}^{N} |Z_t/T_t - 1|^p \qquad \text{usually } p=1$$

(2.8b) $$I(T) = \frac{1}{N} \sum_{t=1}^{N} |Z_t - T_t|^p$$

where

(2.9a) $I_t = Z_t/T_t$

$t=1,\dots,N$

(2.9b) $i_t = z_t - T_t$

are the multiplicative and additive irregularity components, respectively. The initial guess to the iterative process to minimize (2.8a) or (2.8b) for the trend is computed by ordering the S.A.D., separately in each tone of the trend component. The sorting is in ascending or descending order depending on the tone direction. Thus, the first guess of the trend component is just the S.A.D. ordered in a polytone shape. Likewise, their sum is equal by definition. Since the initial guess is so close to the final estimation, we use it (in this report) as a trend. In order that the trend as well as the irregular components would be unique the N numbers I_t are forced to be "around" 1 (or 0) by using the constraint $\frac{1}{N}\sum_{t=1}^{N} I_t = 1$ and $\sum_{t=1}^{N} i_t = 0$ for multiplicative and additive models, respectively.

There are three main options to proceeding with the trend estimations. These options are designated various degrees of smoothness. The three options are:

1) Smooth the trend by symmetric moving-average of length 4 with the weights 1/4(1,2,1). The first and last observation are forecasted based on formula (7.6).

2) Minimized the sum of squares, below, for given $\lambda^2 > 1$

$$(2.10) \quad \underbrace{\sum_{i=1}^{n} (z_i - T_i)^2}_{(A)} + \lambda^2 \underbrace{\sum_{i=3}^{n} (\Delta^2 T_i)^2}_{(B)} .$$

based on experience $\lambda \doteq 10^3$ is chosen. Recall, z_i the input are the seasonally adjusted data. This minimization is a trade-off between (A) the sum of the residuals and (B) the amount of distance of the trend from being local linear. As λ increase, smoother trend is obtained as well as increment in the sum of residuals (A) (smooth is in terms of local linearity).

3) Do option 2 where z_i are the estimated trend which are obtained in option 1.

Briefly, the proposed approach has the following four steps:

1. Estimate the period's length p if not known in advance.
2. Estimate the order m of Polytonicity of the trend.
3. Estimate the seasonal components $\underline{S}_p$, s_p by maximizing $|\mu_m^{(p)}|$.
4. Estimate the trend $[T_t]$ and the irregular component $\underline{I}_t$ simultaneously by minimizing I(T).

The four steps above are executed keeping in mind the parsimony principle of least order of Polytonicity (m) of the trend.

Nonmetric approaches differ from metric in that they do not use a priori metric specifications. The proposed technique is nonmetric since the loss functions that are minimized are based on deviations from Polytonicity shape and not, for example, on a sum of squares from a specified polynom as in filters of moving averages.

Classical models for an economic time-series, which typically use decomposition methods in order to produce periodicity-free (seasonally adjusted) series, estimate the seasonal component and "remove" it from the data. As Kenny (1975) said: "For one thing, it pre-supposed that we have some knowledge of the underlying trend in practice, the trend must be estimated from the same data from which we estimate the seasonal effects. Since estimating the seasonal requires knowledge of the trend, and estimating the trend involves eliminating the seasonal, there are clearly some logical problems."

The present nonmetric technique requires a minimum knowledge of the trend, i.e., that it is Polytone. Presuming a Polytone trend, we estimate the seasonal effects (the 2p coefficients) simultaneously, and assess the extents to which the trend is indeed polytone. For moving seasonality the function f(a) should be estimated as well.

The LPTA procedure could be presented by means of matrices as well. Let us restrict the discussion to the purely multiplicative model below;

$$Y_t = T_t \cdot I_t \cdot S_t = Z_t \cdot S_t \qquad t=1,\dots,N$$

The above equation could be written in the following way:

$$(2.12) \quad \begin{bmatrix} Y_{1+p\cdot 0} \\ \vdots \\ Y_{p+p\cdot 0} \\ \vdots \\ Y_{i+p\cdot a} \\ \vdots \\ Y_{1+p(n-1)} \\ \vdots \\ Y_{p+p(n-1)} \end{bmatrix} = \begin{bmatrix} Z_1 & & 0 \\ & Z_i & \\ 0 & & Z_p \\ \vdots & & \vdots \\ & Z_{i+pa} & \\ 0 & & \\ \vdots & & \vdots \\ Z_{1+p(n-1)} & & 0 \\ 0 & & Z_N \end{bmatrix} \begin{bmatrix} S_1^{(a)} \\ \vdots \\ S_i^{(a)} \\ \vdots \\ S_p^{(a)} \end{bmatrix}$$

or $\underline{Y} = Z\underline{S}$ where

$$(2.12) \quad \begin{bmatrix} S_1^{(a+1)} \\ \cdot \\ \cdot \\ S_i^{(a+)} \\ \cdot \\ \cdot \\ S_p^{(a+1)} \end{bmatrix} = \begin{bmatrix} f_1(a) & & & & \\ & \ddots & & & \\ & & f_i(a) & & \\ & & & \ddots & \\ & & & & f_p(a) \end{bmatrix} \begin{bmatrix} S_1^{(a)} \\ \cdot \\ \cdot \\ S_i^{(a)} \\ \cdot \\ \cdot \\ S_p^{(a)} \end{bmatrix}$$

or $\underline{S}^{(a+1)} = F(a)$ where F is a diagonal matrix. F matrix is a Pattern Transition matrix which transforms the seasonal pattern of any period to the next one.

Periodicity in the strong sense, or fixed seasonality, means that the Pattern Transition matrix F is identify maxtrix, i.e., F=I. In other words, the same seasonal pattern remains along time. In addition, constraints like

$$(2.13a) \quad \sum_{i=1}^{p} f_i(a) = p$$

or

$$(2.13b) \quad \frac{1}{N} \sum_{a=0}^{n-1} f_i(a) = 1 \text{ or}$$

where $f_i(a)$ = constants for all i=1,...,p, namely, $f_i(a)$ is a function of a only should hold.

Periodicity in a weak sense means moving seasonality fashion. Thus the $f_i(a)$ should be constants unequal to 1 or function of the period a and constraint like (2.13) should hold in order to keep the entire set of seasonal patterns on the same scale. The seasonal factors have the same pattern and only the amplitude is changed over time.

If $f_i(a)$ are random numbers and their arithmetic mean equal 1, it seems to me that no seasonal component exists and the series could be decomposed into trend and irregularity only.

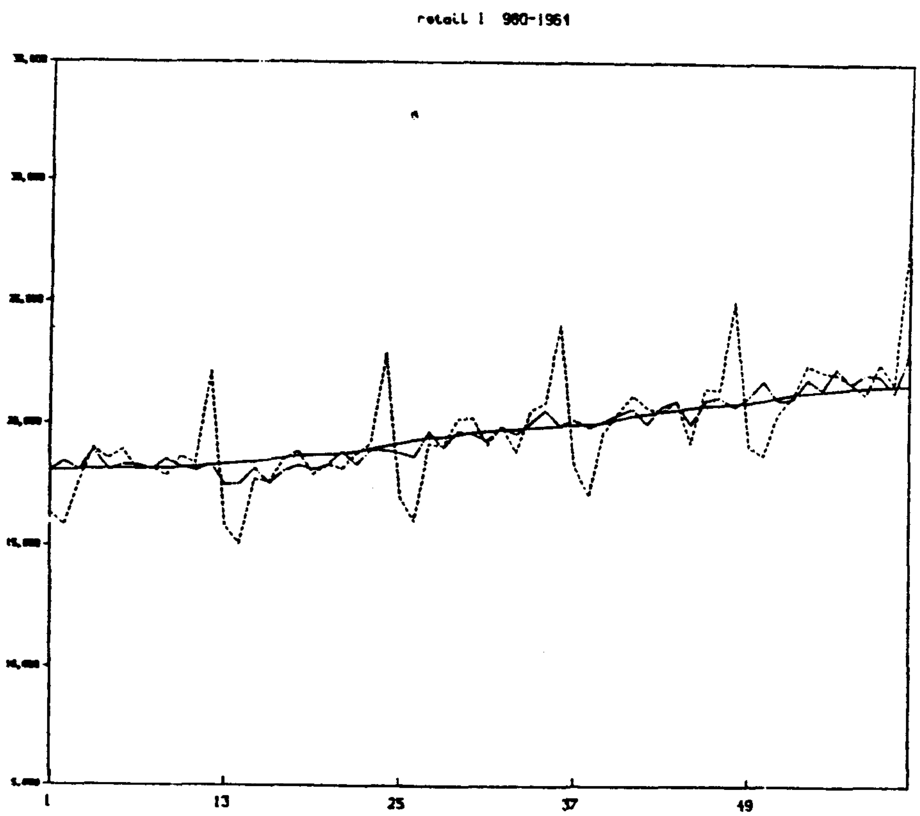
3. NONMETRIC FILTERS FOR FIXED SEASONALITY

In this chapter we restrict the discussion and the analysis of various types of series to fixed seasonal patterns. In order to clarify our approach which uses nonmetric filters let us present some empirical series by example. In the first example the transformation parameters S_i, s_i, are compared with the seasonal factors obtained by the Census X-11 and Burman's program. Some other economic monthly series from the Bureau of the Census data basis are demonstrated as well.

Example 3.1: A Study of a 5 Year Series

Consider the series "U.S. Total Retail sales in Millions of Dollars" for the years 1960-1964. This is a sub-series of the example analyzed by Shiskin et al. (1967) for examplified X-11. In figure (3.1) the graph of the original series $[Y_t]$ and the periodicity-free series, namely S.A.D., $[Z_t]$ are given. The original series is given in Table A in Appendix B.

Figure 3.1: "U.S. Total Retail Sales in millions of $ in the years 1960-1964. ...... original series; -.-. Seasonally Adjusted Data (S.A.D.)---- trend component.



Examination of the graph of $[Y_t]$ makes clear that there is an increasing trend (positive monotonicity). Also, it is easy to see that the seasonal effect is of 12 months period. The seasonal fluctuations are neither perfect proportional nor independent to the trend level and thus indicate the possibility of additive model in an equal case as multiplicative model. The coefficient of monotonicity (polytonicity of order m=1) $\mu_1^{(1)} = 0.720$ indicates a positive trend. By computing maximization of $|\mu_1|^{(12)}$ while using the purely multiplicative model we obtained Max $\mu_1^{(12)} = 0.953$. The criterion for fit $M_1^{(12)} = 0.832$ reflects a good indication that $[Y_t]$ is adequate to a periodically Polytone time-series of period length P=12. While using the additive model a slightly better goodness-of-fit was obtained; Max $\mu_1^{(12)} = 0.956$ and $M_1^{(12)} = 0.841$. The mixed model obtains slightly better results than the additive one, Max $\mu_1^{(12)} = 0.956$, $M_1^{(12)} = 0.842$. In Table 3.1 the vectors of seasonal pattern $\underline{S}_{12}$, $\underline{s}_{12}$ are given for the multiplicative and additive models as well as the "seasonal factors" obtained by X-11 and Burman's methods. In Table 3.1, the arithmetic mean for each month of the multiplicative and additive factors obtained by X-11 are given. The values for the multiplicative model are given in percentage form for comparative purposes.

Table 3.1: The (Seasonal) Periodicity components computed by the 3 methods

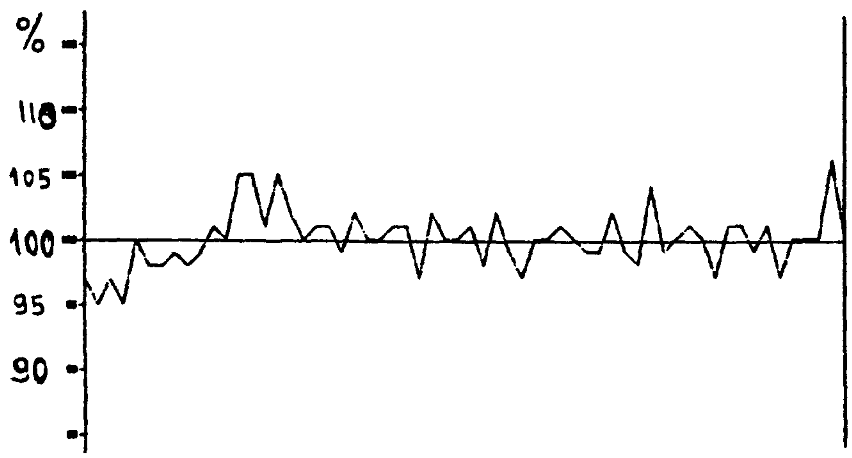
Method	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Multiplicative Model												
Burman	89.3	84.3	98.0	99.6	104.4	103.6	99.1	100.8	96.4	102.1	102.2	120.1
X-11	89.5	84.4	97.5	99.0	103.3	103.0	99.0	100.4	97.0	102.7	103.2	120.8
LPTA	90.1	86.1	97.6	100.9	102.4	103.1	98.7	100.0	96.3	102.3	101.5	120.9
Additive Model												
X-11	-2030	-2922	-462	-200	648	572	-234	72	-620	514	605	4063
Burman	-2093	-2927	-420	-111	844	678	-166	159	-790	530	448	3349
LPTA	-1862	-2737	-478	-124	428	668	-225	24	-720	483	282	4012

The values obtained by the three different methods for the two models seem to be very similar to each other.

It is necessary to keep $[Z_t]$ in the same scale of original data $[Y_t]$ by setting some constraints on seasonal patterns. A natural constraint for additive model is: $\sum_{i=1}^{p} s_i = 0$ that is, the arithmetic mean of s_i equals zero. For multiplicative model, one of the three following constraints on S_i's is suggested: $\sum_{i=1}^{p} [S_i] = p$ that is, their arithmetic mean 1; or $\sum_{i=1}^{p} [S_i]^{-1} = p$ that is, the arithmetic mean of the reciprocals equals 1; or $\prod_{i=1}^{p} S_i = 1$; that is, their geometric mean equals 1. It seems to me that the latter one is the most appropriate. Nevertheless classical methods have not adopted it.

Anyhow, X-11 program uses the second constraints for each 12 months of a calendar year. Adopting the constraints for every 12 consecutive months means that the model should have a fixed seasonal pattern. Next step is to estimate the two other components, trend and irregularity. In Figure (3.1) the monotone trend is exhibited by a solid line. Coefficient of convexity of the trend $\mu_\Delta =$ $= 0.33$ indicating a little bit of convexity trend. More details about convexity measures are given in chapter 6. In Figure (3.2) the irregular component (in percentages) is presented for the multiplicative model. The constraint $1/N \sum_{t=1}^{N} I_t = 1$ (or 100%) is used for N=60 observations. In other words, the arithmetic mean of the irregularities is 100% for multiplicative model.

Figure (3.2): Irregular component (in percentages). The arithmetic mean of the irregularities equal 1 or 100%.



Some Census Series

In order to demonstrate the trend components for longer series, fourteen empirical examples were analyzed by multiplicative model. The various values for the coefficients are given in the Figures and their seasonal patterns are presented in Table (2.5)-(2.23). In these figures the original series denoted by ...... and the trend component by a solid line ____.

The first twelve series are taken from the collection of thirteen series prepared at the Bureau of the Census for the ASA-Census-NBER conference on applied time series analysis of economic data. This conference was held October 13-15, 1981, at Washington D.C.

Table (2.5): Fixed Seasonal Patterns for the various series (Multiplicative Model).

Series	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
LSAGEMEN	91.6	92.5	94.2	99.2	102.5	107.2	107.1	105.6	104.4	103.6	98.4	93.7
BLSUEW16-19	90.7	89.9	90.0	83.0	85.8	150.2	127.4	105.7	103.6	95.3	95.5	82.3
BLSALLFOOD	96.7	95.9	96.0	95.8	96.7	99.9	101.9	106.6	106.9	103.5	101.0	99.0
Demandeposit	102.7	98.1	98.3	100.5	97.4	99.2	100.0	98.9	99.7	100.4	101.1	103.5
Currency M1A	99.6	99.0	99.6	99.4	99.6	99.9	100.4	100.3	99.8	100.1	100.5	101.6
RSWomen	80.7	76.2	93.4	96.2	96.6	92.3	89.5	97.9	100.0	105.6	108.8	162.7
WIGROCERY	99.6	99.2	101.9	100.1	99.3	100.1	98.3	97.9	99.8	101.8	101.4	100.7
RautoDLRS	88.8	92.3	108.8	106.7	111.0	110.1	103.4	103.4	93.1	103.2	93.6	85.5
INS11VS	100.0	104.7	111.4	109.1	108.3	110.0	92.2	91.9	93.7	92.1	93.8	92.7
INS36UO	90.8	99.3	90.9	89.0	90.9	102.2	106.7	114.5	115.8	108.4	95.6	95.6
INS62VS	82.5	88.0	99.9	101.8	104.0	114.5	103.2	104.4	103.8	105.5	101.4	90.9
CON-HSS1F	75.1	84.4	112.7	112.6	113.2	111.2	106.9	109.5	101.8	107.8	89.4	75.5
RVSTOR	70.8	73.4	87.5	90.8	99.3	96.1	92.4	99.4	91.6	97.5	108.3	192.8
UNEMMAN	96.5	98.2	95.3	87.0	80.4	145.5	130.8	97.4	88.7	89.1	95.7	95.2

Figure (2.2): (BLSAGEMEN) Agricultural employment, men, 20 years and older; 1-67 to 10-80. 166 observations, μ = - 0.64, μ_5 = 0.37, Max μ_5 = 0.69. Turning points are: 18,55,85,121.
Goodnessof-fit = M_5^{12} = 0.50. Source: Bureau of Labor Statistics

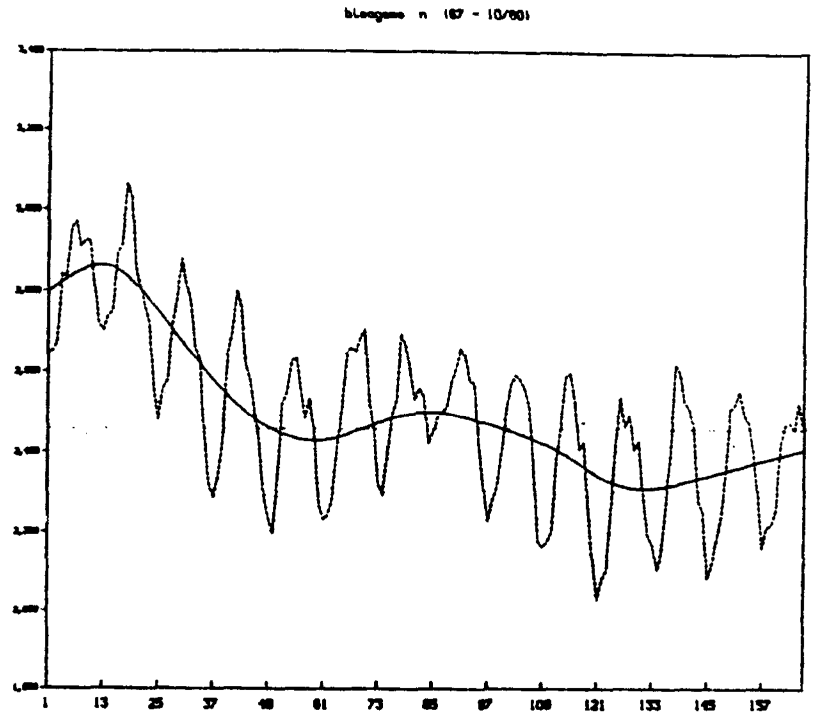


Figure (2.3): (BLSUEW16-19) Unemployment, women, 16-19, CPS data; 1-67 to 10-80. 166 observations, μ = 0.70, μ_2 = 0.58, Max μ_2 = 0.87, Turning point is 108.
Goodness-of-fit = M_2^{12} = 0.70. Source: Bureau of Labor Statistics

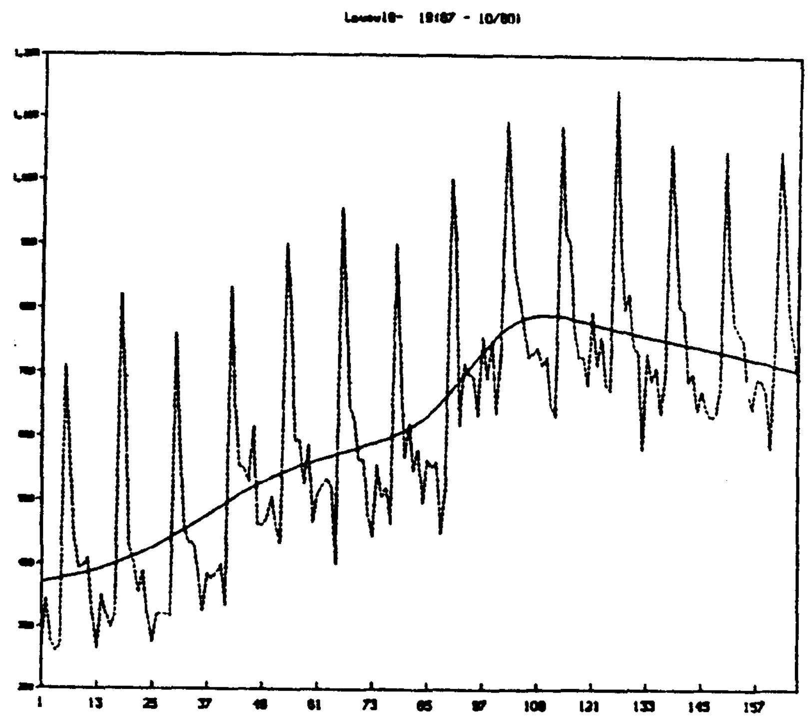


Figure (2.4): (BLSALL FOOD) All employees in food industries; 1-67 to 12-79. 156 observations. $\mu_3 = 0.38$, Max $\mu_3 = 0.90$, Turning points are: 33 and 96. Goodness-of-fit = $M_3^{(12)} = 0.84$. Source: Bureau of of Labor Statistics.

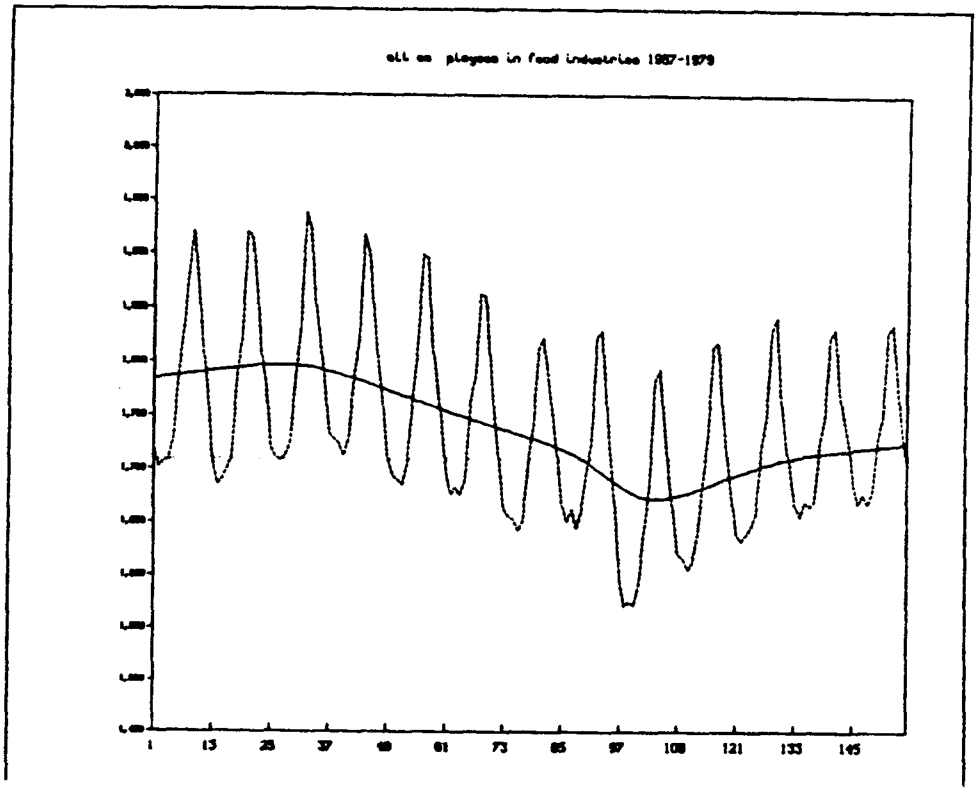


Figure (2.5): (Demandeposit) Demand Deposit component of M-1A Money Supply; 1-68 to 11-80. 155 observations. $\mu_1 = 0.99$, Max $\mu_1 = 0.999$, $\mu_\Delta = 0.04$ (= linear trend.
Goodness-of-fit = $M_1^{(12)} = 0.89$. Source: Federal Reserve Board.

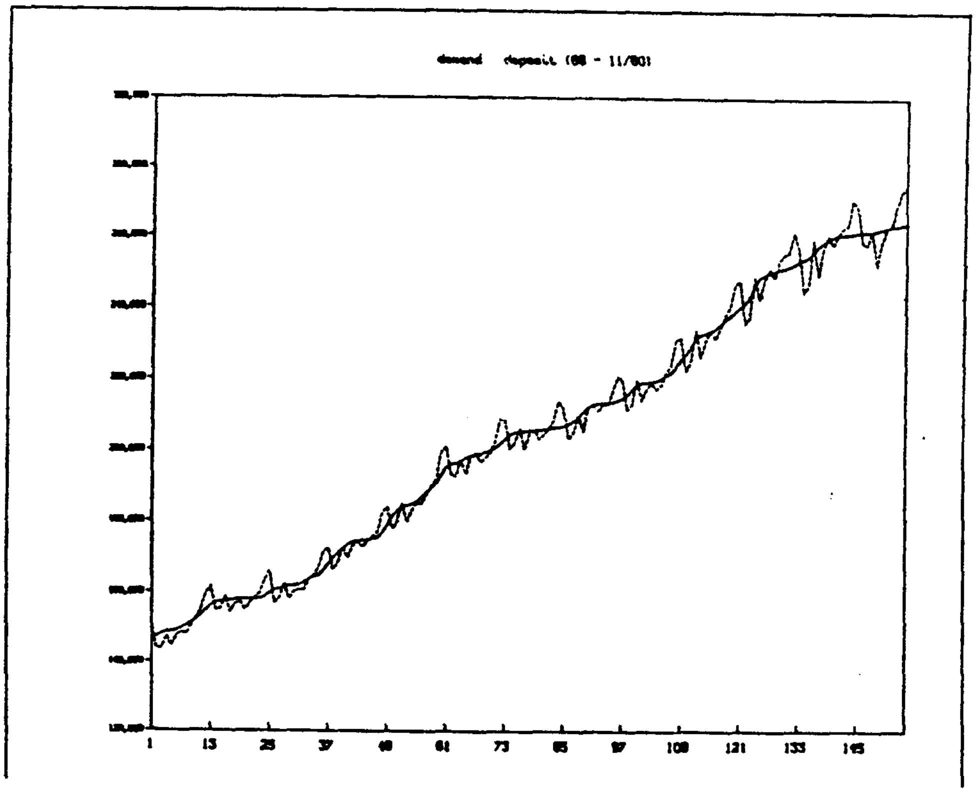


Figure (2.6): (Currency) Currency component of M-1A Money Supply. 1-68 to 11-80. 155 observations. $\mu_1 = 0.99$, Max $\mu_1 = 1.00$, $\mu_\Delta = 0.81$ ($\doteq$ convex trend).
Goodness-of-fit $M_1^{(12)} \doteq 1.0$. Source: Federal Reserve Board

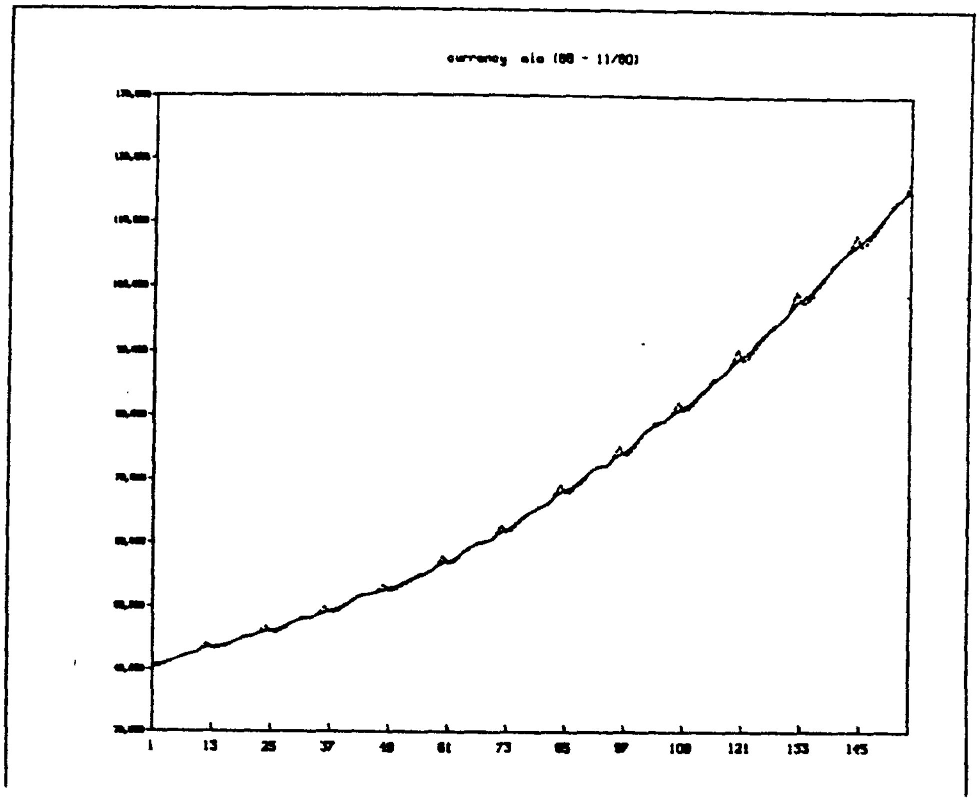


Figure (2.7): (RSWomen) Retail Sales of Women's apparel. 1-67 to 7-80. 163 observations. $\mu_1 = 0.85$, Max $\mu_5 = 0.99$, $\mu_\Delta = 0.42$.
Goodness-of-fit $M_1^{(12)} = 0.95$. Source: Bureau of the Census, Business Division.

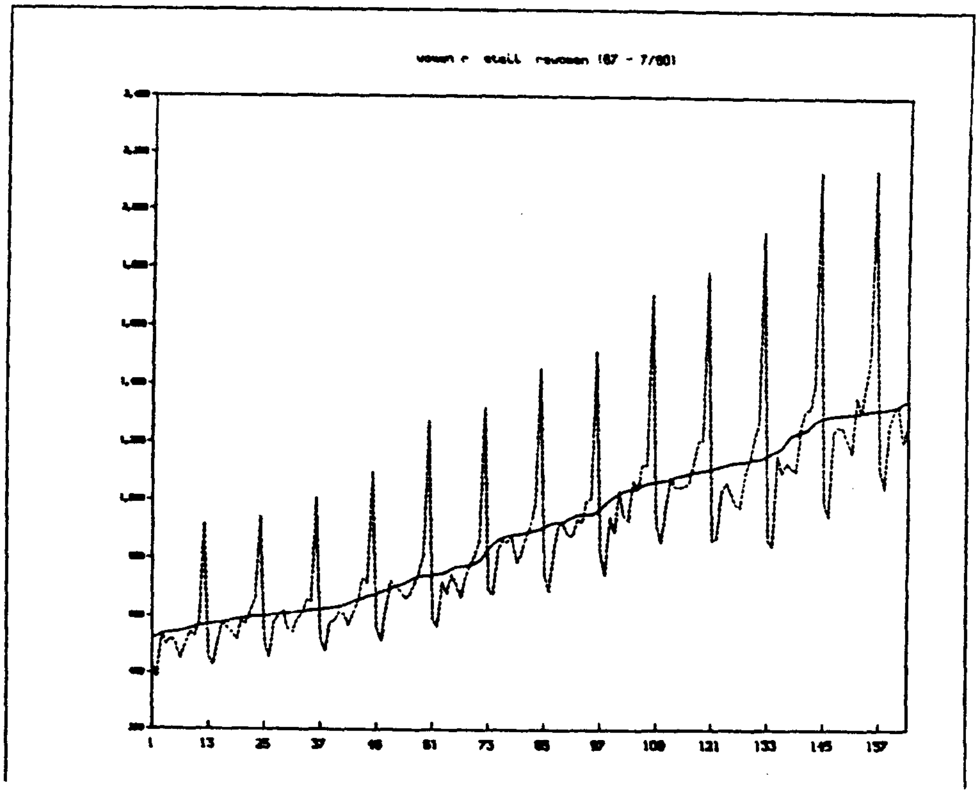


Figure (2.8): (WIGROCERY) Wholesale inventories Grocery Stores, (Mil.of \$). 1-67 to 7-80. 163 observations. μ_3 = 0.996, Max μ_3 = 0.998, Turning points are: 96 and 102.
Goodness-of-fit = $M_3^{(12)}$ = 0.56. Source: Bureau of the Census, Business Division.

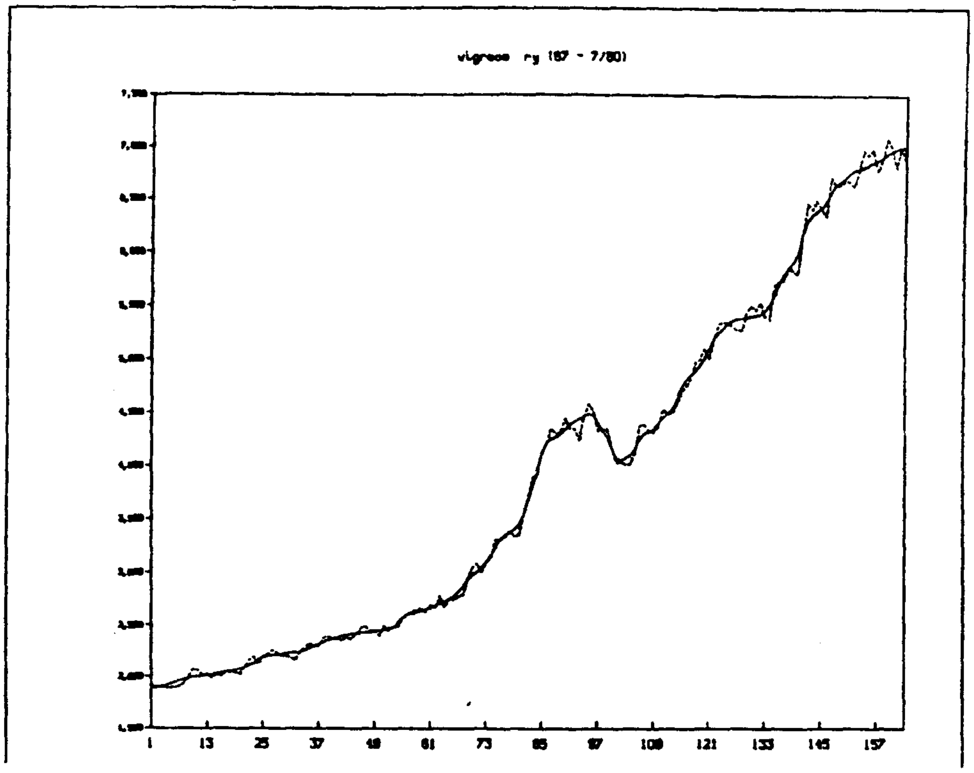


Figure (2.3): Irregular component of WIGROCERY series. (Multiplicative model.)

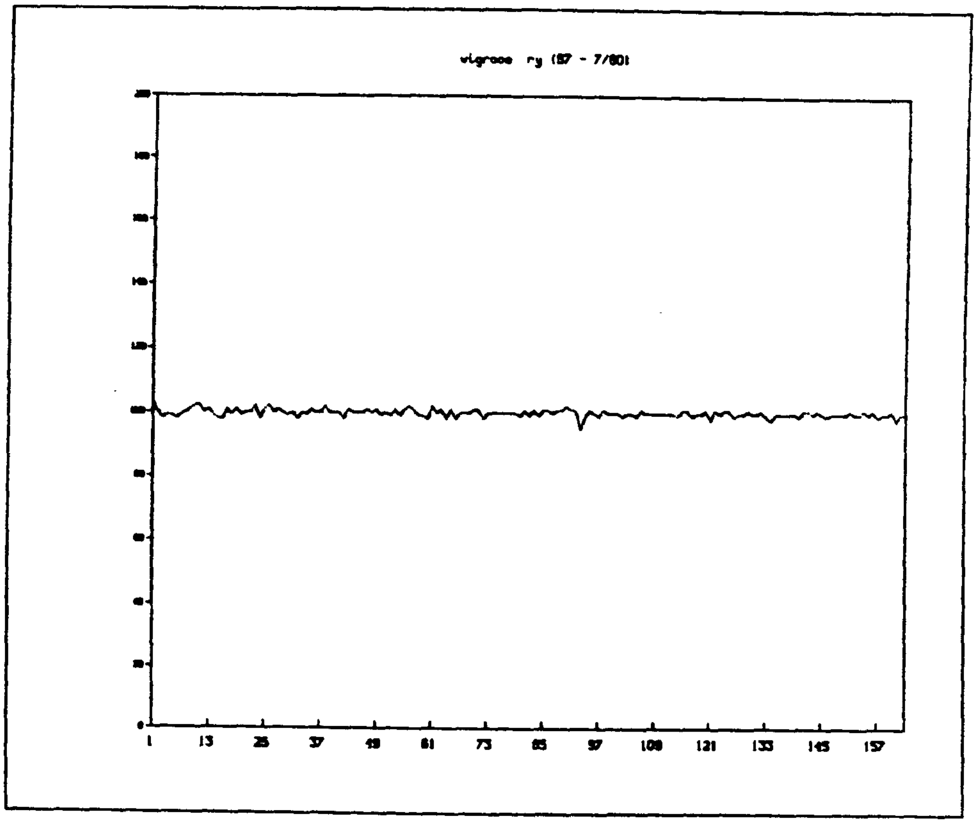


Figure (2.10): WIGROCERY series where the maximum local linearity option was adopted while smoothing trend.

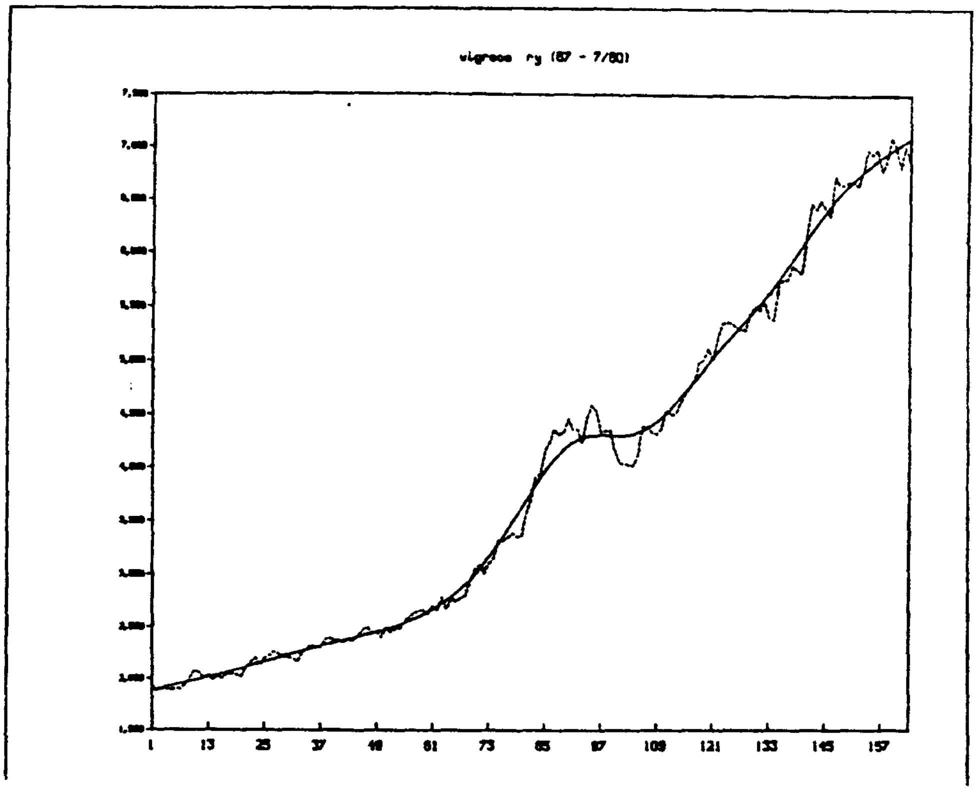


Figure (2.11): Irregular component of WIGROCERY series (local linear trend). (multiplicative model)

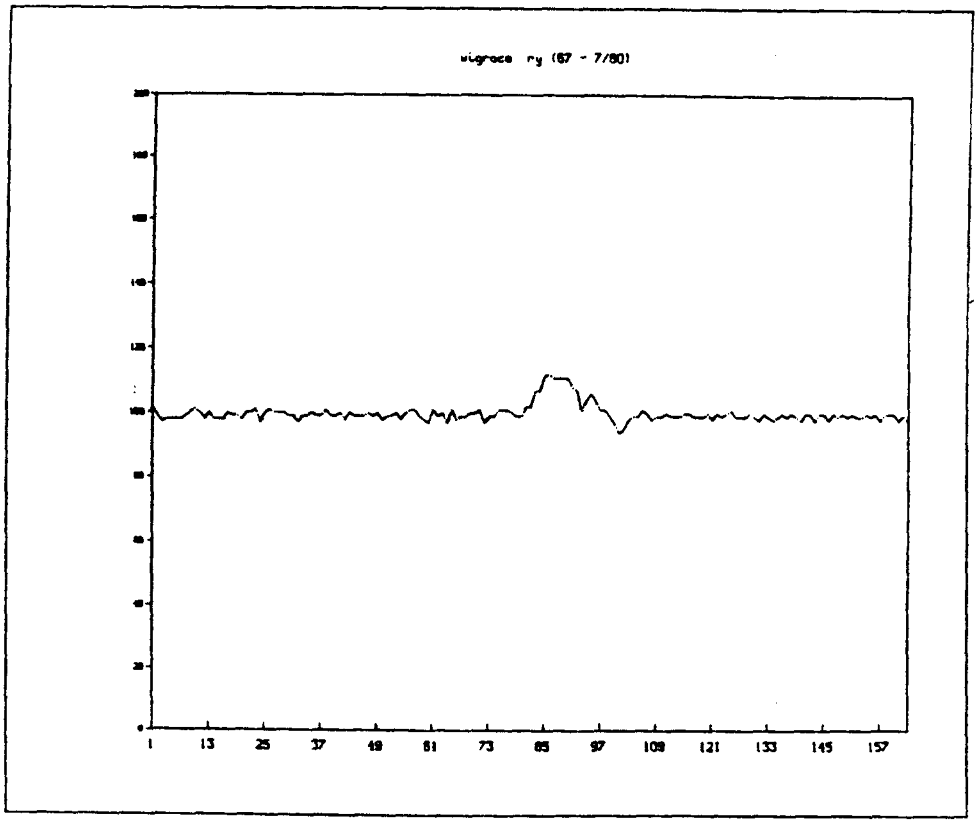


Figure (2.12): (Rautodlrs) Retail Sales of Automotive Dealers. 1-67 to 7-80. 163 observations. μ_2 = 0.86, Max μ_2=0.93, Turning point is: 84 (piecewise monotone of order 2).
Goodness-of-fit = $M_1^{(12)}$=0.48. Source: Bureau of the Census, Business Division.

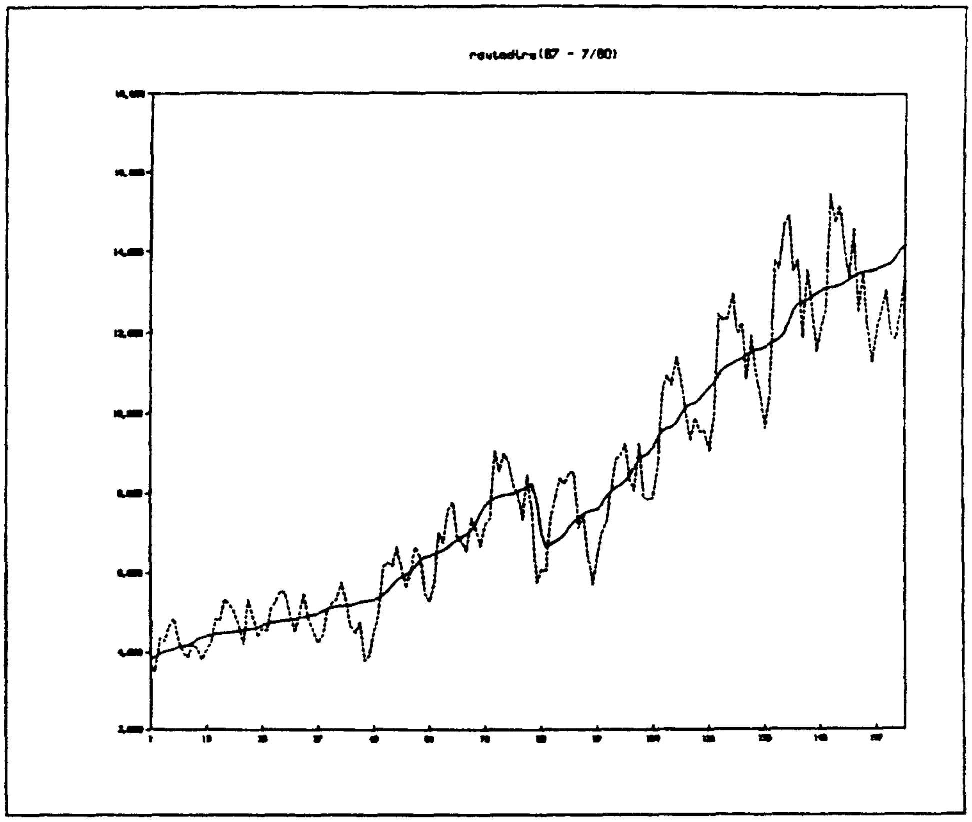


Figure (2.17): Irregular component (multiplicative) of Rautodlrs series.

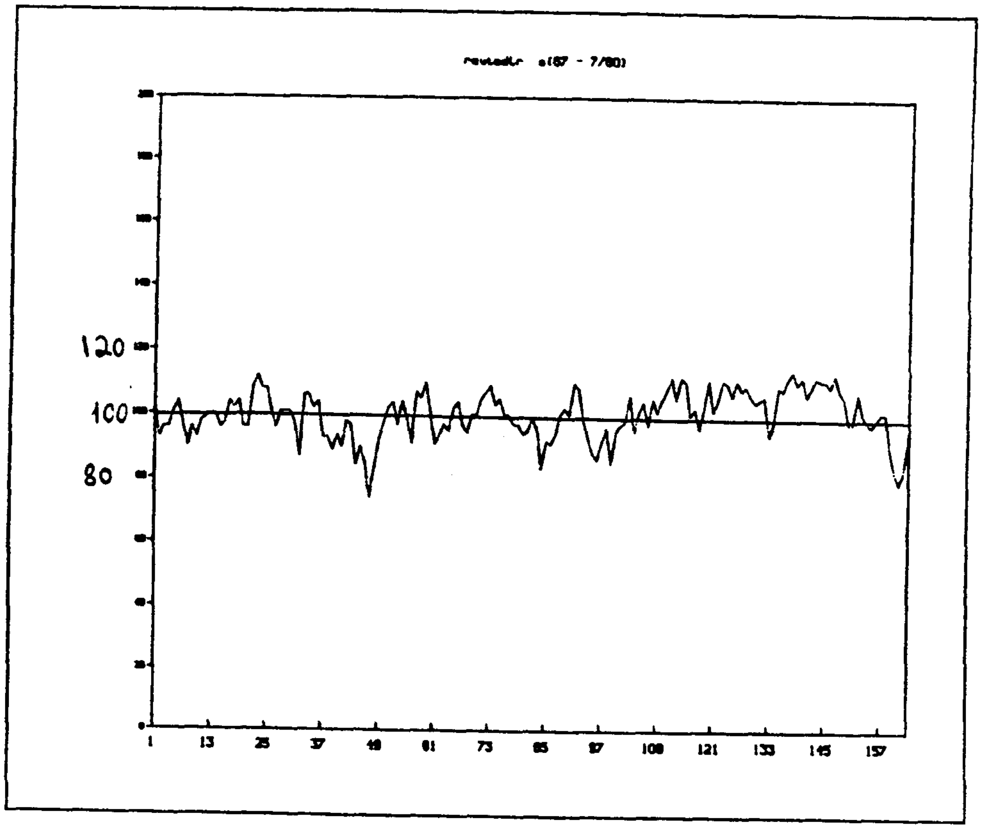


Figure (2.14): (INS11VS) Industry Value of Shipments, Blast Furnaces and Steel Mills (Mil.of \$) 1-58 to 10-80. 274 observations. $\mu_2 = 0.83$, Max$|\mu_2| = 0.86$. The turning point is 202 (piece-wise monotone trend). Goodness-of-fit $= M_1^{(12)} = 0.19$. Source: Bureau of the Census, Industry Division

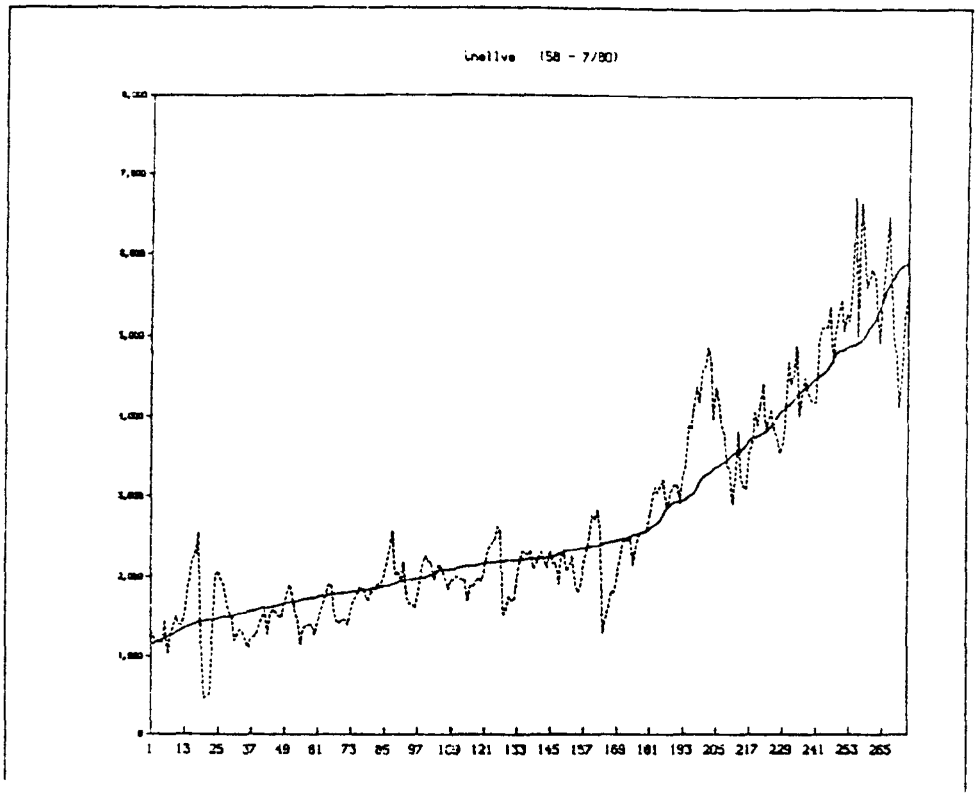


Figure (2.15): (INS36UO) Industry-unfilled orders, Radio and TV. (Mil.of \$) 1-58 to 10-80. 274 observations. $\mu_3 = 0.68$, Max $|\mu_3^{(12)}| = 0.74$. Goodness of fit $= M_1^{(12)} = 0.18$. Source: Bureau of the Census, Industry Division

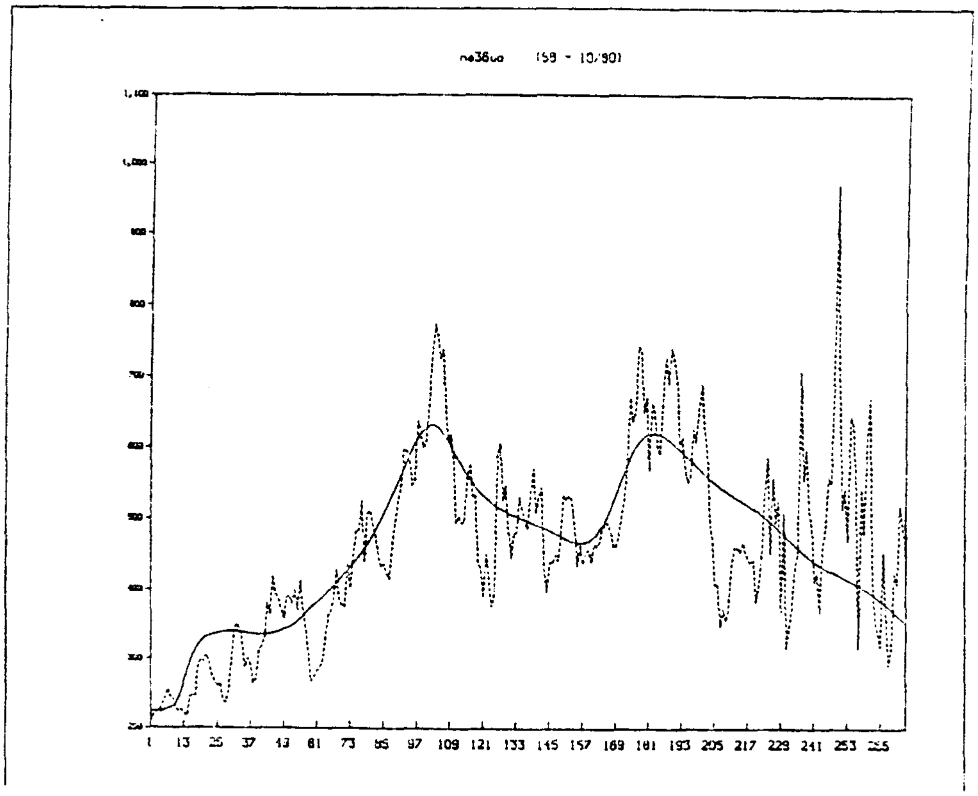


Figure (2.16): (INS62VS) Industry-Value of Shipments, Beverages (Mil.of \$). 1-58 to 10-80. 274 observations. $\mu_1 = 0.986$, Max $|\mu_1^{(12)}| = 0.998$, $\mu_\Delta = 0.91$. Goodness-of-fit = $M^{(12)} = 0.83$. Source: Bureau of Census, Industry Division.

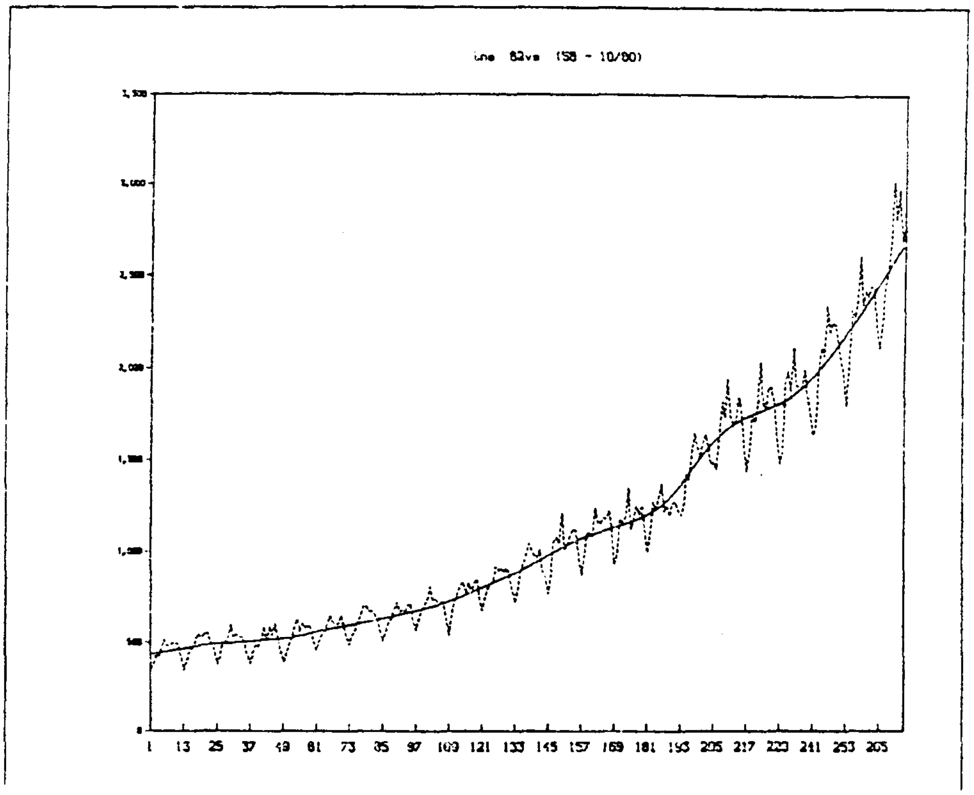


Figure (2.17): (CON-HSS1F) Housing Starts, South, Single family Dwellings. Actual No. units. 1-64 to 10-79. 192 observations. Local linear smoothness. $\mu_5 = 0.63$, Max $|\mu_5| = 0.81$. Turning points are: 72, 103, 133 and 174. Goodness-of-fit = $M_5^{(12)} = 0.48$. Source: Bureau of the Census, Construction Division.

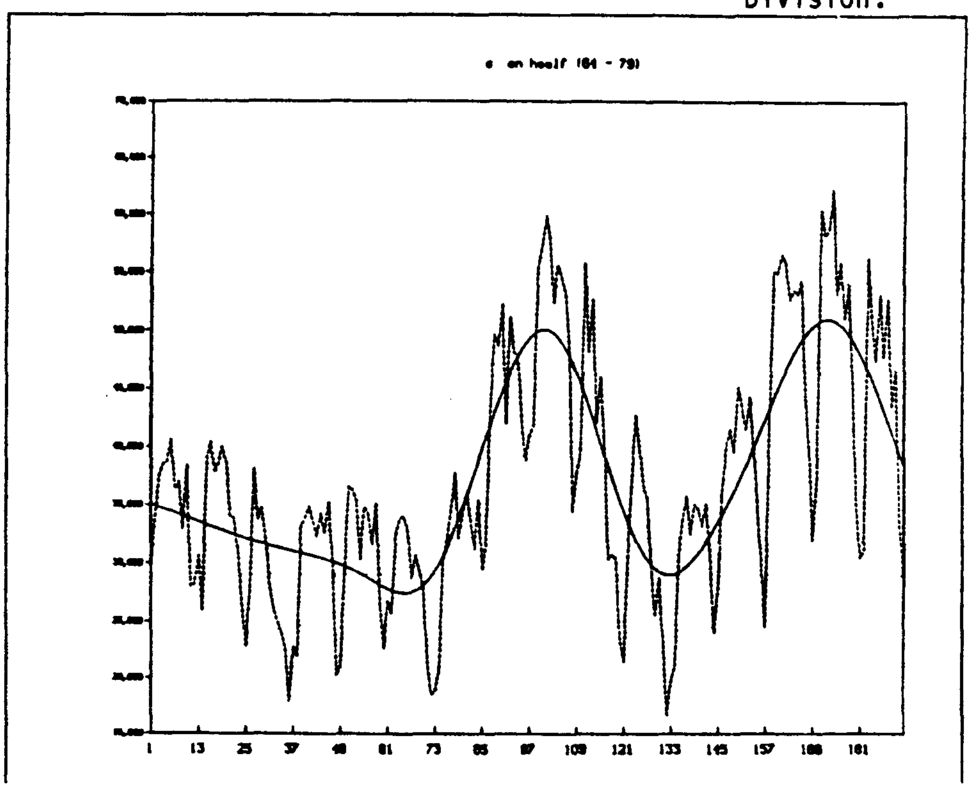


Figure (2.18): Irregular Component of CON-HSS1F series (multiplicative model)

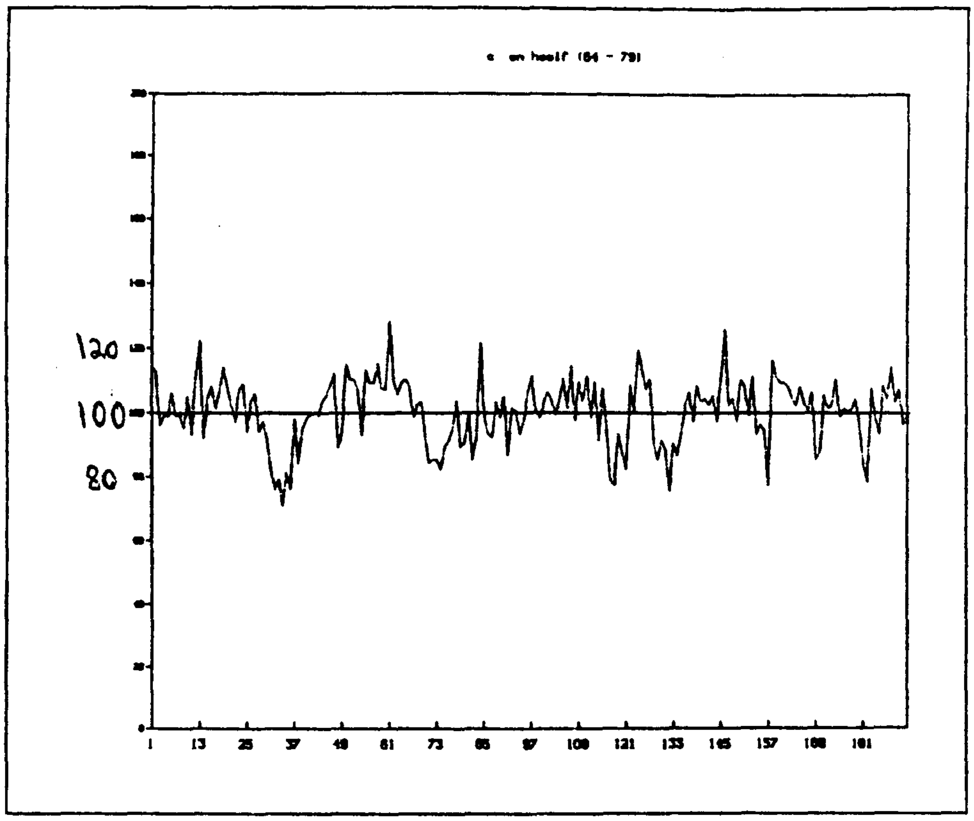


Figure (2.19): (RVSTOR) Retail Sales in Variety Stores. 1-67 to 9-73. 153 observations. μ_2 = 0.53. Max μ_2 = 0.98. The turning is on 111 observation (piece-wise monotone trend).
Goodness-of-fit = $M_2^{(12)}$ = 0.96. Local linear trend. Source: Bureau of Census

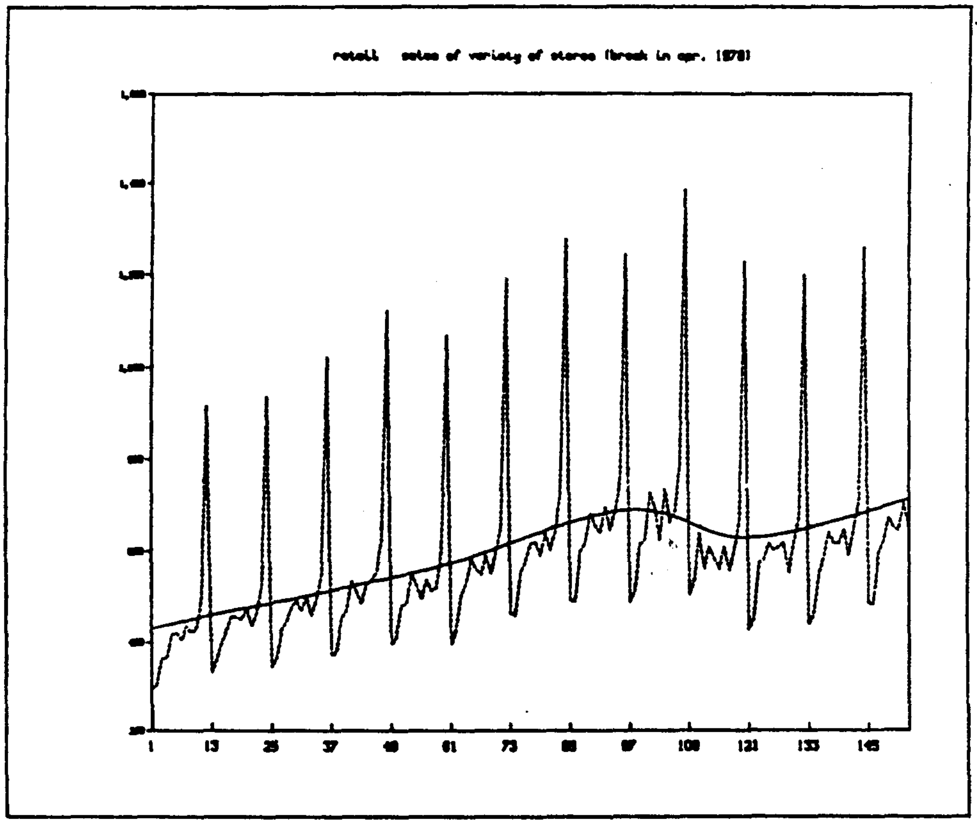


Figure (2:20): (RVSTOR) Series of Figure (2.19) local monotone trend

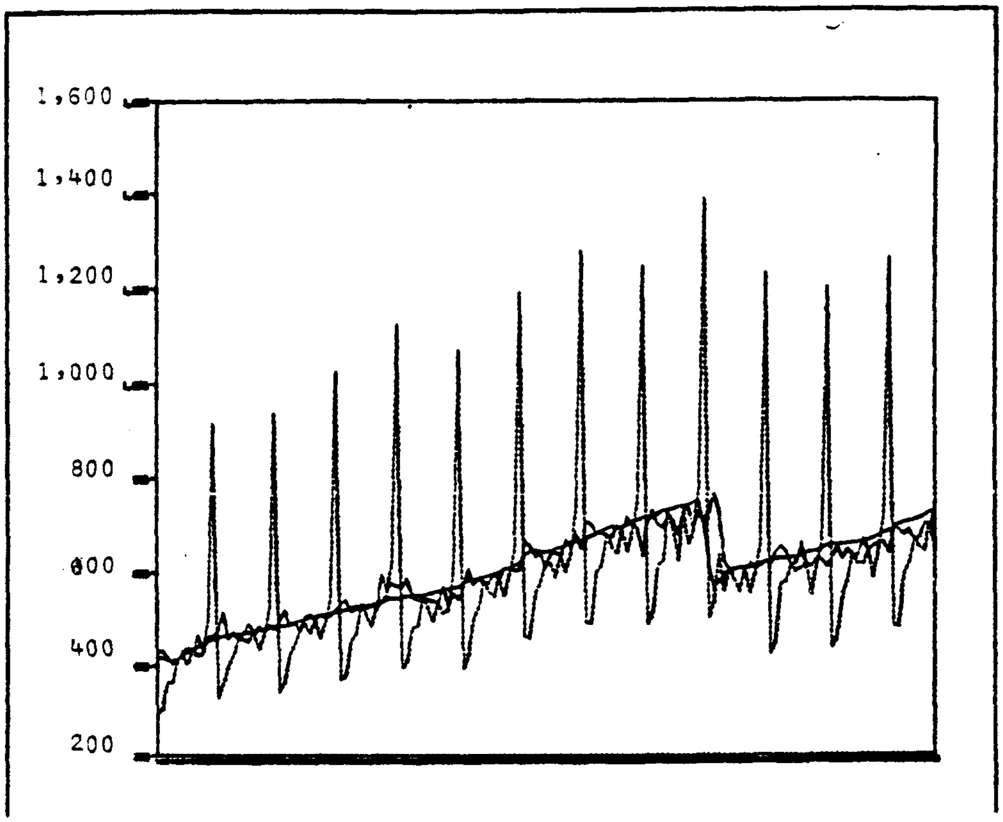


Figure (2.21): Irregularities of RVSTOR of figure (2.19)

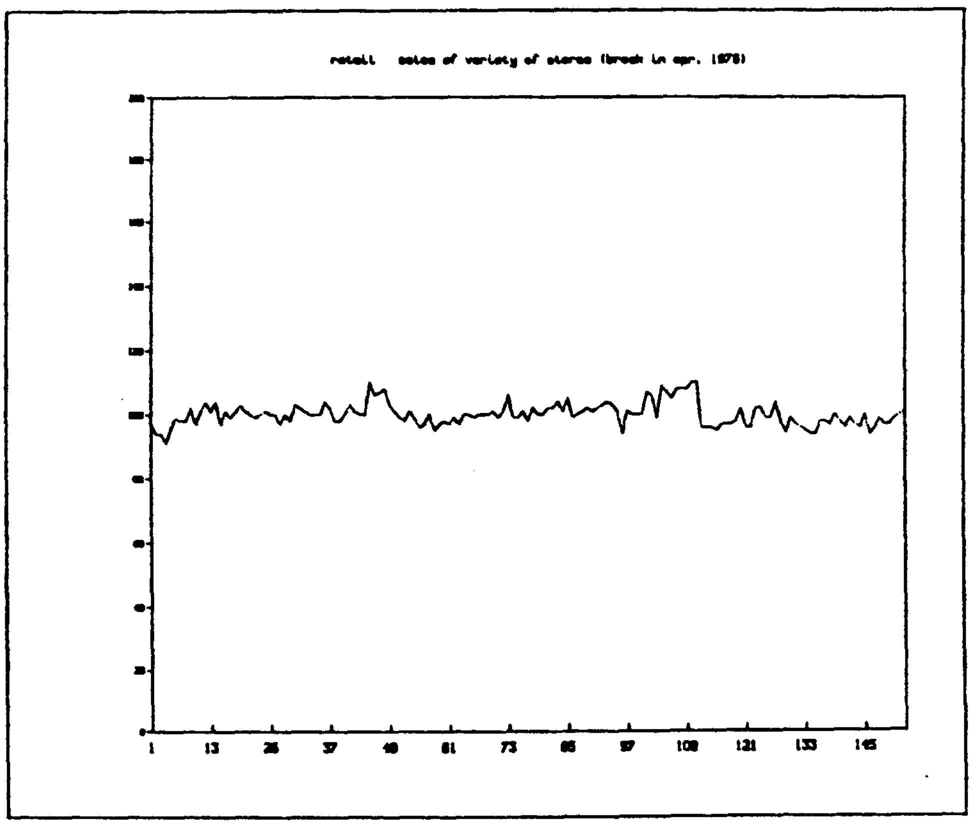


Figure (2.22): Irregularities of RVSTOR of figure (2.20)

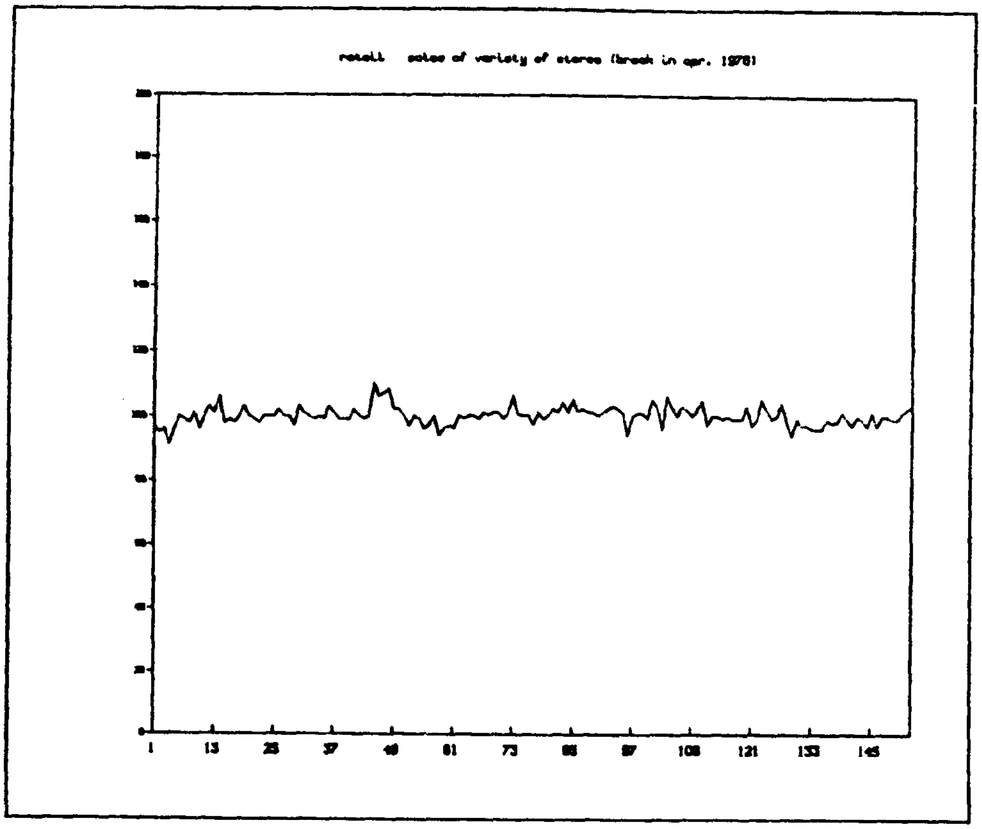
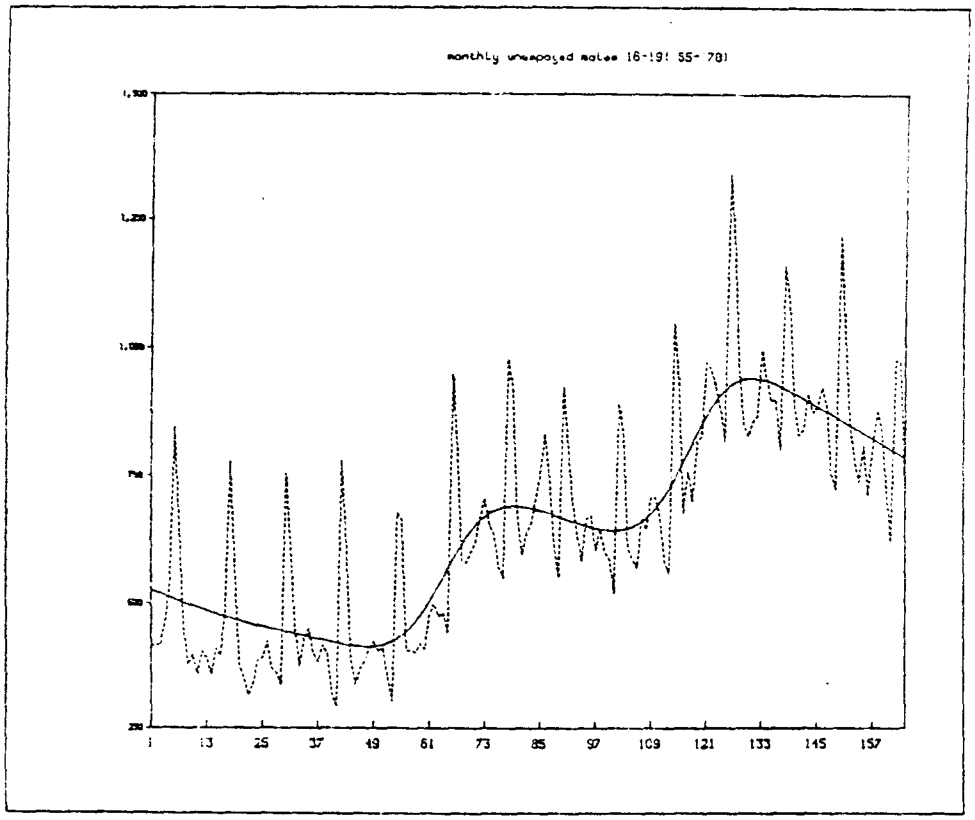


Figure (2.23): (UNEMMAN) Monthly Unemployed Males Aged 16 to 19. 1-65 to 8-73. 164 observations. $\mu_5 = 0.4$, Max $\mu_5^{(12)} = 0.68$. Additive or Moving Seasonality models seem better, see chapter 4. Goodness-of fit = 0.47. Source: Bureau of Labor Statistics



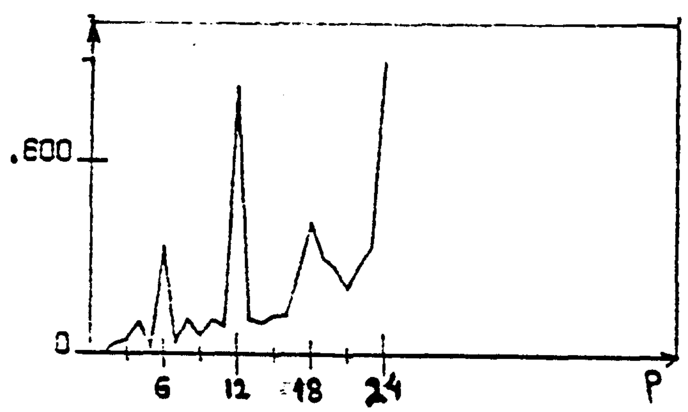
3(a) Estimating the Period's length--Nonmetric Periodogram Analysis.

The definition of period length was given earlier as the minimal P, P=2, ...,N/2 for which the definition of periodicity-polytone time series exists. For the case p=1 the periodicity is degenerate. p=2 describes the minimum period's length, i.e., the fluctuations are between every two consecutive observations. On the other hand, it is usually difficult to distinguish periodicity when there are less than two periods. It is clear that if $[Y_t]$ has period's length p, then kp $k=1,\dots,[\frac{N}{p}]$ is also a period's length.

The graph of the criterion of goodness-of-fit $M_m^{(p)}$ versus p is used to estimate the period's length. $M_m^{(p)}$ varies between 0 and 1. Intermediate values indicate intermediate measures of deviation from the ideal definition of perfect periodicity (with period length p). It is obvious that there always exists $M_m^{(k.p)} \geq M_m^{(p)}$ $k=1,2,\dots,[\frac{N}{p}]$, since the range of the function is wider. In order to estimate the period's length, it is necessary to estimate p that brings $M_m^{(p)}$ close enough to 1 and p is a prime number or the minimal multiplier of a prime number which is the minimal between them. $M_m^{(p)}$ is computed for $p=2,\dots,[\frac{N}{2}]$ (by definition $M_m^{(1)}= 0$). In other words, $M_m^{(p)}$ is a peak value 'sufficiently close' to 1.

In order to exemplify the method of period's length estimation, the values $M_1^{(p)}$ $p=1,\dots,24$ (m=1), for the series of example 1 (using the multiplicative model) are presented in Figure (3.5). The values are given in Table (2.6).

Figure (3.5): Graph of $M_1^{(p)}$ versus p (p=1,...,24).



The peak value $M_1^{(12)} = 0.832$ is very sharp and indicates p=12 as the minimal p which brings $M_1^{(p)}$ closer to 1. Of course, the value $M_1^{(24)} = 0.910$ is greater, but it is not worthwhile in that case to double the number of parameters. Incidentally, the peak value at p=6 is not sufficiently close to 1 and hence one cannot estimate it as a period's length.

The main goal of the periodogram technique and the Spectrum Analysis is to estimate periodicity (simple or complex) in a given series; see Anderson (1971) or Kendall (1973). The estimated period's length (or its reciprocal, the frequency) is achieved at the peak value by looking at the graph of the periodogram (or Spectrum or Autocorrelations) for the plotted values versus the period's length. The main idea is to seek maximum adaptation between the original series $[Y_t]$ and a trigonometric function (usually cos $\frac{2\pi t}{\lambda}$) with a known period's length. For instance, one may plot the expression:

$$S^2(\lambda) = [\frac{2}{N}\sum_{t=1}^{N} Y_t \cos \frac{2\pi t}{\lambda}]^2 + [\frac{2}{N}\sum_{t=1}^{N} Y_t \sin \frac{2\pi t}{\lambda}]^2$$

as a function of the period's length. Its peak value is obtained while $[Y_t]$ has periods of length λ. In the proposed LPTA method, the linear transformations (2.3) are more generalized than trigonometric function for discrete data.

Table (2.6): Values of $M_1^{(p)}$ for p=1,...,14 that were obtained by the LPTA method for the series of example 1.

P	$M_1^{(p)}$	P	$M_1^{(p)}$
1	0.000	13	0.108
2	0.032	14	0.097
3	0.043	15	0.123
4	0.104	16	0.121
5	0.014	17	0.258
6	0.337	18	0.411
7	0.032	19	0.304
8	0.110	20	0.269
9	0.057	21	0.206
10	0.109	22	0.279
11	0.084	23	0.341
12	0.832	24	0.910

3(b) Choosing the Appropriate Type of Seasonality

In this section usage of the figure of merit $M_m^{(p)}$ is demonstrated as a criterion for choosing the appropriate type of seasonality model from among the three following: Multiplicative, Additive, or Mixed.

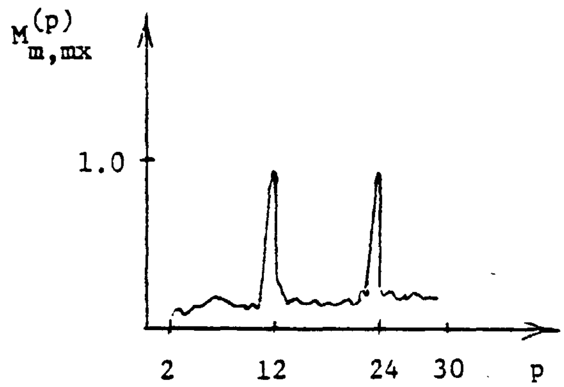
Let us denote the coefficient of goodness-of-fit $M_m^{(p)}$ which has been defined in (2.7) by $M_{m,\ mx}^{(p)}$. Thus it measures the amount of adaptation of a given series to an ideal periodically and polytone series which its seasonality shape is mixed one. The length of the period p and the order of polytonicity m are fixed through the stages of estimating the seasonality patterns. Likewise, let us denote by $M_{m,a}^{(p)}$ and $M_{m,m}^{(p)}$ the coefficients of goodness-of-fit for purely additive and purely multiplicative models, respectively. The multiplicative and additive models use simpler linear periodic transformations of $[Y_t]$ of the form:

$Z_{i+pa} = Y_{i+pa}/S_i^{(p)}$ and

$Z_{i+pa} = Y_{i+pa} - s_i^{(p)}$ $(i=1,\ldots,p;\ a=0,\ldots,n-1)$,

respectively. Obviously, each of these simpler models involve only p parameters. When the period-length p of a series is not known in advance then, our first step is to estimate it. This is done by computing $M_{m,mx}^{(p)}$ for a various periodic lengths p=2,3,... and plotting these coefficients of goodness-of-fit versus p, as presented below in figure (3.b1)

Figure (3.b1): The graph $M_{m,mx}^{(p)}$ versus p=2,3,...,30 for a typical periodic series that has period-length p=12.



The length of period-p is estimated as the smallest p which generates a peak value of $M_{m,mx}^{(p)}$ sufficiently close to 1.0. Thus a typical graph for a periodic series that has period-length p=12 is presented in figure (3.b1). Seasonal adustment methods usually used a specific period-length that can not be changed like the X-11, Shiskin et.al. (1967) that used p=12. For the three seasonality models, above, the following inequalities hold:

$$0 \leq \left\{ \begin{matrix} M_{m,aa}^{(p)} \\ M_{m,mm}^{(p)} \end{matrix} \right\} \leq M_{m,mx}^{(p)} \leq 1$$

One of the three types of models may be appropriate if the respective coefficient is close enough to 1. If $M_{m,mx}^{(p)}$ is as low as zero or nearly zero no such model is appropriate and the series might be decomposed into trend and error only.

In the case that $M_{m,m}^{(p)}$ is greater (lower) than $M_{m,a}^{(p)}$ the purely multiplicative model is better (worse) than the purely additive one.

In the case that $M_{m,mx}^{(p)}$ is only slightly greater than $M_{m,m}^{(p)}$ (or $M_{m,a}^{(p)}$) we choose the simpler model multiplicative or additive (depending on the coefficients) since they use only half the number of parameters for the model of seasonality. In the case that $M_{m,mx}^{(p)}$ is "much greater" than $M_{m,m}^{(p)}$ (or $M_{m,a}^{(p)}$) the mixed model is chosen as the appropriate model. There is no strict rule for computing the amount of difference that $M_{m,mx}^{(p)}$ has to be greater than $M_{m,m}^{(p)}$ for being chosen as the suitable model as well as the impossibility of estimating the exact number of components in principal component analysis.

Some Examples

In this section we use the procedure for choosing the appropriate seasonality models for some known empirical series used in literature in other connections.

Example 3(b1): The Chatfield-Prothero Case-Study

About ten years ago Chatfield and Prothero (1973) (C-P), investigate a case-study of "Sales of a Company X." The authors modeled this series of 77 observations and forecast 6 units ahead by using Box and Jenkins (BJ) approach. The ensuing discussion by 15 known researchers had amplified and illuminated various aspects of time series modeling, estimation and forecasting. One of the main impacts of the fruitful discussion concerned the choice of an appropriate data transformation, see Box and Jenkins (1973) and Wilson (1973). The actual use of the B-J procedure in the time series forecasting is demonstrated in detail.

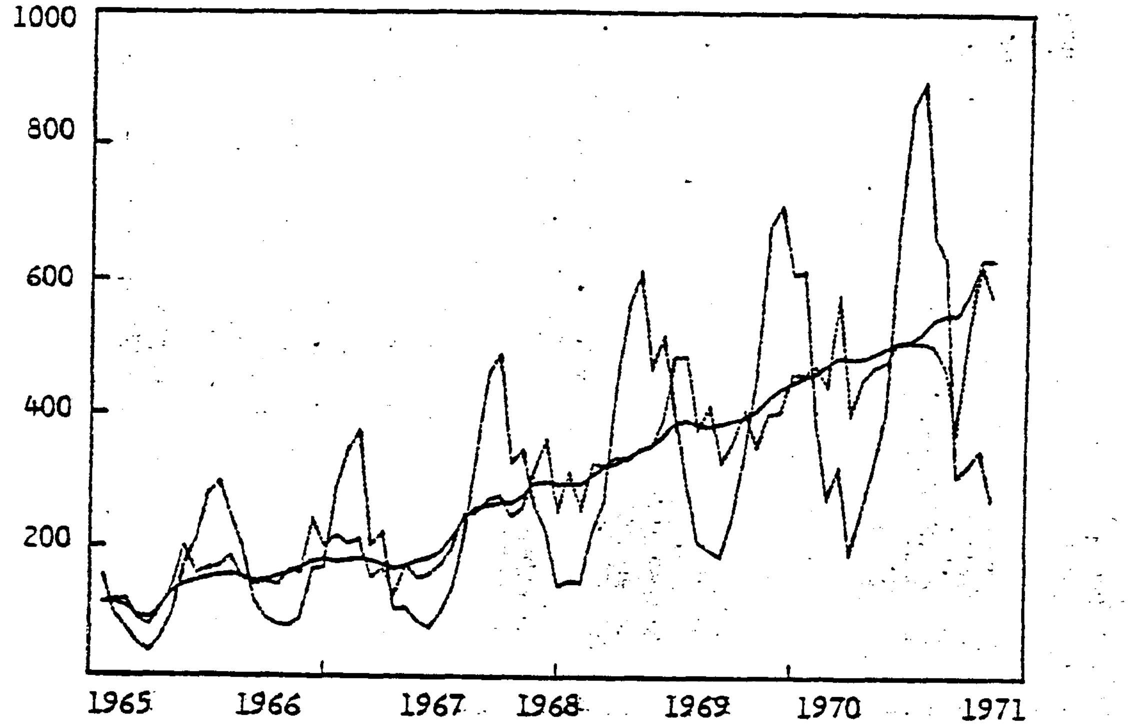
One primary reason for the publication of the case-study was the unsatisfactory results obtained by C-P. As a response, Box and Jenkins have provided a critical appraisal of the C-P paper as well as an alternative and better forecasting. An assumption that the seasonal component is of a multiplicative type was not a controversial one by any of the discussants.

Here, we pointed out that a mixed (multiplicative-additive) seasonality might be more appropriate.

The Chatfield-Prothero case-study is a monthly series of "sales of Company X" from January 1965 till May 1971. This series has 77 observations. In figure 3.b2, a chart of the series is presented and it indicates that the trend is monotone (positive slope) and the fluctuations (seasonality) are in part systematic and increased with time. The original series is given in Table B in Appendix B.

Chatfield and Prothero used the model: ARMA$(1,0)(0,1)_{12}$ on the transformed series $W_t = \nabla\nabla_{12}\log_{10}Y_t$, where Y_t $t=1,\ldots,77$ is the original series. Box and Jenkins on the other hand used the same model for power transformation: $W_t = \nabla\nabla_{12}Y^{.25}$.

Figure (3.b2): The Chatfield-Prothero case study of monthly series. Jan.1965 May 1971. ...... original series, -.-.-. Seasonally Adjusted Data (S.A.D.)------ trend component. The latter two components were obtained by mixed mode.



By using our proposed method with p=12 as the period's length and m=1 as the order of polytonicity, the computed coefficients for goodness-of-fit are given

Table (3.b1): The three types of coefficients of goodness-of-fit $M_m^{(p)}$ for the examples. For these monthly series the period's length assumed to be p=12.

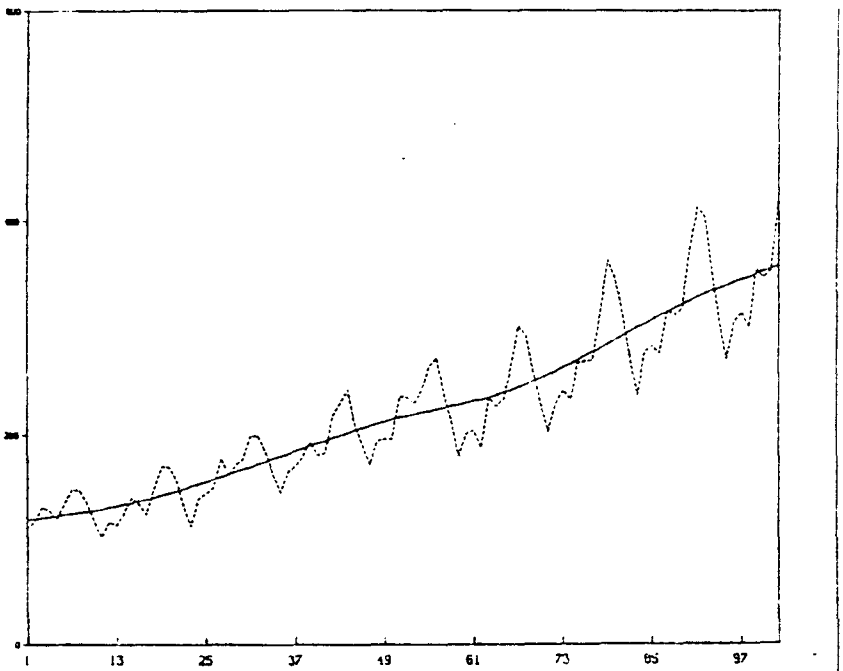
	Type of Model		
Example	Additive-$M_{1,a}^{(12)}$	Multiplicative-$M_{1,m}^{(12)}$	Mixed-$M_{1,mx}^{(12)}$
1	.71	.85	.90
2	.78	.93	.95
3	.88	.88	.89
4	.87	.88	.89
	$M_{5,a}^{(12)}$	$M_{5,m}^{(12)}$	$M_{5mn}^{(12)}$
5	.61	.73	.79

in Table (3.b1). Thus, in this case it seems that the appropriate model is the mixed one, despite the finding that the multiplicative model is substantially better than the additive one as assumed by Chatfield and Prothero and the other discussants.

Example 3b.2: International Airline Passengers in Box & Jenkins (1970, p.304)

The first 102 observations of this series were analyzed. The series is plotted in figure 3.b3 below. The trend is clearly monotone and p=12 is assumed. By looking at Table 3b.1 it is straightforward that the multiplicative model has much greater coefficient of goodness-of-fit than the additive model. The coefficient of the mixed model is slightly greater than that of the multiplicative model and thus it is not so clear which model is more appropriate using the principle of parsimony. We personally prefer the multiplicative model for this well-known series. Box-Jenkins assumed multiplicative model as well.

Figure (3.b3): First 102 observations from the series "Monthly International Airline Passengers."



Example (3.b3): Passenger Miles (Millions Flown on Domestic Services by U.K.

This series was analyzed by Anderson (1976) assuming an additive model. The series includes 119 observations from July 1962 until 1972. The order of polytonicity of the trend is m= 1, i.e., monotonicity. Coefficient of monotonicity of the original series μ_1 = 0.337 and the coefficients of goodness-of-fit for the three seasonality models are given in table 3b1. Thus it seems that both the multiplicative and additive models are equally good and "better" than the mixed model that has only very slightly greater goodness-of-fit but uses double the number of parameters.

Example (3b.4): U.S. Total Retail Sales in Millions of Dollars (Shiskin et al. 1967)

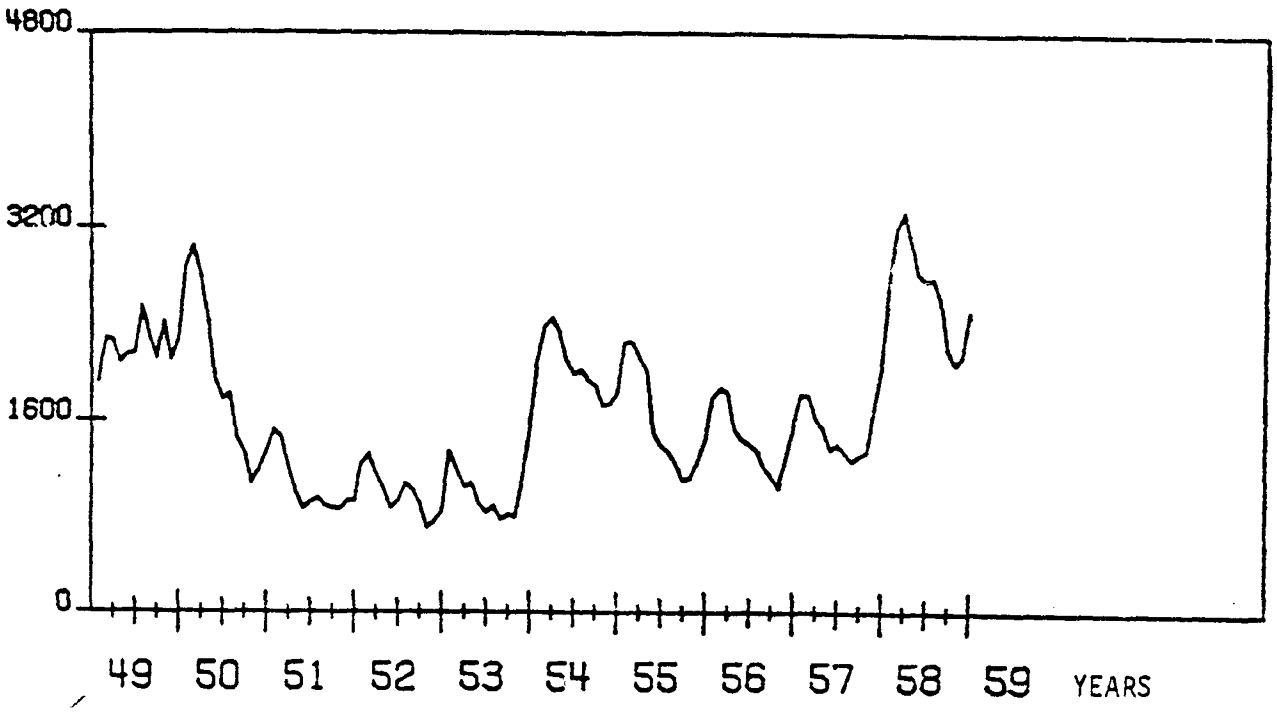
This series of 144 observations was analyzed by Shiskin et al. in their example demonstrating the X-11 method, using the multiplicative version. The coefficient of monotonicity for the original series μ_1 = 0.868 indicate an increasing monotone of the trend. From the coefficients presented in Table (3b.1) it seems that the multiplicative and mixed models have nearly the same goodness-of-fit and so the preferred model is the more simple, i.e., the multiplicative.

Example (3b.5): Unemployed Men in the U.S.A. in the years 1949-58.

This series of 120 observations was anlyzed by B.L.S. (1966) using the multiplicative model. A graph of this series is presented in figure 4. We estimated the order of polytonicity as m=5 and the turning points are 11, 53, 68 and 100 observations. The coefficient of monotonicity μ_1 is about zero but μ_5 = 0.71. The coefficients of goodness-of-fit for the three models are given in Table (3.b.1). These coefficients indicate that the mixed model is supposed to be an adequate model. A multiplicative model is much better than an additive model.

Conclusions: For purposes of demonstration five known examples were analyzed and the appropriate models were chosen. The results are in part similar to previous disuccsions and in part not. The same models were estimated for short subseries of only 35 observations. From the presented examples it seems that sometimes it is difficult to choose the appropriate model in spite of using the principle of parsimony.

Figure (3.b.4). Unemployed men in U.S.A. in the Years 1949-1958.



(3.C) Very Short Series and Missing Data

Classical methods for seasonal adjustment cannot handle series with missing data without substituting estimated values. Likewise they cannot decompose very short series. Thus, X-11, Burman and B.L.S. methods need, for example, at least 36, 60, 96 observations, respectively. The present technique overcomes these two limitations.

For the case of missing observations, zero weights $w_{ij} = 0$ are given for either i or j, the missed observations in eq. (2.5). These weights are combined according to formula (3.c.1).

$$\mu_m^{(p)} = \frac{\sum_{k=1}^{m} \sum_{i>j}^{I_k} (z_i - z_j)\delta_K \cdot w_{ij}}{\sum_{k=1}^{m} \sum_{i>j}^{I_k} |z_i - z_j| \cdot w_{ij}} \qquad (3.c.1)$$

where $w_{ij} = \begin{cases} 0 & \text{either i or j are missing data} \\ 1 & \text{otherwise} \end{cases}$

To illustrate an analysis of a short series with missing data, consider the last two years (periods) of example (3.1), "U.S. Total Retail Sales in Mil. of Dollars" in the years 1963-1964. Chart of the sub-series is exhibited in Figure (3.C.1). Assume that some observations, say, those at time points 2, 6, 7, 8, 9, 10, 11 are missing (or censored for some reason). In table (3.c.1) the series is presented. The values in the parenthesis are the missing observations. By using the multiplicative version of the proposed technique, the following values were obtained: $\mu_1^{(1)} = 0.639$, and $M_1^{(12)} = 1.0$. In Table (3.C.2) the seasonal pattern (in percentages) is given for the three series: (i) the sub-series in the years 1963-1964; (ii) the sub-series with seven missing values; (iii) the series in the years 1960-1964.

Table (3.C.1): The subseries of the two years 1963-1964 of example 1: The values in parentheses are the missing observations.

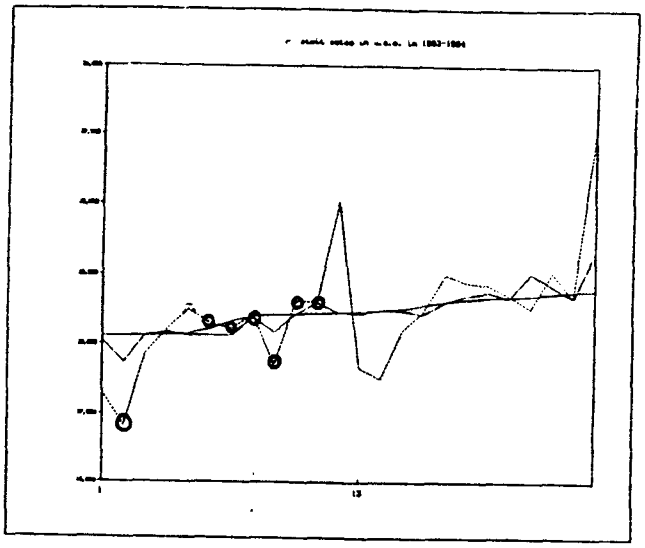
Jan	Feb	Mar	Apr	May	June	Jul	Aug	Sep	Oct	Nov	Dec
18261	(17087)	19653	20518	21228	(20737)	(20540)	(21018)	(19267)	(21528)	(21494)	25104
19154	18758	20502	21186	22508	22242	22145	21778	21313	22605	21720	27719

Table (3.c.2): Periodicity components for the 3 series

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
(i)	90.7	88.2	96.8	100.4	104.6	102.2	101.2	100.1	94.6	102.2	100.2	118.8
(ii)	92.8	89.7	98.0	100.9	104.1	101.8	101.2	99.0	95.0	99.8	95.8	121.7
(iii)	90.1	86.1	97.6	100.9	102.4	103.1	98.7	100.0	96.3	102.3	101.5	120.9

The estimated seasonal patterns for these three series are very similar to each other. This example indicates that the nonmetric technique has the property of stability. An adjustment procedure is said to be stable if the estimated seasonal pattern is not seriously disturbed by updating a series when new data become available. One can generalize this property (which is not defined mathematically) applying it to a series and its sub-series.

Figure (3.C.1): U.S. Retail Sales 1963-1964. ····· Original data, -.-.-. S.A.D. trend component (multiplicative model). The circles are for missing data.



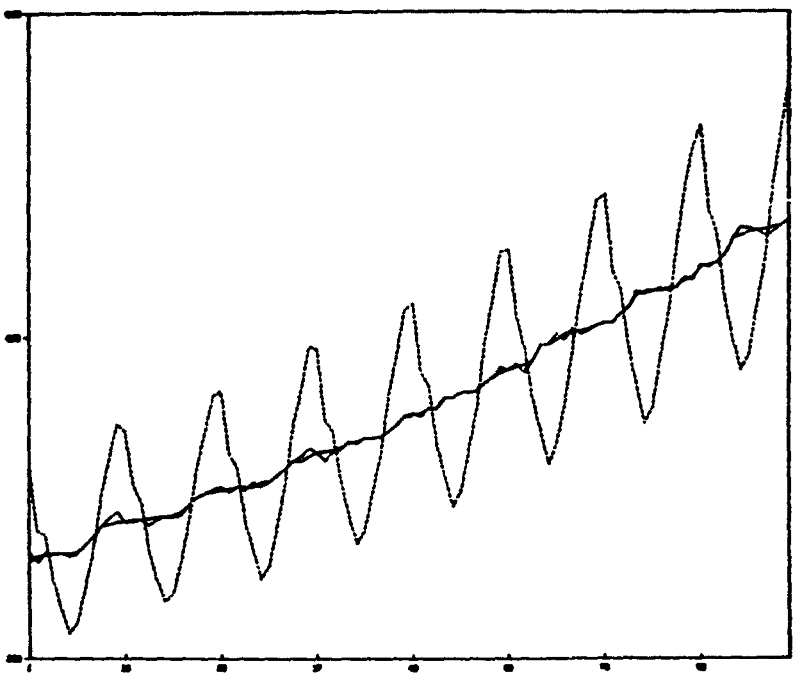
3.(D): Series With Discontinuous Trend

An invisible assumption of decomposition methods is of continuousity of the trend component. For example, it is usually possible to smooth series by moving averages filters since the trend is continuous. The recent papers by Kitagawa (1981), Akaike and Isiguro (1981) and Schlicht (1981), which do not use moving-averages, are based on the assumption that the trend is approximately locally linear, namely, that it is continuous with a specific shape.

In this section we examplify that the LPTA does not need such assumptions on the trend and the process of estimating the seasonal pattern is not distorted as, for instance in X-11 .

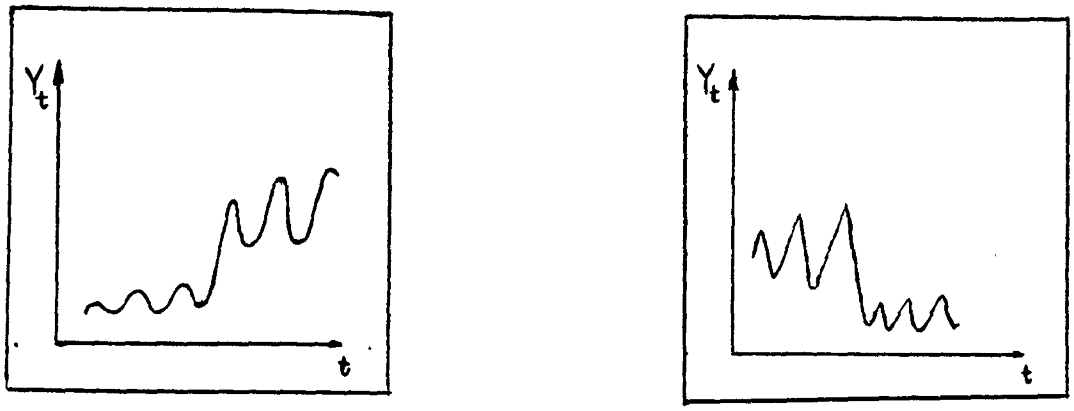
As an empirical example, let us deal with a 'nice' series: "Consumption of Electricity for Public Lighting in the U.S. in the Years 1951-1958." The original series is given in Appendix B, table C, and its graph in Figure 3.D.1. The S.A.D. and trend estimation is presented in Figure 3.D.1 as well.

Figure 3.D.1. Consumption of Electricity for Public Lighting in the U.S. in the Years 1951-1958." ...... Original Data, -.-.-. S.A.D., _____ trend component.



This series has monotone trend (or, to be more specific, this monotone trend is approximately linear), a fixed seasonal pattern, and reasonable irregularity. Let us do the following two transformations: we multiply (a) the last half and (b) the first half of the series by constants k>1 (or add constants k>0). This transforms series and their original are presented in Figure 3.D.2, below.

Figure 3.D.2: A 'nice' series that has monotone trend, a fixed seasonal pattern and reasonable irregularity. In (a) and (b) the last half and the first half of the series have been multiplied by k>1, respectively.



(a) (b)

Seasonality and irregularity have not been changed while using multiplicative (additive) decomposition model with multiplication (additive) transformations. Only the shape of the trend has been changed and thus the same estimations are exptected to the seasonal patterns of the original as well as the two transforms series. For the above series, the constant k=10 have been used for both (a) and (b) transformations. Seasonal factors extimated by X-11 for the series and its two transformations as well. The arithmetic mean of the seasonal factors, separately, for each month is given in Table 3.D.1. The fixed seasonal patterns obtained by our LPTA methods for the same series are given in Table 3.D.1 as well.

Table (3.01): The estimated seasonal (in percentage form) pattern for both the original and the transformed series as obtained by X-11 and the LPA method. (multiplicative model).

METHOD	SERIES	JAN.	FEB.	MAR.	APR.	MAY	JUN.	JUL.	AUG.	SEP.	OCT.	NOV.	DEC.
X-11	Original	119.7	106.5	103.5	93.3	87.0	81.3	83.4	89.8	96.3	106.3	112.8	119.6
	Transformation (a)	124.6	113.6	111.8	99.6	91.1	82.0	79.2	82.1	88.4	99.2	109.2	119.2
	Transformation (b)	119.4	103.1	95.7	84.6	79.8	77.5	83.7	93.9	102.9	115.6	119.8	124.2
LPA	Original	119.8	106.8	103.7	93.3	86.6	81.3	83.5	90.1	96.6	105.9	112.6	119.8
	Transformation (a)	119.5	107.6	103.3	93.2	87.4	81.9	83.9	90.4	96.5	105.4	112.1	119.0
	Transformation (b)	119.8	106.6	103.5	92.6	85.9	80.3	82.8	89.6	96.3	106.8	112.8	119.4

Very similar seasonal patterns were estimated by the LPTA method for the series before and after the multiplication (by K=10) transformations. Very different estimation was obtained by X-11 to the original series and the two transforms series which have the same seasonal pattern and irregular component by definition. The above results mirrored the fact that the LPTA method is robust against an abrupt change in the trend here. On the other hand, X-11 does not have this desirable property. An abrupt change in the trend yields 'strange' estimation results for the other components. X-11 adjusts or removes very differently the very same seasonal patterns when they are combined with different shapes of trends! The amount of distortion in the estimated seasonality (as well as irregularities) is a monotone function of the abrupt change as well.

3.(E) Series with Zero-Value Observations

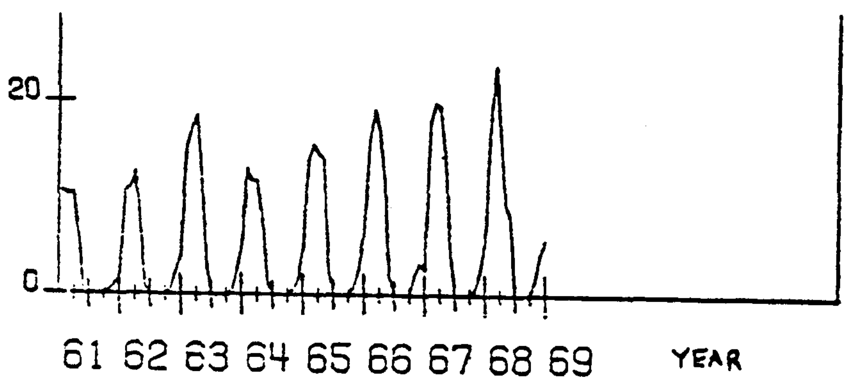
Here analysis of a series that has zero-value observations is presented. For a series with zero-value observations (not missing data) and a trend other than constant, the additive model is inappropriate by definition. This inappropriateness is caused by the fact that the deviations of the zero-value observations are proportional to the level of the trend. In other words, multiplicative model (or mixed one) might be adequate, but not the additive model.

The LPTA technique estimates the periodicity component by using a multiplicative model which the X-11, for example, cannot do. The example is "Export of citrus in millions of $" from Israel in the years 1961-1968. This series has only eight active months in a year. In June, July and August, there is no marketing; in September the marketing is almost zero; hence these four months are omitted. In Table 3.e.1 the original series is given. In Figure 3.e.1 the graph is plotted.

Table (3.E.1): Exports of citrus fruits in millions of $ from Israel in the years 1961-1968 (Original data)

YEAR	Jan	Feb	Mar	Apr	May		Oct	Nov	Dec
1961	10.7	10.3	10.5	6.0	0.0	....	.4	1.0	1.7
1961	10.6	11.1	12.7	7.2	0.8		.5	2.4	4.0
1963	15.1	17.3	18.5	12.5	3.4		.2	2.4	5.3
1964	13.1	11.8	11.8	7.8	1.3		.5	1.4	5.0
1965	13.9	15.6	14.7	14.1	2.9		.8	2.8	6.0
1966	16.2	19.2	16.9	11.7	1.9		2.1	3.3	2.6
1967	17.8	19.9	19.5	14.4	5.2		.5	2.4	5.7
1968	18.9	23.7	16.4	9.3	8.1	....	1.6	4.3	5.7

Figure 3.e.1: The Plotted Series of Table 3.e.1



X-11 program yields inaccurate estimation of series with zero-value observations. This finding will be demonstrated for the "export of citrus" example, later on in Chapter 7. The nonmetric approach decomposes this series using an additive and a multiplicative model. The period's length p=8. The estimated seasonality patterns for both models are given in Table 3.e.2.

<u>Table (3.E.2)</u>: The seasonality pattern components of the two models.

(The values for the multiplivative model are in percentages and for the additive model are absolute.)

Model	$\mu_1^{(1)}$Max	$\mu_1^{(8)}$	$M_1^{(8)}$	Jan	Feb	Mar	Apr	May	...	Oct	Nov	Dec	Mean
Multiplicative	0.18	0.71	0.65	179.6	194.6	183.6	127.6	28.7	...	7.5	26.3	51.3	100.0
Additive	0.18	0.71	0.65	6.95	8.52	7.23	2.31	-5.97	...	-8.09	-6.47	-4.49	0.0

Both models have the same goodness-of-fit based on our figure of merit $M_1^{(8)}$.

4. FILTERS FOR MOVING SEASONALITY

Seasonality, as a concept, means that systematic fluctuations around an unobservable "trend and error" exist. It is natural to think about fixed seasonality where the fluctuations are proportional or not according to the model, either multiplicative or additive, respectively. In order to extend the idea for moving seasonality, it seems that it should be done in a systematic way; otherwise there is no distinction between moving seasonality and white noise. Two ways to extend fixed seasonality are given in this chapter.

(a) Fixed Seasonality That Changed Over Time

The idea of moving seasonality could be interpreted as a fixed seasonality that change over time. In other words, various segments of time have different fixed seasonality. As an example, let us analyze the series UNEMMAN which is presented in Figure (2.23). This series is the monthly Unemployed Males (in U.S.) of Age 16 to 19 in the years between January 1965 and August 1979. This series of 164 observations was analyzed by Hillmer and Tiao (1982), and in chapter 2. In Figure (2.23)the original data and the trend component (additive model) are presented. The time axis was divided into the following 5 segments: 1-60, 48-84, 72-108, 97-132 and 121-164. The directions of the trend are -, +, -, + and -, respectively. For k=5 order of Polytonicity μ_5 = 0.4 and the goodness-of-fit and seasonal patterns for 3 models: Additive, Multiplicative and Mixed, are given in Table (4.a.1). Based on goodness-of-fit, it seems that for fixed seasonality the more appropriate model is the additive one. Let us use the LPTA Procedure for every 5 whole years (60 observations) in a moving way. Thus, we do seasonal adjustment for the sub-series 1965-1969, 1966-1970, 1967-1971 and so on, up to 1974-1978.

Table: (4.a.1): Goodness-of-fit = $M_5^{(12)}$ and seasonal Patterns for 3 models.

Model	$M_5^{(12)}$	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mult.	0.42	96.5	98.2	95.3	87.0	80.4	145.5	130.8	97.4	88.7	89.1	95.7	95.2
Add.	0.54	-32.3	-22.2	-36.2	-82.0	-122.3	305.7	193.8	-17.9	-69.5	-61.0	-30.0	-26.1
Mixed	0.57	197.3	114.7	103.2	96.4	88.6	102.8	104.3	100.0	95.0	88.2	104.5	94.9
		-61.8	-88.2	-49.1	-57.8	-54.0	259.0	151.5	-15.2	-39.8	6.6	-50.8	-0.3

The turning points are those that estimated for the whole series. Since the most interesting part of analyzing is the last part, we analyzed, in addition, the last 4, 3 and 2 whole years, respectively. As a matter of fact, we had 168 observations, namely, from 1.65 to 12.78, and we analyzed these sub-series accordingly. In Table (4.a.2) the estimated seasonal patterns for the various sub-series are given (multiplicative model used). One can verify that over time the seasonal pattern changed gradually and usually the goodness-of-fit for each segment increased.

Table (4.a.2): The estimated seasonal patterns and goodness-of-fit obtained by multiplicative model for various ranges of years.

Years	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	goodness of-fit
65-69	87.46	89.30	88.73	78.95	70.77	169.64	144.40	96.67	87.29	92.71	96.87	97.23	0.27
66-70	90.18	91.30	91.25	80.52	70.71	166.93	143.70	94.29	87.01	91.88	96.32	95.91	0.30
67-71	94.29	94.45	91.85	80.27	71.93	160.71	141.98	94.96	87.09	91.65	95.05	95.76	0.64
68-72	99.92	96.20	94.27	82.01	75.94	145.39	134.85	97.30	89.42	90.80	96.14	97.77	0.50
69-73	99.80	98.13	94.13	85.77	79.60	141.12	128.21	97.42	89.37	89.54	98.00	98.91	0.64
70-74	101.70	100.84	93.61	86.47	79.48	139.88	127.45	96.92	89.55	89.13	97.96	97.81	0.58
71-75	102.16	100.38	95.45	87.55	80.22	138.01	127.32	96.56	89.40	87.52	96.88	98.55	0.63
72-76	102.33	101.61	95.66	90.78	81.51	137.02	123.76	97.01	90.24	88.01	96.63	95.45	0.70
73-77	102.48	102.38	96.98	91.73	84.59	137.48	124.81	99.07	88.53	85.89	93.89	92.17	0.58
74-78	99.99	101.21	97.72	90.63	83.73	135.32	121.60	98.47	90.89	89.77	95.48	94.18	0.72
75-78	101.15	101.25	97.38	90.13	82.94	133.13	121.02	98.16	90.88	90.91	97.60	95.45	0.73
76-78	99.01	103.20	98.55	87.08	81.12	120.04	120.00	97.35	91.93	92.98	100.94	97.81	0.74
77-78	98.37	99.82	96.98	84.82	78.50	129.64	123.18	96.52	93.69	93.94	102.47	102.07	0.60

The entire series

65-78	96.5	98.2	95.3	87.0	80.4	145.5	130.8	97.4	88.7	89.1	95.7	95.2	0.47

(b) Change of the Amplitude

Moving seasonality could be interpreted as a change of the amplitude of the fixed seasonal pattern over time. As an example, let us deal with the Chatfield-Prothero case-study exemplified in chapter 3.b. The monotonicity of the original series is $\mu_1 = 0.68$ and the $M_1^{(12)} = 0.85$ (multiplicative model). The final seasonal factors for fixed seasonality and moving seasonality are given in Table (4.b.1) below:

Final Fixed Seasonal Factors
sum to (1200.00)

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
140.8	94.7	71.9	65.0	54.5	49.7	45.3	84.6	122.3	167.6	173.6	129.9

Final Moving Seasonal Coefficients

144.0	96.9	73.5	66.6	55.7	50.9	46.3	86.6	125.1	171.5	177.6	132.9
127.6	85.8	65.1	59.0	49.4	45.1	41.0	76.7	110.8	151.9	157.3	117.7
142.5	95.8	72.7	65.9	55.1	50.3	45.8	85.7	123.8	169.6	175.6	131.5
147.4	99.1	75.2	68.1	57.0	52.1	47.4	88.6	128.0	175.4	181.7	136.0
151.5	101.8	77.3	70.0	58.6	53.5	48.7	91.0	131.5	180.2	186.7	139.7
142.8	96.0	72.9	66.0	55.2	50.5	45.9	85.9	124.1	170.0	176.0	131.7
131.5	88.4	67.1	60.7	50.9	46.4	42.3	79.0	114.2	156.4	162.0	121.3

The goodness-of-fit for the moving-seasonality (multiplicative) model is $M_1^{(12)} = 0.89$, a slightly less than the mixed (fixed) model. The multipliers for the seven periods, respectively, are: 1.02, 0.91, 1.01, 1.05, 1.08, 1.01, 0.93. Their arithmetic mean is about 1.00

5. COMPLEX SEASONALITY

The numerical nonmetric approach (LPTA) to data analysis of (single) periodic series is extended here for series that have complex seasonality. Complex seasonality means that several periodicity components of different periods interact simultaneously. Thus, the main feature is to simultaneously make adjustments for the different periods.

We use the notations and definition of chapter 3 as much as possible. A generalized formula of (2.1) for multi-seasonal components model either multiplicative or additive is

$$
\begin{aligned}
Y_t &= T_t \cdot I_t \cdot S_{t1} \cdot S_{t2} \cdots S_{tR} + s_{t1} + s_{t2} + \cdots s_{tV} \\
(5.1) \qquad &= T_t \cdot I_t \cdot \prod_{r=1}^{R} S_{tr} + \sum_{v=1}^{V} s_{tv} \\
&= Z_t \cdot \prod_{r=1}^{R} S_{tr} + \sum_{v=1}^{V} s_{tv}\ , \quad t=1,\ldots,N
\end{aligned}
$$

In (5.1) R multiplicative and V additive seasonality components are involved. For the sake of simplicity, let R=V for the remainder of this article, and thus our basic model becomes

$$
(5.2) \qquad Y_t = Z_t \cdot \prod_{r=1}^{R} S_{tr} + \sum_{r=1}^{R} s_{tr} \quad t=1,\ldots,N\ .
$$

Let us call the model expressed by (5.2): A complex periodicity (seasonality) model of order R.

As an example, let us think about daily transportation volume over years. This complex periodicity series of order R=2 has two main periods: The weekly period of length 7 (days) and monthly period of length 12 (months). Estimating first the weekly seasonality and then the monthly seasonality yields different final seasonal adjusted series than estimating first the monthly seasonality and then the weekly. Both ways are not attractive. The more accurate way is to adjust both seasonal patterns simultaneously, which is the aim of this chapter.

Definitions and Notations

In the complex periodicity model of order R, R different components of periodicity (each one has different period's length) are involved. Thus the t-th observation Y_t is within the R different periods. For instance, the 8th observation in the daily volume transportation example is simultaneously located first within the weekly period--first day in the week (in the second period, according to weeks order), and first within the yearly period--first month in the year.

To express $\{Y_t\}$ simultaneously in R periodic terms, it will be useful to replace the observation index t by a complex index of the form $\bigcap_{r=1}^{R} (i_r+p_r a_r)$. R is the complex periodicity order of the model, P_r is the period's length in the r-th period ($p_1<p_2< \ldots <p_r$), i_r is the location of the observation within the r-th period (of length p_r), a_r is the period index in the sequence of periods, with the first indexed 0, the second 1, etc. Thus the observation t is presented in R ways simultaneously:

$$t = i_r + p_r a_r \quad i_r = 1,\ldots,p_r,\ a_r = 1,\ldots,\left[\frac{N}{p_r}\right]$$

for r=1,2,...,R.

It means that the 4th observation is located in the i_r observations within period number a_r that is length is p_r and it holds simultaneously for each of the R (r=1,...,R) components. It is symbolized as previously said:

$$t= \bigcap_{r=1}^{R} (i_r + p_r a_r)$$

It is easy to see that when R=1 (simple seasonality model) t=i+pa.

Given this notation, a series $\{Y_t\}$ t=1,...,N can be written as

$$\{Y_{\bigcap_{r=1}^{R}(i_r = p_r a_r)}\}, \quad i=1,\ldots,p_r, r_r=1,\ldots,[\frac{N}{p_r}], \quad r=1,\ldots,R$$

By a sequence of R linear periodic transformations of $\{Z_t\}$ with $p_1, p_2, \ldots, p_R$ as the period's length we shall mean a series $\{Z_t\}$ whose members are of the form

$$Z_{\bigcap_{r=1}^{R}(i_r+p_r a_r)} = (Y_{\bigcap_{r=1}^{R} i_r+p_r a_r} - \sum_{r=1}^{R} s_{ir}^{(p_r)})/\prod_{r=1}^{R} S_{ir}^{(p_r)} \tag{5.3}$$

$i_r = 1,\ldots,p_r$; $a_r = 1,\ldots,[\frac{N}{p_r}]$, $r=1,\ldots,R$: $(p_1 < p_2 < \ldots < p_R)$ where the transformation coefficients $S_{i1}^{(p_1)},\ldots,S_{iR}^{(p_R)}$ and $s_{i1}^{(p_1)}\ldots,s_{iR}^{(p_R)}$ represent multiplicative and additive periodic coefficients, respectively. When $S_{ir}^{(p_r)} \neq 1$, $r=1,\ldots,R$ and $s_{ir}^{(p_r)} \neq 0$, $r=1,\ldots,R$ then equation (5.3) represents a pure additive or a pure multiplicative model, respectively. Equivalently, (5.3) can be written as

$$Y_{\bigcap_{r=1}^{R}(i_r+p_r a_r)} = Z_{\bigcap_{r=1}^{R}(i_r+p_r a_r)} \cdot \prod_{r=1}^{R} S_{ir}^{(p_r)} + \sum_{r=1}^{R} s_{ir}^{(p_r)}, \tag{5.4}$$

which is similar to formula (2.4)

If $\{Y_t\}$ is not a polytone (monotone) series, it might be possible that R periods of length $p_1,\ldots,p_R$ and coefficients $S_{ir}^{(p_r)}$, $s_{ir}^{(p_r)}$ can be found for which the transformed series $\{Z_t\}$ can be regarded as an underlying (periodicity-free) polytone trend of $\{Y_t\}$ known as S.A.D. and the $2 \cdot p_1 p_2 \ldots P_R$ of coefficients $S_{ir}^{(p_r)}$, $s_{ir}^{(p_r)}$ define the periodic pattern of observations, i.e., the complex periodicity components. To assess the extent to which any series, say $\{Y_t\}$ is polytone (or order m), we shall use the formula (5.5).

$$(5.5) \qquad \mu_m = \frac{\sum_{k=1}^{m} \sum_{i>j}^{I_K} (Y_i - Y_j)\delta_k}{\sum_{k=1}^{m} \sum_{i>j}^{I_k} |Y_i - Y_j|}$$

where $\delta_k = (-1)^{k-1}$ within I_K, $k=1,\ldots,m$. I_k is the k subseries (among m). The inner summation is over all $(i,j) \in I_k$, such that $i>j$. The outer summation is over all m subseries $I_{;}k$, such that $k=1,\ldots,m$. Obviously, $-1<\mu\leq 1$, and $|\mu_m| = 1$ only if the series is perfectly polytone.

The coefficient of polytonicity for the transformed series $\{Z_t\}$, $t=1,\ldots,N$, will be designated by $\mu_m^{(\underline{P})}$ in (5.6) where, as noted earlier, $t = \bigcap_{r=1}^{R} (i_r + p_r a_r)$. The lower index m indicates the order of polytonicity, while the upper index $\underline{P} = (p_1,p_2,\ldots,p_R)$ indicates the period length.

$$(5.6) \qquad \mu_m^{(\underline{P})} = \frac{\sum_{k=1}^{m} \sum_{i>j}^{I_k} (Z_i - Z_j) \cdot \delta_K}{\sum_{k=1}^{m} \sum_{i>j}^{I_K} |Z_i - Z_j|} .$$

The Procedure

The procedure for achieving seasonally (complex) adjusted data from the original series is very similar to that described earlier in chapter 2 and only a sketch is given here.

The main idea is to search for a sequence of R linear periodic transformations of $\{Y_t\}$ with $\underline{P} = (p_1,p_2,\ldots,p_R)$ as the period's length and coefficients $S_{i1}^{(p1)},\ldots,S_{iR}^{(pR)}$ and $s_{i1}^{(p1)},\ldots,s_{iR}^{(pR)}$ $(i_r = 1,\ldots,p_R)$ converting the original series $\{Y_t\}$ into an approximately polytone series Z_t in an optimal manner. That is, bring $|\mu_m^{(\underline{P})}|$ as close to 1 as possible. The closer $\mathrm{Max}|\mu_m^{(\underline{P})}|$ is to 1, the closer is the series $\{Y_t\}$ to being complex periodically polytone.

The maximization of $|\mu_m^{(\underline{P})}|$ as a function of the $2p_1p_2 \ldots p_R$ variables $(S_1^{(p_1)} \ldots S_{p1}^{(p1)} = \underline{S}_1^{(p_1)}, s_1^{(p_1)},\ldots, (S_1^{(pR)},\ldots,S_p^{(p_R)}) = \underline{S}_R^{(p_R)};$

$(s_1^{(p_1)},\ldots, s_{p_1}^{(p_1)}) = \underset{\sim}{s}_1^{(p_1)},\ldots,(s_1^{(p_R)},\ldots,s_{p_R}^{(p_R)}) = \underset{\sim}{s}_R^{(p_R)}$ in the general mixed model (or only $p_1 p_2 \ldots p_R$ variables in the simple pure multiplicative or pure additive model), may be reached by optimization algorithms such as that of Zangwill (1967). For an initial approximation for these coefficients we use the model with no complex seasonal effects, e.g. the $\underset{\sim}{S}_r^{(p_r)} = \underline{1}$ and $\underset{\sim}{s}_r^{(p_r)} = 0$ for, $\ldots,p_R$. For the multiplicative model, the constraints $\underset{\sim}{s}_r^{(p_r)} = 0$ are set up, while for the additive model, the constraints $\underset{\sim}{S}_r^{(p_r)} = 1$ are set up.

For the usual case where $|\mu_m| < 1$ a measure $M_{\underline{m}}^{(P)}$ of goodness-of-fit is defined:

$$(5.7) \qquad M_{\underline{m}}^{(P)} = \frac{\text{Max}|\mu_{\underline{m}}^{(P)}| - |\mu_m|.}{1 - |\mu m|}$$

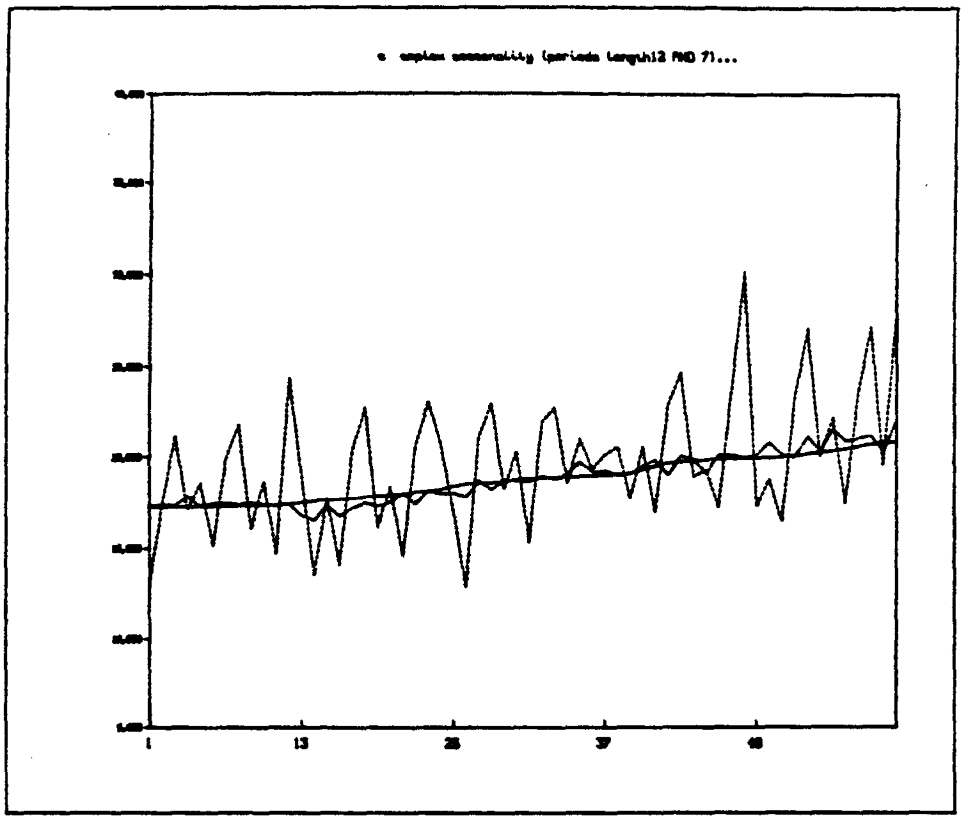
Clearly, $0 \leq M_{\underline{m}}^{(P)} \leq 1$. Further $M_{\underline{m}}^{(P)} = 1$ if and only if the series $\{Y_t\}$ is perfectly complex periodically polytone of order m. That means that model (5.2) is perfectly adequate with $I_t = 1$ for all t=1,...,N (an ideal series without irregularities. $M_{\underline{m}}^{(P)} = 0$ if and only if there are no periodic components and the model $Y_t = T_t \cdot I_t$ (t=1,...,N) is adequate.

The described procedure requires knowledge in advance of the R period's length: $p_1,p_2 \ldots p_r$. Thus, wherever these values are not known in advance, the first step is to estimate them. This is done by selecting as the optimal period's length the vector $\underline{P} = (p_1 \ldots p_r)$ with smallest coordinates $(p_1<p_2\ldots<p_R)$ which generates a $M_{\underline{m}}^{(\underline{P})}$-th peak with the smallest $\underline{P}$ values sufficiently close to 1.

<u>An Example</u> (artificial)

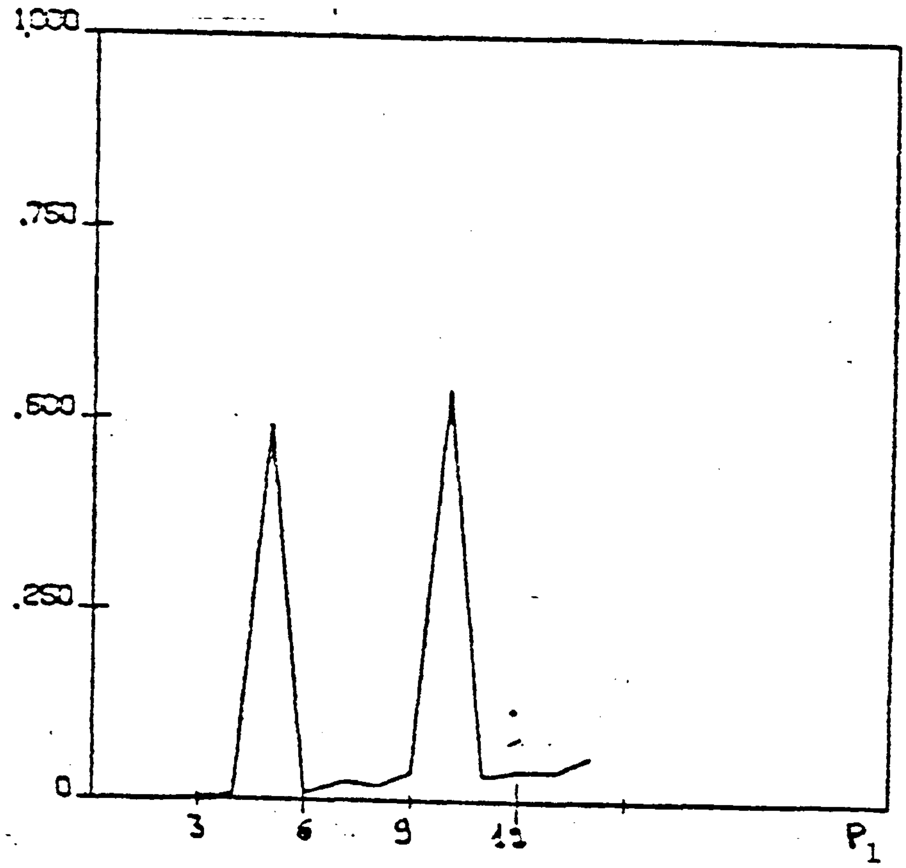
In Figure (5.1) the graph of original series Y_t t=1,...,60 is given. The actual series is in appendix B, Table D.

Figure (5.1): Chart of original series Y_t, t=1,...,60 denoted by ....; S.A.D. denoted by -.-.- and trend component by _____ .



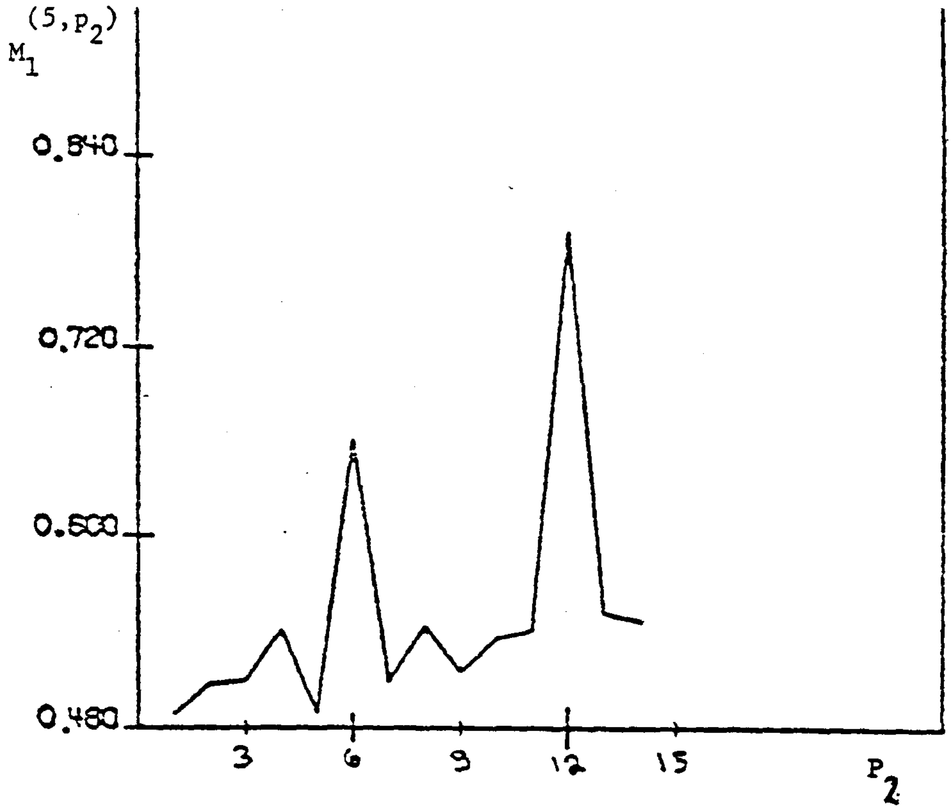
As described earlier, the procedure looks for the optimal coefficients that converts the series Y_t as close as possible to a polytone series Z_t with the minimal order of polytonicity m. By looking at this example it is easy to see that the order of polytonicity of the trend is m=1, namely monotone trend, and μ_1 = 0.46. The second step is to estimate the desired period's length $\underline{P}$ = $(p_1 \ldots p_R)$ by using the multiplicative model: Thus, $\underline{s}_r^{(p_r)}$ = 0 r=1,...,R. For order R=1 of complexity of the periodicity the value $M_1^{(p_1)}$ for p_1 = 1,...,14 were computed and their graph vs. p_1 is given in Figure (5.2)

Figure (5.2): Graph of $M_1^{(p_1)}$ for $p_1 = 1,\ldots,14$.



The sharp peaks at p_1=5,10 indicate a periodicity of length p=5, but Max $\mu_1^{(5)} = 0.726$ and $M_1^{(5)} = 0.489$ are not close enough to 1 (not adequate to model (5.2)), multiplicative type). It might be that a more complex periodicity model would be more fit to the data. Thus a second period's length p_2 is to be estimated. Now $M_1^{(\underline{P})}$ is computed and graphed in Figure (5.3) where $\underline{P} = (p_1,p_2) = (5,p_2)$ for $p_2 = 1,2,\ldots,14$.

Figure (5.3): Graph of $M_1^{(5,p_2)}$ versus $p_2 = 1,2,\ldots,14$.



The graph of Figure (5.3) indicates a sharp peak at $p_2 = 6,12$. Max $\mu_1^{(5,12)} = 0.95$ and $M_1^{(5,12)} = 0.91$ indicated quite high goodness-of-fit. The estimated complex period's length are thus: $p_1 = 5$ and $p_2 = 12$. The computed complex periodicity patterns are in Table (5.1)

Table (5.1): The two multiplicative components of complex periodicity of order R=2 for the lengths 5 and 12, respectively:

Period length		1	2	3	4	5	6	7	8	9	10	11	12
i=1 p_1=5	$\underline{S}_1^{(5)}$:	93.2	89.9	95.4	94.6	130.7							
i=2 p_2=12	$\underline{s}_2^{(12)}$:	89.8	124.4	134.4	102.1	112.4	89.1	80.5	90.9	93.8	95.4	96.7	91.6

Discussion

The complex seasonally adjusted procedure discussed here generalized the LPTA approach presented in chapter 2. The classical methods could not adjust for complex seasonality simultaneously, a property that the proposed procedure has. Besides this desired property the same advantages over classical methods given in Raveh (1981) or later in chapter 8 still exist. Some of these advantages are: Idempotency, possible choosing type of model, any lengths of the periods $p_1, p_2, \ldots, p_R$, and adjust series with missing observations.

6. PERFECTLY MONOTONE SERIES

In the previous chapters we did not deal with series that are perfectly monotone, namely, their $\mu_1=1$ even if seasonal as well as irregular components are hidden. In this chapter we point out two alternative ways to treat such series. In section (a) usages of series of differences will be discussed to treat series that have convex (concave) trend component. In (B) the idea of rotating the time axis is suggested.

(a) Convex (Concave) Series

Up till now, the shape of the trend has been estimated through monotone concepts. The very same concepts could be related to the series of differences (of the original series) of order D. Let the difference between two consecutive observations Y_{t-1}, Y_t be $\Delta_t^{(1)} = Y_t - Y_{t-1}$. Hence, the difference of order d=0 is the original observation $\Delta^{(0)} = Y_t$.

A series of differences of order 1 is $\Delta_t^{(1)}$ t=2,...,N. A series of differences of order d is $\Delta_t^{(d)}$ t=d+1,...,N where $\Delta_t^{(d)} = \Delta_t^{(d-1)} - \Delta_t^{(d-1)}$. Later on we will consider only series of differences of order d=1 and call them, in short, series of differences.

A series Y_t t=1,...,N is convex (concave) if and only if $\Delta_t^{(1)}$ is positive (negative) monotone. If $\Delta_t^{(1)}$ is constant then Y_t is a linear series. In all the definitions and notations that have been used in earlier chapters $\Delta_t^{(1)}$ could replace y_t and convex (concave) should replace the concept monotone. Hence, convexity and concavity of order m can replace polytonicity of the same order, and so on.

The concept convex relates to a specific form of a monotone shape. In order to assess the amount of convexity in an empirical series let us define coefficient of convexity in (6.1), below.

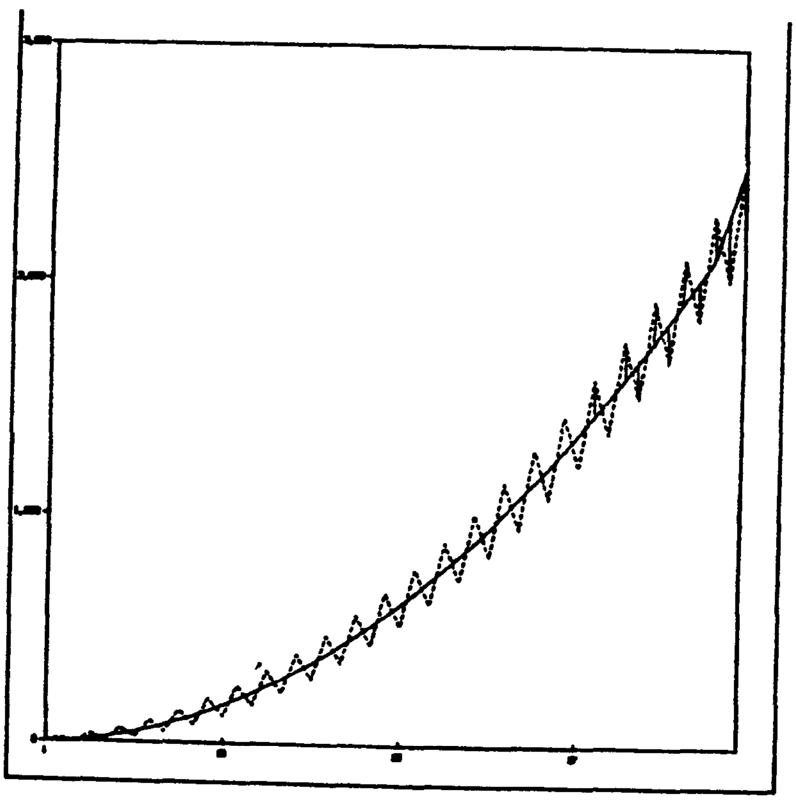
(6.1) $$\mu_{1,1} = \frac{\sum_{i>j} (\Delta_i - \Delta_j)}{\sum_{i>j} |\Delta_i - \Delta_j|}$$

where $\Delta_i = \Delta_i^{(1)}$.

$0 \leq \mu_{1,1} \leq 1$. Thus, if $\mu_1 = \mu_{1,1} = 1.0$ the original series is convex has positive monotone shape. When $\mu_1 = 1$ and $\mu_{1,1} = -1$ the monotone series has a concave shape. $\mu_1 = 1$ and $\mu_{1,1} = 0$ pointed out that the series is approximately a linear one.

As an example, in Figure (6.1) an artificial series is presented. This series has a convex (and monotone) trend and multiplicative seasonality factor. For example, series currency component of M-1A money supply presented in Figure (2.6) has nearly convex trend measured by $\mu_\Delta = 0.84$.

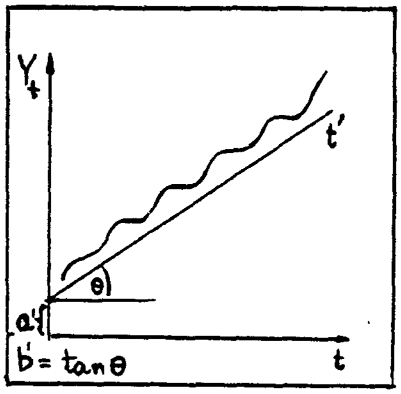
Figure (6.1): An artificial series that has convex trend.



(b) Rotating Time Axis

It is very rare in empirical economic series, yet it may very well be that the trends dominate the other components and the original series is perfectly monotone while seasonality and irregularity still exist. An example is presented in Figure (6.2), below.

Figure (6.2):



For example, the series $Y_t = a + bt + \sin t = b(t + \sin t) + (1-b)\sin t$. The series is purely additive model of linear trend and seasonal components only. For b>1 the series is perfectly monotone. By reduction of artificial linear trend, $t' = a' + b't$ (where $a - a' > 0$ and $1 > b - b' > 0$) from the original series. Y_t' is obtained

$$Y_t' = a + bt - a' + b't + \sin t = (a-a') + (b - b')t + \sin t$$

and Y_t' is not perfect monotone any more for appropriate choice of b', so that $b' < b$ and $(b-b')$ is much smaller than 1 and positive.

7. X-11: SHORT DESCRIPTION AND SOME NOTES

Auerbach and Rutner (1978) presented a case of the U.S. Consumer Price Index (CPI) in which the X-11 Variant program of the Census "seasonally adjusted" out of existence a nonseasonal cycle and apparently mistook this nonseasonal cycle for a seasonal cycle. Similarly, the above property of X-11 is not a quirk peculiar only to the U.S. CPI, but of the procedure itself. To support their conclusion, Auerbach and Rutner demonstrated (an artificial series) a distortion in seasonal adjustment procedure by X=-11 where the resulting series contains a seasonal!

The goal of this note is to support Auerbach and Rutner findings and to shed light on some other properties and assumptions of the Census X-11 program. The properties mentioned here are mainly from the application point of view. In Pierce (1980, p.125) we found that, "The Census Bureau's X-11 seasonal adustment procedure (Shiskin, Young and Musgrave 1967) represented the culmination of a major phase of continuing research in the area of seasonal adjustment." Today, the X-11 program is widely used on economic time-series. Thousands of series are adjusted by it each year and most of them are decomposed by the multiplicative version. Many basic and important properties of X-11 have already been discussed in the literature, see Pierce (1980) and Auerbach and Rutner and their references and Zellner (1978) and Zellner (1982). Studies of the X-11 filters, in the abstract have been done by Wallis (1974, 1981) and Young (1968).

In the next section, a list of some seven properties are given. In the second section, a formula which is based on linear shapes of a series is suggested for prediction purposes. It has been clarified that the built-in-formula used by X-11 for predicting the seasonal factors one year ahead is a special case of the formula we derived. In addition to Auerbach and Rutner (1978) some case studies are demonstrated here as well as abstract properties.

In order to give a common basis for discussion of the X-11, let me cite Plewes (1978, p. 178) on his quick review of the main steps in the program, using as an example the standard multiplicative option. The program: ... "

1. Computes the ratios between the original series and a centered 12 term moving average.
2. Estimates seasonal factors by applying a weighted 5-term moving average to the SI ratios.
3. Adjusts to sum 12.
4. Estimates the irregular component by dividing the factors into the SI ratios.
5. Identifies and removes "extreme" irregulars.
6. Obtains preliminary seasonal factors by applying a weighted 5-term moving average to the SI ratios with extremes replaced.
7. Adjusts to sum 12.
8. Obtains preliminary seasonally adjusted series by dividing these values into the original observations.
9. Obtains estimates of the trend-cycle by applying a 13-term Henderson moving average to the preliminary adjusted series.
10. Estimates new SI ratios, dividing the trend-cycle into the original observations.
11. Estimates seasonal factors by applying a weighted 7-term moving average to the SI ratios.
12. Adjusts to sum 12.
13. Divides seasonal factors into the original series to obtain a seasonal adjusted series.

The X-11 is an iterative procedure. It repeats some of these steps more than once, and in the process, obtains a smoother result. The method gains a good deal of flexibility from the foresighted inclusion of various options which serve to enable the user to more closely approximate the generating mechanisms of the series to be adjusted. Among these that should be singled out for mention are options which:

1. Provide for either multiplicative of additive adjustment.
2. Allow selective (or wholesale) prior adjustments to the input data series.
3. Provide for trading day regressions, to test for the influence of changing days-of-the-week, which are most important in trade series.
4. Enable establishment of various sigma limits for identifying extreme irregular fluctuations.
5. Allow selection of various lengths of terms for the moving averages."

A List of Properties and Assumptions.

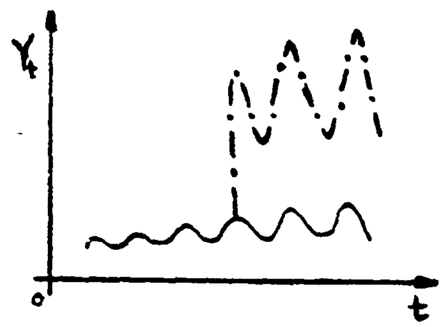
1. Non-Idempotency

A procedure has the desirable property of idempotency when applying it to the seasonally adjusted data (S.A.D.), the estimated seasonal factors are all equal to 100% or 0 for multiplicative and additive models, respectively, Fase et al. (1973) have showed that X-11 as well as other classical methods based on moving average filters that were in use at that time, do not have this property. In contrast, other methods like regression, BAYSEA, see Akaike (1981) or LPTA see Raveh (1981), are idempotent procedures. Non-idempotency means that overestimation or underestimation of the seasonal pattern may very well occur. This is one reason as well as the Slutzky-Yule effect, that significant autocorrelations at seasonal lags (e.g., lag 12) are sometimes revealed. The findings of Auerbach and Rutner (1978) of existence of seasonality in the seasonally adjusted result of X-11 are in part due to the non-Idempotency property of X-11. It might be worthwhile to study the amount of non-idempotency of various procedures. This can be done by applying procedure on S.A.D. and measuring the amount of discrepancy, namely, the amount of deviations of the seasonal factors from 100% or 0 for the appropriate model.

2. Non-Robustness Against an Abrupt Change in the Trend

No moving average method can be expected to produce sensible answers when abrupt change occurs in the trend. An assumption of continuous trend underlies the X-11 program which used moving-average filters. Thus, an abrupt change in the trend yields 'strange' estimation results for the other components. As an example, let us deal with a 'nice' series that has monotone trend (to be more specific, this monotone trend is approximately linear), a fixed seasonal pattern and resaonable irregularity. Let us do the following experiment: Multiply the last half of the series by a constant k>1 (or add a constant k>0). This series and the original one are presented in Figure (7.1), below.

Figure 1: A 'nice' series that has monotone trend, a fixed seasonal pattern and reasonable irregularity.--original series,-.-The last half of the series multiplied by K>1.



Applying the Multiplicative (Additive) model of X-11 yields different estimation results for the seasonal and the irregular components than is obtained for the original series. This means that although seasonality and irregularities have not been changed and only the shape of the trend has been changed, different estimations are obtained! In other words, X-11 adjusts or removes differently the very same seasonal patterns when they are combined with different shapes of trends! The amount of distortion in the estimated seasonality (as well as irregularities) is a monotone function of the abrupt change and/or discontinuity of the trend and the location of the abrupt change as well. An empirical example was given in Raveh (1978). This series is the consumption of electricity in U.S.A. in the years 1951-1958. The original series is given in table A.1 and its graph in Figure A.1. The S.A.D. and trend estimation is presented in Figure

A.1 as well. The last half of the original series was multiplied by K=10. Seasonal factors estimated by X-11 for both series and the arithmetic mean of the seasonal factors, separately, for each month is given in Table (7.1). The LPA method, see Raveh (1981) which is based on nonmetric filters, was applied to these two series. Their fixed seasonal patterns are given in Table (7.1) as well.

Table (7.1): The estimated seasonal (in percentage form) patterns for both the original and the transformed series as obtained by X-11 and the LPA method. (multiplicative model).

METHOD	SERIES	JAN.	FEB.	MAR.	APR.	MAY	JUN.	JUL.	AUG.	SEP.	OCT.	NOV.	DEC.
X-11	Original	119.7	106.5	103.5	93.3	87.0	81.3	83.4	89.8	96.3	106.3	112.8	119.6
	Transform	124.6	113.6	111.8	99.6	91.1	82.0	79.2	82.1	88.4	99.2	109.2	119.2
LPA	Original	119.8	106.8	103.7	93.3	86.6	81.3	83.5	90.1	96.6	105.9	112.6	119.8
	Transform	119.5	107.6	103.3	93.2	87.4	81.9	83.9	90.4	96.5	105.4	112.1	119.0

The seasonal patterns estimated by the LPA method for the series before and after the multiplication by K=10 are very similar. In contrast, the seasonal patterns estimated by X-11 to these two series are very dissimilar although the series have the same seasonal pattern and irregular component by definition.

3. Variant under shifts.

An underlying but not plausible assumption of X-11 is the constraint that seasonal factors for a calendar year add up to 12 (or 1200%) when applying the multiplicative model. The motivation for these constraints is that the annual total for calendar year for the original and the seasonally adjusted data series would be as close as possible. Thus, different shifts of the series backward or forward yield different estimation results for the main components: trend, seasonal and irregular. For example, let a given series cover m whole years from January through December. Then, shift the series six months ahead, or in other words, January is called July, February is called August, etc. Now the series includes (m-1) whole years plus 2 half years. Let us apply X-11 to the shifted series as well as to the original one. It yields different results, namely the original seasonal factors for Januarys are not the same as those for Julys of the shifted series.

The traditional multiplicative model is given in (7.1):

(7.1) $Y_t = T_t \cdot S_t \cdot I_t \quad t=1,\ldots,N$

where T_t, S_t, I_t are the trend, seasonal and irregular of the t^{th} observation, respectively. Y_t denotes the observation at time t. To keep the S.A.D. in the scale of original series, Y_t, for a series that covers only whole years, the following constraints are set by X-11:

(7.2) $$\sum_{K=0}^{11} [S_{i+k}]^{-1} = 12$$

that is, the arithmetic mean of the reciprocals equals 1, where the index i is only for Januarys, e.g. i=1, 13, 25,.... . These constraints are adjusted for incomplete years. Other constraints that might be used are:

(7.3) $$\sum_{K=0}^{11} S_{i+K} = 12$$

that is, their arithmetic mean equals 1; or much more natural constraints are:

(7.4) $$\prod_{K=0}^{11} S_{i+K} = 1$$

that is, their geometric mean equals 1. The natural constraint (7.5) which means that every 12 consecutive seasonal values would add up to the same constant, say 12, implies an assumption of <u>fixed</u> seasonal. This is not the case in X-11.

(7.5) $$\sum_{K=0}^{11} [S_{i+K}]^{-1} = 12 \quad i=1,\ldots,N-12+1$$

It seems that any constraint other than (7.5) is somewhat arbitrary and relates to the vague concept of moving-seasonality which seems a plausible assumption, but is not yet defined clearly in literature. Satisfying the constraints approximately only makes sense if they can be placed in the context of a model. Schlicht (1981), Akaike (1981) and Kitagawa (1981) proposed decomposition methods using additive models so that constraints like (7.6) would be satisfied approximately. Those methods tradeoff between the smoothness of the trend and the varying seasonal patterns with satisfying constraints like (7.6).

$$(7.6) \qquad \sum_{k=0}^{11} s_{i+k} = 0 \qquad i=1,\dots,N-12+1$$

4. Inaccurate estimation of series with zero-value observations.

The multiplicative model of X-11 cannot be used for data which includes zero or negative value observations. For this case the only available option is the additive model. However, applying the additive version of X-11 yields different seasonal factors for the same zero-value observations within the same calendar year! For a series with zero-value observations (not missing data) observations, and a trend other than constant, the additive model is inappropriate by definition. This inappropriateness is caused by the fact that the deviations of the zero-value observations are proportional to the level of the trend, and multiplicative (or mixed) model might be adequate.

The series "Export of citrus in mil. of $" from Israel in the years 1961-68 would be an example. This series has only eight active months each year. In June, July, and August there is no marketing; in September the marketing is almost zero; hence these four months are omitted.

Table (3.E.1) and Figure (3.E.1) present the original data and the graph of this series, respectively. The seasonal factors for each month in each year are given in Table (7.2).

Table (7.2): Seasonal Factors Obtained by the Additive Version of X-11

Year	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
1961	7.58	10.00	8.77	5.14	-2.97	-4.32	-4.81	-4.99	-5.07	-4.69	-3.15	-1.61
1962	7.65	10.14	8.71	5.26	-2.92	-4.35	-4.89	-5.09	-5.19	-4.79	-3.27	-1.56
1963	7.76	10.50	8.70	5.54	-2.88	-4.50	-5.10	-5.30	-5.43	-4.93	-3.45	-1.48
1964	7.99	11.10	8.70	6.04	-2.81	-4.84	-5.45	-5.60	5.74	-5.12	-3.62	-1.36
1965	8.37	11.80	8.73	6.49	-2.73	-5.18	-5.83	-5.95	-6.09	-5.27	-3.77	-1.31
1966	8.80	12.43	8.81	6.93	-2.65	-5.56	-6.19	-6.25	-6.38	-5.39	-3.81	-1.32
1967	9.13	12.84	8.95	7.17	-2.60	-5.80	-6.41	-6.44	-6.57	-5.41	-3.97	-1.42
1968	9.27	13.09	9.07	7.35	-2.59	-5.98	-6.54	-5.63	-6.64	-5.40	-3.72	-1.47
Arithmetic Mean	8.29	11.46	8.77	6.21	-2.74	-5.03	-5.62	-5.74	-5.86	-5.09	-3.54	-1.41

These results are incorrect because all four months have no marketing activity. Within the months of June, July, August, and September there is a gradual decrease of the seasonal values caused by the monotone trend of the series.

Series with particular observations that always have zero values should be modeled differently from other series. The nonmetric approach, LPA in Raveh (1981) can decompose any series that has periods of length $p: 2 \leq P \leq N/4$ by choosing the appropriate model: additive, multiplicative or mixed. Thus, series which are not monthly, quarterly or weekly, e.g., their length of periods is different from the usual 12, 4 or 7, respectively may be decomposed by this procedure. For this example, period's length of eight P=8 and multiplicative model is required.

5. Over sensitivity to outliers.

The sensitivity to outliers of a procedure in estimating the seasonal pattern can be studied from an infinite number of aspects and ways. For more details see Hillmer, Bell and Tiao (1982). The number and amount of the outliers as well as their location can be varied. Here, only one aspect is discussed, that of one outlier located in the middle of an intermediate size of a 'nice' series. Let us deal with a 'nice' series, such as that given in Table A (in Appendix) and plotted in Figure (3.1) and multiply one of its observations, say July, by (or add to) a constant k>1 (or k>0). By using the multiplicative (additive) model on such a series a 'good' procedure is expected to yield slightly higher seasonal values for July's and lower values for the rest of the 11 months in order to agree with the constraints given in (7.2), (7.3) or (7.4). If this is not the case the procedure is defined to be over sensitive to outliers.

As an example the series "U.S. Retail Sales in Millions of Dollars" in the years 1960-1964 has been studied. This is a sub-series of the example given in Shiskin et.al. (1967) and chapter 3, earlier. This sub-series of N=60 observations has nearly fixed seasonality and the monotone trend is nearly linear. The observation in July 1962 was multiplied by K=2.0. As a result, the seasonal values for June, July, August and September were increased, which indicates the undesirable property of over sensitivity to outlier. In Table (7.3) the arithmetic mean of seasonal factors that were computed by X-11 (multiplicative version) are given for the original series and the series with the outlier in July 1962.

Table 3: Arithmetic Mean of Seasonal Factors (Presented in Percent) for the Series "U.S. Retail Sales in Mil. $, in the Years 1960-64" and the Series with July 1962 as an Outlier.

SERIES	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Original	89.5	84.4	97.5	99.0	103.3	103.0	99.0	100.4	97.0	102.7	103.2	120.8
With one Outlier	87.1	82.9	95.3	97.6	101.6	104.6	112.0	102.3	97.4	100.6	99.9	118.3

Two alternative methods suggested by Burman (1965) and Raveh (1981) do not have this drawback of over-sensitivity to outliers, at least for the aspect discussed here. The SABL technique suggested by Cleveland, Dunn and Terpenning (1978) is also insensitive to outliers. The reason is that the smoothers in SABL are based on moving medians as well as moving median-regressions, moving robust-regressions and weighted moving averages.

6. Distinction Between Fixed and Moving Seasonality

Since X-11 has no precise definitions for the various components, there is no distinction between the case of fixed and moving seasonality. For N=36 observations, X-11 automatically produces estimates for fixed seasonality. For other cases, moving seasonality is assumed. There is no clear criterion to decide whether the seasonality is fixed or moving. The F-test is invalid and not usually used.

7. Choosing the Appropriate Type of Models

The lack of definitions and clear critera does not enable the user of X-11 to choose the appropriate model of seasonality: Additive or Multiplicative. Thus, in order to do so, the user needs to use some other techniques such as Durbin & Kenny (1978) or Raveh (1981a). For the case where Mixed (Additive-Multiplicative) model is appropriate, the user is advised to use either Durbin and Murply (1975), or Raveh (1981). In Raveh (1981) the appropriate type of model is chosen based on goodness-of-fit measure and parsimony principle.

8. Absence of Constraints on the Irregular Component

It is desirable that the arithmetic mean of the irregular components should be 1 and 0 for the multiplicative and additive models, respectively. X-11 does not have such above constraints.

Forecasting Seasonal Factors 1 Unit Ahead

Let $S_{i,j}$ denote the estimated seasonal factor for the j^{th} month (j=Jan,..., Dec.) in the i^{th} year. In order to forecast the 1 year ahead factors $\hat{S}_{n+1,j}$ based on the previous n year, X-11 apply eq.(7.7) below.

$$\hat{S}_{n+1,j} = S_{n,j} + 1/2[S_{n,j} - S_{n-1,j}] = 1.5\ S_{n,j} - 0.5\ S_{n-1,j} \qquad (j=1,\ldots,12) \tag{7.7}$$

We will show, in chapter 12, that the formula (7.7) is a special case of our Persistent Structure Principle for prediction with assumption of linear trend.

It is interesting to verify that the classical version of X-11 uses formula (7.7) in order to estimate seasonal factors one year ahead where a=1.5 is chosen as a compromise and not as a function of all the previous seasonal factors. Hence, formula (7.7) is a special case of formula (10.10) with a=1.5. Equation (10.10) is the formula used for prediction purposes when the identified model is an ARIMA(1,1,0) as in eq.(7.8), below:

$$(1 - \theta B)\Delta S_t = a_t \tag{7.8}$$

B is the backshift operator such that $BS_t = S_{t-1}$ and a_t is a white noise process. The parameter θ in (7.8) is equivalent to (a-1) in eq.(10.10). In other words, in order to predict seasonal factors one unit ahead, X-11 uses an ARIMA(1,1,0) model with $\theta = 0.5$ (or a=1.5) as a constant and not as an optimal parameter estimation process. Recently, we found that many economic series from the Bureau of the Census data base have an optimal coefficient very close to 1.5. The optimality is in terms of the above Persistent Structure Principle. The formula (10) might be too dependent on the last two observations. Other conditions for linearity could be used in order to overcome this dependence. Thus, conditions for every four values as in (7.9) could be used instead of (10.6).

$$\frac{S_k - S_l}{k-l} = \frac{S_u - S_v}{u-v} \quad \text{for all } k>l \text{ and } u>v \tag{7.9}$$

The product-moment coefficient of correlation (Pearson's ρ) could be used as a figure of merit as well.

Conclusion

In addition to Auerbach and Rutner (1978) conclusions, we point out here a list of eight properties and assumptions of X-11. These properties are sometimes drawbacks of the Census X-11 program, and users (usually agencies) should be aware of them. Properties 2 and 3 are, as a matter of fact, invisible assumptions of X-11. Some of these properties, like 1,2,3 and 5 are due to the moving-average filters. As Kendall (1973, p.38) pointed out, there is no optimal way to choose the weights as well as the number of elements for the moving-average filters. This makes unclear the trade-off between over and under smoothness of a given empirical series. It was shown that in order to predict Seasonal Factors one unit ahead, X-11 uses a built-in formula based on ARIMA (1,1,0) model with a constant (not function of the data) parameter. It seems that there is not a best seasonal adjustment method which is always the best for any sort of data. The user should be aware of the drawbacks as well as the advantages of the various methods and use the appropriate one according to the data.

8. SOME COMPARISONS BETWEEN LPTA AND X-11

In this chapter a list of some 15 issues are discussed briefly in order to compare the suggested LPTA method with X-11 and Burman's method which is in use in the Bank of England. The discussion as well as some assumptions are summarized in a concise way in Table (8.1). some of the issues were presented in more detail in Chapter 7.

1. Idempotency

LPTA technique has the property of idempotency by definition. This means that applying it to the series from which periodic effects have been removed (i.e., to the seasonally adjusted series) results in the same (adjusted) series again. It is obvious that repeated application of the technique to a series would produce the same result as a single application. Fase et al. (1973) have shown that the two methods, X-11 and Burman (1975) as well as other classical methods based on moving average filters in use at that time do not have the idempotency property.

2. Robustness Against Abrupt Change in the Trend

LPTA technique has the property of robustness against abrupt change in the trend. Some examples are given in chapter 3(D) and the Census series "U.S. Retail Sales in Variety of Stores" exhibits in figure 8. X-11 is lacking this property, see chapter 3(D) and chapter 7.2.

3. Variant Under Shifts

X-11 is variant under shifts. In other words, shifting the series several units (months) backward or forward yields different estimations from the main components. More details are found in chapter 7.3. The LPTA is by definition invariant under shift.

4. Estimation of Series with Zero-Value Observations

The multiplicative model of X-11 cannot be used for data which includes zero or negative value observations. Incorrect estimation is obtained using the additive version, see chapter 7.4. LPTA method deals with zero-value observations the same way as with non-zero observations, see chapter 3(e).

5. Sensitivity to Outliers

A very limited study of sensitivity has been done in this report. In 7.5 it is indicted that X-11 is over-sensitive to outliers, compared to LPTA.

6. Distinction Between Fixed and Moving Seasonality

Since X-11 has no precise definitions for the various components there is no distinction between the case of fixed and moving seasonality. LPTA is based on prespecified definitions. The coefficients of goodness-of-fit enables us to decide the appropriate version between fixed and moving seasonality. See chapter 4.

7. Possible Types of Model

X-11 and Burman methods can use only multiplicative or additive models. The nonmetric approach can use the mixed model too.

8. Choosing the Appropriate Type of Models

No clear criterion is supplied by X-11 to choose between multiplicative and additive models. Some suggestions with examples for choosing the appropriate model are given in chapter 3(b).

9. Constraints on the Irregular Component

X-11 in contrast to LPTA has no constraint on the irregularities. Thus their arithmetic mean may be different from 100% or 0 to multiplicative and additive models, respectively.

10. Minimum Number of Whole Periods Required

Two (possibly less) whole periods are required for either the multiplicative or the additive model, see chapter 3(c). Four periods are required for the mixed model. X-11 and Burman methods require 3 and 5 whole periods, respectively. The B.L.S. (1965) require a minimum of 8 whole periods.

11. Length of the Period - p

Any period length p, $a \leq p \leq [\frac{N}{4}]$ can be used by LPTA. There is a possibility of estimating the period's length if not known in advance, like the periodogram, see 2.(a). The two classical methods X-11 and Burman can analyze only series of periods length p=12. There is a version of X-11, called X-11-Q which is constructed for p=4 as the length of the period. There is no option to choose any desired period's length.

12. Data With Missing Observations

By giving zero weights within the coefficient (2.5), the LPTA approach accepts series with missing observations and there is also a possibility to censor outliers, see chapter 2(c). X-11 and Burman methods lack this property and for missing data the user should substitute appropriate value.

13. Complex Seasonality

LPTA can very easily handle series that have complex seasonality, namely, some seasonal pattern of different period's length are involved simultaneously. X-11 as well as other moving-average methods lack this property. More details are given in chapter 5.

14. Perfect Monotone Series

Series that have dominant trend such that their $\mu = 1$, namely, they are perfect monotone series, can not be directly handled by LPTA. Some transformation of the axis as in chapter 6(b), or using series of differences are needed as in chapter 6(a). X-11 can analyze such series directly.

15. Taking Turning Points into Consideration

While using LPTA, turning points should be checked in advance and the procedure takes into consideration the coefficient of goodness-of-fit. The definition of piece-wise monotone as well as polytonicity is based on this valuable information, see for example the series in Figure (2.4). X-11 as well as Baysea procedure will smooth turning points and their over and under estimation will be obtained near the turning points. On the other hand, X-11 and Baysea do not need any assumption about turning points in advance.

In Table (8.1) these issues and some other distinction between X-11 and LPTA are given.

Table (8.1): Some Comparisons Between the Nonmetric Approach, Least Polytone Trend Analysis (LPTA), and the Known Techniques X-11 and Burman, Respectively.

Issue	LPTA approach	X-11 and Burman, respectively
Types of filters	Nonmetric	Linear (for additive model X-11)
Idempotency	Yes	No (Fase et al. (1973)).
Minimum number of periods required	2 (possibly less) for similar model, either pure multiplicative or pure additive, see chapter 3(c). 4 periods are needed for the mixed model.	3 and 5 respectively.
Possible types of model	Multiplicative or Additive or Mixed	Multiplicative or Additive only
Period's length	Any period length p $2 \leq p\ [\frac{N}{2}]$. The "best period length may be found from the graph of $M_m^{(p)}$. In chapter 3(e) the example was analyzed with p=8.	12 only, must be assumed in advance. There is a version of X-11, called X-11Q which is constructed for p=4 as the length of the period. Anyhow, there is no option to choose desired period's length.
Data with missing observations	Could be handled, see chapter 3(c).	Cannot be handled.
Robustness against an abrupt change in the trend.	Yes, see chapter 3(D).	No, see chapter 3(D). Continuity of the trend is invisible assumption of X-11 as well as other moving average methods.

Table (8.1) continued)

Issue	LPTA approach	X-11 and Burman,
Choice of appropriate model of seasonality	Choice among models can be made on the basis of the best fit, i.e., maximum value for $M_m^{(p)}$. See chapter 3(b).	There is no clear criterion for choosing between the multiplicative and the additive model.
Possibility of estimating the period's length if not known in advance	Yes, see chapter 3(a)	No.
Observations sign	Observations may assume any value in all three models.	In multiplicative model only positive observations can be considered.
Use of data	Use is made of all observations in an equal manner.	By moving-averages technique filters only partial use of information contained by observations near the ends.
Prior assumptions' and arbitrary specifications	Polytonicity of the trend. Initial values of the coefficients ($\underline{S_p} = \underline{1}$; $\underline{s_p = 0}$) in the numerical algorithm. (However, extensive experience shows that results are not sensitive to these).	Period length. Span and weights in the computation of the moving-averages (Kendall (1973), p.38). There are arbitrary limits used to eliminate extreme values.
Shift series backward or forward	invariant - by definition	Variant - see chapter 7.3
Constraint on the irregular component	exist	Does not exist
Defintion of Seasonality	Yes, chapter 5.	No.
Analyzing Perfect Monotone Series	Not possible directly, see chapter 6(a,b). Turning points should be checked in advance, then taken into consideration.	No need to assume turning points in advance and thus do not take them into consideration.

9. DEMOGRAPHIC EXAMPLE: MEASUREMENT AND CORRECTION OF THE TENDENCY TO ROUND-OFF AGE RETURNS.

This chapter presents a usage of the LPTA Procedure for a demographic example. The series are returned age of men and women in various countries obtained in censuses. The purpose of our approach is to measure and correct the tendency to round-off age returns when the 'true' age distribution is unknown.

Our technique has been applied to selected populations and compared to two alternative classical methods, those of Myers and Bachi. The indices of preference or dislike for each of the ten digit units obtained by the proposed technique and the other two classical methods are very similar in results, despite the fact that they involve different methodological strategies.

In addition, our proposed method enables the estimation of the "true" number of persons of age t, T_t, that is, the trend component. Thus, indices for the preference or dislike of each age can be computed, an attractive property that the classical methods do not address.

It is well known that age measurements are often affected by the tendency to rounding: measures with unit digit 0 and 5 and, to a certain extent, those with unit digits 2 and 8 appear with a high frequency, while the number of those ending with other unit digits is understandable. Two main problems arise where data are subject to the tendency to age rounding: how to measure this tendency and how to correct for it. Several decades ago this problem was treated by several researchers. Myers (1940) and Bachi (1951, 1953) developed techniques for measuring such net misstatements only for all age groups ending in the same unit digit i, $i = 0,1,\ldots,9$ when "true" age is unknown. These two authors suggested a general index for age-accuracy, based on the above measures. Another index was offered earlier by Marten (1924) called Whipple's Index and by U.S. (1951). Its main drawback, apart from measuring digit-preference only, is that it measures the preferences for only two digits, 0 and 5. Carrier & Farrag (1959) attacked the problem by adapting an optimal polynomial for the case in which only five year age groups are given.

The purpose of this chapter is to deal with the problem of measuring and correcting for the tendency to round-off age returns at single age t=0,1,2..., 120. In order to solve this problem, the LPTA technique is applied. The method estimates the "true" age distribution by means of trend estimation and provides indices of preference for each single age or age group.

We shall consider the set Y_t, the number of persons enumerated at age t as time-series with a periodicity of length 10 (the ten digits). The trend component T_t is an estimate of the "true" number of persons at age t. It is further assumed that T_t decreases as t increases (e.g., a monotone decreasing) which is the case for many populations. While the other two methods cannot be applied from age 0 (in the examples cited above--Myers (1940) and Bachi (1951, 1953)--which are restricted to the age range 23-72), no such restriction is needed in the method proposed here.

As we pointed out earlier, our data are the number of persons, Y_t, enumerated as of age t. A series Y_t is monotone decreasing if $Y_i < Y_j$ for every i>j. To express Y_t in periodic terms, it will be useful to replace the observation index t by an index of the form 10a+i, where 10 is the proposed period length (of the ten various digits), i is the position (digit) in the sequence of periods, with the first indices 0, the second 1, etc. We denote the number of complete periods by n, so that a=0, 1,...,n-1. Given this notation, a sequence Y_t t=0,...,N can be written as Y_{10a+i} (i=0,...,9; a=0,...,n-1).

By a series of periodic transformations of Y_t (with period length 10) we shall generate a series Z_t whose members are of the form

(9.1) $Z_{10+a+i} = Y_{10a+i}/S_i$ (i=0,1,...,9; a=0,1,...,n-1)

where the 10 transformation coefficients S_i represent multiplicative periodic factors. The series Z_t can be regarded as the preference digits adjusted data in analogous to the known seasonally adjusted data.

It is convenient to keep Z_t in the scale of Y_t by setting the constraints

(9.2) $$\sum_{i=0}^{9} [s_i]^{-1} = 10$$

That is, their arithmetic mean of the reciprocals equals 1. Let us denote by $\mu(Y)$ the coefficient of Monotonicity of a given series $Y_1,\ldots,Y_N$.

(9.3) $$\mu(Y) = \frac{\sum_{i>j} (y_i - y_j) \cdot w_{ij}}{\sum_{i>j} |y_i - y_j| \cdot w_{ij}} \quad 1< <j <i <N \quad \text{where } w_{ij} > 0$$

For simplicity, $\mu(Y)$ will be denoted by μ only. The coefficient of monotonicity when computed for the transformed series Z_t, $t=0,\ldots,N$ will be designated by $\mu(Z)$, where, as noted above, $t=10a+i$.

THE LPTA TECHNIQUE FOR ASSESSING PREFERENCE (DISLIKE) AT EACH UNIT DIGITS EFFECTS (when the "true" age distribution is unknown).

We first assess the monotonicity of the original series Y_t (of the number of persons returned as of age t) by computing μ. The preference and dislike for the ten digits are the periodic fluctuations. In such cases estimation of the ten $S_i, i=0,1,\ldots,9$ of eq. (9.1) are required. This is done by minimizing $\mu(Z)$ toward -1.

The closer Min $\mu(Z)$ is to -1, the closer is the series Y_t to being periodically-monotone. The ten coefficients S_i describe the pattern of variations within periods and will be considered as representing "seasonal" effects -- preference/dislike -- for each digit. We have used as an initial guess the values $S_i=1$, $i=0,1,\ldots,9$. That is, we start with the assumption of no preference/dislike effects.

For the usual case were $\mu > -1$, the measure M of goodness-of-fit is:

(9.4) $$M = \frac{\text{Min } \mu(z) - \mu}{-1 - \mu}$$

THE LPTA TECHNIQUE FOR ASSESSING PREFERENCE (DISLIKE) AT EACH AGE:

The LPTA technique, in contrast to those of Myers and Bachi, enables the assessment of ratios of net misstatement at each age. To do so we have estimated the trend T_t, of the "true" number of persons of age t. This series has to be "smoother", that is, it is more negative monotone than the series z_t which takes into account only fixed effects for each digit unit (an effect which is proportional to the trend).

Our starting point is the series Z_t, which is preference (dislike) adjusted data. We shall search for a perfect monotone series T_t which is as close as possible to Z_t. Thus the loss function to be minimized multiplicative model is

$\sum_{t=1}^{N} |T_t/Z_t - 1| = \sum_{t=1}^{N} (R_i - 1)$ subject to the constraint $T_1 < T_2 < \ldots < T_N$.

The R_i are N residuals fit to each age. An additional constraint is that the arithmetic mean of the reciprocals of R_i equals 1 is adopted.

EXAMPLE

Let us consider some results that are obtained by the LPTA and by that of Myers and Bachi. For purposes of demonstration, consider the series of "Males in Madras (1911) in the 23-72 age range." The original series is given in Table (9.1) and graphed in Figure (9.1). The preference adjusted data as the trend component are presented in Figure (9.1) as well.

Coefficient of monotonicity of the original series Y_t is $\mu = -0.546$, and Min $\mu(Z) = -0.982$ and coefficient of goodness-of-fit is very high, $M = 0.959$. The ten values of S_i obtained at the minimum of $\mu(Z)$ are given in table (9.2).

Thus the series Z_i which is preference (dislike)-free is approximately monotone. The measure of preference or dislike of each digit is defined such that its mean is equal to zero (see for example, Bachi (1953)), and thus is achieved by subtracting 1 from each coefficient. In Table (9.3) the ten measures

Table (9.1): Returned Age of Men in Madras, India 1911. (Age Range 20-79).

Ten Digits	Unit Digits 0	1	2	3	4	5	6	7	8	9
20	3,912	586	1,735	713	1,228	4,118	1,272	675	1,394	468
30	4,964	408	1,176	314	592	3,197	990	446	856	401
40	4,525	395	749	362	395	2,382	549	275	588	302
50	3,302	302	417	235	282	1,186	411	153	306	119
60	2,276	148	237	127	160	518	127	87	129	78
70	496	23	64	20	61	146	72	17	67	120

(indices) of each digit are given for the four populations (men), taken from Bachi (1953, p.6) under the heading R_1. The net percentages of each population returning ages with inaccurate unit digits may be defined as the sum of the positive (or negative) preference indices.

Let us now use our procedure for assessing preference (or dislike) at each age 5, that is, the estimate of the series T_i of the "true" age. This trend component is computed for the same age range 23-71 for men at Madras, 1911.

Now, having on the one hand Y_i and on the other hand T_t, we can compute an index for preference of each age by:

(i) = $(Y_t - T_t)/T_t$ and an index for preference of each digit i, i=0,...,9,

by:

(ii) = $\sum_a Y_{10a+i} - \sum_a T_{10a+i} / \sum_a T_{10a+i}$

Let us compute the preference (dislike) of each age by ratio (i) and, in addition, other indices for the preference of each digit by ratio (ii). These indices are presented in Table (9.3) under the heading R_2. In Table (9.4), these values are presented for the previous example;--that is, Men in Madras 1911 at age range 23-72.

Table (9.2): The Ten Coefficients S_i, i=0,...,9 for the Ten Digits

Digit	0	1	2	3	4	5	6	7	8	9	Average
Value	3.86	.33	.71	.29	.45	2.25	.69	.35	.72	.33	1.0

Some patterns which were unknown just by looking at the data in Table (9.3) now emerge in Table (9.5). Among the ages ending with digit 0 the most preferred age is that of 60 with index of +4.653 which is also the most preferred of all ages. The age with the least index at this range of 23-72 is 30 with index of only 2.237. Among the ages terminating in 5 the index is above 1, except 25 with index of only + 0.946. The most preferred age is 45 (index = 1.784) among the ages terminating with 5. The most disliked age of this population is the age 33 with index of -0.765. Demographers can generate other substantive conclusions by analyzing the results presented in Table (9.4).

Let us now apply our technique to other data that are based on five years age groups: 0-4,5-9,... For purposes of comparisons the Madras example was analyzed at about the same range as the actual series in Table (9.5).

Figure 9.1: Returned Age of Men in Madras, India 1911. ····· Original Data, -·-·-· Preference (dislike) adjusted data, ____ "true" age assumed to be the trend component.

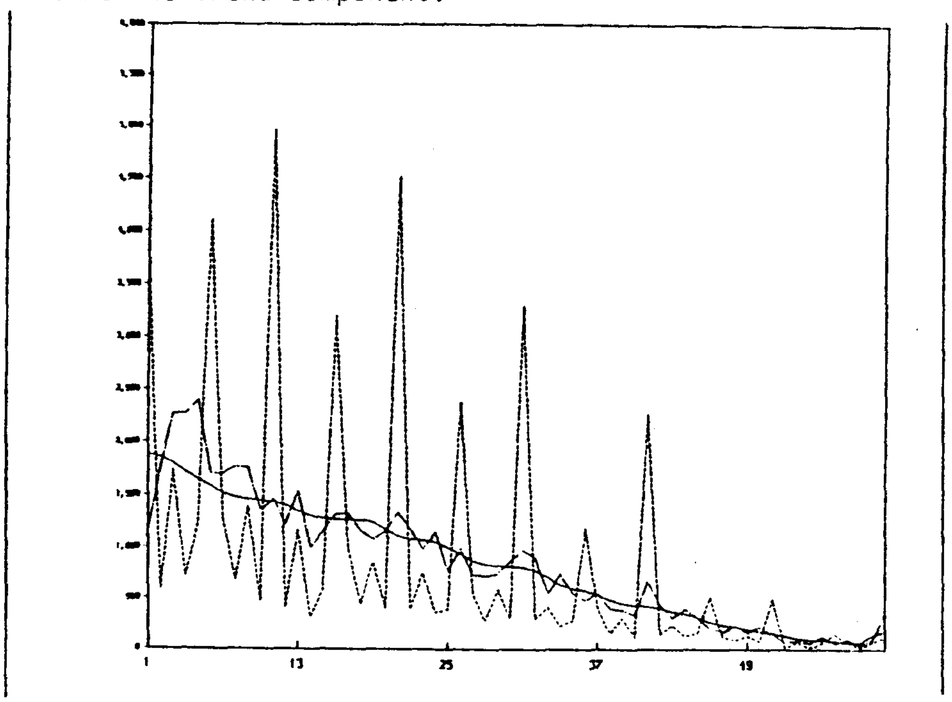


Table (9.3): A) Measurement of the Preference (+) or Dislike (-) for each unit digit in age returns and

B) An Index for measuring the inaccurate unit digit in each population.

These indices are computed for four populations by the three methods of Bachi (B), Myers (M), and LPTA in two ways (R_1,R_2). An explanation for R_2 is given later.

A) Indices of Each Unit digit	Madras, 1911				Egypt, 1927				Spain, 1930				Australia, 1933			
	B	M	R_1	R_2	B	M	R_1	R_2	B	M	R_1	R_2	B	M	R_1	R_2
0	2.80	2.72	2.86	2.90	3.29	3.14	3.76	3.35	.43	.42	.44	.43	.04	.04	.06	.05
1	-.67	-.65	-.67	-.67	-.88	-.85	-.85	-.87	-.22	-.21	-.20	-.20	-.12	-.11	-.11	-.11
2	-.30	-.26	-.29	-.26	-.58	-.51	-.48	-.50	.01	.01	.04	.03	.07	.06	.08	.08
3	-.69	-.69	-.71	-.71	-.71	-.73	-.81	-.83	-.08	-.08	-.08	-.08	.05	.06	.04	.04
4	-.53	-.52	-.55	-.55	-.79	-.78	-.83	-.86	-.04	-.03	-.04	-.04	-.02	-.02	-.02	-.03
5	1.27	1.30	1.25	1.24	2.40	2.38	2.06	1.66	.06	.06	.04	.03	.00	.01	.00	.00
6	-.29	-.30	-.31	-.31	-.74	-.73	-.78	-.75	.01	.01	-.01	-.01	-.03	-.03	-.04	-.02
7	-.65	-.65	-.65	-.65	-.67	-.64	-.71	-.64	-.01	-.09	-.10	-.10	-.04	-.04	-.05	-.05
8	-.27	-.26	-.28	-.26	-.41	-.42	-.49	-.35	.05	.05	.04	.06	.03	.03	.02	.02
9	-.69	-.68	-.67	-.67	-.86	-.86	-.85	-.85	-.12	-.12	-.12	-.13	-.00	-.00	.02	.01
B) Inaccurate rate unit digit	40.8	40.2	41.4	41.4	56.6	55.2	58.1	50.0	5.6	5.4	5.6	5.5	2.0	2.0	2.3	2.0

Table (9.4): Ratio (i) and (ii) for measuring the preference (+) or the dislike (-)for each age and unit digit, respectively, and an index for inaccurate unit digit (which is the sum of positive preference indices in percentages) in age return for the Madras Example.

Tens digits Ratios (i)	0	1	2	3	4	5	6	7	8	9
0										
10										
20				-.734	-.535	.946	-.356	-.642	-.198	-.702
30	2.237	-.728	-.127	-.765	-.550	1.439	-.245	-.645	-.272	-.653
40	3.038	-.645	-.320	-.656	-.608	1.784	-.360	-.676	-.310	-.631
50	3.099	-.624	-.378	-.653	-.536	1.040	-.228	-.675	-.317	-.713
60	4.653	-.635	-.399	-.658	-.519	1.221	-.332	-.616	-.292	-.645
70	2.866	-.676	-.259							
80										
90										
Ratios (ii)	2.900	-.672	-.258	-.714	-.550	1.237	-.312	-.651	-.256	-.673

Index of inaccurate unit digit = 41.37 %

Table (9.5): Group ages: (five years each) 20-24, 25-29,...,70-74 for Men in Madras, 1911

Age Group	20-24	25-59	34-34	35-39	40-44	45-49	50-54	55-59	60-64	65-60	70-74
Actual series	8174	7927	7454	5890	6426	4096	4538	2175	2948	939	664
"True" series	9044	7263	7241	6239	6212	4640	4046	2534	2464	989	640
Index of Preference(+) or dislike(-)	-.10	.09	.03	-.06	.03	-.12	.12	-.14	.20	-.05	.04

In the first step we use formula (9.1) with only two coefficients S_i, i=1,2 for the two different groups, that is,

1. The group of digits ending with 0,1,2,3 or 4 and
2. The group of digits ending with 5,6,7,8 or 9.

We obtain the follwing results: μ = -0.981, Min $\mu(Z)$ = -0.988 and coefficient of goodness-fit M = 0.377. The two measures of preference (+) or dislike (-) for the two groups terminating in 0-4 or 5-9 are: 0.05, -0.05, respectively.

The estimated "true" series T_i for the appropriate groups and the indices of preference (or dislike) for each group is given in Table (9.5)

CONCLUSIONS

The use of the proposed nonmetric technique (using concepts that were taken from the fields of Time Series Analysis) enables the development of a simple and understandable technique for both measuring the tendency to round-off each age, and estimating the "true" distribution of ages. Some substantive information emerges fom the application of this technique, and others may be discovered by the application of the proposed method to other empirical data. The proposed method requires weaker assumptions than the classical methods. It has the advantage of not being limited to a specific age range, thus allowing for the estimation of a very short series (even of 2 whole periods, e.g., 20 observations) of series with missing data (by giving a zero weight). A computer program has been developed and is available to anyone who is interested.

Part 2

FORECASTING QUANTITATIVE SERIES

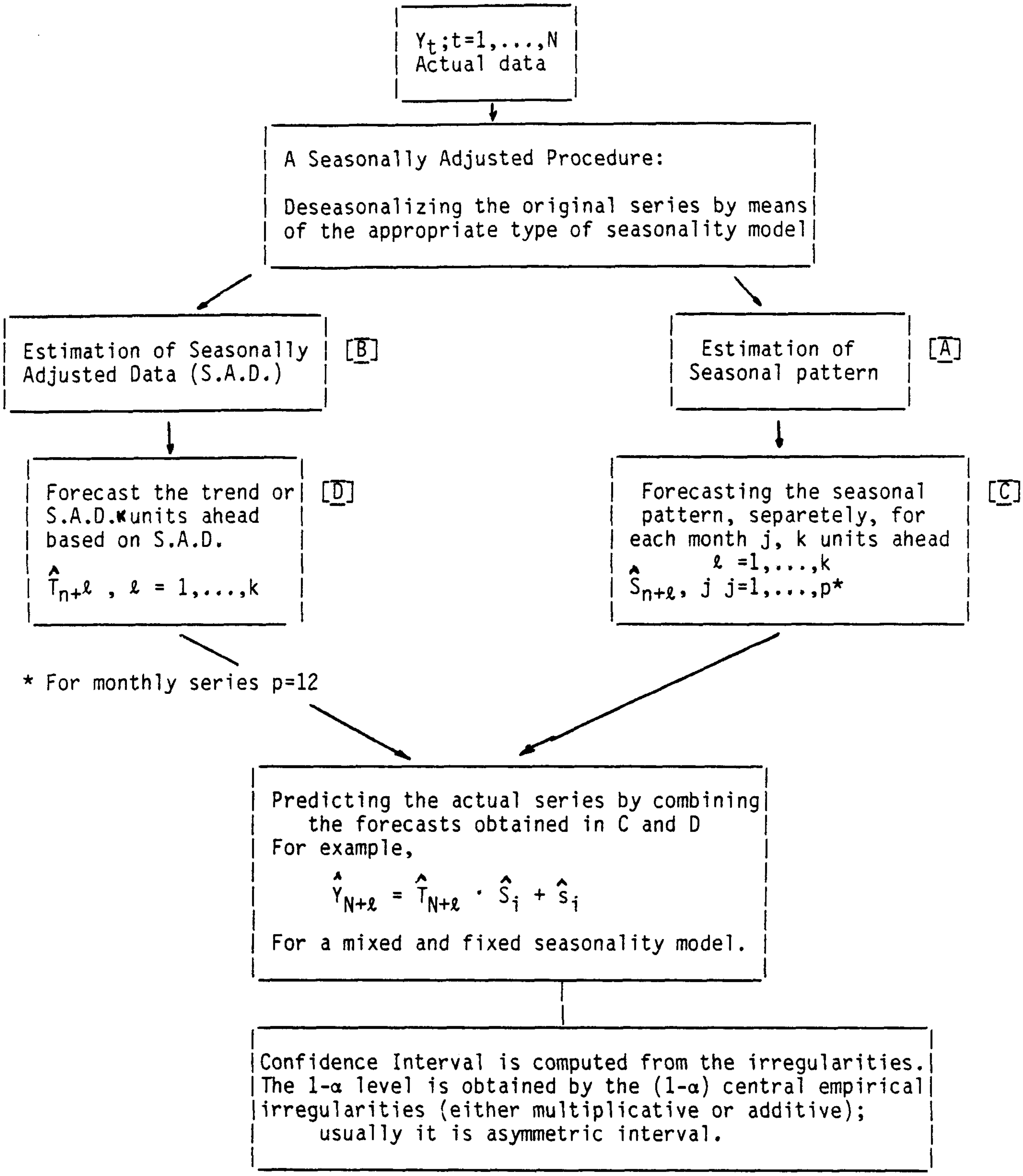
In the first part of this report we discuss approaches, mainly the LPTA, in order to reveal the structure of a given empirical series by means of decompositions methods. In this part we present two different approaches for prediction of quantitative series which are based, in part, on the nonmetric methods discussed in the previous part. Likewise, both approaches are based on the (known) idea of combining seasonal adjustment procedure with a method for prediction of the trend component. In Figure 10 a flowchart of the approaches is presented.

In chapter 10 the Persistent Structure Principle is presented. This principle means that forecast values are estimated in such a way that the values of appropriate coefficients of goodness-of-fit are equal for both the augmented series and the original series. The basic assumption is that the 'structure' of the series remains the same in the forecasting domain as in the past and thus the same goodness-of-fit is obtained. Missing data are treated in a very similar way. Our point of view is that future observations are in some way missing data out of the range of the series.

In chapter 11 The Box-Jenkins approach is combined with our LPTA method for prediction purposes. For both chapters 10 and 11 some well-known examples from the literature and some other examples are demonstrated.

In chapter 12 the Persistent Structure Principle (P.S.P.) is combined with X-11 for forecasting seasonal factors one year ahead. These forecasted factors used for seasonally adjusting current data as is done in the Bureau of the Census Washington D.C., and other agencies as well.

Figure 10: A flowchart for combining a seasonal adjustment procedure with a method for forecast the trend component.



10: PERSISTENT STRUCTURE PRINCIPLE (P.S.P.) FOR PREDICTION OF QUANTITATIVE SERIES

"The thing that hath been, it is that which shall be; and that which is done is that which shall be done, and there is no new thing under the sun."

(Ecclesiastes, 1:9)

The Persistent Structure Principle is suggested for the purposes of forecasting quantitative time series and exemplified by means of economic series. As we already mentioned, the basic assumption is that the 'structure' of the series remains the same in the forecasting domain as in the past, and thus the same goodness-of-fit is obtained.

1: Persistent Structure Principle

Denote by $C(\underline{Y})$ a coefficient of goodness-of-fit for a given series Y_t; t=1, ...,N. This coefficient is supposed to reflect the structure of the series with relation to some specific definitions for trend and seasonal pattern. Let $C(\underline{Y}, \hat{Y}_{N+1})$ be the same coefficient for the augmented series $Y_1,\dots Y_N,\hat{Y}_{N+1}$ where $\hat{Y}_{N+1}$ is the estimated forecast value one unit ahead. If we believe or assume that the 'structure' of the given series in the past is consistent, namely that it remains the same in the very near future, say 1 unit ahead, then their coefficients for goodness-of-fit would be the same, namely

(10.a.1) $$C(\underline{Y}) = C(\underline{Y},\hat{Y}_{N+1})$$

$\hat{Y}_{N+1}$ is the only unknown in equation (10.1). Hence, by solving the equation, the estimated forecast value $\hat{Y}_{N+1}$ is obtained. The reseacher has to choose the coefficient with relation to its prior loss function.

Forecasting the rth unit ahead could be obtained by using the same principle based on data and (r-1) values already have been obtained. Hence, in a recursive way, prediction is achieved for the short-term. Some examples will be given later on.

For a series with missing data, say the i observation, let $\underline{C}_i(\underline{Y}_{-i})=C(Y_1,\dots,$

Y_{i-1}, $Y_{i+1},\ldots,Y_n$) be the coefficient of goodness-of-fit. The Persistent Structure Principle (P.S.P.) suggests that the estimated value Y_i could be obtained by solving equation (10.1) below:

$$(10.1) \qquad C(\underline{Y}_{-i}) = C(Y_i,\ldots,Y_{i-1},Y_{i+1},\ldots Y_n) = C(Y_1\ldots,\hat{Y}_i\ldots,Y_n) = C(\underline{Y}) \ .$$

Let's assume that Y_t $t=1,\ldots,N$ is a time series which can be decomposed into its main three components. Hence $Y_t = f(T_t,S_t,I_t)$ where T_t is the trend, S_t the seasonal component, and I_t is the irregular component. The first two components are systematic, namely, they could be predicted in principle. The latter part is the unsystematic part, called irregularity. Three main models are usually used for decomposition purposes: Purely Additive, Purely Multiplicative, and Mixed model. These models are given below in eq. (10.2), (10.3), and (10.3), respectively,

$$(10.2) \qquad Y_t = T_t + I_t + s_t$$

$$(10.3) \qquad Y_t = T_t \cdot I_t \cdot S_t$$

$$(10.4) \qquad Y_t = T_t \cdot I_t \cdot S_t + s_t$$
$$= Z_t \cdot S_t + s_t$$

where Z_t is the seasonally adjusted data. For a periodic series whose period's length equals p (e.g., p=12 for monthly time series), convert the index t into i+pa where i is the position of the $t^{\underline{th}}$ observation within the period. p is the period's length and a is the period's index. Thus, for monthly series a=0 for the first year, a=1 for the second year, etc. For the sake of simplicity, let Y_t $t=1,\ldots,N$ be a monthly time series, thus $N=12L + K$, $K=0$ where the series has precisely L whole years.

By using the multiplicative model (10.3) for a given series $Y_1\ldots Y_N$ the forecast value 1 unit ahead would be:

$$(10.5) \qquad \hat{Y}_{12L+K+1} = \hat{T}_{12L+K+1} \cdot \hat{S}_{12L+K+1}$$

with the assumption that $E\ I_t = 1$ (multiplicative model). If we believe that the series is consistent in the sense that the same model (multiplicative in the discussed case) is appropriate, $\hat{T}_{12L+K+1}$ and $\hat{S}_{12L+K+1}$ are required.

Here, the persistent structure principle is used to predict the trend component. Presumably since the seasonality is fixed, it is combined with the predicted trend in order to obtain prediction for the series. A prior assumption is needed for the 'structure' of trend and seasonality. Even for the linear case, an infinite number of coefficients of goodness-of-fit could be defined. In the next section the simplest family of coefficients based on the linear assumption will be discussed briefly. In the third section, a monotone shape for a trend of the series will be presented. Three economic examples will be given in the fourth section.

2. The Simplest Linear Case

Let $\underline{Z} = Z_1 \ldots Z_n$ be a series. $\underline{Z}$ is linear series if and only if

(10.6) $$Z_i - Z_{i-1} = Z_{i-1} - Z_{i-2} \text{ for all } i=3,\ldots,N.$$

or $\Delta Z_i = \Delta Z_{i-1}$ for all $i=3,\ldots,N$ where $\Delta Z_i = Z_i - Z_{i-1}$

(10.7) $$\text{or } \Delta^2 Z_i = 0 \text{ for all } i=3,\ldots,N. \text{ or } \sum_{i=3}^{N} \Delta^2\ {}^2Z_i=0 \text{ or } \sum_{i=3}^{N}|\Delta^2 Z_i|^v = 0 \quad v=1,2$$

or any even v. In other words, $\underline{z}$ is perfectly linear series when its slope is constant over time. The series $\underline{Z}$ is the most dissimilar to a linear curve when $\Delta Z_i = -\Delta Z_{i-1}$ or $\Delta^2 Z_i = 2\Delta Z_i$. In other words, $\underline{z}$ is the most dissimilar to a linear one when its slope changes its sign every two consecutive observations. The quantity (10.8) could be used as a basis for a coefficient of goodness-of-fit for linearity.

(10.8) $$K = \frac{\sum_{i=3}^{N} |(\Delta^2 Z_i)|^v}{\sum_{i=2}^{N} |2\Delta Z_i|^v}$$

Let us deal with v=2

(10.9) $$\mathrm{LIN}(Z) = C(\underline{Z}) = 1 - K = 1 - \frac{\sum_{i=3}^{N} (\Delta^2 Z_i)^2}{4 \cdot \sum_{i=2}^{N} (\Delta Z_i)^2}$$

This coefficient of linearity varies between 0 and 1.

$\mathrm{LIN}(\underline{Z}) = C(\underline{Z}) = 1$ $(k=0)$ if and only if $\underline{Z}$ is a perfectly linear series. $\mathrm{LIN}(\underline{Z}) = C(\underline{Z}) = 0$ $(k=1)$ if and only if $\underline{Z}$ is of [a,b,a,b,...,b] type series. Thus the series [a,b,a,b,...,b] where $a \neq b$ is the most dissimilar to a linear series in our definition. $\mathrm{LIN}(\underline{Z}) \doteq 1$ $(K \doteq 0)$ if the series $\underline{Z}$ is locally linear, namely, there are few turning points and in between the series is linear. Since both of these extremes are cases where C=0 or C=1, the series can be predicted without any error.

By using the coefficient (10.9) and the persistent structure principle, $\hat{Z}_{N+1}$ can be computed. Let's equate $C(\underline{Z}) = C(\underline{Z}, \hat{Z}_{N+1})$. By a simple manipulation the required $\hat{Z}_{N+1}$ is obtained:

(10.10) $$\hat{Z}_{N+1} = a \cdot Z_N + (1 - a)\, Z_{N-1}$$

where a can be either $a_{(1)} = (2\sqrt{K}-2)/(2\sqrt{K}-1)$ or $a_{(2)} = (2\sqrt{K}+2)/(2\sqrt{K}+1)$ which are functions of the data values $Z_1,\ldots,Z_N$. For the perfect extreme cases the following results are obtained:

when $C(\underline{Z}) = 1$ then a=2 and $\hat{Z}_{N+1} = 2Z_N - Z_{N-1} = Z_N + (Z_N - Z_{N-1})$,

when $C(\underline{Z}) = 0$ then a=0 and $\hat{Z}_{N+1} = Z_{N-1}$

It is interesting to verify that the classical version of X-11 uses the formula (10.10) in order to estimate seasonal facors 1 year ahead, and chooses a=1.5 as a compromise and not as a function of the previous seasonal factors. The formula (10.10) might be very sensitive to variations of the last two observations. Other conditions for linearity could be used in order to overcome this sensitivity. Hence, conditions for every four values as in (10.11) could be used.

$$(10.11) \qquad \frac{Z_i - Z_j}{(i-j)} = \frac{Z_k - Z_l}{(k-l)} \quad \text{for all } i>j \text{ and } k>l$$

or (10.11) for pairs of observations (Z_i, Y_i) $i=1,\ldots,N$.

$$(10.12) \qquad \frac{Z_i - Z_j}{Y_i - Y_j} = \text{constant} \quad \text{for all } i\leq j<i\leq N.$$

The product-moment coefficent of correlation could be used as a coefficient of goodness-of-fit as well. Thus, for example, we can choose the covariance or a proportional measure to covariance as a coefficient $C(X_1,\ldots,X_n) = \text{cov}(x,i) = \frac{1}{n} \Sigma (X_i-X_j)(i-j)$. In other words, covariance of series $X_1,\ldots,X_n$ and time $t=1,\ldots,n$. Let us equate $C(X_1,\ldots,X_n) = C(X_1,\ldots,X_n,\hat{X}_{n+1})$. After some simple algebra the required X_{n+1} is obtained in formula (10.13).

$$(10.13) \qquad \hat{X}_{n+1} = \frac{1}{\frac{n(n+1)}{2}} \left\{ \frac{2}{n} \sum_{i>j} (Z_i-X_j)(i-j) + \sum_{i=1}^{n} X_i(n+1-i) \right\}.$$

Equation (10.13) is of course a special case of the formula (10.14) below for (X_i,Y_i) $i=1,\ldots,n$ pairs of observations.

$$(10.14) \qquad \hat{X}_{n+1} = \frac{1}{\frac{n(n+1)}{2}} = \left\{ \frac{2}{n} \sum_{i>j} (X_i-X_j)(Y_i-Y_j) + \sum_{i=1}^{n} X_i(Y_{n+1}-Y_i) \right\}.$$

3. The Monotone Case

A series $Z_1,\ldots,Z_N$ is (positive) monotone if and only if

$$Z_i > Z_j \text{ for all } i>j.$$

Thus, one coefficient for goodness-of fit to assess monotone association of series and its order is the following:

$$(10.15) \qquad MON(Z) = \mu = \frac{\sum_{i>j} (Z_i - Z_j)W_{ij}}{\sum_{i>j} |Z_i - Z_j| W_{ij}}$$

where weights $W_{ij} \geq 0$. For sake of simplicity let $W_{ij} \equiv 1$. By using the Persistent Structure Principle (P.S.P.), presuming that the trend for the given series is monotone, we equate $MON(\underline{Z})$ to $MON(\underline{Z},\hat{Z}_{N+1})$. After some algebra the required forecast value is obtained.

$$(10.16)\qquad \hat{Z}_{N+1} = \frac{(1-\mu)\sum_{N_1} Z_i + (1+\mu)\sum_{N_2} Z_i}{(1-\mu)N_1 + (1+\mu)N_2}$$

where μ is the Monotonicity coefficient for the series $Z_1,,\ldots Z_n$, namely a function of the data. $\sum_{N_1} Z_i$, $\sum_{N_2} Z_i$ are the summation over N_1 and N_2 values Z_i such that they are less than $\hat{Z}_{N+1}$ (yet, unknown) or greater than $\hat{Z}_{N+1}$, respectively. The unique solution of (10.16) is obtained by a finite iterative algorithm when $-1<\mu<1$. When $\mu = 0$, $\hat{Z}_{N+1} = N^{-1}\sum_{i=1}^{N} Z_i$, namely, the arithmetic mean. When $\mu = 1$ the solution is not unique any more, and any $\hat{Z}_{N+1} \geq \underset{i}{Max}\,\{Z_i\}$ could be obtained. For such a case of perfect monotone shape of the series, additional constraints should be used. Thus, the same formula could be used for the series of first differences, etc. Hence, the condition for a monotone convex (concave) series is that $\mu = 1$ ($\mu = -1$) for the first differences of the original series.

The coefficient μ could be generalized to any polytone series with m turning points and to a coefficient of local monotonicity as well.

4. The Quadratic Case

A series $Z_1,\ldots,Z_n$ is quadratic if and only if the series of first differences is linear. In other words, $\Delta_2,\ldots,\Delta_n$ is a linear series where $\Delta_i = Z_i - Z_{i-1}$ $i=2,\ldots,n$. Based on the solution obtained in section 2 for the monotone case the prediction formula based on the assumption of a quadratic shape is given in (10.17):

$$(10.17)\qquad \hat{X}_{n+1} = (1+a)X_n + (1-2a)X_{n-1} + (a-1)X_{n-2}$$

where the coefficient a is a function of the data and is computed the same way as for eq.(10.10). A good compromise as is done in X-11 could be for a=1.5. For a=1.5, eq.(10.17) is reduced to the following formula:

(10.18) $\hat{X}_{n+1} = 2.5\ X_n - 2X_n + 0.5 \cdot X_{n-2}$

Similar formulas could be derived to higher order of polynoms. For example, the prediction equation for the Qubic case is given in (10.19) and the reader can very easily derive formula to higher order of polynom.

(10.19) $\hat{X}_{n+1} = (2+a)X_n - 3a\ x_{n-1} + (3a-2)X_{n-2} + (1-a)X_{n-3}$.

The author suggests plugging eq. (19.17) or (10.18) into X-11 program for more complicated case.

5. Prediction by Examples

The persistent structure principle could be used for purposes of prediction of trend and seasonal components. For the examples below, fixed seasonality was assumed, or in other words, $S_{r+12L} = S_{r+12(L-M)}$ for r=1,..., 12 and M=1,...,L. The prediction for the trend was obtained by using the Persistant Structure Principle (P.S.P.) for the seasonally adjusted series Z_t t=1,...,,N. These examples have been decomposed by the nonmetric method LPTA discussed earlier.

a. Consumption of Electricity in the U.S.A. in the years 1951-58

This series of 96 observations is given in the chapter 3.D. From the plotted graph in Figure (3.D1) below it is easy to verify that the series has approximately a fixed seasonal pattern combined with a monotone trend. As a matter of fact, the trend is very close to being linear.

The estimated seasonal pattern for the first 84 observations is given in Table (10.1) for multiplicative and additive models.

Table (10.1): The estimated seasonal patterns for electricity in the U.S.A. using the multiplicative model (presented in percentages) and Additive Model (presented in their absolute values).

Models	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Mul.	119.8	107.6	103.6	92.9	86.3	80.9	83.2	90.0	96.8	106.2	113.0	120.1
Add.	67.1	25.1	12.1	-23.9	-45.9	-63.9	-56.9	-34.9	-10.9	22.1	43.1	67.1

Forecasts for 12 units ahead were computed for the year 1958 with the assumptions of linear and monotone trend for both Multiplicative and additive models. The prediction is based on the previous 7 whole years and are given in Table 2. The coefficient of goodness-of-fit C=0.996. This coefficient is the coefficient of monotonicity (10.15) for the seasonally adjusted series, namely seasonality has been removed. The arithmetic mean of the absolute percent error is lower for the additive model than the multiplicative one. With the monotone and linear assumptions for trend, almost the same results were obtained. Equations (10.16)and (10.10) have been applied recursively for the seasonally adjusted data $Z_1,\ldots,Z_{84},\hat{Z}_{84+1},\ldots,\hat{Z}_{84+12}$. The predicted values are produced by multiplying (adding) the forecasted value for the Trend $\hat{Z}_{84+1},\ldots,\hat{Z}_{84+12}$, with their appropriate fixed seasonal coefficients: $S_1,S_2\ldots,S_{12}$ $(S_1,\ldots,s_{12})$ respectively.

Table (10.2): Actual data, forecasted values and percent error for the year 1958 obtained by the Persistent Structure Principle (P.S.P.) method. Prediction in (a) and (b) is obtained by monotone and linear assumption, respectively, using multiplicative model. In (c) and (d) an additive model was adopted.

	Jan.	Feb.	Mar.	Apr.	May	Jun	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.
Actual Data	529	477	463	423	398	380	389	419	448	493	526	560
Multiplicative Model												
(a) Mon	518	463	448	402	373	350	360	390	419	460	489	520
%	-2.1	-2.9	-3.2	-5.0	-6.3	-7.9	-7.5	-6.9	-6.5	-6.7	-7.0	-7.1
(b) LIN	512	456	441	396	367	344	354	383	412	452	481	511
%	-3.2	-4.4	-4.8	-6.4	-7.8	-9.5	-9.0	-8.6	-8.0	-8.3	-8.6	-8.7
Additive Model												
(c) Mon	512	470	457	421	399	381	388	410	434	467	488	512
%	-3.2	-1.5	-1.3	-0.5	0.2	.3	-.3	-2.1	-3.1	-5.3	-7.2	-8.6
(d) LIN	517	475	462	426	404	386	393	415	439	472	493	517
%	-2.3	.4	-.2	.7	1.5	1.6	1.0	-1.0	-2.0	-4.3	-6.3	-7.7

The arithmetic means for the 12 absolute percent errors for the 4 variants are: (a) 5.8 (b) 7.3 (c) 2.8 (d) 2.4

(b) The Chatfield-Prothero Case-Study: Sales of Company X

This example of 77 observations was discussed earlier in chapter (3.b.1). As we have already mentioned, this series, was analyzed as a case-study by Chatfield and Prothero (1973) and some other 10 discussants in the Journal of the Royal Statistical Society (1973) part A. The original series is given in Table B in Appendix B and its graph is plotted earlier in Figure 3.b.2.

The estimated seasonal pattern based on the 77 observations is given in Table (10.3). The coefficient of goodness-of-fit (10.16) equal to C=.954 is obtained by using multiplicative model. The actual data and forecasts for 6 units ahead obtained by our P.S.P. method are present in Table (10.4). The predicted values obtained by Chatfield & Prothero and Box-Jenkins (1973) are presented as well for comparison purposes. Chatfield & Prothero identified an ARMA(1,0) $x(0,1)_{12}$ model on $W_t = \nabla\nabla_{12}\log_{10}Y_t$ where Y_t is the original series. On the other hand, Box and and Jenkins suggested a different transformation on data and hence they identify the same ARMA model on $W_t = \nabla\nabla_{12}Y^{.25}$.

Table (10.3): Estimated Seasonal Patterns Using Multiplicative and Mixed Models for the 'Sales of Company X'.

Model	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mult.	130.0	82.8	64.7	56.5	43.2	51.0	68.0	86.4	134.5	171.4	178.9	132.6
Mixed $\underline{S}_t$	131.7	108.3	75.4	77.1	43.4	55.4	63.0	75.6	117.6	150.4	152.2	142.7
$\underline{s}_t$	-5.1	-71.3	-39.8	-71.3	-19.7	-17.5	31.9	52.0	52.3	48.5	56.3	-16.3

Table (10.4): Actual data and forecasts value for 6 units ahead starts from June 1971 for the various methods. The values in parentheses are perpercent errors. 50% confidence interval for the mixed model is (92.5, 106.2) percent and for the multiplicative model (93.6, 107.4) percent approximatel

The Method	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	M.A.P.E.*
Actual Data	260	304	390	614	783	872	
Chatfield-Prothero	305 (17.3)	482 (58.5)	673 (72.6)	990 (61.2)	1297 (65.6)	1387 (59.1)	55.7
Box-Jenkins (Using $Y^{.25}$ Transformation)	286 (10.0)	409 (34.5)	511 (31.0)	761 (23.9)	966 (23.4)	1091 (25.1)	24.6
P.S.P. 50% confidence interval**	247 (231,265)	329 (308,353)	417 (391,448)	649 (608,698)	828 (775,889)	864 (809,928)	
Multiplicative Model (monotone trend)	(-5.0)	(8.2)	(6.9)	(5.7)	(5.7)	(0.9)	5.4
P.S.P. 50% confidence interval** (Mixed Model, Monotone Trend)	253 (223,269) (-2.7)	319 (295,339) (4.9)	398 (368,423) (2.0)	619 (572,658) (0.8)	786 (727,835) (0.4)	823 (761,875) (-5.6)	2.7

* Mean of absolute percent error.

** The confidence interval is based on multiplicative irregularities.

(c) International Airline Passenger: Monthly Totals

This very well known series was analyzed by Box and Jenkins (1970, p.305) and has been discussed in chapter 3 as in example 3.b.2. The trend is clearly monotone and for a periods of length p=12 a quite fixed seasonal pattern seems to be a good approximation. Box and Jenkins assumed that the underlying model of seasonality is a multiplicative one and thus they transformed the raw data by natural logarithms. For the transformed series an ARIMA $(0,1,1)x(0,1,1)_{12}$ model was identified (known as the Airline Model). Forecasts for 12 months ahead were made from an arbitrarily selected origin, July 1957. That means that the parameters for the model $\nabla\nabla_{12}\ln Y_t = (1-\theta B)\,(1-\phi B^{12})a_t$ were computed on the first 102 observations. By using the TYMPAC* program the following values were estimated: $\theta = 0.3897$, $\phi = 0.6257$ and $R^2 \equiv 0.983$. The predicted values are given in Table (10.6). Forecasts obtained by the P.S.P. method with the assumption of monotone trend were computed as well as for a short subseries of 30 observations, starts from Jan. 1955. The estimated seasonal patterns are given in Table (10.5).

Table (10.5): Estimated Seasonal Pattern that were computed using Multiplicative Model (a) for the first 102 observations, (b) for the 30 observations prior to July 1957, and (c) as in (b), using additive model.

Series	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
(a)	91.1	89.0	104.0	100.6	98.7	110.0	121.3	119.7	104.7	91.5	79.5	89.8
(b)	90.8	87.8	99.3	98.9	98.2	115.5	126.8	120.5	104.3	90.2	78.3	89.4
(c)	-26.1	-39.4	-1.9	-7.2	-6.2	38.8	85.6	68.5	18.5	-30.5	-67.5	-32.5

* TYMPAX Program (Estimation of parameters in Linear Time Series Models) are owned by the Queen's Statistics Council of Canada and approved by Donald G. Watts.

The estimated Seasonal Patterns given in table 5 for (a) and (b) seem very similar. It indicates that the multiplicative seasonal component is nearly fixed and that the nonmetric approach can estimate seasonal component with as low as 30 observations. The prediction obtained by the P.S.P. method based on the first 102 observations is about the same accuracy as that obtained by Box-Jenkins. However, using P.S.P. on the most recent past (only 30 observation) yields much better forecasts with monotone trend assumption with both multiplicative and additive models.

Table (10.6): The Airlines data: Forecasts for 12 months ahead (starting from July 1957 obtained by (a) Box-Jenkins Approach, (b) P.S.P. method using multiplicative models. In (b.1) and (b.2) monotone and linear assumptions have been assumed. For the same forecasting range the P.S.P. method has been used based on the 30 observations starting from January 1955. In (c.1) and (c.2) the forecasts values are given presumably monotone and linear trend, respectively, using multiplicative model. D.1 and D.2 are similar to (c.1) and (c.2) except that additive model has been used. (The values in parentheses are the percent errors).

The Method	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	M.A.P.
Actual Data	465	467	404	347	305	336	340	318	362	348	363	435	
Box-Jenkins(a)	465	457	405	354	307	355	365	353	413	405	410	477	
	(1.0)	(-2.1)	(0.2)	(2.0)	(.7)	(5.7)	(7.4)	(11.0)	(14.1)	(16.4)	(12.9)	(9.7)	6.8
P.S.P. (b.1)Mon	432	426	373	326	283	320	324	317	371	358	352	392	
	(-7.0)	(-8.7)	(-7.7)	(-6.0)	(-7.1)	(4.8)	(-4.5)	(-.3)	(2.4)	(3.0)	(-3.1)	(-9.9)	5.4
(b.2)Lin	477	476	418	366	318	359	365	356	417	403	395	440	7.1
	(2.7)	(1.9)	(3.5)	(5.5)	(4.4)	(7.0)	(7.3)	(12.1)	(15.1)	(15.7)	(8.9)	(1.3)	
(c.1)Mon	460	437	378	327	284	324	329	319	360	359	356	419	
	(-1.0)	(-6.4)	(-6.4)	(5.8)	(-6.9)	(-3.6)	(-3.2)	(0.3)	(-0.5)	(3.2)	(-1.9)	(-3.7)	3.6
(c.2)Lin	465	443	383	332	288	329	334	323	365	364	361	425	
	(0)	(-5.1)	(-5.2)	(-4.3)	(-5.6)	(-2.1)	(-1.8)	(1.6)	(0.8)	(4.6)	(-0.5)	(-2.3)	2.8
D.1 Mon	467	450	400	351	314	349	355	342	379	374	375	420	
	(0.4)	(-3.6)	(-1.0)	(1.1)	(2.9)	(3.9)	(4.4)	(7.5)	(4.7)	(7.5)	(3.3)	(-3.4)	3.0
D.2 Lin	478	466	418	369	333	368	374	361	399	393	394	439	
	(2.8)	(-0.2)	(3.5)	(6.3)	(-4.0)	(9.5)	(10.0)	(13.5)	(10.2)	(12.9)	(8.5)	(0.9)	6.9

The following are two out of 13 series prepared at the Bureau of the Census for the ASA-Census-NBER (October 1981). These are Bureau of Labor Statistics series. Their original observations are given in Tables E and F in Appendix B.

(d) Agricultural Employment, Men, 20 years and older

This monthly series from Jan. 1967 till October 1980 has 166 observations. It seems that the trend has mainly three turning points: The first 5 years and the next 5 years have decline trend each. In other words, this part of the series is a piece-wise monotone of order 2. The last part of the series has a positive slope trend. By applying multiplicative model the folliwng coefficients of Polytonicity were obtained for monotone assumption: $\mu_1^{(12)} = -0.64$, $\text{Max}|\mu_1| = 0.86$, $M_1^{(12)} = 0.63$. When Polytonicity of order k=3 was assumed $\mu_3 = 0.38$ $\text{Max}|\mu_3|^{(12)} = 0.90$ $M_3^{(12)} = 0.84$. The estimated seasonal pattern is:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
91.6	92.5	94.2	99.2	102.5	107.2	107.1	105.6	104.4	103.6	98.4	93.7

In Figure (2.2), the original series and trend estimation is given.

A long range prediction has been done to the last 24 observations based on only 24 observations prior to the forecasting range. Each predicted value was based on the previous forecasted value in a recursive way. Assumption of fixed seasonality has been used. In table 10.7 the original observations and forecasted values are presented. We tried our P.S.P. approach, multiplicative model with both linear and montone assumptions. For purposes of comparison and for the same range, prediction results obtained by Gersch and Kitagawa (1982) are presented as well. the innovative step in Gersch and Kitagawa report is the maximization of the expected entropy of the predictive distribution interpretation of the minimum AIC procedure. The Modeling and smoothing of series is done using a Kalman prediction/smoother Akaike AIC criterion methodology.

Table 10.7: Original data and forecasts values obtained by Gersch and Kitagawa and the P.S.P. method (multiplicative model) with assumptions of linear and montone trend. 95% confidence interval (multiplicative irregularities are 95.6 and 106.1 percent of each point estimation.

Year	Month	Original	Gersch & Kitagawa Prediction	%Error	P.S.P. Multiplicative MON	%Error	LIN	%Error
1978	11	2277	2333	(2.5)	2344	(2.9)	2294	(0.7)
"	12	2250	2201	-(2.2)	2248	-(0.0)	2189	-(2.7)
1979	1	2084	2148	(3.1)	2229	(6.9)	2166	(3.9)
	2	2117	2167	(2.4)	2157	(1.9)	2095	-(1.0)
	3	2176	2222	(2.1)	2197	(0.9)	2133	-(2.0)
	4	2237	2363	(5.6)	2331	(4.2)	2263	(1.2)
	5	2342	2444	(4.4)	2481	(5.9)	2409	(2.9)
	6	2509	2569	(2.4)	2629	(4.8)	2552	(1.7)
	7	2520	2561	(1.6)	2553	(1.3)	2478	-(1.7)
	8	2554	2511	-(1.7)	2559	(0.2)	2484	-(2.7)
	9	2498	2465	-(1.3)	2469	-(1.2)	2396	-(4.1)
	10	2472	2465	-(0.2)	2488	(0.6)	2415	-(2.3)
	11	2403	2348	-(2.3)	2344	-(2.5)	2276	-(5.2)
	12	2292	2230	-(2.7)	2248	-(1.9)	2182	-(4.8)
1980	1	2160	2186	(1.2)	2229	(3.2)	2164	(0.2)
	2	2213	2208	-(0.2)	2157	-(2.5)	2094	-(5.4)
	3	2217	2262	(2.0)	2197	-(0.9)	2132	-(3.8)
"	4	2255	2401	(6.5)	2331	(3.4)	2263	(0.3)
	5	2422	2479	(2.4)	2481	(2.4)	2409	-(0.5)
	6	2470	2602	(5.3)	2629	(6.4)	2552	(3.3)
	7	2475	2592	(4.7)	2553	(3.1)	2479	(0.2)
	8	2455	2540	(3.5)	2559	(4.2)	2484	(1.2)
	9	2525	2495	-(1.2)	2469	-(2.2)	2396	-(5.1)
	10	2459	2494	(1.4)	2488	(1.2)	2415	-(1.8)
Average of absolute percent error:				2.6		2.7		2.4

(e) All Employees in Food Industries

This series from January 1967 till December 1979 includes 156 observations. The series has 3 main turning points. More specifically, the trend for the first 3 years, the next 5 years, and the last 5 years has increase, decrease and increase tone, respectively. Some of the computed values are: $\mu_3 = 0.38$, $\text{Max}|\mu_3(12)| = 0.90$ and $M_3(12) = 0.84$. The estimated seasonal pattern is:

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
96.7	95.9	96.0	95.8	96.7	99.9	101.9	106.6	106.9	103.5	101.0	99.0

In Figure (2.4), the original series the trend estimation is exhibited.

A long range prediction has been done to the last 24 observations namely for the years 1978 and 1979 based on only 24 observations prior to the prediction range, namely, the years 1976 and 1977. In table 10.8 the actual data and the forecasts results obtained by our P.S.P. approach (multiplicative model and Gersch-Kitagawa approach are presented.

Table (10.8): Original data and forecast values obtained by Gersch-Kitagawa approach and P.S.P. methods (multiplicative model) with assumption of monotone and linear trend. 95% confidence interval (based on multiplicative irregularities) are approximately 98% and 101.4%, respectively.

Year	Month	Original data	Gersch-Kitagawa	% Error	Raveh (P.S.P.)(Multiplicative) MON	% Error	LIN	% Error
1978	1	1665	1650	-(0.9)	1660	-(0.3)	1669	(0.2)
	2	1656	1641	-(0.9)	1654	-(0.1)	1665	(0.5)
	3	1668	1645	-(1.4)	1647	-(1.2)	1658	-(0.6)
	4	1664	1645	(1.1)	1653	-(0.7)	1665	(0.0)
	5	1669	1663	(0.4)	1666	-(0.2)	1678	(0.5)
	6	1722	1722	(0.0)	1726	(0.2)	1738	(0.9)
"	7	1749	1760	(0.6)	1755	(0.3)	1768	(1.0)
	8	1823	1849	(1.4)	1838	(0.8)	1852	(1.6)
	9	1830	1855	(1.4)	1845	(0.8)	1859	(1.6)
	10	1774	1799	(1.4)	1774	(0.0)	1787	(0.7)
	11	1746	1753	(0.4)	1731	-(0.9)	1744	(0.0)
	12	1724	1717	-(0.4)	1693	-(1.8)	1705	-(1.1)
1979	1	1685	1671	(0.8)	1659	-(1.5)	1672	-(0.8)
	2	1666	1660	-(0.4)	1654	-(0.7)	1666	(0.0)
	3	1676	1663	-(0.8)	1646	-(1.8)	1658	-(1.0)
	4	1666	1663	(0.2)	1653	-(0.8)	1665	(0.0)
"	5	1679	1682	-(0.2)	1666	-(0.8)	1678	(0.0)
	6	1728	1741	(0.8)	1726	-(0.1)	1738	(0.5)
	7	1750	1779	(1.7)	1755	(0.3)	1768	(1.0)
	8	1829	1869	(2.2)	1838	(0.5)	1852	(1.3)
	9	1835	1875	(2.2)	1845	(0.5)	1859	(1.3)
	10	1782	1819	(2.1)	1774	-(0.4)	1787	(0.3)
	11	1736	1773	(2.1)	1732	-(0.2)	1744	(0.5)
	12	1706	1737	(1.8)	1693	-(0.7)	1705	(0.0)
Average of absolute percent error=			1.1		0.6		0.6	

11: APPLY BOX-JENKINS APPROACH ON SEASONALLY ADJUSTED DATA.

In this chapter the Box-Jenkins (BJ) approach is combined with our LPTA method for prediction purposes. As mentioned in Flowchart 10, in the first stage LPTA procedure decomposes the original data for the appropriate type of seasonality model. In the second stage, using the Box-Jenkins approach to the S.A.D. gives us forecasts for the trend. Prediction for the series are obtained by combining the trend forecasts with the estimated adequate seasonal pattern. In the following three examples, we limit ourselves to fixed seasonal patterns only. Forecasts for the same prediction range have been done for the first 2 examples by P.S.P. approach in the previous chapter.

Chatfield-Prothero Case Study:

Predictions for 6 units ahead, starting on June 1971, were obtained by Chatfield and Prothero (1973), Box and Jenkins (1975), and others. Some of these results are given in Table (11.1) below. As a different tactic, the series has been decomposed by LPTA to the Box-Jenkins method. Based on the findings of example 3.6.1, the mixed model was chosen as an appropriate one. We identify* an ARIMA (1,1,0) model of the form

$$(11.1) \qquad (1-\phi B)\nabla z_t = a_t \qquad \hat{\phi} = -0.585,\ R^2 = 0.91$$

where a_t is a white noise and R^2 a coefficient of goodness of fit for BJ approach. The ARIMA (1,1,0) model is found by observing in table (11.1) at the 24 first autocorrelations of the original data, the first differences, and the partial autocorrelations of the first differences and the residual autocorrelations of the first differences. The residual autocorrelations are given further in Table (11.2).

* For identification process APCOR program (Auto- and partial correlations) was used. The program is owned by the Queen's Statistics Council of Canada and approved by Donald G. Watts.

Table (11.1): Actual data and forecasts for six months ahead based on 77 observations prior to June 1971.
(The values in parentheses are percent errors.)

1971

The Method	June	July	Aug	Sept	Oct	Nov	Average of 6 absolute percent errors)
Actual data	260	304	390	614	783	872	
CP (Model A)	(17.3)	(58.5)	(72.6)	(61.2)	(65.6)	(59.1)	55.7
Akaike's (Simple-minded forecast)	328	401	465	890	1036	1079	
	(26.2)	(31.9)	(19.2)	(45.0)	(32.3)	(23.7)	29.7
Box & Jenkins using $y^{.25}$ transformation)	286	409	511	761	966	1091	
	(10.0)	(34.5)	(31.0)	(23.9)	(23.4)	(25.1)	24.6
Harrison Muldo Analysis	322	327	484	681	943	932	
	(23.8)	(7.6)	(24.1)	(10.9)	(20.4)	(6.9)	15.6
Medium & Long-term forecasting	255	373	504	749	970	1020	
	(2.0)	(22.7)	(29.2)	27.0)	(23.9)	(17.0)	19.5
Manning (using Holt-Winter)	299	336	458	696	910	968	
	(12.0)	(10.5)	(17.4)	(13.4)	(16.2)	(11.0)	13.4

Table (11.2): Sample autocorrelation and partial autocorrelation of $\{z_t\}$ (Periodicity-free series).

Series	lags	Autocorrelations					
z_t	1-6	.899	.899	.838	.829	.796	.769
	7-12	.720	.689	.647	.597	.582	.524
	13-18	.516	.453	.430	.397	.374	.329
	19-24	.311	.270	.243	.200	.163	.128
∇z_t	1-6	-.526	.322	-.178	-.007	-.054	.087
	7-12	-.095	-.053	-.127	-.265	.401	-.305
	13-18	.205	-.096	.062	-.097	.188	.111
	19-24	.123	.010	.017	.036	-.041	-.120
		Partial Autocorrelation					
∇z_t	1-6	-.526	.063	.018	-.152	-.148	-.067
	7-12	-.023	-.082	.209	-.171	.216	.075
	13-18	.001	.048	.079	-.019	.112	.145
	19-24	.094	.060	.275	.051	.056	-.145

Table (11.4): Residual autocorrelations up to 24 lags*

lags												
1-12	-.01	-.03	-.15	-.19	-.05	.02	-.12	.08	-.00	-.14	.24	-.12
13-14	.06	-.09	-.06	-.01	.13	-.00	.10	.09	.06	-.01	-.17	-.25

* 95 percent limits for correlations are plus and minus .232.

In addition, we include 3 additional sets of forecasts; (A) using ARIMA (1,1,0) with multiplicative seasonality, (B) using ARIMA (1,1,1) with mixed seasonality and (C) using ARIMA (0,1,1) with mixed seasonality. Model D, that of ARIMA (1,1,0) with mixed seasonality is slightly better than Model C which is ARIMA (0,1,1). The R^2 is greater and the sum of squares after the regression is smaller (see Table (11.5)). The cumulative periodogram is smoother (and close to the 45 line) and the chi square of the 24 first residual auto-correlations (if the model is the right one) is also smaller. Still, the forecasts obtained by using the ARIMA (0,1,1) model for the trend look better.

In order to choose the suitable model when several models fit the data equally well we used the MAICE (minimum AIC estimation) procedure which selects a model by using Akaike's Information criterion (AIC). The normalized AIC for the ARIMA model (p,d,q) is given specifically (see Ozaki 1977, p.293) by the equation below

$$\text{AIC}(p,d,q) = N \cdot \log \hat{\sigma}_a^2 + [N/(N-d)]\, 2(p + q + 1 + \delta_{d0})$$

where

$$\delta_{d0} = \begin{cases} 1 \text{ if } d = 0 \\ 0 \text{ if } d \neq 0 \end{cases}$$

N is the number of data points. $\hat{\sigma}_a^2 = S(\phi, \theta)/N$ where $S(\phi, \theta)$ is the minimum sum of squares of the residuals directly from the series $[Z_t]$ $t=1,\ldots,N$ uses a backward forecasting technique as described by Box-Jenkins (1970). The $\hat{\sigma}_a^2$ and AIC computed values are given in Table (11.5). By adopting this MAICE Procedure we select Model D.

Table (11.5): Prediction obtained by combining Box-Jenkins approach and Seasonal Adjustment procedure. Three different ARIMA models were combined with Mixed Seasonality model. The percent errors are in parentheses.

Nonmetric Method Combined with B-J	R^2	σ_a^2	AIC*	June	July	Aug.	Sept.	Oct.	Nov.	Average of absolute percent errors)
Estimated parameters										
A. Multiplicative model combined with ARIMA (1,1,0)				278 (6.9)	395 (29.9)	482 (23.6)	772 (25.7)	967 (23.5)	1024 (17.4)	21.2
$\hat{\phi} = -.55$	0.89	27.37	258.9							
B. Mixed Model Combined with ARIMA (1,1,1)				266 (2.3)	368 (21.1)	419 (7.4)	687 (11.9)	851 (8.6)	896 (2.8)	9.0
$\hat{\phi} = -.65$ $\hat{\theta} = -.10$	0.91	14.54	212.2							
C. Mixed Model Combined with ARIMA (0,1,1)				271 (4.1)	342 (12.6)	413 (5.6)	652 (6.2)	827 (5.6)	860 (-1.4)	5.9
$\hat{\phi} = -.52$	0.90	15.51	215.2							
D. Mixed Model Combined with ARIMA (1,1,0)				265 (1.8)	365 (20.0)	419 (7.3)	681 (10.9)	849 (8.5)	889 (2.0)	8.4
$\hat{\phi} = -.59$	0.91	14.38	209.3							

*AIC: Akaike's Information Criterion

In this example, in contradiction to the previous analysis in J.R.S.S. (1973), part A, it seems to us that the mixed model is an approppriate one in order to remove the systematc (seasonal) fluctuations. In both methods: P.S.P. in the previous chapter and in this chapter, prediction for 6 months ahead, seem much more accurate while using mixed model than mutiplicative, for the same range of prediction; 6 months ahead starting from Jun 1971.

We believe that the main reasons for better result for this case study are:

1) the mixed (fixed) type model for Seasonality is more appropriate than the multiplicative;

2) No transformation like log or power have been used in order to reduce the fluctuations.

While we avoided transforming the data twice (log and antiolog, etc.), the series have not been distorted.

International Airline Passenger: Monthly Totals

For the first 102 observations, the following coefficients were obtained for multiplicative model: μ_1 = 0.94, Max $\mu_1^{(12)}$ = 0.996, $M_1^{(12)}$ = 0.93 and convexity measure of the trend μ_Δ = 0.76 indicate almost convex curve. The estimated seasonal pattern is given below (in percentages):

Month i	1	2	3	4	5	6	7	8	9	10	11	12
$S_i^{(12)}$	91.2	88.8	103.8	100.4	99.4	110.1	121.6	119.5	104.5	91.3	79.5	89.7

An ARIMA (0,1,1) model was identified for the S.A.D. with $\hat{\phi}$ = 0.12 and R^2 = 0.986. The forecasts for 12 units ahead are given in Table (11.6) indicating slightly better results than those obtained by using only Box-Jenkins procedure.

Table (11.6): Forecasts computed by our approach and by BJ for the airlines data and the Women Unemployed.

Series	Methods	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	MAPE*
Airlines Data	Actual							465	467	404	347	305	336	
		340	318	362	348	363	435							
	LPA+ ARIMA (0,1,1)							461	455	398	347	301	339	
		346	338	394	382	377	419							3.5
	ARIMA $(0,1,1)(0,1,1)_{12}$ for logarithm transformation							465	457	405	354	307	355	
		365	353	413	405	410	477							6.8
Women Unemployed	Actual	127.5	128.4	129.5	131.9	120.1	109.1	118.3						
	Mult.LPA+ ARIMA (1,1,0)	126.4	129.0	126.1	122.7	117.2	107.4	111.5						2.8
	Add.LPA+ ARIMA (1,1,0)	125.2	127.0	125.0	122.5	118.8	112.0	114.6						2.9
	ARIMA$(1,1,0)(0,1,0)_{12}$	128.7	135.6	141.7	141.6	136.0	128.5	137.4						10.0

* Mean absolute percent error.

Women Unemployed (1,000's) in U.K. 1-67 to 7-72.

This series was analyzed by Anderson (1976, p.132, series E). Anderson assumed an additive seasonality model and fitted the following model: $(1- 0.349B)\nabla\nabla_{12}Y_t = a_t$, namely an ARIMA $(1,1,0)\text{x}(0,1,0)_{12}$ model. Using this model for the first 60 observations the value $\theta = 0.3376$ was estimated ard the forecasts for seven months ahead are given in Table (11.6). The same series was analyzed by our procedure. $\mu_2 = 0.84$ Max $\mu_2^{(12)} = 0.98$, the trend is estimated as a polytone series of order m=2 with turning point at the 36th observavation. Both the multiplicative and additive seasonality models are equally suitable with $M_2^{(12)} = 0.90$. The seasonality components for each model is presented below:

Month i	1	2	3	4	5	6	7	8	9	10	11	12
$S_i^{(12)}$	104.7	106.4	103.8	101.0	96.4	88.4	91.7	95.3	96.5	104.6	107.6	103.6
$s_i^{(12)}$	4.08	5.48	3.30	.73	-2.95	-9.79	-7.20	-4.39	-2.80	3.98	6.28	3.28

For both models ARIMA (1,1,0) models were estimated for the S.A.D. The estimated parameters for these models were $\theta = 0.339$ and $\theta = 0.372$, respectively. The forecasts for seven months ahead are given for both models in Table (11.6).

The Box Jenkins approach to time series analysis is a significant step in this direction because it provides a general class of ARIMA models which could be fitted to a large number of time series. Our mixed (additive-multiplicative) seasonality model in this line is an extension of the analysis that can be conducted using the BJ approach. If "the proof of the pudding is in the eating" then our results, both in terms of fit and forecasts using the Chatfield Prothero data and two other examples, indicate the possibility of improving forecasting by using deseasonalized procedure (for fixed seasonal monthly series).

12. PERSISTENT STRUCTURE PRINCIPLE COMBINED WITH X-11

The Pertsistent Structure Principle is suggested to be replaced by the built-in formula of X-11 in order to forecast seasonal factors one year ahead. It will be shown that the formula used in X-11 is a special case of the proposed P.S.P. method, see Raveh (1982). Some of the details have already been discussed in chapter 10.2

Let $S_{i,j}$ denote the estimated seasonal factor for the j^{th} month (j=Jan,..., Dec.) in the i^{th} year. In order to forecast the 1 year ahead factors $S_{n+1,j}$ based on the previous n year, X-11 apply eq. (12.1) below.

$$(12.1) \qquad \hat{S}_{n+1,j} = S_{n,j} + 1/2[S_{n,j} - S_{n-1,j}] = = 1.5\, S_{n,j} - 0.5\, S_{n-1,j} \, . \qquad (j-1,\ldots,12)$$

Let $\underline{S}=S_1,\ldots,S_n$ be a series. $\underline{S}$ is <u>linear</u> if and only if $S_K - S_{K-1} = S_{K-1} - S_{K-2}$ for all $K=3,\ldots,n$. or $\Delta S_K = \Delta S_{K-1}$ where $\Delta S_K = S_K - S_{K-1}$ or $\Delta^2 S_K = 0$ for all K or $\sum_{K=3}^{n} \Delta^2 S_K = 0$ or,

$$(12.2) \qquad \sum_{K=3}^{n} [\Delta^2 S_K]^2 = 0$$

The above conditions mean that the slope remains constant over time. The series $\underline{S}$ is <u>most dissimilar</u> to a linear curve when $\Delta S_K = -\Delta S_{K-1}$ or $\Delta^2 S_K = 2\Delta S_K$. In other words, this means that the slope changes its sign every two consecutive data points while its absolute value remains the same over time. Thus, the quantity (12.3) could be used as a measure for linearity.

$$(12.3) \qquad C(\underline{S}) = 1 - \frac{\sum_{K=3}^{n} [\Delta^2 S_K]^2}{\sum_{K=2}^{n} [2\Delta S_K]^2} = 1 - R$$

$C(\underline{S})$ might be a coefficient of linearity which varies between 0 and 1. $C(\underline{S}) = 1$ (R=0) if and only if $\underline{S}$ is a perfect liner series. $C(\underline{S})=0$ (R=1) if and only if $\underline{S}$ is perfectly negatively correlate with its order. Thus, the series [a,b,a,...,b] where $a \neq b$ is the most dissimilar to a linear series in the sense of the above definitions for linearity. For both cases when $C(\underline{S}) = 1$ or $C(\underline{S}) = 0$, perfect prediction is obtained.

A coefficient of association is a numerical value summarizing the strength or degree of relationship for two variables. The numerical value of most measures lies between -1 and +1 (or 0 and 1). If the variables are perfectly associated, according to some criterion of "perfect" the measure achieves its maximum absolute value. As a matter of fact, the various coefficients of association are different versions of loss function. As the amount of deviation of empirical data from a given pefect defined relationsihip (structure) is increased, the loss function increased.

Pearson product-moment coefficient p gets its two extreme values ± 1 for perfect linear association either positive or negative slope, respectively. $\rho=0$ means very little about the shape of the curve.

The Persistent Structure Principle (P.S.P.) means that forecasted values are estimated in such a way that the values of appropriate coefficients of goodness-of-fit would be the same for both the augmented and the original series. Briefly, the predicted value $\hat{S}_{n+1}$ one unit ahead is estimated by solving the equation (12.4), below.

(12.4) $$C(\underline{S}) = (S_1,\ldots, S_n) = C(S_1,\ldots,S_n,\hat{S}_{n+1}) = C(\underline{S},\hat{S}_{n+1})$$

By using as a figure of merit the coefficient (12.3) and the Persistent Structure Principle, $S_{n+1,j}$ could be computed. Let's equate

$$C(S_{1,j},\ldots, S_{n,j}) = C(S_{1,j},\ldots,S_{n,j},\hat{S}_{n+1,j}).$$

By a simple maniuplation the required $\hat{S}_{n+1,j}$ is obtained:

(12.5) $\hat{S}_{n+1,j} = a\ S_{n,j} + (1-a)S_{n-1,j}$ $j=1,\ldots,12$

where a can be either $a_{(1)} = (2\sqrt{R}-2)/(2\sqrt{R}-1)$ or $a_{(2)} = (2\sqrt{R}+2/(2\sqrt{R}+1)$ is a function of the estimated values $s_{1,j},\ldots,S_{n,j}$. For the perfect extreme cases the following results are obtained.

when $C(\underline{S}) = 1$ then $a = 2$ and $\hat{S}_{n+1,j} = 2S_{n,j} - S_{n-1,j}$,

when $C(\underline{S}) = 0$ then $a = o$ and $\hat{S}_{n+1,j} = S_{n-1,j}$.

It is interesting to verify that the classical version of X-11 uses formula (12.5) in order to estimate seasonal factors one year ahead where a=1.5 is chosen as a compromise and not as a function of all the previous seasonal factors. Hence, formula (12.1) is a special case of formula (12.5) with a=1.5. Formula (12.5) might be too dependent on the last two observations. Other conditions for linearity could be used in order to overcome this dependence. thus, conditions for every four values as in (12.6) could be used instead of (12.1) for (S_i,Y_i) $i=1,\ldots,N$ pairs of observations.

(12.6) $$\frac{S_k - S_\ell}{Y_K-Y_\ell} = \frac{S_u - S_v}{Y_u-Y_v} \quad \text{for all } k>\ell \text{ and } u>v$$

The product-moment coefficient of correlation could be used as a figure of merit as well.

With assumption of Quadratic trend the formula (12.7) could be used instead of (12.5)

(12.7) $\hat{S}_{n+1,j} = (1+a)S_{n,j} + (1-2a)S_{n-1,j} + (a-1)S_{n-2,j}$ $j=1,2,\ldots,12$

or a built-in formula (12.8) where a=1.5 chose as a compromise.

(12.8) $\hat{S}_{n+1,j} = 2.5\ S_{n,j} - S_{n-1,j} + 0.5\ S_{n-2,j}$ $j=1,\ldots,12.$

Part 3

QUALITATIVE SERIES

Concepts of Markov Chains process are suggested to describe the relations between consecutive observations of a Qualitative Time Series. For purposes of describing a given series, some measures are defined based on the "average-mode". A graphic presentation to assist analyzation of a given set of series is also suggested later in chapter 17. The technique proposed in chapter 13 enables analysis and forecast of both qualitative (categorical) and quantitative (continuous) series over a given period of time. For the quantitative cases a process for finding an "optimal" division into categories is discussed in chapter 14. Some examples are given for purposes of demonstration.

13. ESTIMATE SEASONAL PATTERN AND PREDICTION

Time-series are characterized by data which is not independent but serially correlated, and the relations between consecutive observations are of interest. The dependence among observations of a qualitative time series will be introduced using both concepts of Markov Chains and data analytic technique. For purposes of description of a given series, some measures are adopted, see Colwell (1974) and Raveh and Tapieo (1980): Periodicity ($P_{(s)}$) for period length s, Constancy (C), and Heterogeneity ($H_{(s)}$). The periodic component is further decomposed into the two others, a constant part and a heterogeneous part such that $P_{(s)} = C + H_{(s)}$. The three measures defined above rely upon combinations of different modes.

A qualitative time series is often used in fields of research such as biology, meteorology, economics, and business administration. Our starting point is a collection of ordered observations, each one fitting one of R given categories; that is, a qualitative time series. Of course a quantitive time series that may appropriately be categorized can be analyzed the same way. The basic contributions of the chapter are two-fold: (a) Estimating the period's length and the pattern of periodicity (seasonal pattern). (b) Estimating Interval Forecasting of quantitative series that might be preferred, compared to usual point estimation.

Definitions and Notations

A qualitative time series is a collection of ordered observations, each of which falls precisely into one category, among R possibilities, at each point in time. Let N be the number of observations in the series.

A periodic qualitative time series with s as the length of the period, is a series $\{Y_t\}$ $t=1,\ldots,N$ in which every one of the j ($j=1,\ldots,s$) sub-series $\{Y_{j+sk}\}$ $k=0,1,\ldots,\{(\frac{N}{s})-1\}$ is a constant series, that is, all the observations fall into the same category.

An example might be the series $Y_1^* = \{A,D,C,A,B,A,D,C,A,B,A,D,C,A,B,A,D,C,A,B\}$. This series of N=20 observations has R=4 different categories. The period's length s=5 and the pattern of periodicity is: <14312>. The five constant sub-series are: {AAAA}, {DDDD}, { CCCC}, {AAAA}, {BBBB}. A qualitative time series is constant if R which is the number of different non-empty categories equal 1, or equivalently the period's length s=1. That is, a constant series is a generate periodic series. In other words, the series is within single category with probability 1. A qualitative series is in the least constant condition where the observations are within every one of the R categories with probability 1/R.

A qualitative time series is Heterogeneous when all the observations at each point of time j within the period -- (call it position j) -- fall into the same category, which is different for each j, j=1,...,s. A necessary condition for a series to be Heterogeneous is that R=s. The series $Y_2^* = \{A,C,B,A,C,B,A,C,B,A,C,B\}$ is an example of a qualitative and heterogeneous time series with R=s=3. All observations of the first position j=1 fall into category A. At position j=2 the category is C, and at position j=3 the category is B.

Consider a qualitative time series with R categories and let s (positions) be the length of the periods. In such a situation the data can be presented in a matrix (contingency table) of R x s. In each cell (i,j), N_{ij} is the number of observations that fall into category i at the jth position - within each period. The matrix denoted by $N = (N_{ij})$ has R rows for the categories and s columns for the s positions in the period.

For the sake of simplicity assume that we are dealing with a series constructed from L whole period; i.e., $N = L \cdot s$, L integer and thus $\sum_{i=1}^{R} N_{ij} = L$ for all j = 1,...,s. Let P_{ij} be the relative frequency to be the i category (i=1,...,R) at the j position (j=1,...,s). Thus, $P_{ij} = N_{ij}/L$ is the empirical probability to be at the cell (i,j). Let us denote $\underline{P} = (P_{ij})$ the R x s matrix of P_{ij}. Let sign by $\overline{P}_j = \underset{1<i<R}{\text{Max}} \{P_{ij}\}$, e.g., the value of the mode of column j in the matrix. The appropriate category is signed by $\overline{i}(j)$. $\overline{P}_j$ is the empirical probability of predicting the category of the j position. The arithmetic mean of the modes in columns is $\overline{P} = \sum_{j=1}^{s} \overline{P}_j/s$, which measures the mean probability of the "true" category prediction at any randomly chosen point in the series. Let us denote q_i the empirical probability of category i, that is, $q_i = \sum_{j=1}^{s} N_{ij}/N = \sum_{j=1}^{s} P_{ij}/s$. Let $q = \underset{1<i<R}{\text{Max}} \{q_i\}$ denote the maximum value of the raw totals, or the model value.

A perfect qualitative time series of period length s can be expressed by matrix $\underline{P}$ that has only one value of 1 and (R-1) zeroes for each column j, j=1,...,s. For example, the matrices $N = (N_{ij})$ and $\underline{P}=(P_{ij})$ for the series Y_1^* are presented in Table (13.1).

Table (13.1): Matrices (a) $N=(N_{ij})$ and (b) $\underline{P} = (P_{ij})$ for the series Y_i^*, R=4, s=5. $\overline{P}=1$, q=0.4.

R = 4 different categories

	1	2	3	4	5	Total
A	4			4		8
B					4	4
C			4			4
D		4				4
Total	4	4	4	4	4	20

s = 5 Points with the period
(a)

	1	2	3	4	5	q_i
A	1			1		.4
B					1	.2
C			1			.2
D		1				.2
$\overline{P}_j$	1	1	1	1	1	

s = 5 points with the period
(b)

3. Artificial Examples

The series Y_1^* can be presented as a Markov Chain that has transition matrix M of order f=5 which is the number of nonempty cells in the $N=(N_{ij})$ matrix. This matrix is given in Table (13.2). There are f=5 different "states": 1A, 2D, 3C, 4A, 5B. The state 1A, for instance, means that the first observation at each period is A.

Table (13.2): Transition matrix M for the qualitaitve time series Y_1^* which is a Markov Chain process.

M =

	1A	2D	3C	4A	5B
1A		1			
2D			1		
3C				1	
4A					1
5B	1				

Notice that although there are four (R=4) different categories: A,B,C, and D, there are five (f=5) different states, in the sense of Markov Chains processes. State 1A is different from state 4A since the transition process is to go from 5B to 1A and from 3C to 4A. Series Y_1^* is deterministic in that the dependence of every two consecutive observations is perfect. This series is an irreducible finite Markov Chain. The states 1A, 2D, 3C, 4A, and 5B are all persistent, non-null states with period f=5 and mean recurrence time 5. By computing its stationary distribution $\underline{\pi}_5 = (\pi_1, \pi^2 \pi_3 \pi_4 \pi_5)$ for the different states, the eigenvector $\underline{\pi}_5 = \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}$ reflects uniform distribution. This solution is obtained by solving the equation (1). The vector for stationary distribution is the eigenvector connected to the eigenvalue $\lambda = 1$ for the transition matrix M.

(13.1) $\underline{\pi}_f \, M = \underline{\pi}_f \quad \text{where} \sum_{j=1}^{f} \pi_j = 1$

The stationary distribution for the 4 different categories: A B C and C is the vector $\underline{q}_R = (q_1, q_2, q_3, q_4) = (0.4, 0.2, 0.2, 0.2)$. The deterministic series Y_1^* can be presented by a flowchart as in Figure (13.1).

Figure (13.1): Flowchart for the 5 different states of series Y_1^*.

```
The f=5 different states:  A--→D--→C--→A--→B
                           ↑               ↑
The s=5 positions          1   2   3   4   5
```

Let series Y_3^* be defined so that every one of the 3 categories A, B, or C happens at probability 1/3 every point of time. Transform this series into 3x3 matrix, i.e., R=s=3. The $\underline{P} = (P_{ij})$ matrix is given in Table (13.3)

This series has 9 different states -- the number of nonempty cells at matrix $\underline{P} = (P_{ij})$. The f=9 (nine) states are: 1A, 2A, 3A, 1B, 2B, 3B, 1C, 2C, 3C. For example, 1A means to be at category A at position 1, in each period. The appropriate transition matrix M is given below at Table (13.4). For example, 2B means that at position 2 the category is B. The probability of the transition from each category to each one is 1/3.

Table (13.3): The matrix $\underline{P} = (P_{ij})$ for the artificial series Y_3^*.

$\underline{P} = (P_{ij}) =$

	1	2	3	q_i
A	1/3	1/3	1/3	1/3
B	1/3	1/3	1/3	1/3
C	1/3	1/3	1/3	1/3
$\bar{P}_j$	1/3	1/3	1/3	

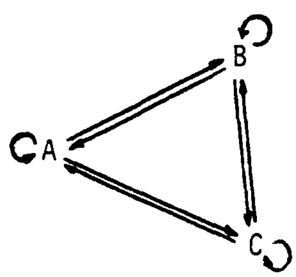
Table (13.4): Transition matrix M of the artificial series Y_3^*.

	1A	2A	3A	1B	2B	3B	1C	2C	3C
1A		1/3			1/3			1/3	
2A			1/3			1/3			1/3
3A	1/3			1/3			1/3		
1B		1/3			1/3			1/3	
2B			1/3			1/3			1/3
3B	1/3			1/3			1/3		
1C		1/3			1/3			1/3	
2C			1/3			1/3			1/3
3C	1/3			1/3			1/3		

The transition matrix M of this completely random series means that the probability to transit from each category to each category is fixed and equal 1/3. Thus, the stationary distribution for the f=9 states is $\underline{\pi}_9 = (1/9,...,1/9)$ and for the R=3 original categories is $\underline{q}_3 = (1/3,1/3,1/3)$.

This completely random series might be presented in the flowchart given in Figure (13.2). Every arrow is an event with probability 1/3.

Figure (13.2): Flowchart of the completely random series Y_3^*.



4. Descriptive measures for qualitative time series

From the definition of Periodic Qualitative Time series above, a qualitative time series is periodic if the conditional distribution at each point in time (within the period) is degenerate and includes only one catyegory. In this situation $\overline{P}_j = 1$ for every j or $\sum_{j=1}^{s} \overline{P}_j = s$. Because $\frac{1}{R} \leq \overline{P}_j \leq 1$ and $\frac{s}{R} \leq \sum_{j=1}^{s} \overline{P}_j \leq s$ a natural measure of Periodicity is $\overline{P}_{(s)} = \sum_{j=1}^{s} \overline{P}_j - \frac{s}{R} / (s-s/R)$ or

(13.2) $P_{(s)} = (R\ \overline{P}-1)/(R-1)$ for $(R \geq 2)$

It is clear that $0 \leq P_{(s)} \leq 1$.

$P_{(s)} = 1$ if and only if the qualitative time series has precisely s as the length of period (the s positions). For example, $P_{(5)} = 1$ for series Y_1^* and $P_{(3)} = 1$ for the series Y_2^*. $P_{(s)} = 0$ if and only if the conditional distribution at every point of time j is uniform. For example $P_{(3)} = 0$ for the series Y_3^*. $P_{(s)} = 1$ and $P_{(s)} = 0$ are the best and worst situations respectively for forecasting observations. All intermediate values $0 < P_{(s)} < 1$ describe different situations closer to or further from periodicity. A necessary conditon for a qualitative time series to be periodic is that $s \geq R$.

A qualitative time series is constant if R=1 or s=1. These two possibilities are equivalent. In mathematical form it means that q=1. We can use the indicator $C = (q-1/R)/(1-1/R)$ because $1/R \leq q \leq 1$ or

(13.3) $C = (R \cdot q-1)/(R-1)$ for $(R \geq 2)$

C=1 if and only if the series is constant. C=0 if and only if the series is in the least constant condition. C=0 is equivalent to $q_i = \frac{1}{R}$ for all i=1,...,R which means that the stationary distribution for the R original categories is uniform as in series Y_2^* and Y_3^*.

A qualitative time series is a heterogeneous series with $\overline{P} = 1$ and q=1/R; thus a natural indicator for such series can be based on the difference $\overline{P}$-q. Since $0 \leq \overline{P}-q \leq 1-1/R$, the coefficient for Heterogeneity is $H_{(s)} = (\overline{P}-q)/(1-1/R)$ or

(13.4) $H_{(s)} = R(P - q)/(R - 1)$. for $(R \geq 2)$

In order to see that $\overline{P}-q \geq 0$ or $\overline{P} \geq q$, write the two expressions for $\overline{P}$ and q:

$$\bar{P} = \frac{1}{s}\sum_{j=1}^{s} \operatorname{Max}_{i} P_{ij} ; \qquad q = \operatorname{Max}_{i} \frac{\sum_{j=1}^{s} P_{ij}}{s}$$

Of course $\sum_{j=1}^{s} \operatorname{Max}_{i} \{P_{ij}\} \geq \operatorname{Max}_{i} \{ \sum_{j=1}^{s} P_{ij} \}$.

$H_{(s)}$ = 1 if and only if the series is heterogeneous, which means two things: (1) All the distributions within the s points of any period are degenerate, namely, there is a different nonzero category for each j; i.e., each column of the matrix $N=(N_{ij})$ has only one non-empty cell.

(2) The non-empty cells are of different categories for each position j (j-1,...,s.

$H_{(s)}$ = 0 if and only if $\bar{P}$=q, which means that the mode values of each position j are all within the same category i. A necessary condition for a series to be heterogeneous is that s=R.

It is easy to verify that

(13.5) $P_{(s)} = C + H_{(s)}$

As $P_{(s)}$ approaches 1, this means that the qualitative time series is more periodic. The size of $P_{(s)}$ is affected by the additivity of constancy (C) and heterogeneity ($H_{(s)}$).

Let us assume that we have a Qualitative Time Series and we want to estimate the length of the period in order to know if the series is really periodic. To do so, we shall try to transform a given empirical series into matrix of Rxs for s=2,...,R,...,$[\frac{N}{2}]$. In each step we compute $P_{(s)}$. The value s that brings $P_{(s)}$ to a peak (local maximum) which is close enough to 1, (the theoretical maximum) is probably the length of the period.

It is obvious that we are interested in the maximum R which fulfills all the above mentioned conditions, because R=1 is a trivial solution. As R increases, the divison of categories becomes finer, and the width of each interval (not necessarily equal) of the original qualitative variables becomes smaller and more useful.

A prediction of a Qualitative Time series of N observations depends as usual on a loss function. For qualitative data, it might be reasonable to choose a method which, if applied to each possibility of states j, will yield the least amount of errors in prediction. Thus in predicting the N + k, k=1,2,... observation (i.e., k units ahead), the amount of error is the proportion of wrong predicted according to the categories of the proposed series, since any method of prediction that we adopt must select a particular category in which to place the N + k obvservation. Hence, if the supposed observation is in position j within the period, and the category with the highest frequency i is selected as the predicted category, then the empirical proportion of error $1-P_j$ will be minimized. Therefore the mode category i(j) is the best predictor for observation N+k that fall into column j using the minimum number of errors as a loss function. Guttman (1941) called this "the principle of maximum probability". If the maximum P_j is found in more than one category, there is no unique method for minimizing the amount of error. The minimum proportion of errors can be seen to still be $1-P_j$. Goodman and Kruskal (1954) adopted this idea for measures of association based on optimal predication and labeled Guttman's coefficients as "λ_a" and "λ_b". See also Bishop, Fienberg and Holland (1975).

Thus for a given point of time N + k we have to compute the appropriate position j within the period. The index j as obtained by the formula

$$j = \begin{cases} k - [k/s]s & \text{when } k > [k/s]s \\ s & \text{when } k = [k/s]s \end{cases}$$

where brackets denote the integer part.

The <u>characterized pattern</u> in a given qualitative time series is thus the vector $(i_{(1)}, i_{(2)}, \ldots, i_{(s)})$ of length s. The s coordinates are the appropriate categories, that is, the caterogies of the modal value in each position. This characterized pattern is the qualitative analogy to the "Seasonal Pattern" used in decompositon method of quantitative time series. The goodness-of-fit of this characterized pattern for a qualitative time series is measured by $P_{(s)}$. $\bar{P}$ is a measure of the mean probability of "true" category prediction at any randomly chosen point in the series. $1-\bar{P}$ is the appropriate amount of error.

To demonstrate our approach we analyze the two following examples:

Example 1. Here we present the first example given by Colwell (1974), dealing with the seasonal pattern of flowering and fruiting in a hypothetical tropical trees species. Say that there are 9yr. L=9 whole cycles of qualitative flowering and fruiting records available for seven individual trees growing in a variety of habitats or regions. Details of scoring and the hypothetical records are given in Table (13.5). The pattern is maximally predictable if the very same seasonal pattern of flowering and fruiting is repeated in all 9 years, as in tree a, b and c in table (13.5). For each of these trees, knowing the season (I, II or III in Table (13.5) tells us with complete reliability the state of the tree. The pattern is designated minimally predictable (tree g) if all phenological states are equally likely in all seasons, so that nothing can be predicted about the state of a tree based on the season.

Colwell discussed the results that were obtained for the various trees in Table (13.5) and computed the corresponding measures. These measures were based on information theory notions and in his article, predictability (P), constancy (C), and contingency (M) are defined. In Table (13.5) we present our measures for the same data.

Table (13.5): The predictability (P), constancy (C), and contingency (M) as defined by Colwell are the same as our periodicity (P), constancy (C), and heterogeneity (M). Values are computed for the seven individual trees. In each matrix, columns represent seasons (I=January-April; II=May-August; III=September-December), and rows represent phenological state (ff-flowering and fruiting; fi=flowering only; nf=no flowering or fruiting).

	I	II	III		Colwell method	our method	Our method with R=2 (only two categories
		Tree a					
ff	9	9	9	P =	1.0	1.0	
fi	0	0	0	C =	1.0	1.0	
nf	0	0	0	M =	0.0	0.0	
		Tree b					
ff	9	0	0	P =	1.0	1.0	
fi	0	0	9	C =	0.0	0.0	
nf	0	9	0	M =	1.0	1.0	
		Tree c					
ff	9	0	0	P =	1.0	1.0	1.0
fi	0	9	9	C =	0.42	0.50	0.33
nf	0	0	0	M =	0.58	0.50	0.67
		Tree d					
ff	1	1	1	P =	0.23	0.50	
fi	2	2	2	C =	0.23	0.50	
nf	6	6	6	N =	0.00	0.00	
		Tree e					
ff	6	0	3	P =	0.61	0.67	
fi	3	0	6	C =	0.00	0.00	
nf	0	9	0	M =	0.61	0.67	
		Tree f					
ff	0	1	4	P =	0.29	0.44	
fi	1	3	3	C =	0.10	0.33	
nf	8	5	2	M =	0.19	0.11	
		Tree g					
ff	3	3	3	P =	0.00	0.00	
fi	3	3	3	C =	0.00	0.00	
nf	3	3	3	M =	0.00	0.00	

Example 2. Here we present the second example given by Colwell (1974), with the climatic patterns, 10 Yr. of monthly precipitations totals for four representative weather stations. The months of the year are thus the time categories (s=12) or, in other words, the period's length. The amount of precipitation was measured on a continuous scale but was transformed to appropriate categories -- the R=12 states. The data are given in Colwell (1974, Table 2: 1151). In Table (13.6) of the present paper, the measures obtained by the two methods are given.

Table (13.6): The measures obtained by Colwell's method and ours for climatic patterns. (The four stations are: Uaupes, Acapulco, Bella Coola, and Miami.)

Station	Colwell's method	Our method	$\bar{p}$
Uaupes	P = 0.75	= 0.64	0.67
	C = 0.66	= 0.48	
	M = 0.09	= 0.16	
Acapulco	P = 0.54	= 0.48	0.53
	C = 0.13	= 0.12	
	M = 0.41	= 0.36	
Bella Coola	P = 0.58	= 0.47	0.52
	C = 0.34	= 0.29	
	M = 0.24	= 0.18	
Miami	P - 0.46	= 0.35	0.40
	C = 0.23	= 0.21	
	M = 0.23	= 0.14	

14: OPTIMAL PARTITION OF THE QUANTITATIVE VARIABLE INTO CATEGORIES

It might sometimes be worthwhile to divide the range of quantitative variables into intervals (not necessarily equal), in order to predict an interval "cover" for k-th units ahead instead of a point estimating (e.g., arithmetic mean or median) plus a confidence interval. One good reason is that the interval might be much narrower than the usual confidence interval based on normal assumption or other non-plausible assumptions. Let us now suggest an optimal division of a continuous variable (Time-Series) into categories. "Optimal" describes a sense of trade-off in both the amount of predictability and the width of the interval.

We pointed out that as $P_{(s)}$ approaches 1, the goodness-of-fit improved. Say that $P_{(s)}$ indicates our ability to predict the suitable category of observation in the future. The expression for $P_{(s)}$ is constructed from the addition of the two components, $H_{(s)}$ and C. $P_{(s)}$ is also a decreasing function of F (the number of different categories), and thus we can always find an $R>R$ so that $P_{(s)} \leq P'_{(s)}$. It is straightforward that we prefer a divison with R as great as possible since it may help us to detect small differences. But while R increases we may also wish to cover the whole range of observations properly, (i.e., q should be as small as possible). Therefore, in this situation we would like to achieve both cases: 1) R and $P_{(s)}$ are as large as possible; 2) q is as small as possible. By this we hope to divide the data into categories such that the Maximum R will give us Maximum $P_{(s)} = C + H_{(s)}$ with minimum C, namely, Maximum $H_{(s)}$ and Minimum C. This situation is similar to finding a division that brings minimum variance "within" and maximum variance "between" groups. To do so, an iterative process is used. In the first step R=2 is chosen with classes defined to produce maximum $H_{(s)}$ and minimum C. In the next step R is increased by 1 to get R=3 and subdivide the data. The process stops at the Maximum R that gives us a sufficiently large $P_{(s)}$ with C as small as possible. This is the desired "optimal" division. The "optimum" means that R cannot be continued to increase without drastically decreasing the value of $P_{(s)}$.

For example, let us look at Table (14.1) below.

Table (14.1): Average of monthly rates of evaporation from class A pan in Eilat mm/day in the years 1960/1067. Source: Gilad (1972).

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1960	5.0	7.2	7.8	10.0	13.4	15.0	15.8	14.8	12.9	10.8	6.7	5.2
1961	4.4	5.8	7.4	10.8	12.7	16.3	15.6	14.5	12.0	9.3	6.6	4.5
1962	4.3	6.0	8.7	9.9	13.7	15.4	14.7	15.0	13.2	8.7	8.2	5.2
1963	5.2	6.9	8.0	9.1	11.4	15.2	15.5	16.5	13.5	10.8	7.8	5.0
1964	4.3	5.6	8.9	10.2	12.9	14.3	14.7	14.0	12.0	10.3	6.6	4.4
1965	3.6	6.2	8.8	9.8	13.8	15.5	15.5	14.6	12.1	9.0	6.2	4.6
1966	4.2	4.7	7.5	10.2	13.1	15.6	14.7	13.3	11.4	8.8	6.2	4.5
1967	4.4	5.1	6.9	9.5	11.3	14.3	14.2	13.1	10.8	7.3	5.5	4.3

For R=2; by division of the amount of average evaporation into two categories: 0-10.8 and 10.81 + the Periodicity measure $P_{(s)}$ = 0.979 is obtained. The division to R=3 by the following categories: 0-5.2, 5.21-10.8, 10.81 + . $P_{(12)}$ = 0.953 is obtained. In the next step, sub-division of the data into R=4 categories $P_{(12)}$ = 0.931 is obtained.

An "optimal" division is given by the division below in Table (14.2) with R=5 categories and $P_{(12)}$ = 0.90.

Table (14.2): The matrix $N = (N_{ij})$ of R=5 categories and s=12 positions (as period's length). The categories are "optimal" division of the amount of average evaporation.

	The "optimal" division of categories	1	2	3	4	5	6	7	8	9	10	11	12	$\sum_{j=1}^{s} N_{ij}/N=q_1$
	0-5.20	8	2	-	-	-	-	-	-	-	-	-	8	18/96
amount of	5.21-7.20	-	6	1	-	-	-	-	-	-	-	6	-	13/96
average	7.21-10.8	-	-	7	8	-	-	-	-	1	8	2	-	26/96
evaporation	10.81-13.9	-	-	-	-	8	-	-	2	7	-	-	-	17/96
26/96	13.91+	-	-	-	-	-	8	8	6	-	-	-	-	22/96
	Total	8	8	8	8	8	8	8	8	8	8	8	8	96/96

There are N=96 observations. The number of whole cycles (years) is L=8. The period's length s = 12. The values of the various descriptive measures are $\overline{P} = 0.92$, q=0.27, $H_{(12)} = 0.81$, C = 0.09, $P_{(12)} = 0.90$.

The division of amount of average evaporation into R=6 categories: 0-5.2, 5.21-7.2, 7.21-9.0, 9.01-11.0, 11.01-14.2, 14.21 +, yields a periodicity measure $P_{(12)} = 0.837$. Thus the "optimum" division is into R=5 categories as presented in Table (14.1). The finer division obtained by R=6 categories decreases the value of $P_{(12)}$, from 0.90 down to 0.84.

The $\underline{P} = (P_{ij})$ matrix obtained for this series is given in Table (14.3) and it has f=17 nonempty cells.

Table (14.3): The $\underline{P} = (P_{ij})$ matrix of the averages of monthly rates of evaporation

R = 5 categories	1	2	3	4	5	6	7	8	9	10	11	12
1	1.0	.25	-	-	-	-	-	-	-	-	-	1.0
2	-	.75	.125	-	-	-	-	-	-	-	.75	-
3	-	-	.875	1.0	-	-	-	-	.125	1.0	.25	-
4	-	-	-	-	1.0	-	-	.25	.875	-	-	-
5	-	-	-	-	-	1.0	1.0	.75	-	-	-	
P_j	1.0	.75	.875	1.0	1.0	1.0	1.0	.75	.875	1.0	7.5	1.0

This series is assumed as a finite ergodic Markov Chain and the transition matrix of order 17 is given in Table (14.4).

Table (14.4): The transition of order f=17 for the Meteorological example. The 17 states are characterized by 2 numbers. The left number is Roman and indicates the different (R=5) categories. The right Arabic number indicates the month, (e.g., the position with the year.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	I1	I2	II2	II3	III3	III4	IV5	V6	V7	V9	V8	III9	IV9	III10	III11	III11	V12
1 I 1		1/4	3/4														
2 I 2				1/2	1/2												
3 II 2					1												
4 II 3						1											
5 III3						1											
6 III4							1										
7 IV 5								1									
8 V 6									1								
9 V 7										1/4	3/4						
10 IV 8												1/2	1/2				
11 V 8													1				
12 III9														1			
13 IV 9														1			
14 III10															3/4	1/4	
15 IIII																	1
16IIIII																	1
17 I 12	1																

The probabilities in this matrix were computed from the original series, (given in Table (14.1) and using the division into R=5 categories presented in Table (14.2). For instance, the probability to transit from state 2 (category 1 in February) into state 4 (category 2 in March) is 1/2. The stationary distribution of the 17 states is $\underline{\pi}$ = 1/96(8,2,6,1,7,8,8,8,8,2,6,1 7,8 6,2,8) which are the probabilities of $\underline{P} = (P_{ij})$ (Table 14.3) divided by L=12 whole periods.

The original categories are unions of these states in the expressions given in Table (14.6)

Each one of the original observations of Table (14.1) was transformed into one of the f=17 different states and the whole series is presented in Table (14.5) below. By computing the Chapman-Kolmogorov equations it seems that the assumption of a finite Markov Chain is quite good.

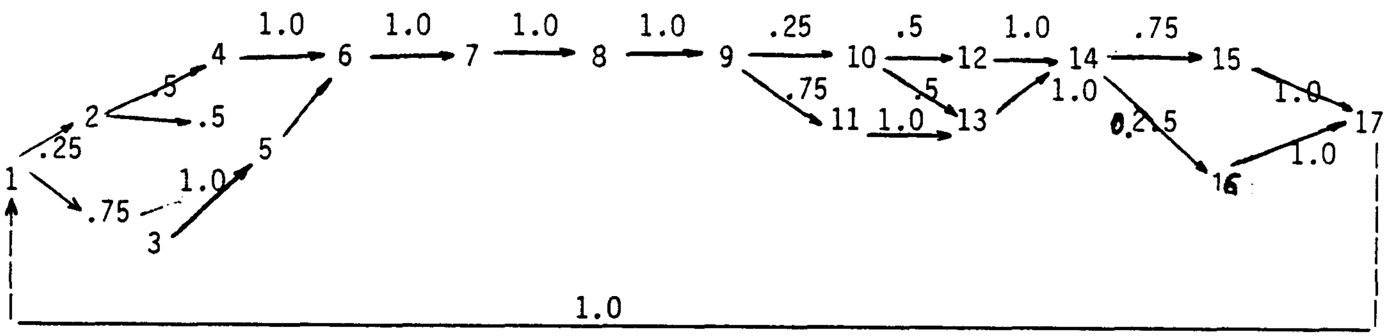
Table (14.5): The original observations after transformation into one of the f=17 different states.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1960	1	3	5	6	7	8	9	11	13	14	15	17
1961	1	3	5	6	7	8	9	11	13	14	15	17
1962	1	3	5	6	7	8	9	11	13	14	15	17
1963	1	3	5	6	7	8	9	11	13	14	16	17
1964	1	3	5	6	7	8	9	11	13	14	16	17
1965	1	3	5	6	7	8	9	11	13	14	15	17
1966	1	2	5	6	7	8	9	10	13	14	15	17
1967	1	2	4	6	7	8	9	10	12	14	15	17

Table (14.6): The decomposition of original categories into states. The stationary distribution of the 5 categories is $\underline{q}_5 = (q_1,q_2,q_3,q_4,q_5) = 1/96$ (18,13,26,17,22). The process may be presented in the flow-chart given in Figure 1 below.

Original category	Union of the Below States	Sum of Probabilities
I (1)	1 + 2 + 17	18/96
II (2)	3 + 4 + 15	13/96
III (3)	5 + 6 + 12 + 14 + 16	26/96
IV (4)	7 + 10 + 13	17/96
V (5)	8 + 9 + 11	22/96

Figure 1: Flowchart of the Meteorological series, the numbers are the 17n states which are given in Table (14.4)



Conclusions

In many kinds of research, data are often of a qualitative (categorical) nature. Even continuous data can be put into categorical form. The Markov Chain process is adopted here to describe a qualitative time series and form a computation of the mean-time of every state -- the stationary distribution. For purposes of description, given time series measures based on modes are defined. These measures are very transparent and are based on easy computations. In this article, Guttman's concepts (Guttman 1941) of the Principle of Maximum Probability is adopted. Indices for goodness-of-fit are measures of the deviations of a given set of empirical data, here, time series from the prespecified definitions. Tools of data analysis are adopted here for finding both the appropriate period's length and to assess three related predefined properties for a given qualitative series: periodicity, constancy, and heterogeneity. A graphic presentation is suggested later in chapter 17 to enable the simultaneous examination of the interrelationships among a given set of series. This might assist a researcher to analyze a set of many series shared by the same life science (precipitation, evaporation, etc.). The process for finding an "optimal" division might be used whenever one wishes to analyze quantitative series by transforming them into categories.

Our approach does not take into account the order of the (whole) periods themselves. This means that the same weight is given to the more distant past as to the recent past. If desired, more weight can be given to the more recent by using only a few last periods, at the expense of decreasing the stability of our measures and predictors. A good tactic may be to use our approach for some subseries (consecutive or not), and to try to study the behavior of the proposed measures. Such an analysis may reveal varying patterns which reflect trends in some or all of the possible states within the period. The method presented here is both very intuitive and based on easy computations. Here concepts of data analysis are adopted, and no classic statistical inference is discussed.

Part 4

SIMILARITY AND DISSIMILARITY OF THE VARIOUS COMPONENTS

The methods discussed in this part are designated to study interrelationships among the various components of series that are obtained usually by one of the many decomposition methods that are in use. In chapter 15, a method of graphic displaying is provided. This method will be exemplified to find common seasonal patterns among series. Chapter 16 deals with a simultaneous analysis of multiple series which are parallel, namely, their slope changes simultaneously over time. Death Rates series will be demonstrated. In Chapter 17, a graphical method is provided to find common patterns of qualitative series like those in part 3.

15. FINDING COMMON SEASONAL PATTERNS AMONG TIME SERIES.

A method for graphically displaying relative distances among a group of seasonal patterns of time series is provided. This chapter is concerned with two aspects:

1. To provide distance measures of different series, or between various periods within the same series.

2. To provide a graphic display technique.

Both goals are not new; nevertheless, the specific application of finding common seasonal patterns among time series is a new one. One particular Multidimensional Scaling (MDS) algorithm was adopted, that is, the nonmetric SSA-1 (Smallest Space Analysis) technique. The group of seasonal patterns is transformed into a symmetric matrix by defining indices of similarity. The symmetric matrix is the input for the graphic display techniques. As a result, a set of objects, either the n various time series or the n periods of a single time series is presented in a map, demonstrating the pairwise interrelationships of the set.

1. Introduction

In the empirical study of time series we usually try to learn about pairwise interrelationships of a given set. Often these series are decomposed into their respective trend, seasonal and irregular, components. This chapter presents a method for comparing time series based on the aspect of seasonality of the proposed series. A series can be decomposed into three components: a trend, a seasonal and an irregular (or random) component where each of these components is empirically identified. For example, a multiplicative monthly time series with constant seasonal pattern of length 12 units of time can be described as follows:

(15.1) $$Y_{aj} = T_{aj} * S_j * I_{aj} \quad a=1,\dots,r; \quad j=1,\dots,12.$$

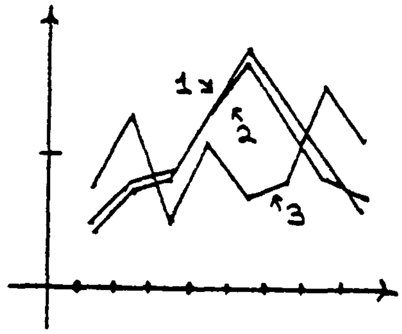
where Y_{aj} is the original observation, T_{aj} is the trend component, I_{aj} the irregular component in the jth month of the ath year, and S_j the seasonal component in the jth month. It is convenient to keep the trend component in the scale of original observations by setting the following constraint:

$$(15.2) \qquad \sum_{j=1}^{p} S_j = p,$$

where p is the period's length and the model is multiplicative. The vector $\underline{S}$ = $(S_1,\ldots,S_p)$ is known in the literature under various names such as "Seasonal Pattern", "Seasonality Index", etc., and is usually expressed in percentages. Extension of the decomposition analysis is made by considering the "seasonal similarity" between time series. That is, the similarity of time series seasonal patterns is investigated once these are decomposed into their respective components. In this chapter given seasonal patterns are treated as vectors for the application of a scaling technique. To do so, the nonmetric SSA-I technique of Guttman (1968) is applied to study the "mutual relationships" of multiplicative seasonal components of time series. The "mutual relationships" are defined more specifically by indices of "distance" or "dissimilarity" between every couple of patterns, i.e., between every constant seasonal pattern of two different time series, or between every two seasonal patterns of a single series which has varying seasonalty.

Given the three seasonal patterns of Figure (15.1), it is seen that the patterns of series 1 and 2 are very similar, whereas that of series 3 is quite different. The proposed method demonstrates this situation by using a graphic display in which the series 1 and 2 are located close to each other and series 3 is far away.

Figure (15.1): Three Seasonal Patterns of Time Series.



To motivate the approach developed here, three examples will be considered. The first two examples present economic monthly time series of "Government income and tax collections" and "Tourist Arrivals in Israel"; the third example demonstrates gradual change in seasonal patterns of the single series "Jewish Marriages" in Israel for the period 1956-1968.

2. Dissimilarity Between Seasonal Patterns

Let us assume n seasonal patterns (vectors of p values) associated with n time series, each with period of length p months, i.e., p=12. Each of these patterns is estimated by one of the available decomposition methods, e.g., X-11 program or LPTA. Let $\underline{S}_i = (S_i,\ldots,S_p)$ be a vector of seasonality patterns of the ith series, i=1,...,n. Each one of the p coordinates corresponds to a jth unit of time within the entire period. For the multiplicative model, we have by definition (expressed in percentages):

$$(15.3) \qquad \sum_{k=1}^{p} S_{ik} = 100 \cdot p \quad (i=1,\ldots,n).$$

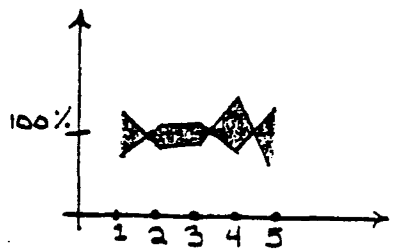
Denote by D_{ij} a dissimilarity coefficient between any two patterns, $\underline{S}_i$ and $\underline{S}_j$. This dissimilarity may be defined in several ways. For example, say dissimilarity is defined as the average absolute distance difference between the two seasonal vectors, i.e.,

$$(15.4) \qquad D_{ij} = \sum_{k=1}^{p} |S_{ik} - S_{jk}|/p \quad (i,j=1,2,\ldots,n).$$

Clearly, (15.4) defined a symmetric matrix ($D_{ij} = D_{ji}$), and therefore the mutual relationship of seasonal patterns is represented by an nxn symmetric matrix with zero elements in the diagonal. Other measures having the symmetric property and expressing dissimilarity may be used as well.

Another suggestion for a dissimilarity coefficient between any two patterns $\underline{S}_j$ and $\underline{S}_j$ is the area B_{ij} between them, e.g., B_{12} is measured by the shaded area in the following Figure (15.2).

Figure (15.2): Two seasonal patterns $\underline{S}_1$, $\underline{S}_2$ of t units and the measure of dissimilarity between them, B_{12} based on area.



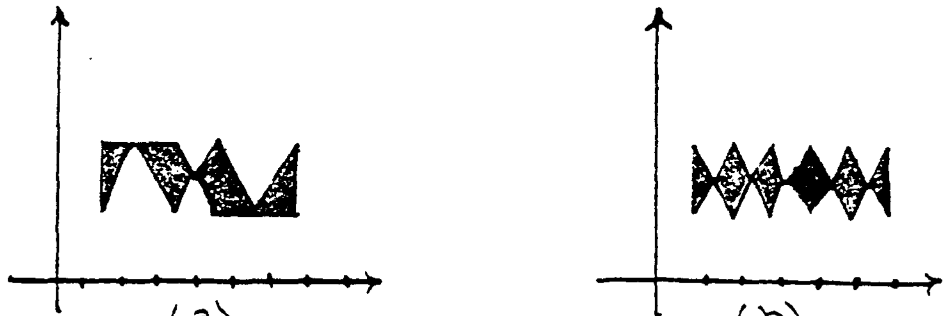
A further approach, that of "Minimum Possible Dissimilarity" is defined by

$$(15.5) \qquad D'_{ij} = \underset{0<\ell<p-1}{\text{Min}} \; \underset{k=1}{} \{ \sum_{i=1}^{p} |S_{ik} - S_{jk'}|)/p \} ,$$

where k' = (k+ℓ) mod p. For example, for ℓ for which the function obtains its minimum is the desired lag.

In other words, D'_{ij} is the minimum absolute mean differences between the two seasonal indices of series i and j where these differences are computed for the 12 (0,...,11) various lags. This approach can be adapted to B'_{ij} which minimized the area between the graphs obtained by two seasonal patterns $\underline{S}_i$ and $\underline{S}_j$. The use of this approach discriminates between the following two cases that are graphed in Figured (15. 3). In both graphs (a) and (b) the coefficients of dissimilarity B_{12} (i=1, j=2) are equal. $B_{12} = 0$ for graph (b) with the desired lag $\ell = 1$. According to the approach of "minimum Possible Dissimilarity". Similarly, $B'_{12} > 0$ for graph (a) for $\ell = 0,\ldots,11$.

Figure (15.3): Seasonal Patterns: S_1 is denoted by ——, S_2 ----
The shaded area is the coefficient of dissimilarity B_{12}.



(a) (b)

These measures were used as examples of possible indices for characterizing the dissimilarity between the seasonal patterns. In order to represent the n(n-1)/2 dissimilarity coefficients D_{ij} (or B_{ij}) simultaneously in a graphic display, the SSA-I technique is used.

3. The SSA-I Technique: A Brief Review

The family of methods "Smallest Space Analysis" or briefly, SSA techniques, were developed and constructed by Guttman (1968), see also Lingoes (1973). Here the use of the first member of the family SSA-I is demonstrated, since it is particularly suitable for data analysis of symmetric matrices.

SSA-I provides a graphic presentation of pairwise interrelationships of a set of objects, here, n times series. Each series i (here, a seasonal pattern of the i series) i=1,...,n is represented as a point in a space and the SSA-I technique seeks the Euclidean space with minimum dimensions that can monotonely reproduce the original dissimilarities D_{ij}. That is, the dimensionality m is sought to be as small as possible for which the Euclidean metric will satisfy the monotonicity condition below:

(15.6) $d_{ij} < d_{k\ell}$ if $D_{ij} < D_{k\ell}$ for every i,j,k,ℓ

D_{ij}: The dissimilarity coefficient (input) between series i and j.

d_{ij}: The Euclidean distance (output) in the graphic map between the points i and j.

The matrix of dissimilarities expresses interrelationships between items, while the computer program enables the analysis of these interrelationships in a simple yet comprehensive manner. The main part of the SSA-I output is in the space diagram, in which each variable is represented by a point. If the dissimilarity coefficient between i and j is larger than the dissimilarity coefficient between k and ℓ, then the Euclidean distance bweween i and j is larger than between i and ℓ .

The goodness-of-fit of an output space to the input dissimilarities is assesssed by a coefficient of alienation. This coefficient is defined by $\theta = \sqrt{1 - \mu^2}$ and varies between 0 and 1 where μ is a coefficient of monotonicity of (D_{ij}, d_{ij}) and $(i,j=1,\ldots,n)$, and is defined by

$$(15.7) \qquad \mu = \frac{\sum_{i,j=1}^{n} D^{*}_{ij} \cdot d_{ij}}{\sqrt{\sum_{i,j,=1}^{n} D^{*2}_{ij} \cdot \sum_{i,j=1}^{n} d^{2}_{ij}}}$$

The D^{*}_{ij} are the rank images of the d_{ij}, that is the d_{ij} rearranged in the rank order of D_{ij}, allowing the untying of ties. Thus, the graphic presentation, namely, the Euclidean space that obtained as an output by SSA-I technique is invariant under monotone transformation of the elements in the input matrix. A perfect fit is represented by $\theta = 0$, and the worst possible fit is given by $\theta = 1$. Intermediate values of the coefficient represent intermediate degrees of goodness-of-fit. How small should the coefficient of alienation be for the fit to be satisfactory? This is a question to which there is no absolute answer, see Guttman (1977, p.89). As a rule-of-thumb, a coefficient of alienation of less than 0.15 is considered a good candidate for being "satisfactory"; a more complete answer requires fit to theory. A value of $\theta = 0.15$ is equivalent of $\mu = 0.99$ which often seems quite high. It is obvious that for a given matrix the coefficient decreases as m, the dimensionality of the space, increases. It is always possible to represent n points in (n-1) dimensionality space keeping conditions (15.6) and thus $\theta = 0.0$. A configuration of N points in m dimensional space when m<<n may prove to be a useful tool in data analysis. In the real examples below, two (m=2) and one (m=1) dimensions for the space diagram were sufficient to represent the data while giving small enough values of θ.

4. An Arificial Example

Let us take the two time series below:

(i) $Y_{1t} = \sin \frac{2\pi}{p} t$

(ii) $Y_{2t} = \cos \frac{2\pi}{p} t$ $\qquad (t=1,2,...)$

These two series might be decomposed into the three components. The trend and the irregular components equal $T_t = I_t = 1$. This means that the trend has no slope and there is no randomness. $\underline{S_1} = (\sin \frac{2\pi}{p}, ..., \sin 2\pi$ for given p. For instance, when p=4

$$\underline{S_1} = (\sin \frac{\pi}{2}, \sin \pi, \sin \frac{3\pi}{2}, \sin 2\pi) = (1,0,-1,0)$$

$$\underline{S_2} = (\cos \frac{\pi}{2}, \cos \pi, \cos \frac{3\pi}{2}, \cos 2\pi) = (0,-1,0,1)$$

$$D_{12} = \sum_{k=1}^{4} |S_{1k} - S_{2k}|)/4 = 1.0$$

The coefficient of "Minimum Possible Dissimilarity" vanishes, i.e., $D_{12}^{\ell} = 0$, when $\ell = 3$ is the desired lag.

5. The Method

Given a set of n time series points, the proposed method is described by a sequence of steps:

step 1: Decompose each series into its components using one of the decomposition methods mentioned above in section 1. Thus from each series $[Y_{aj}]$ a=1,...,r j=1,...,p, $\underline{S} = (S_1,...,S_p)$ is obtained.

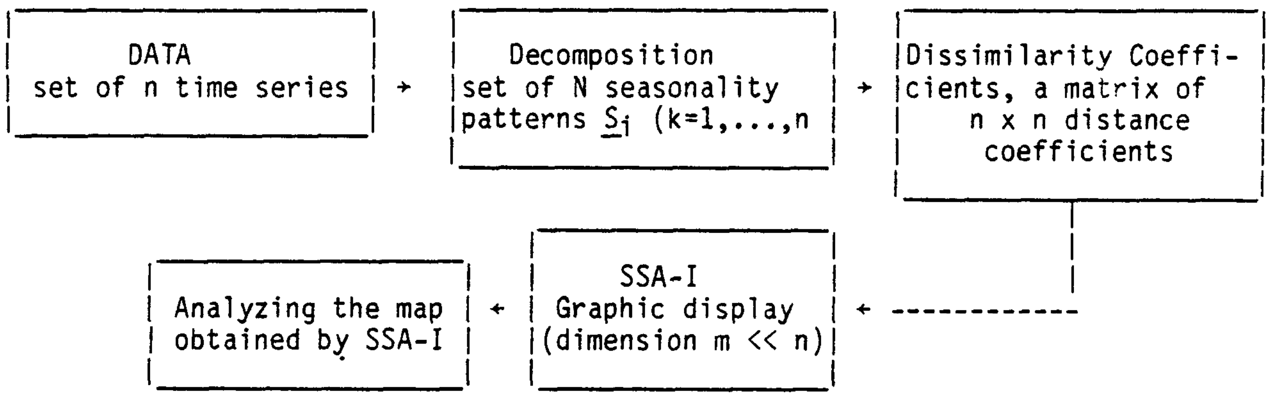
step 2: For each pair of series a coefficient of dissimilarity as mentioned in section 2 is computed. Thus from n seasonal patterns $\underline{S}_i$ i=1,...,n n(n-1)/2 coefficients are obtained.

step 3: The nxn symmetric matrix of coefficients is used as the input to the nonmetrics tehcnique SSA-I. As a result a graphic display is obtained. This presentation is demonstrated in m-dimension space diagram.

step 4: To analyze the output of the SSA-I, it might be possible that characterization of the various seasonal patterns would be revealed within a low m dimensionality and quite low θ also. Thus common seasonal patterns among time series might be found.

These steps are presented in a flowchart in figure (15.4).

Figure (15.4): Flowchart of the various steps of the proposed method.



6. Applications

Below, three examples are considered which demonstrate both the usefulness and applicability of the method. For convenience, the first two are primarily an economic example for which data are available and where the relationship between the seasonality of time series is important. The second, a demographic example, demonstrates a very specific structure of matrix called simplex. These three real examples are described by reference to the steps in the previous section.

Example 1. Seasonality of government income tax collections

We consider n=13 time series of (Israeli) government income from taxes and compulsory payments. As pointed out earlier, the series are monthly time series, each with period's length p=12. The series are labelled and defined below: (1) taxes and compulsory payments-total; (2) income tax-total; (3) income tax from employees; (4) income tax from self-employed; (5) advance income tax payments from self-employed; (6) income tax from companies; (7) income tax advance payments from companies; (8) income from customs and excise; (9) income from customs; (10) purchase tax-total; (11) purchase tax on local production; (12) purchase tax on imports; (13) foreign travel tax.

These are monthly time series for the period 1962-1969, and are available in Bar On (1973). In table (15.1) we represent the 13 seasonality indices estimated for 1968. The analysis was performed by the X-11 method, multiplicative model [see Shiskin et al. (1967)], and it provides moving seasonality factors.

Next we apply the SSA-I in three steps:

(1) The 78 distances are computed using eq.(15.4) with n=13 series, p=12 period's length. The n(n-1)/2 distances are the inputs to the symmetric matrix of SSA-I. We apply the technique for m=2 dimensions (which is found to be satisfactory) and obtain the space diagram presented in fig. (15.5). Here we note a clear discrimination for 3 regions of similar patterns of seasonality. The first region A, consists of foreign travel tax. This series has a specific pattern of seasonality and is very different than the other 12 series. There is a high peak in summer months July and August, concurring with Israelis traveling abroad. In the winter months, there is a sharp trough, expressed in low seasonality indices.

The second region B, contains two series 6 and 7. The seasonality indices of these series have 4 peaks in March, June, September, December. This occurs since payments are made quarterly.

Finally, the region C contains the other 10 series. This region is characterized by a peak in March and a trough in April, except for series 5 that has a very sharp trough in April. Series 5 is in some ways similar to the other 9 series but is slightly different. In the next step this difference will be amplified.

The members in the map indicate the geometric location of the series respectively. The center (arithmetic average) is represented by the cross.
(2) After accepting a clear discrimination for 3 regions, we amplify the study of region C by running 10 corrresponding series again in SSA-I. The resultant, more detailed map is given in fig. (15.7). We note that series 5 is separated from the other 9 because of her different characteristic 2 peaks in January and November.
3) We pursue the SSA-I analysis of region C-5 of fig. (15.6). The resulting map is given in fig. (15.7) where we distinguish between 3 regions of 3 strips. In the middle strip (B) the series are: taxes and compulsory payments - total and purchase tax on local production. The first series (1) is an 'average' seasonality of the other series seasonalities.

Series of the same kind belong to the same strip. The customs series, for example, are on the left strip (C), where payments arise on taxes from foreign products. Series 10 is on the 'border' of strips B and C because it has payments from both local production and foreign imports.

'Income' series are on the right side of the map in fig. (15.7) (strip A) and series 2 is between her 2 sub-series, 3 and 4. The location of series 10 is also between her 2 sub-series, 11 and 12.

The coefficients of alienation in these three maps (figs. (15.5), (15.6) and (15.7)) were small enough to warrant our stating that there is a high goodness-of-fit, meaning that reproduction of the original matrix by SSA-I has not distorted the data.

Table 15.1: Seasonal indices of the 13 times series on government income and taxes (multiplicative model).

Series	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	103.0	94.0	121.3	84.9	89.9	103.6	99.6	96.9	102.3	96.0	102.0	106.7
2	101.8	94.1	136.6	72.4	86.2	107.6	96.2	93.4	108.1	95.4	97.7	110.7
3	98.9	95.2	133.9	85.1	94.3	101.2	98.8	98.9	97.0	99.9	100.7	96.1
4	114.0	107.0	120.6	51.3	85.4	97.8	100.9	100.7	97.0	99.7	111.2	114.7
5	121.0	112.6	94.4	2.9	76.3	106.4	114.4	110.5	106.6	115.4	121.8	116.2
6	80.8	67.3	173.4	54.0	61.3	157.9	81.8	70.8	150.8	83.8	75.9	142.7
7	77.3	60.6	222.5	16.9	33.9	152.0	81.1	54.2	169.5	79.8	74.8	176.5
8	103.3	92.4	108.1	96.7	95.9	98.5	102.2	101.3	96.8	97.7	104.7	102.4
9	109.3	93.1	108.4	99.0	101.2	89.3	91.9	95.8	97.6	97.3	111.4	105.4
10	110.1	98.0	118.0	91.4	92.3	92.0	94.9	92.8	94.8	103.3	107.0	107.5
11	107.8	99.5	131.9	80.4	85.7	95.0	96.5	101.7	97.6	100.8	100.2	103.2
12	108.5	98.1	109.9	105.2	96.4	95.1	91.9	87.2	87.2	107.5	111.4	103.8
13	47.6	50.7	79.0	66.6	99.9	158.8	213.7	181.5	115.8	72.6	61.9	52.5

Figure (15.5): The space diagram in m=2 dimensions for matrix of distance, θ = 0.01.

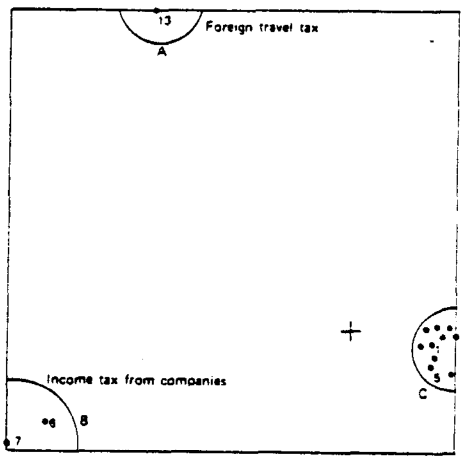


Figure (15.6): The space diagram of region C with coefficient of alienation, θ = 0.04

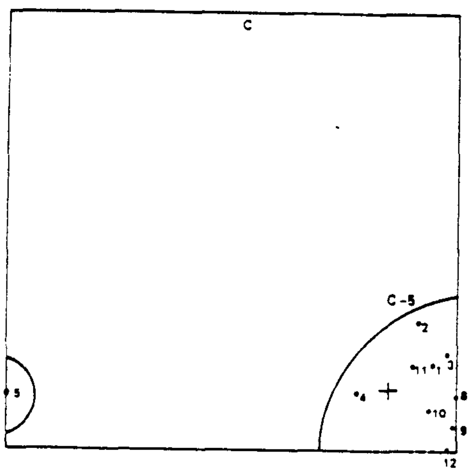
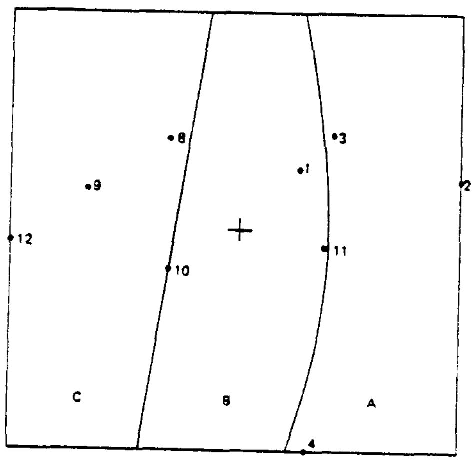


Figure (15.7): Space diagram C-5, θ = 0.08



Example 2: Tourist Arrivals in Israel

In this example sixteen series of "Tourist Arrivals in Israel" from 1974 to 1976 are presented. The series 1 through 12 are tourist arrivals by air from 12 countries or groups of countries. Series 13 is a Tourist arrivals total. Series 14 and 15 are tourists arriving by land and by air, respectively. Series 16 is Tourists arriving from countries other than those included in series 1-12. In step 1, sixteen constant seasonality patterns were computed by LPTA technique using a multiplicative model. The seasonal patterns are given in Table (15.3).

In step 2 the distance between each pair of seasonal patterns is measured by the coefficient of dissimilarity defined in equation (15.5). The period's length is p=12. Thus (16 x 15)/2 = 120 coefficients obtained and given in Table (15.3). The desired lag, that is, those lags where the D^{ℓ}_{ij} obtains its minimum are given within parentheses, just above the D^{ℓ}_{ij}.

In step 3, the 16 x 16 matrix of distances is used as the input to SSA-I. Applying the SSA-I technique for m=2 dimensions, the resultant graphic display is presented in Figure (15.8). The points in the map indicate the geometric location of the series. Each point presents a time series, or more accurately, its seasonal pattern. Two pairs of points are closer on the map than other pairs of points as their similarity is greater. The coefficient of alienation θ = 0.16 is small enough to indicate that there is a high goodness-of-fit, meaning that reproduction of the original matrix by SSA-I has not distorted the data.

Table (15.2): Seasonal Indices of the 16 Time Series on Tourist Arrivals to Israel (Multiplicative Model).

Tourists' Arrivals From	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1. U.S.A.	98	130	139	110	101	112	145^{p}	73	63	106	63	58
2. England	57	60	121	177^{p}	114	84	149	124	101	90	48	73
3. France	42	78	133	85	77	86	226	252	63	53	40	65
4. Scandinavia	67	98	175	176^{p}	95	100	82	59	65	103	76	104
5. Canada	52^{t}	108	126	127	141	102	142	78	82	109	55	78
6. Germany and Austria	39$_{t}$	69	183	214^{p}	77	64	97	89	91	137	57	82
7. Switzerland	47^{t}	84	131	208^{p}	100	54	133	57	99	150	56	80
8. Holland	43^{t}	71	117	164^{p}	111	97	175	81	110	115	55	62
9. Italy	49^{t}	75	120	144	86	72	100	146	167	79	57	105
10. S. America	101^{t}	46^{p}	52	127	93	76	109	94	145	136	74	146^{p}
11. Oceania	116	71	100	126	110	92	94	89	81	83	68	168
12. S. Africa	171^{p}	84	85	135	105	93	74	95	89	79	54	134
13. Total	56^{t}	73	116	134	93	97	153	122	90	109	69	87
14. By Land	54^{t}	58	117	112	89	118	174	128	87	123	68	71
15. By Air	57^{t}	79	121	137	95	94	151	118	87	106	67	88
16. Other Countries	50^{t}	64	118	174	87	89	160	135	96	78	61	89

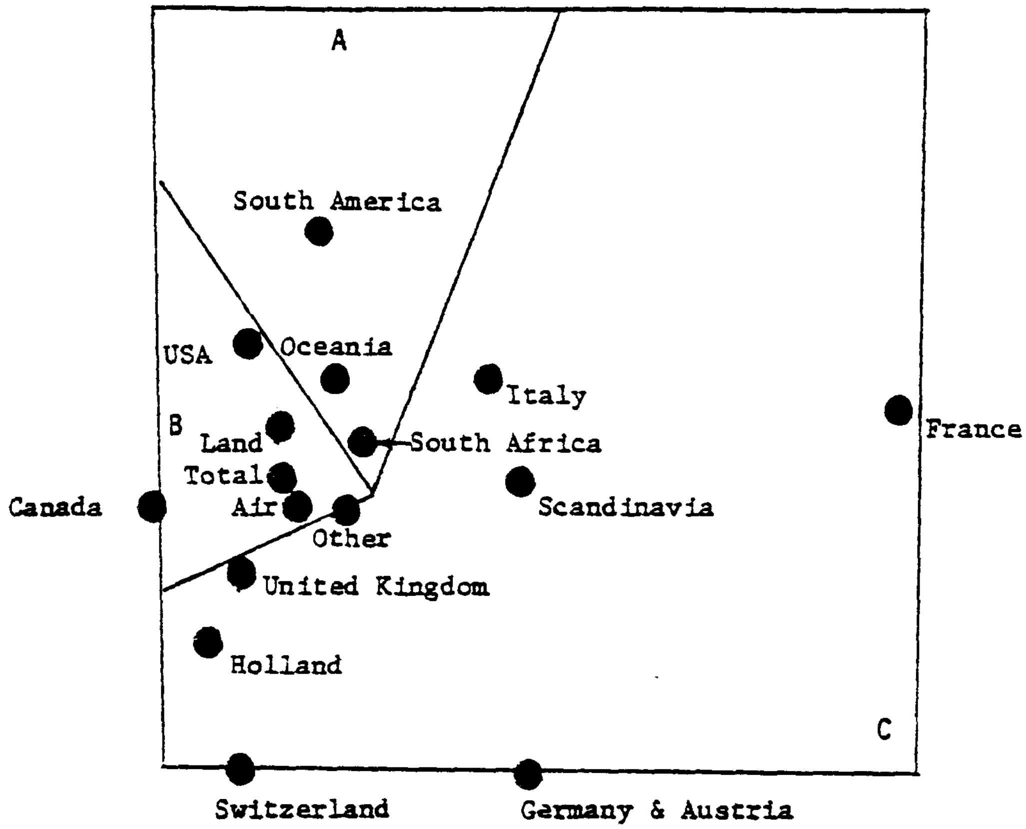
^{p} indicates the peak of the seasonality.

^{t} indicates the trough of the seasonality

Table (15.3): The distance Matrix Among 16 Seasonal Patterns of the Time Series as Computed from Table (15.2) using formula (15.5) with p=12. The values in parentheses are the suitable lags in which the minimum is obtained.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1.	0															
2.	(11) 24	0														
3.	(11) 37	(0) 39	0													
4.	(0) 27	(1) 31	(4) 33	0												
5.	(0) 17	(0) 21	(0) 44	(0) 27	0											
6.	(11) 40	(0) 29	(4) 36	(0) 25	(0) 33	0										
7.	(0) 33	(0) 23	(4) 43	(0) 28	(0) 24	(0) 17	0									
8.	(0) 27	(0) 15	(0) 41	(9) 28	(0) 18	(0) 28	(0) 22	0								
9.	(10) 19	(0) 24	(11) 31	(7) 21	(10) 29	(7) 27	(7) 26	(0) 30	0							
10.	(2) 23	(3) 16	(7) 41	(2) 25	(2) 26	(3) 28	(3) 28	(3) 25	(8) 26	0						
11.	(6) 24	(7) 31	(10) 44	(6) 25	(7) 26	(6) 36	(6) 36	(9) 32	(11) 22	(4) 26	0					
12.	(10) 24	(4) 21	(3) 43	(3) 25	(3) 26	(3) 27	(4) 33	(4) 29	(4) 24	(0) 15	(0) 24	0				
13.	(11) 21	(0) 14	(0) 37	(8) 22	(0) 16	(0) 29	(0) 26	(0) 16	(0) 23	(9) 18	(2) 25	(9) 17	0			
14.	(0) 25	(0) 19	(0) 35	(8) 21	(0) 22	(0) 35	(9) 30	(0) 18	(0) 31	(6) 24	(2) 25	(9) 22	(0) 10	0		
15.	(11) 20	(0) 13	(0) 37	(8) 24	(0) 15	(0) 29	(0) 25	(0) 17	(0) 22	(9) 18	(2) 26	(9) 18	(0) 3	(0) 13	0	
16.	(11) 24	(0) 10	(0) 35	(8) 27	(0) 25	(0) 28	(0) 26	(0) 17	(0) 19	(9) 13	(1) 29	(8) 24	(0) 11	(0) 17	(0) 12	0

Figure (15.8): The space Diagram in m=2, θ = 0.16



In the map we note the following results: The series 13 and 15, "Tourists Arriving" totals and "By Air" totals have very similar patterns of seasonality. The series 15 contains 88% of the total tourists. The location of these series is approximately in the center, and in fact represents the "average" weighted seasonality of the other series. "Tourists Arriving From France" (3) has quite a different seasonality than that of the other series and is located further from other countries. It is easy to distinguish similar series in the same sectors. The series of "Tourists Arriving From the Southern Hemispheres" are located in sector A. "Tourists Arriving from North America" are located in sector B. In sector c we find "Tourists Arriving fom Europe and other countries."

Example 3: Jewish Marriages -- A Demographic Example

In Table (15.4) below, moving seasonal patterns of "Jewish Marriages" in Israel in 1956-1968 obtained by X-11 are presented for each year. The matrix of distance $D_{ij}^{\prime}$ is given in Table (15.5). All the suitable lags are equal to 0.

Table (15.4): Seasonal Indices of the 13 Years of the Series "Marriages of Jews" in Israel (1956-1968)

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1956	90.4	86.4	134.1	88.4	80.3	111.9	81.2	142.6	109.1	92.0	86.1	97.8
1957	90.2	85.8	134.1	88.0	80.0	122.6	80.1	144.9	110.1	90.8	86.6	97.5
1958	90.0	84.4	134.6	86.7	79.8	114.2	77.8	148.8	111.5	88.7	88.2	96.9
1959	88.9	83.0	135.6	85.1	78.9	115.7	74.8	154.4	112.4	87.1	89.8	95.4
1960	88.5	81.8	135.6	83.0	77.8	117.5	71.8	160.5	114.2	85.9	90.4	94.0
1961	88.0	81.1	136.7	81.1	74.5	118.9	70.4	165.9	115.2	86.5	90.4	92.5
1962	88.0	80.5	136.9	79.0	71.5	120.0	71.3	167.8	117.7	86.5	90.6	92.1
1963	87.1	80.3	136.1	77.7	68.4	121.8	73.7	167.3	119.3	86.9	91.0	93.1
1964	86.1	79.8	133.3	76.8	67.5	123.6	76.2	164.4	123.2	86.1	91.3	94.2
1965	85.1	79.1	130.0	77.5	67.0	124.7	78.0	162.1	125.4	85.5	91.8	94.9
1966	84.1	78.2	126.7	78.4	68.4	124.7	79.0	159.6	130.6	84.1	92.8	95.1
1967	83.3	77.8	123.3	79.9	69.9	124.6	78.8	159.0	132.4	83.7	93.0	95.0
1968	83.1	77.7	121.1	80.5	71.0	125.2	78.2	158.8	132.8	83.7	93.0	94.0

In this series, there is a gradual change from year to year in the varying patterns of seasonality. This implies that the distance betweeen any two patterns is smallest with neighboring yearly patterns. Thus, the distances near the diagonal are the smallest and the distance in the upper right and lower left corners are greatest. Therefore, a one dimensional structure called a simplex is accepted as a representation of the space diagram, see Guttman (1950, 1954) and Anderson (1959).

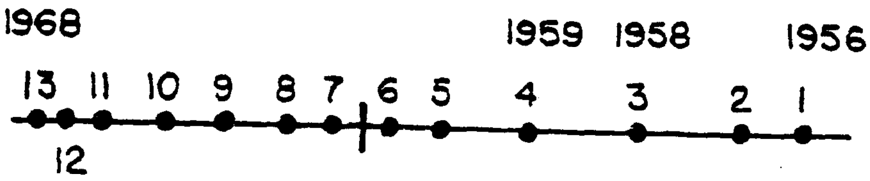
Table (15.5): The Distance Matrix of the 13 Years as Computed From Table (15.4)

COEFFICIENTS MATRIX

	1	2	3	4	5	6	7	8	9	10	11	12	13
1.	0.00												
2.	0.72	0.00											
3.	2.14	1.43	0.00										
4.	3.95	3.23	1.81	0.00									
5.	5.67	4.96	3.53	1.72	0.00								
6.	7.14	8.42	5.00	3.19	1.57	0.00							
7.	8.07	7.35	5.92	4.12	2.48	1.08	0.00						
8.	8.42	7.70	6.28	4.47	3.23	2.09	1.20	0.00					
9.	8.62	7.90	6.56	5.15	3.99	3.41	2.75	1.62	0.00				
10.	9.07	8.35	7.04	5.90	4.83	4.66	4.00	2.87	1.37	0.00			
11.	9.55	8.83	7.69	6.55	5.66	5.76	5.10	4.08	2.77	1.48	0.00		
12.	9.85	9.13	7.96	6.82	6.01	6.11	5.60	4.93	3.62	2.35	0.92	0.00	
13.	10.08	9.36	8.08	6.94	6.08	6.18	5.77	5.29	3.97	2.76	1.44	0.54	0.00

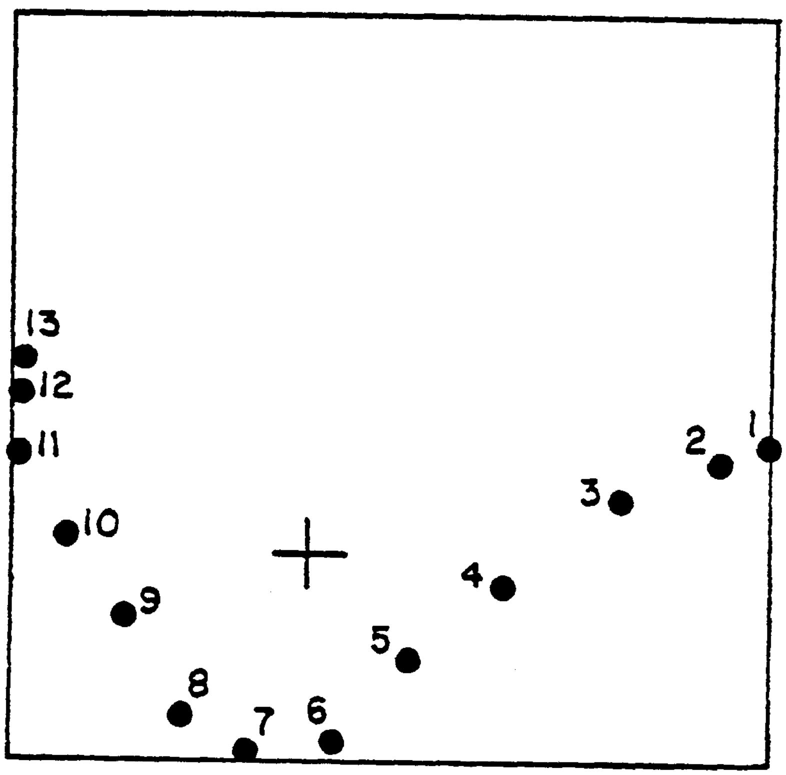
A simplex structure matrix of distances has the property that its elements incline away from the diagonal; often changes of rows or columns are needed. By transforming such a matrix into a map, applying the SSA-I technique and using one dimension (M=1), a (nearly) perfect order structure is obtained as in figure (15.9). The coefficient of goodness-of-fit is $\theta = 0.09$, indicating a very good fit, i.e., there are only small deviations of the original matrix from a perfect simplex. The center (Arithmetic-average) is presented by the cross.

Figure (15.9): The Space Diagram in m=1 dimension, θ = 0.09.



The matrix of distances given in Table (15.5) presents another interesting property of a simplex structure which was revealed by Guttman (1950, p.323) or Guttman (1954, pp.319-324) and is presented in Figure (15.10) using m=2 dimensions. This property means that the second component is a U-shaped function of rank order, i.e., the first component. The bending point is located on the arithmetic-average of the first component denoted by the cross. The coefficient of goodness-of-fit is θ = 0.01, indicating a very good fit.

Figure (15.9): The Space Diagram in m=2 dimensions, θ = 0.01



In this example, the internal order of the points is determined in terms of time. The most different seasonal patterns are those for the years 1956 and 1968. The most similar years in that sense of seasonality are those of 1967 and 1968.

7. Discussion

A method of graphical display of the pairwise interrelationships of a set of seasonal patterns of time series is provided. There are many possibilities of an empirical study of a given set of time series. The proposed method concentrates on the simultaneous comparison of one aspect of each series, that of the seasonality. The method suggested here is based on a sequence of four steps: decomposing each time series; computing a coefficient of dissimilarity between every pair of seasonal patterns; using a graphic technique for the matrix of distances; analyzing and interpreting the map obtained in the previous step.

None of the four above steps is new; however, the application specifically to time series is new. The method makes it possible to find common seasonal patterns among time series. Either n various time series as in example 1 and 2 or the n periods of a single time series as in example 3 could be used for raw material. Other coefficients for dissimilarity among other components, such as trend and irregularity with suitable modification, might be defined as in the second step as well. Thus the method can be generalized to other aspects of comparison in given sets of time series.

The SSA-I technique is invariant under monotone transformations of the elements of the input matrix that are computed in step 2. Thus the same graphic display would be obtained for any monotone transformation of D_{ij} or D'_{ij}, etc.

16. COMMON TREND OF MULTIPLE SERIES

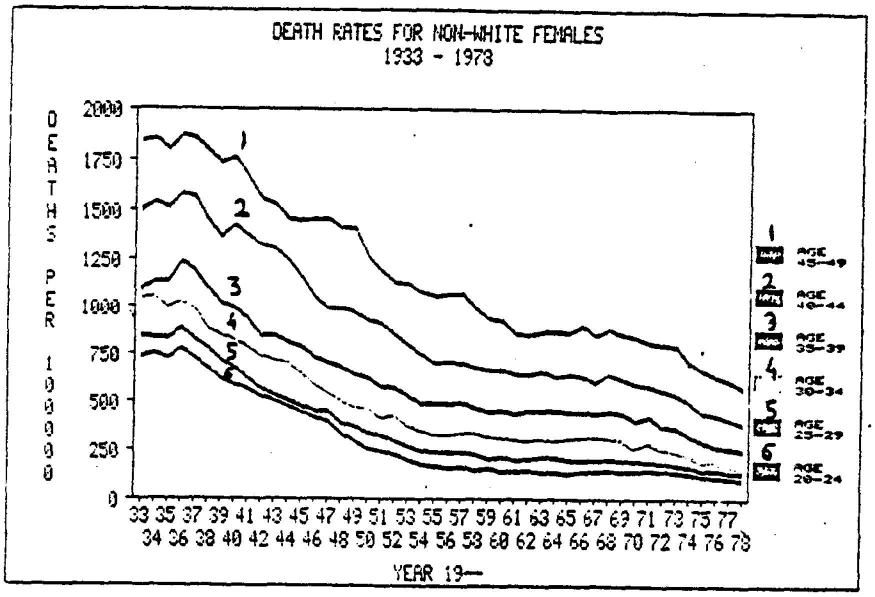
This chapter deals with a simultaneous analysis of multiple series by means of univariate Seasonal Adjustment and Prediction Approaches. These series are approximately 'parallel', namely, their slope changes simultaneously over time. Series such as these can be unified into a single series and treated as it has trend and seasonal components. Our proposal is one way of reducing the number of parameters needed to describe the trend and fluctuations of deaths. The analysis is similar to Factor Analysis in the sense that the main and common factor is the trend component which has negative and monotone slope. The series underlying this study is Death Rates for Non-White Females. In our point of view, the series is "trend and error" component.

INTRODUCTION

In this chapter there is an attempt to analyze a set of time series which have some properties in common as a single unified series. In order to study sets of time series, tools for analyzing unidimensional time series are adopted. For similar purposes, Ledermann and Breas (1959) used factor analysis technique. The main idea is to reduce the number of parameters needed to describe the trend and fluctuations of deaths. A definition for parallel series is given as well as a data analysis tool designated to measure deviations of empirical series from being parallel.

The series presented in this paper are Death Rates for Non-White Females in the years 1933-1978 for 6 age groups. The numerical values are deaths per 10000. The age groups are: 20-24, 25-29, 30-34, 35-39, 40-44, and 45-49. The original data is given in Figure (16.1), below.

Figure (16.1): Death Rates for Non-White Females 1933-1978 for 6 age groups



The six Death Rates series are unified in a special way in order to get one series as is obtained in Figure (16.2). The unified series is decomposed into its three main components: Trend, Seasonality and Irregularity. The period's length p=6, the Seasonal Pattern seems nearly fixed using a purely multiplicative model. The trend seems to be monotone with negative slope. In the next section, the formal definition for parallel series is presented, as well as analysis of a specific example,--Death Rates of Non-white Females.

Parallel Series - a definition and an example.

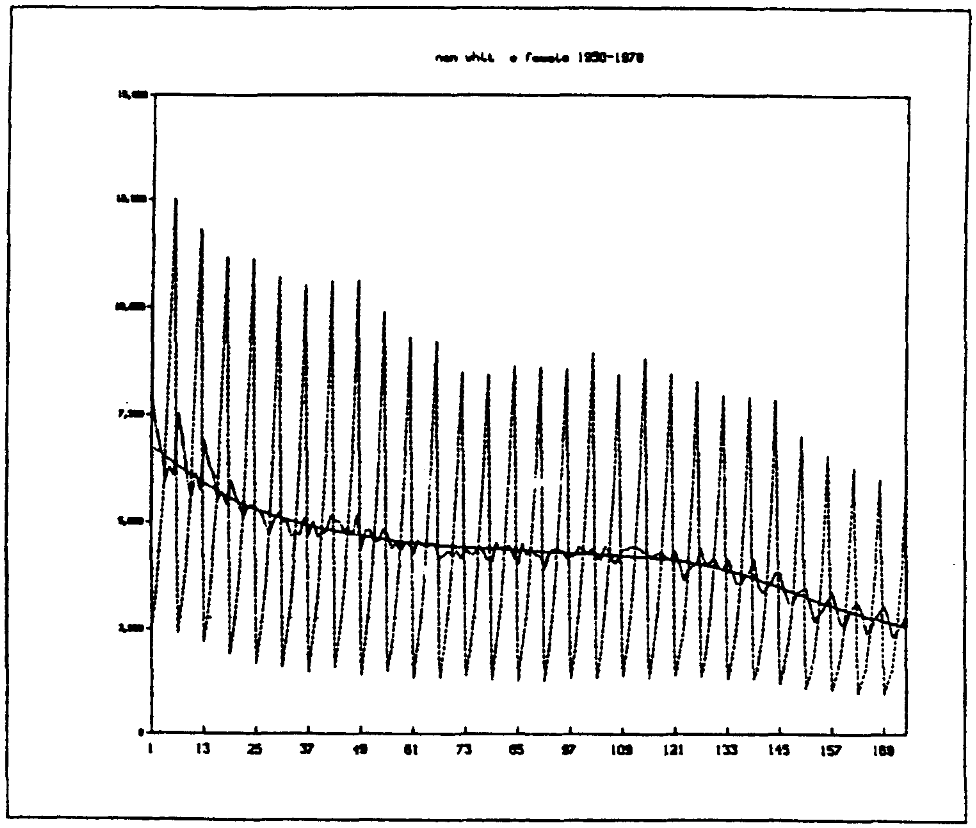
Let Y_{tj} $t=1,\ldots,n$ $j=1,\ldots,p$ be a set of p numerical (quantitative) time-series each having n observation. Series k and l are said to be Additively parallel if $Y_{tk} = Y_{tl} + \Delta_{kl}$ for all $t=1,\ldots,n$. Where the unknown parameter Δ_{kl} is not a function of the t index. Let us denote these two series by $Y_{.k} \underset{a}{||} Y_{.l}$.

Series k and l are said to be multiplicatively parallel if $Y_{tk} = \Delta_{Kl} \cdot T_{tl}$ where $\Delta_{kl}>0$ is independent of t. Let us denote these two series by $Y_{.k} \underset{m}{||} Y_{.l}$. Of course a log transformation of multiplicatively parallel series yields additively parallel series.

If a set of series is parallel it indicates at least common shapes of a trend. In other words, they have the same slope which is usually changing over time. The various series have a fixed relation to an unobservable common trend component. Their quotient or difference is fixed depending on whether the model is purely multiplicative or purely additive. If seasonal components exist they should have the same pattern either in the multiplicative or additive sense, depending on the type of parallelism.

Let us unify a group of p series which looks like a parallel set into a one series in the following way: pick up the first observation in each series one after another such that p ordered observations are obtained. Then repeat it to the second observations, and third observations, and so forth. At last, a series of np observations is obtained. In Figure (16.2) the unified series of the six series of Figure (16.1) is presented. If the initial p series are parallel then the unified series has a fixed Seasonal component pattern combined with an "average" trend. The average trend is the unobservable common trend.

Figure (16.2): The Death Rates of six age groups unified into one series.;...Series ----Seasonally Adjusted data of the unified series.——Trend Component. This unified series is based on data from 1950 through 1978.



Death Rates of Non-White Females:

The unified series has 276 observations. By applying our LPTA approach that uses nonmetric filters (Raveh 1981), fixed seasonal factors were obtained and presented in Table (16.1), below. We adopt purely multiplicative version, taking into consideration the declining fluctuations along with the negative monotone trend. The period's length is p=6 in the proposed data.

In order to test the null hypothesis: $Y._1||Y._2||...||Y._p$ that the p series are perfectly parallel and have common trend, Max $|\mu^{(p)}|$ could be used as a descriptive statistic. If the null hypothesis is true then Max $|\mu^{(p)}|$ should be equal to 1, otherwise it declines toward 0.

The unified series of Data Rates has a coefficient = -.15 indicating that the trend is apparently negative. The fluctuations reduce the absolute value of μ drastically. Thus, Min $\mu^{(6)}$ = -.95 for S.A.D. pointed out the very closeness of the series to being parallel. The fixed seasonal factors reflects the relative ratios of the various age groups to the overall average -- their trend. In order to predict the various Death Rates of the age groups one year ahead we predict 6 units ahead (one whole period) of the unified series. We used the Persistent Structure Principle method was used in order to predict the Death Rates for the (last) year 1978 based on the 54 observations prior to this year, namely the years 1969, 1970, up to 1977. In table (16.2) the predicted values as well as the actual data and percent error is given.

Table (16.1): Fixed Seasonal Patterns presented in percentages using multiplicative model.

	Group Age						Average
	20-24	25-29	30-34	35-39	40-44	45-59	
Fixed Seasonality (in percentages)	32.2	44.2	70.8	100.2	146.9	205.6	100 %

Table (16.2): Actual data for the six age groups in the year 1978, the predicted values based on 54 observations prior to 1978 as well as the absolute percent error. Multiplicative models have been used.

	Group Age					
	20-24	25-29	30-34	35-39	40-44	45-59
Actual data	972	1285	1597	2448	3809	5701
Predicted Values	964	1275	1816	2828	4172	5941
Percent Error	(-.8)	(.8)	(-12.0)	(13.4)	(9.5)	(4.2)

Other decomposition methods or prediction approaches than those used here could be applied for the unified data, see for example Kitagawa (1981) and Akaike and Ishiguro (1981). The basic idea is that the concepts of parallel series and seasonality are related to each other intimately. Thus, series which are nearly parallel could be treated simultaneously by using conventional time-series methods after unifying them appropriately. Other examples of collections of series that could be treated the same way are: In biology, growth rates of various animals; in economics, a collection of yearly taxes series of collection of trend components of series that have something in common.

For a set of p series in which only (p-k) of them are parallel, e.g., k=1,2,...,p-2, then the above treatment could be done to this sub-set only. In some sense, the analysis is similar to Factor Analysis of series. Here, the main and common factor is the trend component which has negative and monotone slope.

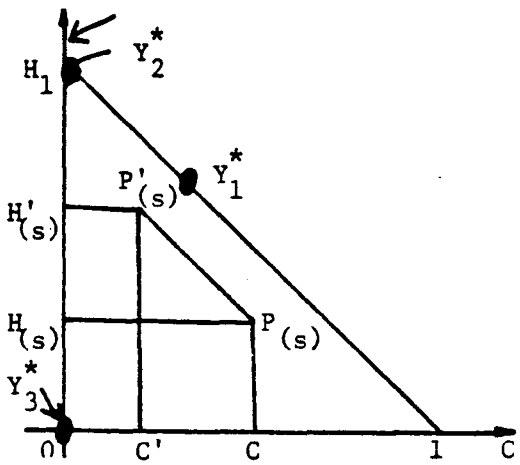
A classic tool to deal with p 'parallel' series is Analysis of Covariance approach. Let $Y_{tk} = f_k(t) + C_k + e_{tk}$ where $f_k(t)$ is a known function of t, usually a line, C_k a constant and $e_{tk} \sim$ i.i.d. $N(0,\sigma^2)$. One would test for parallelism by testing $f = f_1 = \ldots = f_p$. An advantage of the proposed approach is is that $f_k(t)$ is replaced with a very general trend T_{tk} and the data suggests the trend.

For the case of parallelism, its relation to a unified series with fixed seasonality is discussed. Lack of parallelism that is showing up in given series may relate to a unified series with moving seasonality.

17. GRAPHIC PRESENTATION OF QUALITATIVE SERIES

In this chapter, a graphical presentation is suggested for finding common behaviors of qualitative series based on definition given in part 3. The idea is to locate a given series with $P_{(s)}$ as its periodic measure in a right angle triangle as in Figure (17.1) below. $P_{(s)}$ can be decomposed into constancy (c) and Heterogeneousity $H_{(s)}$ which are used as the two axis in the graphic display.

Figure (17.1): $P_{(s)} = C + H_{(s)}$. Two different periodic phenomena with the same $P_{(s)}$ which decomposes into different C and $H_{(s)}$.



In Figure (17.1) the horizontal axis is the measure C and the vertical axis is for the measure $H_{(s)}$. Within this triangle each line parallel to the hypotenuse presents the series with the same predictability value - $P_{(s)}$. Each series located on a line parallel to the hypotenuse has two different components C and $H_{(s)}$. A graphic presentation of a given set of series enables the simultaneous examination of the interrelationships between each couple of series.

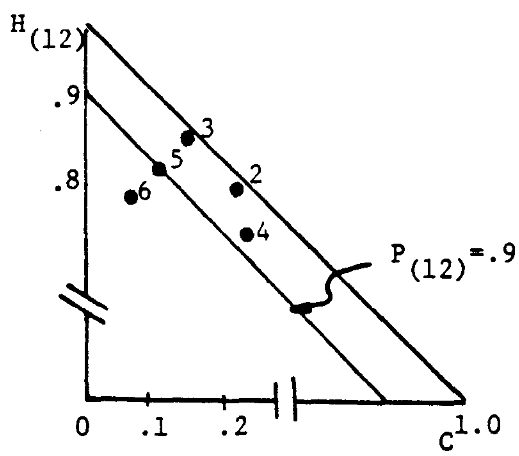
To demonstrate our data analysis approach the computed measures: $P_{(s)}$, C, $H_{(s)}$ and P for the three artificial series Y_1^*, Y_2^*, and Y_3^* of chapter 13 are given in Table (17.1). These series are presented in Figure (17.1) as well.

Table (17.1): The descriptive measures $P_{(s)}$, C, $H_{(s)}$ and $\bar{P}$ for the three artificial series Y_1^*, Y_2^*, and Y_3^*.

The Series	Number of categories=R	Period's length=s	$P_{(s)}$	C	$H_{(s)}$	$\bar{P}$
Y_1^*	4	5	1	.25	.75	1
Y_2^*	3	3	1	0	1	1
Y_3^*	3	3	0	0	0	$\frac{1}{3}$

Real Example: Let us locate the indices obtained for the example in chapter 14 in the right angle triangle in Figure (17.2) below. Each point represents one of the divisions obtained in the process.

Figure (17.2): Graphic presentation of the 5 different divisions (R=2,3,...,6).



Series which fall closer in the triangle have similar behavior in the sense of our definitions.

18. CONCLUSIONS

In this short chapter I would like to indicate the main differences of the LPTA and P.S.P. procedures from classic methods for seasonal adjustment and prediction, respectively. Likewise, I would like to indicate some general conclusions based on my own opinion and experience for the above topics.

About Seasonal Adjustment:

1: In contrast with X-11 and most of the other classic methods, the LPTA estimates the Seasonal Pattern first and there is no need for prior estimation of trend. The LPTA is idempotent procedure or, in other word, one stage procedure. Thus, there are no didfferent estimations of the various components over the estimation process.

2: The three model options for seasonality: Additive, Multiplicative and Mixed are acting on the actual data and there is no need to transform the data as is often done for multiplicative model.

3: A minimal assumptions needed for the trend: either of a monotone shape or polytone curve of order k. For the latter case (K-1) turning points have to be guessed.

4: Estimating the trend is done basically by reordering the Seasonally Adjusted data. First the turning points are selected and then a monotone trend is produced between the turning points by ordering the data - smallest to largest for montone increasing segments and largest to smallest for monotone decreasing segments. This trend is a first guess and a very good approximate solution to the problem (for a montone increasing segment) Min $\sum_{k}^{\ell} (z_t - T_t)^2$ subject to $T_K < T_{K+1} < \dots , < T_\ell$ where z_t is the seasonally adjusted data and T_t the trend at time t. There is an option to further smooth this trend.

5: It seems that any method uses any kind of transformation or smoothing process in order to estimate the trend. Such transformation could be differencing the data or using power transformation on actual data, or estimating an 'optimal' polynom of fixed degree in a global or local (moving-average) way. The re-ordering transformation we use in LPTA is the simplest in the sense of mathematical point of view. We just remove the location of the various observations taking into account their quantative value.

6: When the practitioner has information about turning points the LPTA Procedure allows him to take it into consideration in advance.

7: LPTA procedure enables the user to choose the appropriate type of model for the 3 main components:

Component	Various Options	Criterion to choose
Seasonality Model	Fixed or Moving either fixed or moving could be: Additive, Multiplicative, Mixed or Complex.	Goodness-of-fit = $M_m^{(p)}$
Irregularities Model	Additive or Multiplicative	Coefficient of Monotonicity Graphic display of irregularities, their autocorrelations, etc.
Trend Smoothness	maximum local of monotonicity or linearity or quadratic shape, etc.	Graphic display of irregularities.

8: It seems that the main problem in decomposing any given time series is that at least n+p-1 parameters are needed to be estimated based on at most n data points. In order to solve that problem in as objective a manner as possible some constraints should be adopted and they should relate to pre-defined components.

9: Most of the methods that are in use try to estimate the various components without a given proper definition for them. Thus, usually we can not find a coefficient of goodness-of-fit for various models for seasonality and irregularities. The moving seasonality is a very vague conception and I could not find any definition for it except as to define moving seasonality as fixed seasonality (for a given fixed range of data) that changed over time.

10. For qualitative series, the mode category plays the same role of Seasonal Pattern as in quantitative series.

11: The LPTA is used keeping in mind the parsimony principle. This means that we have to trade-off between the number of parameters and the goodness-of-fit. In order to select turning points and appropriate models and degrees of smoothness of trend, one has to take into consideration simultaneously the magnitude of μ_1, Max$|\mu^{(p)}|$, $M_m^{(p)}$, p=period's length, k=order of polytonicity, μ_Δ, the coefficient of monotonicity of irregularies, the percent change from unit to unit of original data, seasonally adjusted data and trend component. Likewise it is suggested that one look on the charts of original data, S.A.D. and iregularities.

Prediction Purposes

Persistent Structure Principle is (in some way) the other way around the classical methods for prediction purposes. Usually, one fits a model to data in an optimal way and projects it ahead. In contrast, the Persistent Structure Principle fits forecasted data in order to agree with previous behavior measures by goodness-of-fit criteria. The P.S.P enables the using of various shapes for the data: Linear, Monotone, Convex, Quadratic and Exponential are some examples. Persistent Structure Principle can be used for estimating missing data as well. In a sense, Prediction is an estimation process for missing data in future.

Our recommendations for predicition purposes are the following:

1) Decompose the series into S.A.D. and seasonal patterns. Try to choose the most appropriate model of seasonality and irregularities with appropriate smoothness of the trend.

2) Project the S.A.D. or trend and Seasonal Pattern (in case that moving seasonality is assumed) ℓ units ahead. Use P.S.P. or Box-Jenkins approach or any other reasonal prediction method.

3) Combine the forecasted value obtained in step 2 in order to get point estimation for actual data.

4) Use the empirical $(1-\alpha)\%$ central estimated irregularities in order to get $(1-\alpha)\%$ confidence interval to the forecasted value in step 3. Use either multiplicative or additive model. This confidence interval is usually asymmetric.

The four steps above enable one to do forecasting without the need for transforming the data (like power or logarithm transformation). My experience shows that it is usually better to use S.A.D. rather than trend estimation in step 2.

If there are no outliers in the very recent past, use a small number of observations for the projection in step 2, but use the irregularities and seasonality (for fixed model) that was obtained from the entire series. For a polytone trend, the recommendation is to use only the last monotone segment (tone).

Appendix A

ABOUT THE COEFFICIENTS OF MONOTONICITY & POLYTONICITY

The family of coefficients presented here are designed to measure monotone association between two (at least) ordered variables. Q of Yule, γ of Goodman & Kruskal, Kendall's τ and Spearman's r_S are special cases of the family as well as other coefficients. The coefficients are based on the Absolute Value Principle which will be discussed in the next section. More details are given in Raveh (1982b). One special case of this family is used in this report to measure the monotone association between quantitative variable and ordered variable, usually time. The family is extended to be able to assess polytone association where the turning points are or are not given in advance.

<u>Absolute Value Principle for Monotonicity Coefficients</u>

In order to measure the amount of monotone association two main aspects have to be taken into consideration:

(i) The shape of the monotone relationship

(ii) The weight given to each deviation from perfect relatioships.

The Absolute Value Principle means that $U\,\mathrm{sgn}(U) = |U|$ or $-|U| \leq U \leq |U|$, where sgn(U) is the sgn of the real number U. That is,

$$\mathrm{sgn}(u) = \begin{cases} 1 \text{ if } U>0 \\ 0 \text{ if } U=0 \\ -1 \text{ if } U<0 \end{cases}$$

Below, a formal definition for a wide family of coefficients is given based on the above principle.

Let x and y be any two <u>ordered</u> sets (numerical or not), and let (x_i, y_i) $(i=1,2,\ldots,n)$ be pairs of observations, where x_i x and y_i y $(i=1,2,\ldots,n)$. Let

$$\alpha_{ij} = \begin{cases} 1 \text{ if } x_i > x_j \\ 0 \text{ if } x_i = x_j \\ -1 \text{ if } x_i < x_j \end{cases} \qquad \beta_{ij} = \begin{cases} 1 \text{ if } y_i > y_j \\ 0 \text{ if } y_i = y_j \\ -1 \text{ if } y_i < y_j \end{cases} \quad (i,j=1,2,\ldots,n)$$

or in other words: $\alpha_{ij} = \mathrm{sgn}(x_i-x_j)$, $\beta_{ij} = \mathrm{sgn}(y_i-y_j)$. In order to consider the shape of the monotone curve, let

$$\theta_{ij}^{(1)} = |\alpha_{ij}\beta_{ij}| \,,$$
$$\theta_{ij}^{(2)} = |\alpha_{ij}| \,,$$
$$\theta_{ij}^{(3)} = |\beta_{ij}| \,,$$
$$\theta_{ij}^{(4)} = |\alpha_{ij}| + |\beta_{ij}| - |\alpha_{ij}\beta_{ij}| \,.$$

Then always $\theta_{ij}^{(m)} = 0$ or 1 for m=1,2,3,4.

Furthermore, for all i and j

$$\theta_{ij}^{(1)} \leq \left\{ \begin{array}{c} \theta_{ij}^{(2)} \\ \theta_{ij}^{(3)} \end{array} \right\} \leq \theta_{ij}^{(4)} , \tag{1}$$

and

$$-\theta_{ij}^{(m)} \leq \alpha_{ij}\beta_{ij} \leq \theta_{ij}^{(m)} \quad \begin{array}{l}(i,j=1,2,\ldots,n) \\ m=1,2,3,4\end{array} \tag{2}$$

The upper index m designates for the four types (shapes) of monotonicity. They are called: weak, semi-weak, semi-strong and strong monotonicity, respectively. To have an equality in (2) holds for all (i,j) for fixed m is to have perfect

$$\left\{ \begin{array}{l} \text{weak} \\ \text{semi-weak (semi-strong)} \\ \text{strong} \end{array} \right\} \text{ monotonicity acccording as } m = \left\{ \begin{array}{c} 1 \\ 2(3) \\ 4 \end{array} \right\} .$$

The sign being + or -, according as to equality is on the right or left of (2).

The second aspect, namely, weighting any pair of observations that deviate from conditions (2) is done by a weight function $W_{ij}^{(m)}$ i,j=1,...,n.

Let $W_{ij}^{(m)}$ be any non-negative numbers not vanishing if $\theta_{ij}^{(m)}$ does not vanish:

$$\mathrm{sgn}(W_{ij}^{(m)}) \geqq {}_{ij}^{(m)} ,$$

then, the family of coefficients is defined by quantity (3)

$$\mu_W^{(m)} = \frac{\sum_{i=j}^{n} \sum_{j=1}^{n} W_{ij}^{(m)} \alpha_{ij}\beta_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij}^{(m)} \; {}_{ij}^{(m)}} . \quad (m=1,2,3,4) \tag{3}$$

It is easy to verify that always

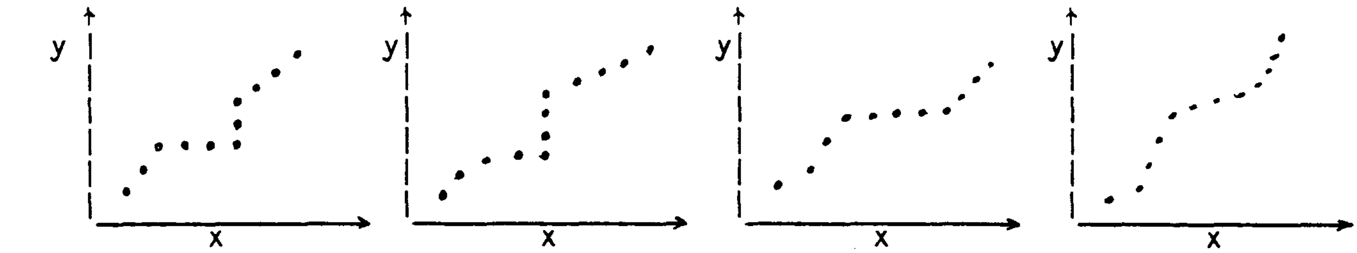
$$-1 \leqq \mu_W^{(m)} \leqq 1, \tag{4}$$

and an equality will hold in (4) if and only if perfect monotonicity of type m holds.

(a) <u>Four Types of Monotonicity</u>

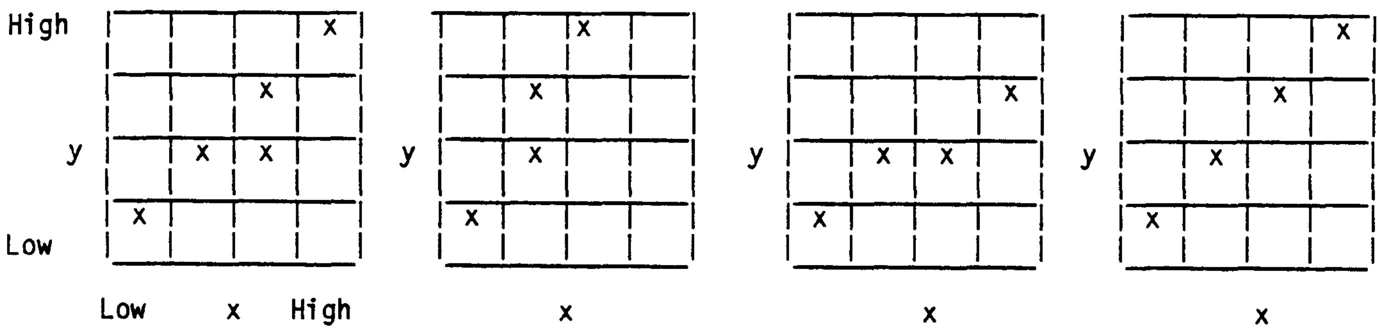
As mentioned earlier, four types of monotonicity are recognized: weak (m=1), semi-weak (m=2), semi-strong (m=3) and strong (m=4). These four types have different definitions for monotonicity. In Figure 1 four curves for 2 ordered quantitative sets of points are plotted. These four curves have perfect fitness to the four types of monotonicity, respectively. Four types of Perfect monotonicity for qualitative variables are presented below in Figure 2. The data grouped in two-dimensional cross classification table.

Figure 1: Four curves that have perfect adequation for the 4 types of monotonicity: (a) weak, (b) semi-weak, (c) semi-strong, (d) strong. Below the graphs, the conditions for these types of monotonicity are given for all i, j=1,...,n.



(a) weak	(b) semi-weak	(c) semi-strong	(d) strong
$\alpha_{ij}\beta_{ij}=\|\alpha_{ij}\beta_{ij}\|=\theta_{ij}^{(1)}$	$\alpha_{ij}\beta_{ij}=\|\alpha_{ij}\|=\theta_{ij}^{(2)}$	$\alpha_{ij}\beta_{ij}=\|\beta_{ij}\|=\theta_{ij}^{(3)}$	$\alpha_{ij}\beta_{ij}=\|\alpha_{ij}\|+\|\beta_{ij}\|-\|\alpha_{ij}\beta_{ij}\|=\theta_{ij}^{(4)}$
$\mu^{(1)}=1$	$\mu^{(2)}=1$	$\mu^{(3)}=1$	$\mu^{(4)}=1$

Figure 2: Four 4x4 tables that present perfect fitness for the various definitions of monotonicity.



(a)	(b)	(c)	(d)
x is not a unique function of y and vice versa.	x is a function of y but <u>not</u> vice versa.	y is a function of x but <u>not</u> vice versa.	x is a one-to-one function of y.

Necessary Conditions

NO	y has not less categories than x.	x has not less categories than y.	x and y have the same number of categories.

(b) Three Types of Weighting

For each m, different choices of $W_{ij}^{(m)}$ lead to different families of coefficients. We distinguish among three major families: $\mu_w^{(m)}$ w=0,1,2 where w indicates the number of numerical variables that are involved. As mentioned earlier, the upper index indicates the type of monotonicity and the lower index indicates the form of weighting each pair of observations. The three types of weights, w=0,1,2, are loosely to designate for nominal (qualitatives), ordered and interval (quantitatives) variables, respectively.

Below, the condition for weights for the three major families are given:

1. Coefficients designate for qualitative data and thus based on order only: Here, the first family is demonstrated, namely w=0.

   $\mu_w^{(m)} = \mu_0^{(m)} \quad w_{ij}^{(m)}$ C where C is any positive constant.

This case is appropriate for two qualitatives variables (often two-dimensional cross-classification table) where the weights are equal. Every pair of observations deviates from the monotonicity condition and obtains the same weight. In other words, this means that the number of deviations from monotonicity condition is the loss function. The existance of a deviation is important and not the amount of deviation. By substituting $w_{ij}^{(m)} \equiv C$ into eq.(3) $\mu_0^{(m)}$ is obtained.

$$\mu_0^{(m)} = \frac{\sum_{i=j}^{n} \sum_{j=1}^{n} \alpha_{ij}\beta_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n} \theta_{ij}^{(m)}} \qquad m=1,2,3,4 \qquad (5)$$

A special case of $\mu_0^{(m)}$ is when m=1, namely $\mu_0^{(1)}$. This coefficient is the very well-known Gamma (γ) of Goodman & Kruskal (1954, p.749), which is designated for Rxc cross-classification table. For 2x2 contingency table Q of of Yule is a special case of $\mu_0^{(1)} \equiv \gamma$.

Another special case of $\mu_0^{(m)}$ where m=4 and there are no ties is Kendalls τ .

2. Coefficients Designated to Ordinal or Numerical Data.

In this section the lower index w=1 is adopted. This family is designated either for two ordinal variables or for one ordinal and one numerical (quantitative) variable. For this family the weights are given to only one of the variables.

Three possible kinds of weights are discussed.

(i) $w_{ij} = |i-j|$

(ii) $w_{ij} = |y_i-y_j|$

(ii) $w_{ij} = |y_i-y_j|^2$

(i) Using the first kind of weights as a special case of $\mu_1^{(m)}$ is the very well-known coefficient-r_s of Spearman. The weights $w_{ij} = |i-j|$ is the difference in ranks of the sorted observations, x_i and x_j respectively.

μ_1 for Time Series

A special case of $\mu_1^{(m)}$ for m=1 or 3 and weight $w_{ij} = |Y_i-Y_j|$ is used extensively in this report, where x is the time axis, namely $x_i = i$, $i=1,\ldots,n$. Recall that $\beta_{ij}|Y_i-Y_j| = (Y_i-Y_j)$ after some manipulation it can be shown that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij}\cdot\beta_{ij}w_{ij} = 2\sum_{i=1}^{n} Y_i(\sum_{j=1}^{n}\alpha_{ij}) = 2\sum_{i=1}^{n} Y_i\alpha_{i.}$$

where $\alpha_{i.} = \sum_{j=1}^{n}\alpha_{ij}$

and in a similar way

$$\sum_{i=1}^{n} \sum_{j=1}^{n} |\alpha_{ij}\cdot \beta_{ij}|w_{ij} = 2\sum_{i=1}^{n} Y_i(\sum_{j=1}^{n}\beta_{ij}|\alpha_{ij}|)$$

By substituting the above expressions into (3) one finds that

$$\mu_1^{(1)} = \frac{\sum_{i=1}^{n} Y_i\alpha_{i.}}{\sum_{i=1}^{n} Y_i(\sum_{j=1}^{n}\beta_{ij}|\alpha_{ij}|)} \qquad (6)$$

For the case where variable x has no ties as is usually the case in quantitative time series, then $|\alpha_{ij}| = 1$ for $i \neq j$ and (14) has a simpler formula:

$$\mu_1^{(1)} = \frac{\sum_{i=1}^{n} Y_i \alpha_{i.}}{\sum_{i=1}^{n} Y_i \beta_{i.}} = \frac{COV(Y,R(X))}{COV(Y,R(Y))} \quad (7)$$

and the coefficients of monotonicity is proportional to the quotient of two covariances. These covariances are of the Y variable with the time ranks and its ranks, respectively.

The coefficient $\mu_1^{(1)}$ in eq. (7) can be expessed in a different way:

$$\mu_1^{(1)} = \frac{\sum_{i>j} (Y_i - Y_j)}{\sum_{i>j} |Y_i - Y_j|} \quad (8)$$

Formula (8) is used for the purpose of decomposition of economic time series in our LPTA procedure and are used for prediction purposes as well.

(iii) By substituting the weight $w_{ij} = (y_i - y_j)^2$ the coefficient $\mu_1^{(1)}$ has been used by Johnson (1975) for Nonmetric Analysis of Variance.

3. Coefficient for two numerical variables

This section deals with w=2, or in other words, both variables are numerical. The weights $w_{ij}^{(m)}$ should be such that $w_{ij}^{(m)} < w_{ij}^{(m)}$ if $x_i < x_j < x_k$ and $Y_i \leq Y_j \leq Y_k$ or vice versa. The weight should take into account both variables. Two various kinds of weights are:

(i) $w_{ij} = |x_i - x_j| \cdot |Y_i - Y_j|$

(ii) $w_{ij} = (x_i - x_j)^2 + (Y_i - Y_j)^2 = d_{ij}^2$

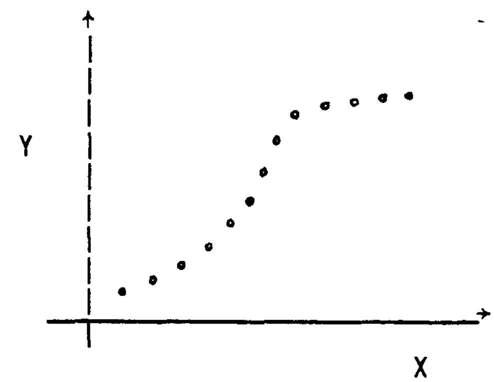
$\mu_2^{(1)}$ Versus r_{xy} Of Pearson

The first weight $w_{ij} = |x_i - x_j|\ |Y_i - Y_j|$ means that the "area" between every pair of observations is taken into account. Substituting this type of weighting into (3) with m=1 yields the equation below (9)

$$\mu_2^{(1)} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-x_j)(Y_i-Y_j)}{\sum_{i=1}^{n}\sum_{j=1}^{n}|x_i-x_j||Y_i-Y_j|} = \frac{\sum_{i>j}(x_i-x_j)(Y_i-Y_j)}{\sum_{i>j}|x_i-x_j||Y_i-Y_j|} \qquad (9)$$

It is easy to verify that always $|\mu_2^{(1)}| \geq |r_{xy}|$ where r_{xy} is Pearson's product-moment coefficient. The equality holds only if $|r_{xy}|$ is 0 or 1. The same inequality was found for Pearson's coefficient and Q of Yule for the case of qualitative variables. If $|r_{xy}|=1$ then $|\mu_2^{(1)}|=1$ but not vice versa as in figure 3.

Figure 3: A perfect weak monotonicity association exists, ($\mu_2^{(1)}=1$), while linear relationship does not imply ($r_{xy}<1$).



Both coefficients have proportional denominator and thus, sgn $(\mu_2^{(1)})=\text{sgn}(r_{xy})$.

$\mu_2^{(1)}$ Using $w_{ij} = d_{ij}^2$

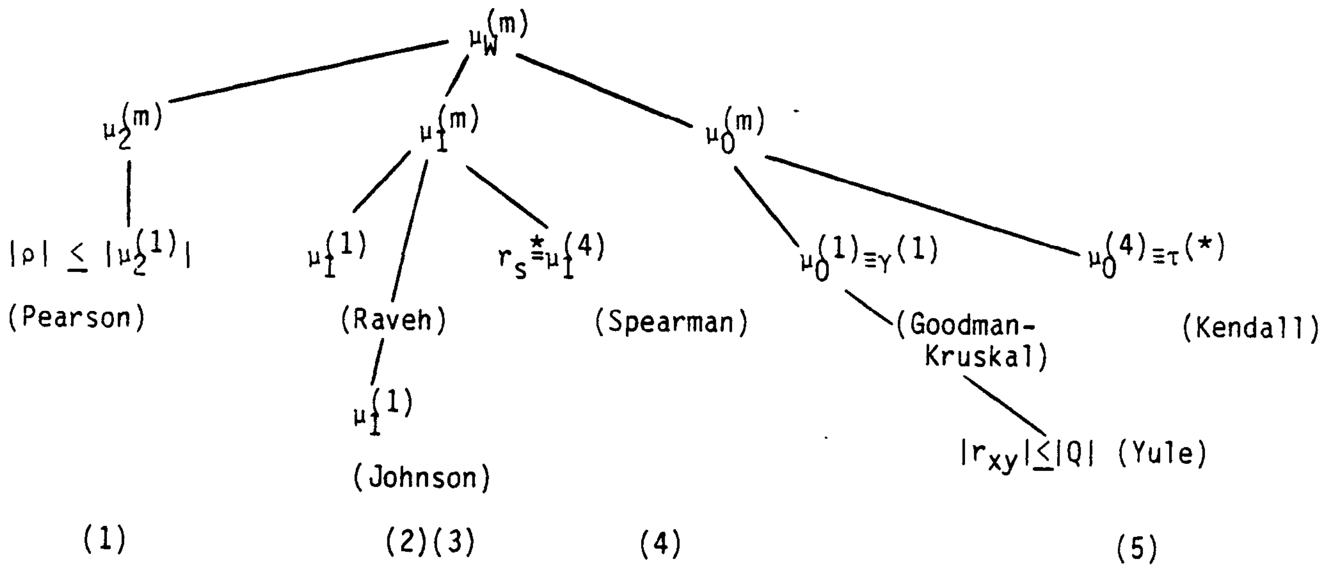
A special version a $\mu_2^{(1)}$ is achieved by using as a weight the square euclidean distance between every pair of observations. This weight $w_{ij} = d_{ij} = (X_i-X_j)^2 + (Y_i-Y_j)^2$ does not vanish for each one of the observations in Figure 3. Substituting these weights into eq. (3) yields eq. (10) below:

$$\mu_2^{(1)} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_{ij}\beta_{ij}\ \{\ (X_i-X_j)^2 + (Y_i-Y_j)^2\ \}}{\sum_{i=1}^{n}\sum_{j=1}^{n}|\alpha_{ij}||\beta_{ij}|\ \{\ (X_i-X_j)^2 + (Y_i-Y_j)^2\ \}} \qquad (10)$$

Conclusions

The various coefficients of association like the γ of Goodman and Kruskal, Kendall's τ and Spearmans's r as well as other coefficients, are various coefficients of monotonicity based on the Absolute Value Principle. Each coefficient uses its own particular combination of the weight given to each pair of observations and the desired shape of the monotone relationship. In Figure 4 below, a schematic graph is presented for the various special cases. The family of coefficients is used as a definitional basis for a variety of data analysis techniques. Some of the applications are given in the figure.

Figure 4: A schematic graph of the various special cases of Coefficients of Monotonicity. Some suggested applications are given as well.



Some Possible Applications:

1- Multidimensional Scaling (MDS): Mapping an nxn matrix of 'things' (say, correlations) into Space (say, Euclidean) with as small dimensionality as possible in order to keep the following conditions: If $r_{ij} > r_{kl}$ then $d_{ij} < d_{kl}$ to as many indices i,j,k,l as much as possible. The procedure maximizes $\mu_{r,d}^{(d)}$ or in other words, maximizes the coefficient of

(*) with the assumption of no ties.

monotonicity between, say, correlations (input) and the Euclidean dis-distances (d_{ij}-output) on the Space diagram.

2- Analysis of Variance (ANOVA) in a Nonmetric Approach.

3- Analysis of quantitative time-series: Seasonal Adjustment and Forecasting.

4- Coefficient of correlation of ranks (two ordinal variables).

5- Measuring Association of Contingency tables

6- Coefficient of association between 2 ordered variables.

Coefficients of local Monotonicity (Polytonicity)

Polytone curve of order K is a collection of K monotone curves (tones) unified together. The (k-1) points in between the monotone tones are defined as turning points. In this sense, monotone association is a special case of polytonicity of order K, namely, there are no turning points. In order to assess the polytone association quantitatively, we will define coefficients of polytonicity in two ways.

1) When the turning points are known in advance: The time axis would be divided accordingly. The coefficient of Polytonicity (||) is extension of eq.(3)

$$\mu_{w,K}^{(m)} = \frac{\sum_{k=1}^{K} \sum_{i>j}^{I_K} w_{ij}^{(m)} \alpha_{ij}\beta_{ij} \delta_{ij}^{(k)}}{\sum_{k=1}^{K} \sum_{i>j}^{I_k} w_{ij}^{(m)} \theta_{ij}^{(m)}} . \quad (m=1,2,3,4) \tag{11}$$

where $\delta_{ij} = (-1)^{(k-1)}$ is the sign of the direction of the related segment of time-axis, where both i and j belong to, either positive or negative. Usages of (11) are presented earlier in (2.5) and (2.6).

2) When the turning points are not known in advance: The conditions for weak monotonicity are:

$(Y_i-Y_{i-1})(Y_{i-1}-Y_{i-2}) \geq 0$ for all i=3,...,N, where $\Delta Y_i = Y_i-Y_{i-1}$ or $\Delta Y_i \Delta Y_{i-1} = |\Delta Y_i \Delta Y_{i-1}|$ for all i=3,...,N.

Thus, series can be measured by coefficient of local monotonicity given in equation (12) below.

$$\text{Mon}(Y) = \frac{\sum_{i=3}^{N} \Delta Y_i \cdot \Delta Y_{i-1}}{\sum_{i=3}^{N} |\Delta Y_i \cdot \Delta Y_{i-1}|} \tag{12}$$

Mon(Y) = 1 if and only if the series is perfectly weak monotone with either positive or negative slope. If very few turning points exist (relative to the length of the series) and between them the series is monotone then Mon(Y) $\doteq$ 1, and we call it local monotonicity. The least monotone series would be obtained for the following series: a,b,a,b,...,b where $a \neq b$. For such a series Mon(Y) = -1 and local linear series yields Mon(Y) = 1. More details are given in Raveh (1982c).

APPENDIX B

Some of the actual Series Analyzed in the Report

Table A: "U.S. RETAIL SALES IN MILLIONS OF $" IN THE YEARS 1960-1964.

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1960	16312	15829	17632	18973	18548	18918	18066	18153	17848	18648	18385	22153
1961	15803	15071	17714	17618	18532	18907	17992	18325	18158	18761	19224	22881
1962	17007	16042	19193	19097	20226	20254	19138	19920	18863	20576	20911	24127
1963	18261	17087	19653	20518	21228	20737	20540	21018	19267	21528	21494	25104
1964	19154	18758	20502	21186	22508	22242	22145	21778	21313	22605	21720	27719

Table B: Chatfield-Prothero Case-Study. 1-65 to 5-71.

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
154	96	73	49	36	59	95	169	210	278	298	245
200	118	90	79	78	91	167	169	289	347	375	203
223	104	107	85	75	99	135	211	335	460	488	326
346	261	224	141	148	145	223	272	4454	560	612	467
518	404	300	210	196	186	247	343	464	680	711	610
613	392	273	322	189	257	324	404	677	858	895	664
628	308	324	348	272							

Table C: Public Consumption of Electricity in the U.S. in the Years 1951-1958

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1951	318	281	278	250	231	216	223	245	269	302	325	347
1952	342	309	299	268	249	236	242	262	288	321	342	364
1953	367	328	320	287	269	251	259	284	309	345	367	394
1954	392	349	342	311	290	273	282	305	328	364	389	417
1955	420	378	370	334	314	296	305	330	356	396	422	452
1956	453	412	392	362	341	322	335	359	392	427	454	483
1957	487	440	429	393	370	347	357	388	415	457	491	516
1958	529	477	463	423	398	380	389	419	448	493	526	560

Table D: Original Data of the Artificial Series (Has complex seasonality).

13050	17412	21158	17076	18548	15134	19873	21784	16063	18641	14708	24368
18964	13564	17714	14094	20385	22688	16130	18325	14526	20637	23069	20593
17007	12834	21112	22916	18203	20254	15310	21912	22636	18518	20911	19302
20087	20504	17688	20518	16982	22811	24648	18916	19267	17222	23643	30125
17239	18758	16402	23305	27010	20018	22145	17442	23444	27126	27719	27719

Table E: Agricultural Employment, Men, 20 years and older

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2638	2648	2672	2844	2836	2955	2970	2909	2924	2922	2814	2718
2702	2739	2750	2891	2915	3064	3030	2859	2809	2757	2717	2559
2479	2557	2579	2713	2786	2879	2815	2766	2663	2627	2447	2324
2283	2330	2423	2636	2696	2801	2759	2614	2578	2500	2418	2286
2233	2194	2324	2518	2546	2627	2633	2556	2484	2531	2440	2266
2230	2243	2287	2417	2500	2642	2660	2647	2682	2703	2532	2464
2319	2289	2388	2488	2524	2694	2664	2596	2528	2558	2536	2420
2448	2483	2503	2508	2571	2609	2655	2634	2574	2570	2415	2311
2226	2282	2310	2401	2499	2569	2591	2579	2557	2514	2362	2177
2163	2174	2202	2379	2468	2588	2596	2531	2405	2424	2248	2125
2030	2081	2106	2259	2423	2536	2464	2492	2406	2427	2283	2192
2171	2105	2145	2274	2393	2617	2599	2525	2512	2462	2277	2250
2084	2117	2176	2237	2342	2509	2520	2554	2498	2472	2403	2292
2160	2213	2217	2255	2422	2470	2475	2455	2525	2459		

Table F: All Employees in Food Industries

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1720	1702	1707	17081	1727	1789	1829	1880	1920	1872	1811	1771
1706	1685	1690	1700	1711	1788	1819	1919	1914	1867	1804	1776
1719	1710	1707	1711	1724	1784	1827	1936	1921	1853	1820	1778
1731	1727	1723	1711	1727	1787	1817	1916	1900	1844	1795	1758
1708	1693	1689	1685	1705	1761	1811	1898	1895	1816	1783	1744
1697	1675	1682	1675	1690	1761	1785	1862	1860	1803	1747	1707
1661	1654	1651	1641	1651	1705	1738	1811	1821	1786	1749	1711
1666	1649	1661	1643	1659	1696	1722	1822	1828	1867	1707	1660
1594	1571	1575	1572	1593	1643	1681	1780	1791	1742	1692	1655
1619	1615	1605	1612	1637	1685	1728	1813	1817	1754	1710	1672
1638	1632	1638	1645	1658	1717	1747	1829	1840	1765	1725	1698
1665	1655	1668	1664	1669	1722	1749	1823	1830	1774	1746	1724
1685	1666	1676	1666	1679	1728	1750	1829	1835	1782	1736	1706

REFERENCES

Akaike, H. (1981). "Seasonal Adjustment by a Bayesian Modeling". J. Time Series Analysis Vol.1, No.1, 1-14.

Akaike, H. and Isiguro, M. (1981). "Comparative Study of the X-11 and Baysea Procedures of Seasonal Adjustment". In Applied Series Analysis of Economic Data, ed. Arnod Zellner, Washington D.C: U.S. Department of Commerce, Bureau of the Census.

Anderson, T.W. (1959). "Some Stochastic Process Models for Intelligence Test Scores", in Mathematical Methods in the Social Sciences, Arrow, Karlin & Suppes, eds., Stanford Univ. Press, 205-220.

Anderson, T.W. (1971). The Statistical Analysis of Time Series. John Wiley & Sons, London.

Anderson, O.D. (1976). Time Series Analysis and Forecasting: The Box-Jenkins Approach. Butterworths, London & Boston.

Anderson, O.D. (1977). A commentary on 'Survey of Time Series'. International Statistical Review 45, 273-297.

Auerbach, R.D. and Rutner, J.L. (1978). "The Misspecification of a Nonseasonal Cycle as a Seasonal by the X-11 Seasonal Adjustment Program." The Review of Economics and Statistics, Vol.LX, No.4, 601-603.

Bachi, R. (1951). The Tendency to Round-off Age Returns: Measurement and Correction, "Bulletin of the International Statistical Institute", Vol.XXXIII, part IV; 195-222.

_________ (1953). Measurement of the Tendency to Round-off Age Returns, Proceedings of the International Statistical Congress, Rome.

Bar On, R.R.V. (1973). "Analysis of Seasonality and Trends in Statistical Series-Methodology and Applications in Israel", Technical Publication No. 39.

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975). Discrete Multivariate Analysis: Theory and Practice. The MIT Press, Cambridge, Massachusetts, and London, England.

B.L.S., (1966). "The BLS Seasonal Factor Method". Washington D.C., Bureau of Labor Statistics, U.S. Department of Labor

Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis: Forecasting and Control. Holden Day, San Francisco.

_________ (1973). "Some Comments on a Paper by Chatfield and Prothero and on a Review by Kendall." J.R.Statist.Soc. A, 136, part 3, 337-352.

Burman, J.P. (1965). Moving Seasonal Adjustment of Economic Time Series, J.R. Statist.Soc., A. 128, part 4, 534-558.

Carrier, N.H. and Farrag, A.M. (1959). The Reduction of Errors in Census Populations for Statistically Underdeveloped Countries, "Population Studies", 12; 240-285.

Chatfield, C. and Prothero, D.L. (1973). "Box-Jenkins Seasonal Forecasting: Problems in a Case-Study." J.R.Statist.Soc. A, 136, part 3, 295-315.

Cleveland, W.S., Dunn, D.M., and Terpenning, I.J., (1978). "SABL--A Resistant Seasonal Adjustment Procedure with Graphical Methods for Interpretation and Diagnosis, in Seasonal Analysis of Economic Time Series, ed. Arnold Zellner, U.S. Department of Commerce, Bureau of the Census, Washington, D.C: 201-231.

Colwell, R.K. (1974). Predictability, constancy and contingency of Periodic Phenomena, Ecology 55, 1148-1153.

Durbin, J. and Murphy, M.J, (1975). Seasonal Adjustment Based on a Mixed Additive Multiplicative Model. J.R.Statist.Soc. A, 138, part 3, 385-410.

Durbin, J. and Kenny, P.B. (1978). "Seasonal Adjustment when the Seasonal Component Behaves Neither Purely Multiplicatively Nor Purely Addditively." In Seasonal Analysis of Economic Time Series, ed., Arnold Zellner, U.S. Department of Commerce, Bureau of the Census, Washington D.C., 173-188.

Fase, M.M.G., Koning, J. and Volgenant, A.F. (1973). An experimental look at seasonal adjustment. De Economist 121, 441-480.

Gersch, W. and Kitagawa, G. (1982). Interim ASA/Census Time Series Research Report. (Preliminary Draft: The Prediction of Time Series with Trends and Seasonalities".)

Gilad, D. (1968). Evapotranspiration in salt-marsh of Yotvata area. Hydrological Service of Israel (in Hebrew). p. 49.

Goodman, L.A. and Kruskal, W.H. (1954). Measures of association for cross-classification. J.Amer.Statist.Assoc 49, 732-764.

Guttman, L. (1941). An outline of the statistical theory of prediction, supplementary study B-1 (pp.253-318) in Horst, Paul and others, eds. The Prediction of Personal Adjustment, Bulletin 48, Social Science Research Council, New York.

_________. (1950). "The Principal Components of Scale Analysis", in Stouffer, S.A. ed. Measurement and Prediction, Vol.4, Princeton, New Jersey, Princeton University Press, pp.312-361.

_________. (1954). "A New Approach to Factor Analysis: The Radex", in Mathematical Thinking in the Social Sciences, Lazarsfeld, P.F., ed., The Free Press, Glencoe, Illinois; 258-348.

_________ (1968). "A General Nonmetric Technique for Finding the Smallest Coordinate Space for a Configuration of Points", Psychometrika, 33, 469-506.

_________ (1977)., "What is Not What in Statistics". The Statistician, Vol.6, Number 2; 81-107.

Hillmer. S.C., Bell, W.R., and Tiao, G.C. (1982). "Modeling Considerations in the Seasonal Adjustment of Economic Time Series. In Applied Time Series Analysis of Economic Data, Arnold Zellner, ed., Washington, D.C. October 1981. To appear.

Hillmer, S.C., Tiao, G.C., (1982). An ARIMA-Model-Based Approach to Seasonal Adjustment. J. of the American Statistical Association, Vol.77, No.377, 63-70.

Kendall, M.G. (1973). Time Series. Hafner Press, New York.

Kenny, P.B. (1975). Problems of Seasonal Adjustment. Statistical News No.29, 29.1-29.6.

Kitagawa, G. (1981). "A nonstationary time series model and its fitting by a recursive filter". J. Time Series Analysis, Vol.2, No.2, 103-116.

Kuiper, J. (1978). "A Survey and Comparative Analysis of Various Methods of Seasonal Adjustment," in Seasonal Analysis of Economic Time Series, Arnold Zellner, ed ., Washington D.C: U.S. Department of Commerce, Bureau of the Census, 59-76.

Ledermann, M.S. and Breas, M.J. (1959). "Les Dimensions De La Mortalité Population, Vol.14, No.4, 637-682.

Lingoes, J. (1973). "The Guttman-Lingoes Nonmetric Program Series." Mathesis Press, Ann Arbor, Michigan.

Makridakis, S. (1976). A Survey of Time Series. International Statistical Review 44, 29-70.

Marten J.F. (1924). Census of India, 1921, Vol.1, part I; Calcutta:126-127.

Myers, R.J. (1940). Errors and Bias in the Report of Ages in Census Data, Actuarial Society of America, Transactions, Vol.XLI:pp.411-415.

Ozaki, T. (1977). "On the Order Determination of ARIMA Models." Applied Statistics, 26, No.3, 290-301.

Pierce, D.A. (1980). "A Survey of Recent Developments in Seasonal Adjustment." The American Statistician, Vol 34, No. 3, 125-134.

Plewes, T.J.(1978). "Seasonal Adjustment of the U.S. Unemployment Rate: Introductory Remarks". The Statistician, Vol.27, Nos.2,3,4, 177-179.

Powell, M.Y.D. (1964) An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives, Computer Journal, 7: pp.155-162.

Raveh, A. (1975). Analysis of a 'Monotone-Periodic' Time Series. Proceedings of the National Conference on Data Processing 2, 747-760 (in Hebrew).

_________. (1978). Finding Periodic Pattern in Time Series with Monotone trend: A New Technique. In Theory Construction and Data Analysis in the Behavioral Sciences. S. Shye, ed. Chap. 15, 371-390. Jossey-Baass, San Francisco.

Raveh, A. and Tapiero, C. (1980). Cyclicity, Constancy, Homogeneity and the categories of Qualitative Time-Series. Ecology, 61 (3), pp.715-719.

Raveh, A. (1981). "A Numerical Nonmetric Approach for Analyzing Time Series Data:. Communications in Statistics - Theory and Methods, A10(8), 809-821.

_________. (1981a). "Choosing the Suitable Type of Seasonality", Submitted for Publication.

_________. (1982). Forecasting Economic Series Using the Persistent Structure Principle, submitted for publication.

_________. (1982a), "Comments on Kenneth F. Wallis 'Models for X-11 and X-11 Forecast Procedures for Preliminary and Revised Seasonal Adjustment'". In Applied Time Series Analysis of Economic Data, Arnold Zellner (ed.). Washington D.C. October 1981. To appear.

_________. (1982b). "Guttman's Coefficients of Monotonicity. Submitted for publication.

Schlicht, E. (1981) "A Seasonal Adjustment Principle and a Seasonal Adjustment Method Derived from this Principle". Journal of the American Statistical Association, Vol.76, No.374, 374-378.

Shiskin, J., Young, A.H., and Musgrave, J.C. (1967), "The X-11 Variant of the Census Method-II Seasonal Adjustment Program," Technical paper No. 15, U.S. Bureau of the Census.

United Nations (1955). Measurement of Age-Accuracy by Means of an Index, ST/SOA/ Series A, "Population Studies", n.23; United Nations; New York.

Wallis, K. (1974), "Seasonal Adustment and Relations Between Variables," Journal of the American Statistical Association, 69, 18-31.

Wallis, K. (1981), "Models for X-11 and 'X-11-Forecast' Procedures for Preliminary and Revised Seasonal Adjustments". In Applied Time Series Analysis of Economic Data, ed. Arnold Zellner, Washington D.C: U.S. Department of Commerce, Bureau of the Census.

Wilson, G.T. (1973). Discussion of Paper by Dr. Chatfield and Dr. Prothero. J.R.Statist.Soc. A, 136, part; 3, 315-319.

Young, A.H., (1968), "Linear Approximation to the Census and BLS Seasonal Adjustment Methods, Journal of the American Statistical Association, 63, 445-457.

Zangwill, W. (1967). Minimizing a Function Without Calculating Derivatives, "Computer Journal", 10; pp.293-296.

Zellner, A. (1978), "Seasonal Analysis of Economic Time Series" (ed.) Washington, D.C: U.S. Department of Commerce, Bureau of the Census.

Zellner, A. (1982), "Applied Time Series Analysis of Economic Data" (ed.) Washington D.C: U.S. Department of Commerce, Bureau of the Census (to appear)