BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-80/01

1979-80-CENSUS SEASONAL ADJUSTMENT PROJECT

FINAL REPORT ON RESEARCH ACTIVITIES

by

Steven Hillmer
Research Fellow

William Bell
Research Trainee

U.S. Bureau of the Census
Washington, D.C. 20233

This series contains research reports, written by or in cooperation with staff members of the Statistical Research Division, whose content may be of interest to the general statistical research community. The views reflected in these reports are not necessarily those of the Census Bureau nor do they necessarily represent Census Bureau statistical policy or practice. Inquiries may be addressed to the author(s) or the SRD Report Series Coordinator, Statistical Research Division, Bureau of the Census, Washington, D.C. 20233.

Recommended by: Kirk M.Wolter

Report completed: August 31, 1980

Report issued: September 1, 1983

1979–80 ASA–Census Seasonal Adjustment Project

Final Report on Research Activities

by

Steven Hillmer
Research Fellow

and

William Bell
Research Trainee

Introduction

The Census X-11 seasonal adjustment method is a widely used procedure that has been difficult to understand from a theoretical point of view. In an attempt to understand how X-11 might fit into the statistical framework of signal extraction, Cleveland and Tiao (1976) found an ARIMA model for the observed unadjusted series for which the standard options of X-11 is appropriate. However, many series that are adjusted by X-11 do not follow the Cleveland and Tiao model; consequently an appropriate question is: given an ARIMA model for the unadjusted data, what is an appropriate way to seasonally adjust the series? In an attempt to answer this question, Box, Hillmer and Tiao (1978) began to develop an ARIMA model based approach to seasonal adjustment. The basic idea behind this approach is that there is information in the unadjusted series about how the seasonal adjustment should be carried out and that this information should be fully exploited when seasonally adjusting a series. The ideas in Box, Hillmer and Tiao (1978) are largely theoretical in nature and prior to this project have not been tested on a large number of actual series. Therefore, the main purpose of this year's project has been to complete an empirical comparison involving the model based procedure given in Box, Hillmer and Tiao (1978).

Before proceeding we shall briefly indicate some of the reasons that motivated us to complete the empirical study. First, a comparison of the model based approach with X-11 is important in order to evaluate the relative merits of the two methods. Second, as George Box has frequently pointed out, we believe that the most rapid progress can be achieved if we iterate between theory and practice. Since we have an available model based theory for seasonal adjustment, an important next step is to apply that technique to real data. At the beginning of this project we expected to find that the current modeling methods were too narrow to apply to all of the Census Bureau series and that these modeling techniques would have to be expanded and their implications for seasonal adjustment explored. In addition, we expected to learn more about how the model based approach fits into the idea of seasonal adjustment so that improvements can be made. Finally, we felt that it was important to demonstrate that a model based approach is feasible for the Census Bureau. Even though a model based approach is expensive in terms of the time required to model each time series, we believe that the advantages far outweight the expenses. Modeling the data is always a good idea. It forces the model builder to learn more about the data so that the data is seen in a different perspective. Model building frequently provides new insights about a data set that can be helpful in whatever the data is to be used for. We have a bias against automatic statistical methods because too often an automatic method allows the user to do something without thinking; whenever this happens we are likely to lose some information that may be important. We believe that one of the biggest advantages to a model based seasonal adjustment technique is that it is not automatic and in contrast to an automatic method it forces the user to think about the data.

We next outline the structure of the remainder of this report. In part 1 we provide the theoretical details behind the ARIMA model based seasonal adjustment method used in the empirical study. In part 2 we describe the ways in which we have extended the ARIMA time series models to more appropriately model some of the series that were considered in the study. In part 3 we report the results of the empirical study. Finally, in part 4 we discuss a number of issues related to seasonal adjustment that arose during the year. We have written each part so that they are self contained as much as possible. In particular, the equation numbers are relevant only within each individual part.

Part 1: An ARIMA Model Based Approach to Seasonal Adjustment

I. Introduction

Business and economic time series frequently exhibit seasonality; this may be described as regular periodic fluctuations which reoccur with about the same intensity each year. Many people argue that seasonality should be removed from economic time series so that the underlying trend is more clearly discernable. As a result of this belief, a number of procedures to seasonally adjust data have been developed, the most widely used is the Census X-11 procedure described in Shiskin, Young and Musgrave (1967). The X-11 program can be viewed as an emperically based method developed over many years. The purpose of this paper is to develop a model based approach to seasonal adjustment based in part upon the years of experience implicit in the X-11 procedure.

It is assumed that an observable time series at time t, Z_t, can be represented as

$$Z_t = S_t + T_t + N_t \tag{1.1}$$

where S_t, T_t and N_t are mutually independent seasonal, trend and noise components. It may be the case that a more accurate representation for Z_t would be the product of S_t, T_t and N_t. In this situation, however, the model (1.1) would be appropriate for the logarithms of the original series. We shall also assume that Z_t follows the multiplicative ARIMA model. (Box and Jenkins, 1970)

$$\phi(B)\,\Phi(B^s)Z_t = \theta(B)\;\Theta(B^s)a_t \tag{1.2}$$

where B is the backshift operator, $BZ_t = Z_{t-1}$, $\phi(B)$ is a polynomial in B of degree p and $\Phi(B^s)$ is a polynomial in B^s of degree P both having their zeros on or outside the unit circle, $\theta(B)$ is a polynomial in B of degree q and $\Theta(B^s)$ is a polynomial in B^s of degree Q both having their zeros outside the unit circle, $\phi(B)\;\;\Phi(B^s)$ and $\theta(B)\;\;\Theta(B^s)$ have no common zeros, s is the seasonal period and the a_t's are independent and identically distributed as $N(0,\sigma_a^2)$. In what follows we shall denote $\phi(B)\ \Phi(B^s)$ by $\phi^*(B)$ and $\theta(B)\ \Theta(B^s)$ by $\theta^*(B)$. We assume in this paper that the parameters in (1.2) are known. The reason for restricting Z_t to be generated by an ARIMA model is that, Box, Hillmer and Tiao (1978) have argued that the class of ARIMA models are flexible enough to describe the behavior of many actual economic series and that ARIMA models have been used to successfully model a wide variety of time series data. In addition, Box and Jenkins (1970) have described methods to build ARIMA models from actual data. In practice there are situations where ARIMA models may not be flexable enough to adequately approximate a particular data set, for example a set of data may be affected by a strike. However, in these situations ARIMA models can frequently be appropriately modified to better approximate reality, for instance intervention analysis, Box and Tiao (1976), can be used to allow for strikes.

Based upon (1.1) and (1.2), we propose a procedure to estimate S_t and T_t uniquely. Properties of the procedure are explored. The procedure is illustrated on actual time series and the results are compared to those obtained by the Census X-11 method.

2. Properties of Seasonal and Trend Components

If in (1.1) the stochastic structures of S_t, T_t and N_t are known then estimates of S_t and T_t can be easily obtained (see Whittle, (1963) and Cleveland, (1972). In practice, however, neither S_t nor T_t are observable so that it is impossible to completely specify their structures. In contrast, since the Z_t's are observable, the stochastic structure of this component can be accurately determined. It is therefore reasonable to expect that the known stochastic structure of Z_t will at least partially determine the stochastic structures of S_t and T_t. This idea is more fully developed in section 3, however, we first investigate the properties that we expect the seasonal component and trend components to have.

It is well known that the Census X-11 procedure may be approximated by a linear filter, for instance see Young (1960) and Wallis (1974). One important feature of the filter weights for both the trend and the seasonal components implicit in the X-11 method is that the weights applied to more removed observations from the current time period decrease. This feature was probably incorporated into the X-11 program because many series have both stochastic seasonal and stochastic trend components. In other words, the trend and seasonal components tend to change over time so that the information about the current trend or seasonal is contained in the values of Z_t close to current time. Therefore, in developing a seasonal adjustment procedure we must allow for stochastic trend and stochastic seasonal components.

2.1 Stochastic Trend

We shall assume that the trend component, T_t, follows a model in the ARIMA class.

$$\phi_T(B)T_t = \eta(B)c_t \tag{2.1}$$

where $\phi_T(B)$ and $\eta(B)$ are polynomials in B and c_t are i.i.d. $N(0, \sigma_c^2)$. To allow for a stochastic trend component it is required that T_t be a nonstationary model or equivalently that $\phi_T(B)$ have zeros on the unit circle. Box and Jenkins (1970) have shown that if $\phi_T(B) = (1-B)$ the forecast function of (2.1) is an updated level and if $\phi_T(B) = (1-B)^2$ the forecast function of (2.1) is a first order polynomial whose level and slope are updated each period. Furthermore, it is well known that realizations of nonstationary time series wander through time with no fixed mean level.

We next consider the trend component from the frequency domain. For stationary time series, the spectral density function of a trend component should be large for the low frequencies and relatively smaller for the high frequencies. Now the spectral density function of (2.1) if $\phi_T(B) = \gamma(B)(1-B)^d$ is strictly speaking not defined, however, we can define a pseudo spectral density function (p.s.d.f.) for (2.1) by

$$f_T(w) = \sigma_c^2 \frac{\eta(e^{iw})\,\eta(e^{-iw})}{\phi_T(e^{iw})\,\phi_T(e^{-iw})} \qquad (2.2)$$

Now the p.s.d.f. (2.2) is infinite for $w = o$ and very large for small w. This is consistent with what could be viewed as a stochastic trend component.

2.2- Stochastic Seasonal

Initially, it is more difficult to specify a mathematical model to describe the seasonal component than to describe the trend component. However, judging from considerations in the X-11 program it is evident that the seasonal component should (i) be capable of evolving over time and (ii) be such that for an additive model, the sum of any s consecutive seasonal components should be close to zero. It is again assumed that the seasonal component is generated by an ARIMA model. In particular, we shall show that the model

$$(1 + B + \ldots + B^{s-1})\, S_t = U(B)S_t = \Psi(B)b_t \qquad (2.3)$$

where $\Psi(B)$ is a polynomial in B and b_t are iid. $N(0,\sigma_b^2)$ satisfies the requirements (i) and (ii).

Since the polynomial U(B) has all of its zeros on the unit circle, (2.3) is the model for a nonstationary time series and the seasonal component will evolve over time. Also, $E[U(B)S_t] = E[\Psi(B)b_t] = 0$; consequently the expected value of the sum of any s consecutive seasonal components is zero. If in addition the variance of $U(B)S_t$ is relatively small, then requirement (ii) will be satisfied by (2.3).

It is also informative to consider the psdf, $f_s(w)$, of the model in (2.3)

$$f_s(w) = \sigma_b^2 \frac{\Psi(e^{iw})\Psi(e^{-iw})}{U(e^{iw})U(e^{-iw})} \ . \qquad (2.4)$$

It can be shown that $f_s(w)$ has the following properties: (i) $f_s(w)$ is infinite at the seasonal frequencies $w = \frac{2k\pi}{s}$ for $k = 1, \ldots, [\frac{s}{2}]$ where $[x]$ denotes the greatest interger less than or equal to x. (ii) $f_s(w)$ has relative minimum near the frequencies $w = 0$ and $w = \frac{(2k-1)\pi}{s}$ for $k = 2, \ldots, [\frac{s}{2}]$. Therefore, the p.s.d.f of (2.3) has infinite power at the seasonal frequencies and relatively small power away from the seasonal frequencies.

Based upon the preceding discussion we have the following requirements for the p.s.d.f, $f_s(w)$ of a stochastic seasonal component. (i) $f_s(w)$ is infinite at the $m = [\frac{s}{2}]$ seasonal frequencies $w = \frac{2k\pi}{s}$ for $k = 1, \ldots, m$. (ii) $f_s(w)$ has exactly m relative minimum near the frequencies $w = 0$ and $w = \frac{(2k-1)\pi}{s}$ for $k = 2, \ldots, [\frac{s}{2}]$. Note that (ii) guarantees that $f_s(w)$ will exhibit a "smooth" behavior around each relative minimum. Therefore, we desire that $f_s(w)$ exhibit monotonically decreasing behavior as w moves from one seasonal frequency to the local minimum and the exhibit monotonically increasing behavior as w approaches the next seasonal frequency.

3. Model Based Seasonal Adjustment

Decomposition Weights for Known Component Models

In what follows, we assume the additive structure (1.1) and that the observable Z_t's follow the ARIMA model (1.2). In addition, the unobservable seasonal component, S_t, follows the ARIMA model

$$\phi_S(B)S_t = \psi(B)b_t, \tag{3.1}$$

the unobservable trend component, T_t, follows the ARIMA model

$$\phi_T(B)T_t = \eta(B)c_t, \tag{3.2}$$

and the unobservable noise component N_t follows the MA model

$$N_t = \alpha(B)d_t\ . \tag{3.3}$$

Then Cleveland and Tiao (1976) show that the optimal estimates of the seasonal and trend components at time t are respectively

$\hat{S}_t = \sum_{j=-\infty}^{\infty} W_j Z_{t-j} = W(B)Z_t$ and $\hat{T}_t = \sum_{j=-\infty}^{\infty} h_j Z_{t-j} = h(B)Z_t$. For values of t near the middle of the observable data, the weight functions W(B) and h(B) are

$$W(B) = \frac{\sigma_b^2}{\sigma_a^2} \quad \frac{\phi_T(B)\psi(B)\phi_T(F)\psi(F)}{\theta^*(B)\theta^*(F)} \tag{3.4}$$

and

$$h(B) = \frac{\sigma_c^2}{\sigma_a^2} \quad \frac{\phi_S(B)\eta(B)\phi_S(F)\eta(F)}{\theta^*(B)\theta^*(F)} \quad . \tag{3.5}$$

Also, Cleveland (1972) and Bell (1980) have shown that the asympotic weight functions (3.4) and (3.5) can be applied near the ends of the observed time series by obtaining minimum mean squarred error forecasts of the future and past of the Z_t series and using the forecasts as if they were actual observations in the asympotic formula.

Because in practice the S_t, T_t and N_t series are unobservable, it is usually unrealistic to assume that the models (3.1) - (3.3) are known as a result, the generating functions (3.4) and (3.5) cannot be obtained and the estimates $\hat{S}_t$ and $\hat{T}_t$ cannot be calculated. We can, however, get an accurate estimate of the model (1.2) from the observable Z_t series. Consequently, it is of interest to investigate to what extent a known model for Z_t will determine the models for the component series.

3.1 Restrictions upon the Component Models

Now (1.1) and (1.2) imply that $\theta^*(B)a_t = \phi(B)\Phi(B^s)S_t + \phi(B)\Phi(B^s)T_t + \phi(B)\Phi(B^s)N_t$. By taking the covariance generating functions of both sides of this equation it follows that $\sigma_a^2\theta^*(B)\theta^*(F) = \sigma_b^2 \dfrac{\phi^*(B)\psi(B)\phi^*(F)\psi(F)}{\phi_S(B)\phi_S(F)} +$

$$\sigma_c^2 \quad \frac{\phi^*(B)\eta(B) \ \phi^*(F)\eta(F)}{\phi_T(B) \ \phi_T(F)} + \sigma_d^2 \ \phi^*(B)\alpha(B) \ \phi^*(F)\alpha(F) \tag{3.6}$$

We assume that $\phi_S(B)$ and $\phi_T(B)$ have no common zeros since in practice the only zeros that we would expect to be common to $\phi_S(B)$ and $\phi_T(B)$ would lie on the unit circle and Pierce (1979) gives reasons to rule these out. In this case it follows from (3.6) that $\phi^*(B) = \phi_S(B)\ \phi_T(B)$. If $\phi_S(B)\ \phi_T(B)$ does not include all of the zeros of $\phi^*(B)$ then $\theta^*(B)$ will have at least one zero in common with $\phi^*(B)$ violating an assumtion in section 1. Conversely, let X be a zero of $\phi_S(B)\ \phi_T(B)$ but not of $\phi_*(B)$. Then (3.6) implies

$$\sigma_a^2\theta^*(B)\theta^*(F)\phi_T(B)\phi_S(B)\phi_T(F)\phi_S(F) = \sigma_b^2\phi^*(B)\phi_T(B)\Psi\ (B)\phi^*(F)\phi_T(F)\Psi(F)$$
$$+ \sigma_c^2\phi^*(B)\phi_S(B)\eta(B)\phi^*(F)\phi_S(F)\eta(F) + \sigma_d^2\phi^*(B)\phi_T(B)\phi_S(B)\phi^*(F)\phi_T(F)\phi_S(F) \tag{3.7}$$

Now if X is a zero of $\phi_S(B)$ and by assumption not a zero of $\phi_T(B)$ then substituting X in (3.7) implies that $\Psi(X) = 0$ which contradicts the assumtion that $\phi_S(B)$ and $\Psi(B)$ have no common zeros. A similiar argument can be made if X is a zero of $\phi_T(B)$. Therefore, it is evident that given the model (1.2) for Z_t that the models for S_t and T_t are restricted so that the product of their autoregressive polynomials, $\phi_S(B)\phi_T(B)$, is equal to $\overset{*}{\phi}(B)$.

Therefore, (3.6) reduces to

$$\sigma_a^2\theta^*(B)\theta^*(F) = \sigma_b^2\phi_T(B)\Psi(B)\phi_T(F)\Psi(F) + \sigma_c^2\phi_S(B)\eta(B)\phi_S(F)\eta(F)$$
$$+ \sigma_d^2\phi_T(B)\phi_S(B)\alpha(B)\phi_T(F)\phi_S(F)\alpha(F). \tag{3.8}$$

3.2 A Particular Model for Z_t

To facilitate the developments that follow we first consider the case where Z_t follows the particular model

$$(1-B)(1-B^s)Z_t = (1-\theta_1B)(1-\theta_2B^s)a_t. \tag{3.9}$$

We know that the product of the autoregressive polynomials of S_t and T_t must be equal to $(1-B)(1-B^s)$. Therefore, a particular factorization of $(1-B)(1-B^s)$ must be chosen. Based upon the discussion in sections 2.1 and 2.2 we take the model for the stochastic trend to be

$$(1-B)^2 T_t = \eta(B)c_t \tag{3.10}$$

and the model for the stochastic seasonal to be

$$U(B)S_t = \Psi(B)b_t \; . \tag{3.11}$$

It follows from (3.6) that

$$\sigma_a^2(1-\theta_1 B)(1-\theta_2 B^s)(1-\theta_1 F)(1-\theta_2 F^s) = \sigma_b^2(1-B)^2\Psi(B)(1-F)^2\Psi(F)$$
$$+ \sigma_c^2\, U(B)\eta(B)U(F)\eta(F) + \sigma_d^2(1-B)(1-B^s)\alpha(B)(1-F)(1-F^s)\alpha(F) \tag{3.12}$$

Observe that for $\eta(B)$, $\Psi(B)$ and $\alpha(B)$ to be consistent with the model (3.9) these polynomials must be chosen so that they satisfy equation (3.12). Any polynomials that satisfy (3.12) will be called <u>acceptable polynomials</u> and the resulting decomposition will be called an <u>acceptable decomposition</u>. Note that the largest power of B on the left hand side of (3.12) is $s+1$, thus it follows that in general the degree of $\Psi(B)$ will be $s-1+k$, the degree of $\eta(B)$ will be $2+k$ and the degree of $\alpha(B)$ will be k. Also, if $k>0$ then at least two of the polynomials $(1-B)^2\Psi(B)$, $U(B)\eta(B)$ and $(1-B)(1-B^s)\alpha(B)$ must have orders larger than $s+1$ and furthermore the polynomials $\Psi(B)$, $\eta(B)$ and $\alpha(B)$ must be chosen in a manner so that the powers of B larger than $s+1$ on the right hand side of (3.12) exactly cancel. Therefore, even though strickly speaking $k>0$ is possible, this case seems unrealistic and we shall require the order of $\Psi(B)$ to be less than or equal to $s-1$, the order of $\eta(B)$ to be less than or equal to 2, and the order of $\alpha(B)$ equal to zero for the particular model (3.9).

If both sides of (3.12) are divided by $(1-B)(1-B^s)(1-F)(1-F^s)$ we obtain

$$\frac{\sigma_a^2(1-\theta_1 B)(1-\theta_2 B^s)(1-\theta_1 F)(1-\theta_2 F^s)}{(1-B)(1-B^s)(1-F)(1-F^s)} = \frac{\sigma_b^2\Psi(B)\Psi(F)}{U(B)U(F)} +$$

$$\frac{\sigma_c^2\eta(B)\eta(F)}{(1-B)^2(1-F)^2} + \sigma_d^2 \; . \tag{3.13}$$

In equation (3.15) the left hand side is known and we wish to determine the elements of the right hand side. One way to proceed is to do a partial fractions decomposition of the left hand side of (3.15). For instance, a unique partial fractions expansion is

$$\frac{\sigma_a^2(1-\theta_1 B)(1-\theta_2 B^s)(1-\theta_1 F)(1-\theta_2 F^s)}{(1-B)(1-B^s)(1-F)(1-F^s)} =$$

$$\frac{\psi^*(B,F)}{U(B)U(F)} + \frac{\eta^*(B,F)}{(1-B)^2(1-F)^2} + \bar{\epsilon}_3 \tag{3.16}$$

where expressions for $\psi^*(B,F)$, $\eta^*(B,F)$ and ϵ_3 are given in the appendix. Let $f_s^*(w) = \dfrac{\psi^*(e^{iw}, e^{-iw})}{U(e^{iw})U(e^{-iw})}$ and $f_T^*(w) = \dfrac{\eta^*(e^{iw}, e^{-iw})}{(1-e^{iw})^2(1-e^{-iw})^2}$. Then we note the following: (i) For the partial fractions decomposition to correspond to stochastic models for S_t, T_t and N_t and hence an acceptable decomposition, it is required that $f_s^*(w) \geq 0$, $f_T^*(w) \geq 0$ for $0 \leq w \leq \pi$ and $\epsilon_3 \geq 0$. (ii) Other possible decompositions can be derived from (3.16) by adding constants to any or all of $\dfrac{\psi^*(B,F)}{U(B)U(F)}$, $\dfrac{\eta^*(B,F)}{(1-B)^2(1-F)^2}$ or ϵ_3 subject to the restriction that the net amount added to all three expressions be zero. Consequently, if an initial partial fractions decomposition yields values of $f_s^*(w)$ or $f_T^*(w)$ or ϵ_3 which are unacceptable then it may be possible to add constants to these elements so that they are all positive. In particular, if we let $\epsilon_1 = \min_{0 \leq w \leq \pi} f_s^*(w)$ and $\epsilon_2 = \min_{0 \leq w < \pi} f_T^*(w)$ then it follows that it is possible to create at least one acceptable decomposition from an initial partial fractions decomposition if and only if $\epsilon_1 + \epsilon_2 + \epsilon_3 \geq 0$. (iii) If $\epsilon_1 + \epsilon_2 + \epsilon_3 > 0$ then it follows that there are an infinite number of ways to modify (3.16) and obtain acceptable decompositions. In this case the known model for Z_t does not specify a unique decomposition.

In summary, if we are given a model for Z_t then we can perform an initial partial fractions expansion and find ϵ_1, ϵ_2 and ϵ_3. If $\epsilon_1 + \epsilon_2 + \epsilon_3 < 0$ then there does not exist an acceptable decomposition which is consistent with the known model for Z_t and the restrictions we have imposed on S_t, T_t and N_t. If $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$ then there is a unique acceptable decomposition. If $\epsilon_1 + \epsilon_2 + \epsilon_3 > 0$ then there are an infinite number of acceptable decompositions consistent with the given model for Z_t.

Properties of the Seasonal and Trend Derived from the Partial Fractions

From the previous discussion it is evident that the partial fractions expansion (3.16) will determine the general shape of the pseudo spectral density functions $f_S^*(w)$ and $f_T^*(w)$. It is of interest to examine if the partial fractions approach leads to psdf's which are similiar to those of the stochastic trend and stochastic seasonal discussed in sections 2.1 and 2.2. The details involved for the development of this portion of the paper are included in the appendix. The main conclusions follow.

The Case $\theta_1 = 1$

We first discuss the case where $\theta_1 = 1$ so that the model (3.9) for Z_t reduces to

$$(1-B^s)Z_t = (1-\theta_2 B^s)a_t. \tag{3.17}$$

It is shown in the appendix that the following are true based upon the model (3.17) for Z_t. (i) The trend psdf is infinite at $w = 0$ and monotonically decreasing for any θ_2 in the range $-1 \leq \theta_2 < 1$. (ii) The seasonal p.s.d.f. is infinite at the seasonal frequencies and has exactly $m = [\frac{s}{2}]$ relative minimum at $w = 0$ and near $w = \frac{(2k-1)\pi}{s}$ for $k = 2, \ldots, m$. (iii) It is possible to derive an acceptable decomposition when Z_t follows (3.17) as long as $\theta_2 \geq \dfrac{-(10s^2-4)-(96s^4-96s^2)^{\frac{1}{2}}}{2(s^2+2)}$.

Values of the minimum possible θ_2 for selected values of s are given in the following tabulation.

s	2	4	6	8	10	12	∞
min. θ_2	−.1716	−.1170	−.1080	−.1049	−.1035	−.1027	−.1010

Therefore, there are values of θ_2 for which the model (3.17) is not consistent with an additive decomposition as we have defined it, however a value of $\theta_2 > -.1010$ will always lead to an acceptable decomposition with intuitively pleasing p.s.d.f's.

$(0,1,1)\ (0,1,1)^s$ Model

In the situation where Z_t follows the model (3.9) the following results are derived in the appendix. (i) The p.s.d.f. of the trend is infinite at $w = 0$ and is monotonically decreasing for θ_1 and θ_2 satisfying the inequality

$$\frac{1}{6}(1-\theta_1)^2(1-\theta_2)^2+2\theta_2(1-\theta_1)^2+\frac{1}{6}s^2\ (1-\theta_2)^2(1+\theta_1)^2+4(1-\theta_2)^2(1+\theta_1{}^2)\geq 0. \quad (3.18)$$

Furthermore, if $\theta_2 > -.1010$ then any value of θ_1 in the range $-1 \leq \theta_1 < 1$ will satisfy (3.18) for any s. (ii) The p.s.d. of the seasonal is infinite at the seasonal frequencies and has a relative minimum at $w = 0$. (iii) If we require that the p.s.d. of the trend is monotonically decreasing then the p.s.d. of the seasonal has unique relative minimum near $w = \frac{(2k-1)\pi}{s}$ for $k = 2, \ldots, m$.

In summary, for the two models of Z_t considered, the shape of the p.s.d.'s behave in a reasonable manner for a large range of the possible parameters. However, in both examples we conclude that there are values of θ_1 and θ_2 corresponding to the model for Z_t which are not consistent with decomposing the Z_t series as we have defined the decomposition.

3.3 A Canonical Decomposition

In this section we assume that an admissable decomposition corresponding to the model for Z_t exists. In the absence of prior knowledge about the stochastic structure of the trend and seasonal components, all of the information in the known model of Z_t, (1.2), about S_t and T_t is embodied in (3.8). However, as indicated previously this information is not sufficient to uniquely determine the models for S_t and T_t. Therefore, we must rely upon additional information or another principle to determine these models.

If the observed series follows (1.2) than $\phi_S(B).\phi_T(B) = \phi^*(B)$ where $\phi_S(B)$ is the seasonal autoregressive polynomial and $\phi_T(B)$ is the trend autoregressive polynomial and

$$\frac{\sigma_a^2\,\theta^*(B)\,\theta^*(F)}{\phi^*(B)\,\phi^*(F)} = \frac{\theta_r^*(B,F)}{\phi^*(B)\,\phi^*(F)} + \alpha^*(B,F) \qquad (3.19)$$

where $\alpha^*(B,F)$ and $\theta_r(B,F)$ are respectively quotient and remainder where the numerator of the left hand side of (3.19) is divided by the denominator. Then the right hand side of (3.19) can be expanded by partial fractions as follows.

$$\frac{\sigma_a^2\,\theta^*(B)\,\theta^*(F)}{\phi^*(B)\,\phi^*(F)} = \frac{\psi^*(B,F)}{\phi_S(B)\,\phi_S(F)} + \frac{\eta^*(B,F)}{\phi_T(B)\,\phi_T(F)} + \alpha^*(B,F). \qquad (3.20)$$

If we let $\varepsilon_1 = \min\limits_{0 \le w \le \pi} \dfrac{\psi^*(e^{iw}, e^{-iw})}{\phi_S(e^{iw})\,\phi_S(e^{-iw})}$, $\varepsilon_2 = \min\limits_{0 \le w \le \pi} \dfrac{\eta^*(e^{iw}, e^{-iw})}{\phi_T(e^{iw})\,\phi_T(e^{-iw})}$,

$\varepsilon_3 = \min\limits_{0 \le w \le \pi} \alpha^*(e^{iw}, e^{-iw})$ then it follows from the developments in section 3.1 that for an acceptable decomposition to exist it is necessary and sufficient that $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \ge 0$. In what follows we assume that an acceptable decomposition exists, then other acceptable decompositions can be derived from (3.20) by adding constants to the components subject to the restrictions that for all $0 \leq w \leq \pi$,

$$\frac{\psi(e^{iw}, e^{-iw})}{\phi_S(e^{iw})\phi_S(e^{-iw})} + \xi_1 \geq 0, \quad \frac{\eta(e^{iw}, e^{-iw})}{\phi_T(e^{iw})\ \phi_T(e^{-iw})} + \xi_2 \geq 0, \quad \alpha(e^{iw}, e^{-iw}) + \xi_3 \geq 0 \text{ and}$$

$\xi_1 + \xi_2 + \xi_3 = 0$. Equivalently we require that $-\epsilon_1 \leq \xi_1 \leq \epsilon_2 + \epsilon_3$, $-\epsilon_2 \leq \xi_2 \leq \epsilon_1 + \epsilon_3$, $-\epsilon_3 \leq \xi_3 \leq \epsilon_1 + \epsilon_2$, and $\xi_1 + \xi_2 + \xi_3 = 0$.

In the general situation it is evident that an infinite number of acceptable decompositions corresponding to (1.2) may exist and in order to perform the seasonal adjustment we must pick one decomposition based upon information other than that contained in the observable series Z_t. Intuitively, it seems reasonable to extract as much white noise as possible from the seasonal and trend components subject to the restrictions in (3.8). This will maximize the error variance σ_d^2 and yield the most deterministic seasonal and trend components. Therefore, we define the <u>canonical decomposition</u> as the decomposition which maximizes σ_d^2 subject to the restrictions in (3.8). Some properties of this decomposition are now discussed.

(i) From (3.20) and the restrictions upon ξ_1, ξ_2, and ξ_3 it is evident that every admissible combination of ξ_1, ξ_2 and ξ_3 defines a unique acceptable decomposition. Now from (3.20) it follows that

$$\sigma_a^2\theta^*(B)\theta^*(F) = \psi^*(B,F)\phi_T(B)\phi_T(F) + \eta^*(B,F)\phi_S(B)\phi_S(F) + \alpha^*(B,F)\phi^*(B)\phi^*(F). \quad (3.21)$$

Using a result of Hannan (1970, p. 137) we have that

$$\ln \sigma_d^2(\xi_3) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln \ f_3(w,\xi_3)\ dw \qquad (3.22)$$

where $f_3(w, \xi_3) = [\alpha^*(e^{iw}, e^{-iw}) + \xi_3]\ |\phi^*(e^{iw})|^2$. Now $f_3(w,\xi_3)$ does not depend on ξ_3 if $\phi^*(e^{iw}) = 0$ and is otherwise strictly increasing in ξ_3. Thus it follows that σ_d^2 is maximized when $\xi_3 = \epsilon_1 + \epsilon_2$. However, from the restriction that $\xi_1 + \xi_2 + \xi_3 = 0$ and the restrictions on ξ_1 and ξ_2 we have that for the canonical decomposition $\xi_1 = -\epsilon_1$ and $\xi_2 = -\epsilon_2$. It follows that the canonical decomposition is unique.

(ii) The canonical decomposition minimizes σ_b^2, the innovation variance of the shocks driving the seasonal component, and σ_c^2, the innovation variance of the shocks driving the trend component. To see this result note that

$$\ln \sigma_b^2(\xi_1) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln [f_1(w, \xi_1)]dw \qquad (3.23)$$

and

$$\ln \sigma_c^2(\xi_2) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln [f_2(w, \xi_2)]dw \qquad (3.24)$$

where

$$f_1(w, \xi_1) = \sigma_a^2|\theta^*(e^{iw})|^2 - \eta^*(e^{iw}, e^{-iw})|\phi_s(e^{iw})|^2 - \alpha^*(e^{iw}, e^{-iw})|\phi^*(e^{iw})|^2$$

$+ \xi_1|\phi_T(e^{iw})|^2$ and $f_2(w, \xi_2) = \sigma_a^2|\theta^*(e^{iw})|^2 - \Psi^*(e^{iw}, e^{-iw})|\phi_T(e^{iw})|^2$ $- \alpha^*(e^{iw}, e^{-iw})|\phi^*(e^{iw})|^2 + \xi_2|\phi_s(e^{iw})|^2$. Since $f_1(w, \xi_1)$ either does not depend on ξ_1 or is strictly decreasing as ξ_1 decreases and a similiar observation holds for $f_2(w, \xi_2)$ it is clear that σ_b^2 is minimized when $\xi_1 = -\epsilon_1$ and σ_c^2 is minimized when $\xi_2 = -\epsilon_2$. Since these values correspond to the canonical decomposition the stated result is true. These particular properties of the canonical decomposition are intuitively pleasing since the randonness is S_t arises from the sequence of b_t's and the randonness in T_t arises from the sequence of c_t's. Thus minimizing σ_b^2 and σ_c^2 makes the seasonal and trend components as deterministic as possible while remaining consistent with the information in the observable Z_t series.

(iii) We let $\Psi(B)$ denote the moving average polynonical of the seasonal and $\eta(B)$ the moving average polynonical of the trend in the cannonical decomposition. Then from (3.20) we have that $\Psi(B)\Psi(F) = \Psi^*(B,F) - \epsilon_1\phi_s(B)\phi_s(F)$ and $\eta(B)\eta(F) = \eta^*(B,F) - \epsilon_2\phi_T(B)\phi_T(F)$. From the definitions of ϵ_1 and ϵ_2 it is evident that both $\Psi(B)$ and $\eta(B)$ have at least one zero on the unit circle. Thus for the canonical decomposition both the model for S_t and the model for T_t are not invertible. When S_t and T_t are stationary it is known that for large n the eigenvalues of the covariance matrix of S_t approach $2\pi f_s(-\pi + 2\pi j^\lambda/n, -\epsilon_1)$ and those of T_t approach $2\pi f_T(-\pi+2\pi j^\lambda/n,-\epsilon_2)$.

Then, asympotically, for both S_t and T_t at least one of the eigenvalues will approach zero implying a linear dependence in each component.

3.4 More General ARIMA Models

The developments in section 3.1 were based on the assumption that the model for Z_t was (3.9). However, the properties of the canonical decomposition were based on the general ARIMA model (1.2). A close look at the derivations in the previous sections reveals that as long as an acceptable decomposition is achievable, there is no restrictions upon the form of the moving average polynomial $\theta^*(B)$ in the model based method. In addition, it is an easy manner to enlarge the possible autoregressive polynomials over the $(1-B)$ $(1-B^s)$ polynomials considered in section 3.1. All that is required is to chose a particular factorization of the autoregressive polynomial for the trend and seasonal components. In particular a model for Z_t which has the autoregressive operator $\phi(B)$ $(1-B^s)$ should be factored so that $\phi(B)$ $(1-B)$ is the trend polynomial and $U(B)$ is the seasonal polynomial. With these extensions the model based seasonal adjustment techniques that have been described should cover a wide range of actual time series.

Appendix

In this appendix we develop some properties of the shape of the p.s.d.f of the trend and seasonal components based upon a partial fractions expansion for the models (3.9) and (3.17). Without loss of generality we take $\sigma_a^2 = 1$. After quite a bit of algebra it can be shown that the partial fractions expansion based upon (3.9) is

$$\frac{(1-\theta_1 B)(1-\theta_2 B^S)(1-\theta_1 F)(1-\theta_2 F^S)}{(1-B)(1-B^S)(1-F)(1-F^S)} = \theta_1\theta_2 +$$

$$\frac{1}{(1-B)^2(1-F)^2}\left\{\frac{(1-\theta_1)^2(1-\theta_2)^2}{S^2} + [\theta_2(1-\theta_1) + \frac{\theta_1(1-\theta_2)}{S^2} + \frac{(S+1)(S-1)}{12S^2}\quad(1-\theta_1)^2(1-\theta_2)^2]\right.$$

$$\left.(1+B)(1+F)\right\} + \frac{(1-\theta_2)^2}{U(B)U(F)}\left\{[\frac{(S+1)(S)(S-1)}{6S^2}\quad\theta_1 + \{\frac{(S+1)(S-1)(S+1)(S)(S-1)}{72S^2}\right.$$

$$- \frac{(S+2)(S+1)(S)(S-1)(S-2)}{120S^2}\}\quad(1-\theta_1)^2] + [\frac{S(S-1)(S-2)}{6S^2}\quad\theta_1 + \{\frac{(S+1)(S-1)(S)(S-1)(S-2)}{72S^2}$$

$$- \frac{(S+1)(S)(S-1)(S-2)(S-3)}{120S^2}\}\ (1-\theta_1)^2\ ]\quad(B+F)\quad + \ldots + [\frac{(4)(3)(2)}{6S^2}\quad\theta_1 +$$

$$\{\frac{(S+1)(S-1)(4)(3)(2)}{72S^2} - \frac{(5)(4)(3)(2)(1)}{120S^2}\}\ (1-\theta_1)^2\ ]\quad(B^{S-3} + F^{S-3}) +$$

$$\left.[\frac{(3)(2)(1)}{6S^2}\quad\theta_1 + \frac{(S+1)(S-1)(3)(2)(1)}{72S^2}\quad(1-\theta_1)^2]\ (B^{S-2} + F^{S-2})\right\} \qquad (A.1)$$

Also it can be shown that the partial fractions decomposition based upon the model (3.17) is

$$\frac{(1-\theta_2 B^S)(1-\theta_2 F^S)}{(1-B^S)(1-F^S)} = \theta_2 + \frac{\frac{1}{S^2}(1-\theta_2)^2}{(1-B)(1-F)} + \frac{(1-\theta_2)^2}{U(B)U(F)}\left\{\frac{(S+1)(S)(S-1)}{6S^2} +\right.$$

$$\left.\frac{(S)(S-1)(S-2)}{6S^2}\quad(B+F) + \ldots + \frac{(3)(2)(1)}{6S^2}\quad(B^{S-2} + F^{S-2})\right\} \qquad (A.2)$$

Now consider first the expansion (A.2) corresponding to the model (3.17). It is evident that the trend p.s.d.f is infinite for $w=0$ and monotonically decreasing in w since the term $\frac{1}{S^2}(1-\theta_2)^2 \geq 0$ for any possible θ_2. Also, we conjecture that the p.s.d.f of the seasonal component has its minimum at $w=0$. This conjecture is difficult to verify analytically but has been verified by numerical means for S from 3 to 20. Consequently, we have that $\epsilon_1 = \frac{1}{4S^2}(1-\theta_2)^2$, $\epsilon_2 = \frac{(S+1)(S-1)}{12S^2}(1-\theta_2)^2$ and $\epsilon_3 = \theta_2$ so that for an acceptable decomposition to exist it is required that $\epsilon_1 + \epsilon_2 + \epsilon_3 = (S^2+2)\theta_2^{\,2} + (10S^2-4)\theta_2 + (S^2+2) \geq 0$ or equivalently

$$\theta_2 \geq \frac{-(10S^2-4) + (96S^4 - 96S^2)^{\frac{1}{2}}}{2(S^2+2)} \quad . \qquad \text{(A.3)}$$

In summary, for the model (3.17) any θ_2 which satisfies (A.3) will via partial fractions lead to an acceptable decomposition with a trend p.s.d.f that is infinite at $w=0$ and monotonically decreasing. The characteristics of the seasonal p.s.d.f for this model are similiar to that of model (3.9) given below. They follow from the fact that the trend p.s.d.f is monotonically decreasing.

For the model (3.9) and the corresponding partial fractions decomposition (A.1) we first consider if the trend p.s.d.f is monotonically decreasing. From (A.1) the trend p.s.d.f is

$$f_T(w) = \frac{\alpha_1}{4[1-\cos(w)]^2} + \frac{\alpha_2}{2[1-\cos(w)]} \qquad \text{(A.4)}$$

where $\alpha_1 = \frac{1}{S^2}(1-\theta_1)^2(1-\theta_2)^2$ and $\alpha_2 = \theta_2(1-\theta_1) + \frac{1}{S^2}\theta_1(1-\theta_2) + \frac{(S+1)(S-1)}{12S^2}(1-\theta_1)^2(1-\theta_2)^2$. Now for $f(w)$ to be monotonically decreasing in w we require that $f'(w) \leq 0$ for all $0 \leq w \leq \pi$. This then requires that we have $\alpha_1 + \alpha_2[1-\cos(w)] \geq 0$ for all $0 \leq w \leq \pi$. If $\alpha_2 > 0$ then this inequality is satisfied for all w and if $\alpha_2 < 0$ the extreme case is at $w = \pi$. To guarantee that the inequality is satisfied for all w we require that $\alpha_1 + 2\alpha_2 \geq 0$. Therefore, the region for which the trend p.s.d.f is monotonically decreasing is

$$\frac{1}{6}(1-\theta_1)^2(1-\theta_2)^2 + 2\theta_2(1-\theta_1)^2 + \frac{1}{6S^2}\{(1-\theta_2)^2(1+\theta_1^2) + 4(1-\theta_2)^2(1+\theta_1^2)\} \geq 0. \tag{A.5}$$

Note that the last term in (A.5) is positive for all values of θ_1 and θ_2 and it's influence decreases as S becomes large. It follows that if $\theta_2 > -.1010$ then the sum of the first two terms in (A.5) is positive and thus any $\theta_2 > -.1010$ will lead to a trend p.s.d.f which is monotonically decreasing.

Finally, we consider the shape of the seasonal p.s.d.f in the case where the trend p.s.d.f is monotonically decreasing. We first consider the p. s.d.f. of Z_t

$$f(w) = \frac{|1-\theta_1 e^{iw}|^2 . |1-\theta_2 e^{isw}|^2}{|1-e^{iw}|^2 . |1-e^{isw}|^2} .$$

Now $f(w) = f_1(w).f_2(w)$ where $f_1(w) = \dfrac{(1+\theta_1)^2 - 2\theta_1\cos(w)}{2[1-\cos(w)]}$ and $f_2(w) =$ $\dfrac{(1+\theta_2^2)-2\theta_2\cos(sw)}{2[1-\cos(sw)]}$. Furthermore, we have that

$$f_2'(w) = -\frac{\sin(w)(1-\theta_1)^2}{2[1-\cos(w)]^2} \leq 0 \text{ for } 0<w \leq \pi$$

and

$$f_2'(w) = -\frac{s\sin(sw)(1-\theta_2)^2}{2[1-\cos(sw)]^2}$$

so that the sign of $f_2'(w)$ depends upon the value of $\sin(s.w)$. Now if $s.w = (2k-1)\pi$ for $k=1,2,\ldots,m$ with $m = [\frac{S}{2}]$ then $\sin(s.w) = 0$ and $\cos(sw) = -1$ so that $f_2'(w) = 0$. This implies that there are m relative minimum of $f_2(w)$ at $w = \frac{1}{S}(2k-1)\pi$ for $k=1, \ldots, m$. In additiou,since $-\sin(sw)$ is a monotonic increasing function in $\frac{1}{S}(2k-2)\pi < w < \frac{1}{S}(2k)\pi$ for $k=1, \ldots, m$ then the value of $f_2'(w)$ is monotonically increasing in these intervals.

Next we have that $\bar{f}'(w) = f_1'(w).f_2(w) + f_1(w).f_2'(w)$ so that $f'(w) = 0$ if and only if $f_1'(w)f_2(w) = - f_1(w)f_2'(w)$. From the previous discussion it follows that $f_1'(w).f_2(w) \leq 0$ for all w and $f_1(w).f_2'(w)$ is monotonically increasing for $\frac{1}{S}(2k-2)\pi < w < \frac{1}{S}(2k)\pi$ and k=1, ..., m. Thus, there are m unique relative minimum of f(w) and f'(w) is monotonically increasing for $\frac{1}{S}(2k-2)\pi < w < \frac{1}{S}(2k)\pi$ k=1, ..., m.

Now we have that $f(w) = f_T(w) + f_s(w)$ where $f_s(w)$ is the p.s.d.f. of the seasonal and $f_T(w)$ is the p.s.d.f. of the trend. We consider the case where $f_T(0) = \infty$ and $f_T(w)$ is monotonically decreasing. We let the set $\Omega = w: \{w = \frac{1}{S} 2k\pi \quad k=0, \ldots, m\}$, then it follows that $f_s'(w) = f'(w) - f_T'(w)$ if $w \notin \Omega$ and $f_s'(w) = 0$ if and only if $f'(w) = f'_T(w)$. From the fact that $f'_T(w) \leq 0$ for all $0 < w < \pi$ and the properties of f'(w) derived above we have that $f_s(w)$ has m-1 unique relative minimum in the range $\frac{1}{S}(2k-2)\pi < w < \frac{1}{S}(2k)\pi$ k = 2, ..., m. Finally, it is an easy matter to verify directly from the form in (A.1) that $f_s'(w) = 0$ for w = 0, so that $f_s(w)$ has an additional relative minimum at w = 0.

Part 2: Extensions of ARIMA Models

A time series model builder armed with Box-Jenkins ARIMA methods discovers very quickly that ARIMA models alone are insufficient to deal with many of the Census Bureau's time series. Simple ARIMA models do not specifically allow for the possibility of outliers, trading day variation and variation due to the placement of Easter; however, a large percentage of Bureau series contain one or more of these characteristics. Consequently, if we are advocating a model based approach to seasonal adjustment, then it becomes necessary to develop methods which allow one to model time series with these characteristics. During the course of this year we have begun to develop these models. Our attempts should not be regarded as the final product but rather a first step in the right direction. Even though we frequently used some simple minded models, we have found that our efforts have appeared to be sucessful in many cases. In this section, we shall briefly describe the techniques we have developed to deal with outliers, trading day variation, and holiday variation in time series. We acknowledge that there are other types of problems that occur in Bureau series (e.g., strikes), but thus far we have not considered these. One obvious approach to many other problems is to use the intervention analysis technique developed by Box and Tiao (1975).

Outliers

When dealing with outliers in ARIMA time series models, it is important to understand the ways in which potential outliers impact the original time series. These issues can probably be best illustrated by means of a simple example. Suppose that an observed time series, Z_t, follows a random walk model

$$(1-B)Z_t = a_t. \qquad (2.1)$$

Then suppose we have an outlier at time $t = t_0$. It will be convenient to define the "pulse" variable

$$P_t^{(t_0)} = \begin{cases} 1 & \text{if } t = t_0 \\ 0 & \text{otherwise.} \end{cases} \qquad (2.2)$$

We let Z_t denote the value of the time series at time t excluding the effect of the outlier and $\tilde{Z}_t$ denote the value of the time series at time t including the effect of the outlier. Then there are at least two ways to allow for the possibility of an outlier at $t = t_0$. We can let model be

$$(1-B)Z_t = a_t \quad \text{and} \quad \tilde{Z}_t = Z_t + w \cdot P_t^{(t_0)} \tag{2.3}$$

or formally

$$\tilde{Z}_t = w \cdot P_t^{(t_0)} + \frac{a_t}{1-B} . \tag{2.4}$$

The outlier model (2.3) clearly allows for $\tilde{Z}_{t_0} = Z_{t_0} + w$ and $\tilde{Z}_t = Z_t$ if $t \neq t_0$ so that the result of (2.3) is a pertubation at time t_0 in the original time series. We shall call outliers generated in this manner "observation outliers".. Another possible model is

$$(1-B)Z_t = a_t \quad \text{and} \quad (1-B)\tilde{Z}_t = (1-B)Z_t + w \cdot P_t^{(t_0)} \tag{2.5}$$

$$(1-B)\tilde{Z}_t = w \cdot P_t^{(t_0)} + a_t \tag{2.6}$$

Now from (2.5) it follows that $\tilde{Z}_t = Z_t$ if $t < t_0$. However, $\tilde{Z}_{t_0} = Z_{t_0-1} + w \cdot P_t^{(t_0)} + a_{t_0} = Z_{t_0} + w$, $\tilde{Z}_{t_0+1} = \tilde{Z}_{t_0} + a_{t_0+1} = Z_{t_0+1} + w$ and $\tilde{Z}_{t_0+k} = Z_{t_0+k} + w$ for $k \geq 0$. Therefore, the effect of (2.5) is a pertubation of w in every value of the original time series after $t = t_0$. Since in (2.6) the outlier can be viewed as a shift in a_{t_0} we call outliers generated in this manner "innovation outliers".

The choice of the way in which we wish to model outliers should depend in part upon knowledge about the time series being modeled. However, for most cases if would seem to be unreasonable to expect an outlier to affect all of the observations subsequent to its initial impact. Thus, in the absence of any other knowledge it seems more reasonable to use the observation outlier concept. But as an illustration that the observation outlier concept should not be unversally applied, we found while modeling the time series "retail sales of variety stores" that there were problems when the observation outlier concept was applied. After some investigation we were told that at the time of the apparent outlier a large variety store chain went out of business and, as a result, many of the variety stores sales were probably transferred to the category of department stores sales. Therefore, in this particular series the innovations outlier formulation is more appropriate.

We next describe the procedure that was usually followed to identify and model outliers for the time series analyzed during this year. First an ARIMA time series model which allowed for trading day and holiday effects where appropriate was identified and the parameters were estimated. The estimated residuals, $\hat{a}_t$, were obtained and if $\hat{a}_{t_o}$ was larger than three times its estimated standard error a potential outlier at time t_o was identified. The model

$$Z_t = w \cdot P_t^{(t_o)} + N_t$$

where N_t is the previously identified ARIMA model was then fit and the estimated value of w, $\hat{w}$, was compared to its standard error. If $\hat{w}$ was larger than twice its standard error then the outlier was taken to be signifigant and left in the model; if $\hat{w}$ was smaller than twice its standard error the outlier was not included in the model.

There are at least two potential problems with the above procedure that need to be investigated further. First, using $\hat{a}_t$ to judge the timing of an outlier is appropriate if we are using the innovations outlier concept. But because for most cases we want to allow for observation outliers, it is not clear that examination of the $\hat{a}_t$'s will necessarily lead to the correct specification of t_o. For instance, an observation outlier at time t_o following model (2.1) can lead to large values of $\hat{a}_{t_o}$ and $\hat{a}_{t_o-1}$. The point is that unless we are careful we may try and include more outliers in a model than are truly present in the data or we may exclude a potential outlier from the model. This issue clearly deserves further investigation. Secondly, while the probability of observing, say, a normal random variable more than three standard errors away from its mean is small, it is not zero. Therefore, we may be fitting outliers and deciding that potential outliers are significant when in fact they were just random occurances. Again this area needs further study.

Modeling Trading Day Effects

A signifigant proportion of the Census Bureau's time series are affected by the fact that every month except February has a variable number of the different kind of days of the week. This arises in part because of the fact that individuals buying behavior and some institutional patterns are based upon the days of the week rather than the month. Time series models such as ARIMA models which attempt to describe the correlation pattern between months without allowing for these trading day effects are inadequate for widespread use at the Census Bureau. In what follows we describe the way in which we have expanded the class of ARIMA models to allow for trading day variation.

We assume that after appropriately accounting for the seasonality in a time series that the residual effect of trading day changes can be approximated by a deterministic model. We let TD_t denote the trading day variation of month t. Then TD_t should be a function of the number of distinct types of days in month t. In particular, we assume that

$$TD_t = \sum_{i=1}^{7} \beta_i \cdot X_{it} \qquad (2.7)$$

where X_{it} $i = 1, \ldots, 7$ are respectively the number of Mondays, Tuesdays, Wednesdays, Thursdays, Fridays, Saturdays and Sundays in month t and $\beta_1, \ldots, \beta_7$ are parameters. Now since (2.7) represents the trading day portion of the time series and since we are developing the models with the idea of seasonal adjustment, we want to impose the restriction that the average trading day effect over the long term is zero. Otherwise, TD_t could be viewed as including a portion of the trend. Therefore, for large n we desire that

$$0 = \sum_{t=1}^{n} TD_t = \sum_{t=1}^{n} \sum_{i=1}^{7} \beta_i X_{it} = \sum_{i=1}^{7} \beta_i \sum_{t=1}^{n} X_{it} \quad . \tag{2.8}$$

Now, we note that $\sum_{t=1}^{n} X_{1t}$ is the total number of Mondays in the n months and similarily for $\sum_{t=1}^{n} X_{it}$ i=2,...,7. In addition, for given large n the value of $\sum_{t=1}^{n} X_{it}$ will be approximately the same for each i. Therefore, from (2.8) if we want $\sum_{t=1}^{n} TD_t = 0$ then we must have $\sum_{i=1}^{7} \beta_i = 0$. Incorporating this restriction into (2.7) we may write,

$$TD_t = \sum_{i=1}^{6} \beta_i D_{it} \tag{2.9}$$

where $D_{it} = X_{it} - X_{7t}$ i=1,...,6.

Now we assume that seasonality apart from TD_t can be modeled by ARIMA models; therefore an overall model for a time series including trading day variation is

$$Z_t = TD_t + N_t \tag{2.10}$$

where N_t is an appropriately chosen ARIMA model. We note that (2.10) is a member of the class of regression models with ARIMA errors.

Now it is well known (see e.g. Pierce (1971)) that under some relatively weak restrictions as n gets large the parameter estimates of $\underset{\sim}{\beta}' = (\beta_1, \ldots, \beta_6)$ in (2.10) are approximately normally distributed and are distributed independently of the parameter estimates in the ARIMA model. Therefore, inferences about $\underset{\sim}{\beta}$ can be made separately from inferences about the ARIMA parameters and standard normal theory can be applied. However, the estimates of the parameters, $\hat{\underset{\sim}{\beta}}$, are correlated so that inferences about the parameters $\underset{\sim}{\beta}$ should be made jointly. For example, we would like to test the hypothesis H_0: $\beta_1 = \beta_2 = \ldots = \beta_6 = 0$ since if that hypothesis cannot be rejected than there is no evidence that the trading day variables are necessary in the model (2.10). Now if we let A denote the covariance matrix of $\underset{\sim}{\beta}$ then assuming H_0 is true it follows from standard normal theory that $Q = \underset{\sim}{\beta}' A^{-1} \underset{\sim}{\beta}$ has asymptotically a chi-squared distribution with 6 degrees of freedom. Consequently, this distribution theory can be used to find a critical value for which H_0 will be rejected when Q is too large. In our modeling this year we viewed the parameters $\underset{\sim}{\beta}$ as a set and either included all six parameters in the model or did not include any of the six. There are undoubtedly cases where there is a signifigant trading day effect but six distinct parameters are not necessary; however, for the purposes of seasonal adjustment we see no need to reduce the number of parameters.

Identification and Estimation of Models which include trading day variation

In the model (2.10) the nature of TD_t has been specified above; however, it is necessary to identify a particular ARIMA process for the model. One approach that has worked sucessfully is to examine the sample autocorrelation function of the original time series in order to determine the degree of differencing. Suppose that it is appropriate to difference the data so that the sample autocorrelation function of $W_t = (1-B)^d(1-B^{12})^D Z_t$ dies out. It is frequently the case that the sample autocorrelation function of W_t exhibits a confused pattern because of the trading day influence. Therefore, what we have done is to consider the sample autocorrelation function of the residuals from the regression of W_t on $(1-B)^d(1-B^{12})^D \; D_{it}$ $i=1, \ldots, 6$. The basic idea is to remove the influence of the trading day factors by a preliminary regression so that the pattern in the sample autocorrelation function of the regression residuals can be clearly seen. Note, it is necessary to first determine the degree of differencing since otherwise the estimates of the β_i's in the preliminary regression will be inconsistent and the sample autocorrelation patterns may be misleading.

The method described above has seemed to be successful; but it should be examined from a more theoretical viewpoint in order to have a better understanding of the process. Also, it is of interest to know how the trading day factors affect the sample autocorrelation function.

Once a model has been tentatively identified, the parameters $\underset{\sim}{\beta}$ and the ARIMA model parameters can be simultaneously estimated using the TSPACK package. In addition, TSPACK can be used to perform diagnostic checks on the fitted model. Once an adequate model has been found it is an easy matter to trading day adjust the data: (i) The estimated trading day factors are calculated by $\hat{TD}_t = \sum_{i=1}^{6} \hat{\beta}_i D_{it}$ where $\hat{\beta}_i$ are the estimates of β_i for i=1, ..., 6. (ii) The trading day adjusted series for month t is then $Z_t - \hat{TD}_t$.

Easter Holiday Effects

For a number of retail sales series, the level of the series can be significantly changed because of consumer buying behavior and stores marketing behavior around the date of Easter. This phenomenon creates a different kind of problem than many other holidays (such as Christmas) because the date of Easter changes each year so that its impact upon March and April also changes each year depending upon its placement relative to these months. We next consider one method to model the effects of Easter on a time series.

For illustration we assume that a series of hypothetical daily sales will increase by a constant unknow amount δ for a fixed number of days, say m days, prior to Easter. We acknowledge that this is a simple minded assumption; however, because we can only observe monthly data, it is impossible to empirically determine the daily movements around Easter. Furthermore, the effects of our simple assumption upon the final results are probably very similar to the effects of more complex assumptions about the daily behavior.

Under the above assumption, in order to account for the effect of Easter we need to properly allocate the effects from Easter to the months of March and April. To accomplish this we define

$$E_j = \begin{cases} S & \text{if Easter falls on April } S \text{ in year } j \\ S-31 & \text{if Easter falls on March } S \text{ in year } j \end{cases}$$

and further define for month t in year j

$$Y_t = \begin{cases} 0 & \text{if the month is not March or April} \\ 1 & \text{if } E_j \le 1 \text{ and the month is March} \\ 0 & \text{if } E_j \ge m+1 \text{ and the month is March} \\ \dfrac{m-E_j+1}{m} & \text{if } 2 \le E_j \le m \text{ and the month is March} \\ 0 & \text{if } E_j \le 1 \text{ and the month is April} \\ 1 & \text{if } E_j \ge m+1 \text{ and the month is April} \\ \dfrac{E_j-1}{m} & \text{if } 2 \le E_j \le m \text{ and the month is April.} \end{cases}$$

Note that for each month Y_t represents the proportion of the m days assumed to be affected by Easter in that month. Based upon these developments, the model describing the Easter holiday effect is

$$H_t = (m \cdot \delta) \cdot Y_t = \Delta \cdot Y_t \ .$$

where Δ is an unknown parameter which will be estimated from the observed data. If we assume that the Easter effect is fixed, that there are trading day effects and a residual ARIMA model to describe the seasonality (N_t) then an appropriate model is

$$Z_t = \Delta \cdot Y_t + \sum_{i=1}^{6} \beta_i \cdot D_{it} + N_t \ . \qquad (2.11)$$

When modeling series with changes due to Easter we have fit model (2.11). However, it is necessary to specify a value of m in order to calculate Y_t. What we have done to determine m is to fit model (2.11) using m = 0, 7, 14 and 21 and then choose the value of m which yeilds the smallest sum of squares. In this way we let the value of m be determined from the series being modeled. This procedure has seemed to work successfully for most of the series that we have modeled. However, we feel that the best way of modeling Easter variation should be studied further.

Part 3: The Empirical Study

As we stated in the introduction to this report, our main goal this year was to conduct a moderate sized empirical study comparing the model based seasonal adjustment method to Census X-11 and Statistics Canada X-11 ARIMA. The reasons for including X-11 in the study are fairly obvious. (i) X-11 is the method currently being used at the Census Bureau. (ii) X-11 has been used for many years and most people involved with seasonal adjustment have had experience using X-11, therefore X-11 is a natural method to use as a standard of comparison. (iii) Apparently people are somewhat happy with the results of X-11 so that any new method must not give results which are radically different than X-11 for the majority of series or the new method is probably suspect. X-11 ARIMA was included in the study because we felt that there was a chance that it might replace X-11 in the near future; consequently, comparisons of the model based and X-11 ARIMA were important. Finally, the particular model based procedure was included because of the research fellow's involvement in developing it and the theoretical advantages cited in part 1 of this report.

From a purely theoretical viewpoint there are several reasons why the model based procedure could be considered superior to the other two methods in the comparison. The model based method uses information about the individual time series being adjusted in a rigorous way when deriving the moving average filters to be used for adjusting the series. In this sense, the model based adjustment is therefore consistent with the structure of the data while X-11 and X-11 ARIMA are not necessarily consistent with this information. The filters in the model based approach are derived from the theory of optimal signal extraction so that in particular the filters applied at the ends of the data are more appropriate than those of X-11. As opposed to X-11 and X-11 ARIMA, all of the statistical assumptions for the model based procedure are rigorously specified and the degree of arbitrariness is exactly known. Because the idea behind the model based approach is to make use of the information available in the series to be adjusted when deriving the estimates of seasonal factors, in principle we might expect the model based procedure on the average to give a better seasonal adjustment than X-11 or X-11 ARIMA.

Even though in theory a model based adjustment may be better than an adjustment from either of the other two methods, it is of interest to determine if in practice there are any major differences in the various methods. If, for instance, for all practical purposes the model based approach yields the same results as X-11 then there would be no advantage in using the model based approach even though in theory it may be superior. Determining the degree to which the three methods differ in practice is largely an empirical matter. In addition, there are several questions about the model based approach that can only be answered empirically. First, as is evident from the theoretical discussion of the model based approach given in part 1 of this report, there do exist ARIMA models for which a seasonal decomposition does not exist. An important empirical question is whether or not these types of ARIMA models occur in data series that are commonly seasonally adjusted at the Census Bureau. Secondly, we have chosen to use the cannonical seasonal defined in part 1 of this report as the model for the seasonal component. Since this can be regarded as an arbitrary choice, it is important to find out if the results from making this choice are consistent with what experts consider an adequate seasonal adjustment. Again, one way to judge this is empirically. Finally, an empirical study of this kind should raise questions about seasonal adjustment in general and about model based seasonal adjustment in particular. We address some of the issues that resulted from this empirical study in part 4 of this report.

One of the problems in conducting an empirical study comparing seasonal adjustment methods is in deciding how to judge the relative merits of the various methods. It is not difficult to eliminate a procedure if it is grossly inadequate; for example, if there was substantial residual seasonality in the seasonally adjusted series or there was a portion of the trend in the seasonal component. On the other hand, because of the arbitrary nature of seasonal adjustment, there is not a "correct seasonal component" that can be used as an absolute standard of comparison. Therefore, judging the virtues of the methods in this study is difficult and to a large extent is a matter of personal opinion about what constitutes a good seasonal adjustment.

For the purposes of this discussion we assume that the three methods are not grossly inadequate. To compare the three methods in this situation we have chosen to calculate measures of revisions in the seasonally adjusted series, a measure of smoothness in the seasonally adjusted series and a measure of the extent to which the level of the unadjusted series is preserved by the adjusted series. It may be valuable to look at these measures even if they are not used to judge the adequacy of a seasonal adjustment. This is because the measures may provide information as to how the methods behave. For example, they may indicate that one method is buying smoothness at the expense of high revisions when compared with another method. We next discuss our reasons for using these particular measures and the results of the study.

Revisions

The current seasonal adjustment practice at the Census Bureau is at the beginning of each year, to adjust a particular series using data through the months of December of the previous year and using the same data forecasts of the seasonal factors for the next 12 months are produced. We call the forecasted seasonal factors the year ahead factors. These year ahead seasonal factors are then used to derive the official seasonally adjusted series as the unadjusted data becomes available. After an additional year's worth of unadjusted data becomes available, the process is repeated. However, because there is information about the current year's seasonal factors in current and future data, as more data becomes available we expect that the estimated seasonal factors will be changed. These changes in seasonal factors then lead to revisions in the seasonally adjusted series. In particular, if we let

X_t^f = the seasonally adjusted value for month t based upon the forecasted seasonal factors and let X_t^i = the seasonally adjusted value for month t based upon i years of additional data (i = 1,2,3); then the revisions in the seasonally adjusted value for month t after i years of new data becomes available are $X_t^i - X_t^f$ for i = 1,2,3. For each series considered in the study, we have taken as a measure of revisions the average relative absolute revisions

$$R^i = \frac{1}{12} \sum_{t=1}^{12} \frac{|X_t^i - X_t^f|}{X_t^i} \qquad i = 1,2,3. \qquad (3.1)$$

Then for each series R_i for i = 1,2,3 is a measure of the relative amount of revision in the seasonally adjusted series after one, two and three years of additional data become available.

Note that for each series R^i is a measure of the average amount of revision relative to the level of the most recent estimate of the seasonally adjusted series. If everything else were equal we would prefer small values of R^i. Therefore, one way to compare the three seasonal adjustment methods is, for each series to compute R^i i = 1,2,3 for the three different seasonal adjustment methods and determine which approach if any gives smaller value of R^i.

In this empirical study we used 76 time series. A brief description of these series together with the abbreviations used to refer to each series are given in table A.1 of the appendix. In addition, the dates of each series that were used in this study are given in table A.2 of the appendix. For the purpose of calculating the measures of revisions in this study, we used data from the starting date up to three years from the ending date to calculate the year ahead seasonal factors and the values of X_t^f. We then added one, two and three years of data in order to compute respectively X_t^1, X_t^2 and X_t^3 also the values R^1, R^2 and R^3 were computed. We note that R^1, R^2 and R^3 were computed for 69 of the 76 series and only R^1 was computed for 7 series because the length of time for which data was available was too short to compute R^2 and R^3 with these 7 series. These computations were done for each of the three seasonal adjustment methods (the standard options were used for X-11 and X-11 ARIMA). Then for each series, the values of R^i for the model based adjustment and for X-11 ARIMA were divided by the values of R^i for X-11. The values of these relative ratios for each series are reported in table A.3 of the appendix. Note a value of the relative ratio which is less than 1.00 indicates the approach did better than X-11 and a value greater than 1.00 indicates the approach did worse than X-11. In order to get an idea of the overall performance of the three methods we calculated the average of the relative ratios of table A.3. These averages are reported in table 3.1.

Based upon table 3.1 we draw the following conclusions. On the average there is about a 40 percent reduction in the relative absolute first, second and third year revisions of the model based method over X-11. In addition, the model based method had a smaller measure of revisions than X-11 in over 85 percent of the series for each of the first, second and third year measures. There is about a 50 percent reduction in the relative absolute first year revisions and about a 40 percent reduction in the second and third year revisions when using the model based approach rather than X-11 ARIMA. Note that the averages for X-11 ARIMA are divided into the subset of series in which the program automatically picked a model and the subset of series in which a model was picked by the user (model forced). This was done to determine if forcing a model in the ARIMA program would lead to larger revisions than the cases where models were chosen by the program. In this study, the forced models had about the same or a smaller amount of revisions than did the automatically chosen models. In summary, from the results of the study, we conclude that on the average there is a substantial reduction in first, second and third year revisions when the model based approach is used rather than X-11 or X-11 ARIMA.

Table 3.1

Average Relative Ratio of Revision Measures for

Model Based and X-11 ARIMA

First Year Revisions

Adjustment Method	n	Average Relative Ratio	Number Better Than X-11
Model Based	76	.62	66
X-11 ARIMA	76	1.25	28
X-11 ARIMA - automatic	60	1.26	23
X-11 ARIMA - forced	16	1.20	5

Second Year Revisions

Adjustment Method	n	Average Relative Ratio	Number Better Than X-11
Model Based	69	.60	61
X-11 ARIMA	69	1.06	36
X-11 ARIMA - automatic	53	1.08	27
X-11 ARIMA - forced	16	1.05	9

Third Year Revisions

Adjustment Method	n	Average Relative Ratio	Number Better Than X-11
Model Based	69	.57	63
X-11 ARIMA	69	.96	41
X-11 ARIMA - automatic	53	.95	34
X-11 ARIMA - forced	16	.98	7

It is of interest to attempt to understand why in our study there were smaller revisions when using the model based procedure rather than the other methods. One way to do this is to examine the plots of the series for the model based and X-11 seasonal adjustments in the appendix. In particular, for each series it is helpful to compare the estimated seasonal factors obtained with the two methods. It is evident from these plots that the seasonal factors for the model based approach evolve more slowly than do the seasonal factors for X-11. In addition, if we examine the form of the smoothing formulas given in part 1 of this report it is evident that the length of the moving averages is determined by the parameters in the moving average polynomial of the model for Z_t. In particular, the magnitude of the seasonal moving average parameter is the most important factor in determining the length of the seasonal filter. Now Cleveland and Tiao (1976) have found that the X-11 program with the standard options is in some sense assuming the observable series has a seasonal moving average parameter equal to about .45. This is to be contrasted with the estimated seasonal moving average parameters for the series in this study. These are summarized in table 3.2. In all but one of the 76 series the estimated θ_{12} was larger than .45 and in most cases the estimated θ_{12} was substantially greater than .45. Although the collection of Bureau series chosen was not a random sample, the series were in no way selected with an eye toward large θ_{12} values. These facts imply that the length of the moving average filter for estimating the seasonal component in the model based method was almost always longer than the length of the corresponding X-11 moving average. Thus the estimated seasonal component for the model based procedure will not change as rapidly as the estimated seasonal component for X-11.

Table 3.2

Estimated Values of θ_{12}

	<.5	.5-.6	.6-.7	.7-.8	.8-.9	.9-.1
Number of Occurances	1	5	3	11	38	18
Percentage of Occurance	1.3	6.6	4.0	14.5	50.0	23.7

Intuitively, this seems to at least partially explain why the model based procedure had smaller revisions than X-11. A method which yeilds seasonal factors that evolve slowly should have relatively small revisions. The thing that is signifigant about these observations is that the value of θ_{12} can be estimated from the observable data so that one implication of this study is that the moving average filters for the seasonal component should be longer than the lengths of filters presently used in the standard version of X-11 for most of the Bureau series.

We might conjecture that the cannonical decomposition that we are using in the model based method would yield smaller mean squared errors in revisions than any other acceptable decomposition because we have argued in part 1 of this report that the cannonical decomposition is the "most deterministic" representation for the seasonal component which is also consistent with information in the data. We have found that this conjecture is very difficult to verify analytically; however using the results of Pierce (1980) it is possible to numerically calculate the expected mean square error in revisions for any acceptable decomposition. Consequently, we considered the situation when the model for the overall series is

$$(1-B)(1-B^{12})Z_t = (1-\theta_1 B)(1-\theta_{12}B^{12})a_t$$

For all possible combinations of $\theta_1 = .1, .3, .5, .7, .9$ and $\theta_{12} = .1, .3, .5, .7, .9$ we computed the expected mean square error in revisions for a grid of eleven possible acceptable decompositions including the canonical decomposition. We assumed that the revisions were computed from the initial adjusted values derived from the year ahead seasonal factors. For all of the combinations of θ_1 and θ_{12} considered we found that the canonical decomposition gave the smallest expected mean squared error in revisions. Therefore, these findings seem to support our intuition; however, additional work needs to be done in this regard.

In summary, we have shown in an empirical study that a substantial reduction in the amount of revisions can be achieved if the model based seasonal adjustment approach is used rather than either X-11 or X-11 ARIMA. We have argued that one reason for this result is that the majority of Bureau series have seasonality which evolves rather slowly. Furthermore, the model based method provides a way to discover when the seasonal pattern of a series is slowly changing so that this fact can be used to obtain more appropriate estimates of the seasonal components.

Measures of smoothness and level preservation

As a means of comparing the broad characteristics of the model based adjustment method and X-11 for this empirical study we have computed the measures

$$SM = \sum_{t=3}^{n} (X_t - 2X_{t-1} + X_{t-2})^2 \qquad (3.2)$$

and

$$F = \sum_{t=12}^{n} \Big[\sum_{j=t-11}^{t} (Z_t - X_t) \Big]^2 \qquad (3.3)$$

where Z_t and X_t denote respectively the original value and the seasonally adjusted value for month t. Note that (3.2) is a crude measure of the smoothness of the seasonally adjusted series and (3.3) is a crude measure of how closely the moving twelve month totals of the unadjusted and adjusted data agree. We calculate (3.3) using the metric for which the additive representation $Z_t = S_t + X_t$ is appropriate.

For the purpose of comparison we have calculated the values of SM and F for each of the 76 series in our study and for the model based and X-11 methods. We excluded X-11 ARIMA because we view X-11 ARIMA as an attempt to improve the adjustment of the ends and consequently X-11 ARIMA's broad characteristics should be similiar to those of X-11. For each series the values of SM and F for the model based approach were divided by the values for X-11. These ratios are reported in table A.4 of the appendix. Based on this table it is evident that X-11 has on the average a smoother seasonally adjusted series than does the model based approach and that the model based adjusted series preserves the level of the twelve month moving sums better than the X-11 adjusted series. These observations are consistent with the earlier observations about the two methods.

Resource requirements of the model based approach

One of the important things that must be considered before any model based seasonal adjustment procedure can be adopted is the amount of resources required for its use. Based upon our experience with this empirical study we have a rough idea about what would be required to implement the particular model based approach used in the study.

First, some kind of Box-Jenkins ARIMA modeling software package is necessary. For Census Bureau series it is essential to be able to handle intervention models and regression type models with ARIMA error structures. In our opinion, the best available software package is the TSPACK routine which we used this year. That routine is available at the Bureau.

Another software requirement is a program to compute the seasonal adjustment based upon the particular model for the data. This program was partially written before the start of the project this year and during the course of the year it has been checked and rewritten to include more general models. However, before the program can be used extensively it probably should be rewritten so that its input can be simplified and its output made to conform more closely to Census needs.

Finally, the largest resource committment, at least initially, is the time required to model the individual series. The time required to model a series depends both on the experience of the model builder and the difficulty of the series. Our opinion is that someone with a moderate amount of experience should be able to adequately model an easy to moderately difficult series (at least 75% of those in this study were easy to moderately difficult) in less than 2 hours. Modeling a large number of series thus involves a large time committment. However, we must also recognize that for most series this is an initial investment that will not have to be repeated. Once a model is built for most series the model probably will not change, and updating may only involve reestimation of the model parameters. An additional advantage is that a model based approach forces us to become intimately involved with the time series being adjusted. We therefore would expect that a model based approach may not only lead to better seasonal adjustment, but also to a better understanding of the unadjusted series. Consequently, a model based seasonal adjustment approach may lead to improvements in areas of interest at the Bureau other than seasonal adjustment.

Part 4: Other Issues Related to Seasonal Adjustment

During the course of conducting the empirical study several issues that have not yet been discussed arose. In this part of the report we briefly discuss these issues. To a certain extent what we will consider in this part are some interesting but only partially answered questions. Our hope is that these questions may generate some additional research interest.

Deciding when a series should not be adjusted

One question that is important to be able to answer is: how can we decide if a particular series should not be seasonally adjusted? This question can be addressed from a number of different angles, however we shall consider the question from a statistical modeling viewpoint. We assume that the unadjusted series can be approximated by an ARIMA model. Now for monthly data if the particular ARIMA model representing an unadjusted series does not have any seasonal part then it is clear that the unadjusted data is not seasonal and the series should not be seasonally adjusted. As an example, during the course of the empirical study we found that the model

$$Z_t = Z_{t-1} + a_t \qquad (4.1)$$

was an appropriate model to describe the behavior of the log of the industrial inventory series S45TI (ship building). Because there is no seasonality implicit in a series which follows model (4.1) we conclude that this series should not be seasonally adjusted.

The above example was clear cut; however we can also consider the example of the industrial inventory series S07TI (glass containers) which follows the model

$$W_t = .26W_{t-12} + a_t \qquad (4.2)$$

where W_t is the first difference of the logs of the unadjusted data. Now the model (4.2) does allow for some seasonality since W_t is related to the value twelve months ago. However, the seasonality implicit in (4.2) can be considered very weak. First, the values of W_t for each different month vary around zero rather than a distinct monthly mean. As a result the values of W_t will stay close to zero. Second, in successive years the values of say, January, are positively correlated; but that correlation is very small. Therefore, it would not be unusual to have a positive value of W_t for one January and in a year or two to have a negative value of W_t and a year or two later another positive value. These remarks are also valid for the other months. Therefore, we can compare the expected behavior of a series following the model (4.2) with a definition given by Kallek (1978) that seasonality is "regular periodic fluctuations which recur every year with about the same timing and intensity". It is evident that realizations of a series that has as its model (4.2) do not exhibit the behavior in this definition. Hence, we can argue that based upon statistical modeling considerations the series S07TI should not be seasonally adjusted.

From the last example it appears that correlation at a twelve month lag is not a sufficient reason to justify seasonal adjustment of a series. Consequently, we need to have some other criterion to judge whether or not to seasonally adjust a series. With this in mind we consider the model

$$W_t = W_{t-12} + N_t \qquad (4.3)$$

where W_t is the value of a series (possibly transformed or differenced) at month t and N_t follows a stationary zero mean model. The model (4.3) implies that the value of the series for say, January, is equal to the value for last January plus an error term that varries around zero. Thus, a series following (4.3) will exhibit regular monthly fluctuations which recur with about the same intensity provided the variability in N_t is not large compared to the monthly values of W_t. Now the model (4.3) is seasonally nonstationary because it will typically not have a level that is time invariant but rather the level will depend upon the particular month. The monthly intensities in (4.3) will change but as long as the variance of N_t is not too large the changes will be gradual.

Based upon the above considerations we call a monthly series strongly seasonal if it is necessary to include a twelfth difference in its model and we call a monthly series weakly seasonal if it has significant autocorrelations at multiples of lag 12 but no twelfth difference in its model. If we base our decision only on statistical modeling considerations, we feel that only strongly seasonal time series should be adjusted. Of course, in practice the modeling considerations may be only part of what is used to decide whether or not to adjust a particular series, but we feel that they may help in making the decision. During the process of modeling series for the empirical study we found several series that were not strongly seasonal; they are listed in table 4.1.

Table 4.1

Series that are weakly seasonal or not seasonal

Series I.D.	Description
S07TI	industry inventories - glass containers
S13TI	industry inventories - nonferrous metals
S24TI	industry inventories - construction mining material handling
S25TI	industry inventories - metal working machinery
S29TI	industry inventories - general industry machinery
S37TI	industry inventories - commercial equipment
S38TI	industry inventories - electrical components
S45TI	industry inventories - ship building
S76TI	industry inventories - paperboard containers
S83TI	industry inventories - industrial chemicals
SX2TI	industry inventories - electrical transmition & distribution equip.
SX5TI	industry inventories - aircraft, missiles, parts
TI501	wholesale inventories - motor vehicles, automotive parts and supplies
TI502	wholesale inventories - furniture and home furnishings
TI508	wholesale inventories - machinery, equipment and supplies

Detection of Deterministic Seasonality

When modeling and analyzing seasonal time series it is sometimes possible to represent the seasonality in the series as a fixed or deterministic component like monthly means. Of course not all seasonal time series can be modeled using deterministic seasonality, but when that is possible there are several advantages to representing the seasonality in that way. First, the behavior of deterministic seasonality is easy to explain. Second, in constrast to the situation where we have stochastic seasonality, if there is fixed seasonality then there is no unresolved ambiguity about how to do the seasonal adjustment. Therefore, it can be important to develop techniques to discover when using a fixed seasonal is appropriate.

In the context of ARIMA modeling there is an informal way to discover when fixed seasonality is appropriate. Suppose that an observed time series, Z_t, is equal to the sum of a seasonal component S_t and a white noise component a_t. Furthermore, we assume that the S_t are fixed, distinct monthly means so that $S_t = S_{t-12}$ for all t. Now if we perform a standard Box-Jenkins analysis on Z_t we will be lead to consider

$$(1-B^{12})Z_t = (1-B^{12})S_t + (1-B^{12})a_t = (1-B^{12})a_t. \qquad (4.4)$$

Therefore, in theory the series $W_t = (1-B^{12})Z_t$ follows a stationary moving average process with seasonal moving average parameter equal to 1. If we appropriately estimate the parameter θ_{12} in the model

$$W_t = a_t - \theta_{12}a_{t-12}$$

then for the above situation we would get an estimate of θ_{12} near 1. We note in passing that this situation can create some rather difficult estimation problems. The point of this illustration is that if it is necessary to seasonally difference a series and if the resulting model includes a seasonal moving average parameter whose estimate is "close to 1", then there is some indication that the data could be modeled using a fixed seasonal representation.

The signifigance of this is from table A.2 in the appendix there are a number of series for which the estimate of θ_{12} is larger than .75. In addition, the method used in TSPACK tends to give estimates of θ_{12} which are less than 1 in the case where the true value is equal to 1. A better way to estimate θ_{12} is by using an exact likelihood procedure. In order to check if there is evidence for deterministic seasonality in the series modeled for the empirical study, we took 24 series that had estimates of θ_{12} calculated by the TSPACK routing to be larger than .75 and estimated the parameters using an exact likelihood method available in the Wisconsin Multiple Time Series package. Note that we would have preferred to use an exact likelihood procedure originally but we needed to have the capability to simultaneously estimate outlier, trading day and holiday effects along with ARIMA parameters. The results of this comparison are reported in table 4.2. From this table we see that the informal procedure does not indicate the presense of deterministic seasonality except for the series HST5 and possible S23TI. In fact for most of the 24 series the exact likelihood estimates of θ_{12} were smaller than the TSPACK estimates of θ_{12}. We suspect that in cases where the true θ_{12} is not close to 1 TSPACK estimates tend to be larger than exact likelihood estimates of θ_{12}.

Therefore, for most of the series considered there is not substantial evidence, based upon the cancellation argument, that using deterministic seasonality is appropriate. Surprisingly, we have found for some series (e.g., HST1 and ENM20) that we can model the data using monthly means and apparently account for all of the seasonality in the data, even if the cancellation argument does not indicate deterministic seasonality. This raises some interesting questions. It would seem to be important to investigate the theoretical implications of the model with deterministic seasonality compared to the model with stochastic seasonality to decide if one should be preferred over the other on theoretical grounds. In addition, it would be useful to develop a formal test to decide which of the two representations are more appropriate for a given time series. Finally, it would be important to know in what circumstances using the two models will result in significantly different conclusions.

Table 4.2

Estimates of θ_{12} in candidate series for deterministic seasonality

Series I D	Backforecasted Estimate	Exact Likelihood Estimate
SHIP	.80	.74
S60TI	.85	.77
S65TI	.78	.71
TI506	.90	.74
TI507	.87	.74
EAM20	.88	.74
ENF20	.90	.82
INVE	.91	.84
ENM20	.88	.81
S62TI	.79	.72
TI500	.90	.79
S35TI	.88	.76
S36TI	.91	.85
HST1	.82	.70
UF16	.74	.63
ENM16	.87	.72
ENF16	.87	.74
HST5	.92	1.00
HSTT	.89	.87
S16TI	.90	.84
EAM16	.90	.84
EAF16	.77	.68
S23TI	.91	.93
S21VSU	.88	.88

Multiplicative vs Log Additive Seasonal Adjustment

The most common way in which people conceive of an observed time series, Z_t, in terms of its components is the multiplicative model

$$Z_t = S_t \cdot X_t \tag{4.5}$$

where S_t is a seasonal component and X_t is a nonseasonal component. Some reasons for using the representation (4.5) are: (i) it is an empirical fact that the seasonal variability tends to increase as the level of the series increases for many series, and (ii) many series are measured in dollars so that both the level of the series and the magnitude of the seasonality can be affected by inflation. If the representation (4.5) is appropriate than it follows that the additive representation

$$\ln Z_t = \ln S_t + \ln X_t \tag{4.6}$$

is appropriate for the logarithms.

Now if we assume that seasonal adjustment can be thought of as a signal extraction problem, then it is necessary to use representation (4.6) rather than (4.5) because the theory of signal extraction is developed in terms of an additive structure. In addition, in the situation where S_t is deterministic then some form of regression analysis would be appropriate. For example, if $\ln X_t$ were white noise (4.6) is the standard regression model. However, again we must use the additive representation because the usual regression assumptions are not satisfied in the multiplicative framework. Of course if the analysis is done in the log metric then the results must be transformed back into the metric of the original series for publication purposes.

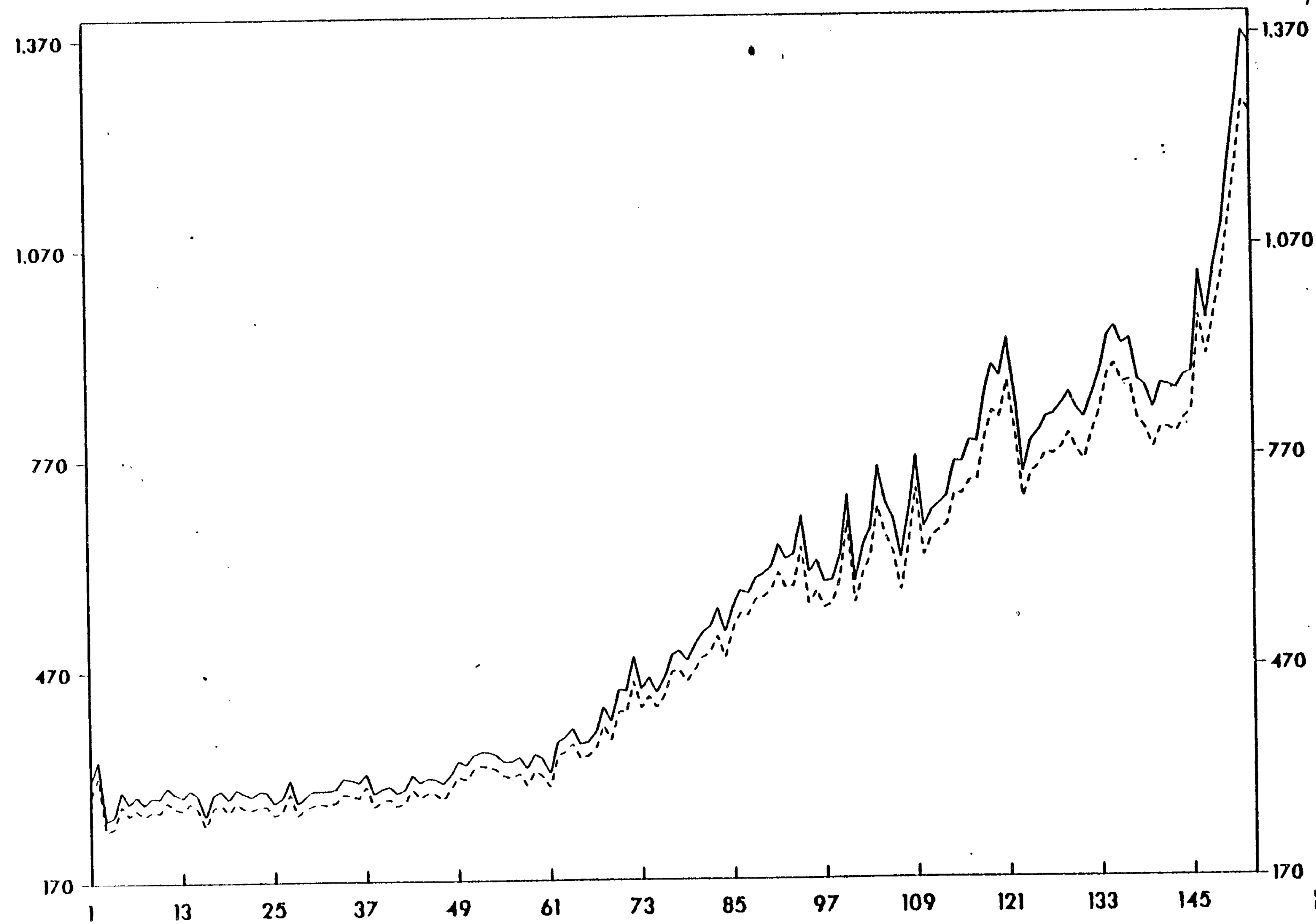
The above discussion is relevant because for a multiplicative representation the current verison of X-11 performs its analysis upon the original data using arithemetric averages (multiplicative adjustment) when a strong theoretical case can be made for first transforming the data by taking logarithms and then treating the resultant series as if it were an additive model (log additive adjustment). For the majority of series adjusted there are only minor differences between the results of the X-11 multiplicative adjustment and the X-11 log additive adjustment; however, there are a few cases where the two alternatives do give different results. As an example, in figure 4.1 we have plotted the seasonally adjusted series for the series 598000, the retail sales for fuel oil dealers, liquefied petroleum dealers and fuel and ice dealers. The multiplicative seasonally adjusted series is plotted in the solid line and the log additive seasonally adjusted series is plotted in the dashed line. As is evident from the plot, there is approximately a constant difference between the two adjusted series. It is also clear that at least one of the methods must be ruled out.

After some thought and an examination of the original data it is relatively easy to explain the difference in the two methods. The multiplicative version of X-11 tends to make the yearly arithmetic average of the seasonal components close to 1 $\left(\frac{1}{12}\sum_{i=1}^{12} S_i = 1\right)$. In contrast, the X-11 log additive adjustment tends to make the arithmetic average of the logarithms of the seasonal factors equal to 0 $\left(\frac{1}{12}\sum_{i=1}^{12} \ln S_i = 0\right)$. Equivalently, for the log additive approach we have that the geometric mean of the seasonal factors is about 1 $\left(\left[\prod_{i=1}^{12} S_i\right]^{1/12} = 1\right)$. Now for the particular series under consideration it happens that the seasonal components range from as low as .6 to as high as 1.7 so that the S_i are not close to 1. In this case it is easy to show that the geometric mean is substantially smaller than the arithmetic mean. Therefore, for this series the seasonal factors for the two methods differ by about a constant amount and these relatively constant differences are reflected in the seasonally adjusted series.

The series 598000 was the case that showed the most extreme difference between the X-11 multiplicative and X-11 log additive adjustment. In order to get an idea of the magnitude of difference in the two approaches for some other series we have included plots for five other series that showed some differences (in figures 4.2 through 4.6). Keep in mind that these examples were chosen because they represent

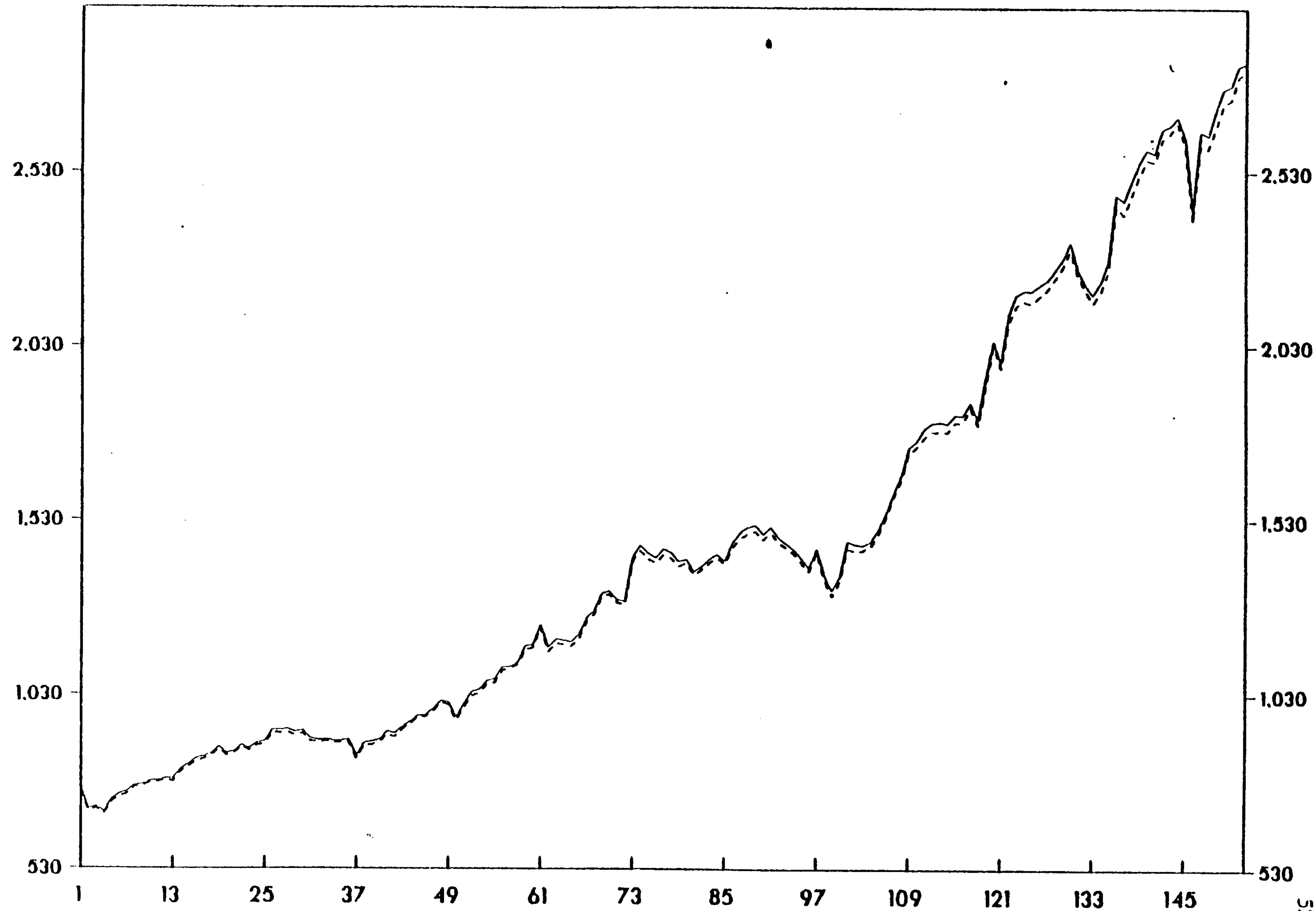
RETAIL 598000 X-11 MULTIPLICATIVE AND X-11 LOG ADDITIVE

Figure 4.1



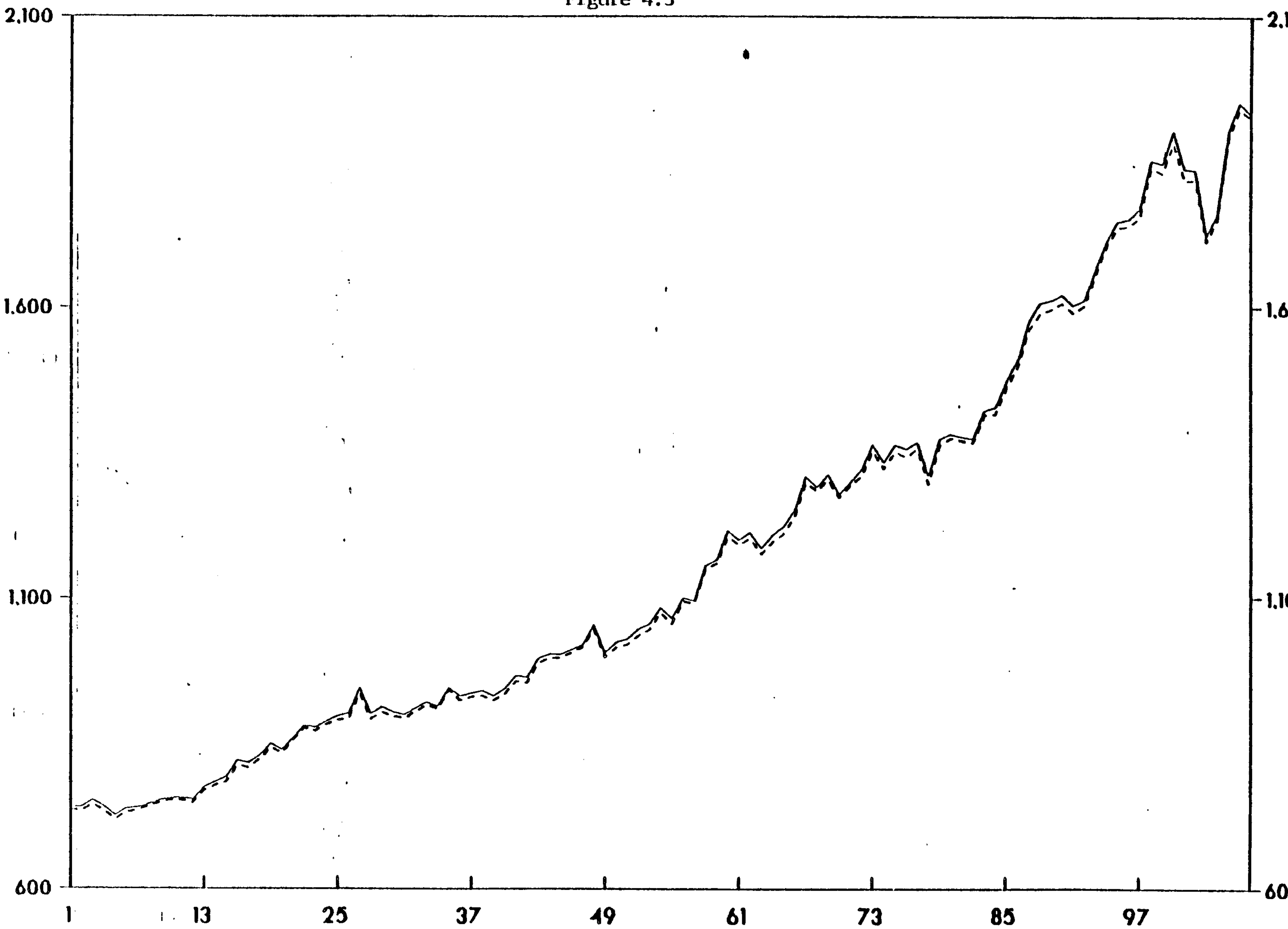
RETAIL 520002 X-11 MULTIPLICATIVE AND X-11 LOG ADDITIVE

Figure 4.2



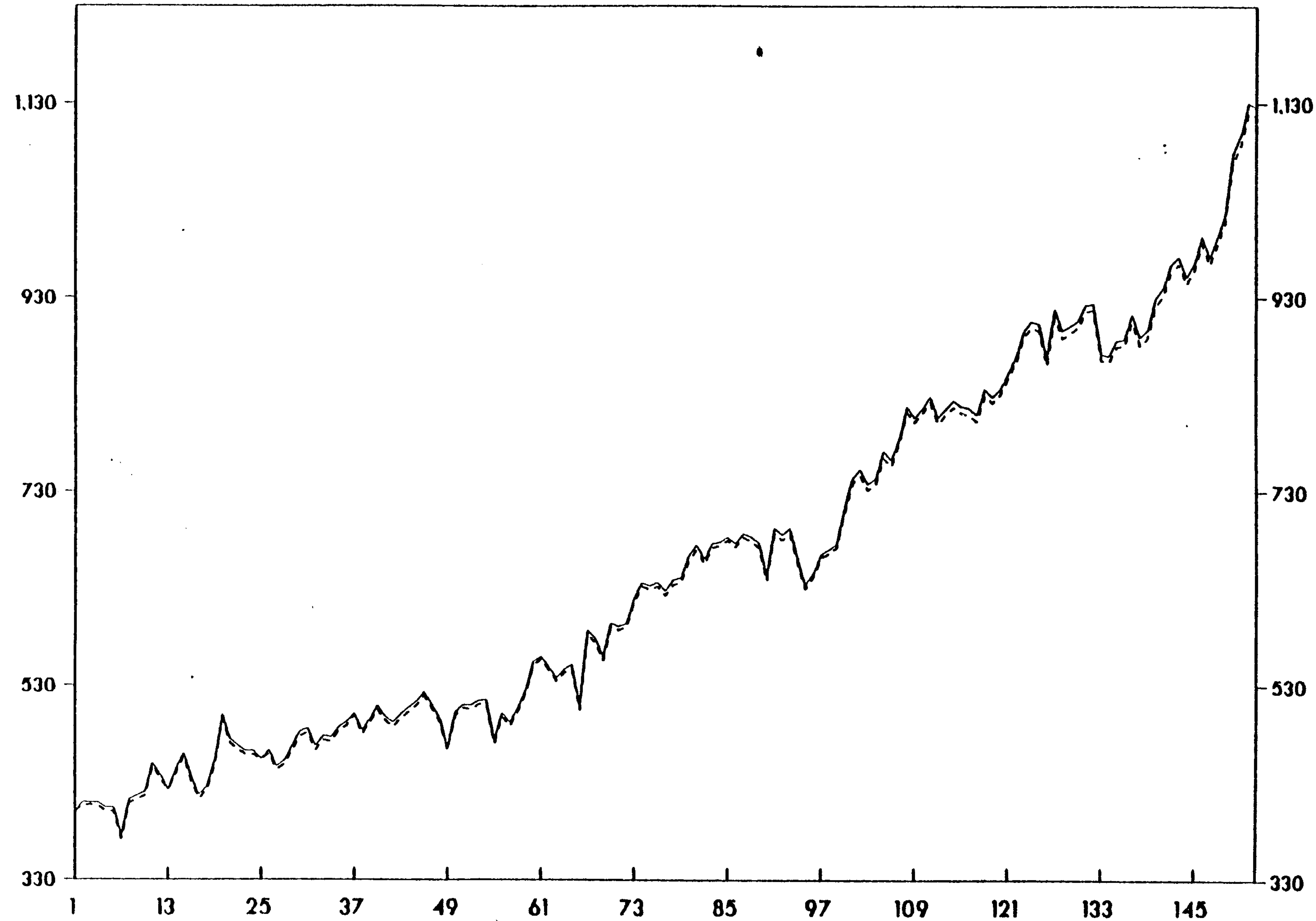
SERVICE 701000 X-11 MULTIPLICATIVE AND X-11 LOG ADDITIVE

Figure 4.3



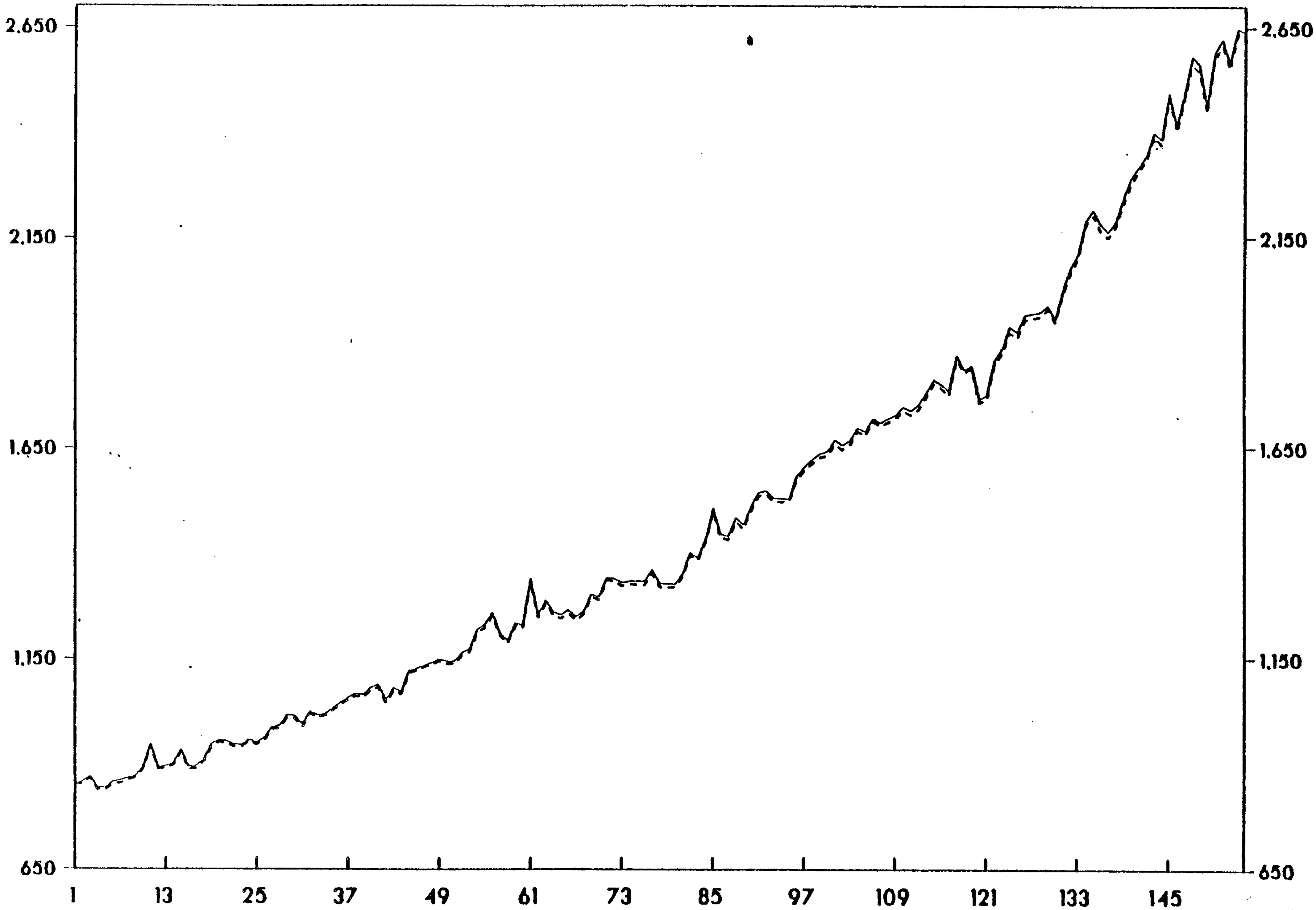
RETAIL 570002 X-11 MULTIPLICATIVE AND X-11 LOG ADDITIVE

Figure 4.4



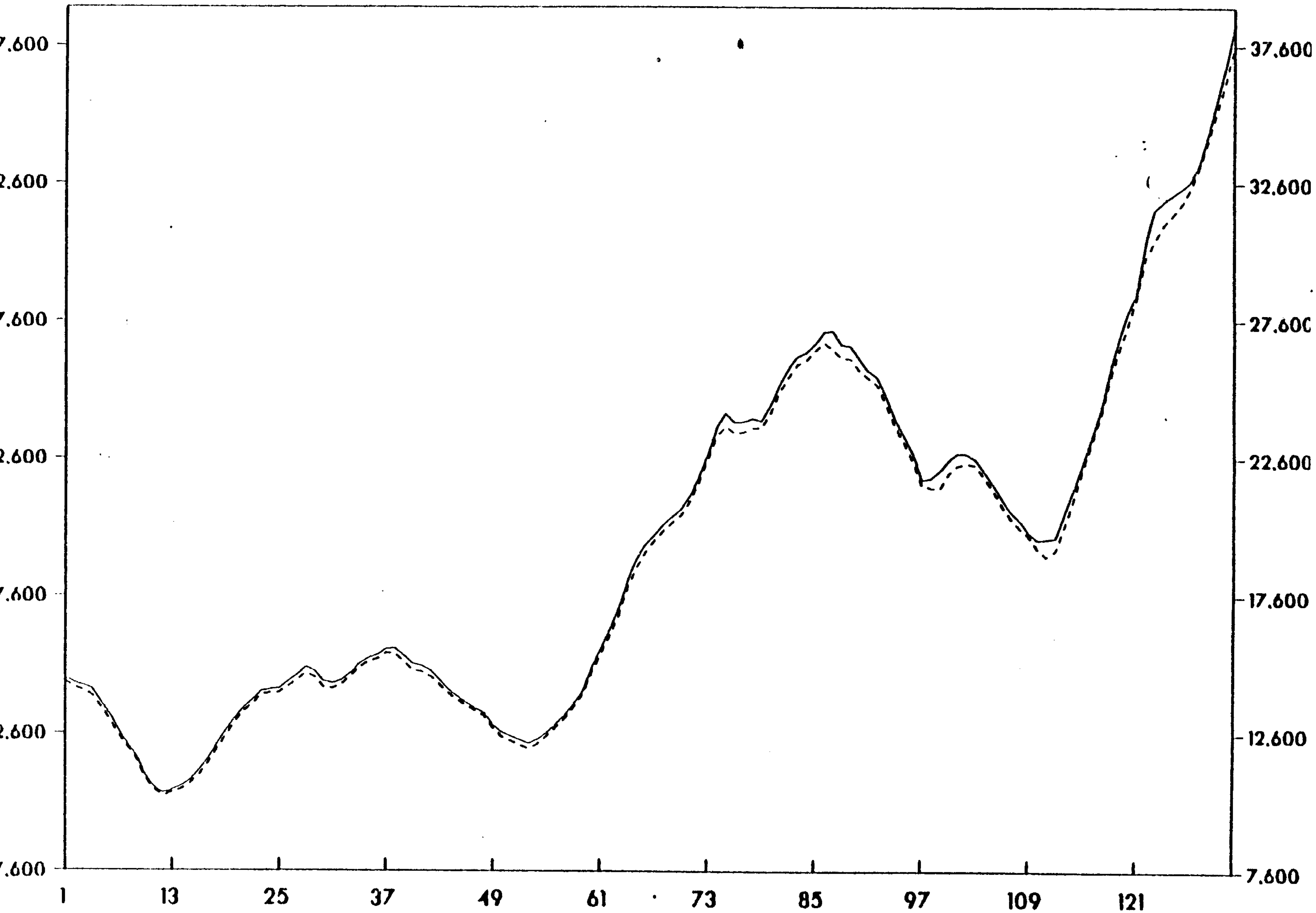
WHOLESALE 518000 X-11 MULTIPLICATIVE AND X-11 LOG ADDITIVE

Figure 4.5



VIP X-11 MULTIPLICATIVE AND X-11 LOG ADDITIVE

Figure 4.6



examples of cases where there are differences in the two procedures and for the majority of series we considered there would not be discernable differences.

Because of the discrepancy between the multiplicative and log additive methods, it may be necessary to determine which of the two approaches is preferrable. The arguments in support of the log additive are: (i) from the point of view of signal extraction the analysis is more appropriately performed for an additive representation; (ii) from the view of modeling the original series, the usual ARIMA model assumptions are going to be more appropriate in terms of the logarithmic metric since for instance the data in the original metric will not have a constant variance if (4.5) is true, (iii) the multiplicative approach of X-11 is inconsistent in that it is mixing arithmetic averages with data assumed to have proportional seasonality. For multiplicative X-11 it would be more appropriate to use geometric averages instead of arithmetic averages. On the other hand, we can only think of one possible reason to support the multiplicative approach. That is that the restriction implicit in the multiplicative approach, $\frac{1}{12}\sum_{i=1}^{12} S_i = 1$, tends to make the yearly totals of the unadjusted series more nearly, equal to the yearly totals of the adjusted series than does the restriction, $[\prod_{i=1}^{12} S_i]^{1/12} = 1$. Some additional thought is needed about whether or not the yearly sums of the adjusted and unadjusted series should be approximately equal.

Bell, W.R. (1980) "Multivariate Time Series - Smoothing and Backward Models" Unpublished Ph.D Thesis, Statistics Department, The University of Wisconsin-Madison.

Box, G.E.P., Hillmer, S.C. and Tiao, G.C. (1978) "Analysis and Modelling of Seasonal Time Series" in Seasonal Analysis of Economic Time Series edited by Arnold Zellner U.S. Department of Commerce.

Box, G.E.P. and Jenkins, G.M. (1970) Time Series Analysis: Forecasting and Control. San Francisco, Holden Day.

Box, G.E.P. and Tiao, G.C. (1975) "Intervention Analysis with Applications to Economic and Environmental Problems" JASA, 70, 70-79.

Cleveland, W.P. (1972) "Analysis and Forecasting of Seasonal Time Series" Unpublished Ph.D Thesis, Statistics Department, The University of Wisconsin-Madison.

Cleveland, W.P and Tiao, G.C. (1976) "Decomposition of Seasonal Time Series: A Model for the Census X-11 Program" JASA, 71, 581-587.

Kallek, S. (1978) "An Overview of the Objectives and Framework of Seasonal Adjustment" in Seasonal Analysis of Economic Time Series edited by Arnold Zellner U.S. Department of Commerce.

Pierce, D.A. (1971) "Least Squares Estimation in the Regression Model with Autoregressive- Moving Average Errors." Biometrika, 58, 299-312.

Pierce, D.A. (1980) "Signal Extraction Error in Nonstationary Time Series" The Annals of Statistics, 1, 1303-1320.

Pierce, D.A. (1980) "Data Revisions with Moving Average Seasonal Adjustment Procedures" forthcoming in The Journal of Econometrics.

Shiskin, J., Young, A.H. and Musgrave, J.S. (1967) "The X-11 Variant of Census Method II Seasonal Adjustment Program" U.S. Department of Commerce, Technical Paper 15, Washington, D.C., Government Printing Office.

Wallis, K.F. (1974) "Seasonal Adjustment and Relations Between Variables", JASA, 69, 18-31.

Whittle, P. (1963) Prediction and Regulation by Linear Least-Square Methods, Princeton, D. Van Nostrand Co.

Young, A.H. (1968) "Linear Approximations to the Census and BLS Seasonal Adjustment Procedures", JASA, 63, 445-471.

Appendix to the Report

In this appendix we present some details about the individual series that were modeled during the year. Table A.1 gives a legend of the series ID's that are used throughout the report. Table A.1 gives a list of the series used in the empirical comparison grouped by the type of ARIMA model which was built for the series. Parameter estimates for the ARIMA model parameters are given along with an indication of whether or not outliers were included in the model. Note that the remaining tables and plots of the series in the empirical study are arranged in the same order as this table. Table A.3 gives a summary of the first, second and third year revision measures for each individual series. Table A.4 gives the smoothness and fit measures for the individual series. For each of the series considered in the empirical study we have included two diagrams. The first diagram includes for the model based approach a plot of the original series and seasonally adjusted series (in a dashed line) on one graph and a plot of the seasonal factors for the model based adjustment on a separate graph. The second diagram includes the analogous plots for the X-11 adjustment. Finally, during this year some series that included Easter holiday effects were modeled but not included in the empirical study. All of these series were modeled from 1/67 through 9/79. The models for these series are given in table A.5.

Table A.1

Lengend of Series ID's

S07TI:	inventories - glass containers
S11TI:	inventories - blast furnaces
S13TI:	inventories - nonferrous metals
S16TI:	inventories - metals, cans, barrels, drums
S23TI:	inventories - farm machinery and equipment
S24TI:	inventories - construction mining material handling
S25TI:	inventories - metal working machinery
S29TI:	inventories - general industry machinery
S35TI:	inventories - household appliances
S36TI:	inventories - radio and T.V.
S37TI:	inventories - commercial equipment
S38TI:	inventories - electircal components
S45TI:	inventories - ship building
S60TI:	inventories - meat products
S62TI:	inventories - beverages
S63TI:	inventories - fats and oils
S64TI:	inventories - all other nondurable products
S65TI:	inventories - tobacco
S76TI:	inventories - paperboard containers
S83TI:	inventories - industrial chemicals
S85TI:	inventories - drugs, soap, toiletries
SX2TI:	inventories - electrical transmition and distribution equipment and industrial apparatus
SX4TI:	inventories - motor vehicle and parts
SX5TI:	inventories - aircraft, missiles, parts
TI500:	wholesale inventories - U.S. total
TI501:	wholesale inventories - motor vehicles, automotive parts and supplies
TI502:	wholesale inventories - furniture and home furnishings
TI503:	wholesale inventories - lumber and other construction materials
TI506:	wholesale inventories - electrical goods
TI507:	wholesale inventories - hardware, plumbing, heating equipment and supplies

TI508:	wholesale inventories - machinery, equipment and supplies
TI517:	wholesale inventories - petroleum and petroleum products
501000:	wholesale sales - motor vehicles, automotive parts and supplies
502000:	wholesale sales - furniture and home furnishings
503000:	wholesale sales - lumber and other construction materials
504000:	wholesale sales - sporting, recreational, photographic goods
505000:	wholesale sales - metals and minerals except petroleum
506000:	wholesale sales - electrical goods
507000:	wholesale sales - hardware, plumbing, heating equipment and supplies
508000:	wholesale sales - machinery, equipment, and supplies
511000:	wholesale sales - paper and paper products
512000:	wholesale sales - drugs, drug proprietaries, drugest' sundries
513000:	wholesale sales - apparel, piece goods and notions
514000:	wholesale sales - groceries and related products
515000:	wholesale sales - farm product raw materials
516000:	wholesale sales - chemicals and allied products
517000:	wholesale sales - petroleum and petroleum products
518000:	wholesale sales - beer, wine and distilled alcoholic beverages
701000:	retail service receipts - hotels, motels, and tourist courts
721000:	retail service receipts - laundries, laundry services and cleaning and dyeing plants
723000:	retail service receipts - beauty shops
724000:	retail service receipts - barber shops
731000:	retail service receipts - advertising
750000:	retail service receipts - automotive repair
753000:	retail service receipts - automotive repair shops
760000:	retail service receipts - misc. repair services
S07VSU:	value shipped - glass containers
S11VSU:	value shipped - blast furnaces
S13VSU:	value shipped - nonferrous metals
S16VSU:	value shipped - metals, cans, barrels, drums
S21VSU:	value shipped - steam engines and turbines
S23VSU:	value shipped - farm machinery and equipment
S24VSU:	value shipped - construction, mining, material handling
S25VSU:	value shipped - metal working machinery
S29VSU:	value shipped - general industry machinery
S35VSU:	value shipped - household appliances

S36VSU:	value shipped - radio and T.V.
S38VSU:	value shipped - electrical components
S48VSU:	value shipped - scientific and engineering
S50VSU:	value shipped - photographic good
S60VSU:	value shipped - meat products
S62VSU:	value shipped - beverages
S63VSU:	value shipped - fats and oils
S64VSU:	value shipped - all other nondurable products
S76VSU:	value shipped - paperboard containers
S83VSU:	value shipped - industrial chemicals
S85VSU:	value shipped - drugs, soap, toiletries
SHIP:	total industry shipments
INVE:	total industry inventories
VIP:	value put in place
EAF16:	employed argicultural females 16-19
ENM20:	employed non-agricultural males 20 and up
EAF20:	employed agricultural females 20 and up
EAM20:	employed agricultural males 20 and up
UM16:	unemployed males 16-19
UF16:	unemployed females 16-19
ENM16:	employed non-agricultural males 16-19
ENF20:	employed non-agricultural females 20 and up
EAM16:	employed agricultural males 16-19
HPT:	total housing permits
HST1:	total single family housing starts
HSTT:	total housing starts
HST5:	total five unit housing starts
520002:	retail sales - lumber, building materials, paint, glass wallpaper
525100:	retail sales - hardward stores
550001:	retail sales - motor vehicles dealers, boat dealers, recreational and utility trailer dealers, motorcycle dealers
553100:	retail sales - auto and home supply stores
570001:	retail sales - furniture stores, floor covering stores, drapery curtain and upholstery stores, misc. home furnishing
570002:	retail sales - household appliance stores, radio and television stores
531100:	retail sales - department stores

533100:	retail sales - variety stores
539900:	retail sales - misc. general merchandise stores
541100:	retail sales - grocery stores
554100:	retail sales - gasoline service stations
561100:	retail sales - mens and boys clothing and furnishing stores
566100:	retail sales - shoe stores
560001:	retail sales - women's ready to wear stores, women's accessary and specialty stores, furriers and fur shops
580000:	retail sales - restaurants and lunchrooms, social caterers, cafeterias, refreshment places, contract feeding, ice cream and frozen custard stands, drinking places
591200:	retail sales - drug stores and proprietary stores
592100:	retail sales - liquor stores
596101:	retail sales - mail order houses
594400:	retail sales - stationary stores
598000:	retail sales - fuel oil dealers, liquefied petroleum gas dealers, fuel and ice dealers

Table A.2

The 76 series used in the Empirical Comparison Grouped by Model Type

(* under outliers indicates outliers were modeled)

Model: $(1-B)(1-B^{12})(\ln Z_t - TD_t) = (1-\theta_1 B)(1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\theta}_1$	$\hat{\theta}_{12}$	Outliers
501000	1/67 - 11/79	.37	.79	
503000	1/67 - 11/79	0.	.87	
504000	1/67 - 11/79	.29	.90	*
505000	1/67 - 11/79	0.	.88	
506000	1/67 - 11/79	0.	.87	*
507000	1/67 - 11/79	.20	.74	
508000	1/67 - 11/79	.29	.88	*
511000	1/67 - 11/79	0.	.91	
512000	1/67 - 11/79	.26	.54	
513000	1/67 - 11/79	.34	.87	
514000	1/67 - 11/79	.41	.91	*
515000	1/67 - 11/79	0.	.91	
516000	1/67 - 11/79	.27	.84	*
517000	1/67 - 11/79	0.	.86	*
518000	1/67 - 11/79	.49	.70	*
520002	1/67 - 9/79	0.	.86	
550001	1/67 - 9/79	0.	.87	*
598000	1/67 - 9/79	.38	.87	*
570001	1/67 - 9/79	.27	.73	
S13VSU	1/63 - 12/78	0.	.92	*
S16VSU	1/63 - 12/78	.64	.88	*
S29VSU	1/63 - 12/78	.34	.91	*
S35VSU	1/63 - 12/78	.31	.88	
S38VSU	1/67 - 12/78	0.	.95	*
S50VSU	1/63 - 12/78	.53	.89	
S60VSU	1/63 - 12/78	.22	.88	*
S62VSU	1/67 - 12/78	.51	.87	
S63VSU	1/67 - 12/78	0.	.88	*
S64VSU	1/63 - 12/78	.26	.86	*

S76VSU	1/63 - 12/78	.31	.92	*
S83VSU	1/63 - 12/78	0.	.56	*
701000	1/71 - 11/79	.24	.84	*
723000	1/71 - 11/79	0.	.83	*
724000	1/71 - 11/79	0.	.88	
731000	1/71 - 11/79	.30	.59	*

Model: $(1-B)(1-B^{12})(\ln Z_t - TD_t) = (1-\theta_1 B-\theta_2 B^2)(1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_{12}$	Outliers
S24VSU	1/63 - 12/78	.17	.19	.77	*
750000	1/71 - 11/79	0.	.45	.86	*
753000	1/71 - 11/79	.29	.46	.87	*
570002	1/67 - 9/79	.17	.34	.88	*

Model: $(1-\phi B)(1-B)(1-B^{12})(\ln Z_t - TD_t) = (1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\phi}$	$\hat{\theta}_{12}$	Outliers
S25VSU	1/63 - 12/78	-.42	.91	

Model: $(1-B)(1-B^{12})\ln Z_t = (1-\theta_1 B)(1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\theta}_1$	$\hat{\theta}_{12}$	Outliers
S16TI	1/58 - 6/79	-.19	.90	*
S60TI	1/58 - 6/79	0.	.85	
S63TI	1/70 - 6/79	0.	.87	*
S65TI	1/58 - 6/79	0.	.78	
SX4TI	1/58 - 6/79	0.	.68	*
TI503	1/67 - 4/79	0.	.64	
TI506	1/67 - 4/79	0.	.90	
TI507	1/67 - 4/79	0.	.87	*
SHIP	1/58 - 8/78	0.	.80	*

Model: $(1-B)(1-B^{12})\ln Z_t = (1-\theta_1 B - \theta_2 B^2)(1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_{12}$	Outliers
S21VSU	1/63 - 12/78	.60	.25	.88	

Model: $(1-\phi B)(1-B)(1-B^{12})\ln Z_t = (1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\phi}$	$\hat{\theta}_{12}$	Outliers
ENM20	1/65 - 8/79	.26	.88	
INVE	1/58 - 8/78	.66	.91	*
VIP	1/66 - 12/77	.88	.80	*
S36TI	1/60 - 6/79	.26	.91	*
S62TI	1/58 - 6/79	.13	.79	*
S64TI	1/58 - 6/79	.32	.79	
S85TI	1/58 - 6/79	.35	.66	*
TI500	1/67 - 4/79	.39	.90	*

Model: $(1-B)(1-B^{12})Z_t = (1-\theta_1 B)(1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\theta}_1$	$\hat{\theta}_{12}$	Outliers
EAF20	1/65 - 8/79	.32	.58	*
UM16	1/65 - 8/79	.31	.82	
UF16	1/65 - 8/79	.61	.74	
ENM16	1/65 - 8/79	.25	.87	
ENF16	1/65 - 8/79	.22	.87	
EAM20	1/65 - 8/79	0.	.88	*
ENF20	1/65 - 8/79	0.	.90	*
HST5	1/64 - 8/78	.45	.92	
HSTT	1/64 - 8/78	.28	.89	*
HST1	1/64 - 8/78	.25	.87	
721000	1/71 - 11/79	.25	.86	

Model: $(1-\phi B)(1-B^{12})Z_t = (1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\phi}$	$\hat{\theta}_{12}$	Outliers
EAM16	1/65 - 8/79	.54	.90	*
EAF16	1/65 - 8/79	.51	.77	*

Model: $(1-\phi B)(1-B)(1-B^{12})Z_t = (1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\phi}$	$\hat{\theta}_{12}$	Outliers
S35TI	1/60 - 6/79	.40	.88	*

Model: $(1-\phi_1 B - \phi_2 B^2)(1-B)(1-B^{12})Z_t = (1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\theta}_{12}$	Outliers
S23TI	1/58 - 6/79	.27	.20	.91	*

Model: $(1-B)(1-B^{12})(Z_t^{\frac{1}{2}} - TD_t) = (1-\theta_1 B)(1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\theta}_1$	$\hat{\theta}_{12}$	Outliers
525100	1/67 - 9/79	.30	.36	
502000	1/67 - 11/79	.39	.87	

Model: $(1-B)(1-B^{12})(Z_t^{1/3} - TD_t) = (1-\theta_1 B)(1-\theta_{12}B^{12})a_t$

Series ID	Time Period Available	$\hat{\theta}_1$	$\hat{\theta}_{12}$	Outliers
553100	1/67 - 9/79	.31	.56	

Table A.3

Revision Measures for Individual Series

(* under X-11 ARIMA indicates that an ARIMA model was forced)

Series ID	First Year Revisions		Second Year Revisions		Third Year Revisions	
	Model Based	X-11 ARIMA	Model Based	X-11 ARIMA	Model Based	X-11 ARIMA
501000	.59	1.15	.76	1.16	1.06	1.18
503000	.42	.88	.43	.67	.61	.72
504000	.88	1.06	.49	.81	.47	1.74
505000	.40	.55*	.30	.69*	.30	.75*
506000	.29	.56	.40	.44	.49	.94
507000	.86	1.72	1.21	1.68	.96	1.08
508000	.76	1.49	.61	1.25	.50	1.11
511000	.46	1.67	.27	.91	.28	.85
512000	1.25	.82	.82	.49	1.16	.75
513000	.33	.66	.55	.81	.60	.90
514000	.52	1.22	.41	1.19	.43	1.25
515000	.79	1.59	.42	1.02	.37	.83
516000	.77	1.19	.60	.83	.47	.76
517000	.62	1.35*	.45	.90*	.48	.87*
518000	.51	.93	.71	.65	.94	.74
520002	.58	1.25	.84	1.21	.76	1.01
550001	.20	.99	.28	.66	.28	.76
598000	1.70	2.27	1.03	1.33	.72	.86
570001	.41	1.02	.67	.87	.73	.74
S13VSU	1.07	1.08*	.68	1.10*	.63	1.07*
S16VSU	.35	.80	.34	.90	.32	.98
S29VSU	.83	1.65	.44	1.01	.43	1.21
S35VSU	.91	2.41	.48	1.41	.46	1.38
S38VSU	1.07	2.21	.41	.95	.23	1.00
S50VSU	.33	1.34	.30	1.21	.23	.88
S60VSU	.42	.81	.35	1.04	.49	1.19
S62VSU	1.72	1.67	1.19	1.13	.78	.99
S63VSU	1.31	2.28*	.84	1.88*	.62	1.22*

Table A.3 (con't.)

Series ID	First Year Revisions		Second Year Revisions		Third Year Revisions	
	Model Based	X-11 ARIMA	Model Based	X-11 ARIMA	Model Based	X-11 ARIMA
S64VSU	.70	1.44	.86	1.81	.76	1.76
S76VSU	.93	3.32	.66	2.21	.46	1.51
S83VSU	1.52	1.73*	1.32	1.40*	1.03	1.04*
701000	.35	.76	—	—	—	—
723000	.72	.99	—	—	—	—
724000	.71	1.24	—	—	—	—
731000	.40	1.18	—	—	—	—
S24VSU	1.99	1.03*	1.26	.89*	.75	.88*
750000	.31	.81	—	—	—	—
753000	.33	.60	—	—	—	—
570002	.64	.93	.52	.83	.47	.79
S25VSU	.23	.95	.25	.99	.27	1.02
S16TI	.39	1.33*	.30	.82*	.39	.89*
S60TI	.64	1.67	.52	1.01	.45	.65
S63TI	.26	.98	1.19	2.22	.60	1.14
S65TI	.77	1.33	.52	.90	.56	.86
SX4TI	1.14	.97	.76	.71	.64	.64
TI503	.86	1.13*	.57	.78*	.45	.49*
TI506	.63	2.06	.52	1.19	.53	.86
TI507	.49	1.72*	.80	1.38*	.67	1.06*
SHIP	.45	1.01	.38	.91	.47	.79
S21VSU	.31	.86*	.26	.65*	.39	.73*
ENM20	.38	1.02	.68	1.18	.27	1.05
INVE	.17	1.04	.28	.81	.47	.95
VIP	.43	1.64	.35	.97	.45	1.10
S36TI	.30	1.61	.32	1.03	.44	1.13
S62TI	.41	.61*	.76	.87*	.72	.73*
S64TI	.51	1.02	.56	1.24	.71	1.17
S85TI	.64	.59*	1.00	.98*	1.08	1.05*
TI500	.25	1.10	.47	1.40	.35	.90
EAF20	.93	1.29*	.80	1.04*	.71	.86*

Table A.3 (con't.)

Series ID	First Year Revisions		Second Year Revisions		Third Year Revisions	
	Model Based	X-11 ARIMA	Model Based	X-11 ARIMA	Model Based	X-11 ARIMA
UM16	.45	1.20*	.64	.94*	.63	1.15*
UF16	1.02	1.58*	1.02	1.35*	1.26	1.32*
ENM16	.93	2.33	.64	1.09	.54	.83
ENF16	.72	1.57	.77	1.14	.62	.89
EAM20	.19	.92	.27	.75	.28	.68
ENF20	.31	1.31	.45	1.26	.61	1.07
HST5	.13	.81	.34	.73	.36	.63
HSTT	.24	.98	.25	.69	.56	.74
HST1	.44	.86	.62	.89	1.04	1.11
721000	.41	1.36	—	—	—	—
EAM16	.28	.83	.51	1.02	.56	.89
EAF16	.33	.91*	.52	1.07*	.77	1.57*
S35TI	.36	.35	.35	.83	.35	.70
525100	1.03	1.25	1.01	.97	1.00	.93
502000	.56	.65	.50	.75	.41	.71
553100	.57	1.36	.76	.93	.86	.87

Table A.4

Smoothness and Fit Measures for Individual Series

Series ID	SM	F
501000	1.07	.70
503000	1.16	.80
504000	1.18	.68
505000	1.13	.82
506000	1.12	.71
507000	1.05	.70
508000	1.05	.73
511000	1.00	.80
512000	.83	.90
513000	1.07	.71
514000	1.20	.63
515000	1.04	.69
516000	1.06	.82
517000	1.10	.60
518000	.99	.93
520002	1.13	.67
550001	1.13	.70
598000	1.05	.72
570001	1.10	.47
S13VSU	1.06	.54
S16VSU	1.09	.86
S29VSU	1.11	.66
S35VSU	1.11	.68
S38VSU	1.24	.44
S50VSU	1.13	.72
S60VSU	1.03	.63
S62VSU	2.39	.72
S63VSU	1.19	.48
S64VSU	1.11	.75
S76VSU	1.23	.78
S83VSU	.93	.87
701000	1.10	.46
723000	1.05	.54

Table A.4 (con't.)

Series ID	SM	F
724000	1.00	.51
731000	1.13	1.11
S24VSU	1.32	.21
750000	1.09	.60
753000	1.12	.69
570002	1.56	.09
S25VSU	1.13	.53
S16TI	1.13	.16
S60TI	1.16	.40
S63TI	1.02	1.69
S65TI	1.06	.62
SX4TI	.93	2.88
TI503	.96	1.62
TI506	1.15	.16
TI507	.95	.60
SHIP	1.06	.58
S21VSU	1.24	.10
ENM20	1.13	.29
INVE	1.08	.10
VIP	1.07	.21
S36TI	1.09	.14
S62TI	1.07	.69
S64TI	1.04	.56
S85TI	.75	4.23
TI500	1.22	.13
EAF20	.94	1.31
UM16	1.02	.67
UF16	.98	.88
ENM16	1.06	.27
ENF16	1.04	.29
EAM20	1.12	.23
ENF20	1.07	.14
HSTS	1.14	.09
HSTT	1.08	.19

Table A.4 (con't.)

Series ID	SM	F
HST1	1.05	.29
721000	1.18	.72
EAM16	1.07	.11
EAF16	1.02	.50
S35TI	.96	1.48
S23TI	1.09	.08

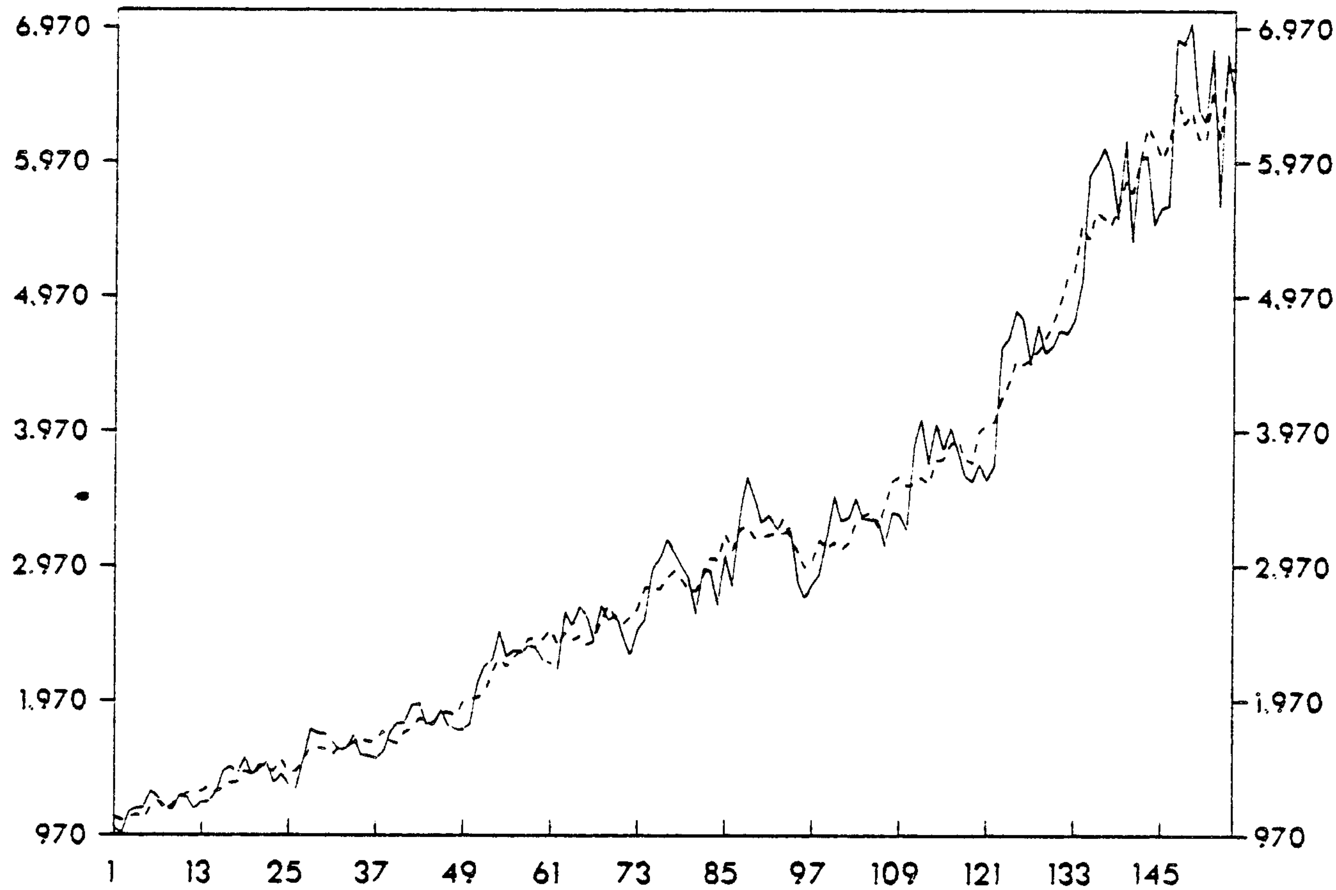
Table A.5

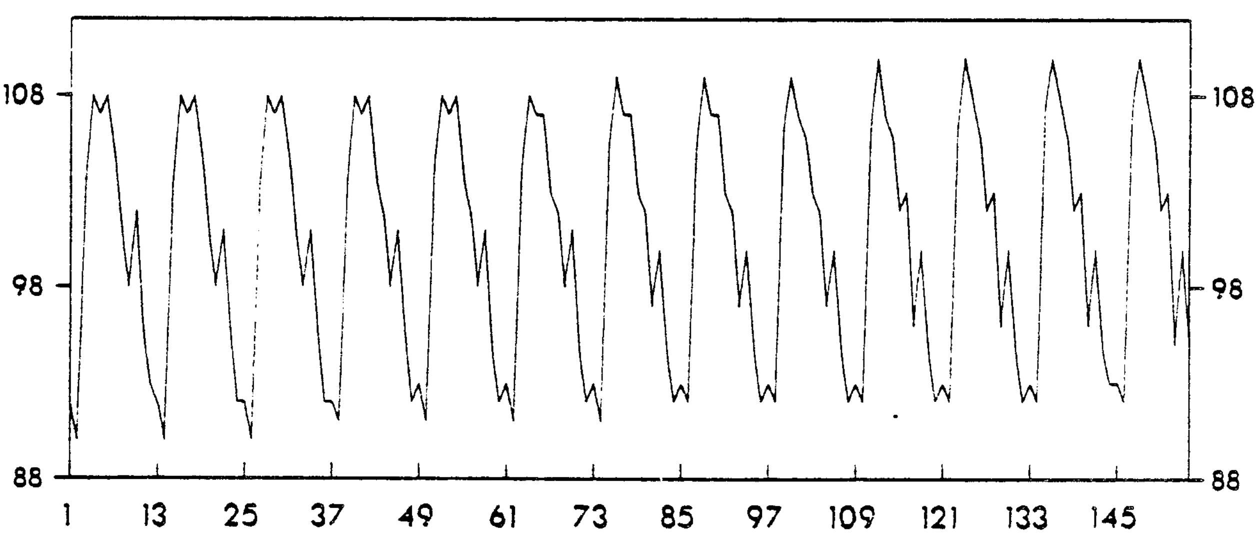
Easter Holiday Series Modeled

Series ID	Length of Holiday	Noise Model	Outliers	Intervention
580000	no effect	$(0,1,2)x(0,1,1)_{12}$	*	
566100	7 day	$(0,1,1)x(0,1,1)_{12}$	*	
561100	14 day	$(0,1,2)x(0,1,1)_{12}$		
531100	14 day	$(0,1,1)x(0,1,1)_{12}$		
533100	7 day	$(2,1,0)x(0,1,1)_{12}$	*	*
539900	7 day	$(0,1,1)x(0,1,1)_{12}$	*	
541100	7 day	$(2,1,0)x(0,1,1)_{12}$		
554100	no effect	$(0,1,0)x(0,1,1)_{12}$	*	
560001	14 day	+	*	

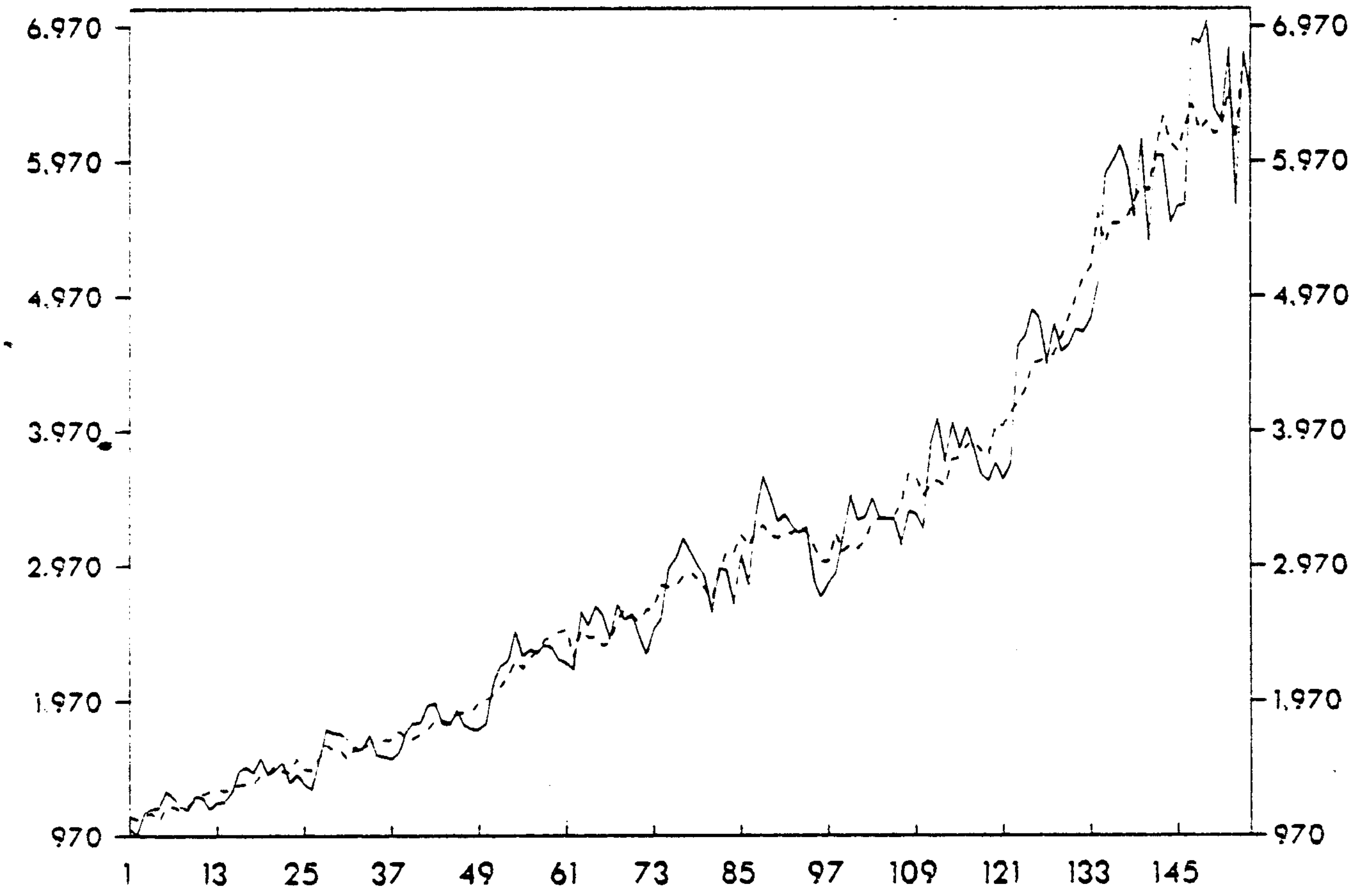
+ The noise model for this series was $(1-B)(1-B^{12})N_t = (1-\theta_1 B-\theta_{12}B^{12}-\theta_{13}B^{13})a_t$

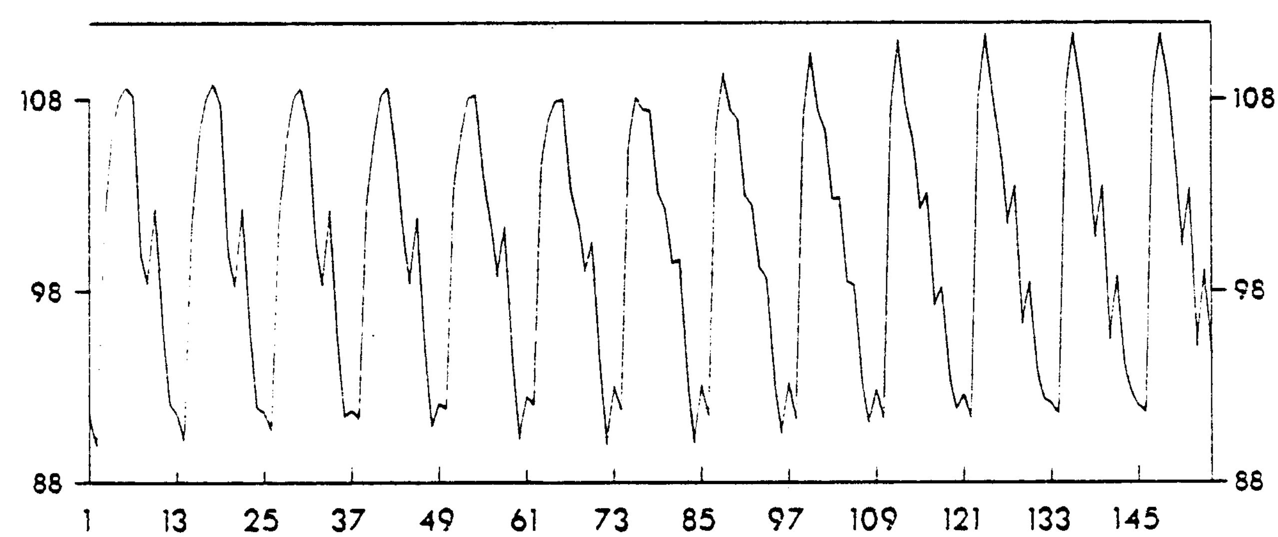
WHOLESALE SALES 501 MODEL BASED



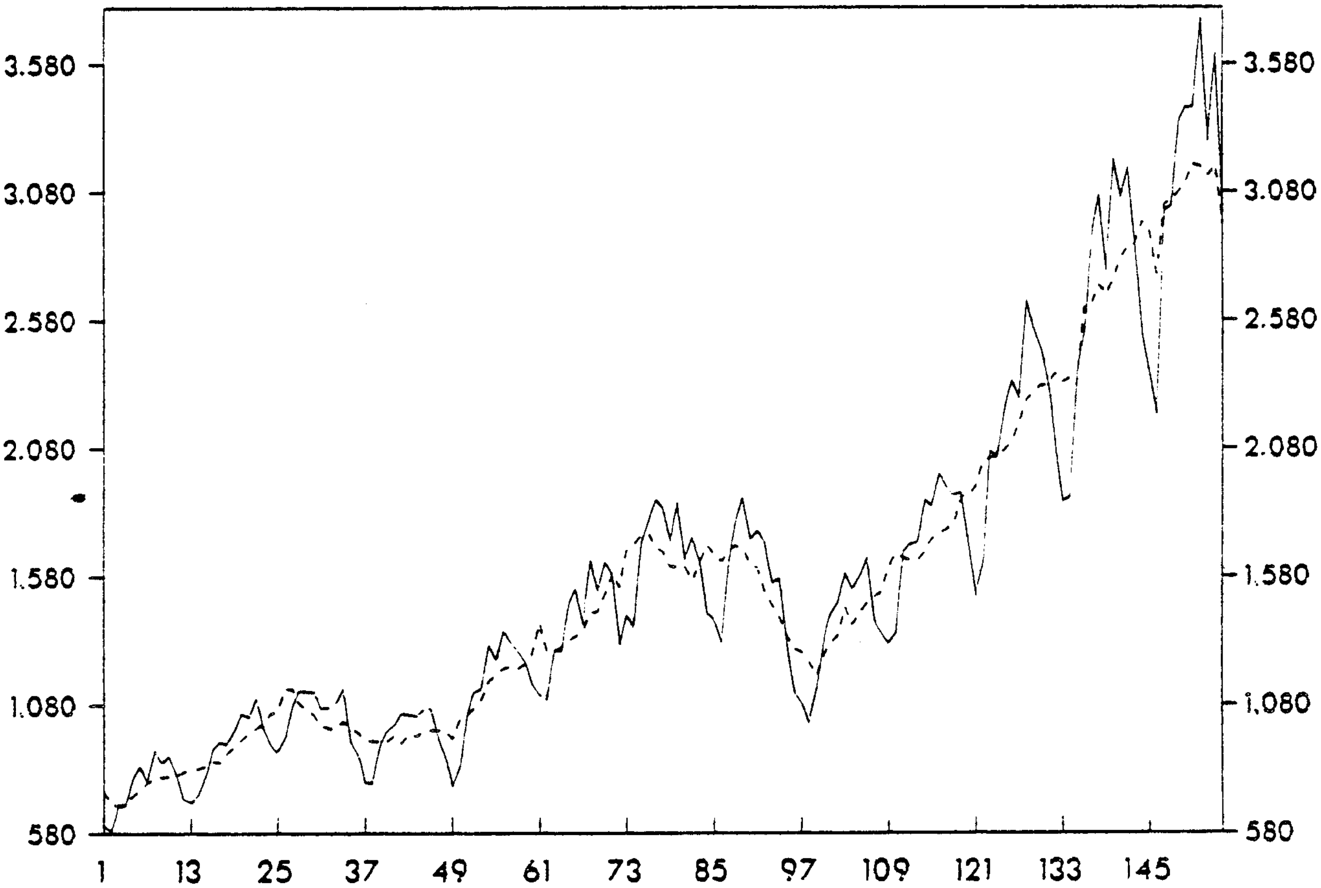


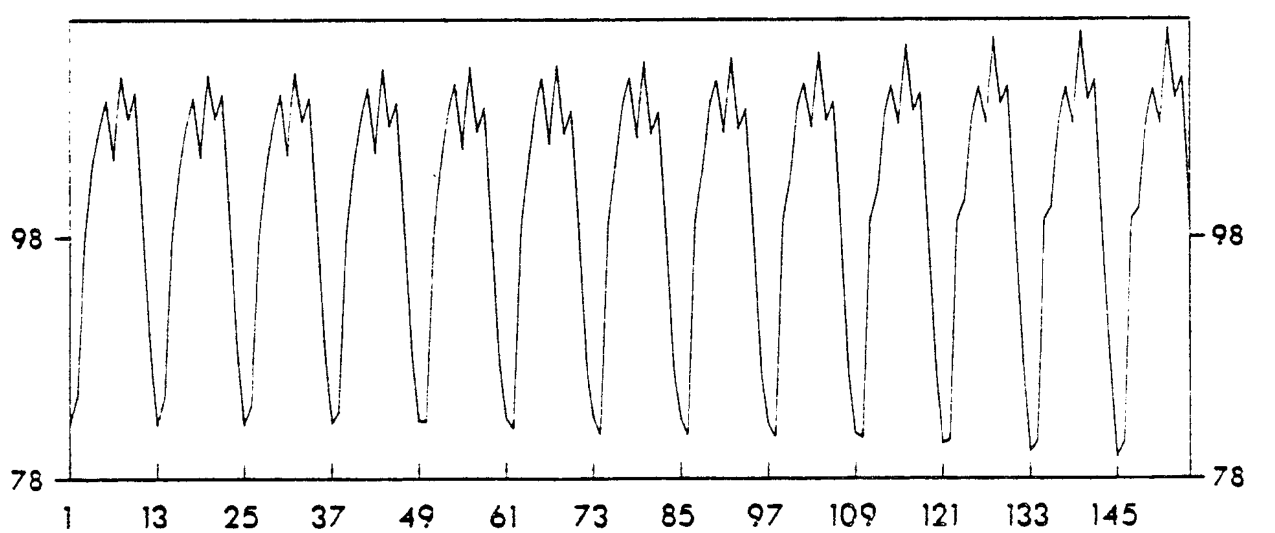
WHOLESALE SALES 501 X-11



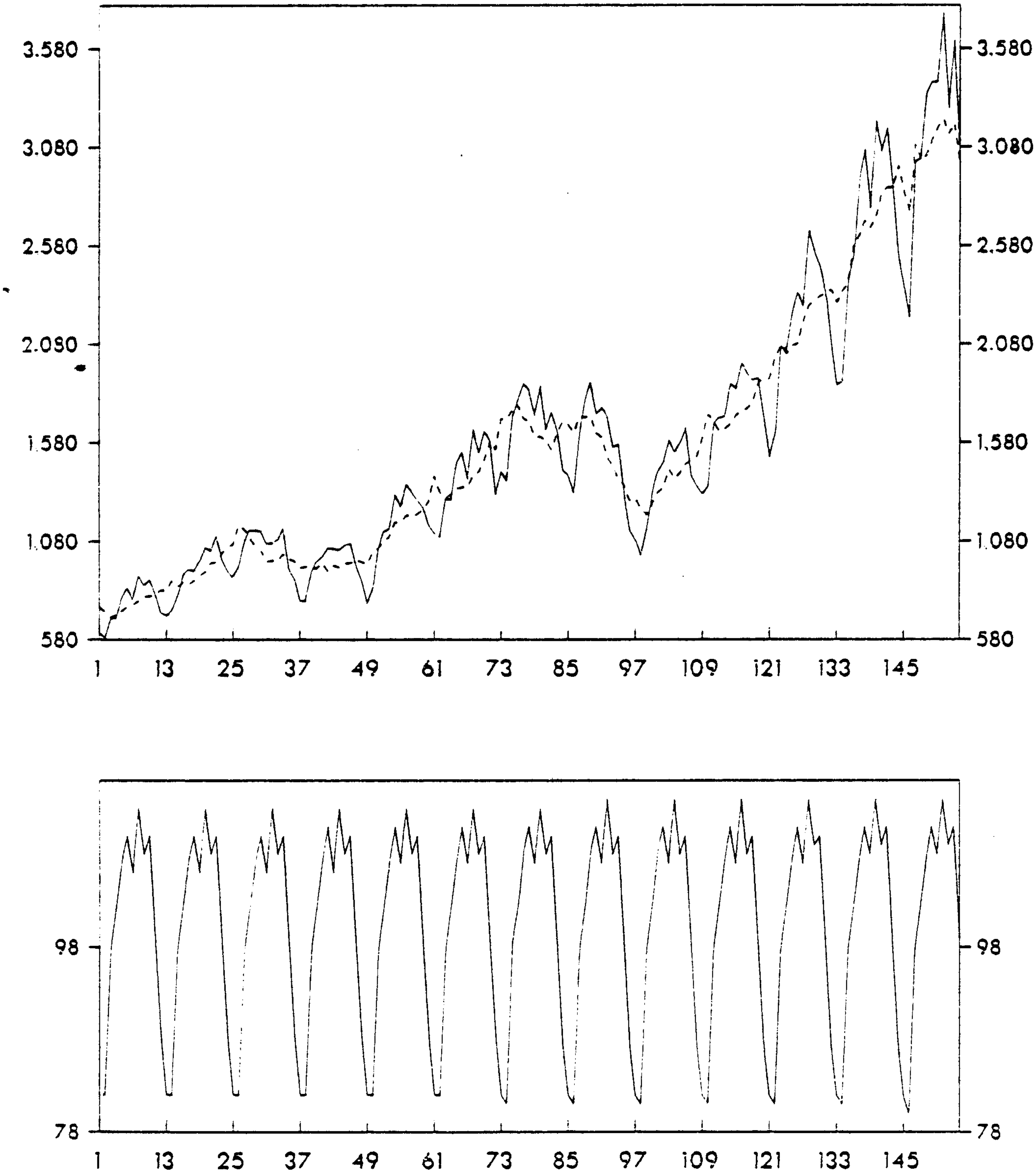


WHOLESALE SALES 503 X-11

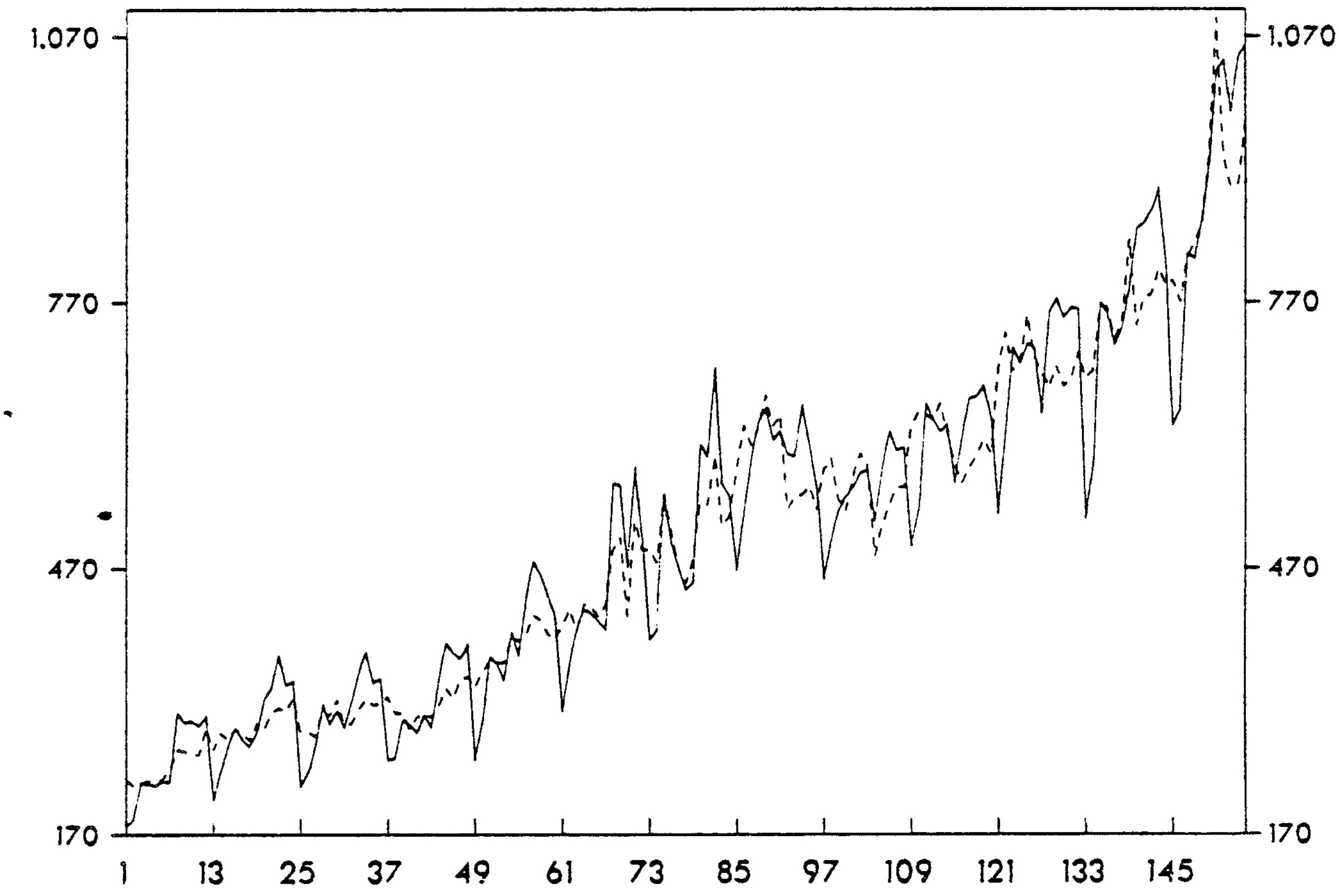


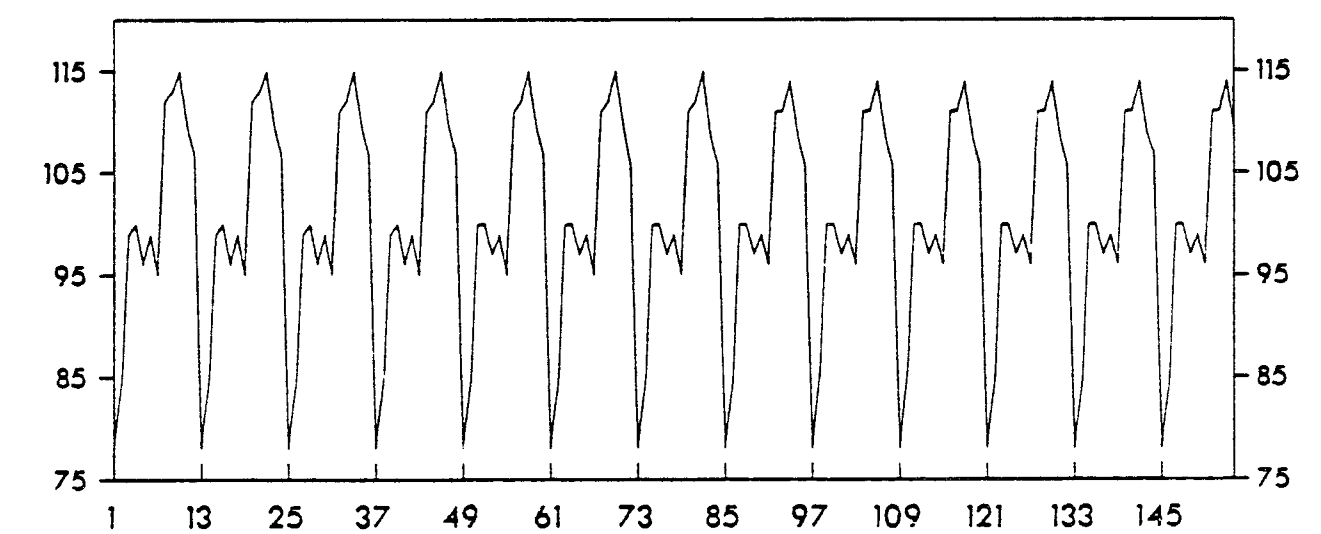


WHOLESALE SALES 503 MODEL BASED

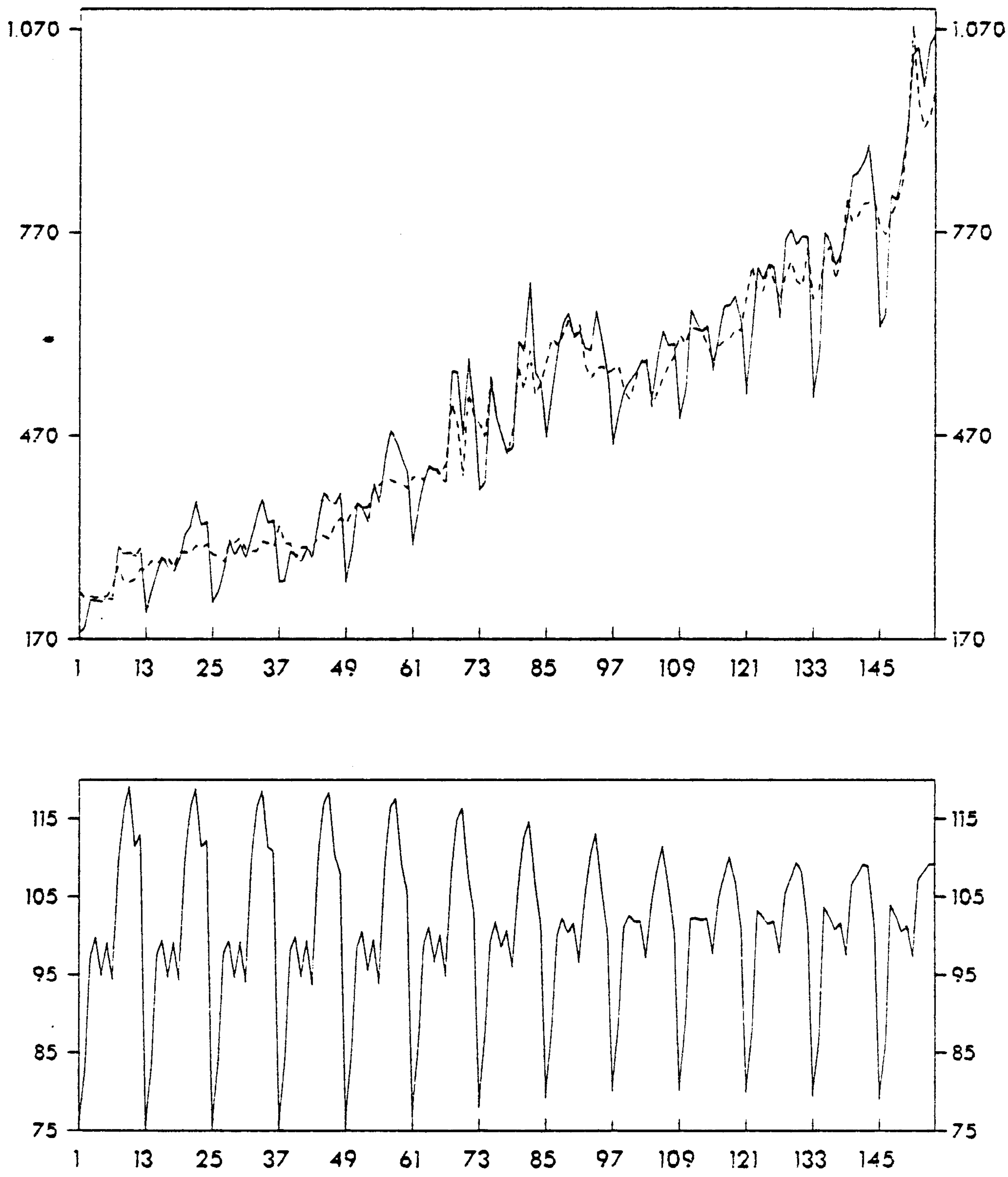


WHOLESALE SALES 504 MODEL BASED

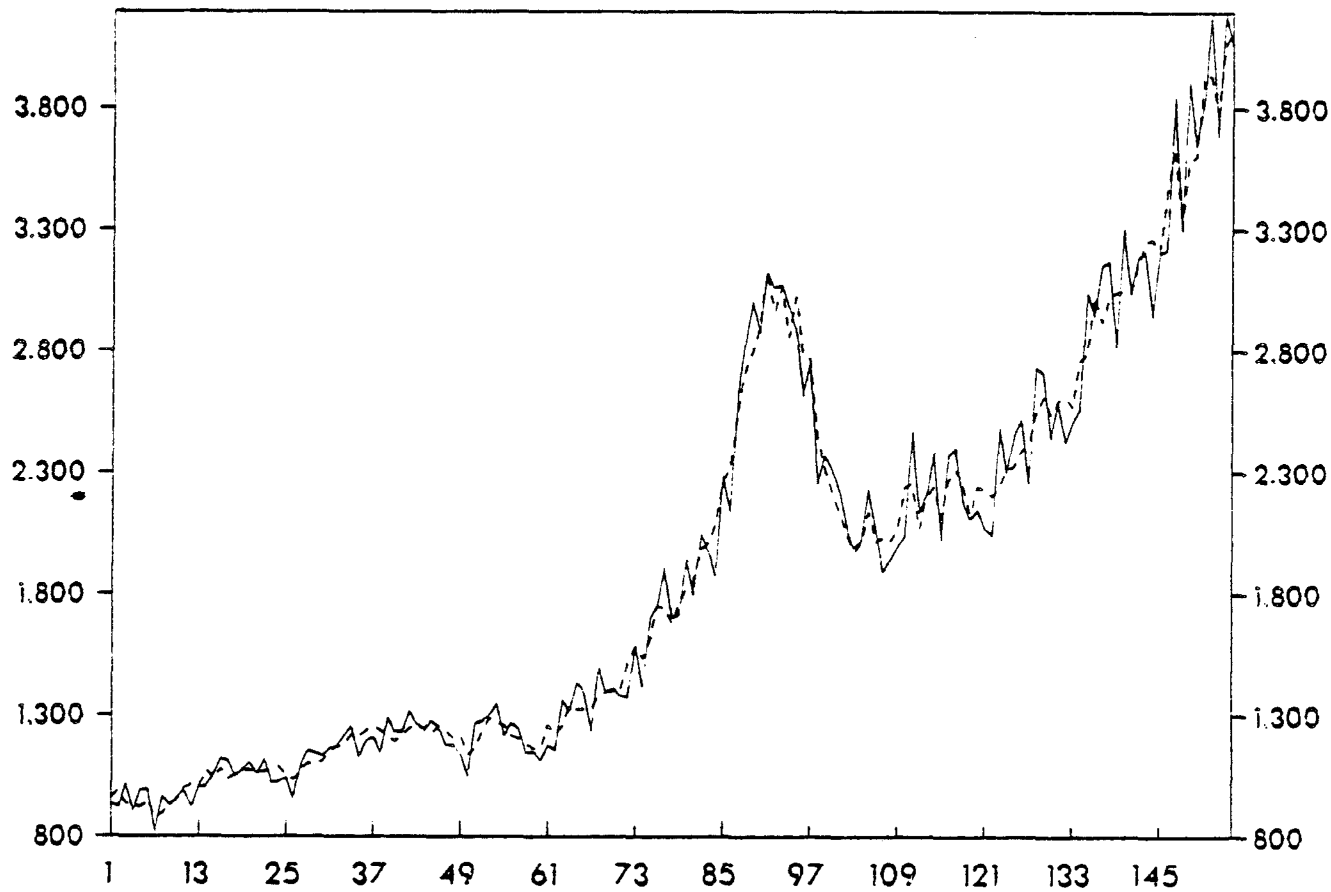


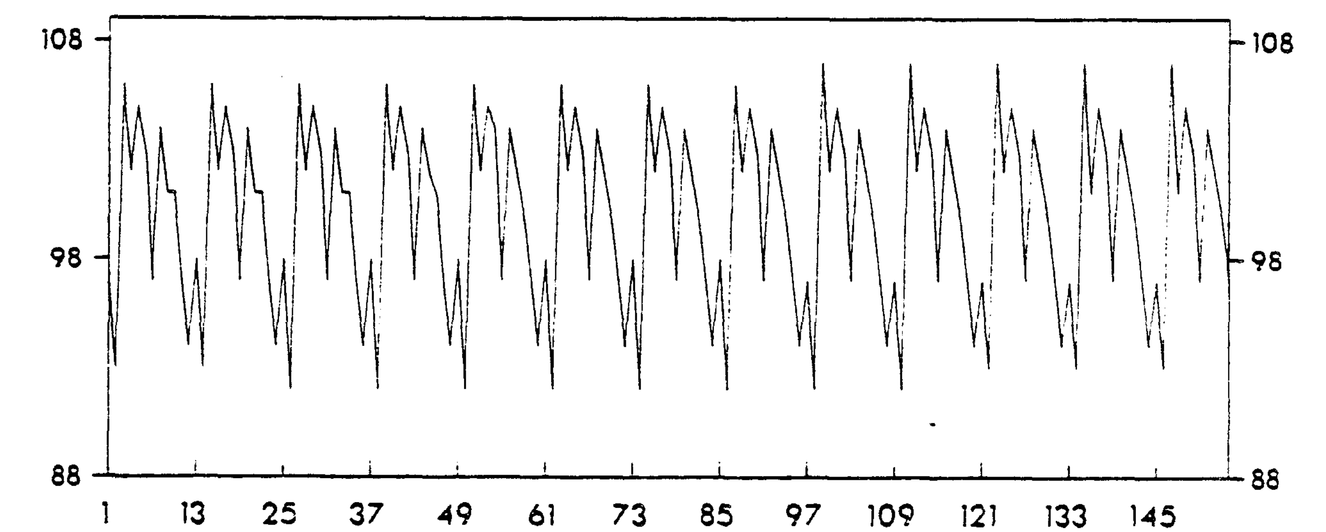


WHOLESALE SALES 504 X-11

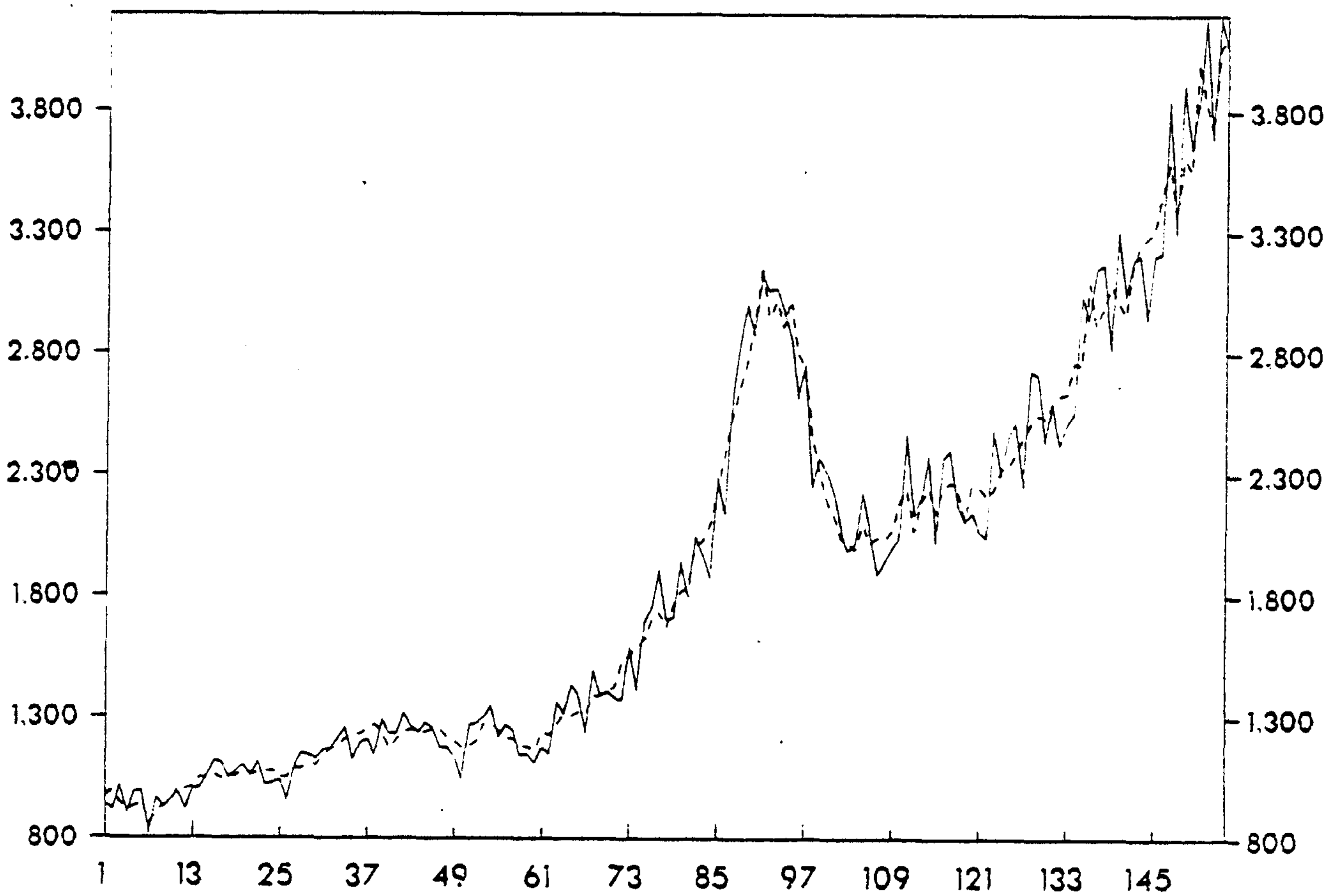


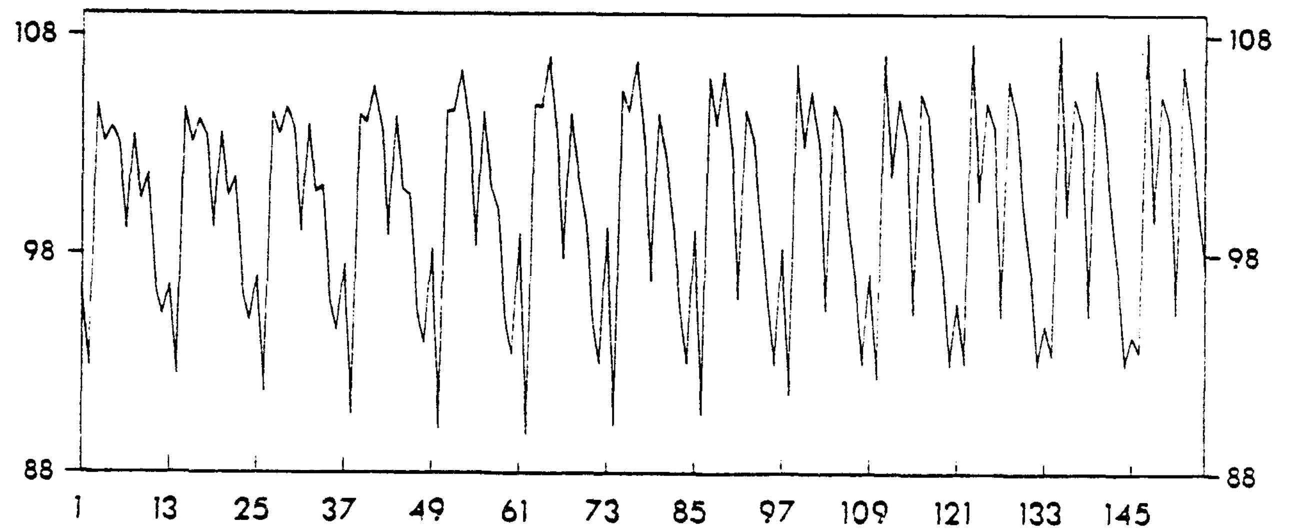
WHOLESALE SALES 505 MODEL BASED



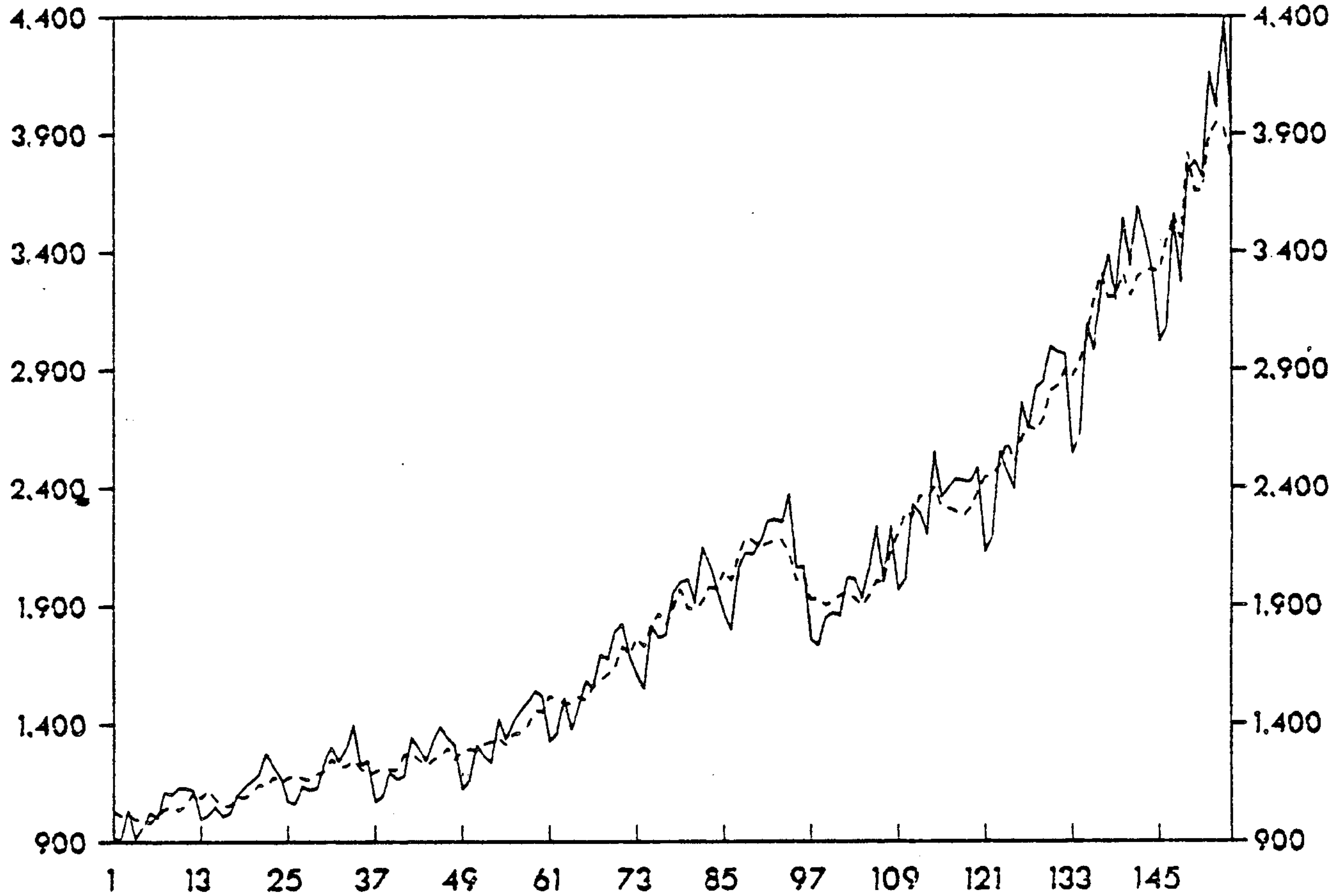


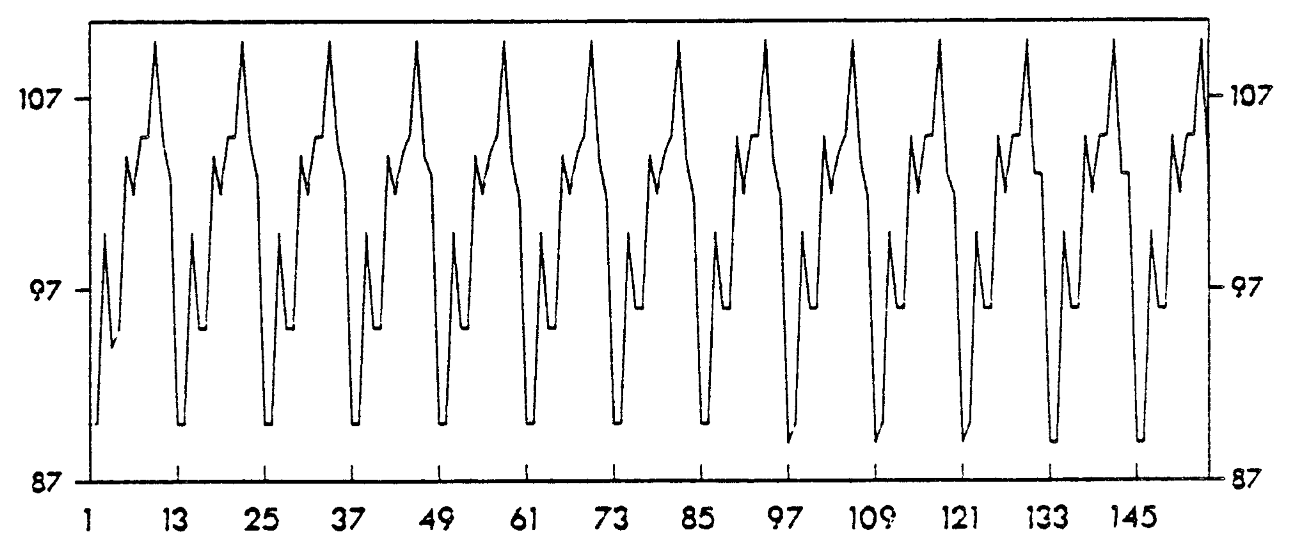
WHOLESALE SALES 505 X-11



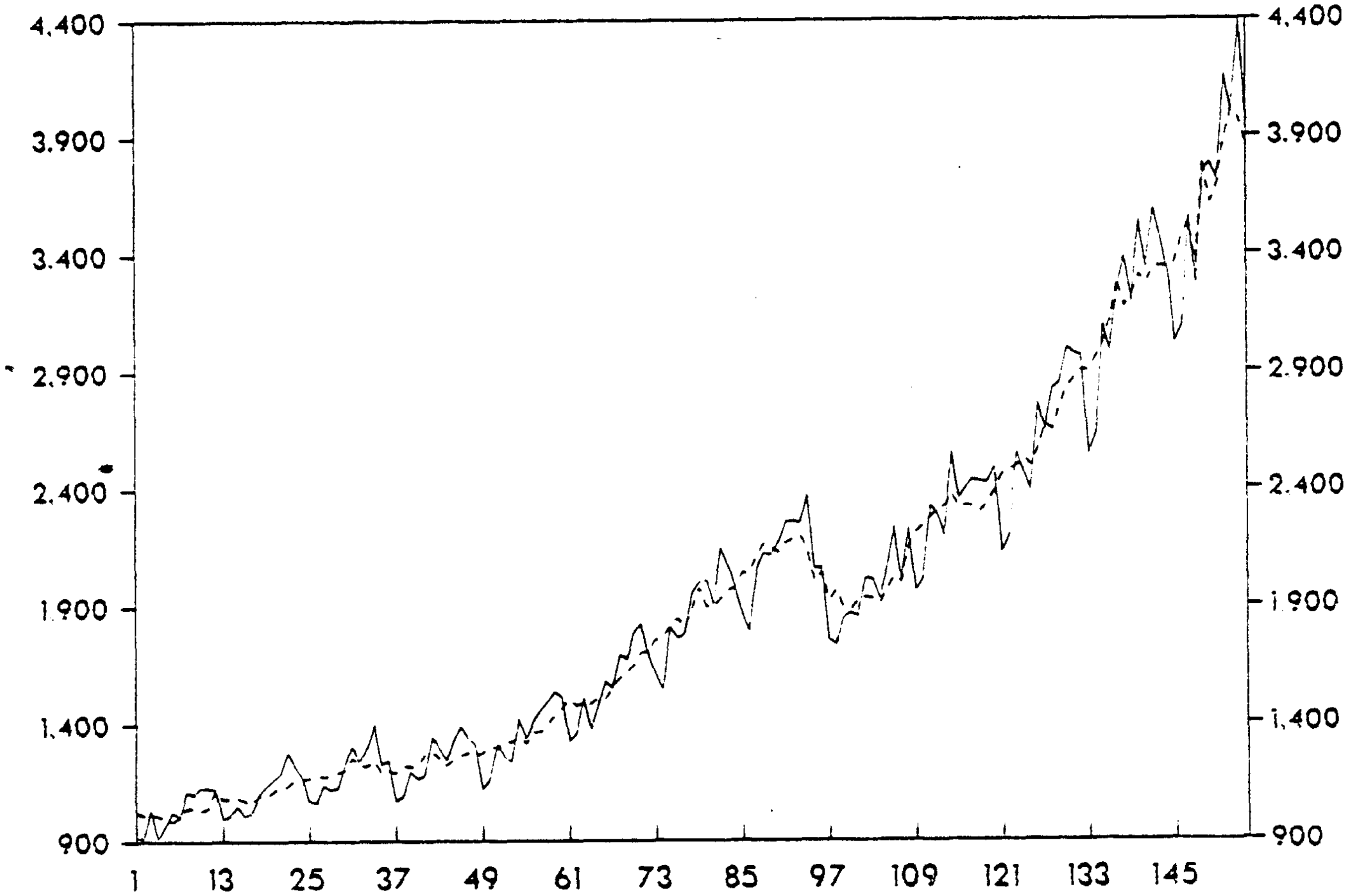


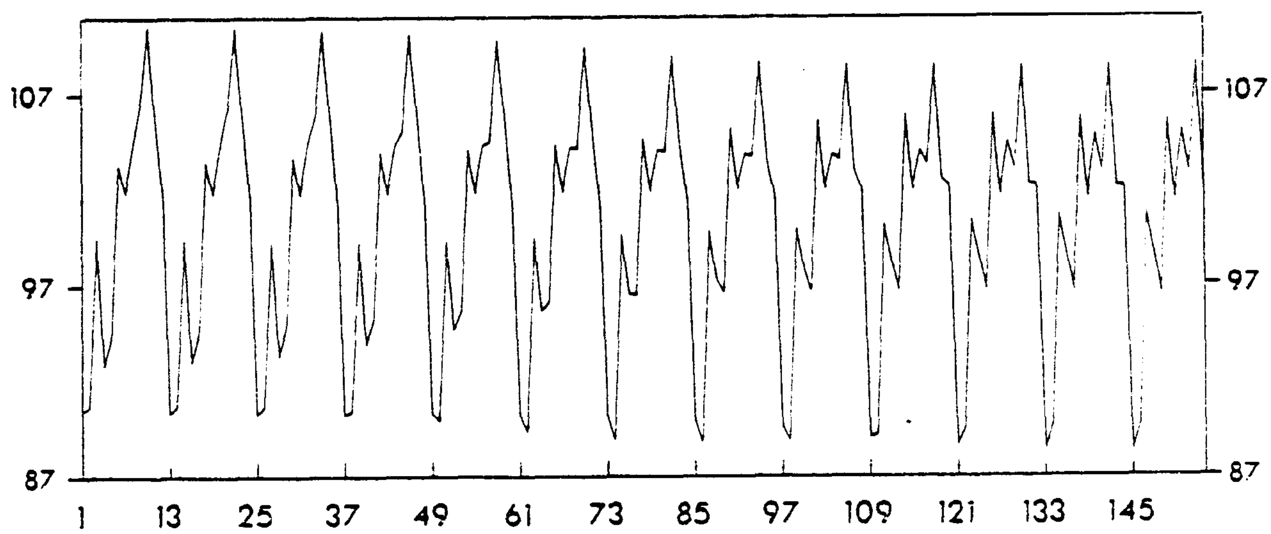
WHOLESALE SALES 506 MODEL BASED



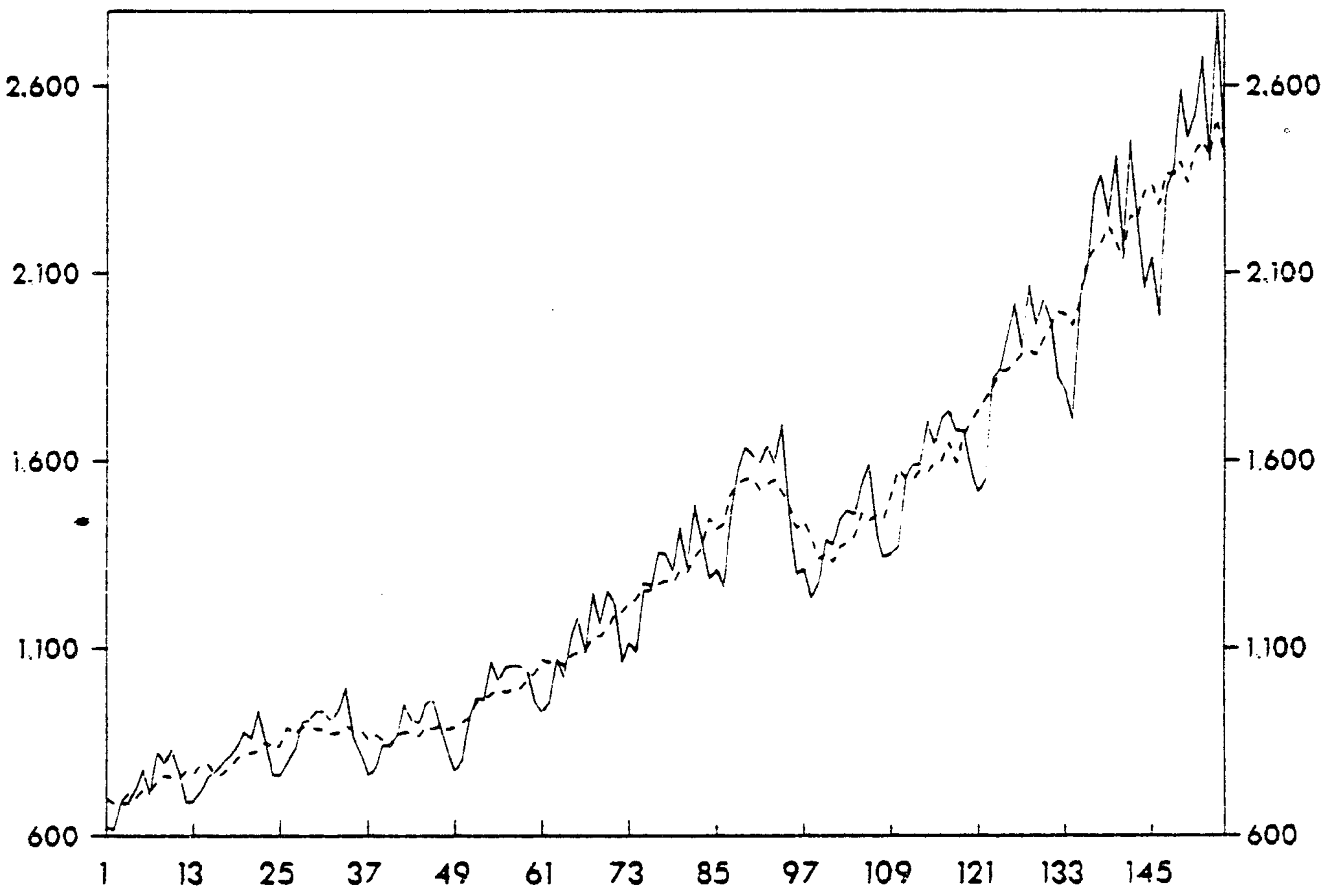


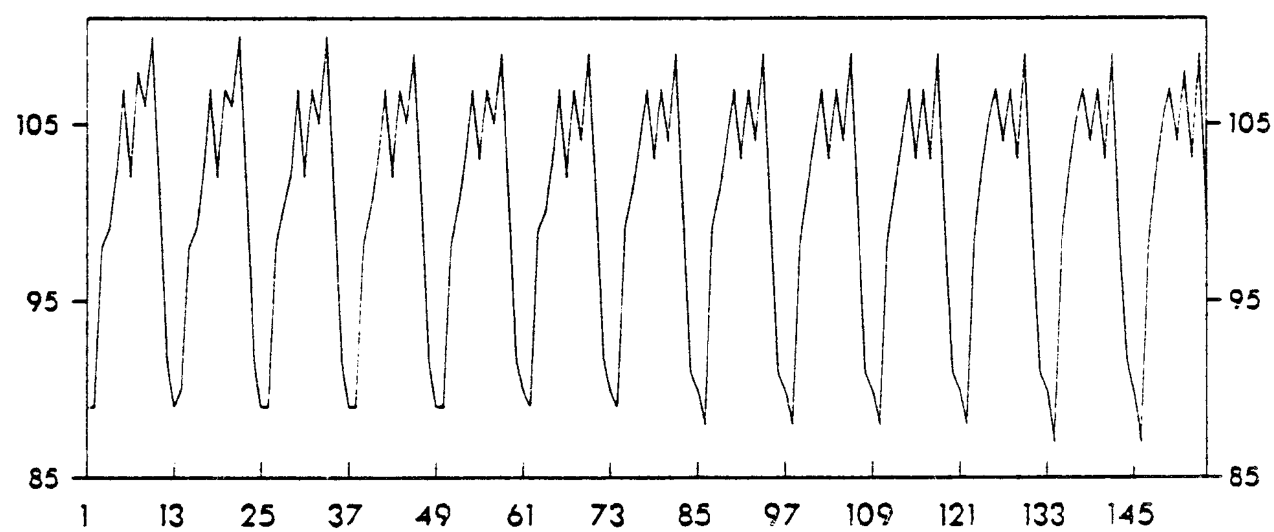
WHOLESALE SALES 506 X-11



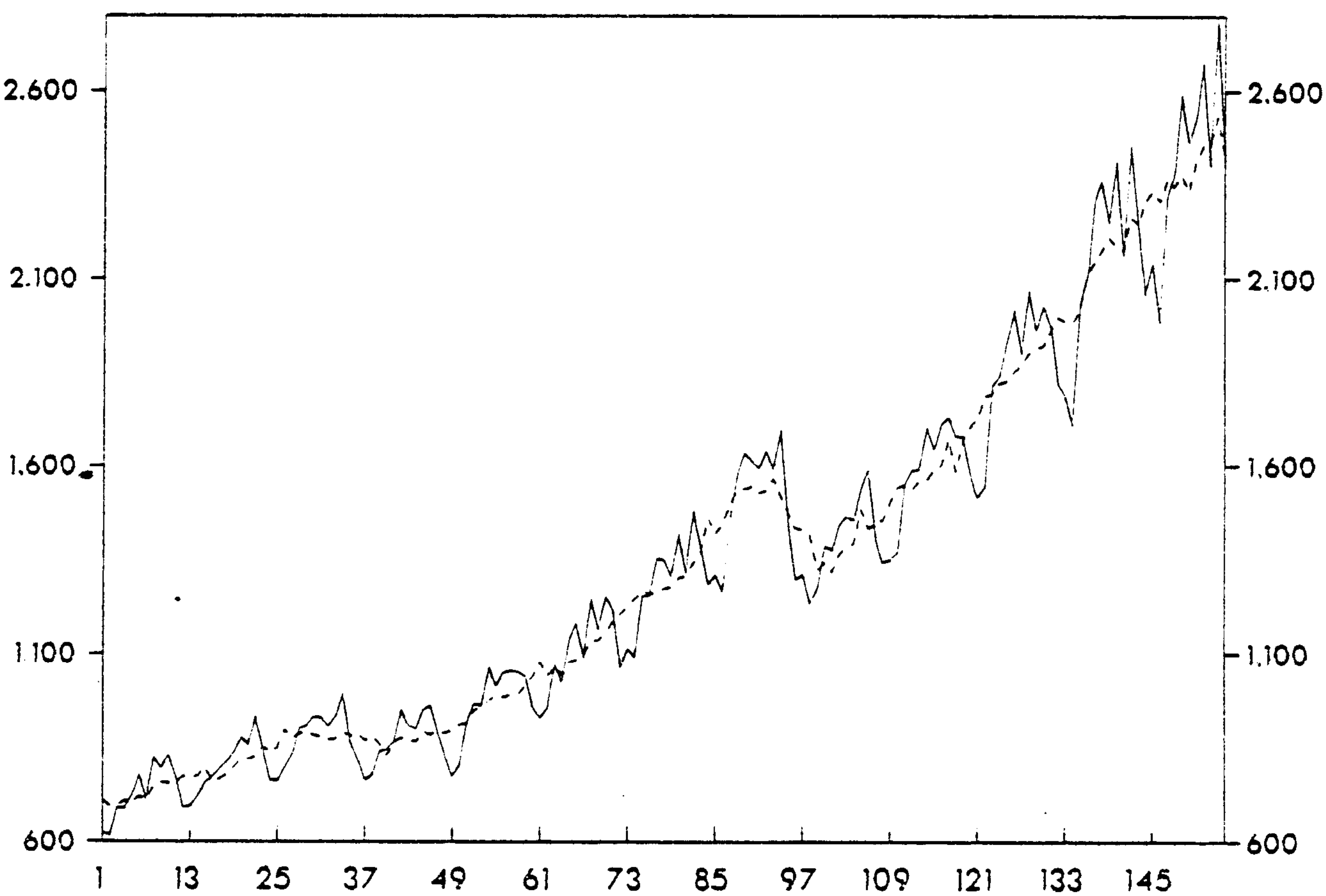


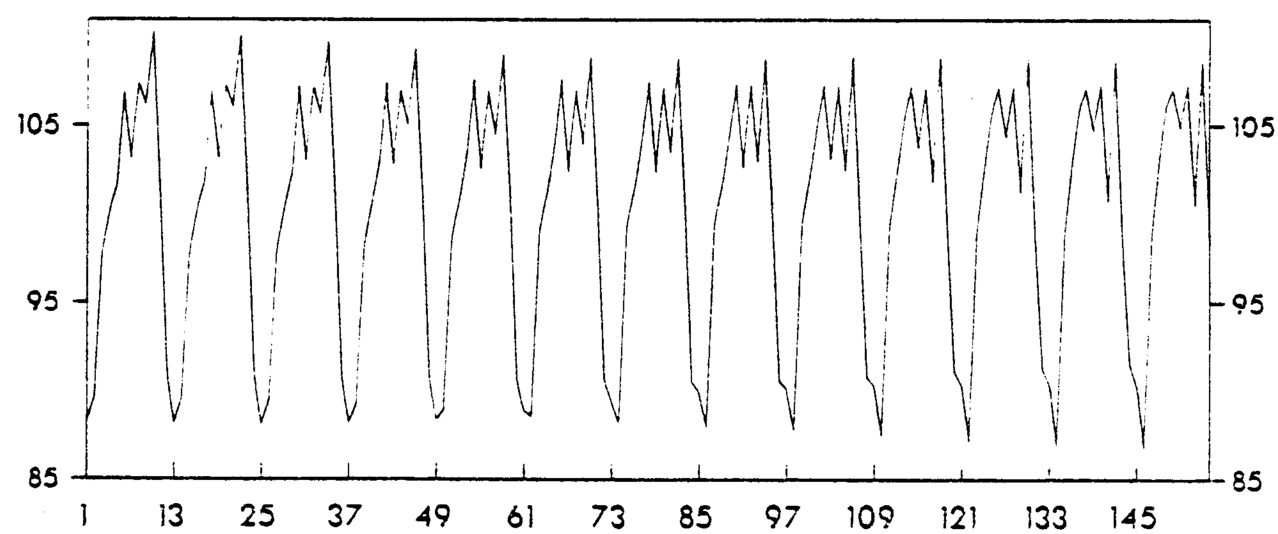
WHOLESALE SALES 507 MODEL BASED



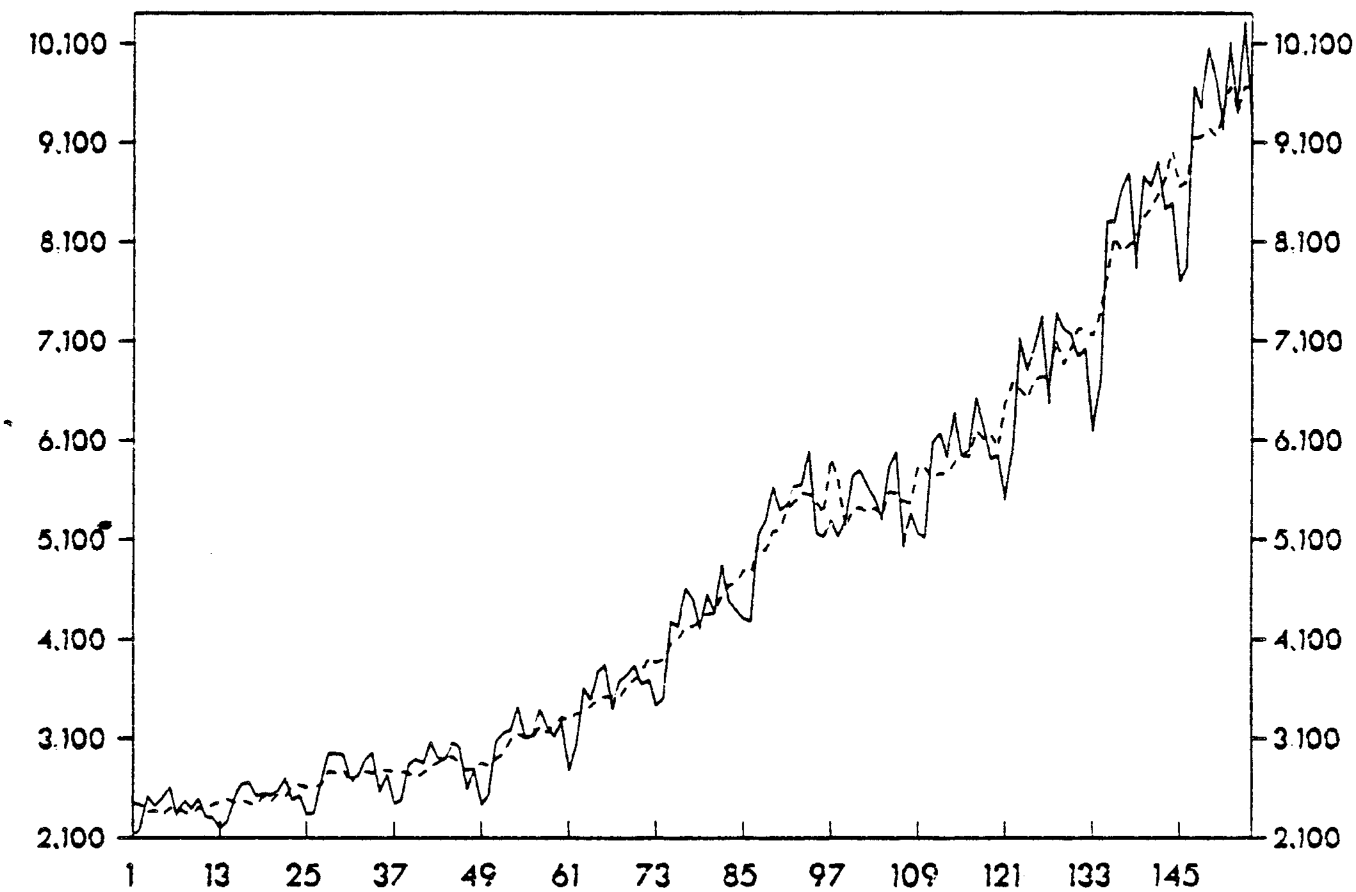


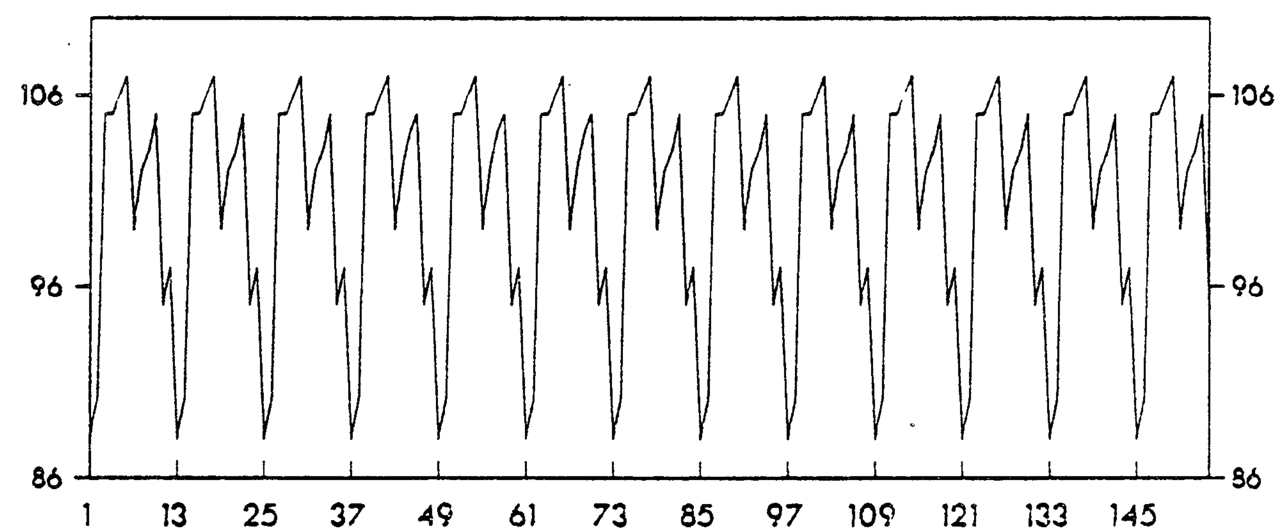
WHOLESALE SALES 507 X-11



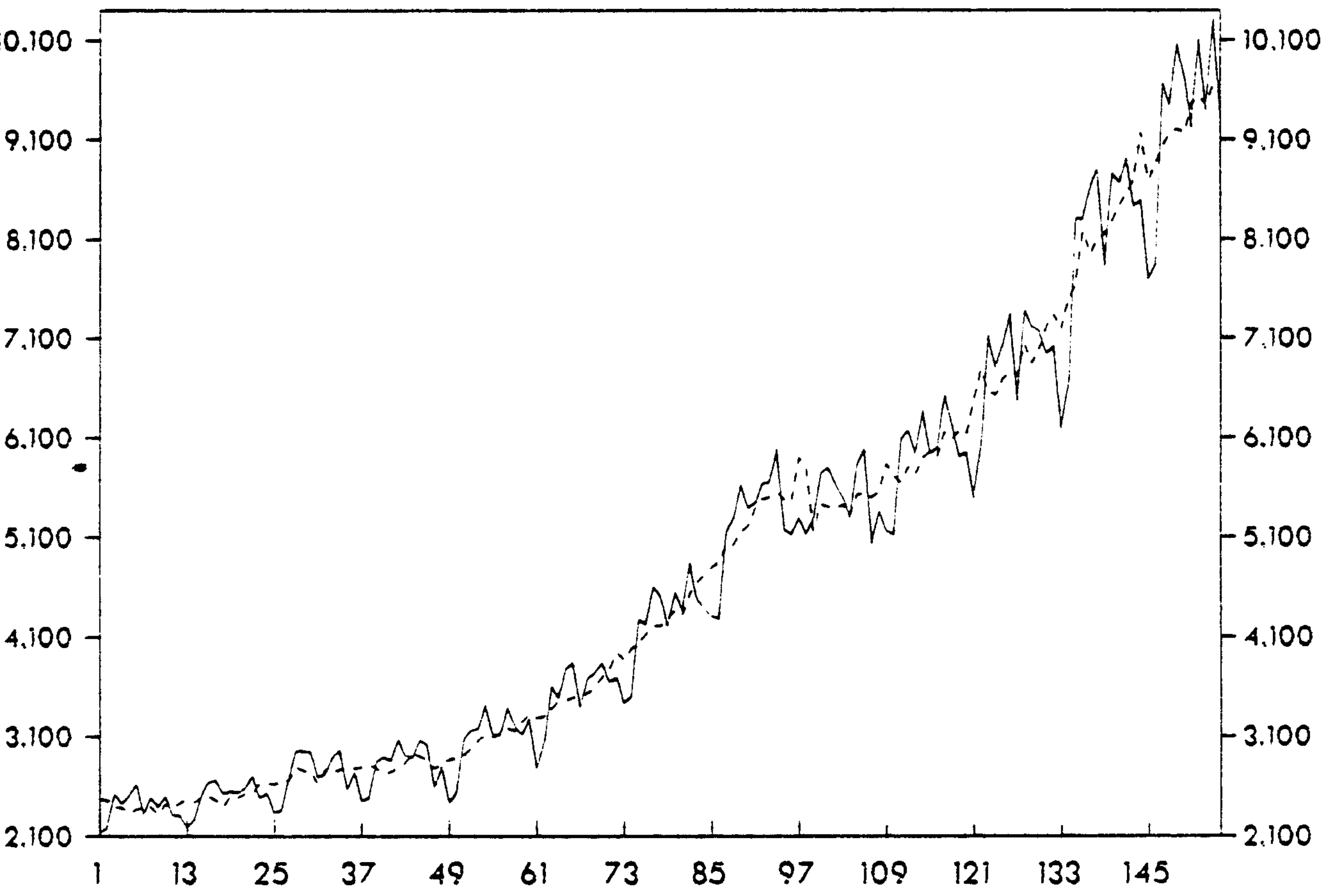


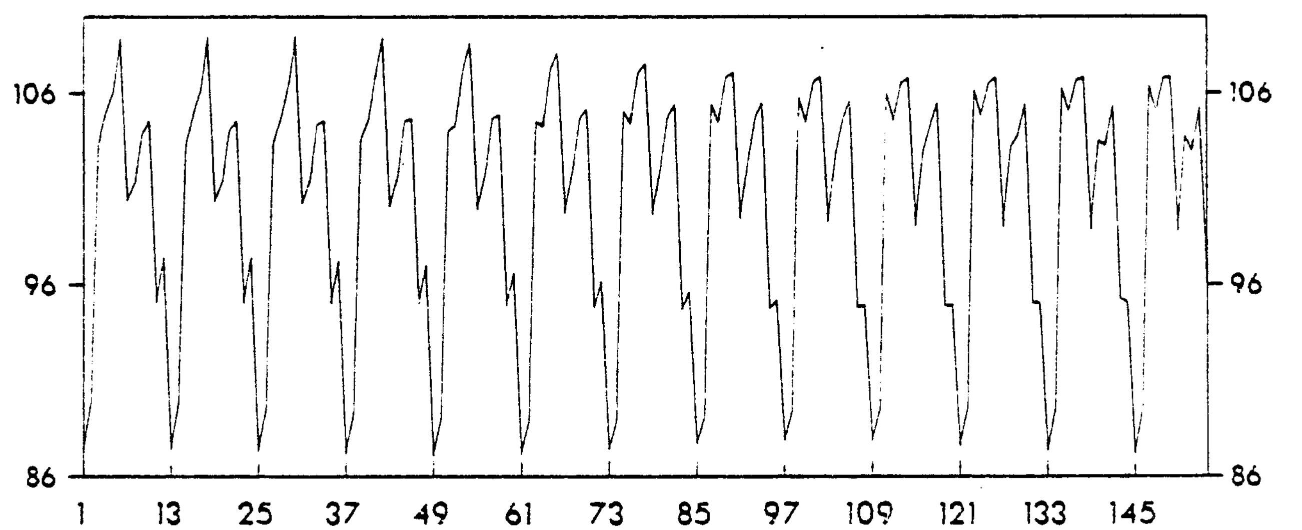
WHOLESALE SALES 508 MODEL BASED



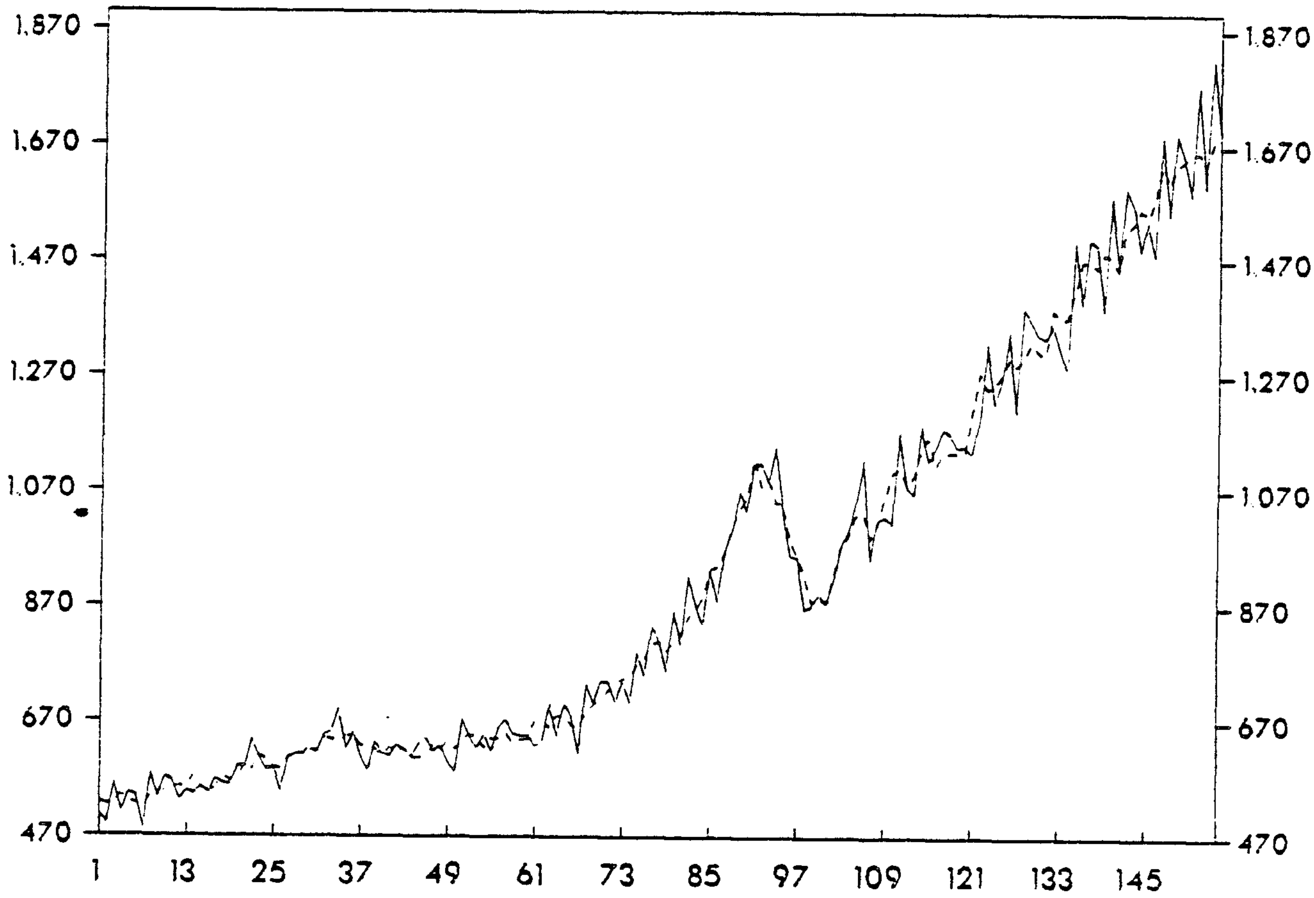


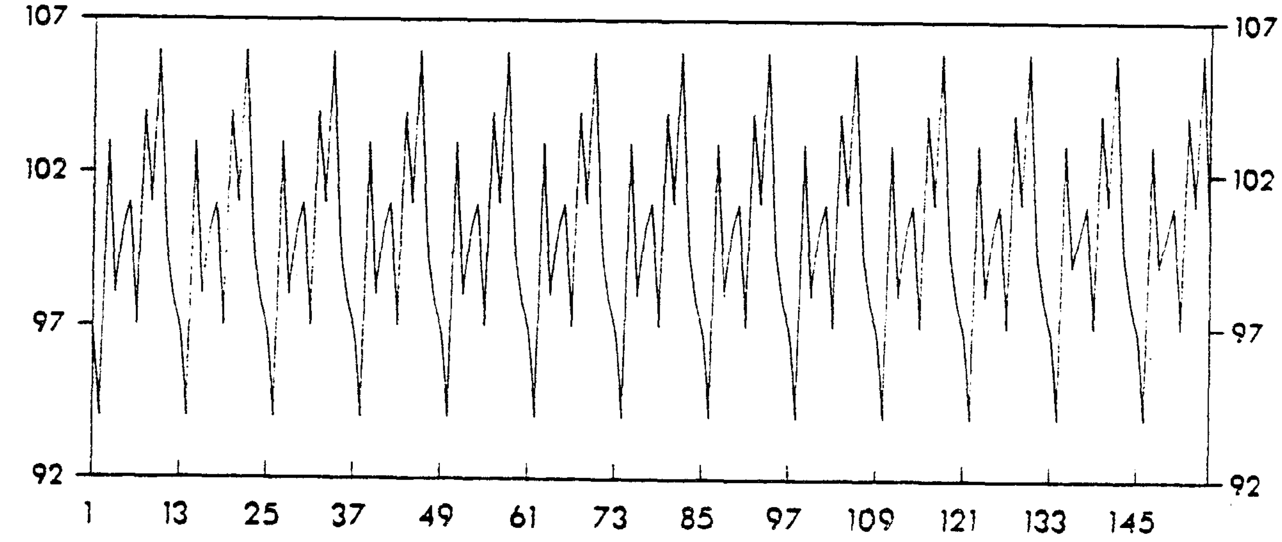
WHOLESALE SALES 508 X-11



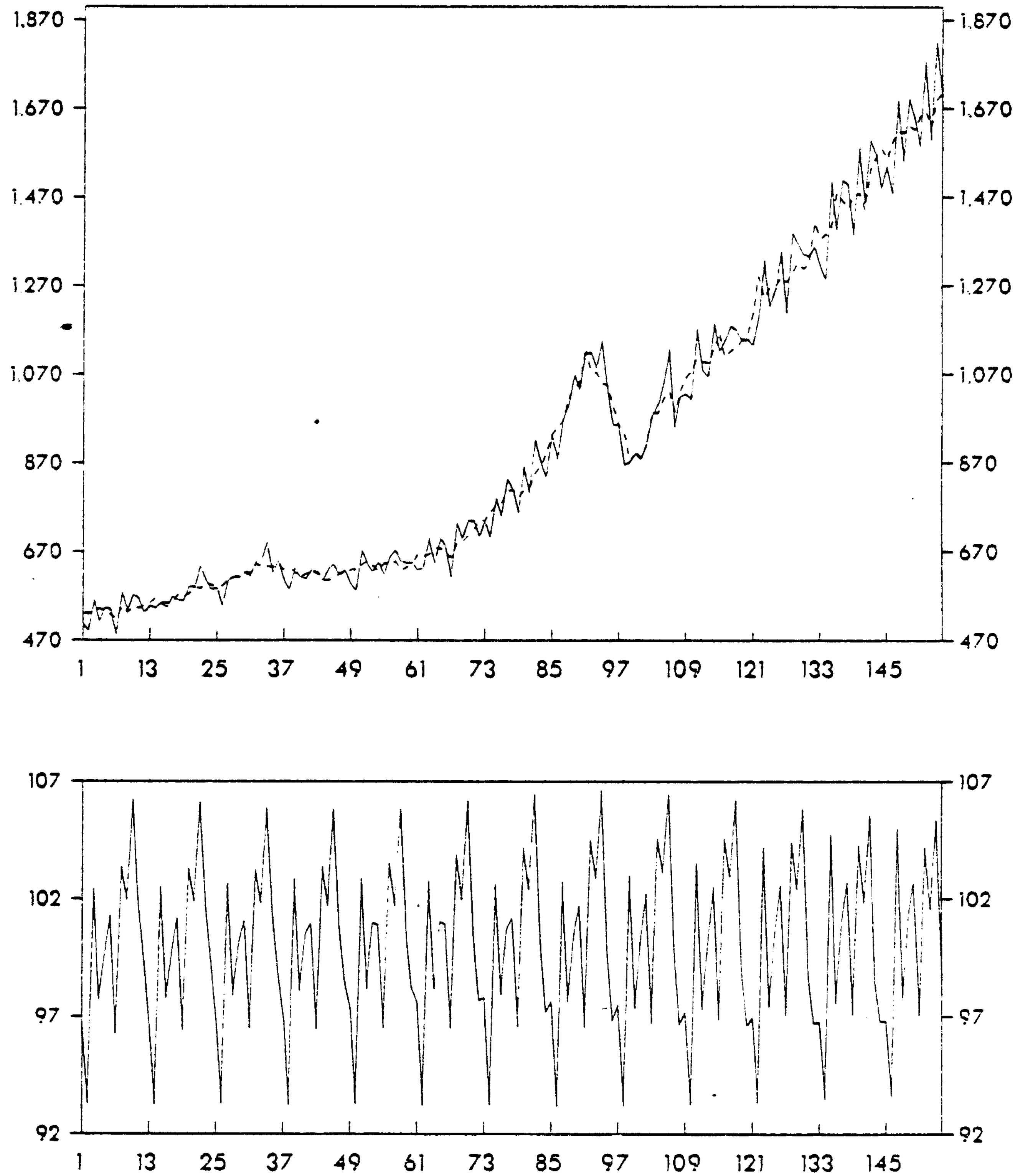


WHOLESALE SALES 511 MODEL BASED

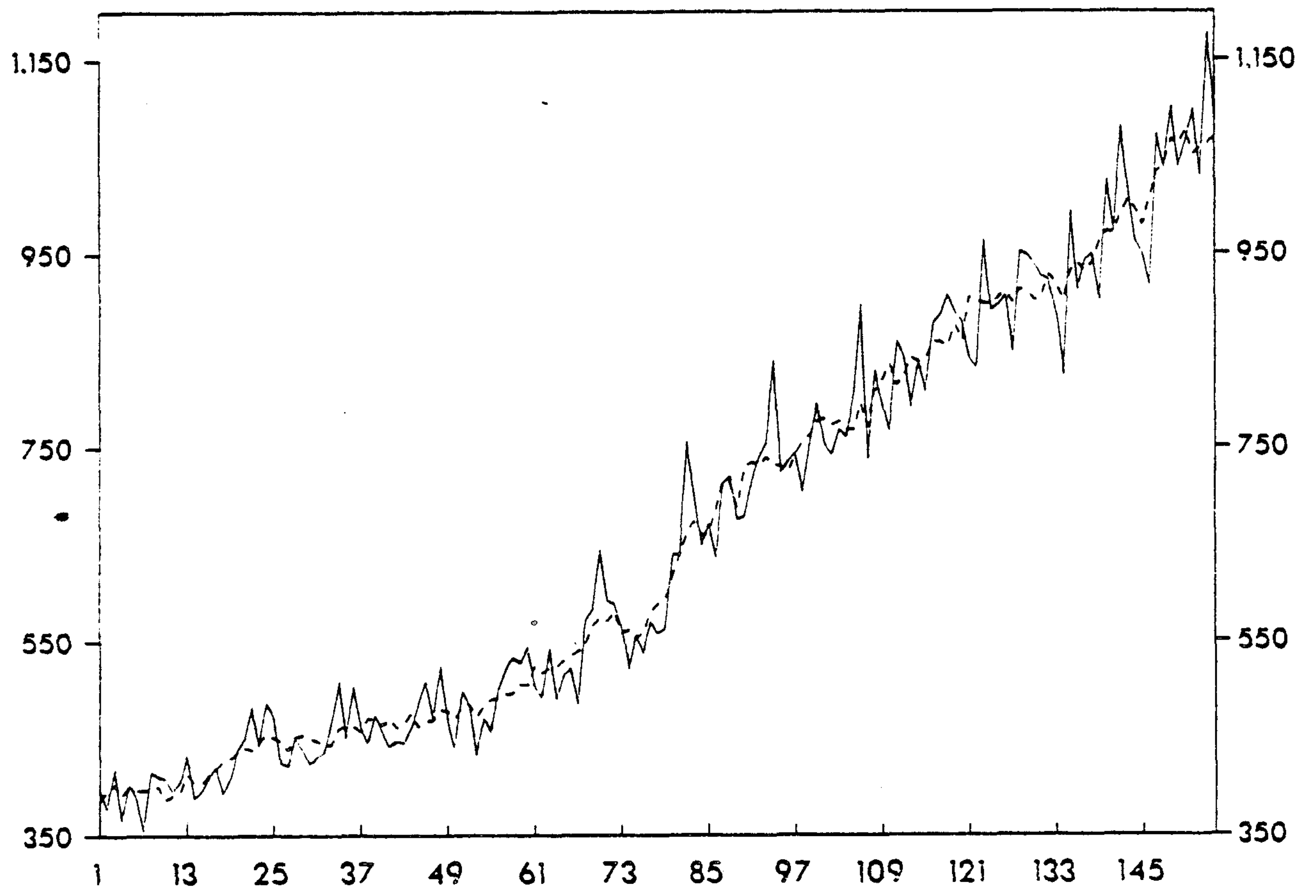


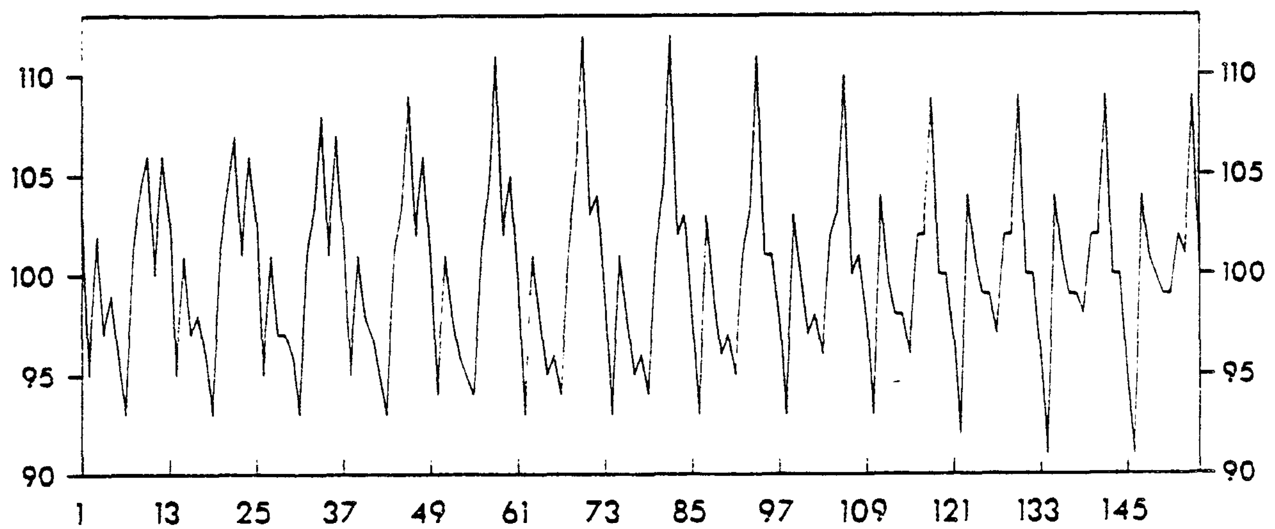


WHOLESALE SALES 511 X-11

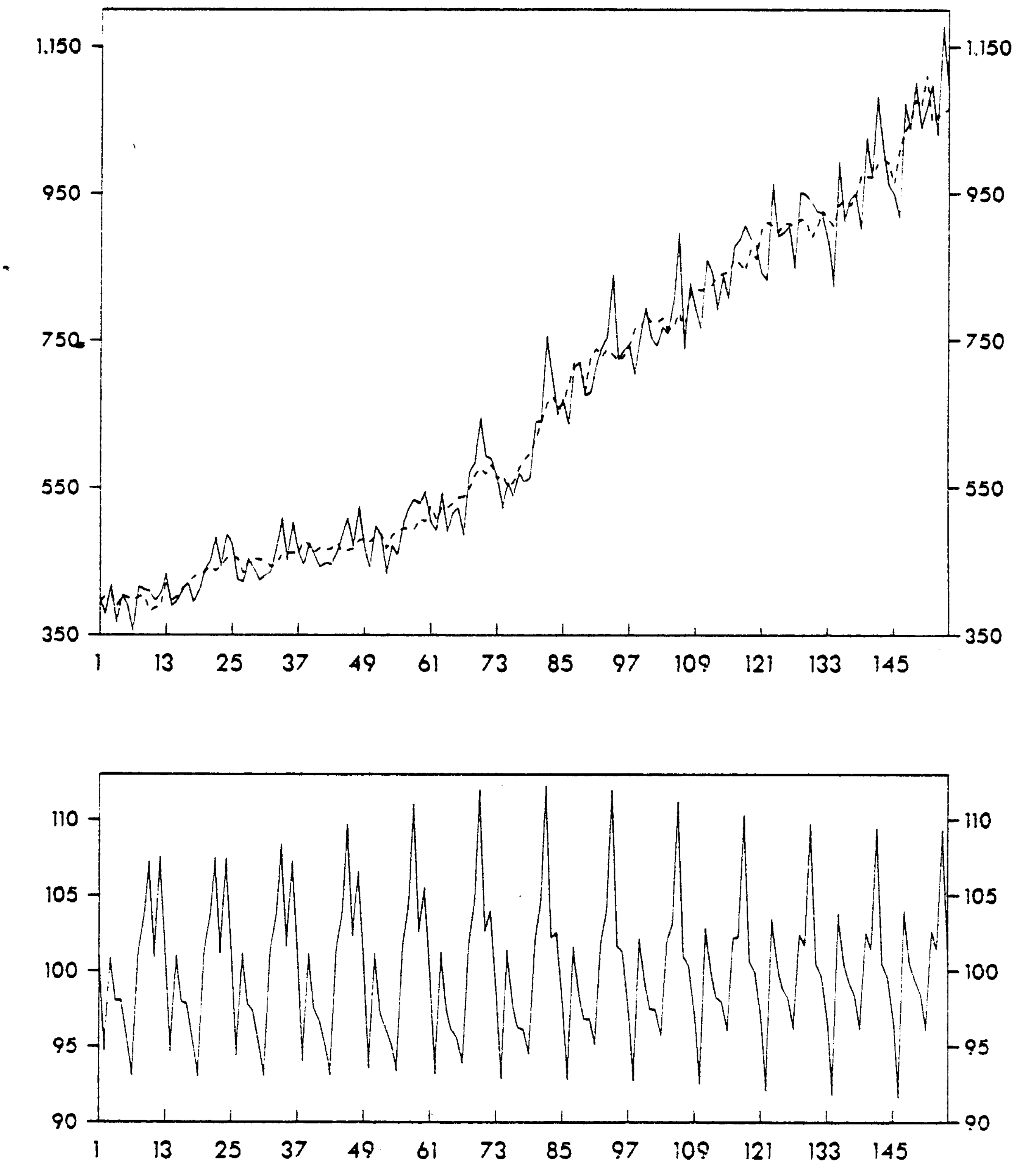


WHOLESALE SALES 512 MODEL BASED

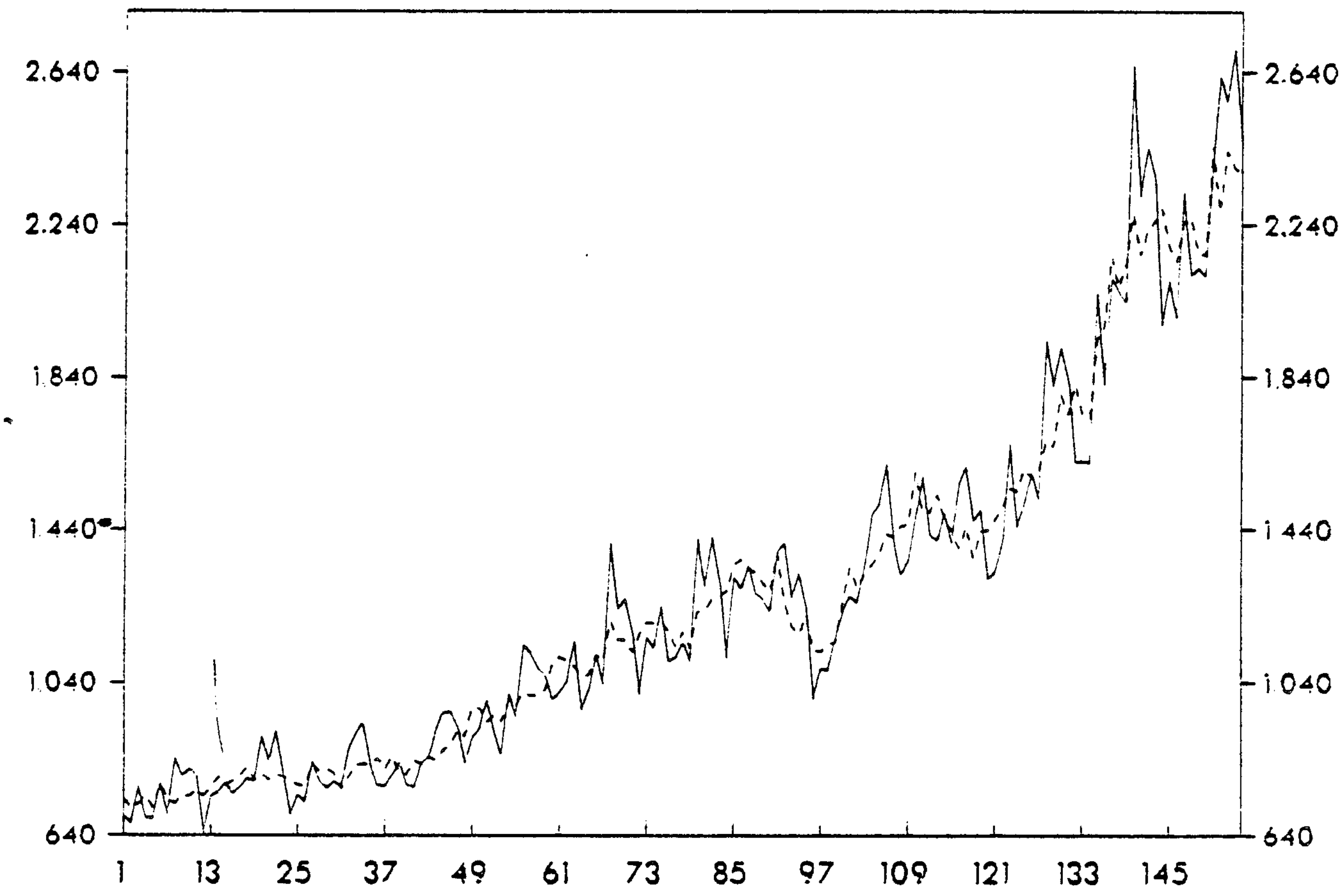


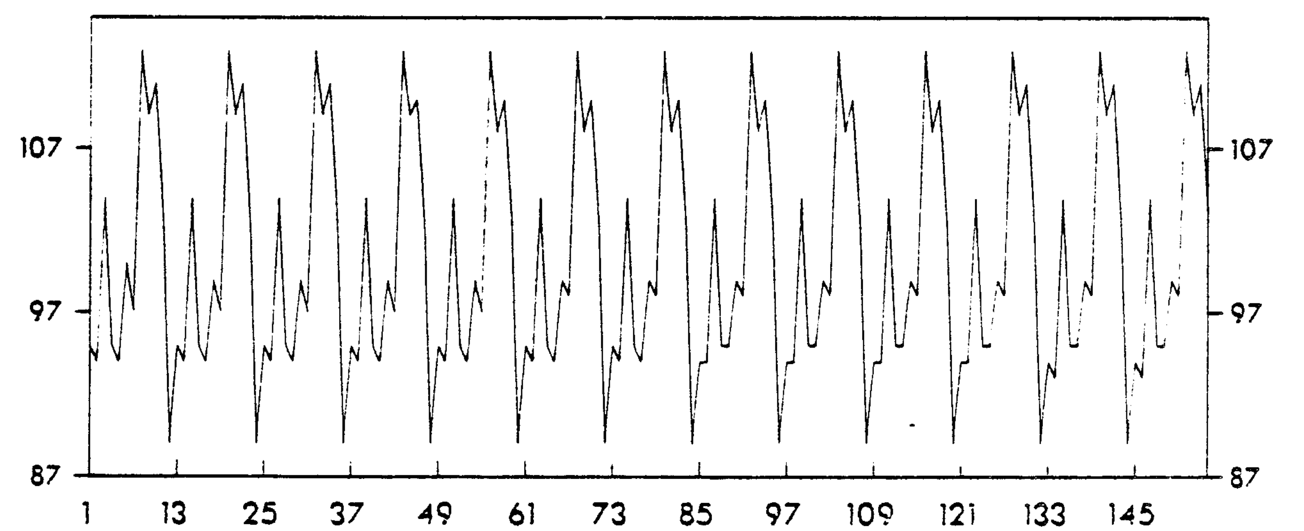


WHOLESALE SALES 512 X-11

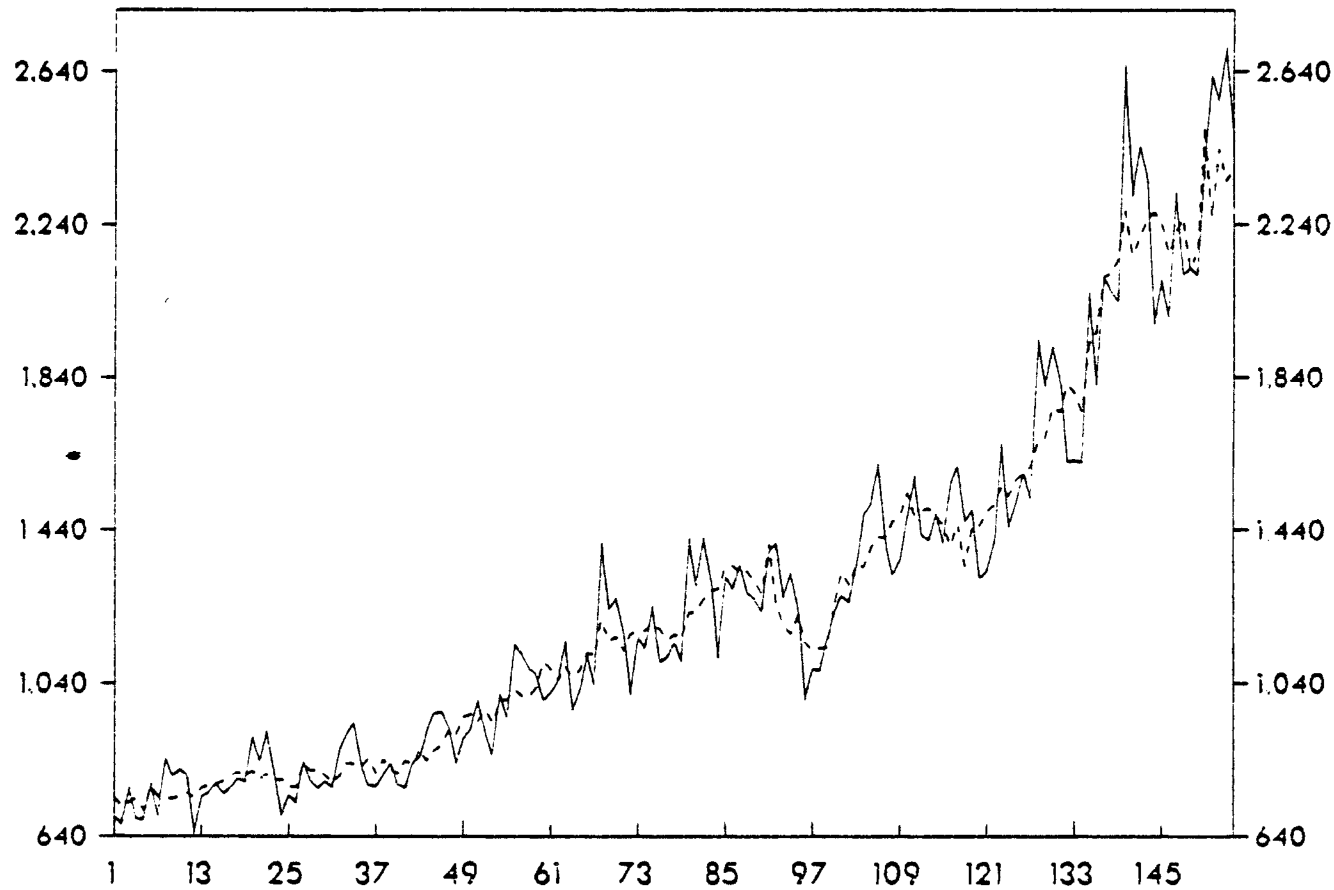


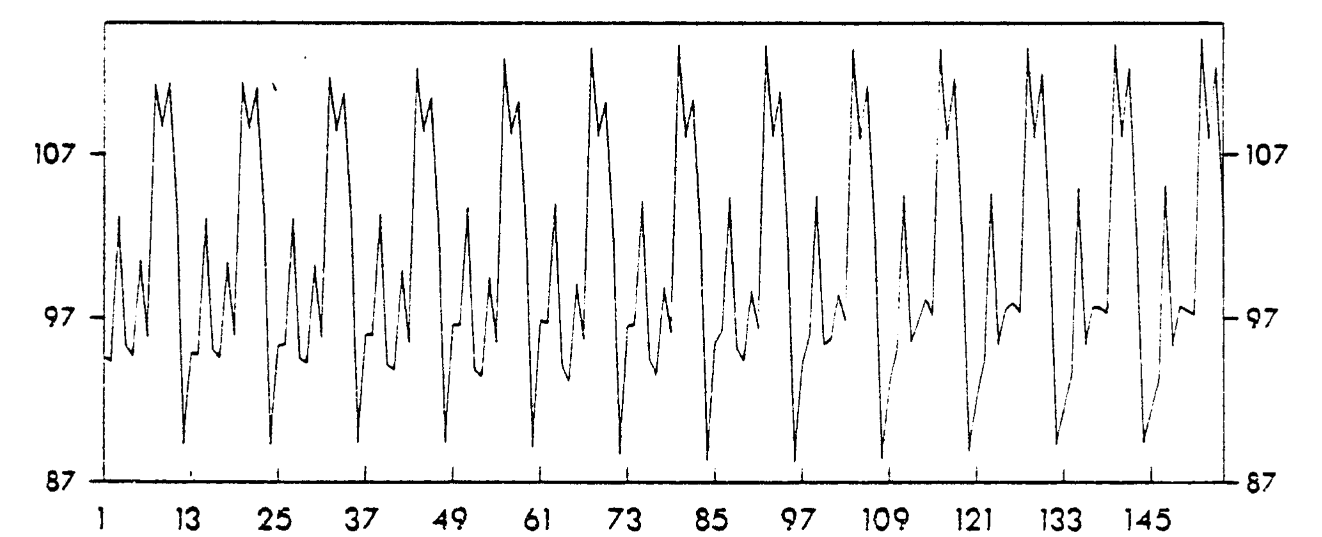
WHOLESALE SALES 513 MODEL BASED



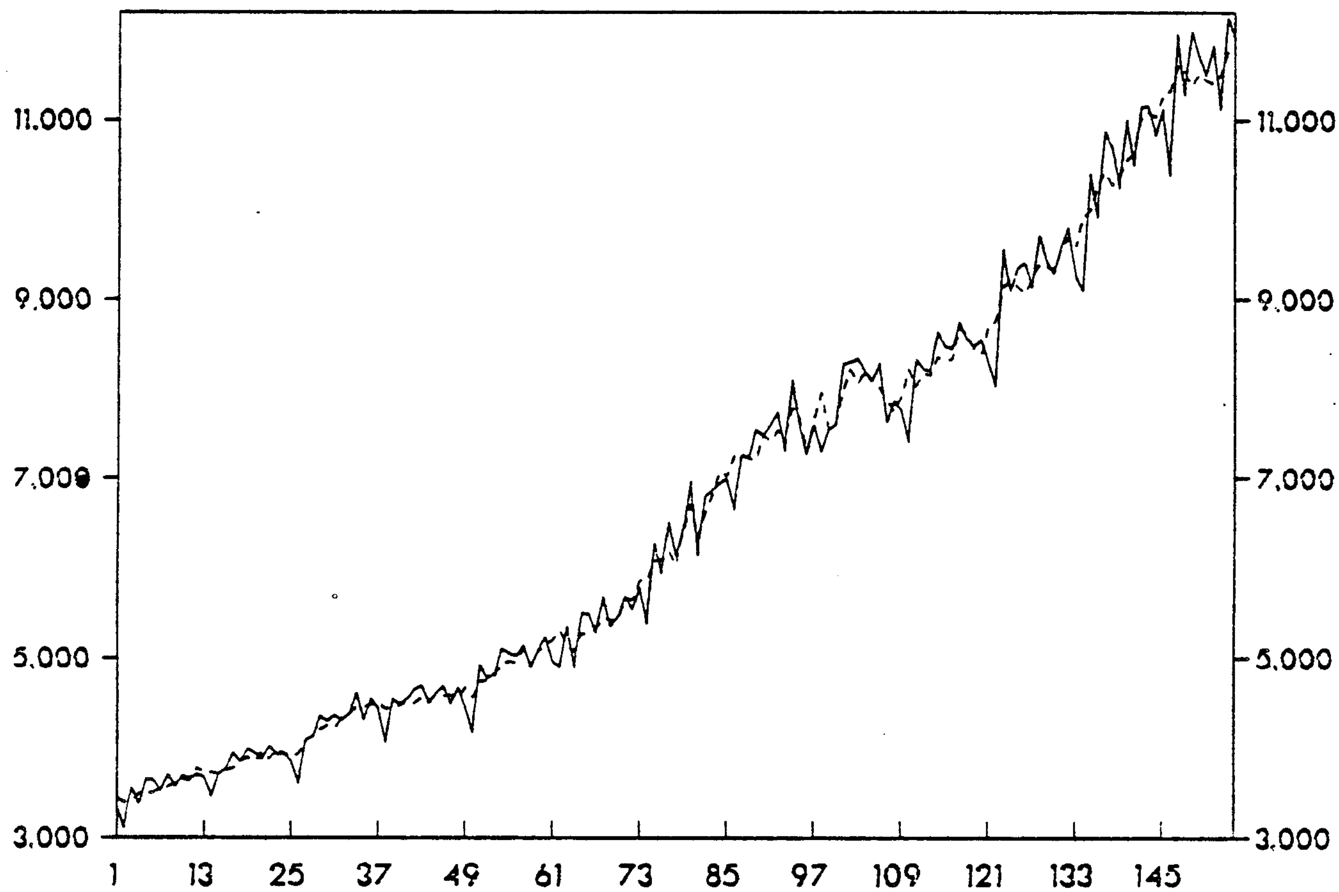


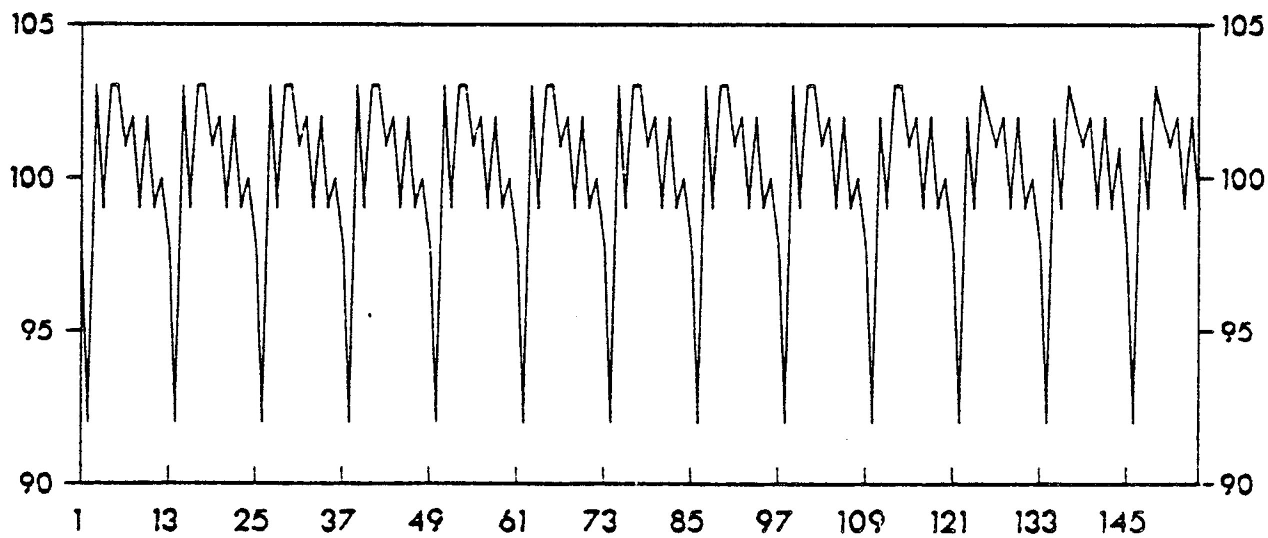
WHOLESALE SALES 513 X-11



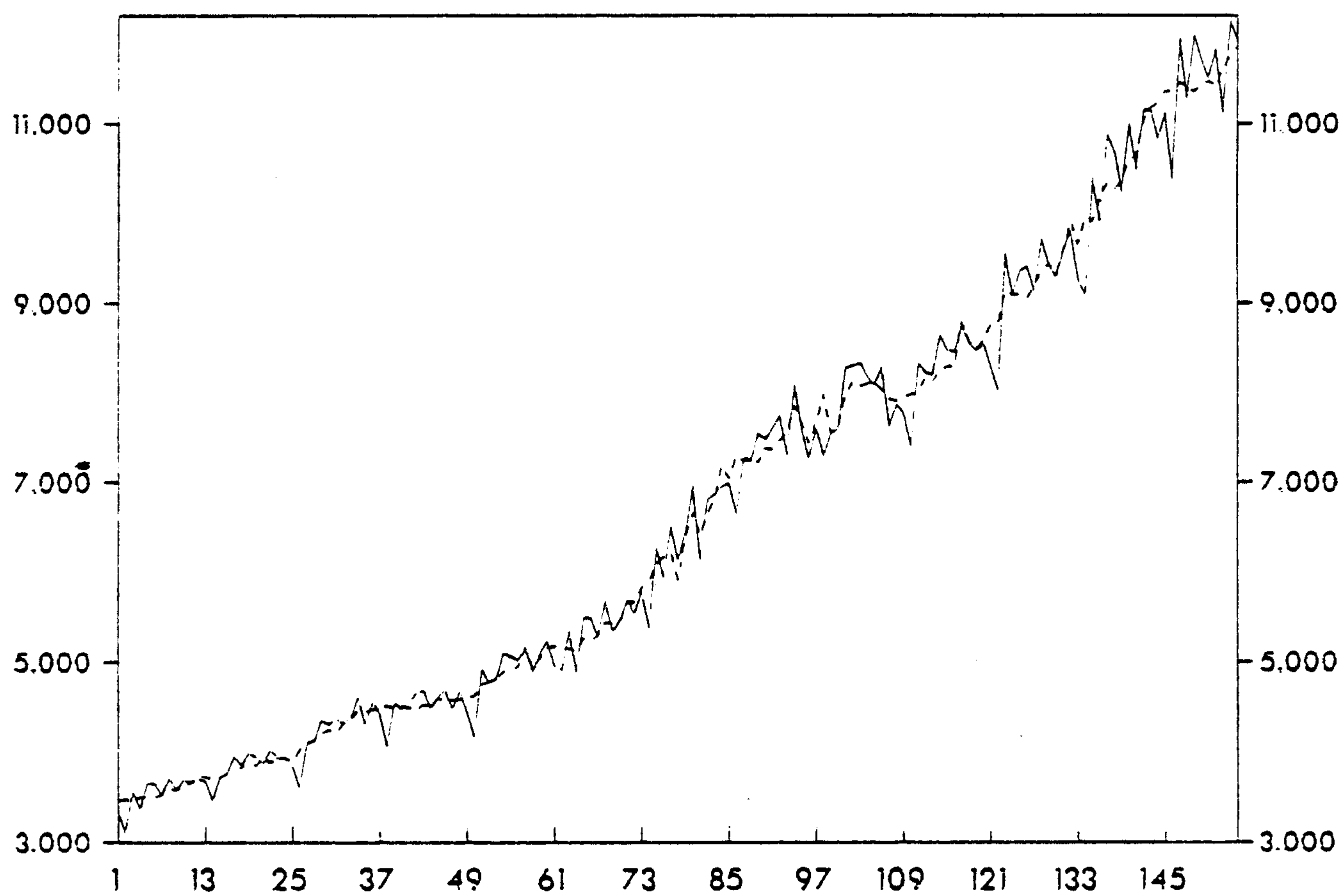


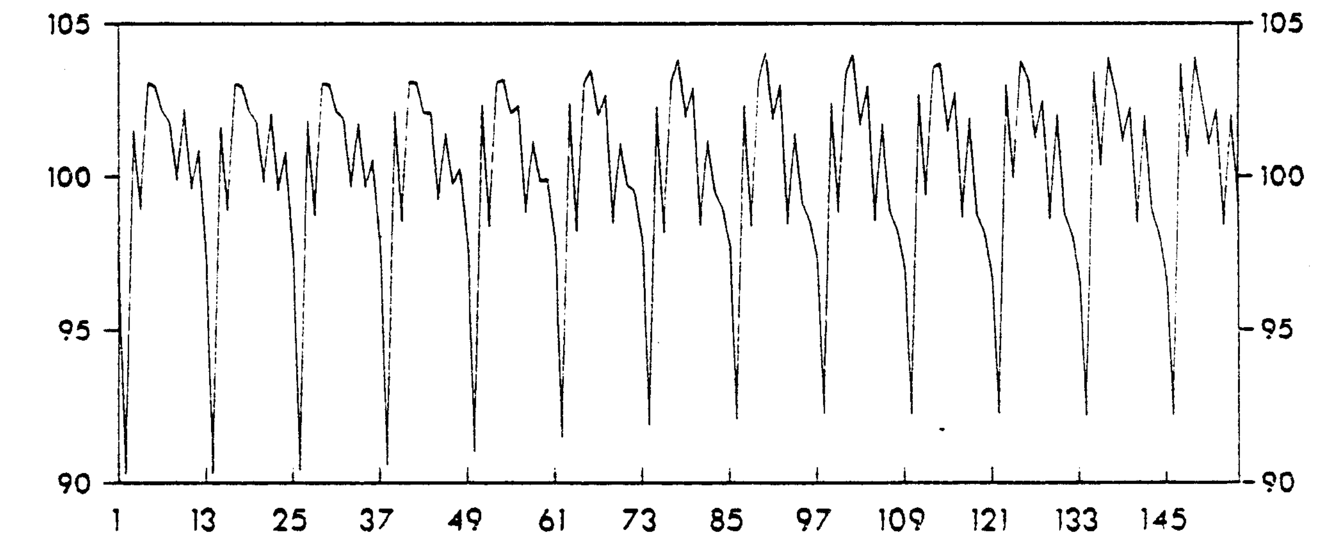
WHOLESALE SALES 514 MODEL BASED



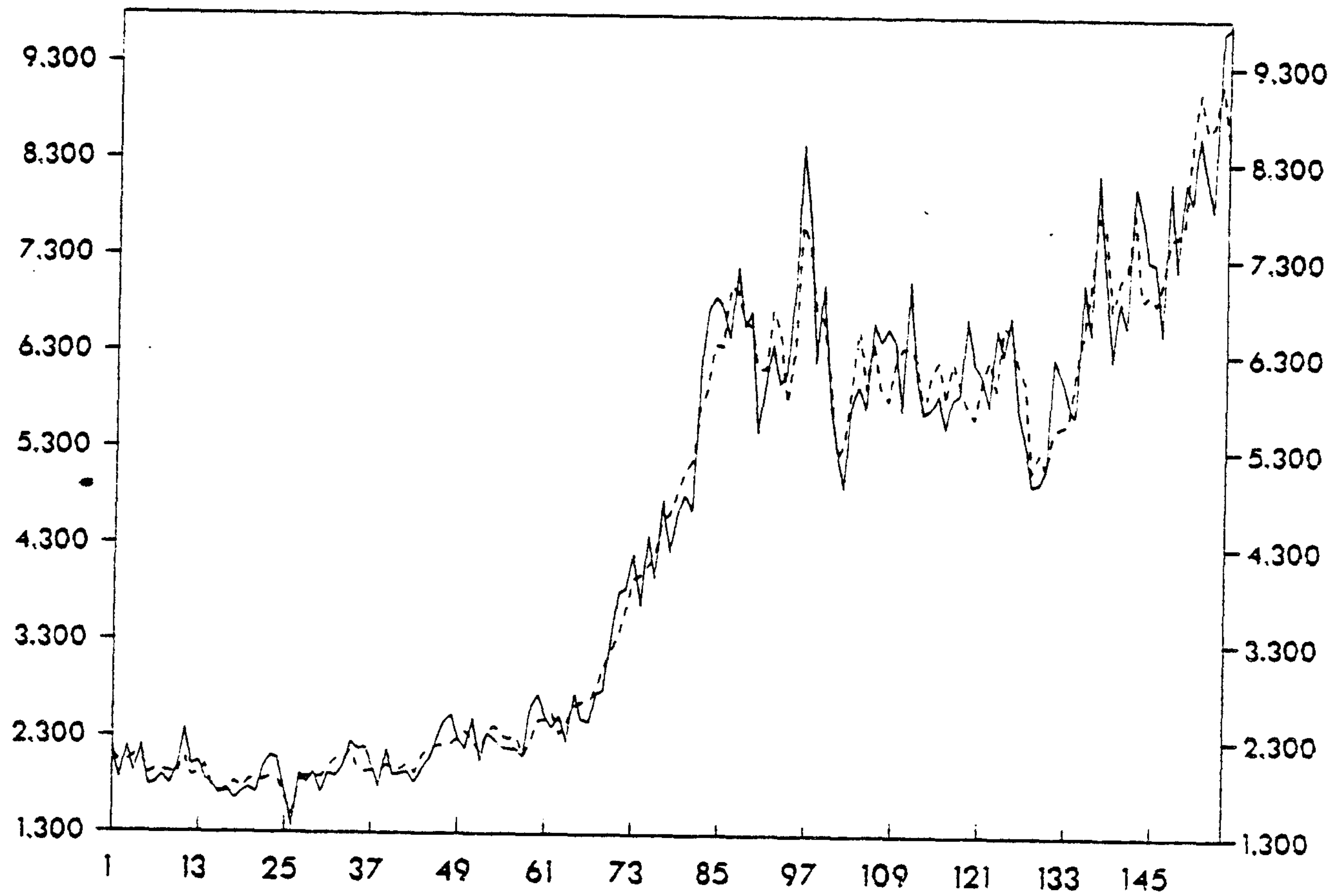


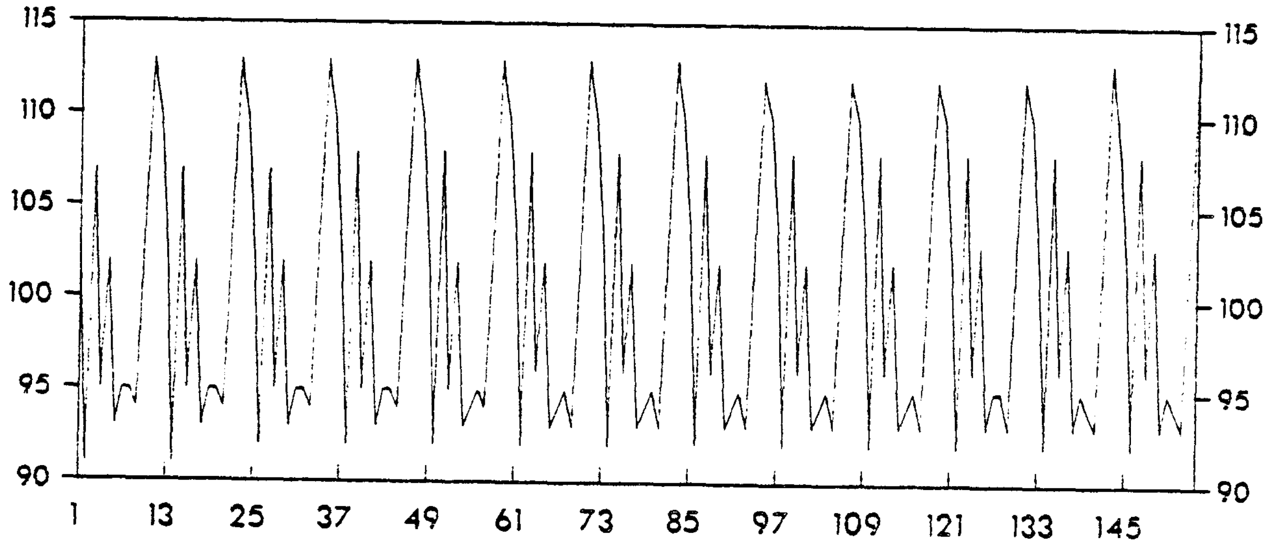
WHOLESALE SALES 514 X-11



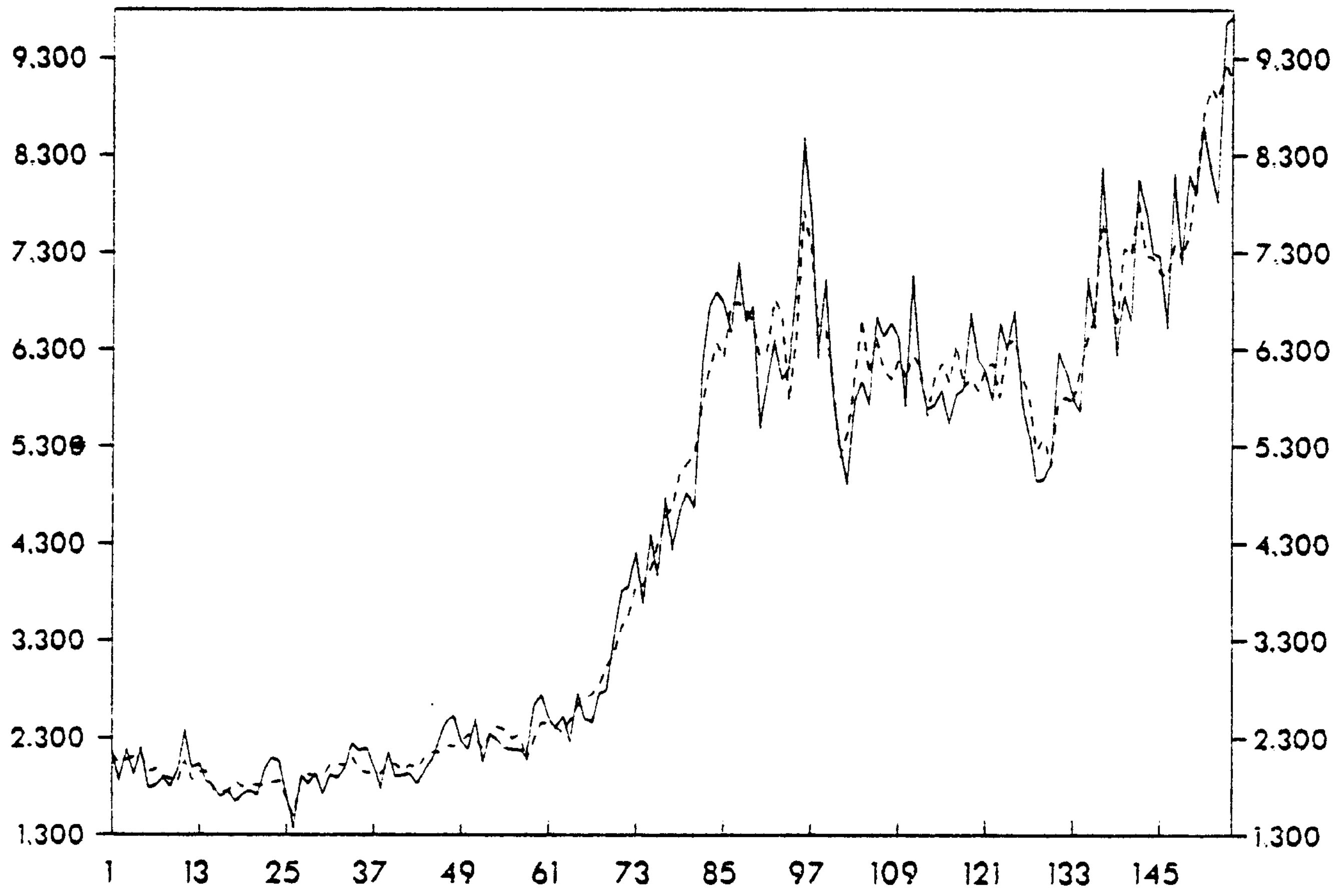


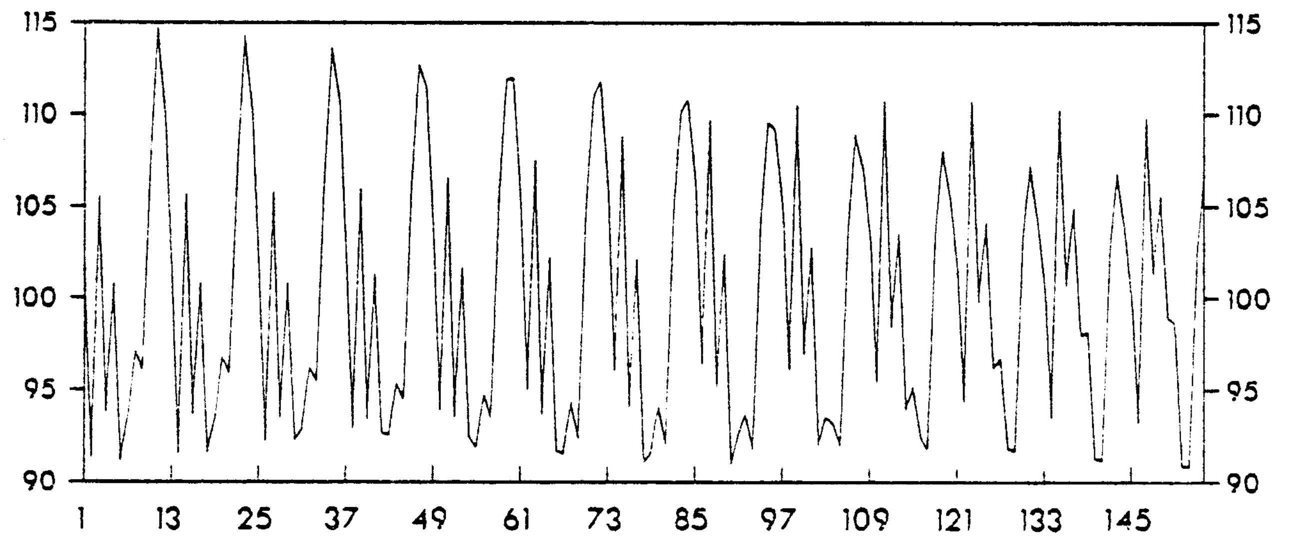
WHOLESALE SALES 515 MODEL BASED



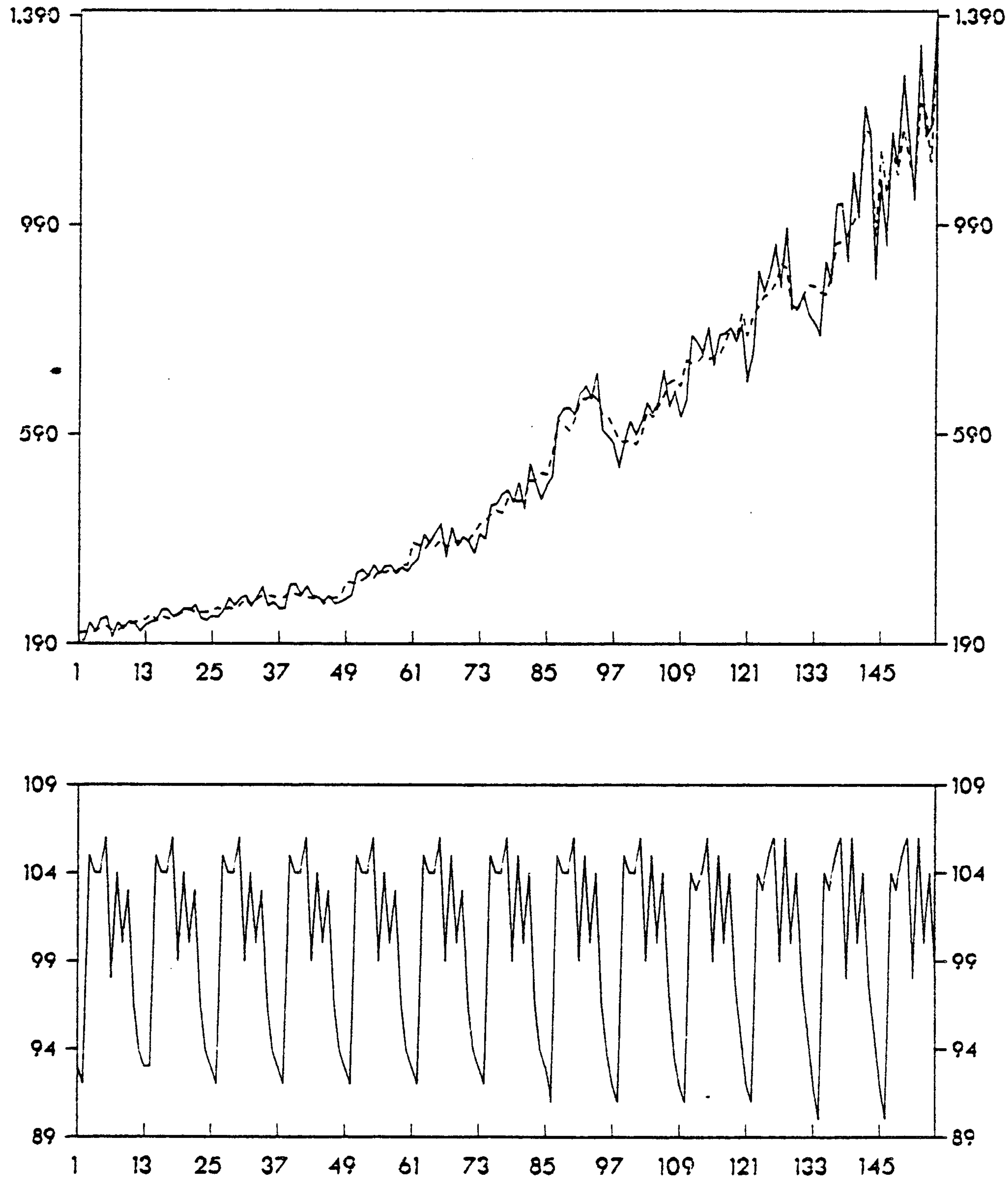


WHOLESALE SALES 515 X-11

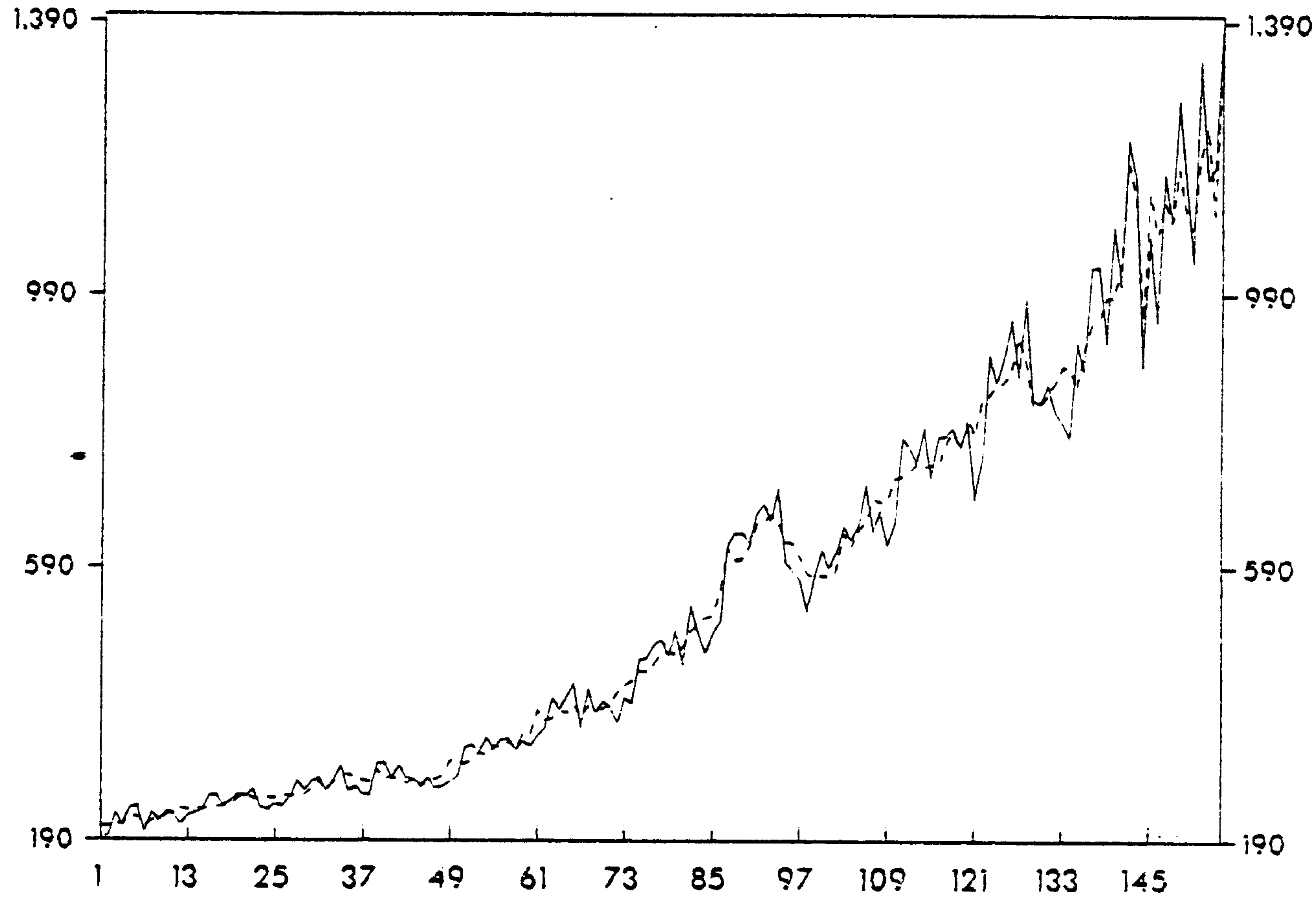


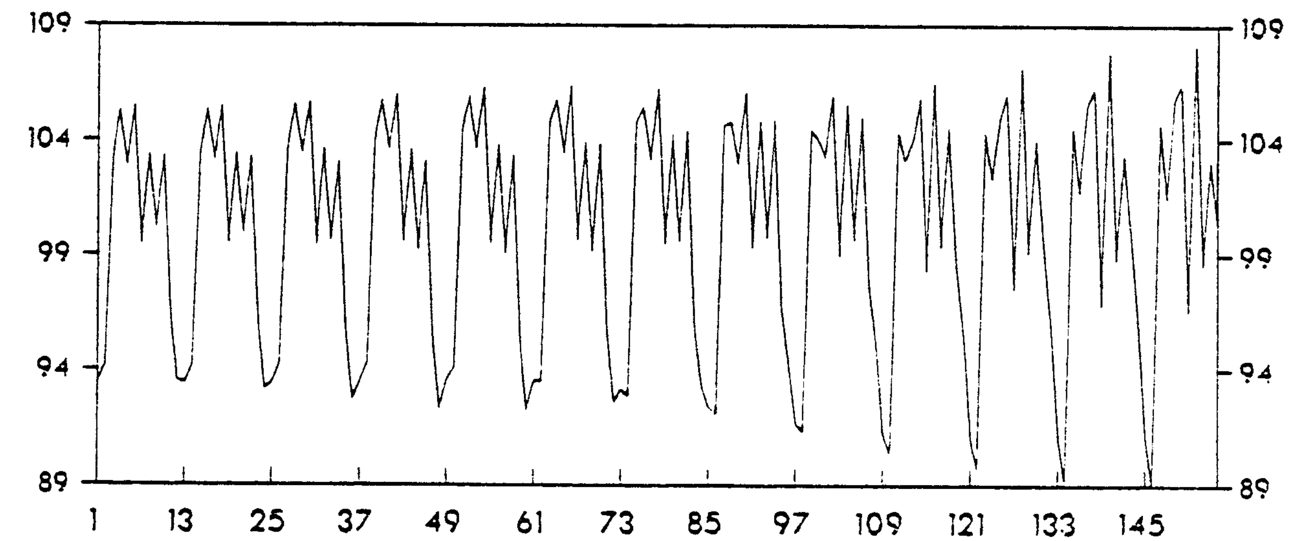


WHOLESALE SALES 516 MODEL BASED

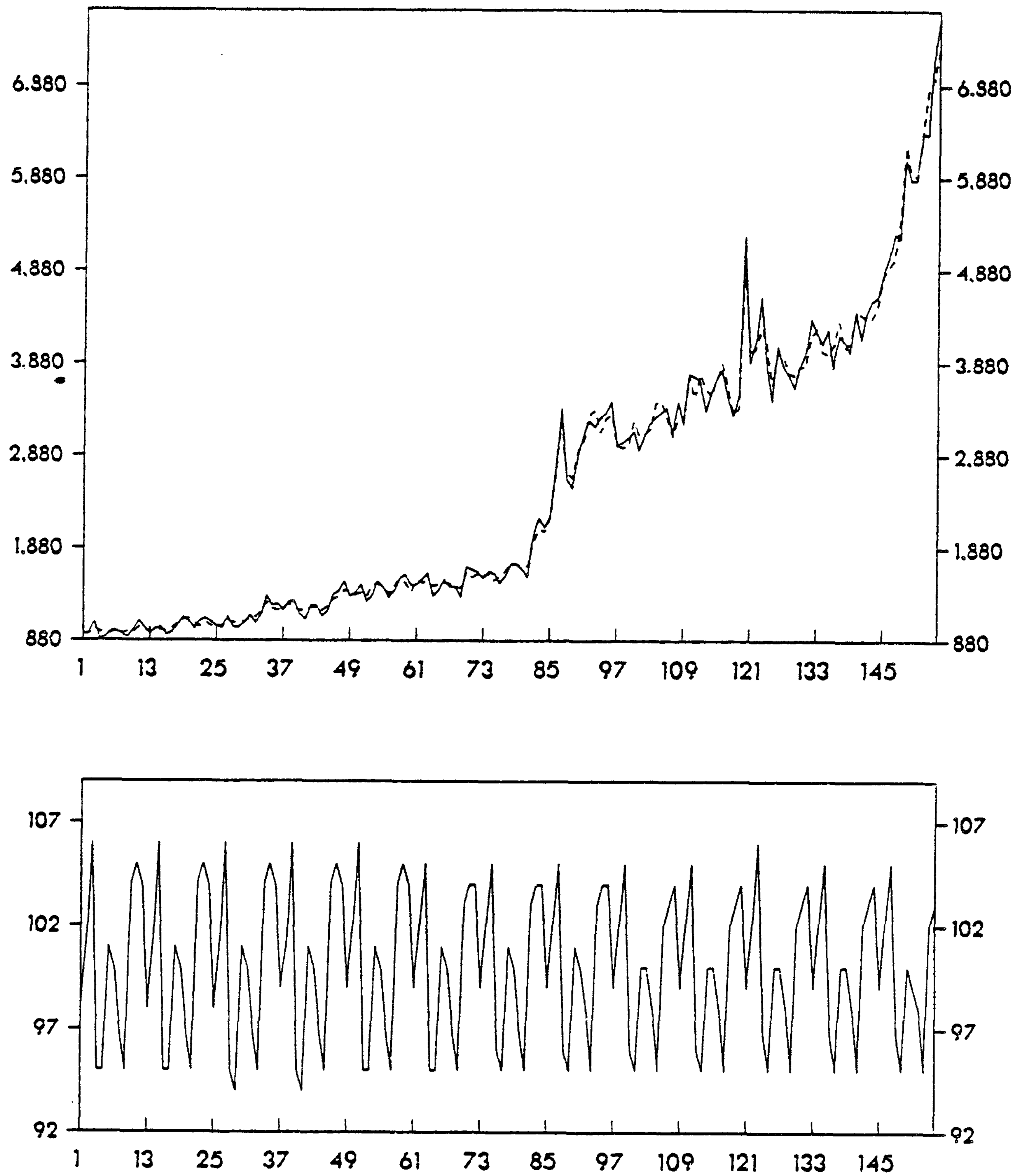


WHOLESALE SALES 516 X-11

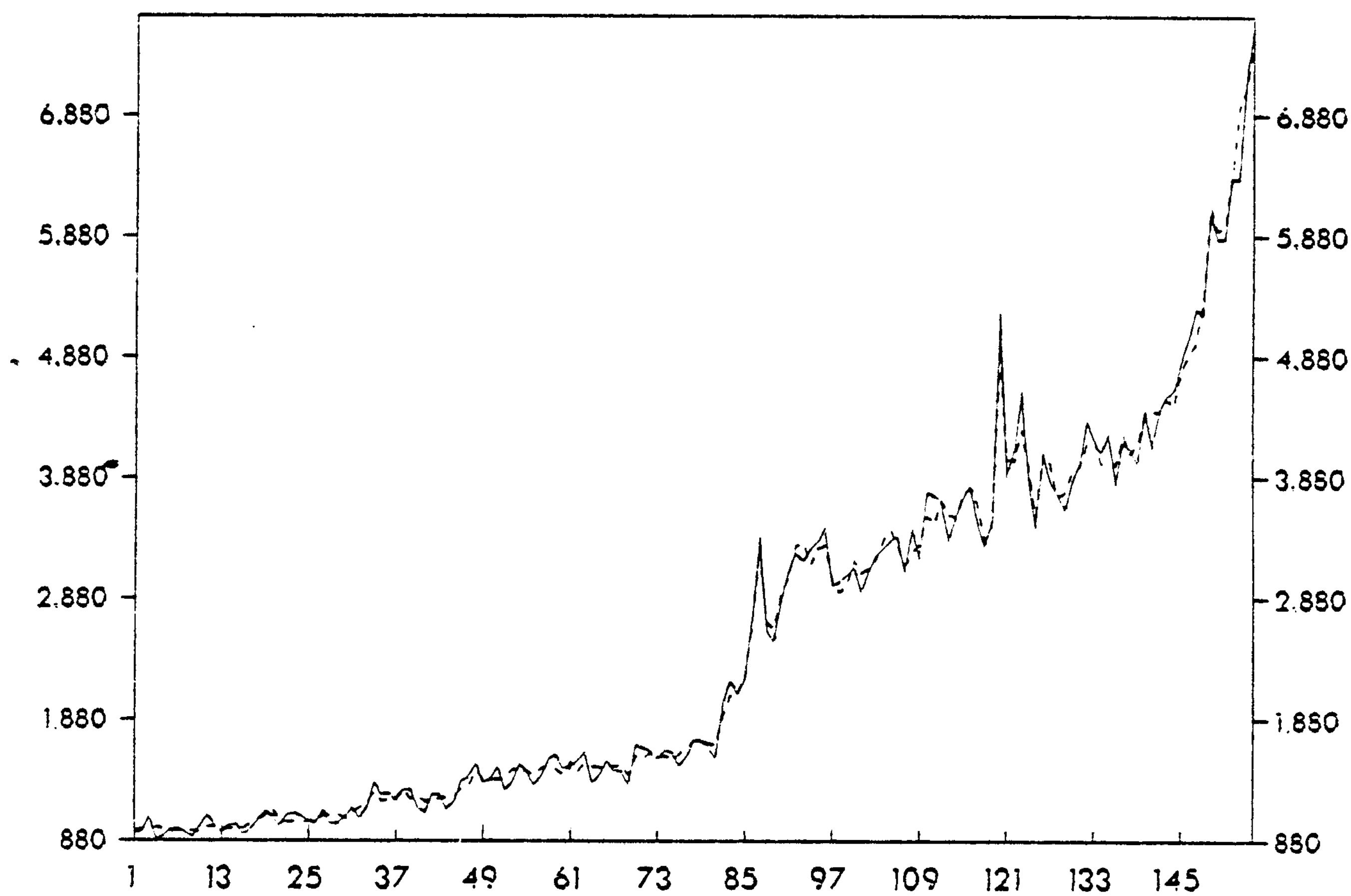


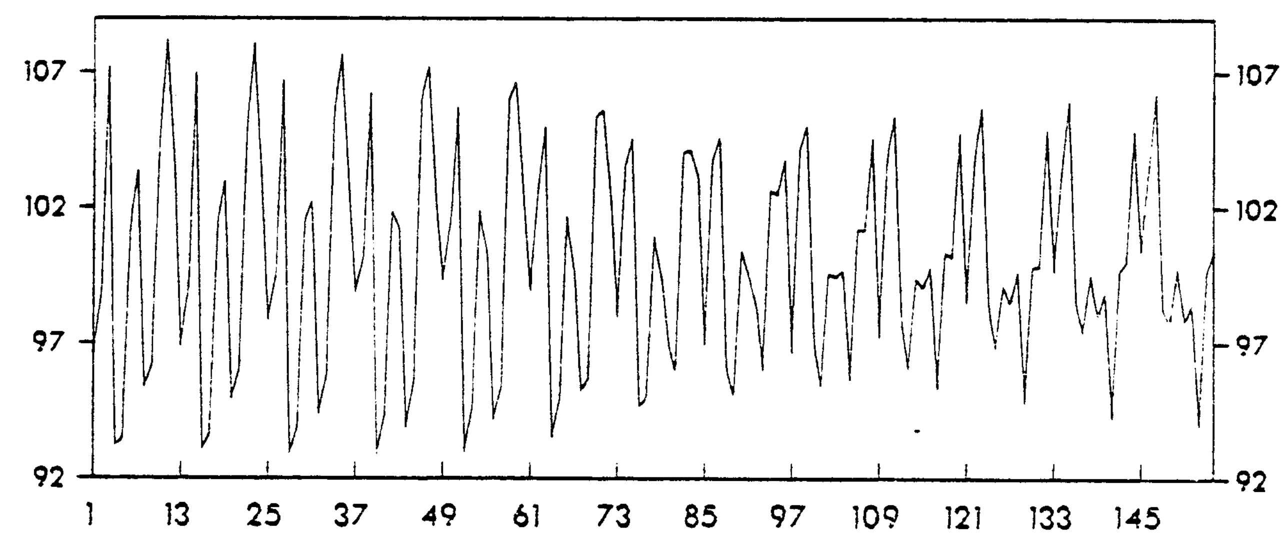


WHOLESALE SALES 517 MODEL BASED

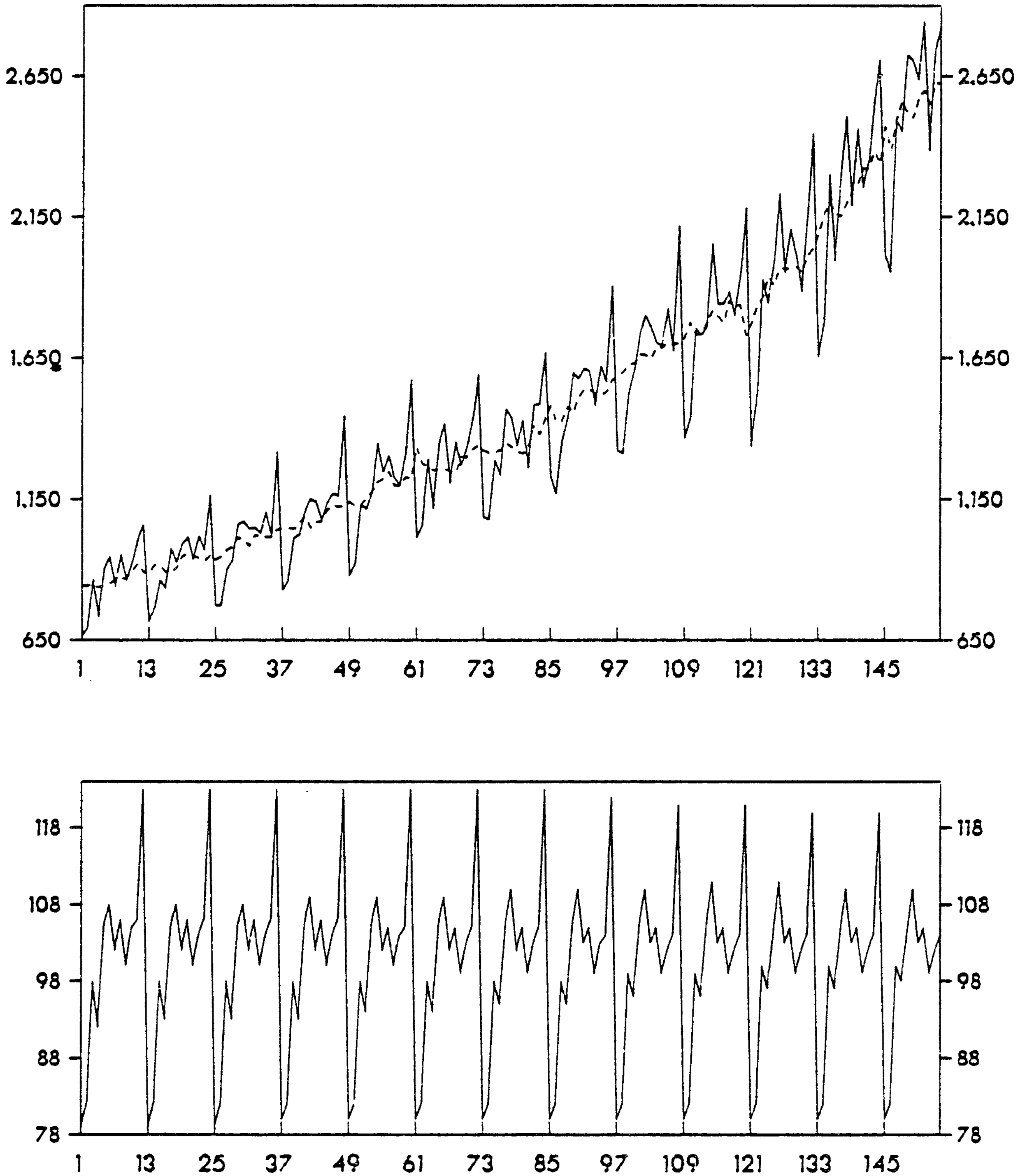


WHOLESALE SALES 517 X-11

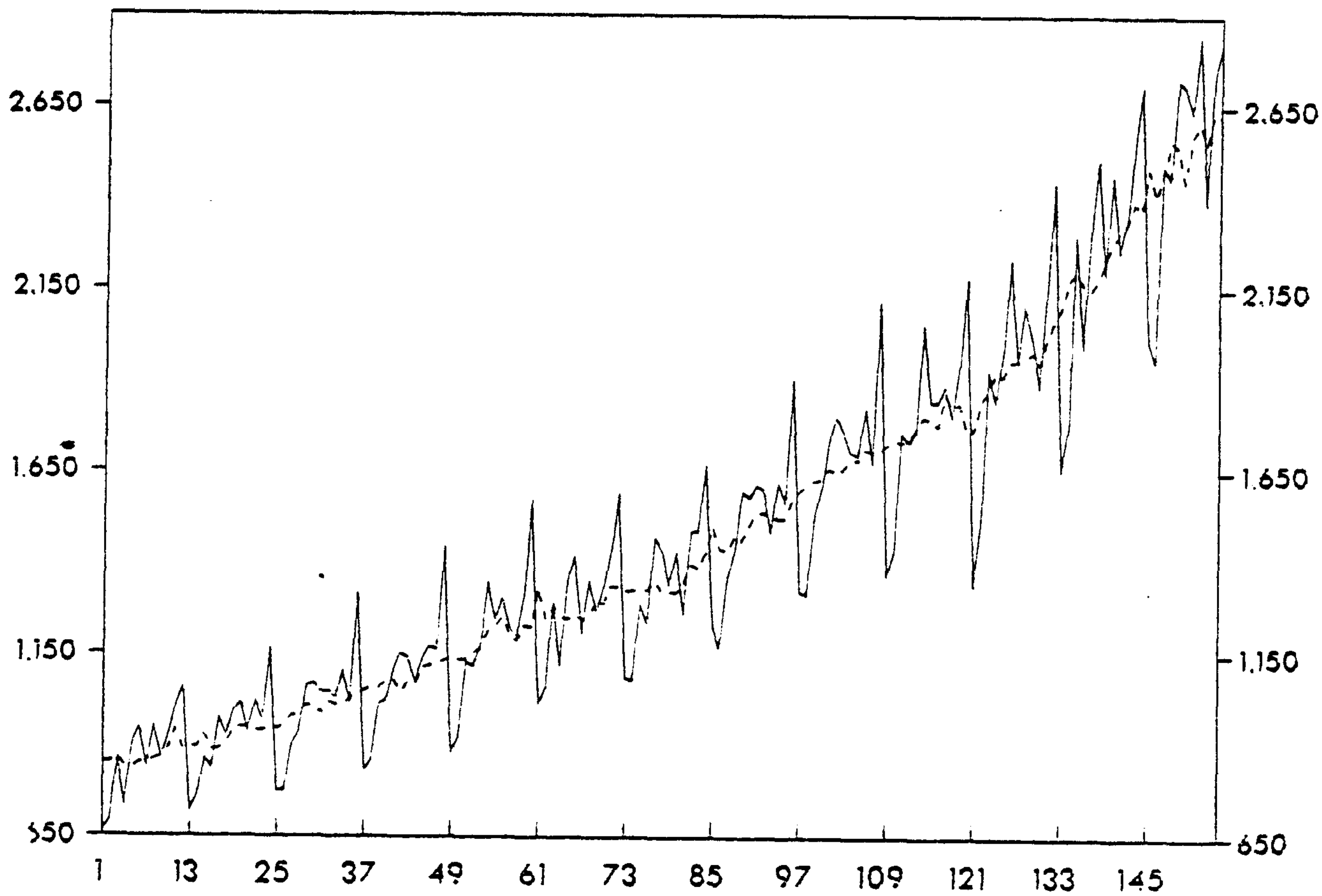


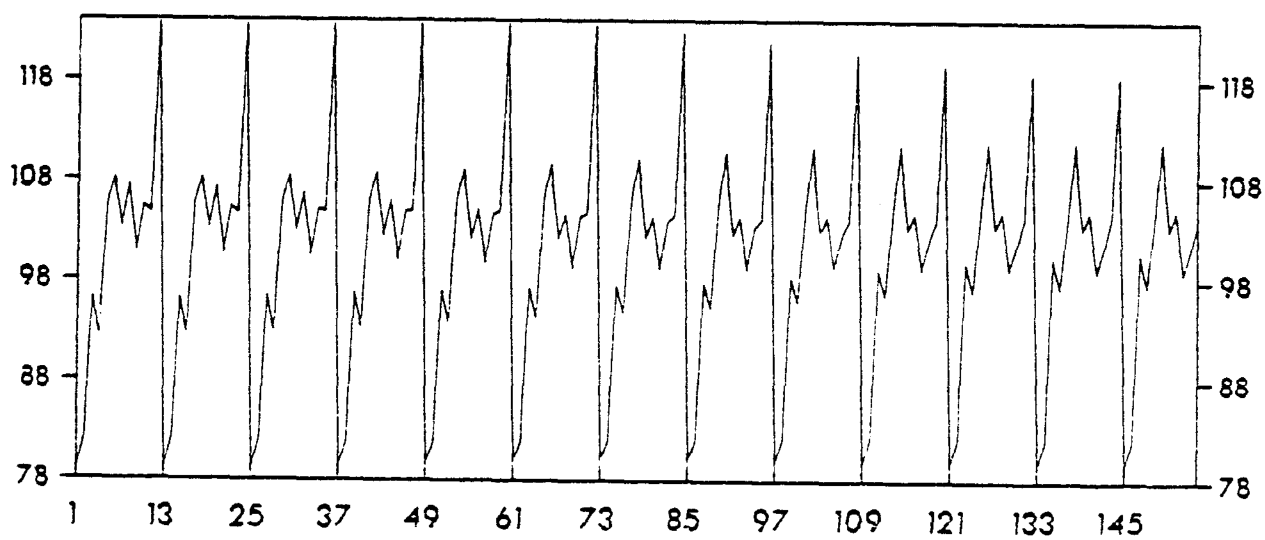


WHOLESALE SALES 518 MODEL BASED

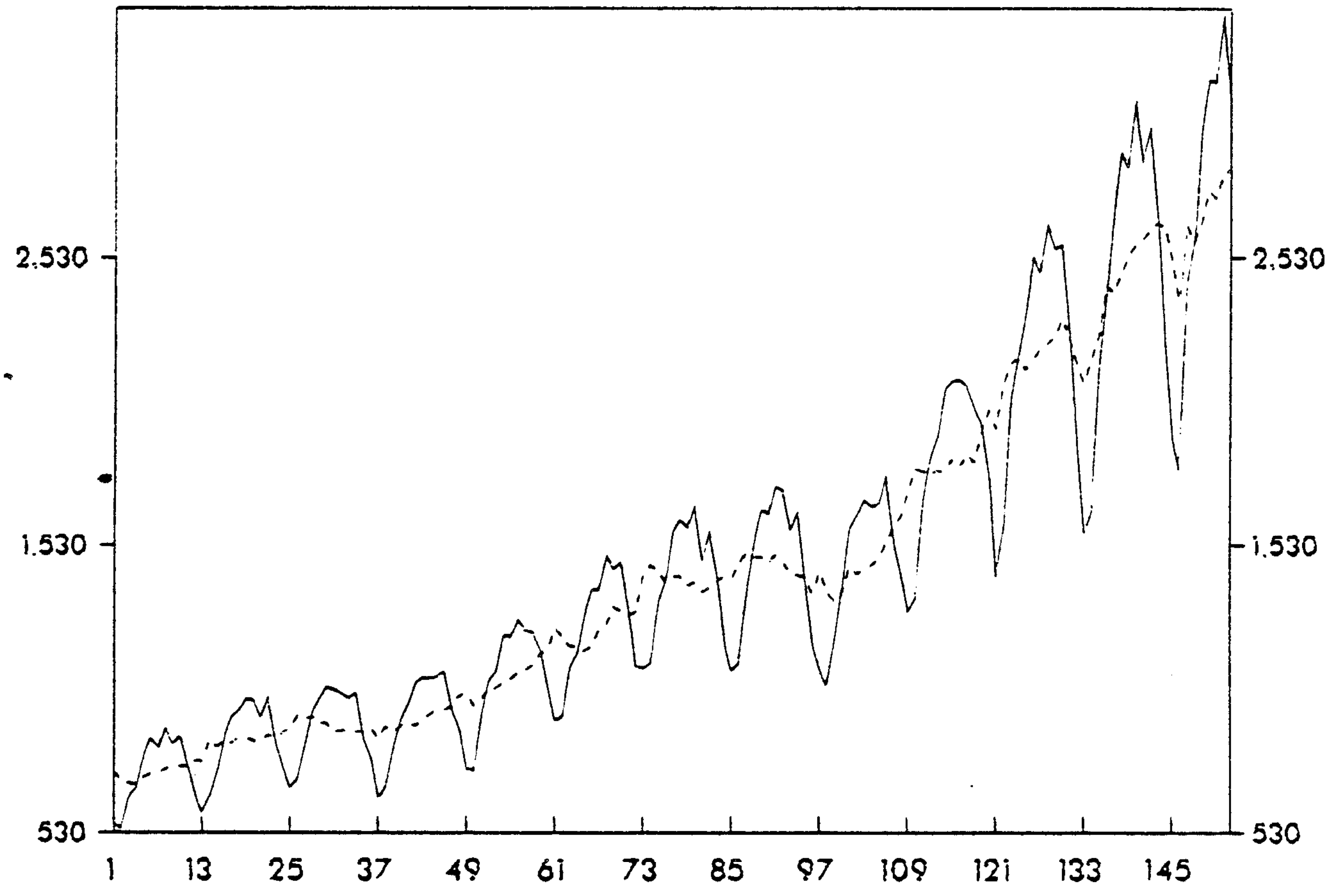


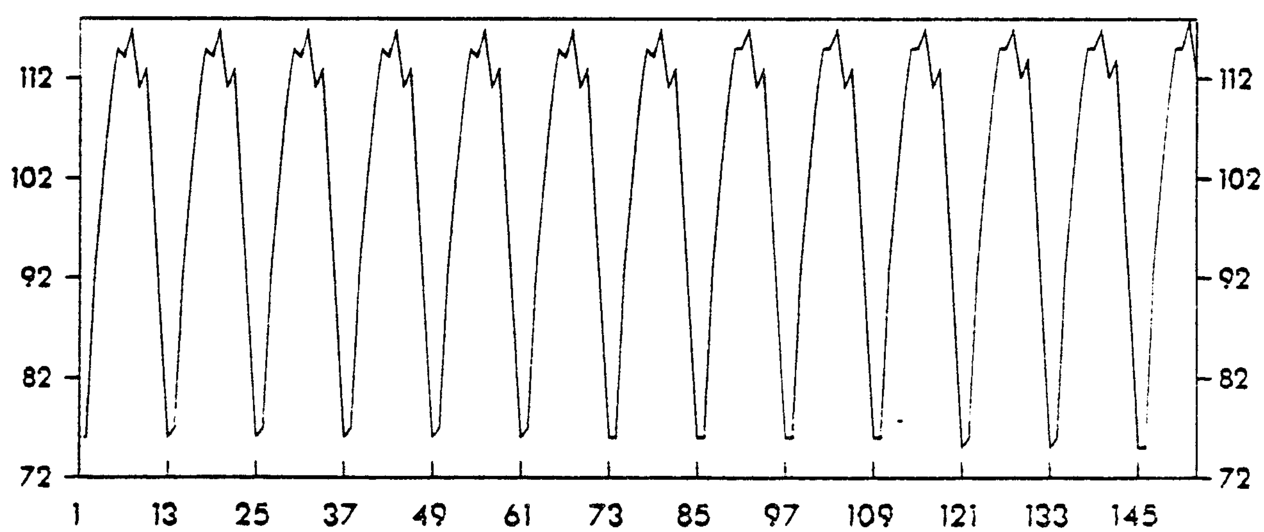
WHOLESALE SALES 518 X-11



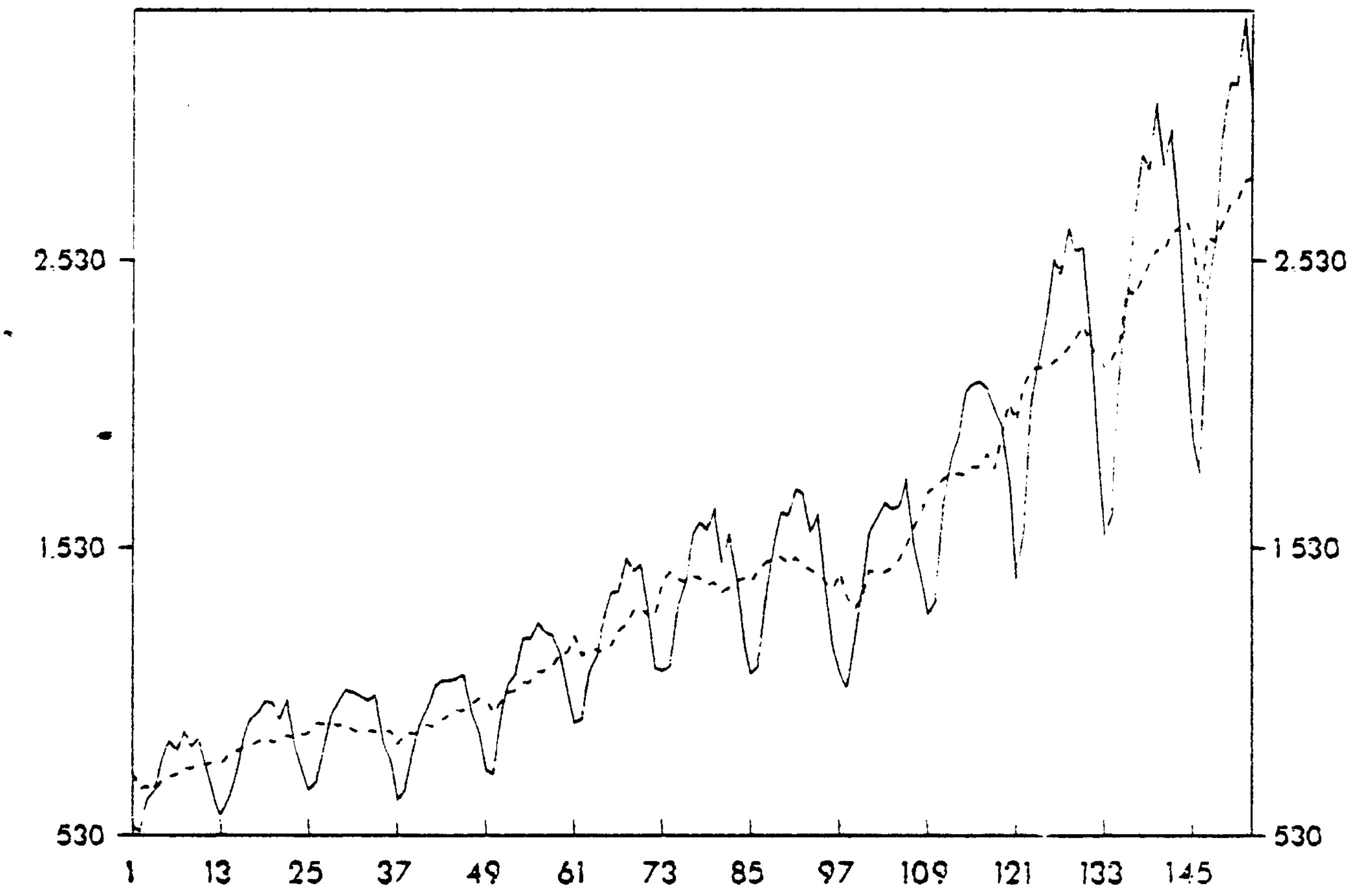


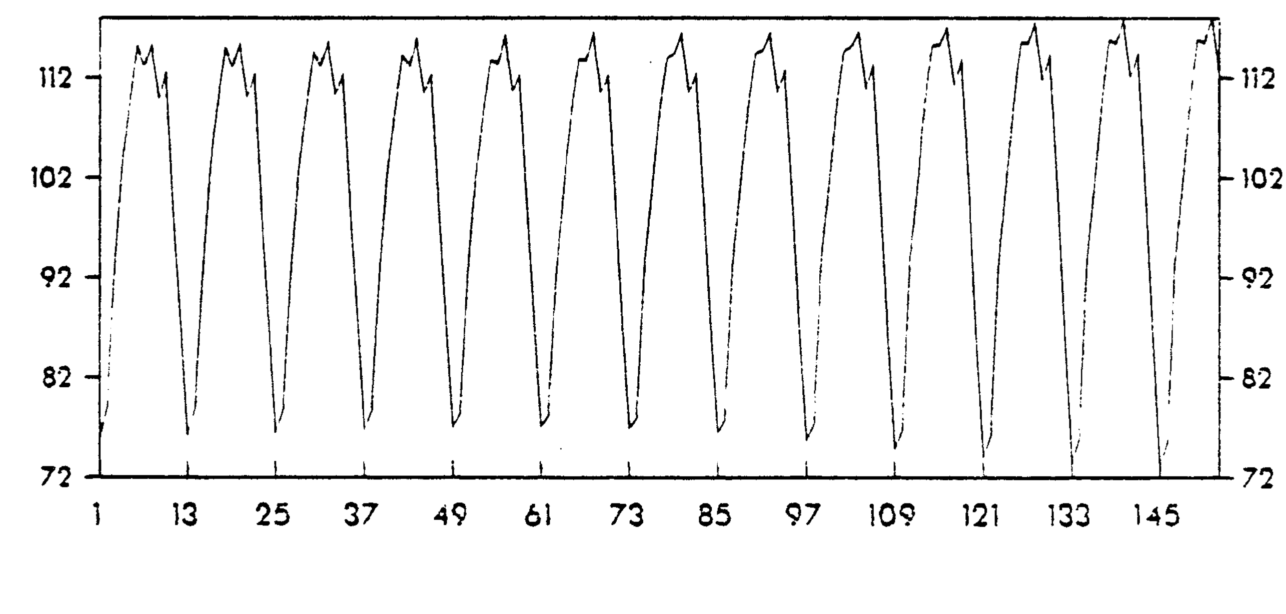
RETAIL 520002 MODEL BASED



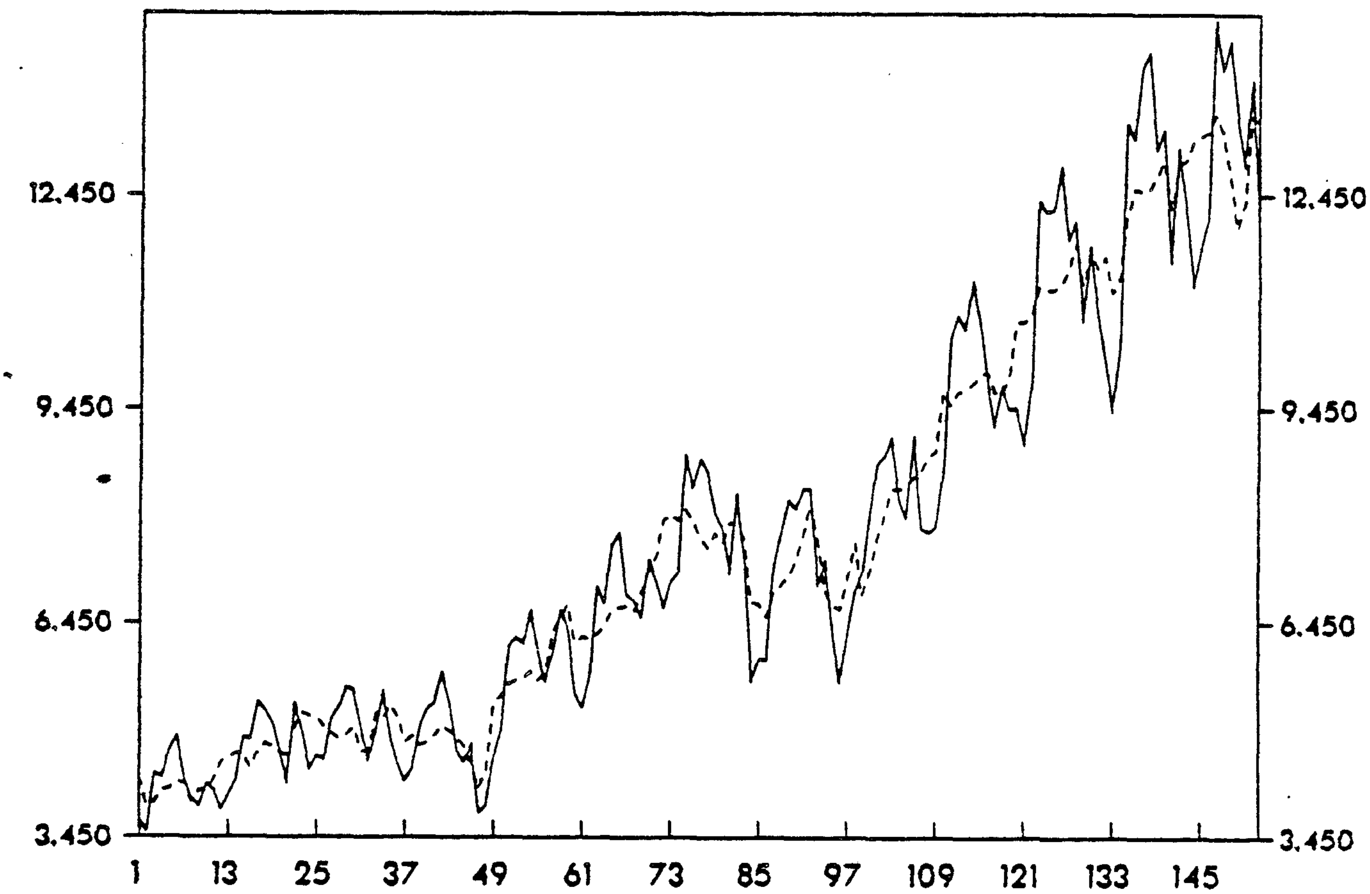


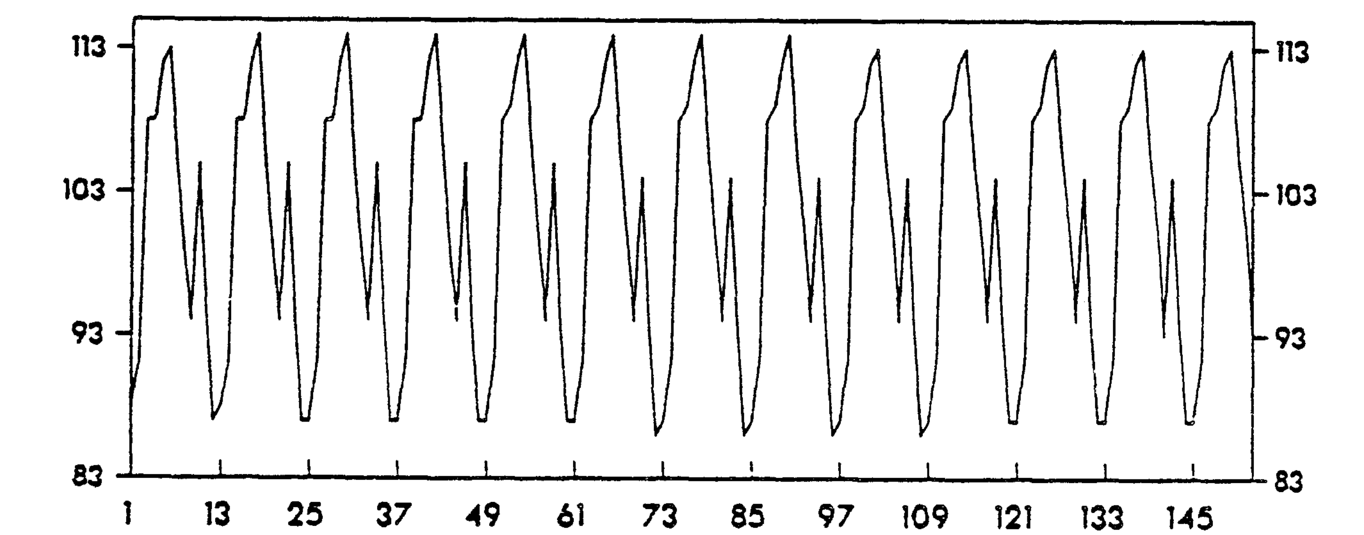
RETAIL 520002 X-11



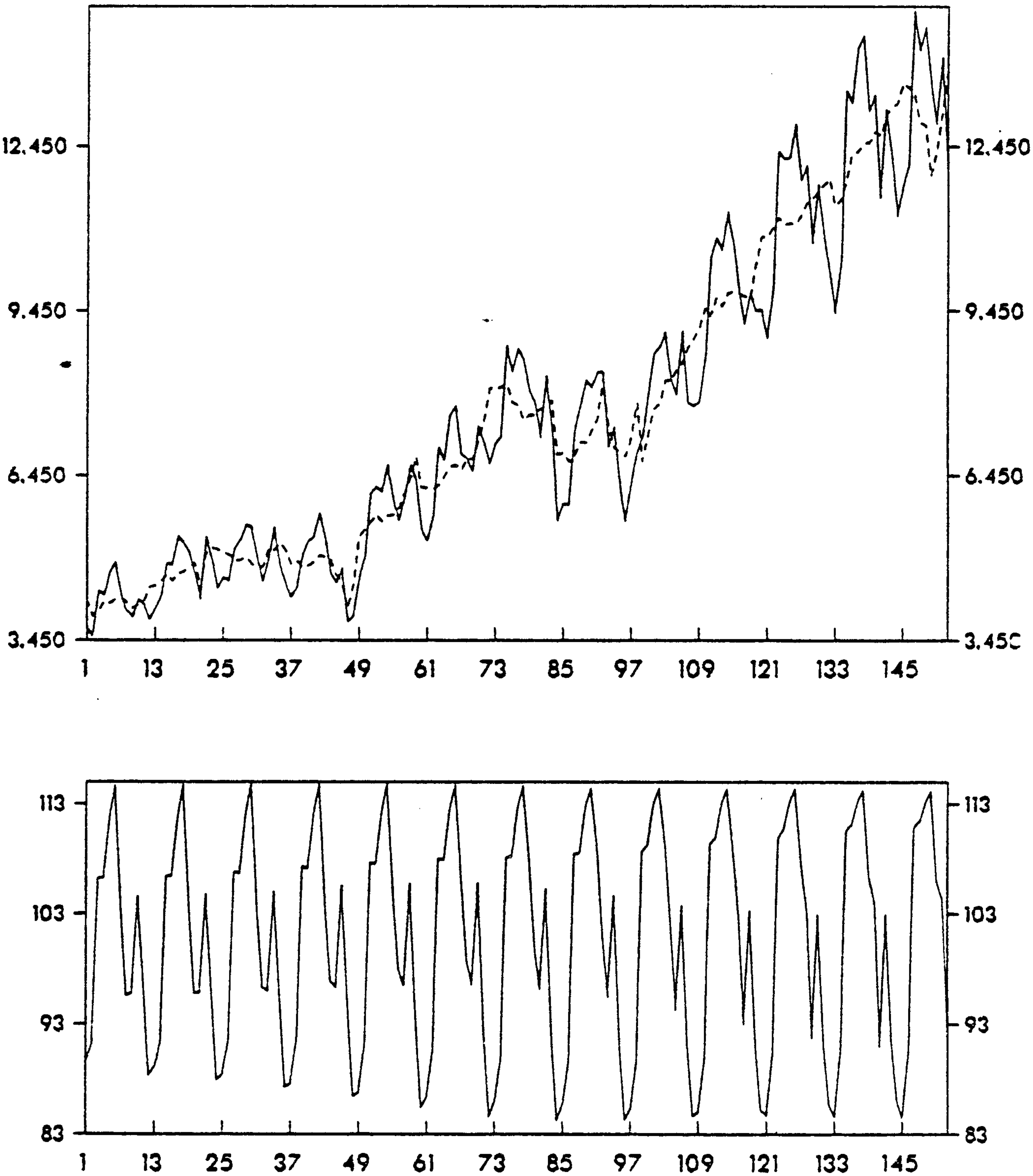


RETAIL 550001 MODEL BASED

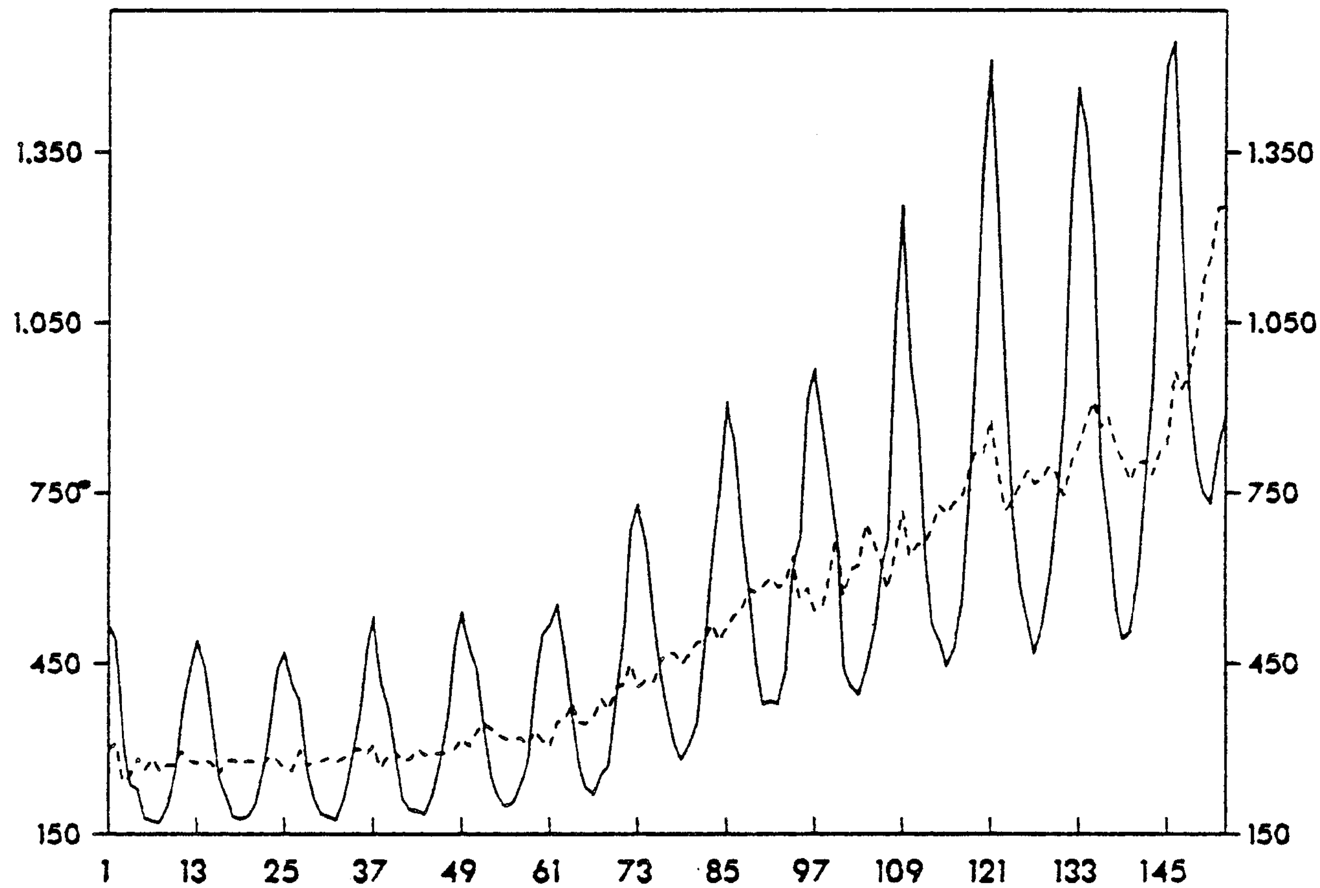


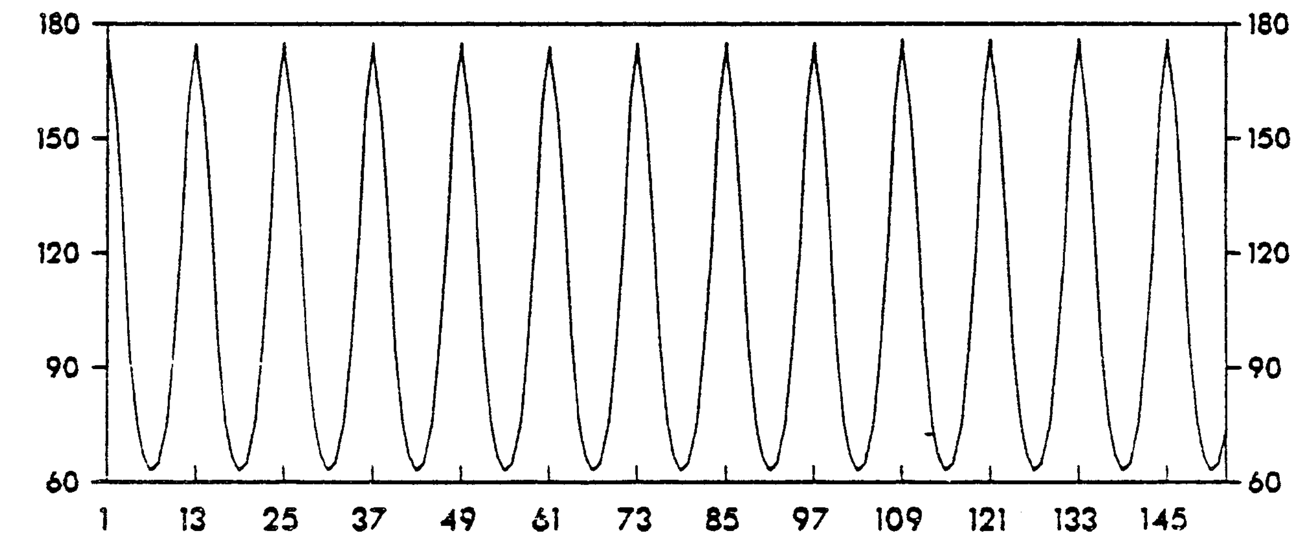


RETAIL 550001 X-11

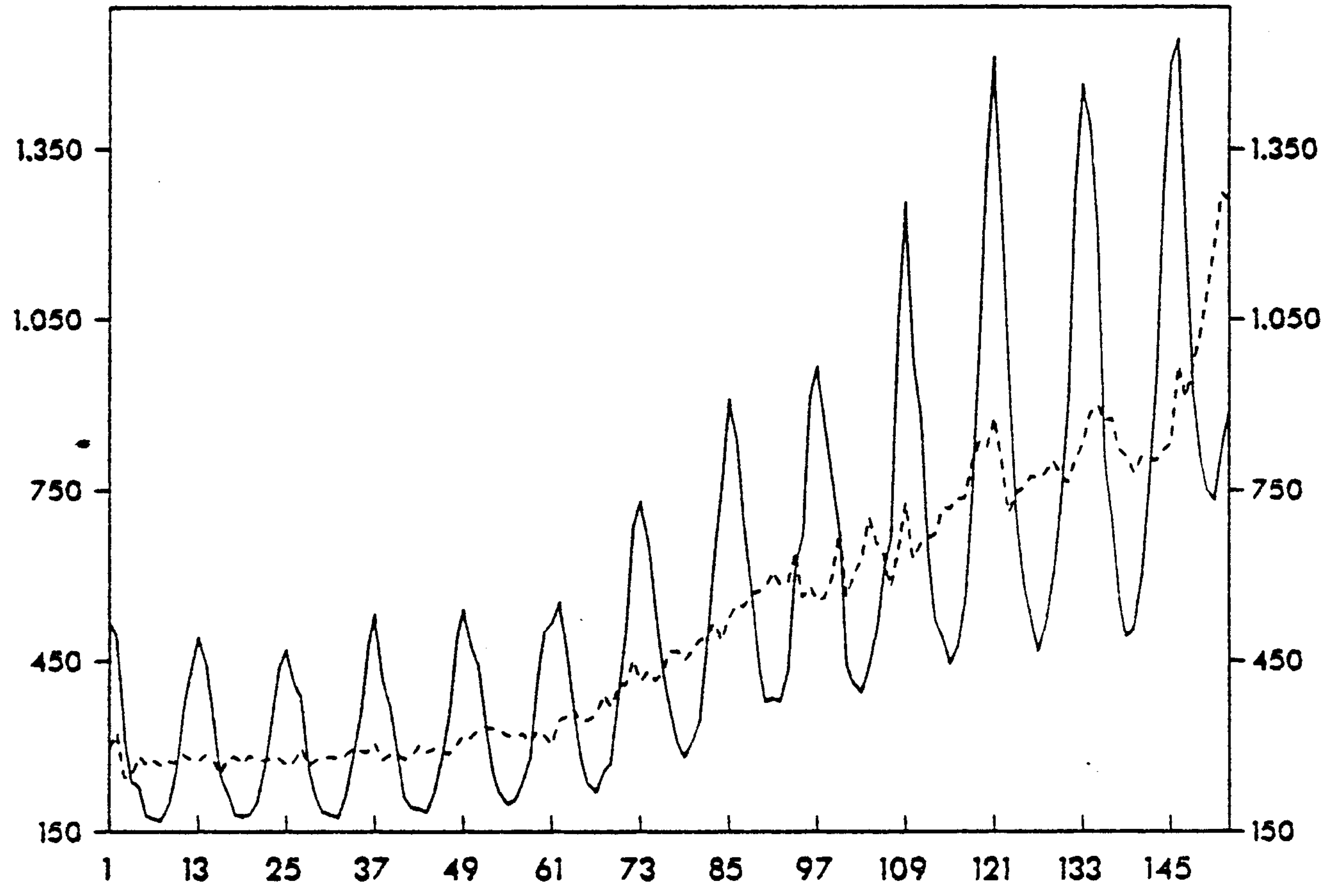


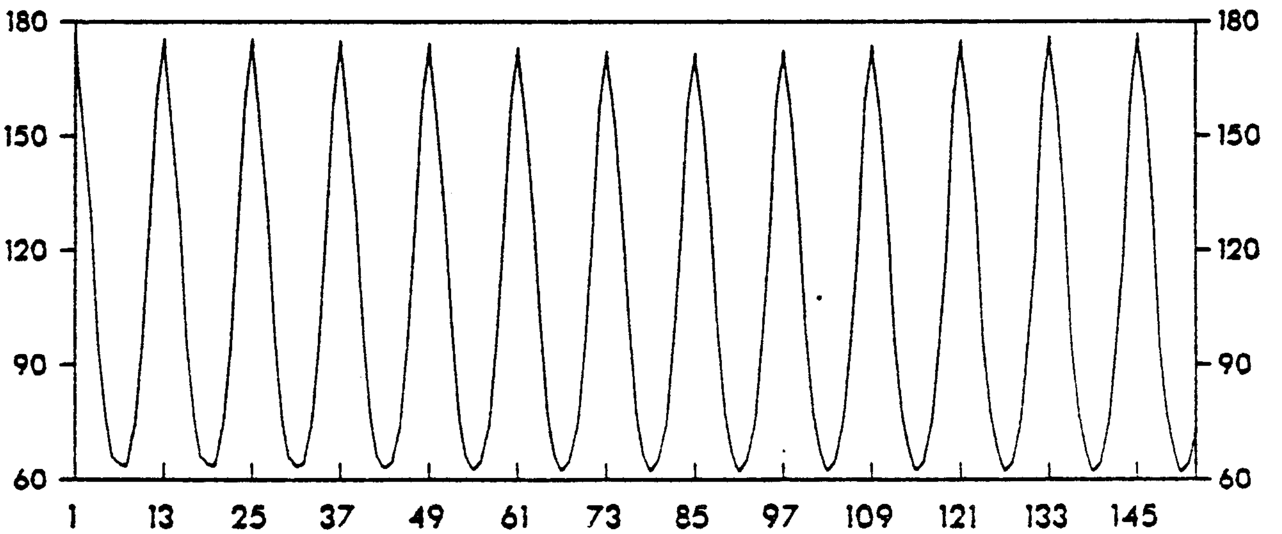
RETAIL 598000 MODEL BASED



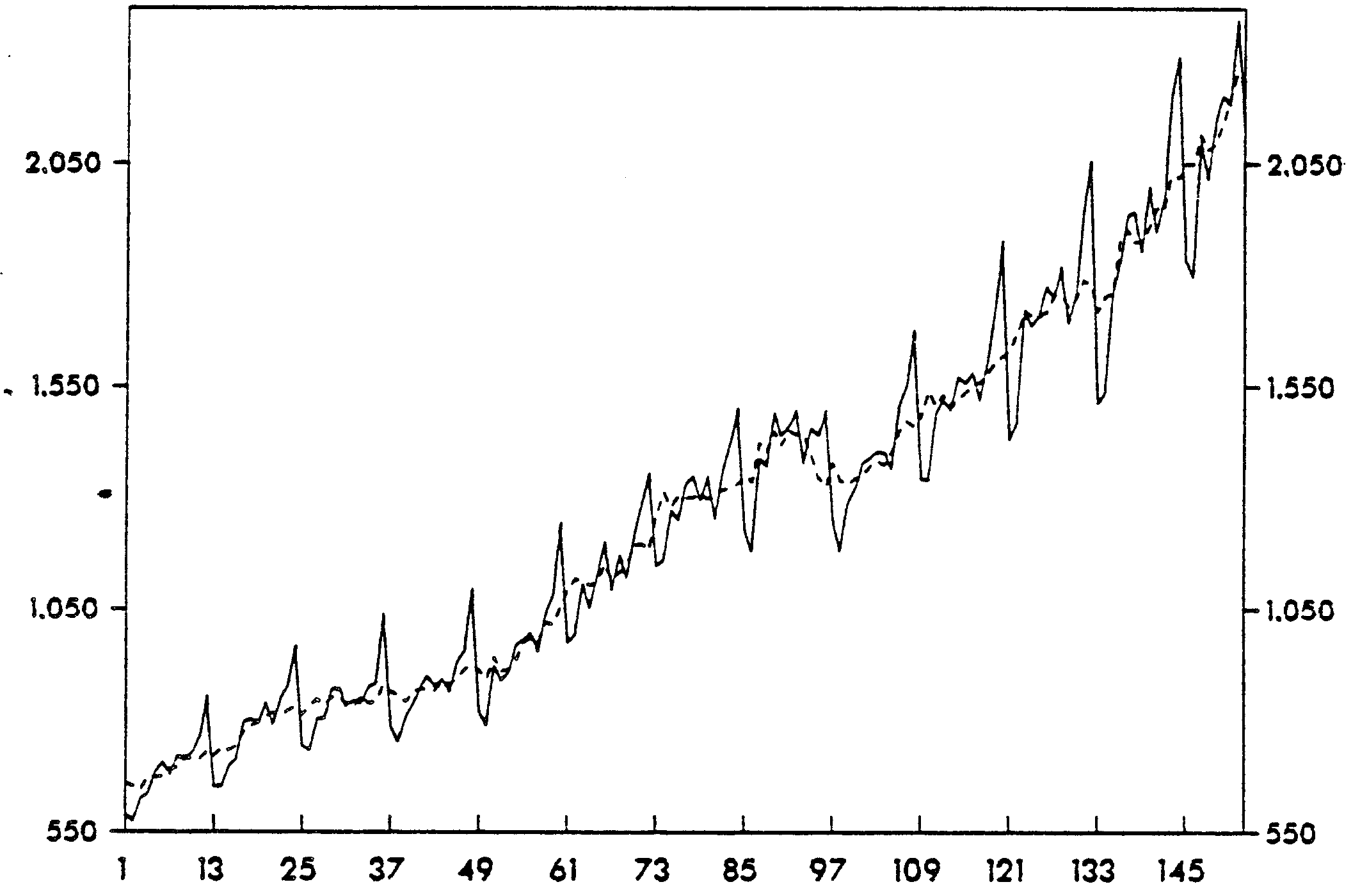


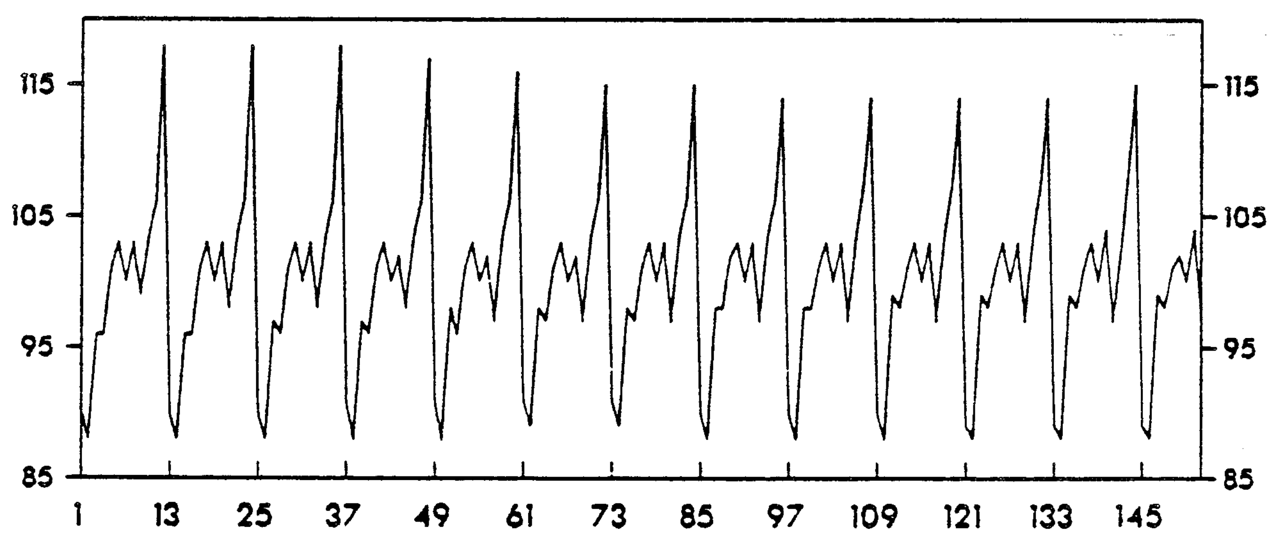
RETAIL 598000 X-11



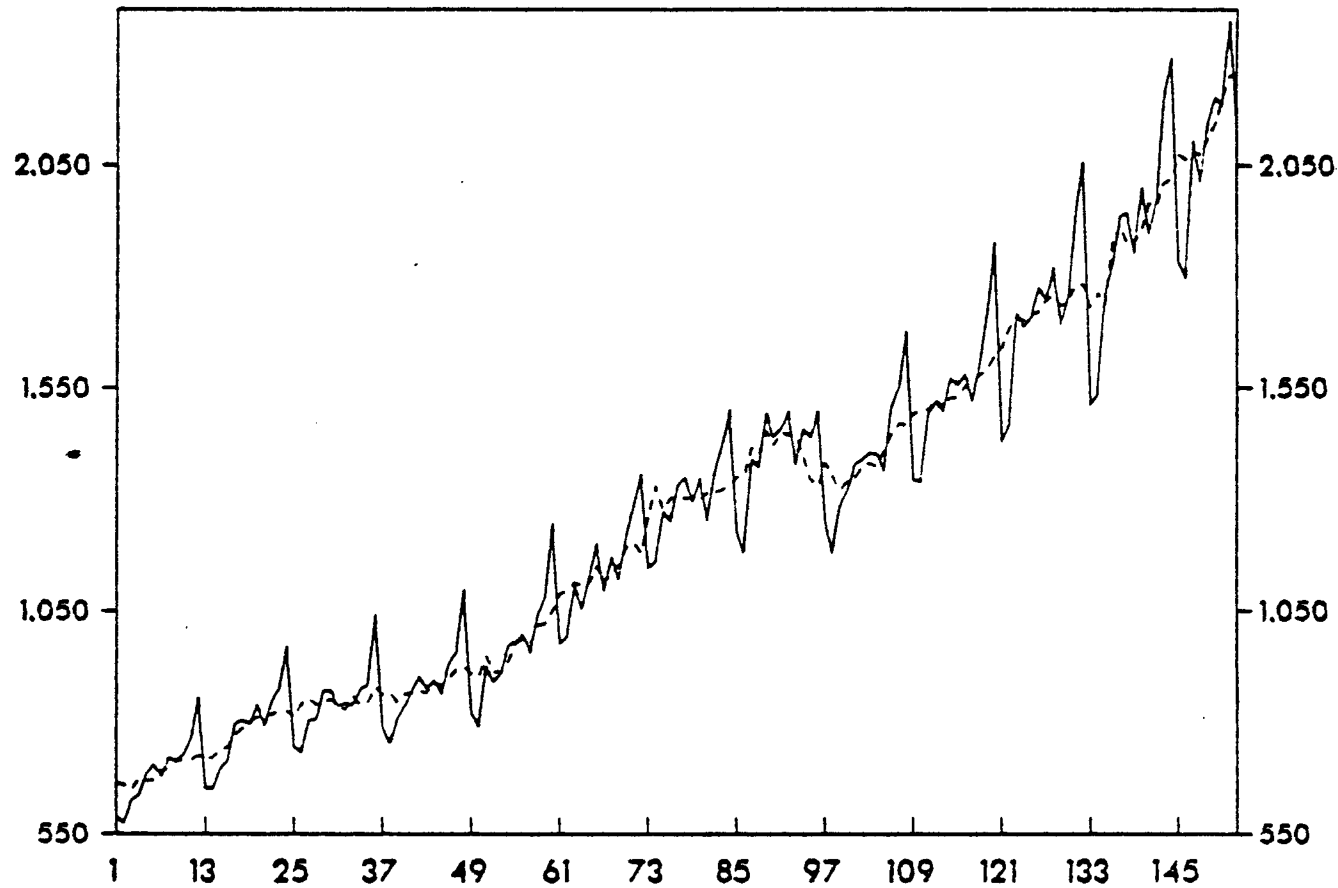


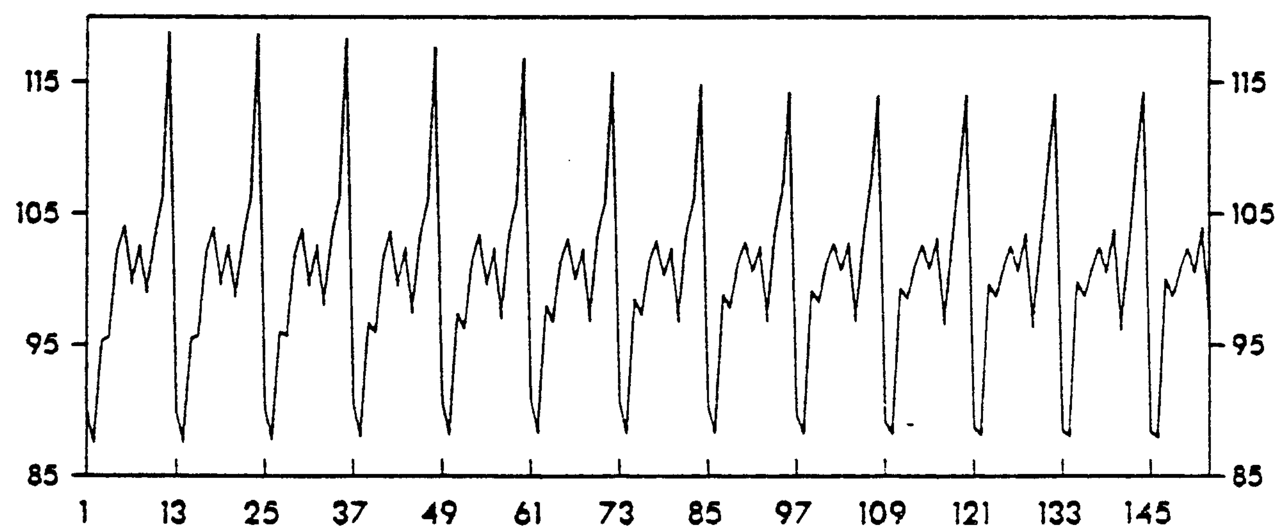
RETAIL 570001 MODEL BASED



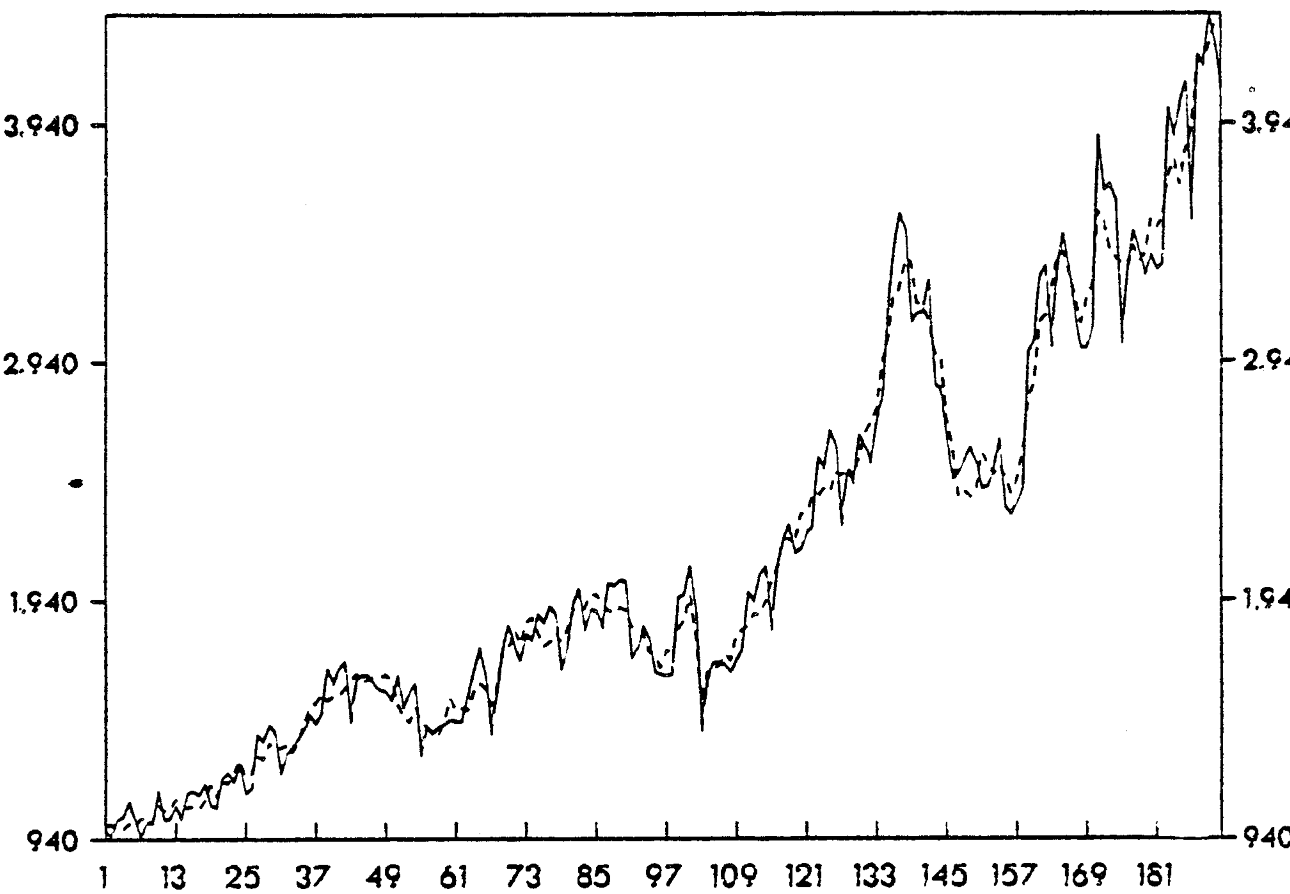


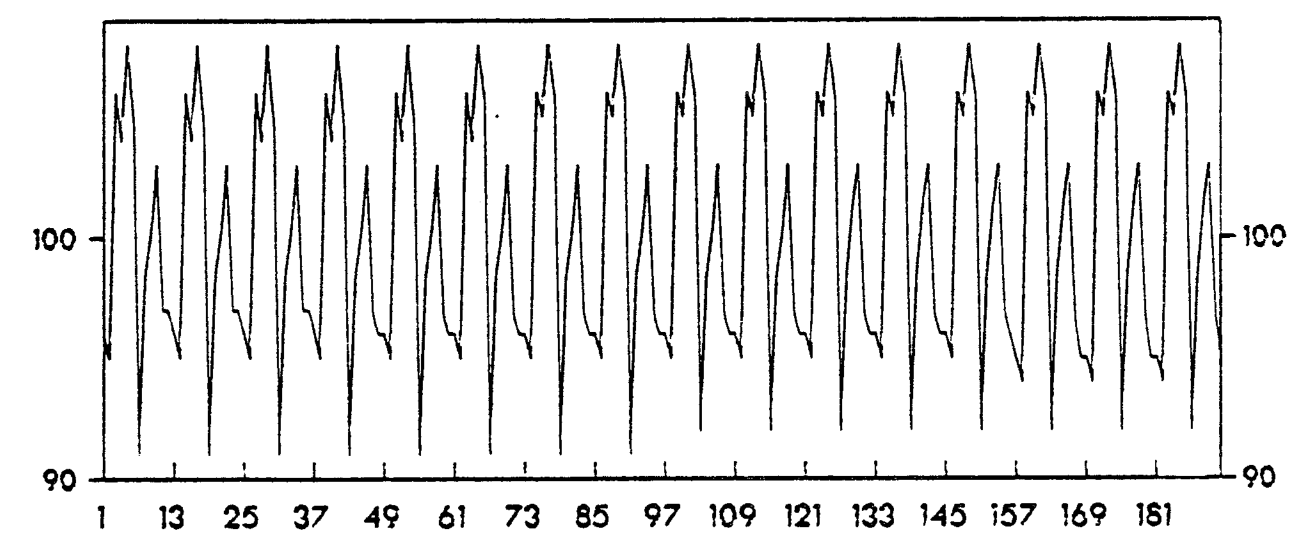
RETAIL 570001 X-11



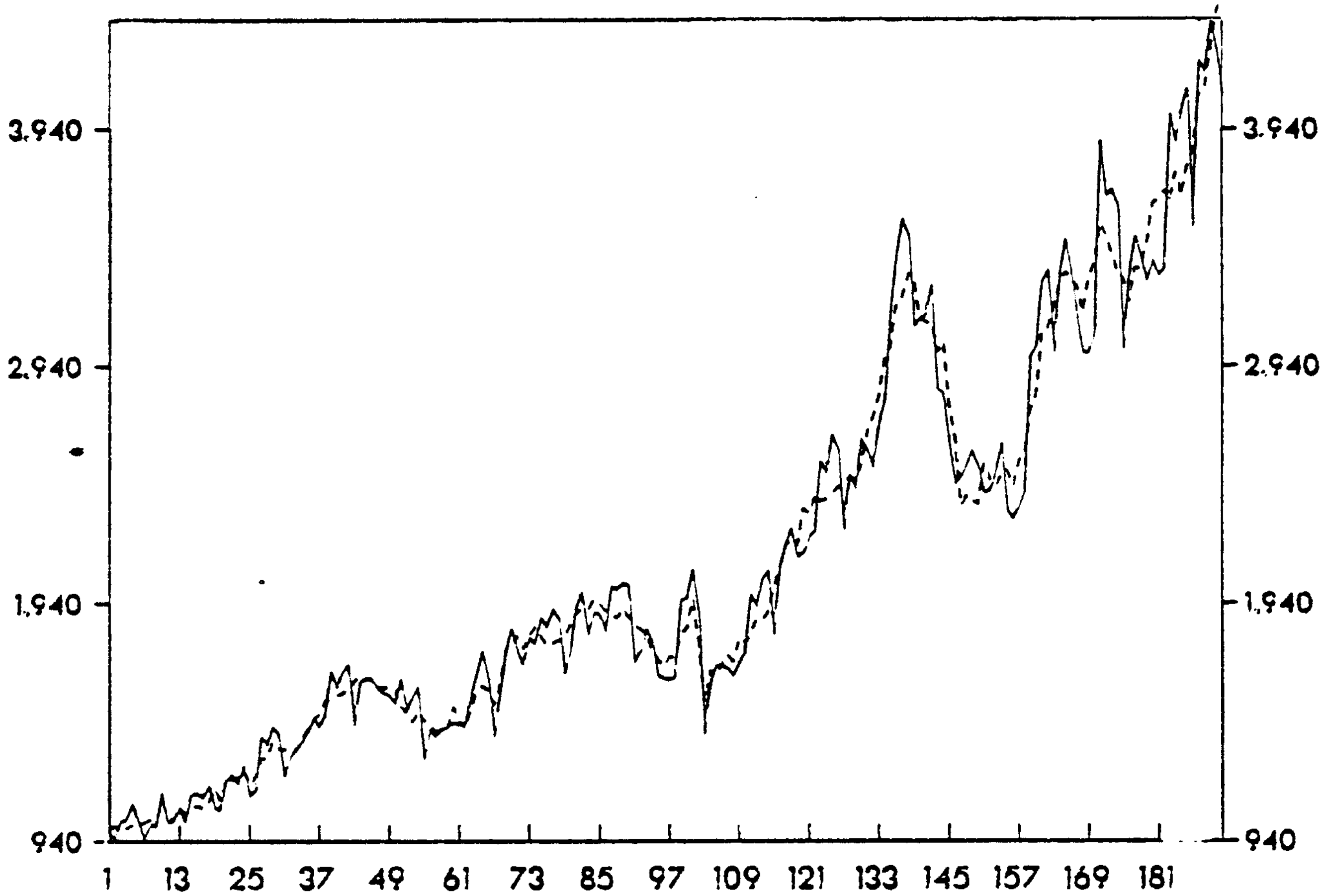


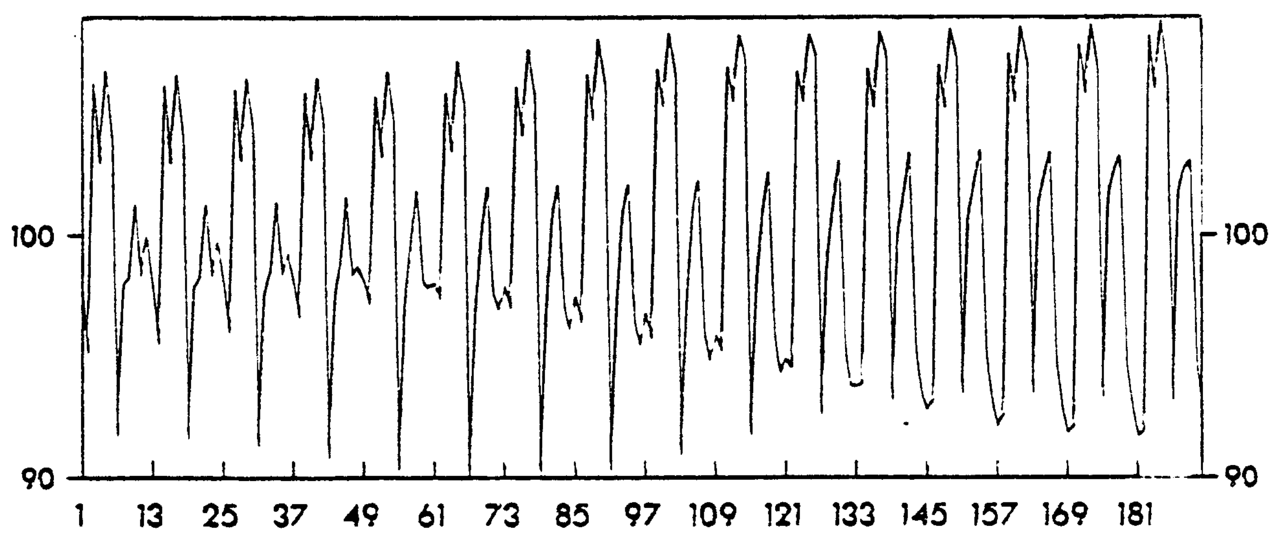
INDUSTRY S13VSU MODEL BASED



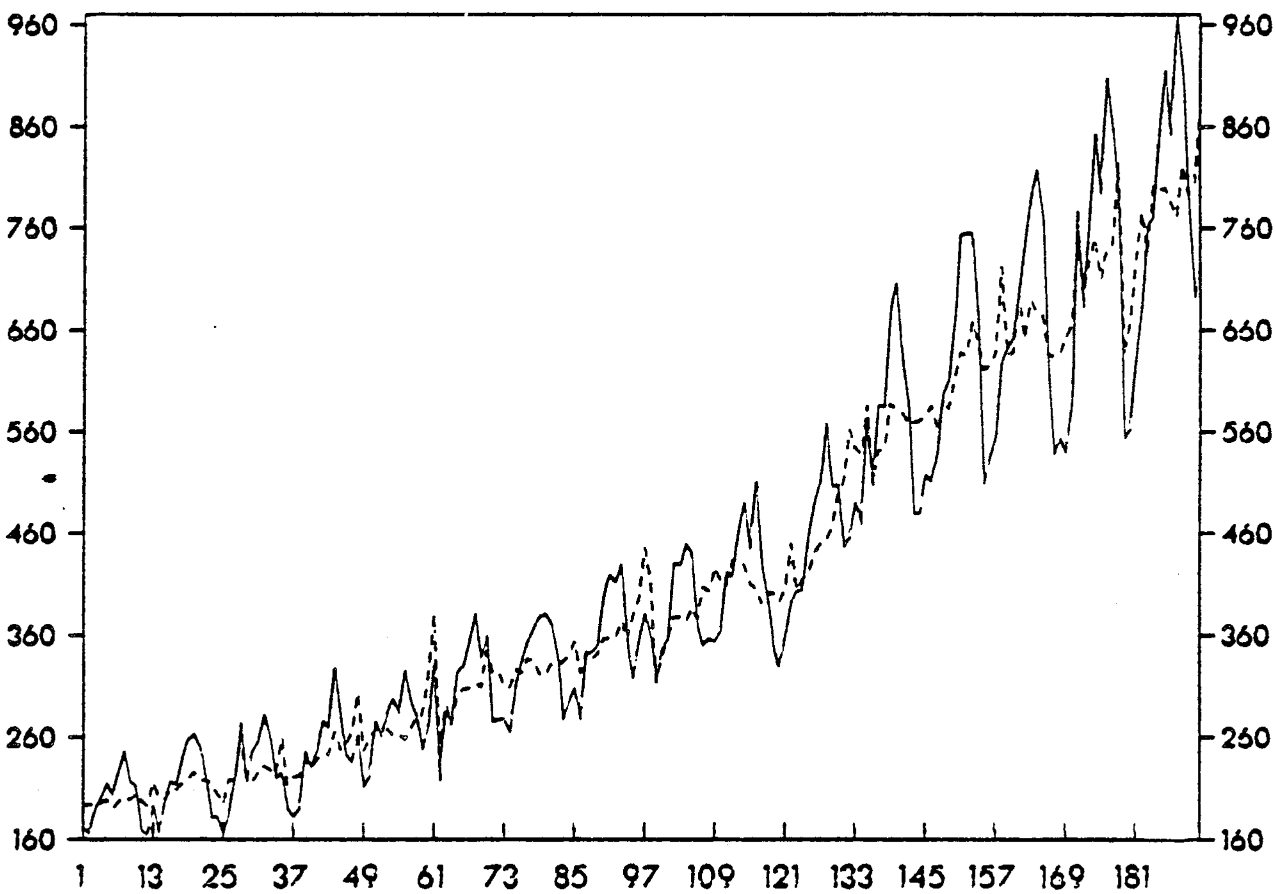


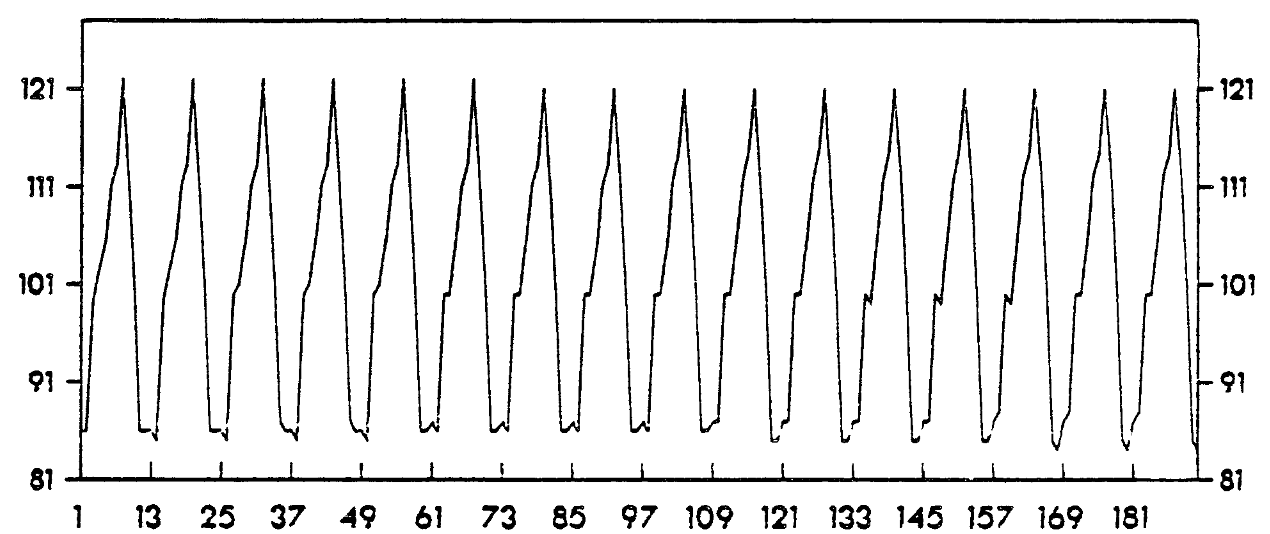
INDUSTRY S13VSU X-11 LOG



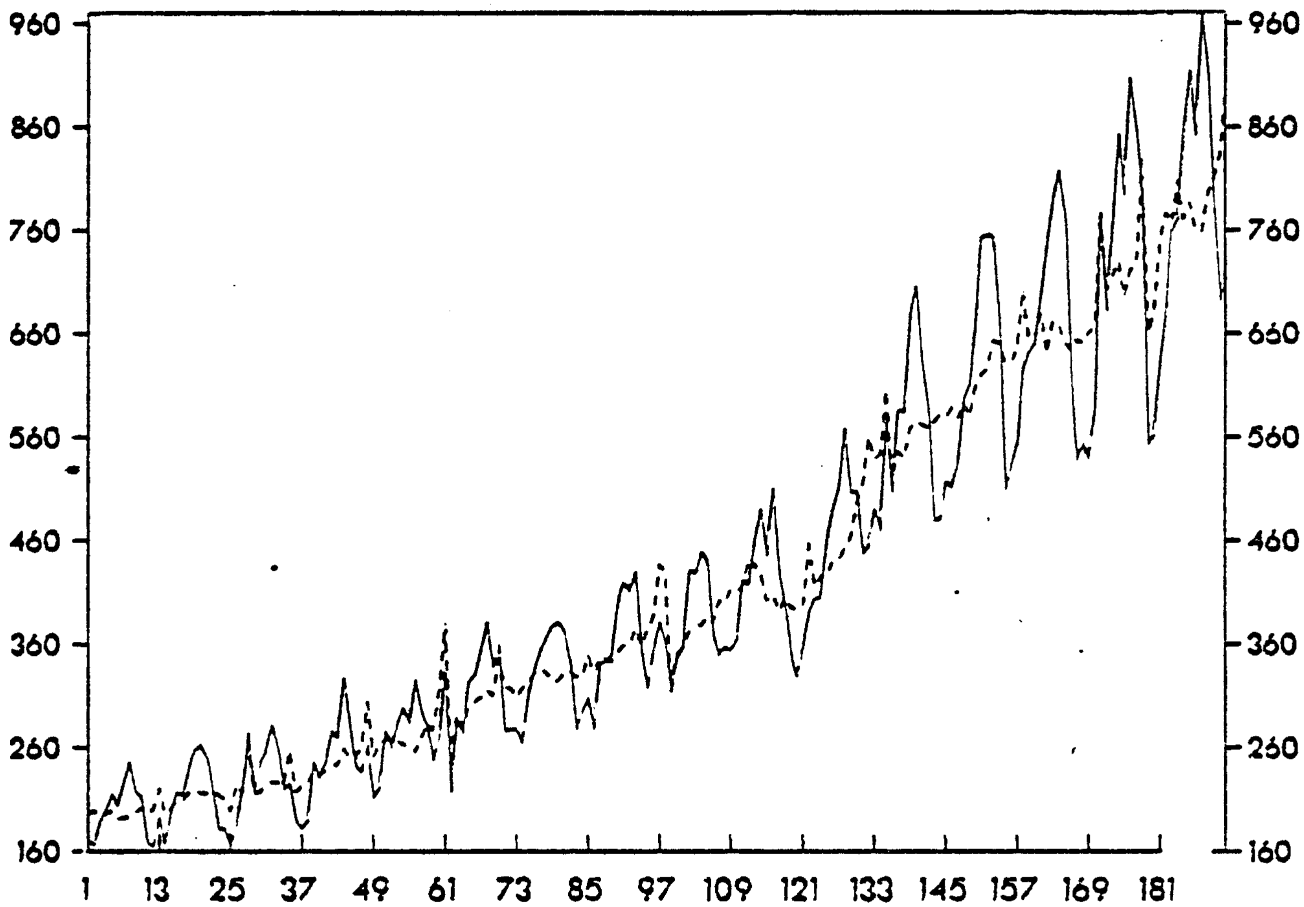


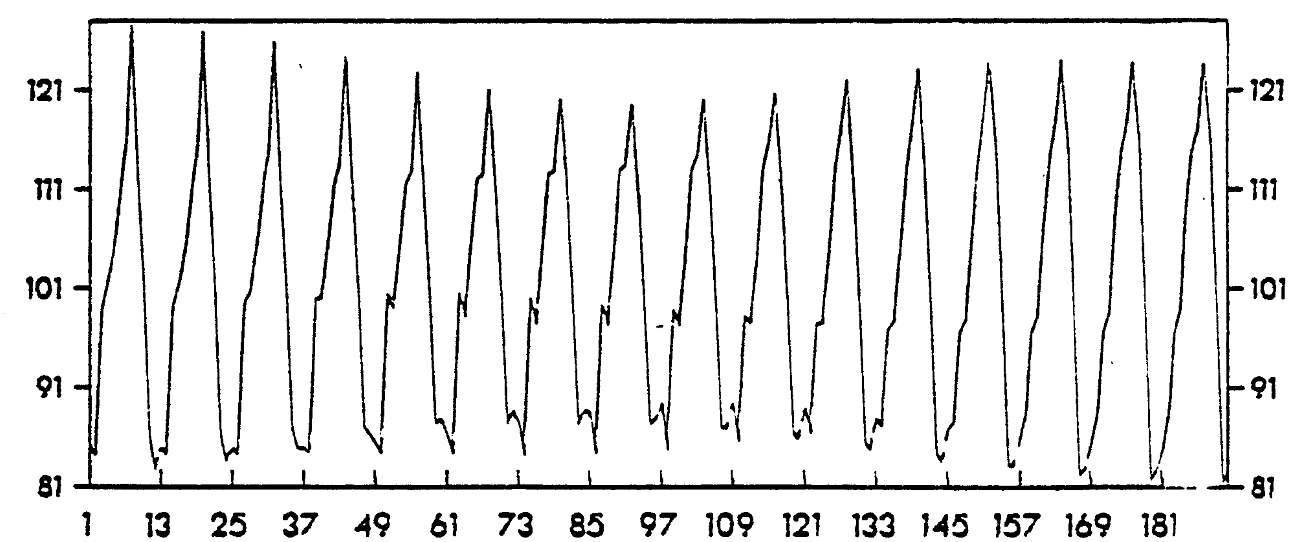
INDUSTRY S16VSU MODEL BASED



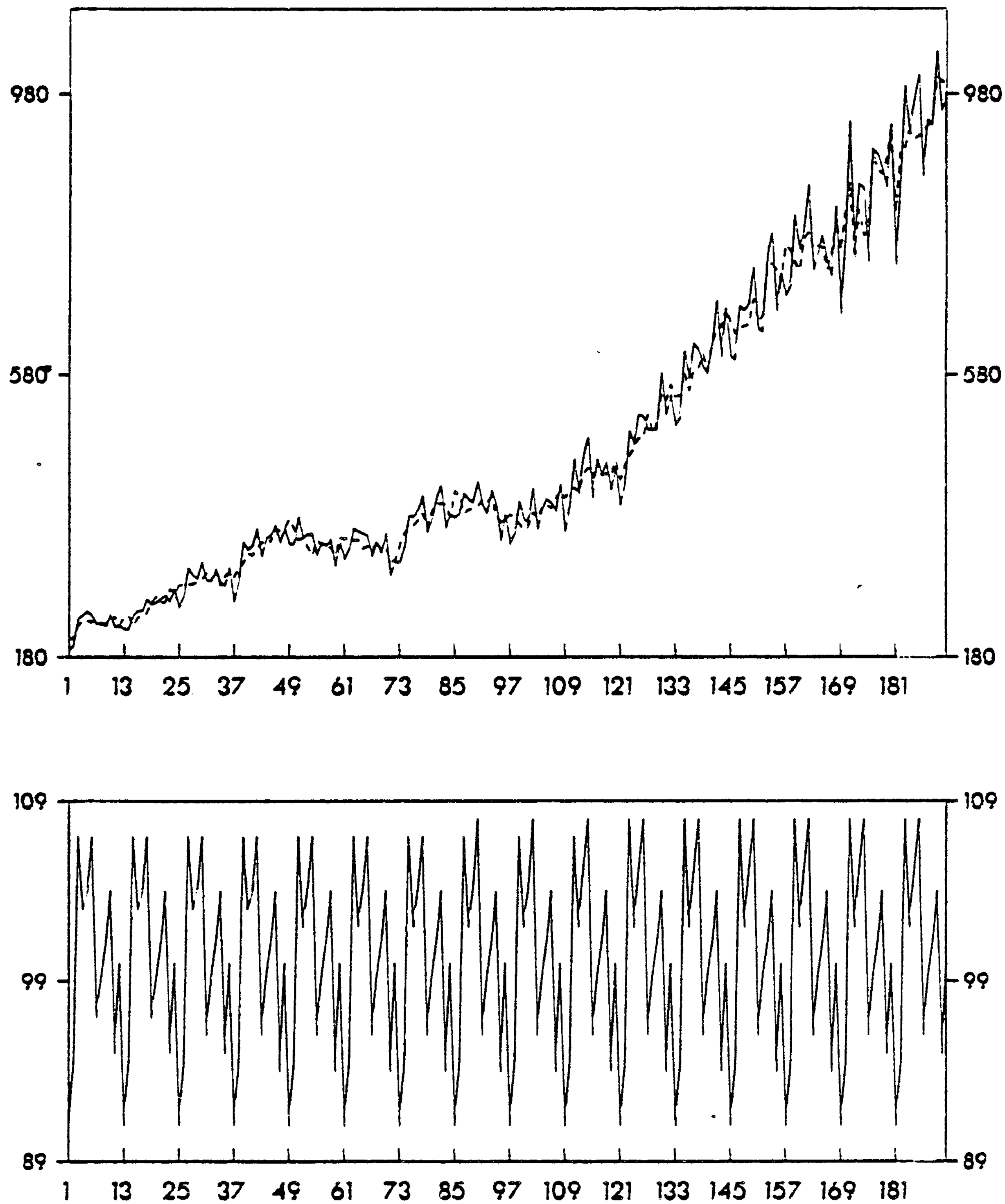


INDUSTRY S16VSU X-11 LOG

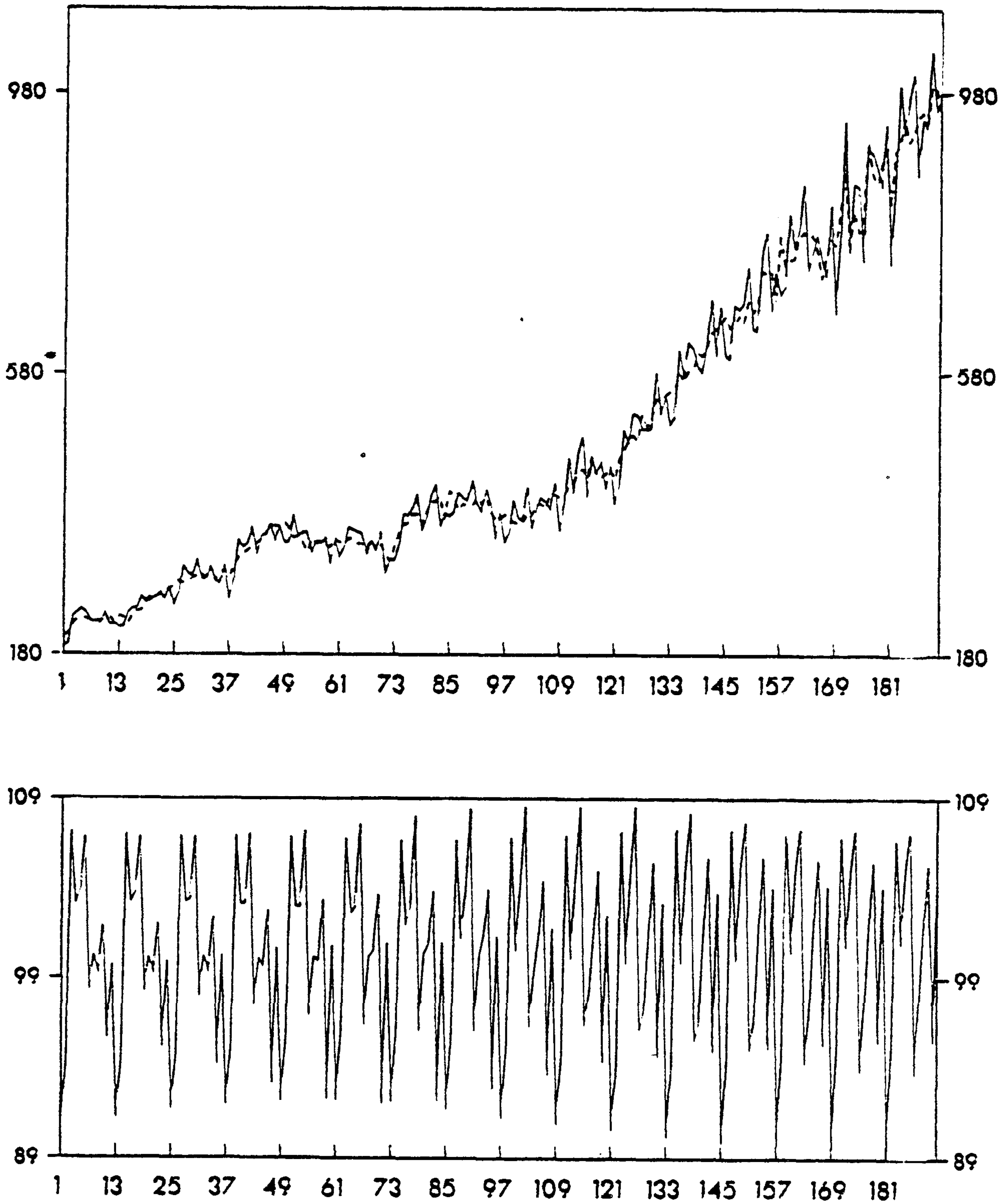




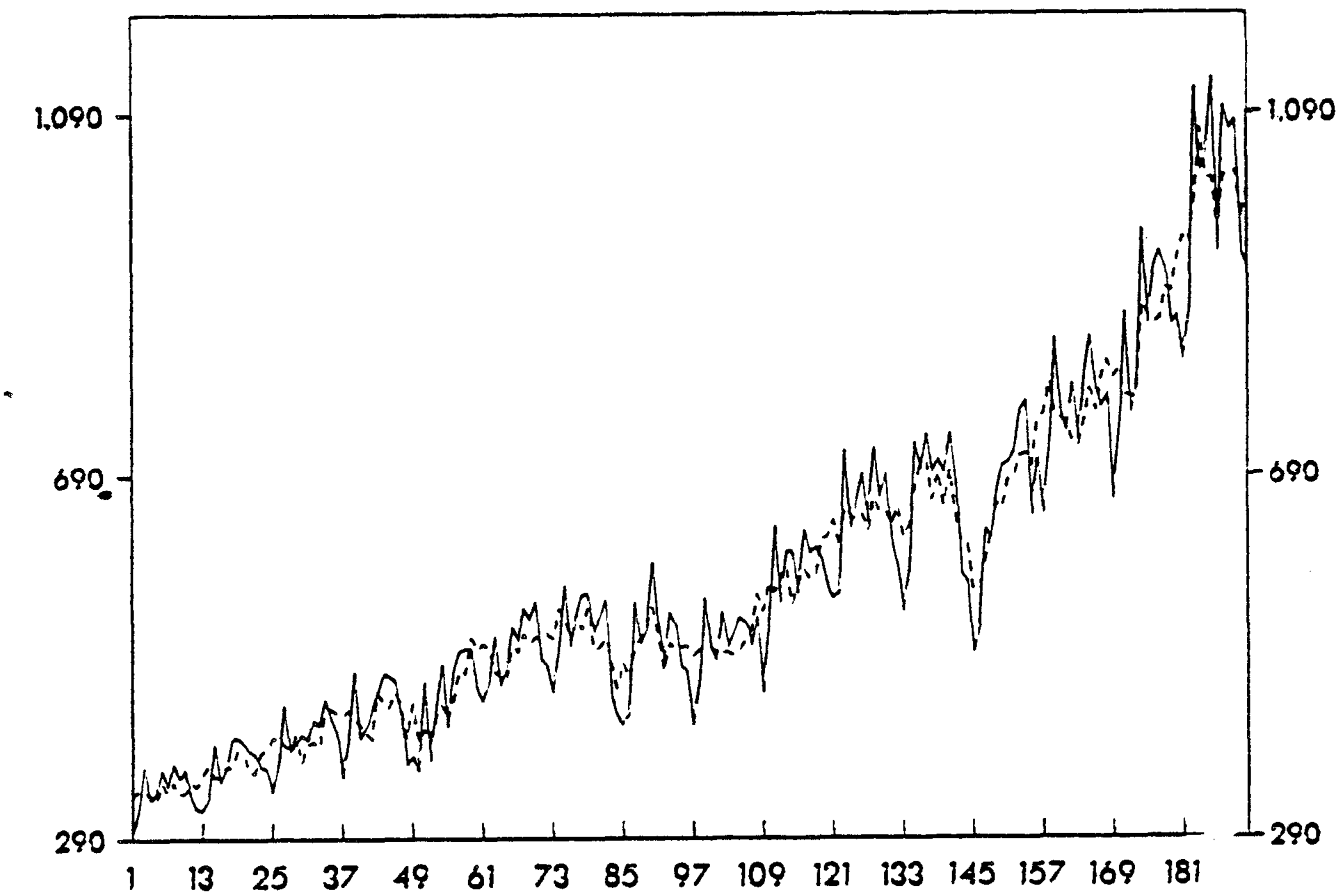
INDUSTRY S29VSU MODEL BASED

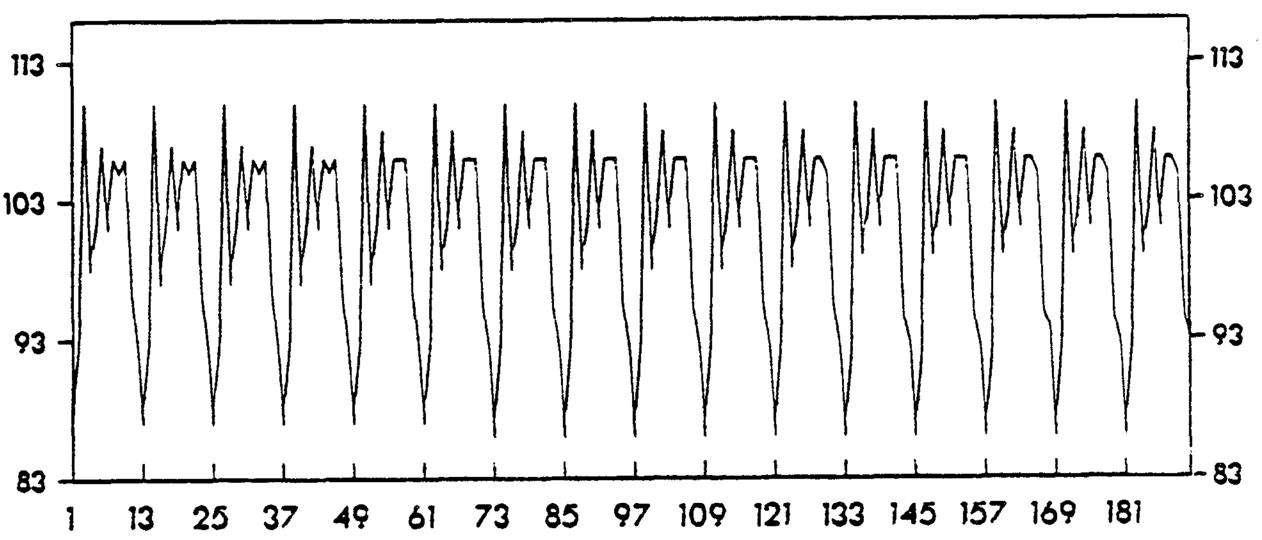


INDUSTRY S29VSU X-11 LOG

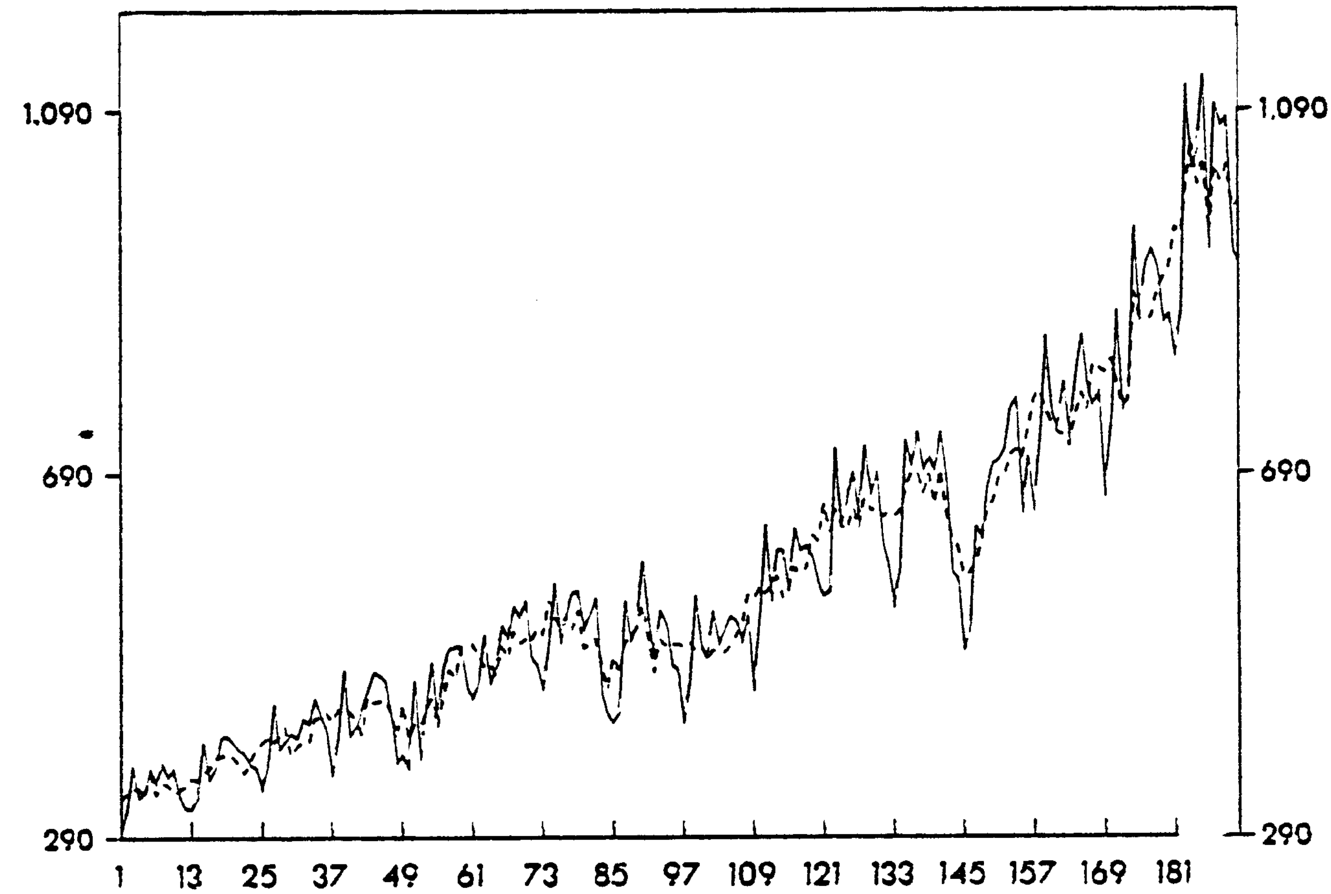


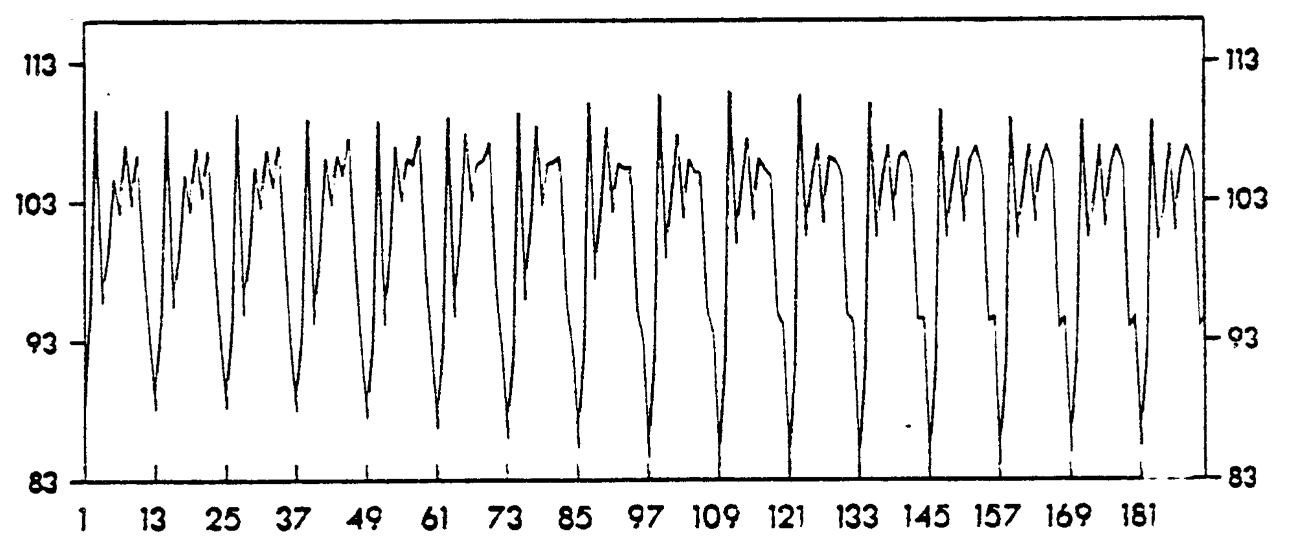
INDUSTRY S35VSU MODEL BASED



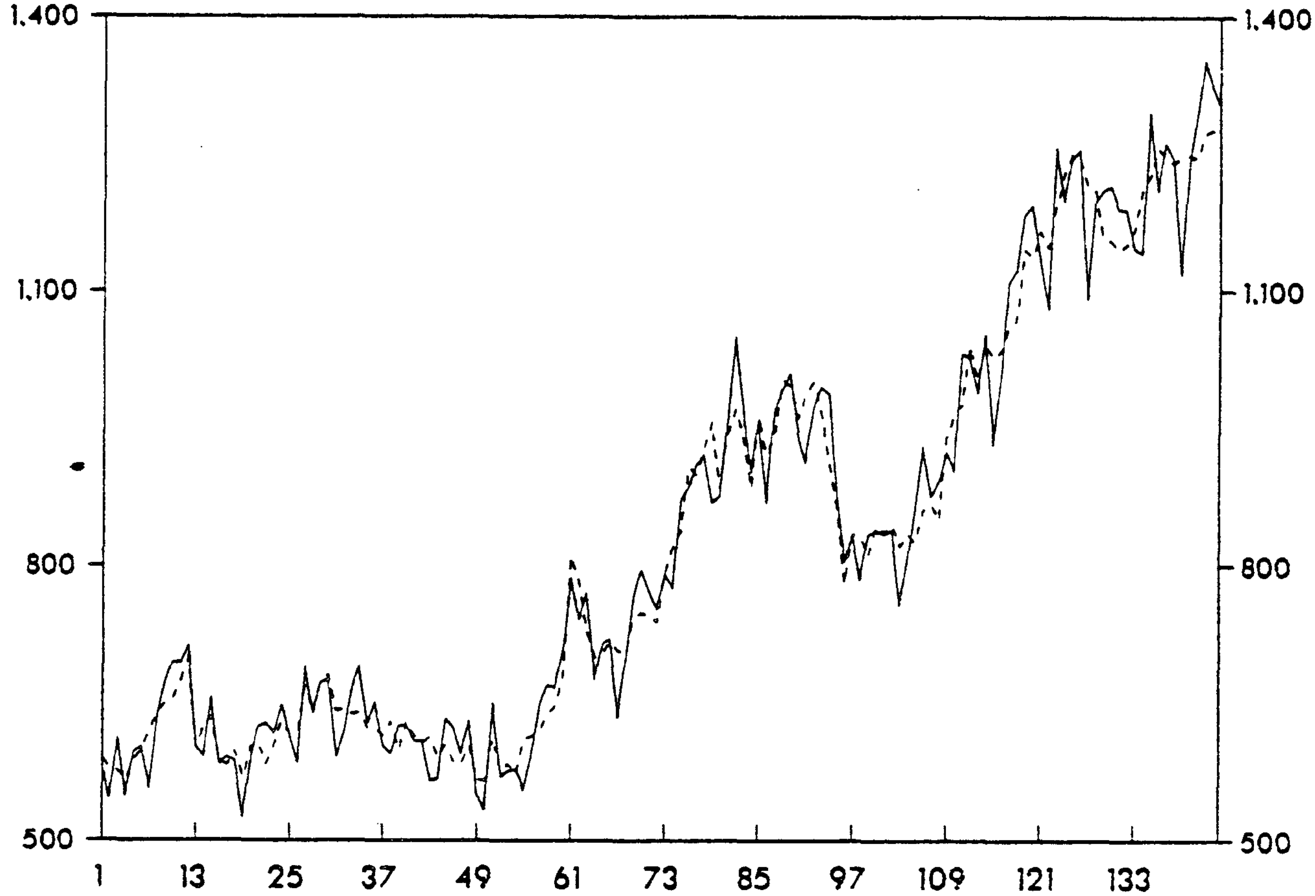


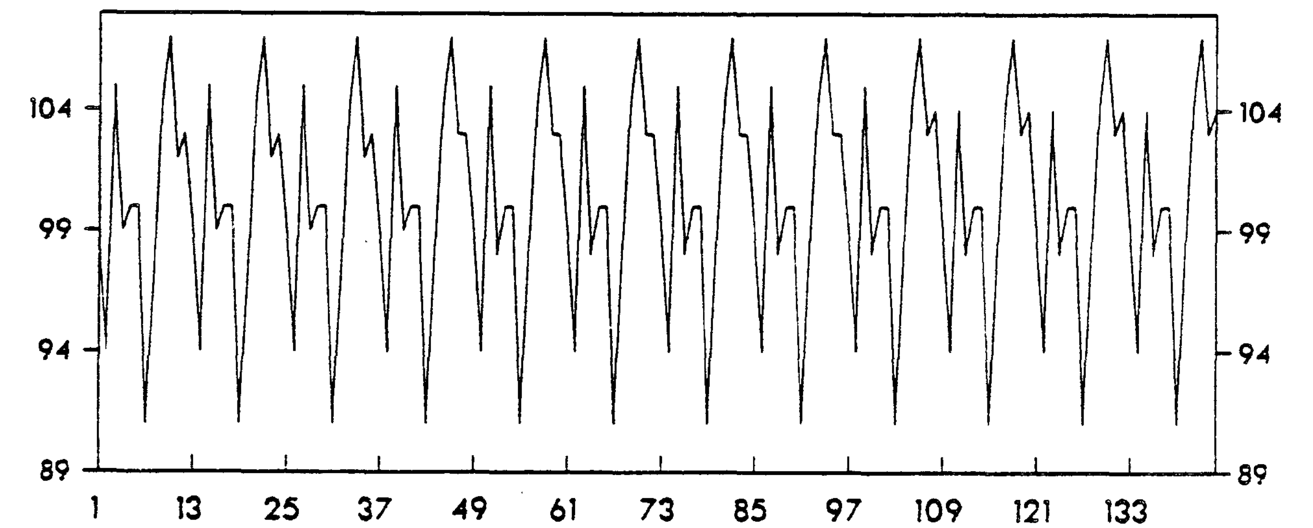
INDUSTRY S35VSU X-11 LOG



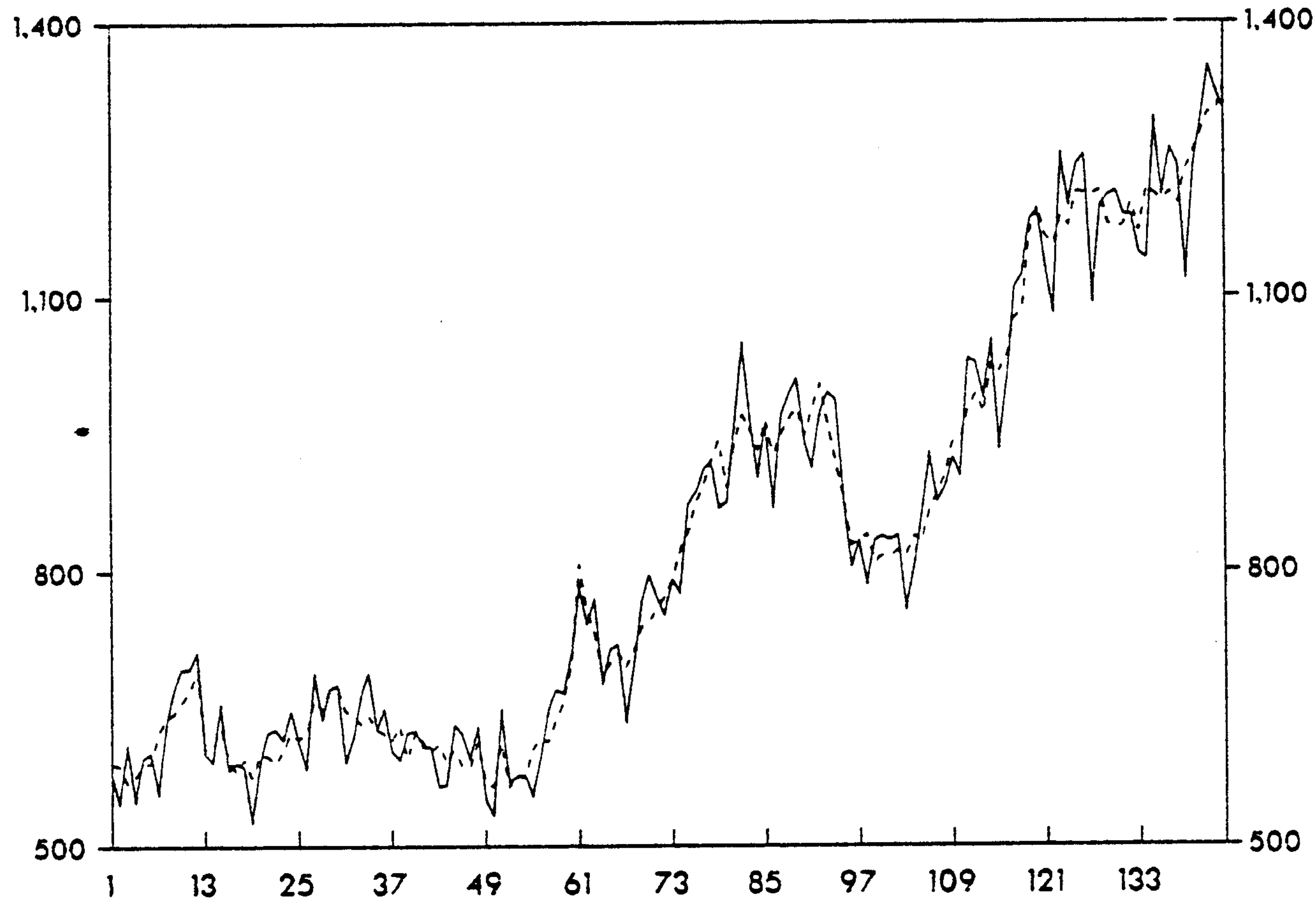


S38VSU MODEL BASED

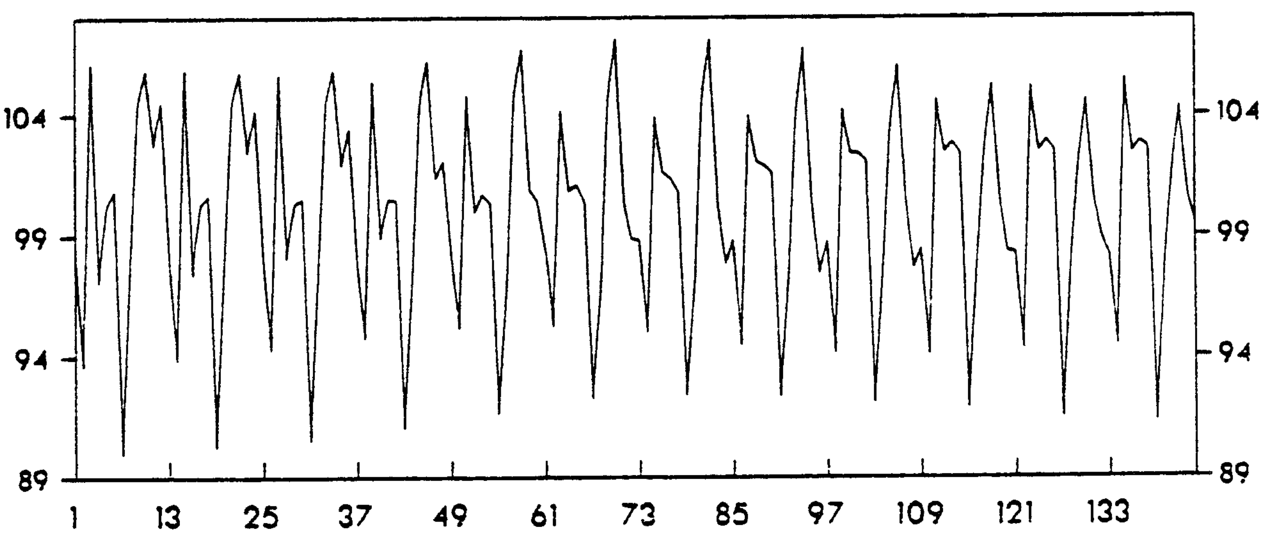




S38VSU X-11

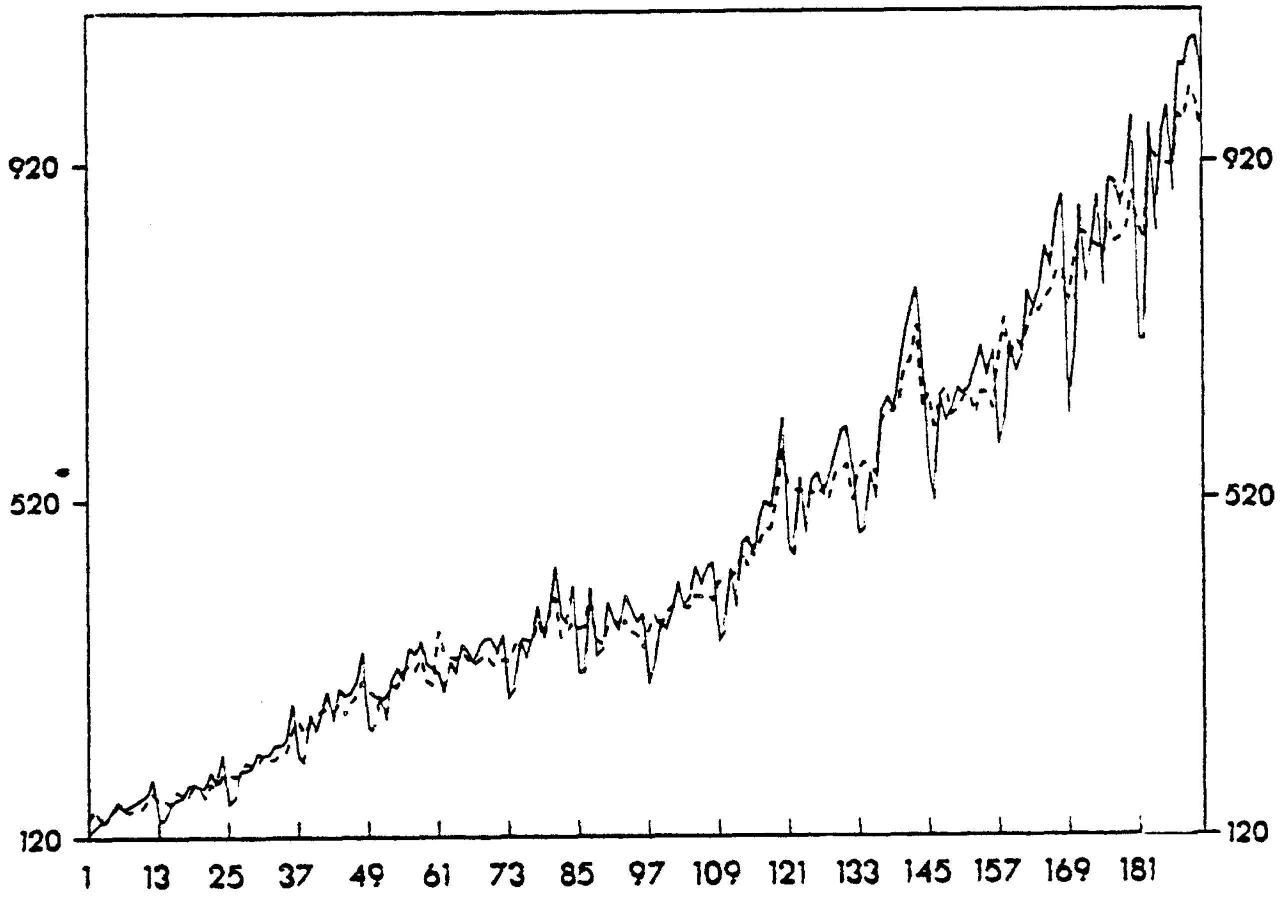


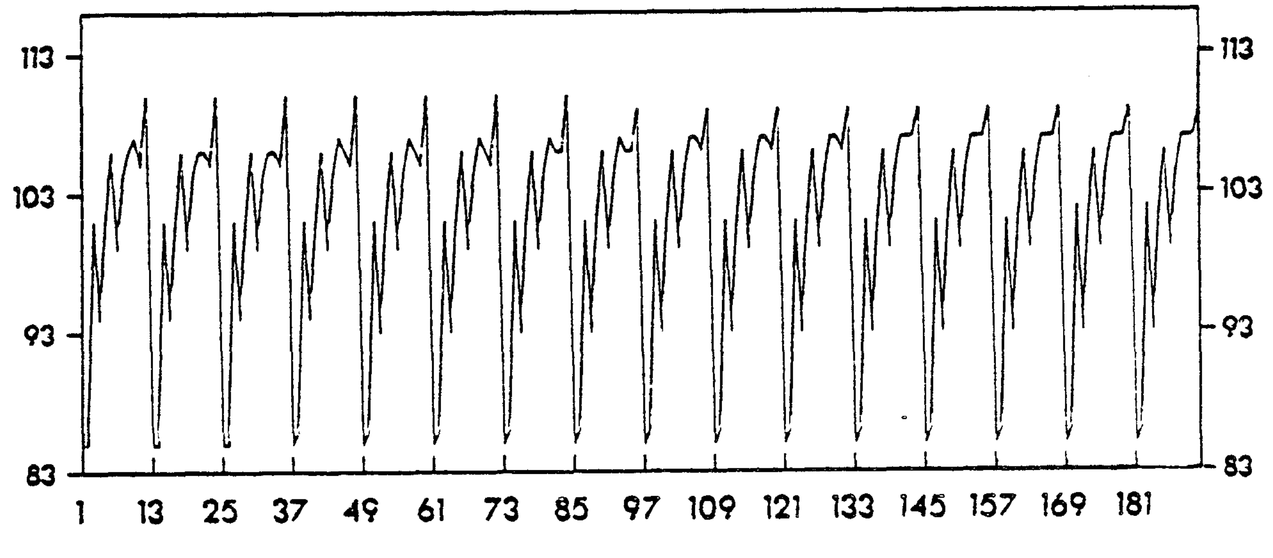
1,400
1,100
800
500
1
13
25
37
49
61
73
85
97
109
121
133



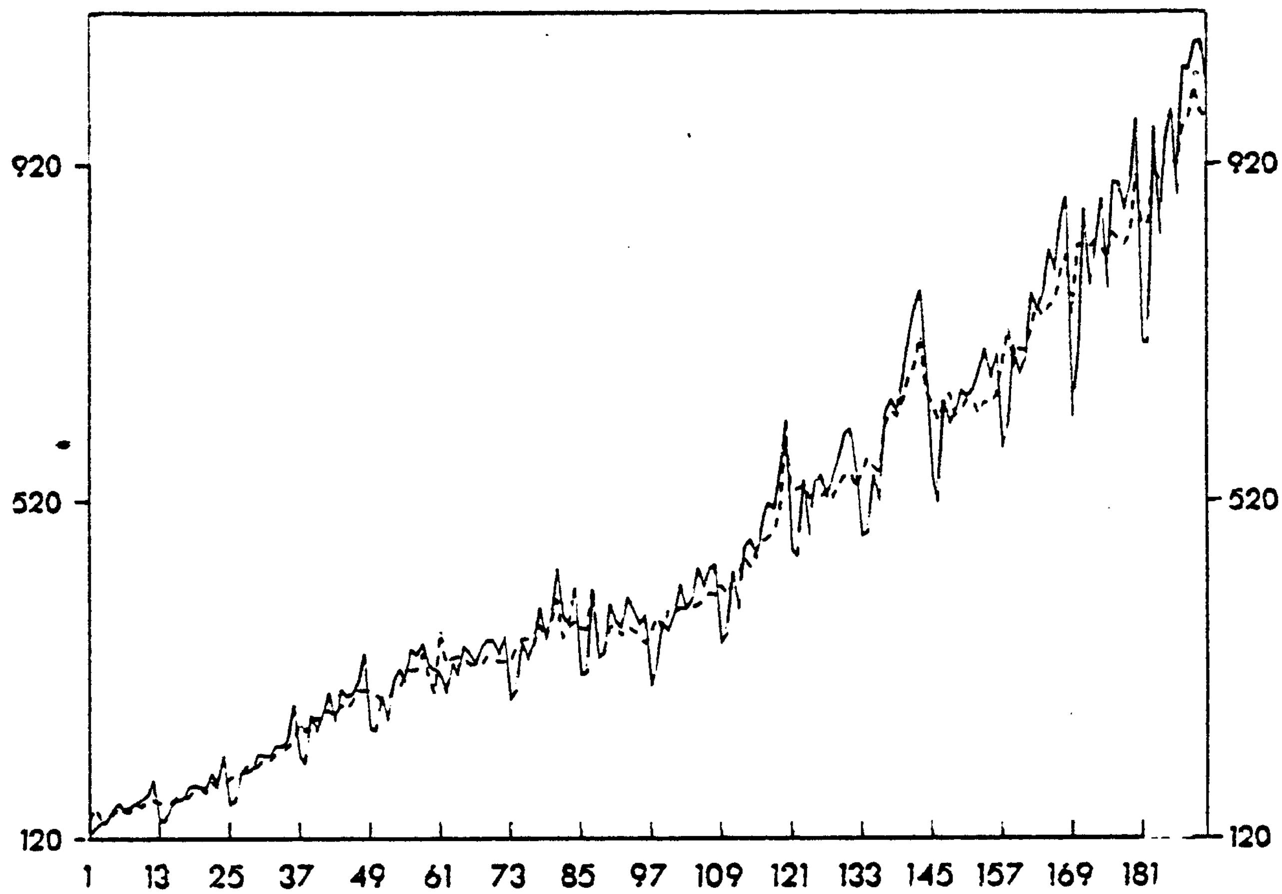
104
99
94
89
1
13
25
37
49
61
73
85
97
109
121
133

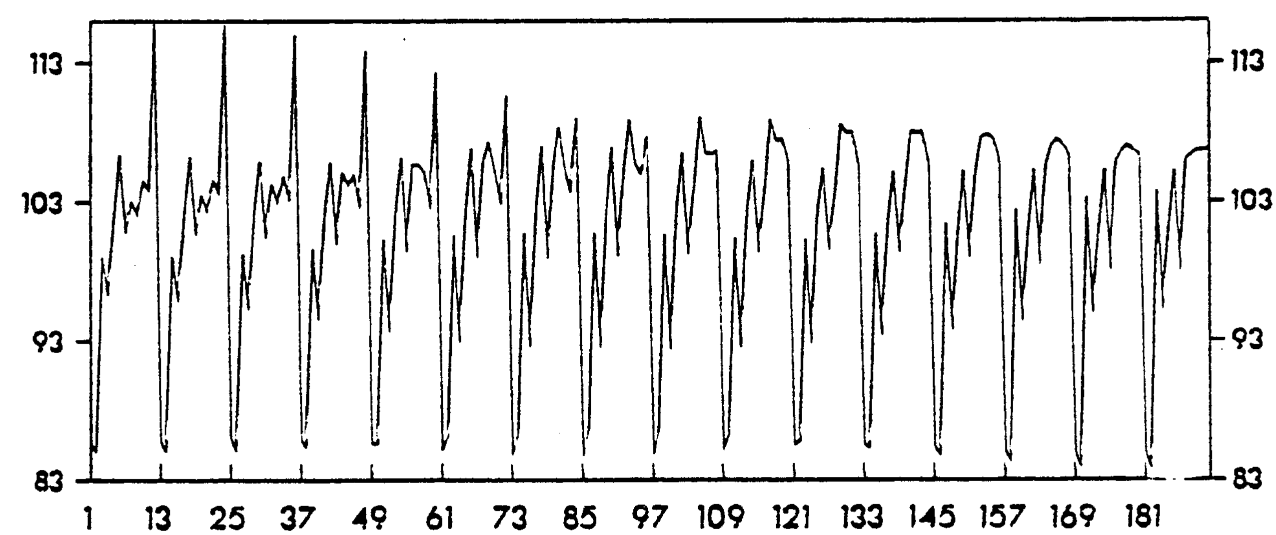
INDUSTRY S50VSU MODEL BASED



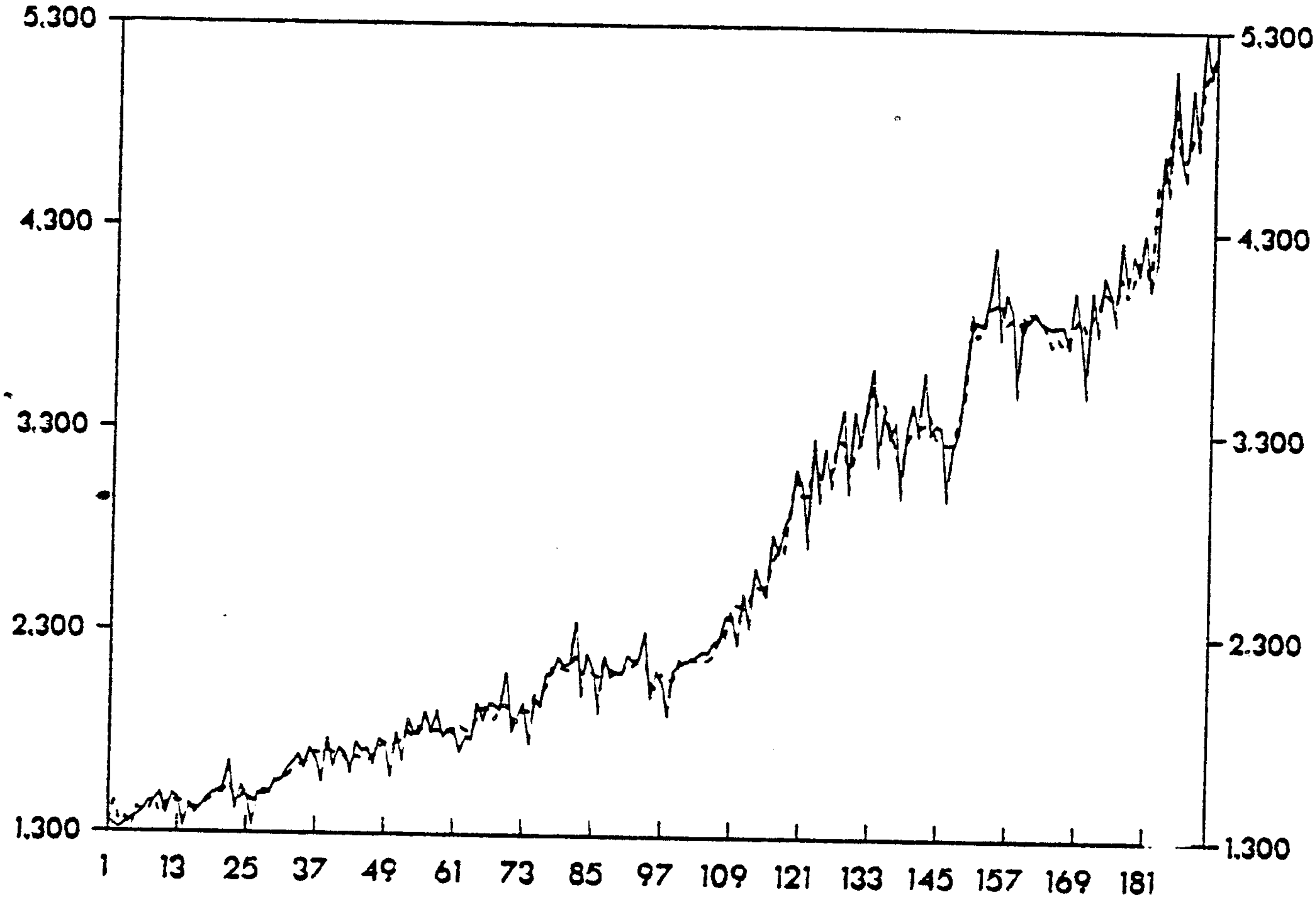


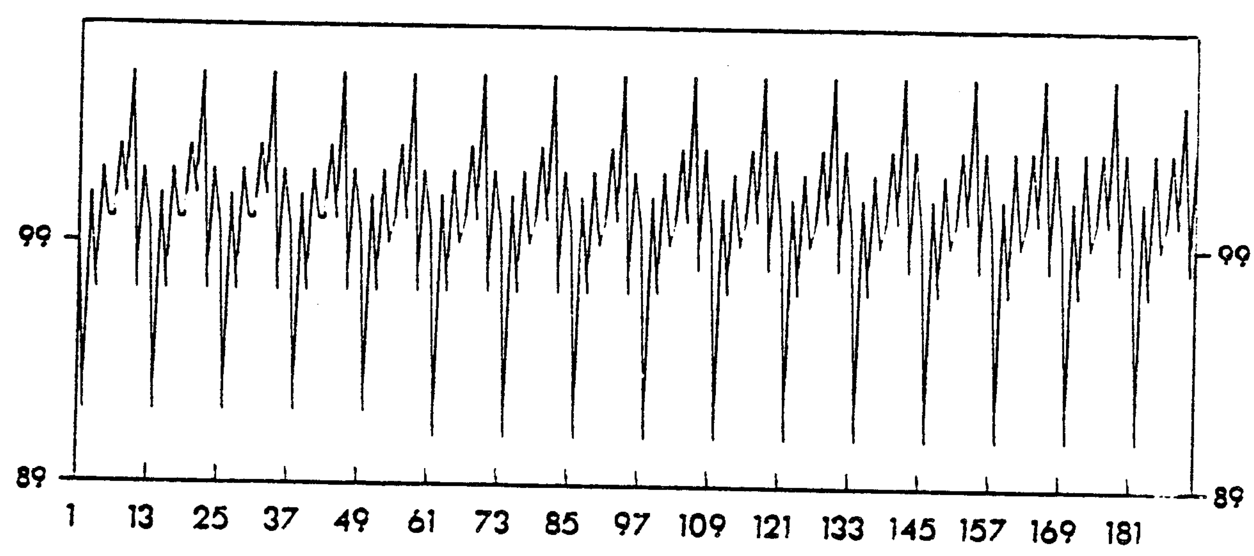
INDUSTRY S50VSU X-11 LOG



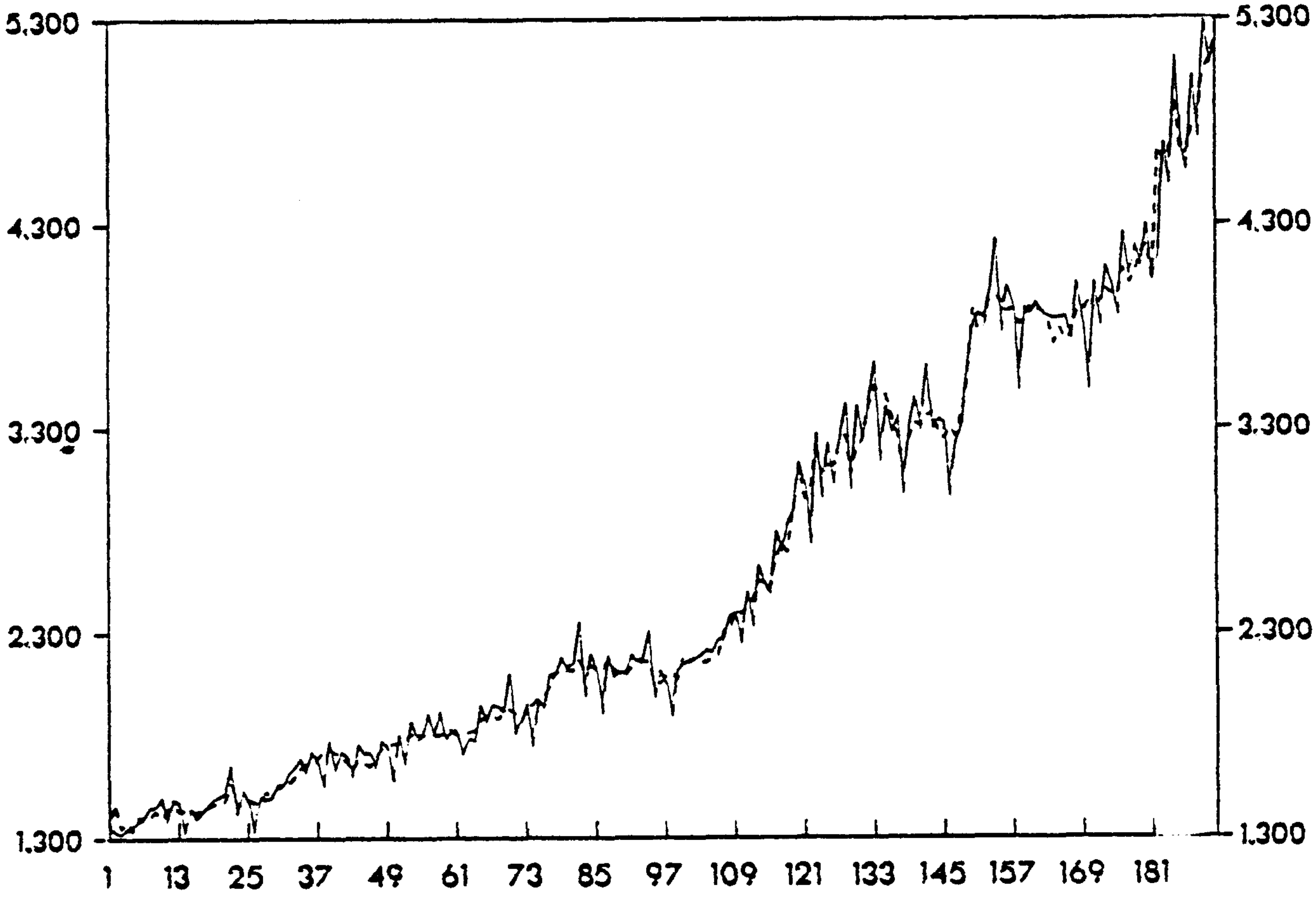


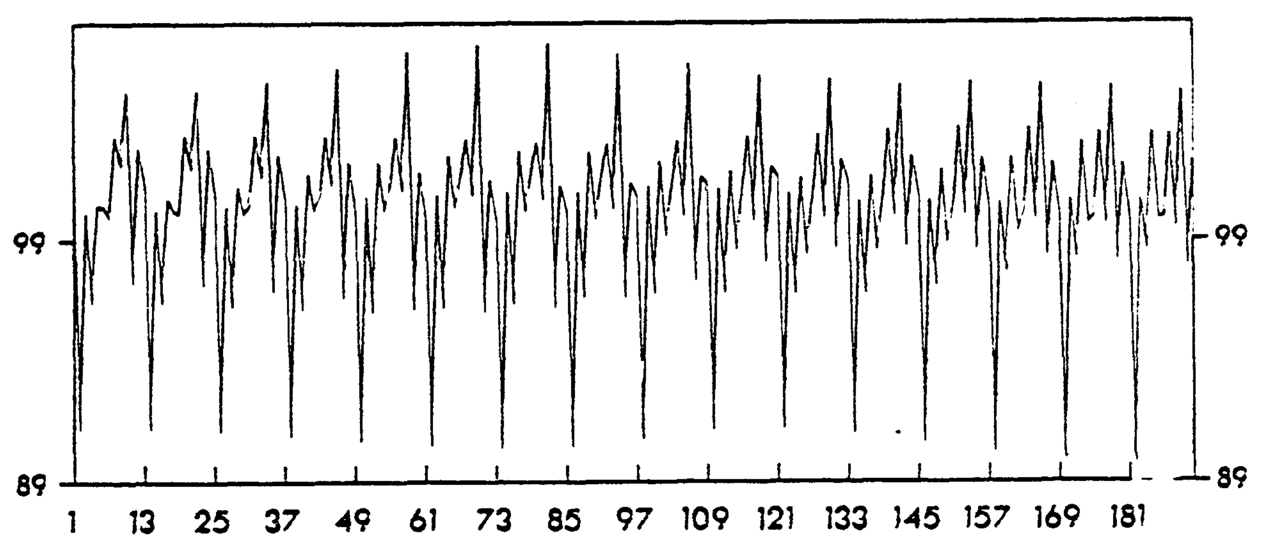
INDUSTRY S60VSU MODEL BASED



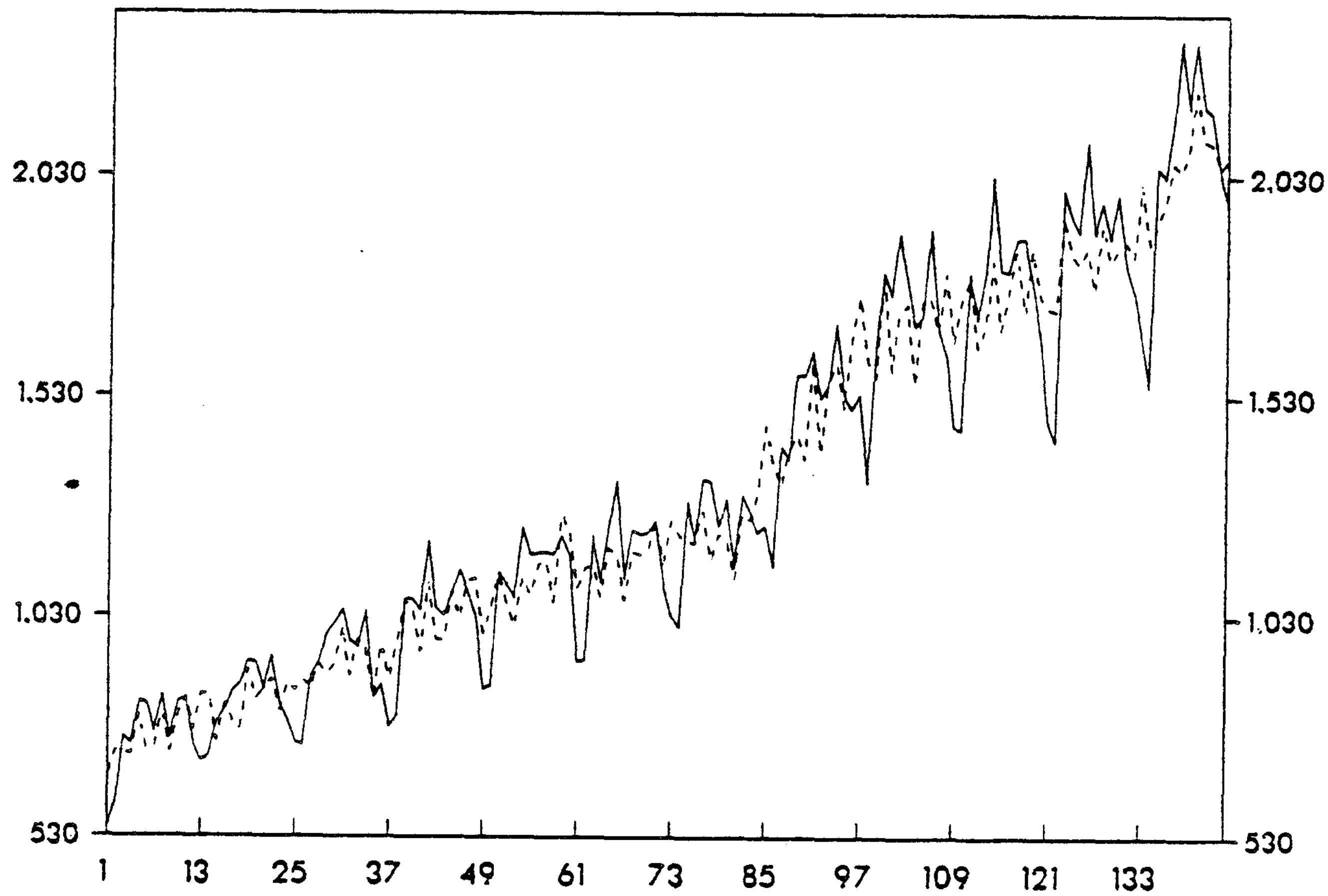


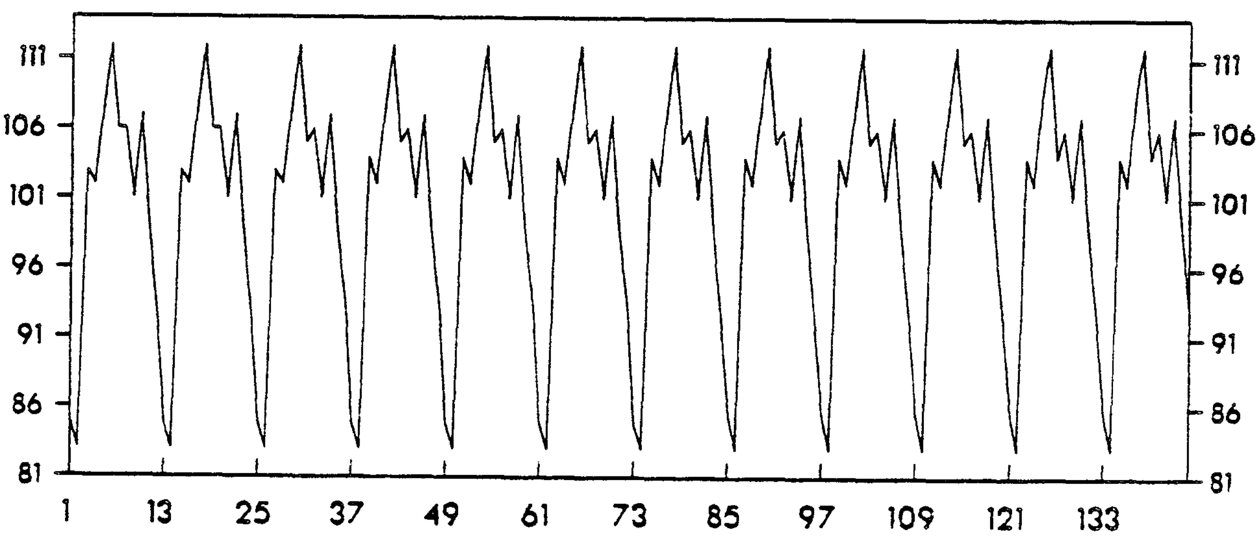
INDUSTRY S60VSU X-11 LOG



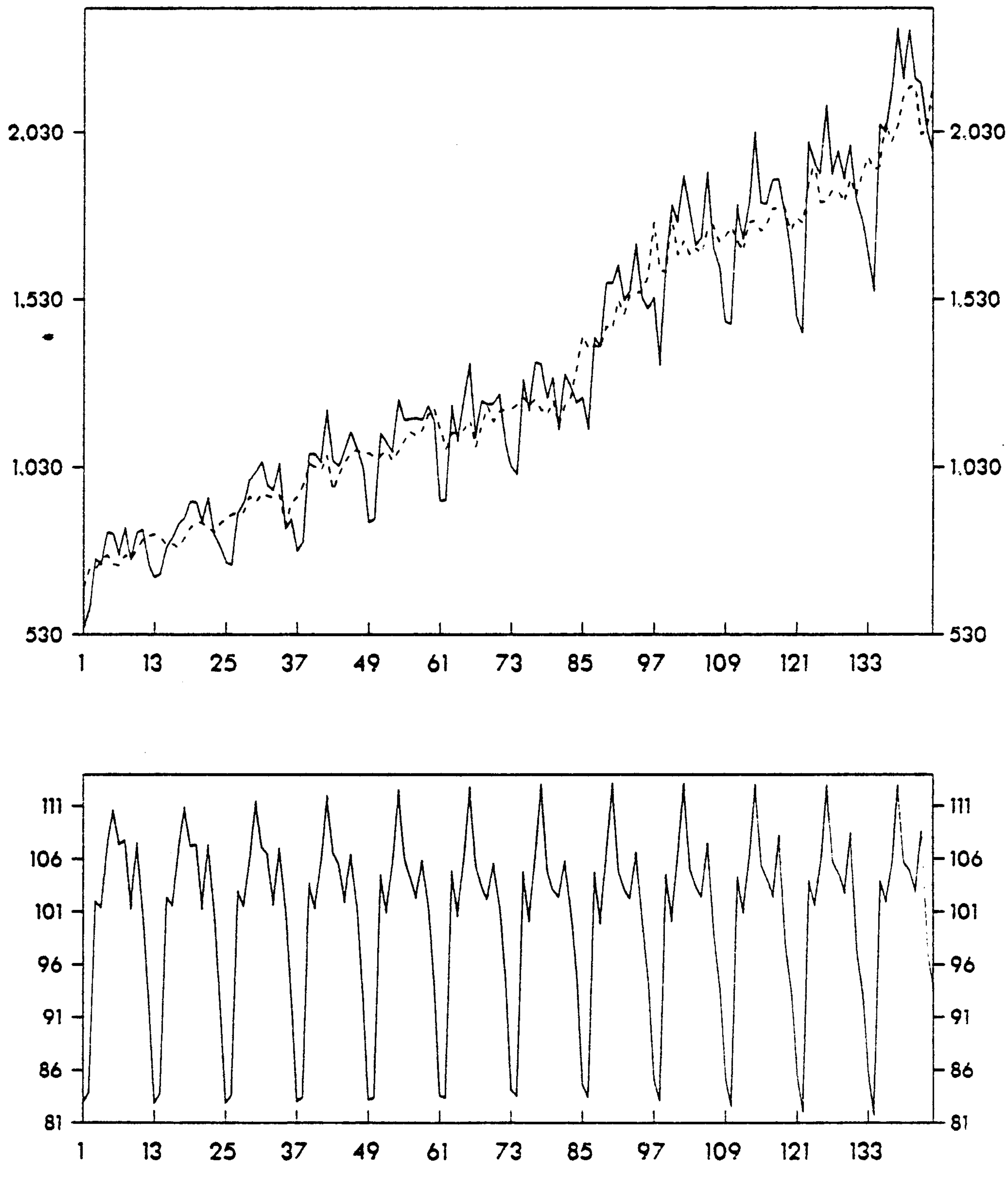


S62VSU MODEL BASED

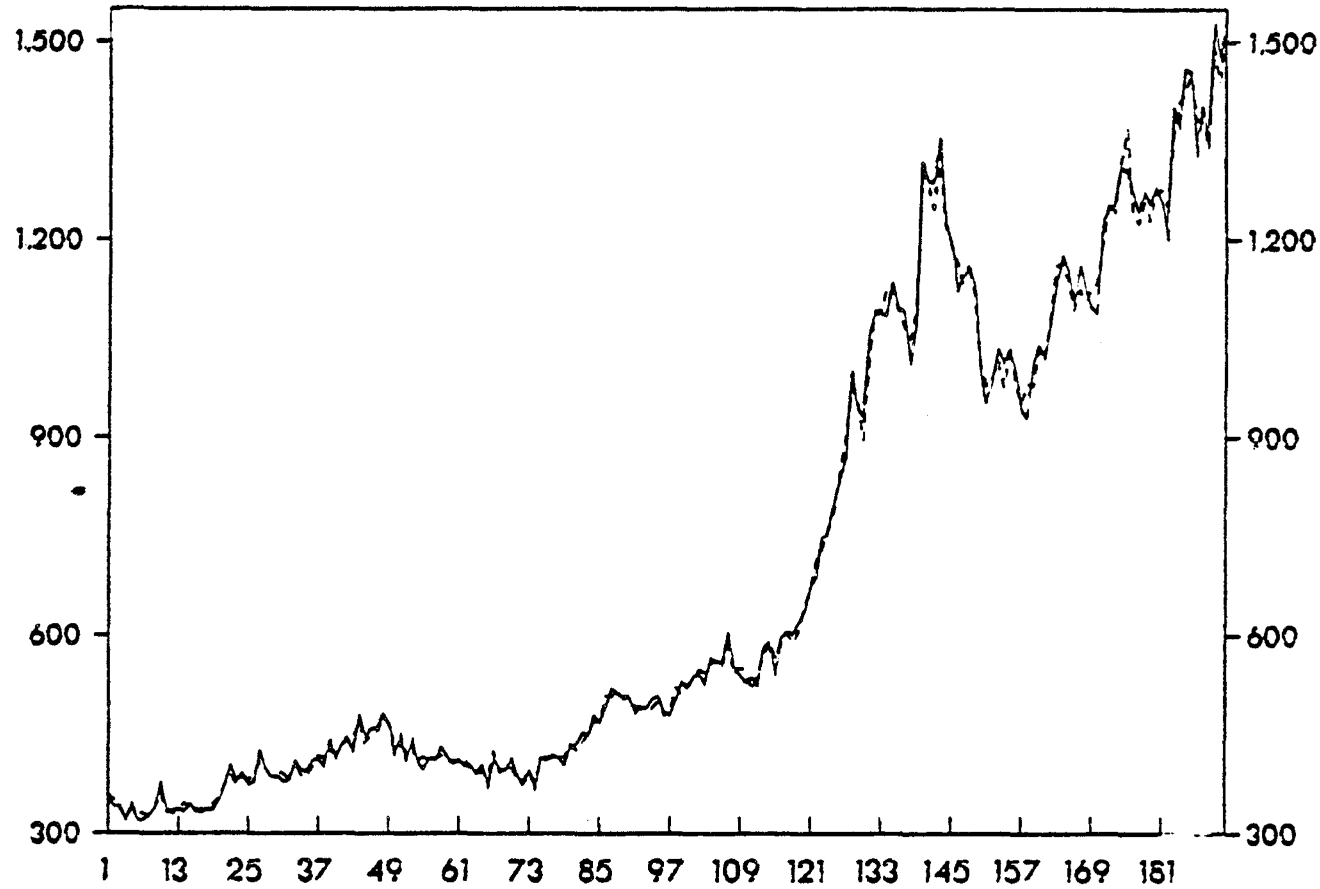


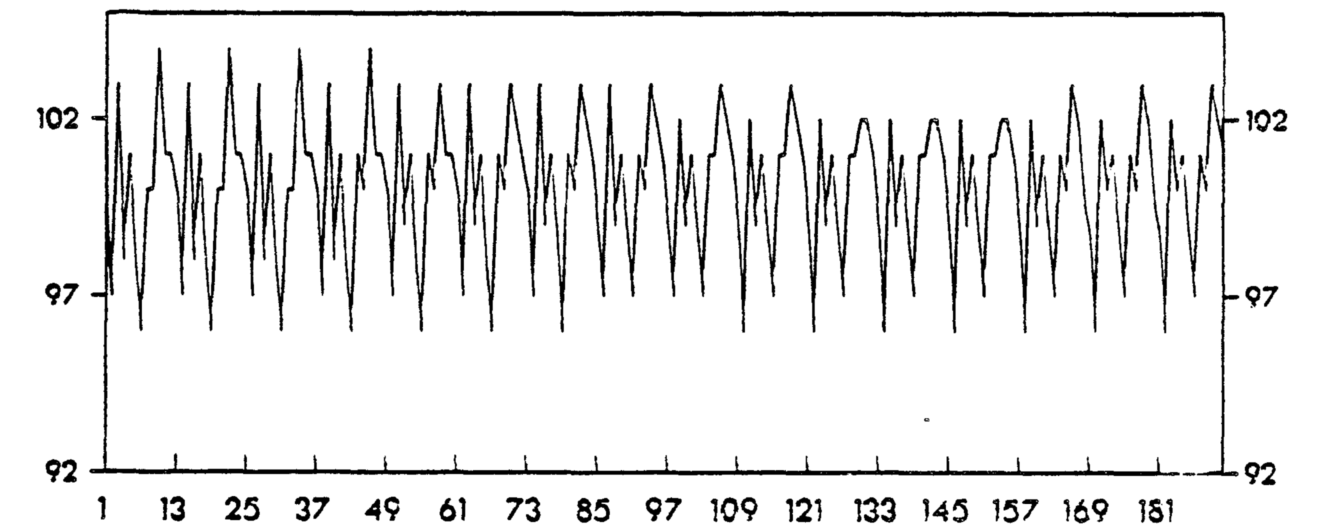


S62VSU X-11

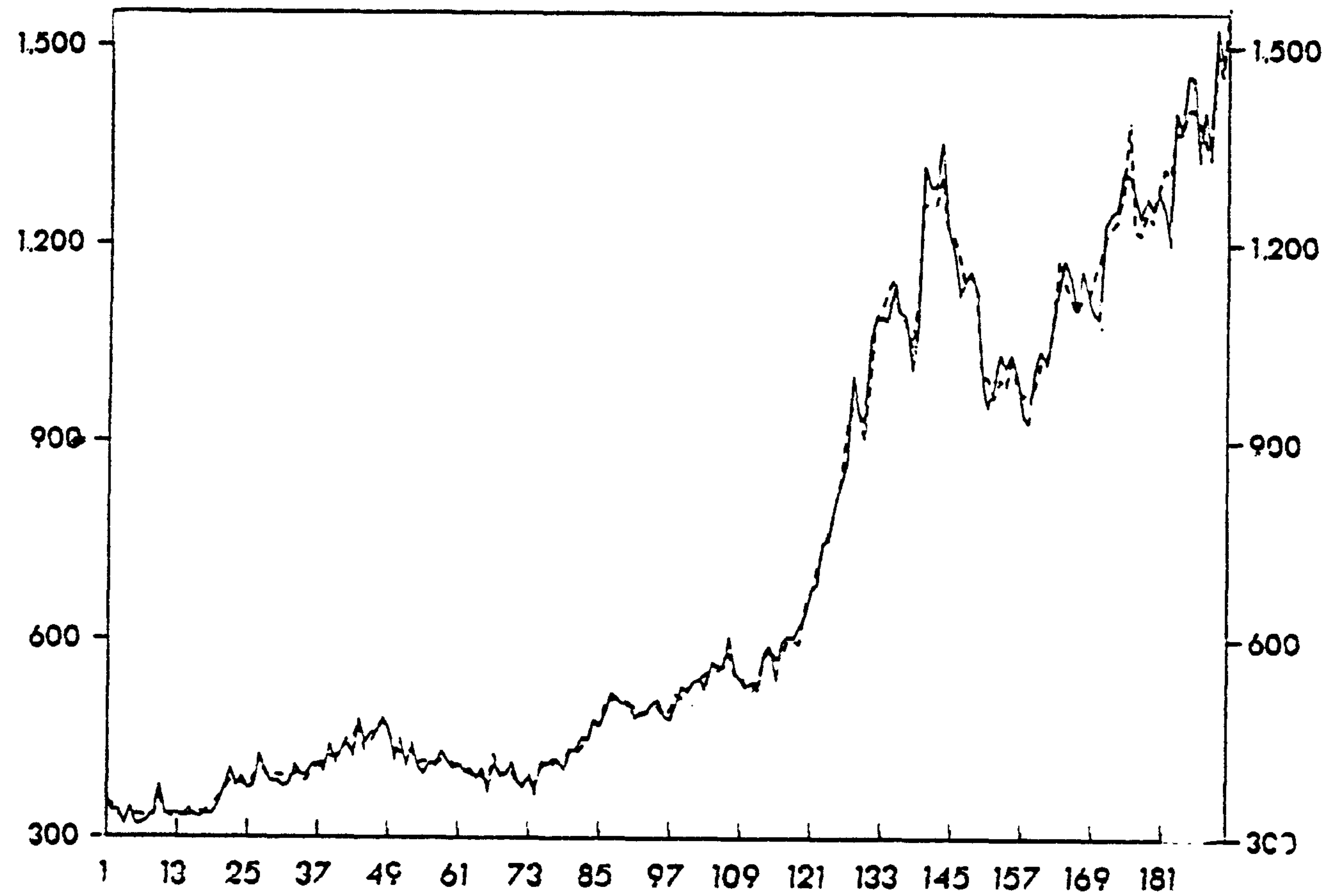

2,030
1,530
1,030
530
1
13
25
37
49
61
73
85
97
109
121
133
111
106
101
96
91
86
81

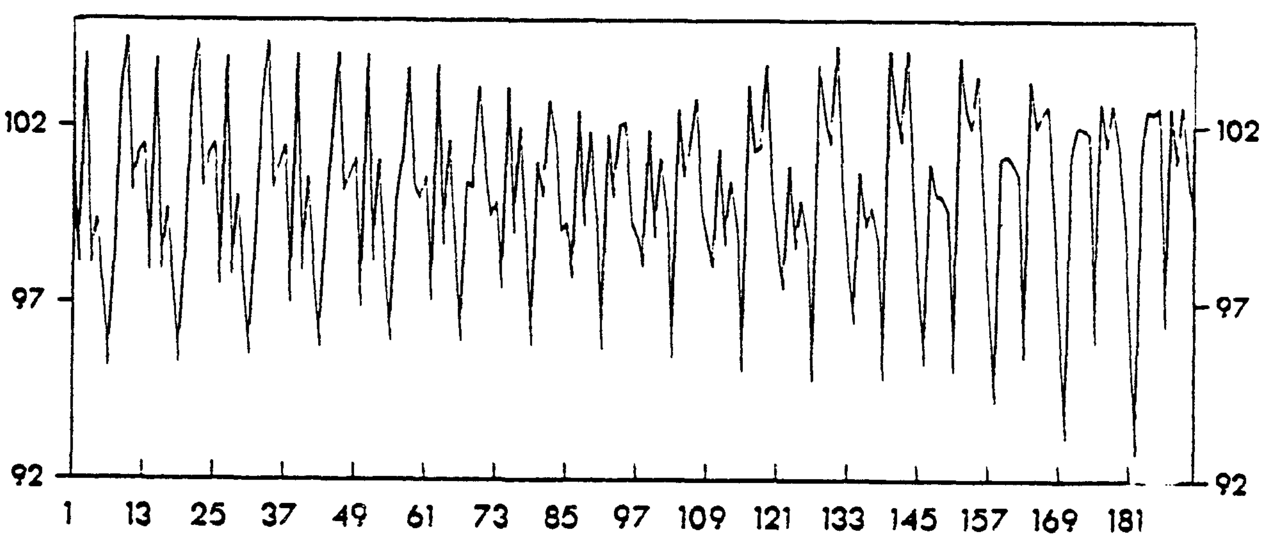
INDUSTRY S63YSU MODEL BASED



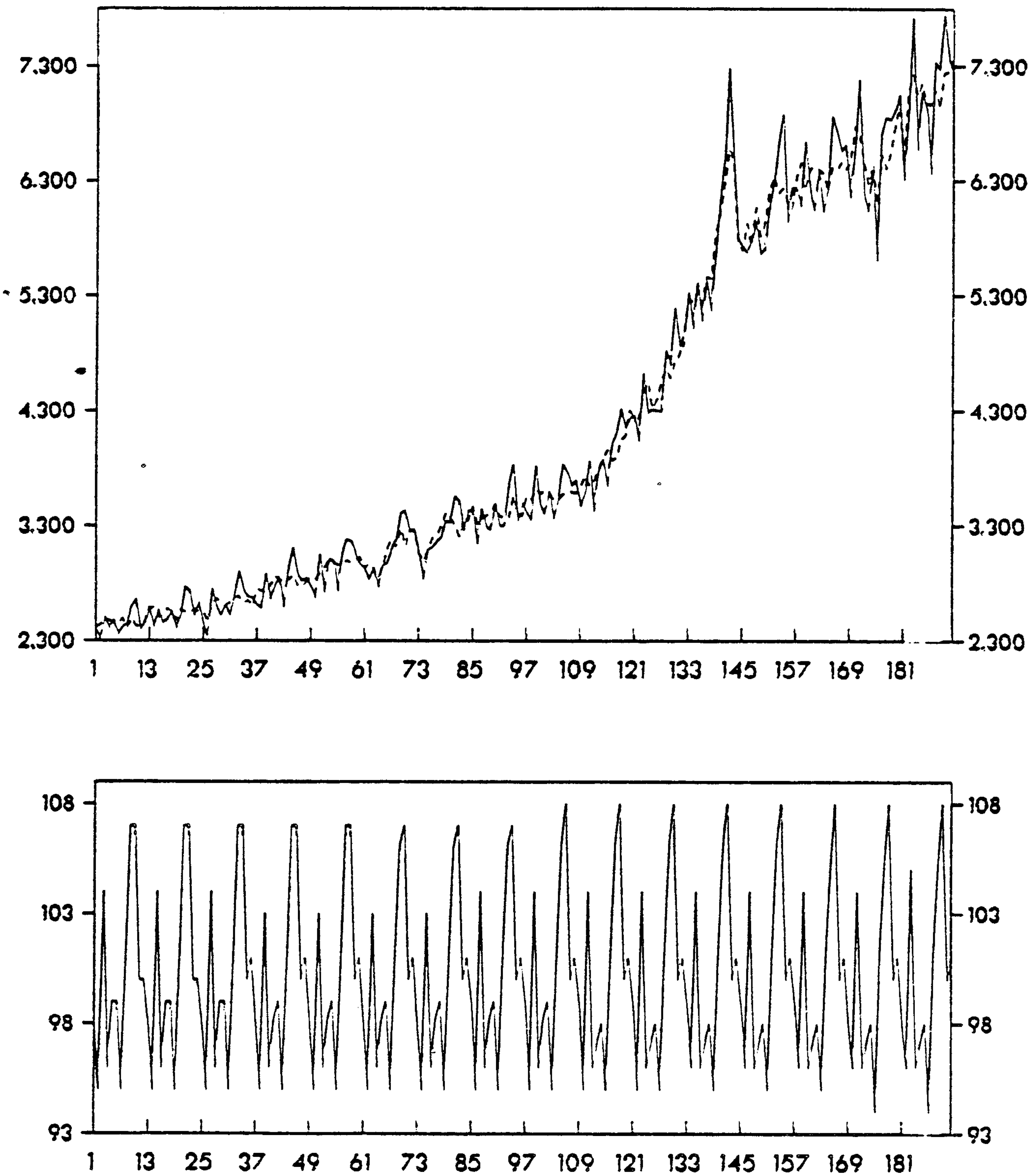


INDUSTRY S63VSU X-11 LOG

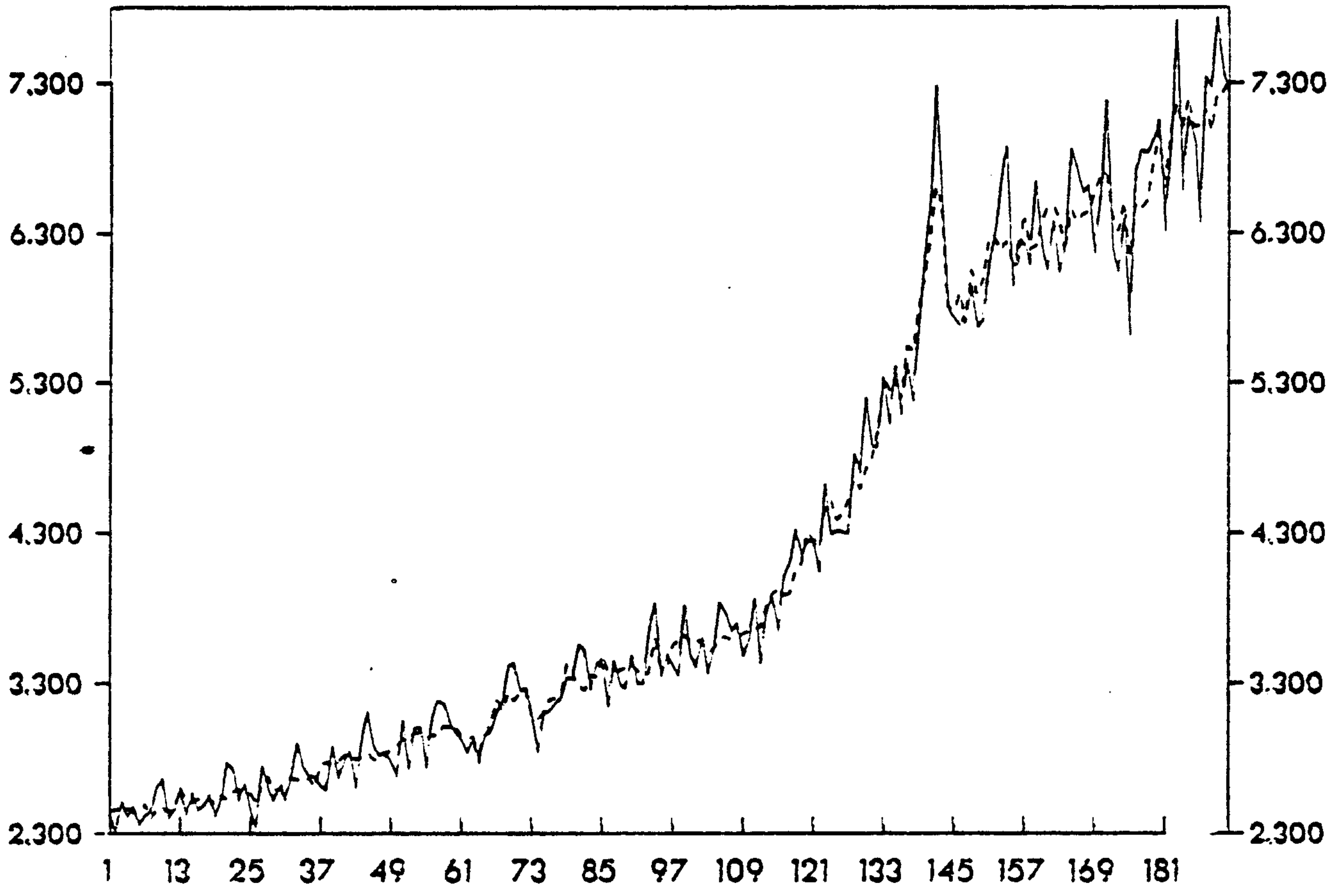


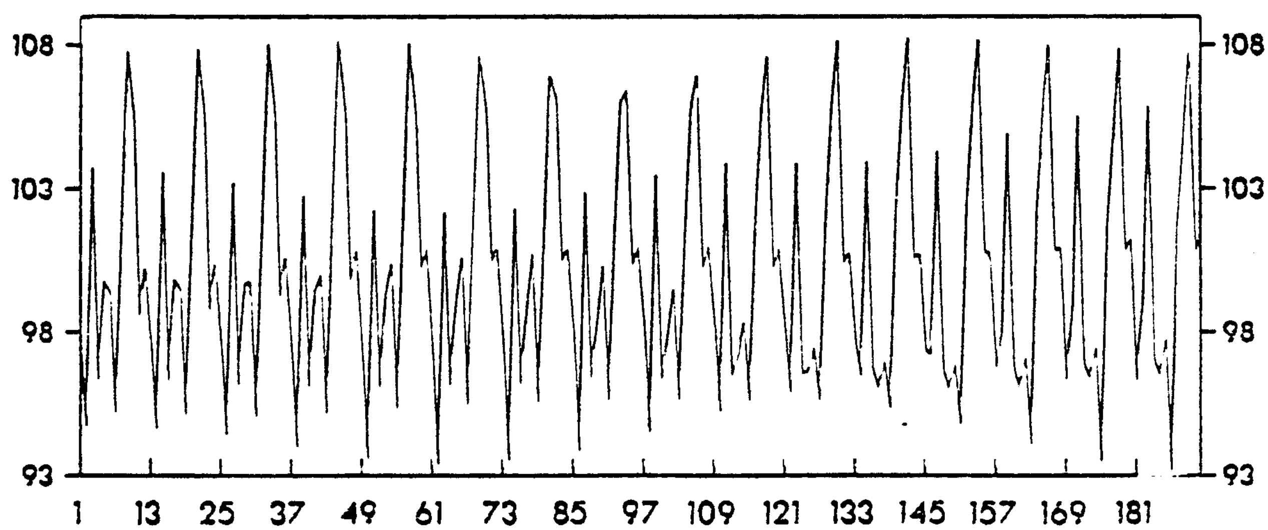


INDUSTRY S64VSU MODEL BASED

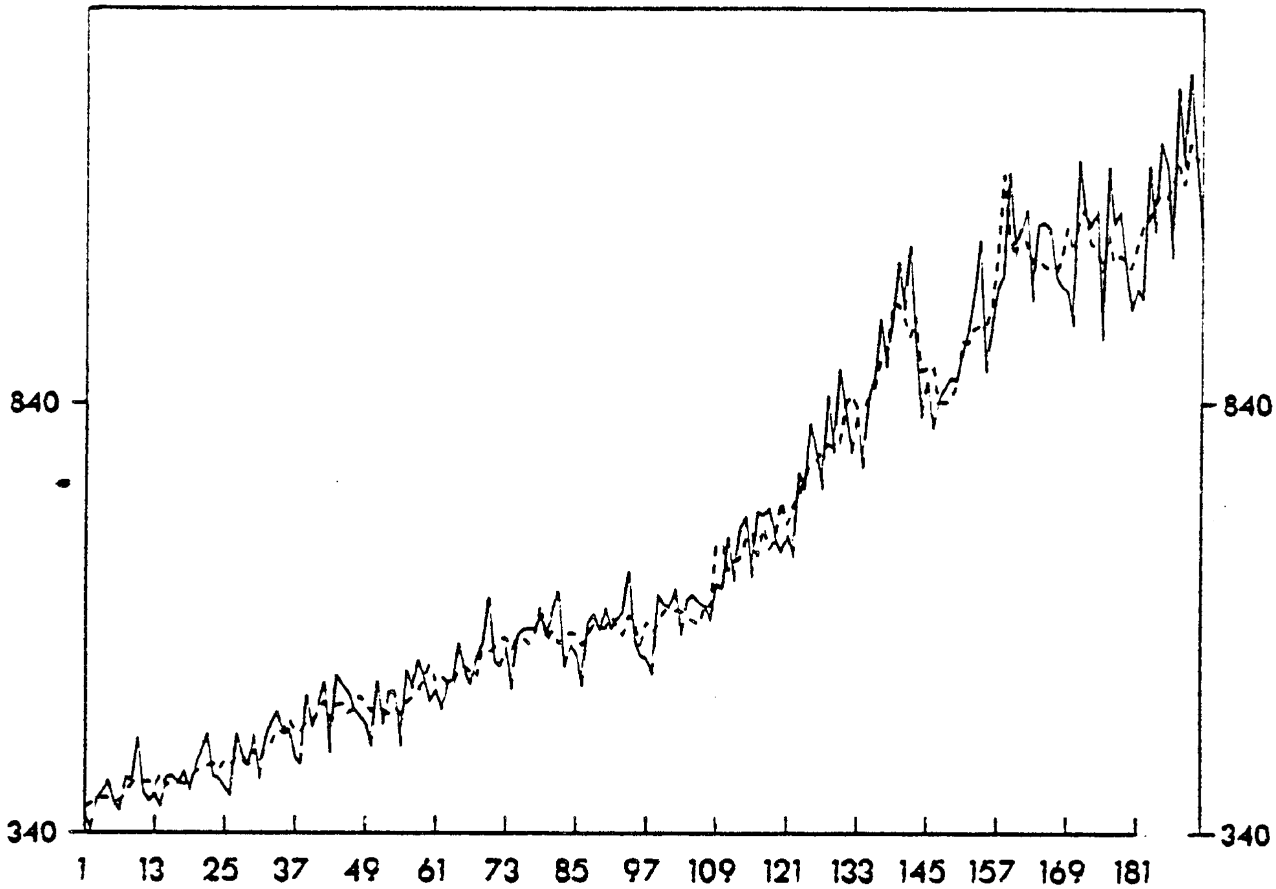


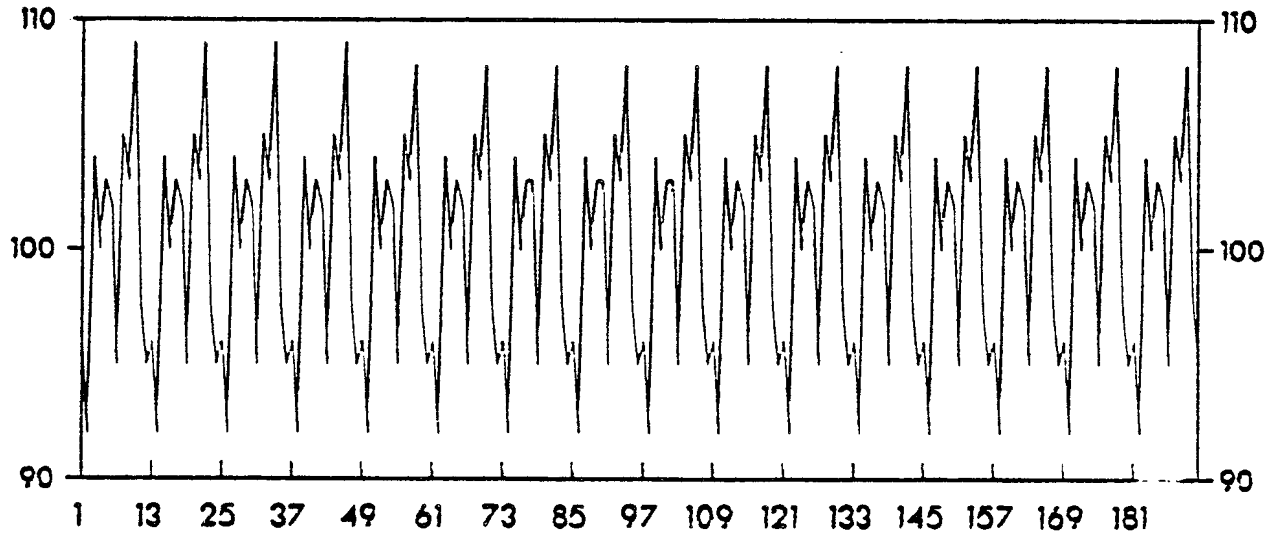
INDUSTRY S64VSU X-11 LOG



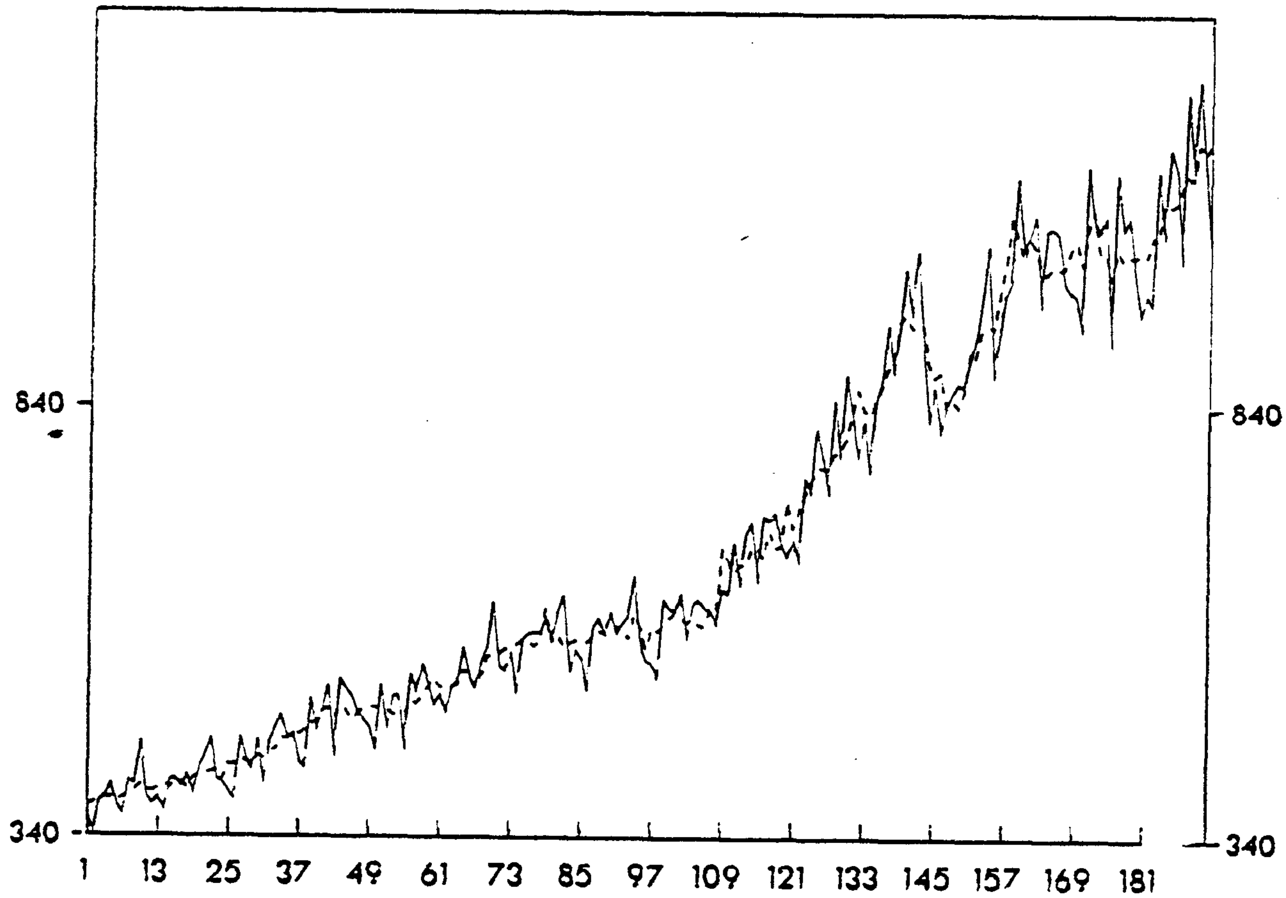


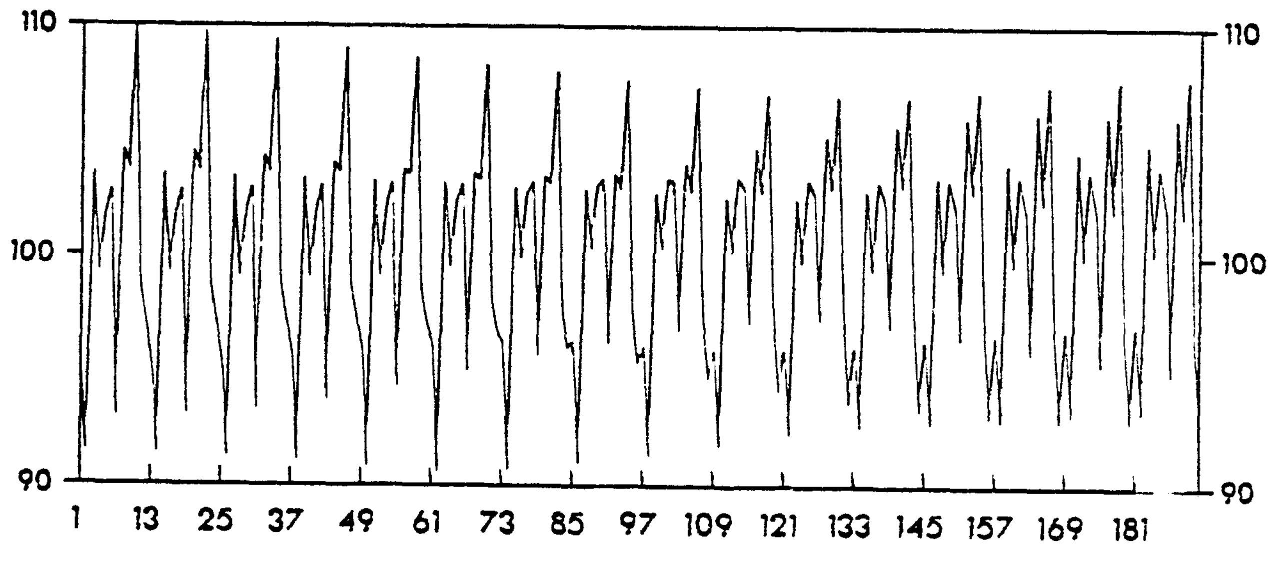
INDUSTRY S76VSU MODEL BASED



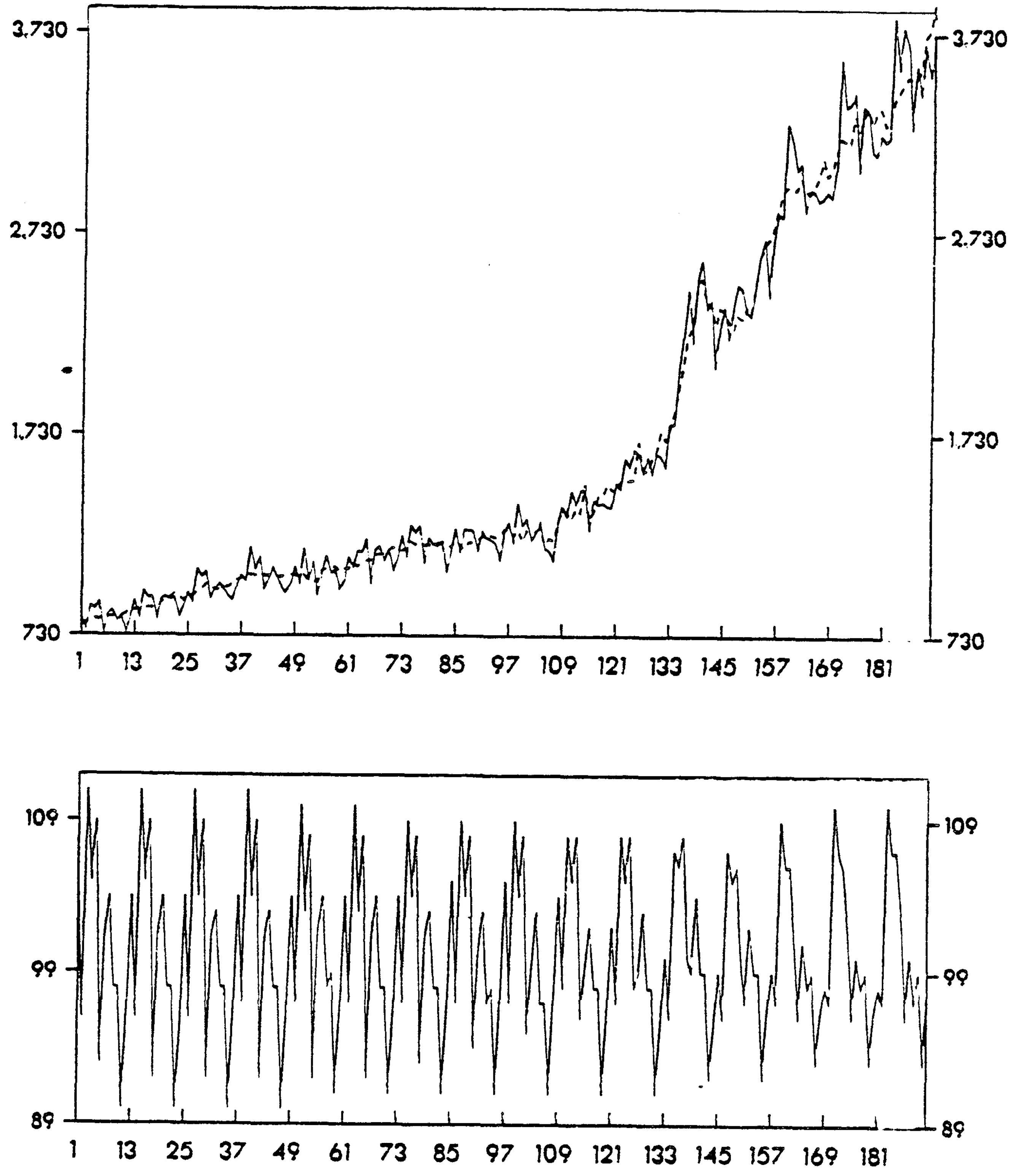


INDUSTRY S76VSU X-11 LOG

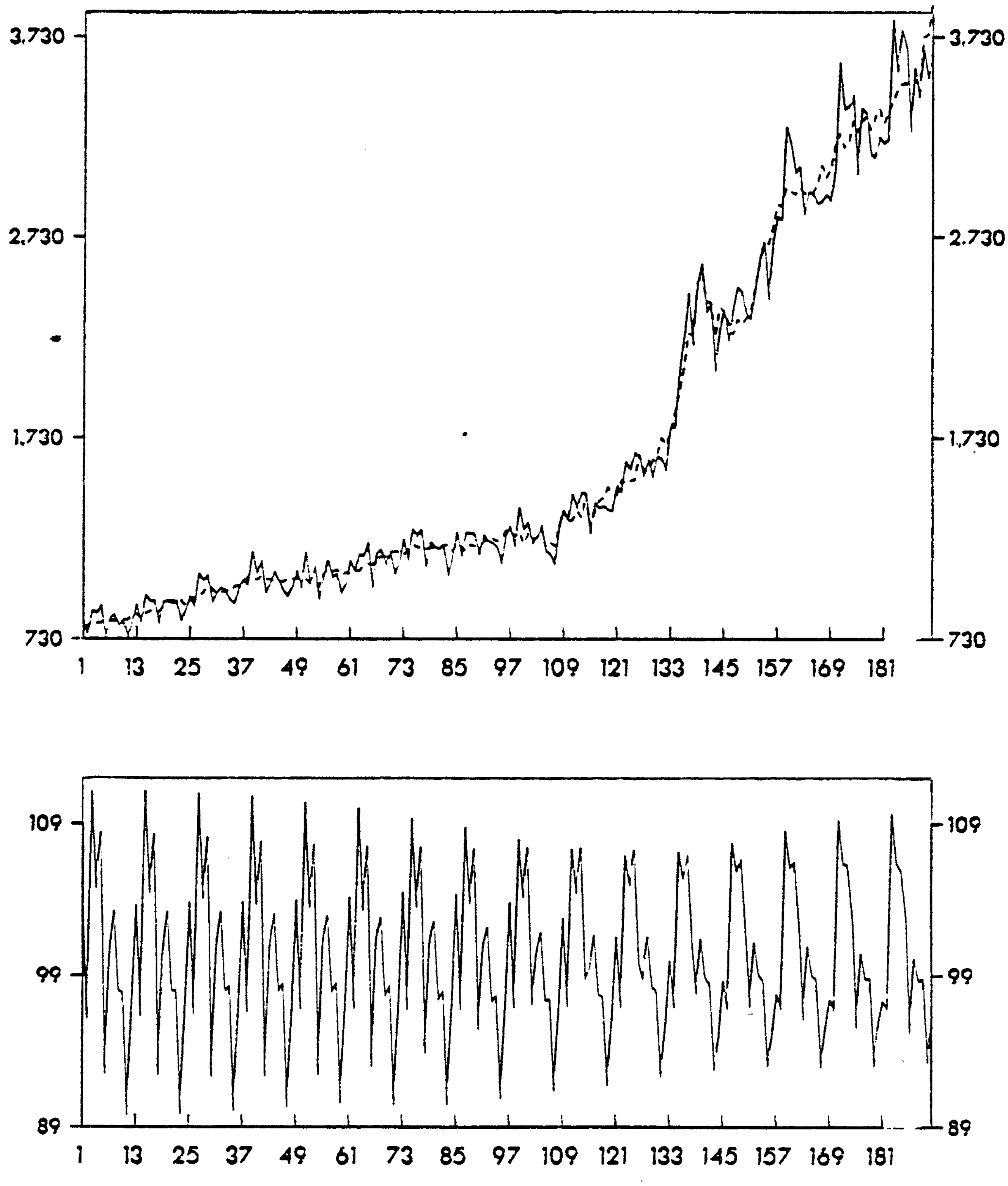




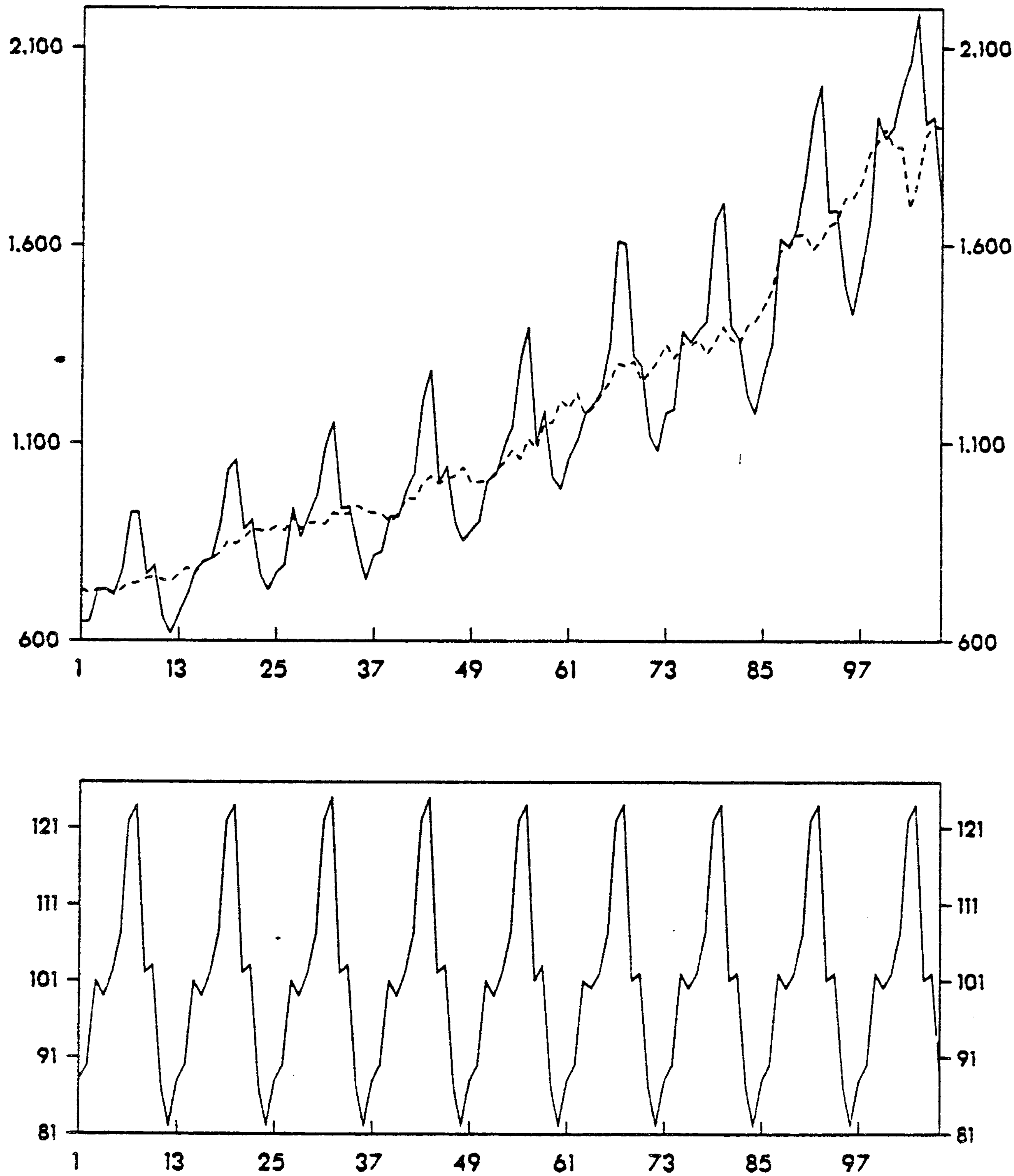
INDUSTRY S83VSU MODEL BASED



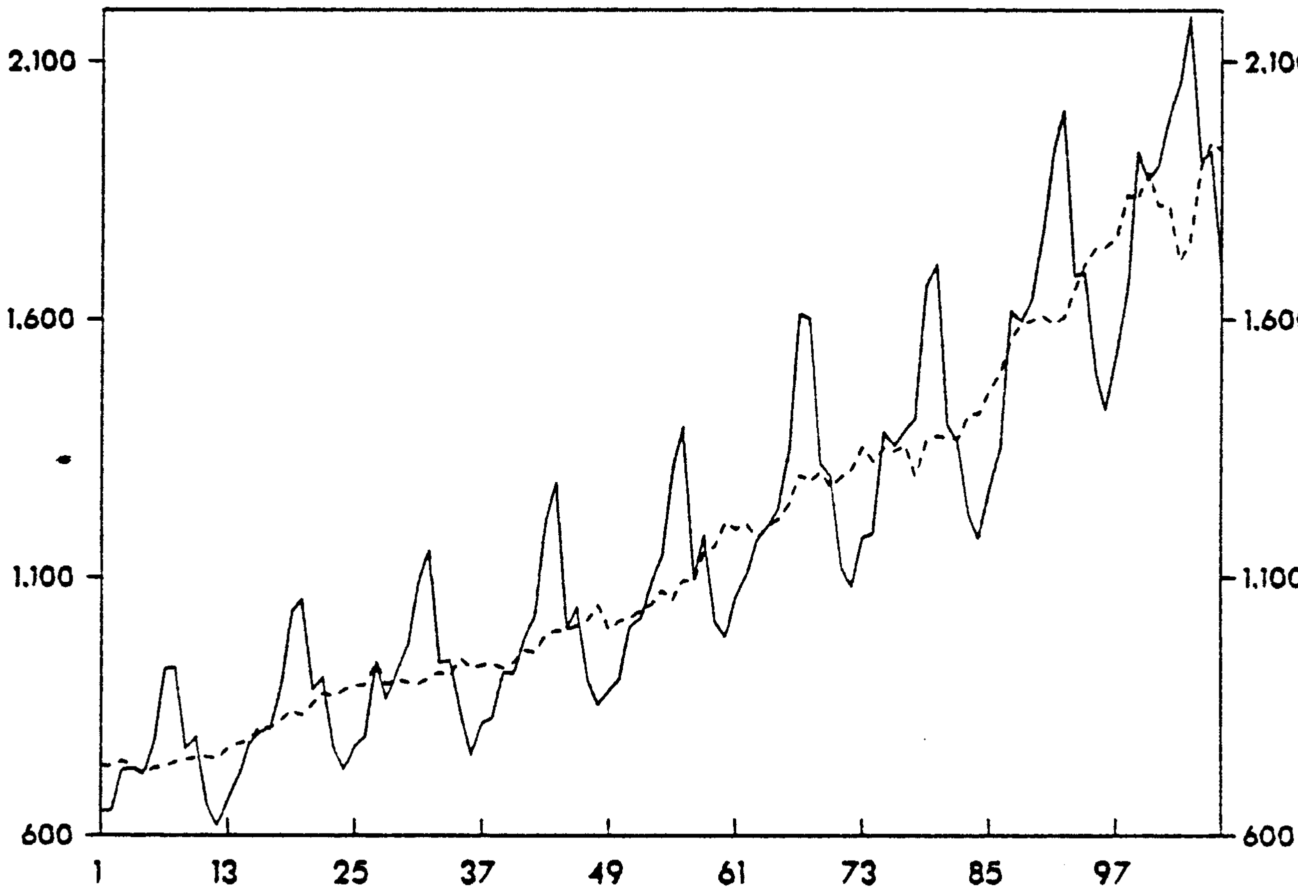
INDUSTRY S83VSU X-11 LOG

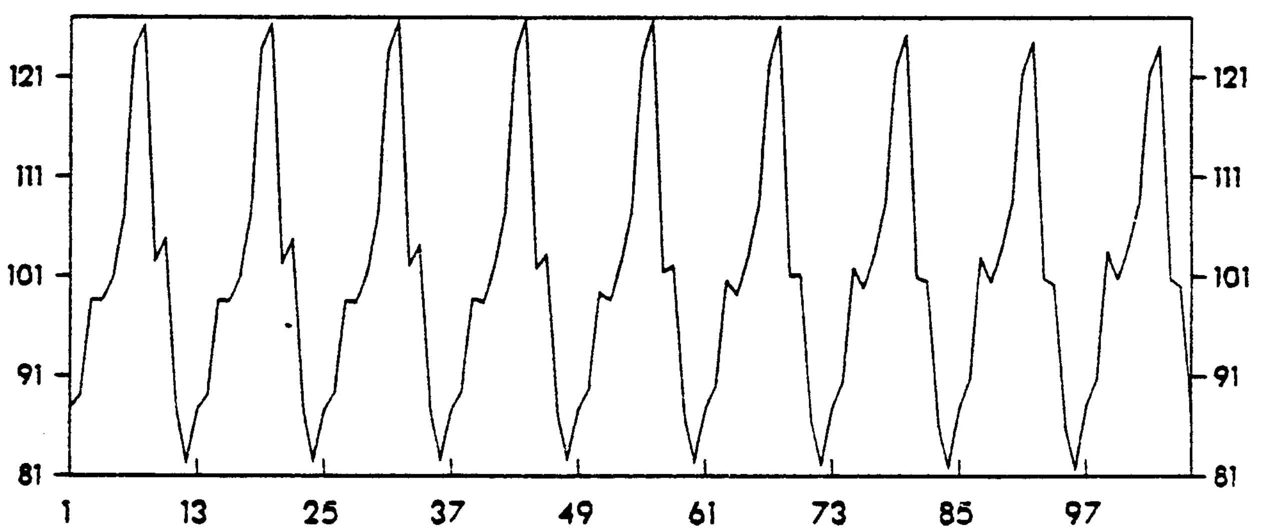


RETAIL 701000 MODEL BASED

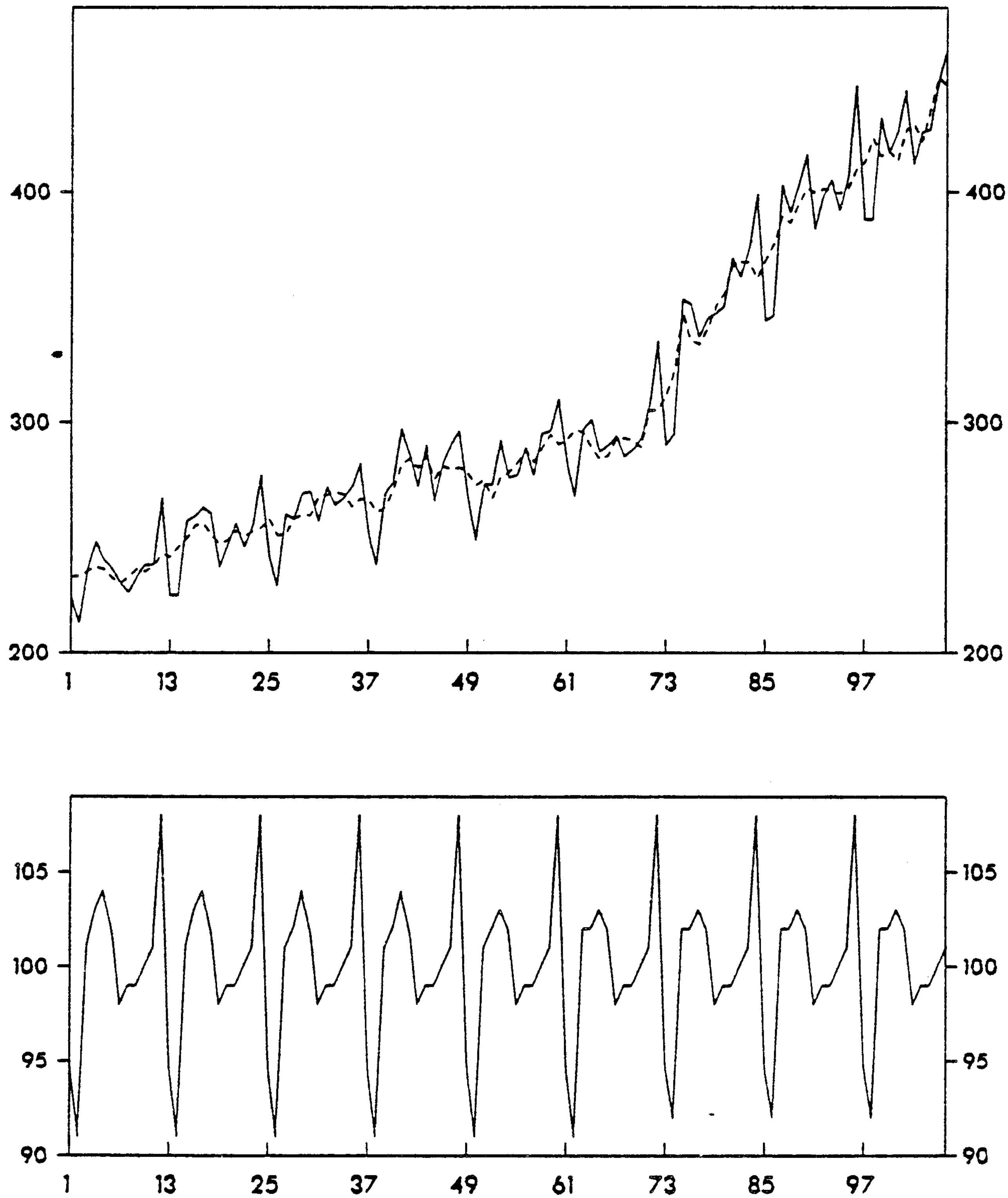


RETAIL 701000 X-11

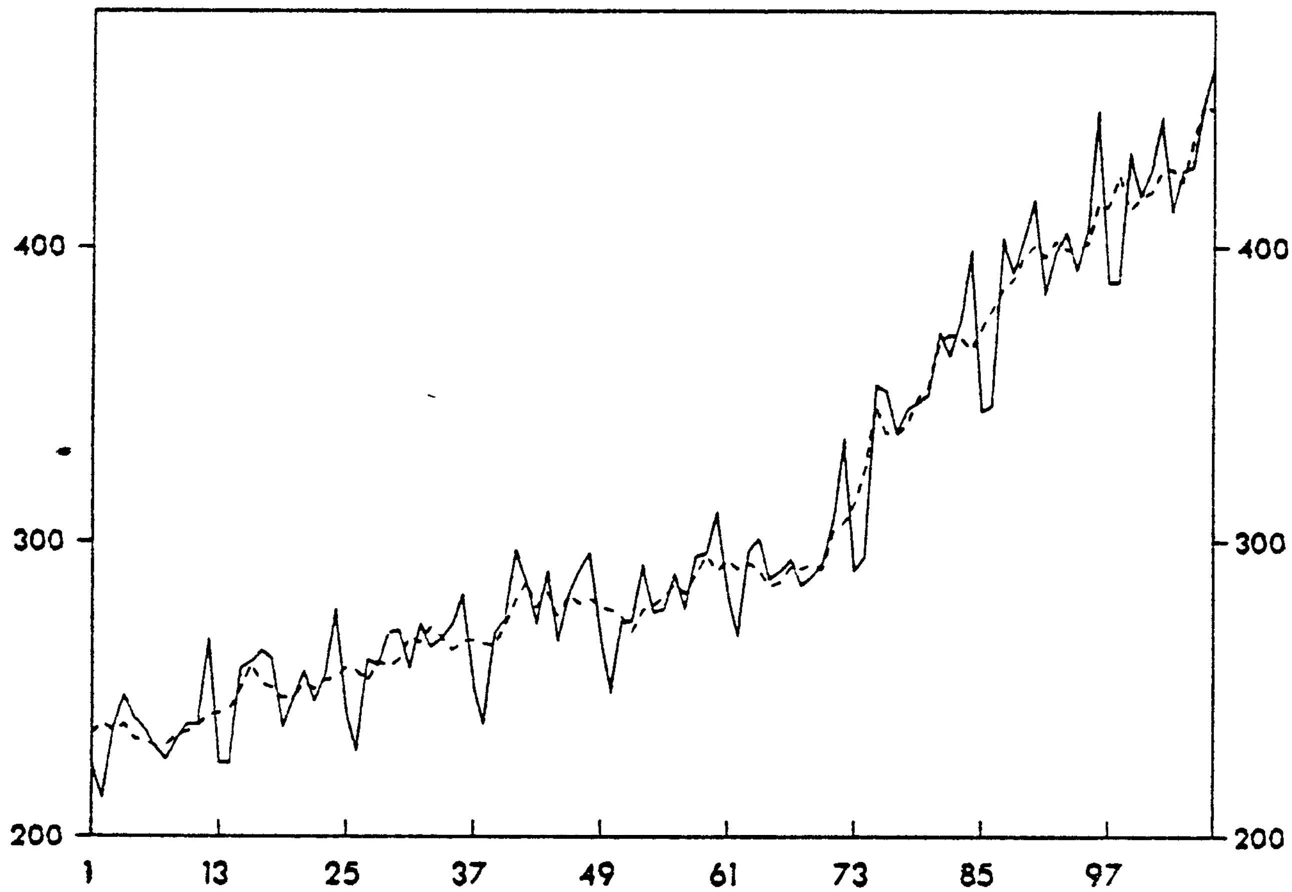


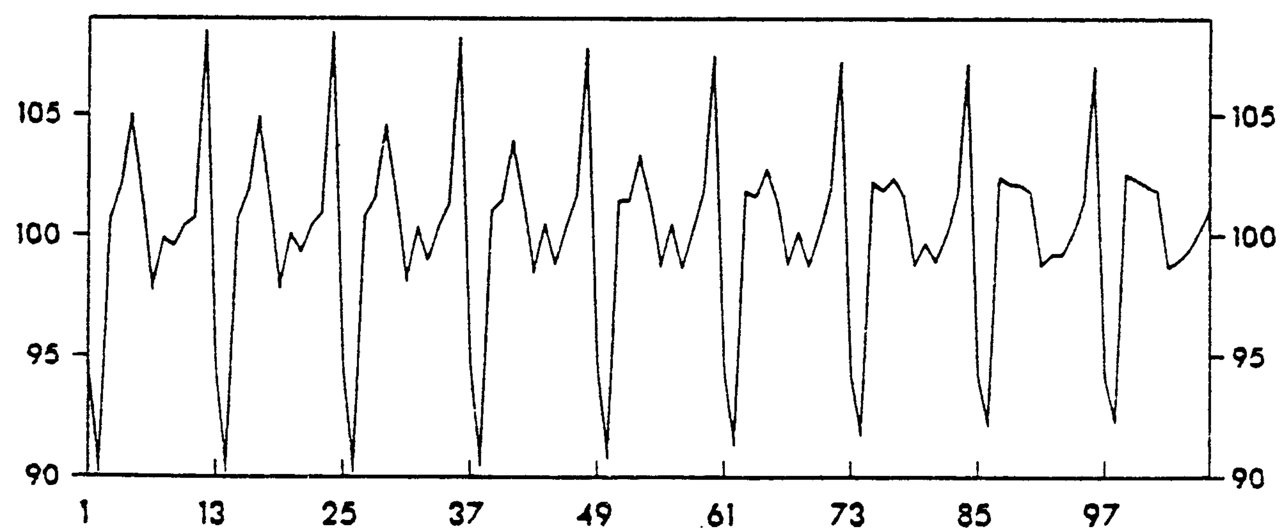


RETAIL 723000 MODEL BASED

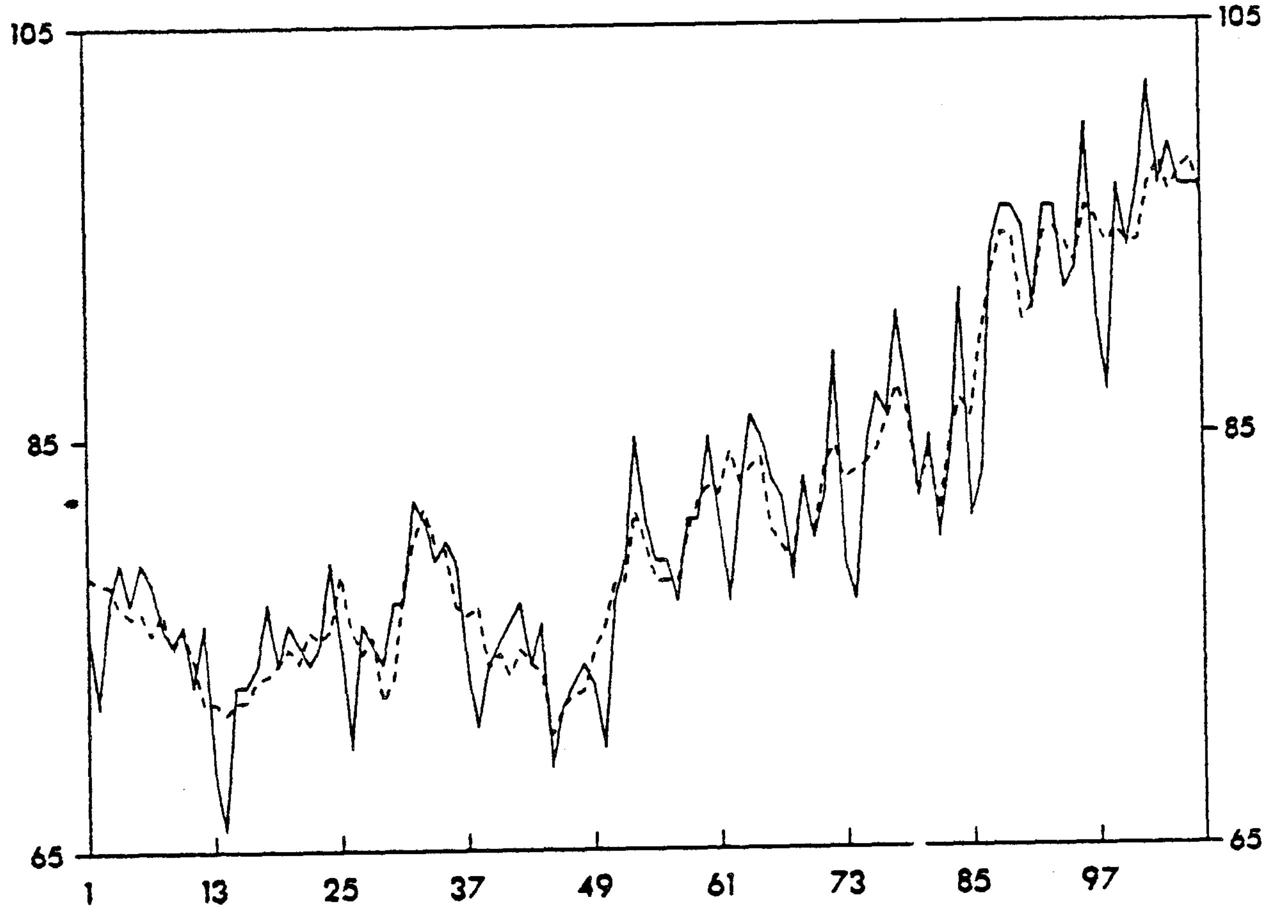


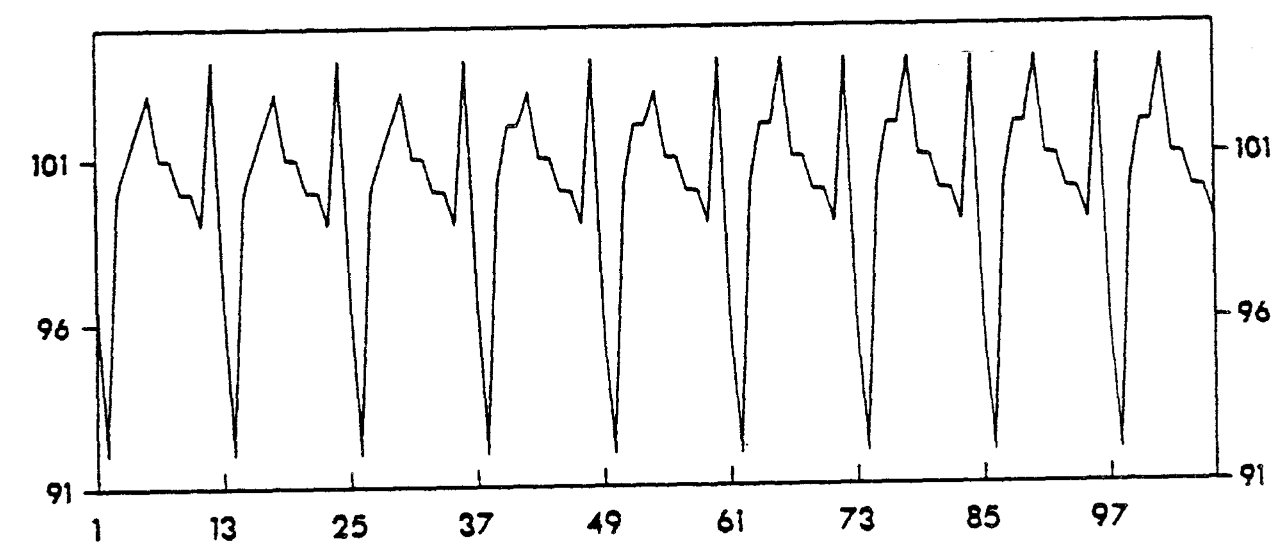
RETAIL 723000 X-11



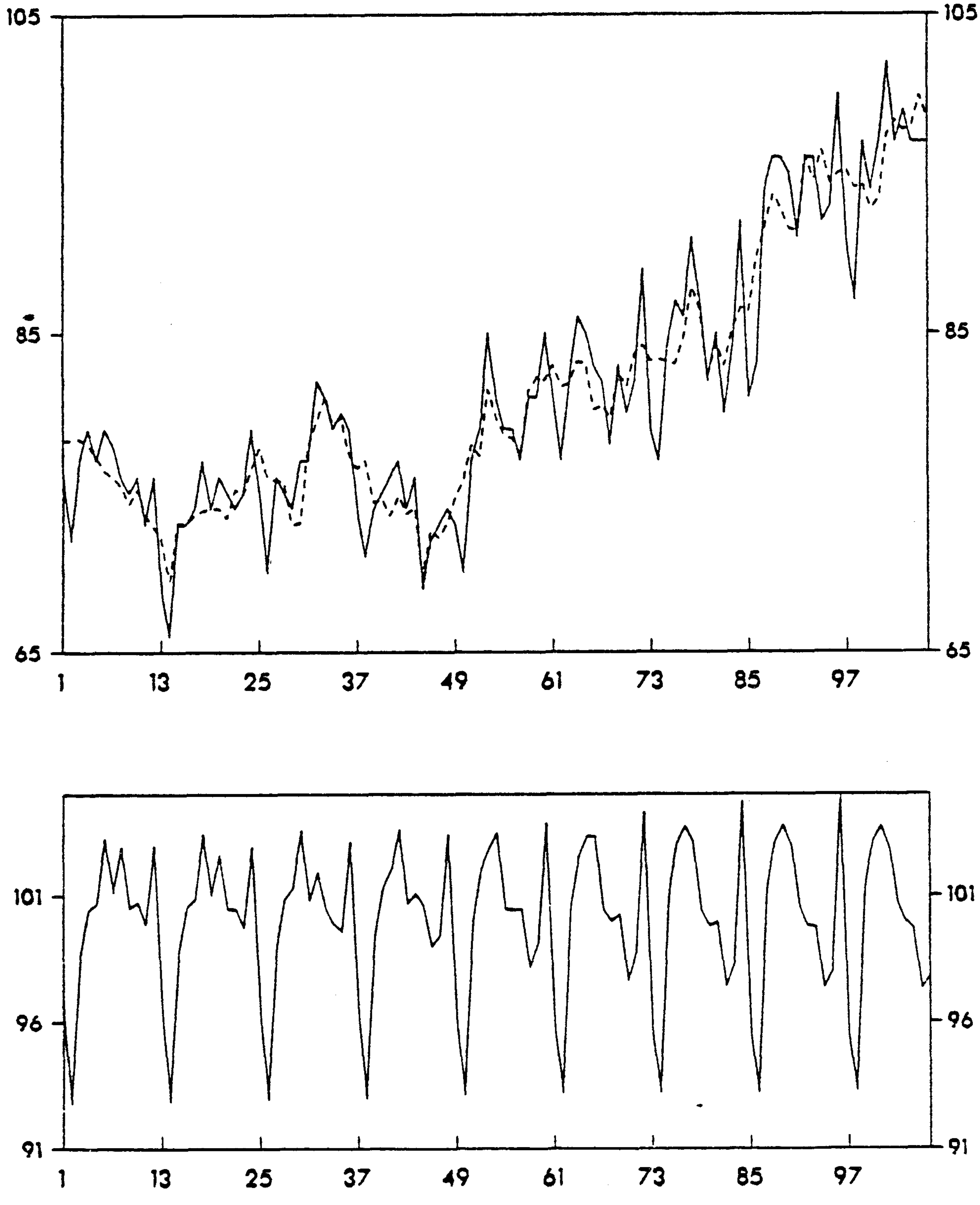


RETAIL 724000 MODEL BASED

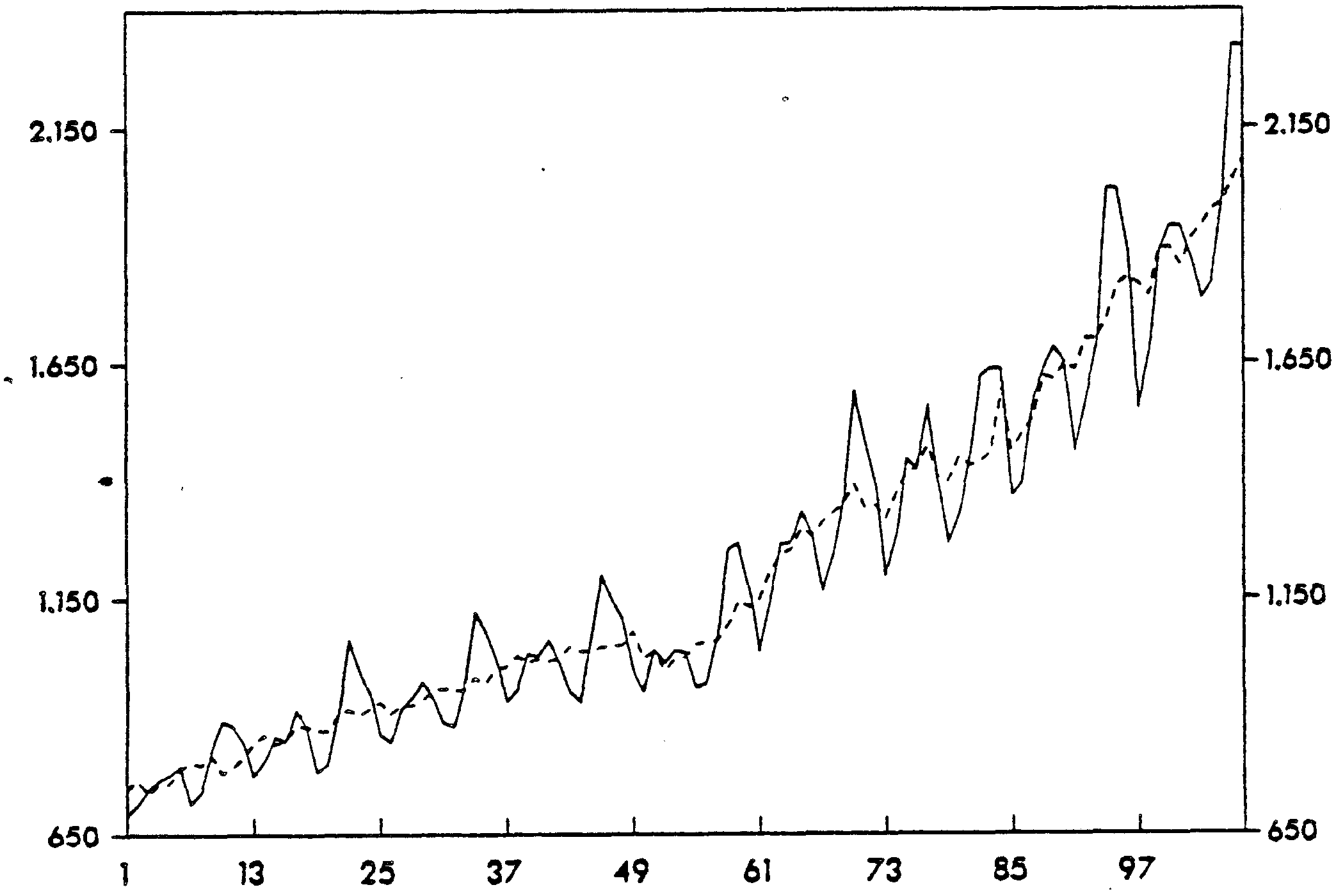


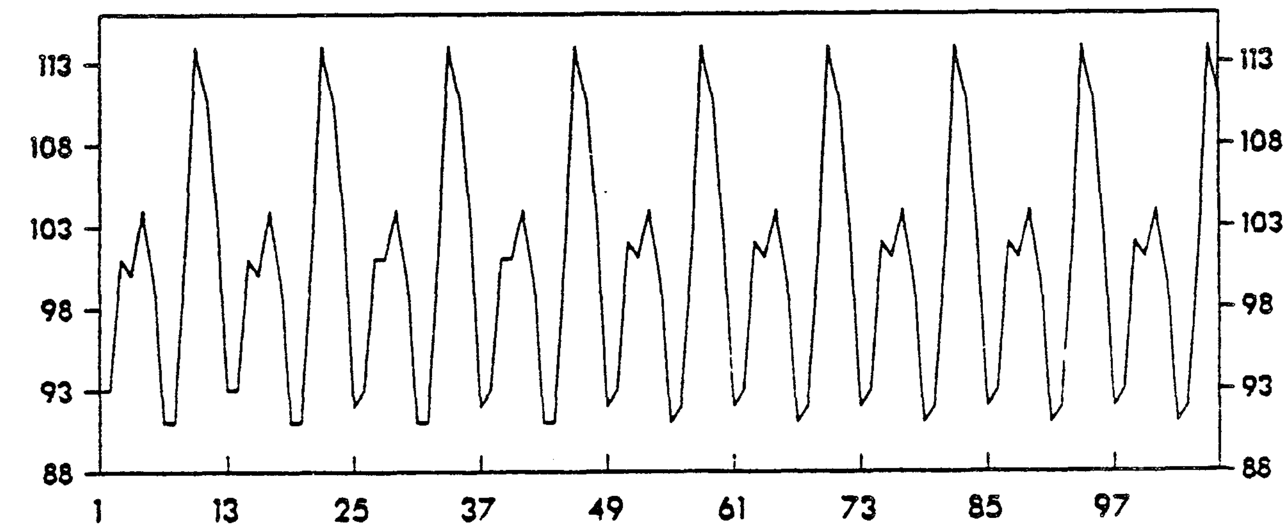


RETAIL 724000 X-11

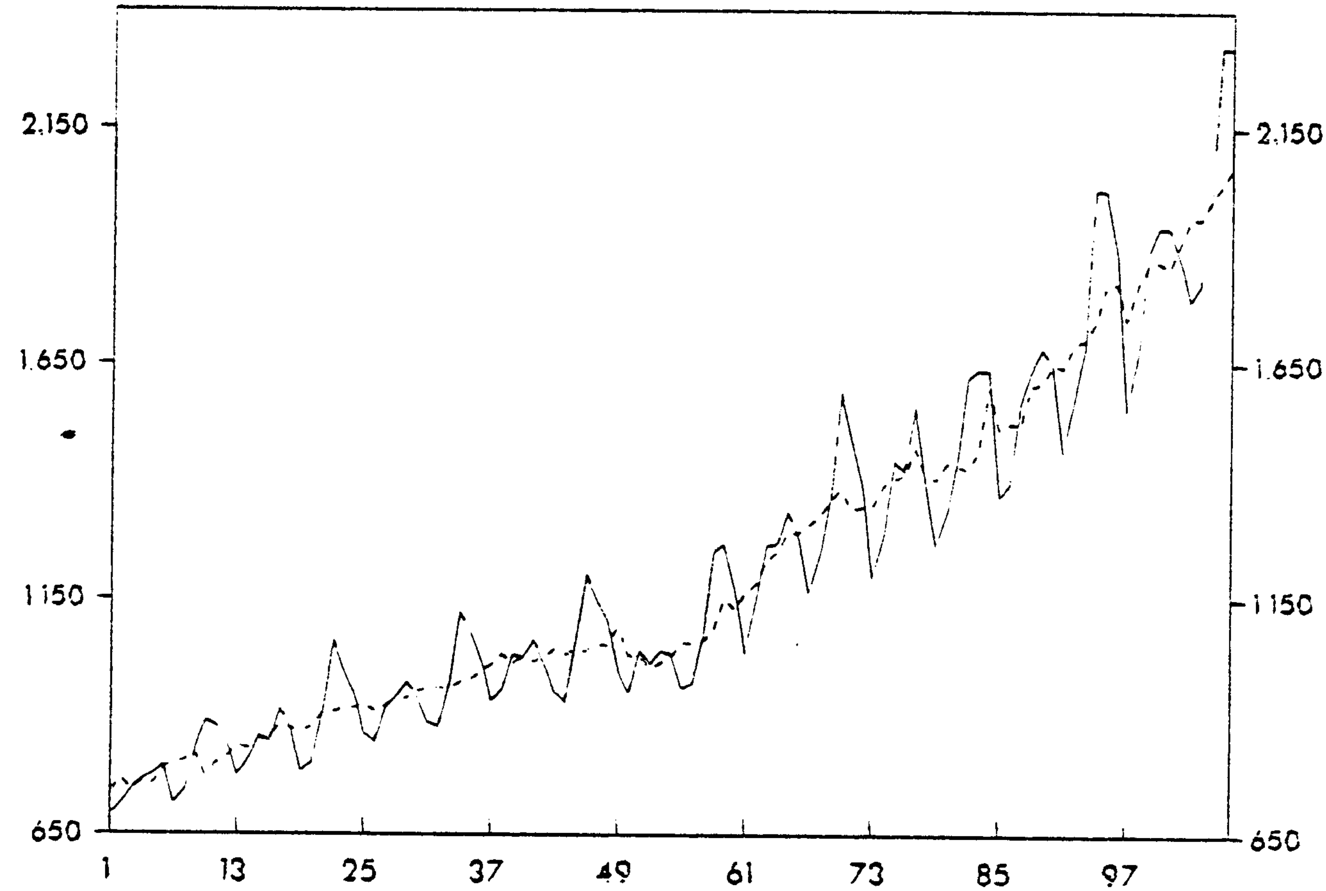


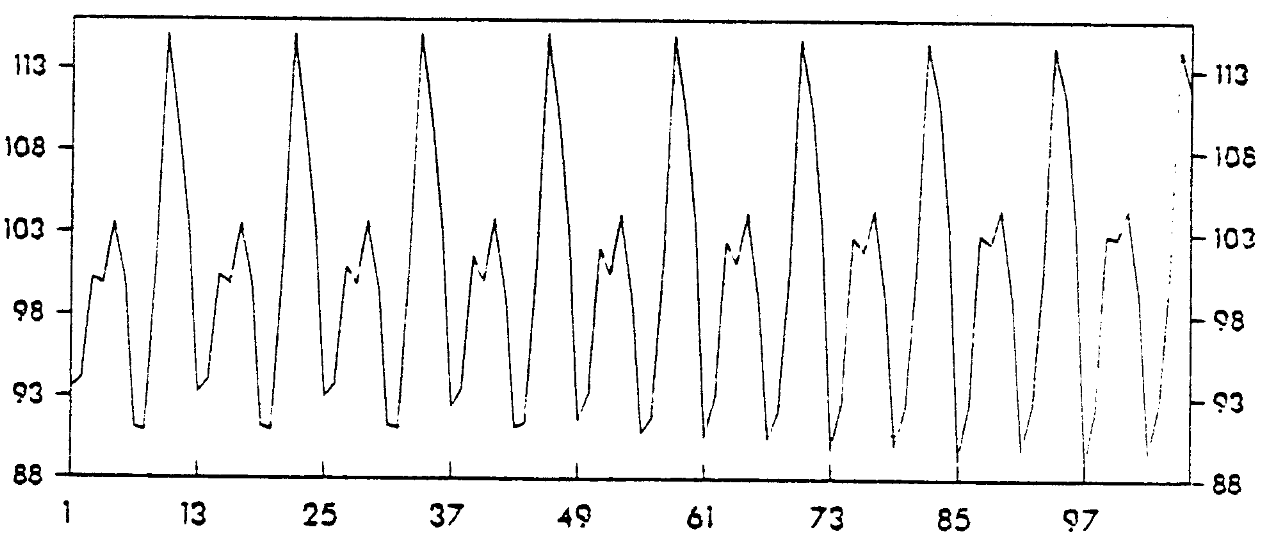
RETAIL 731000 MODEL BASED



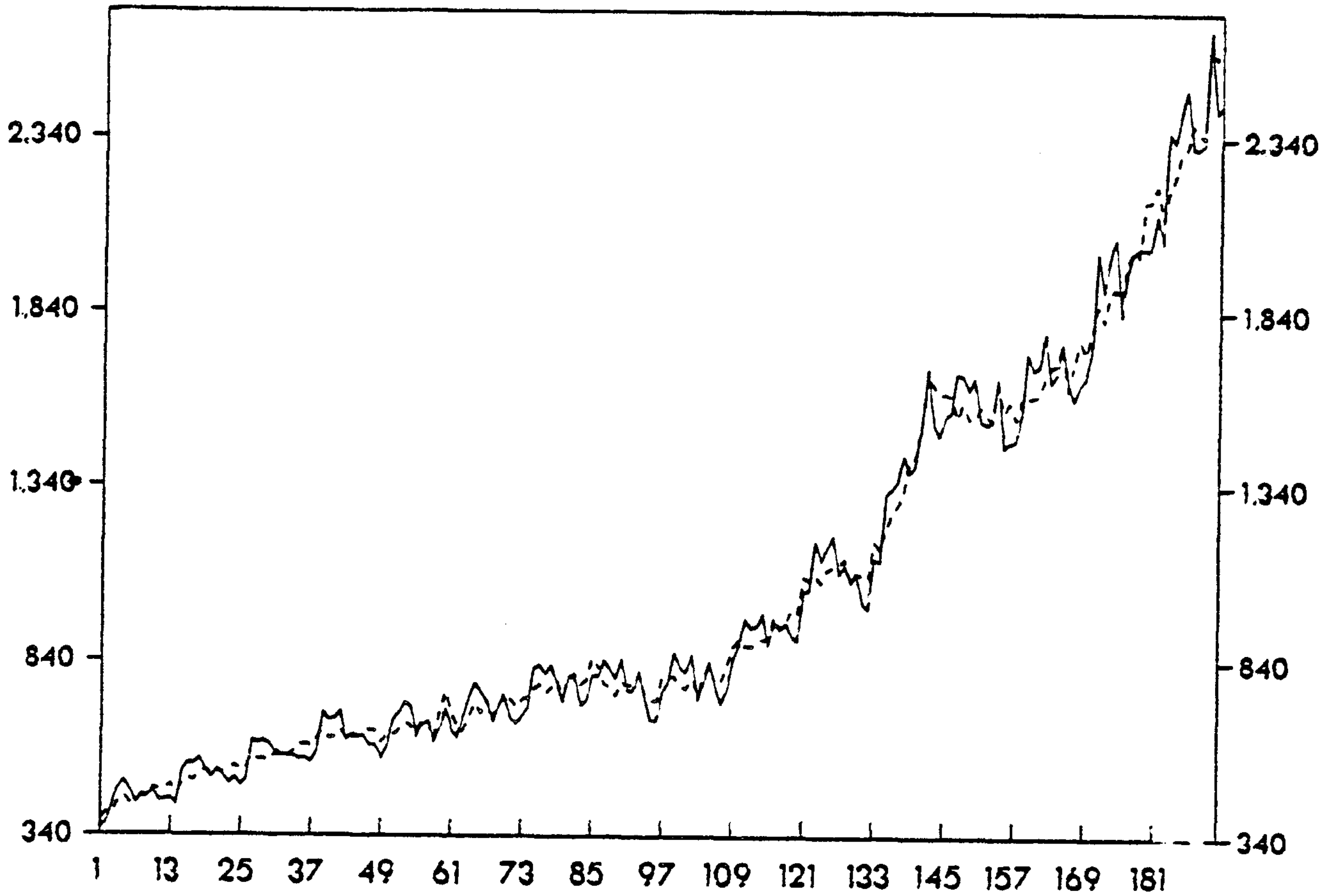


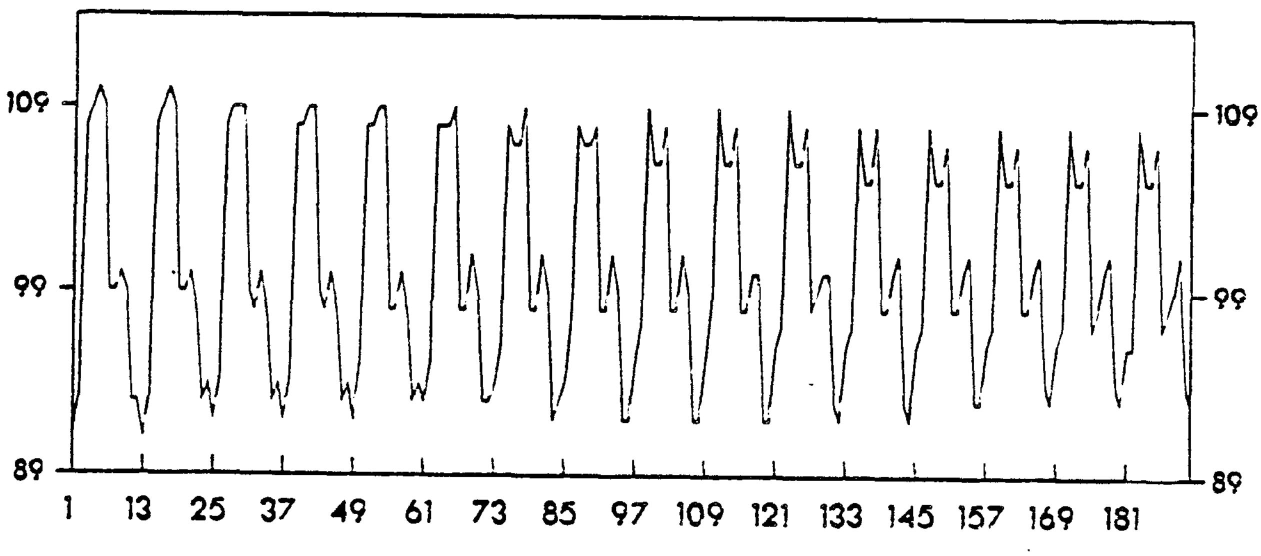
RETAIL 731000 X-11



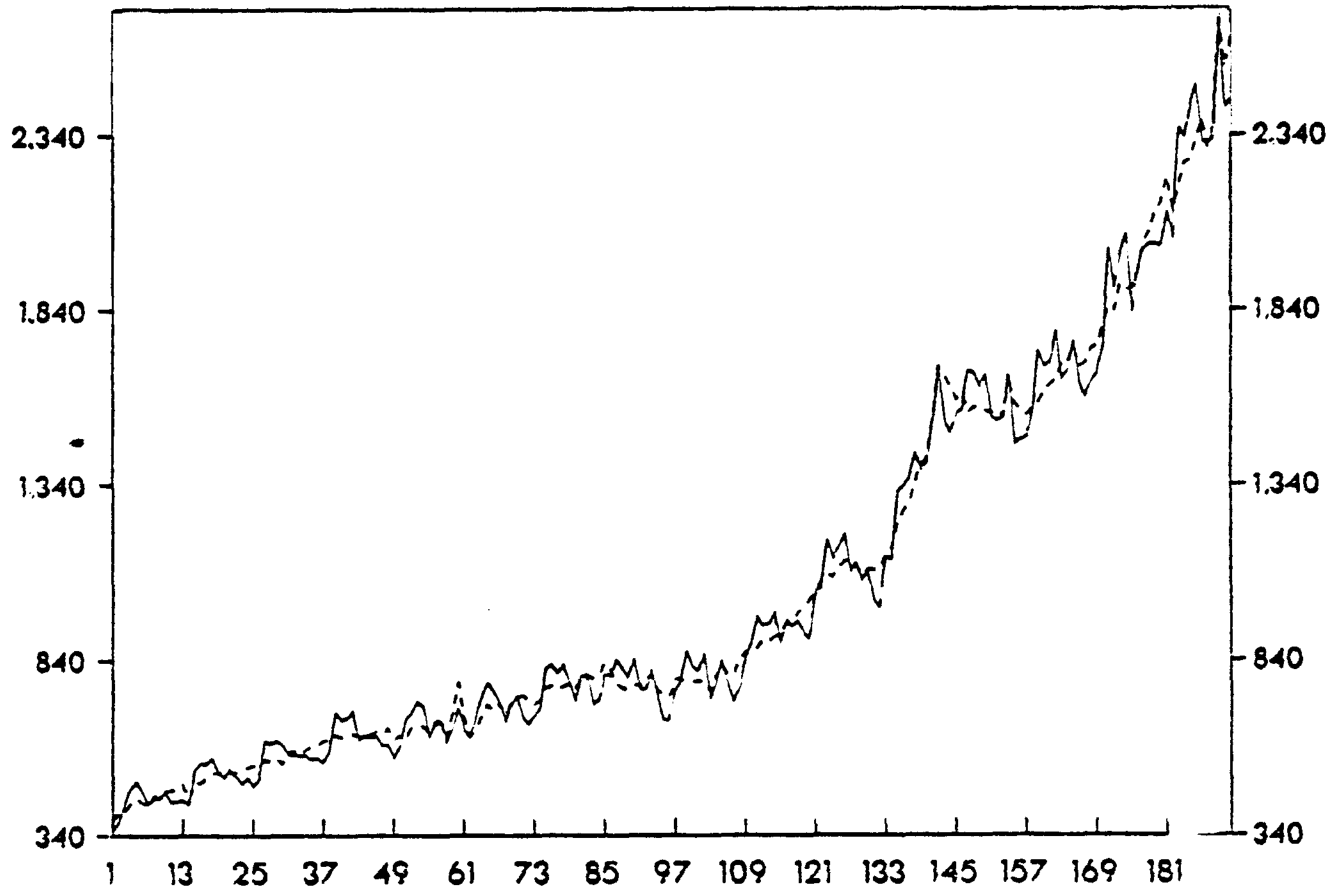


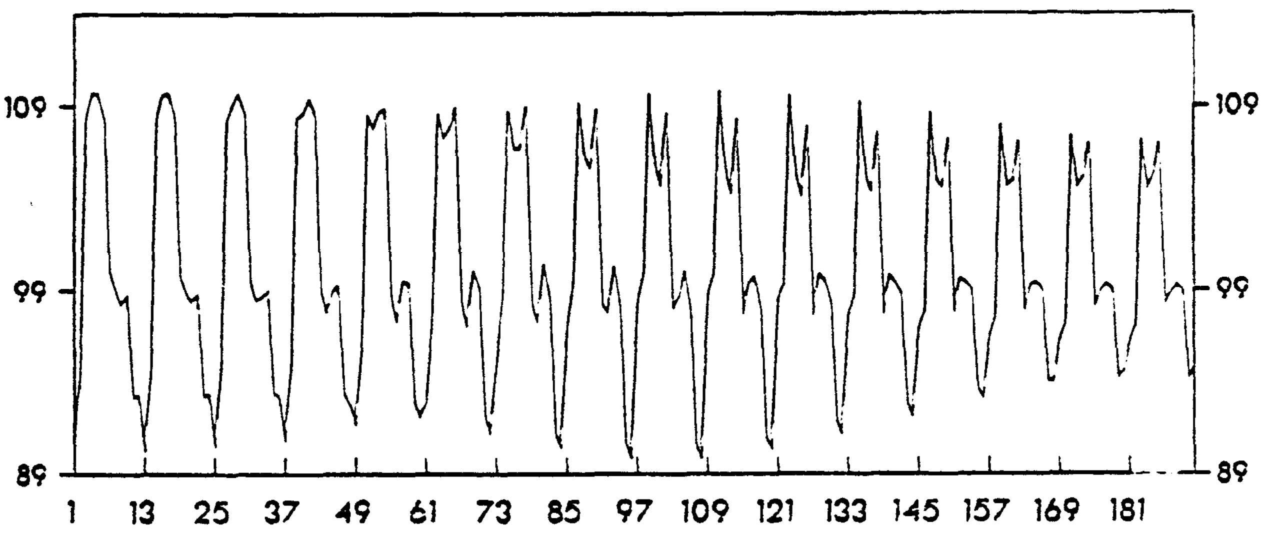
INDUSTRY S24YSU MODEL BASED



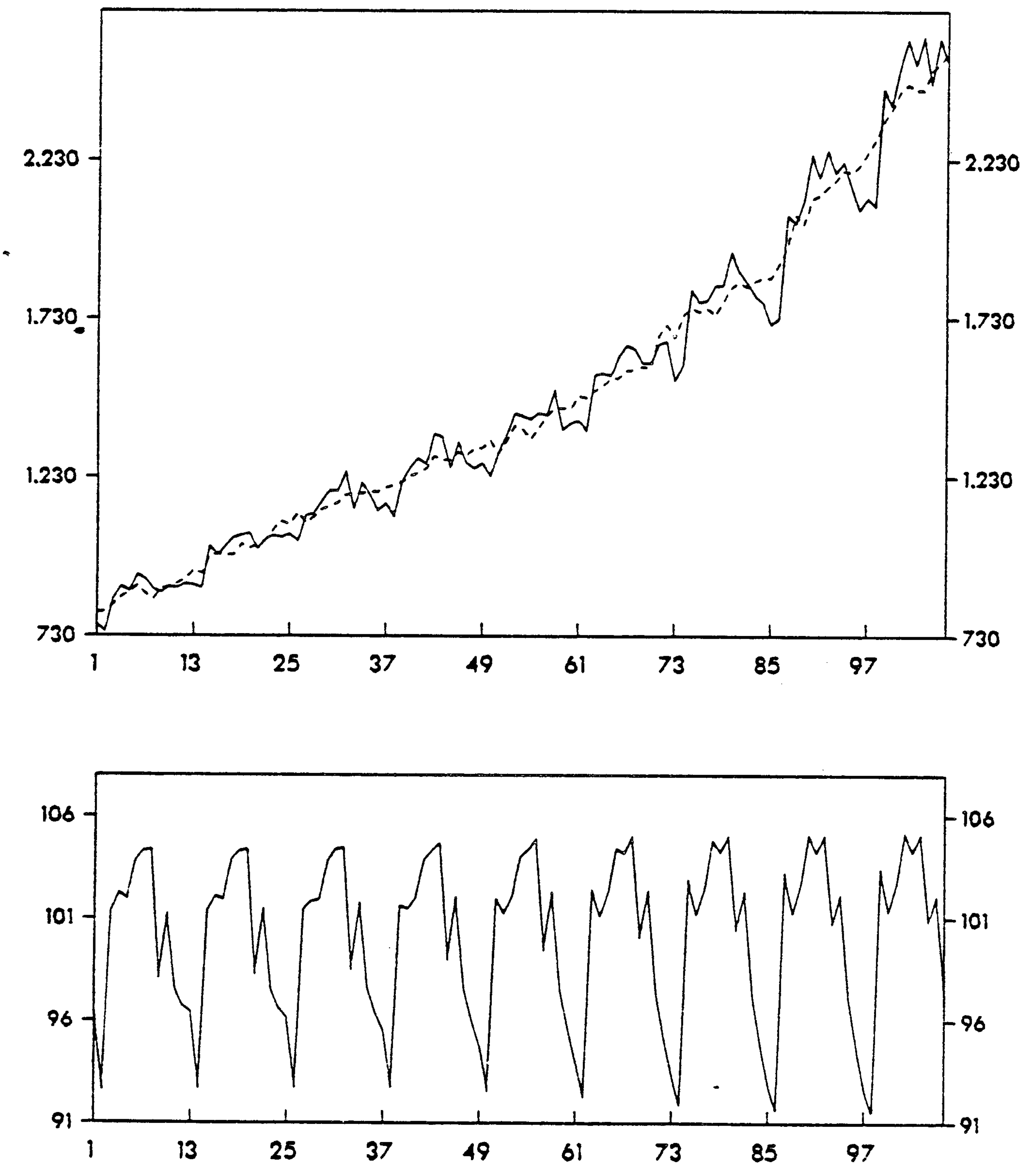


INDUSTRY S24VSU X-11 LOG

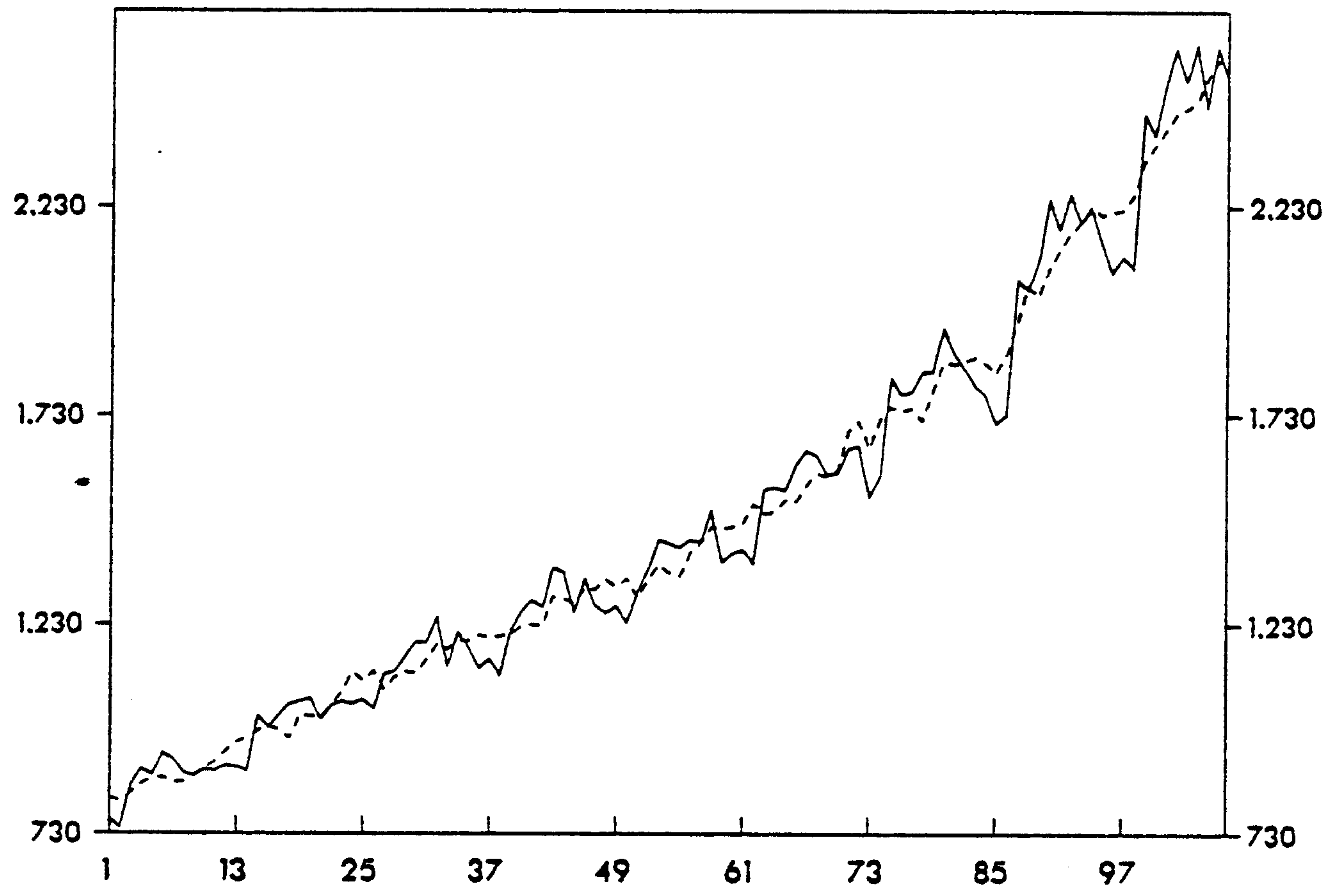


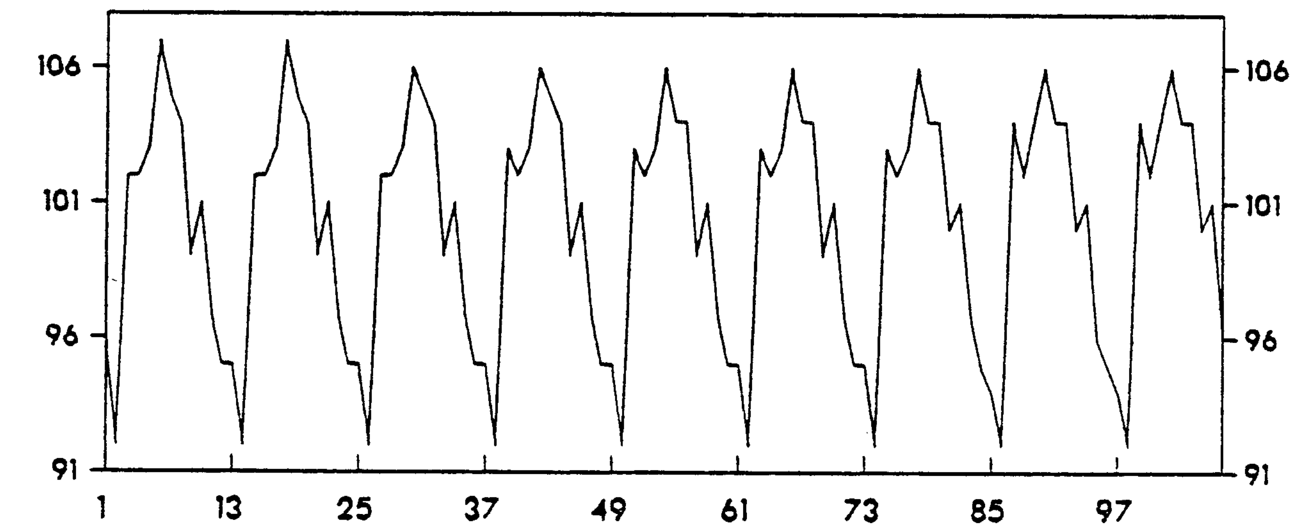


RETAIL 750000 X-11

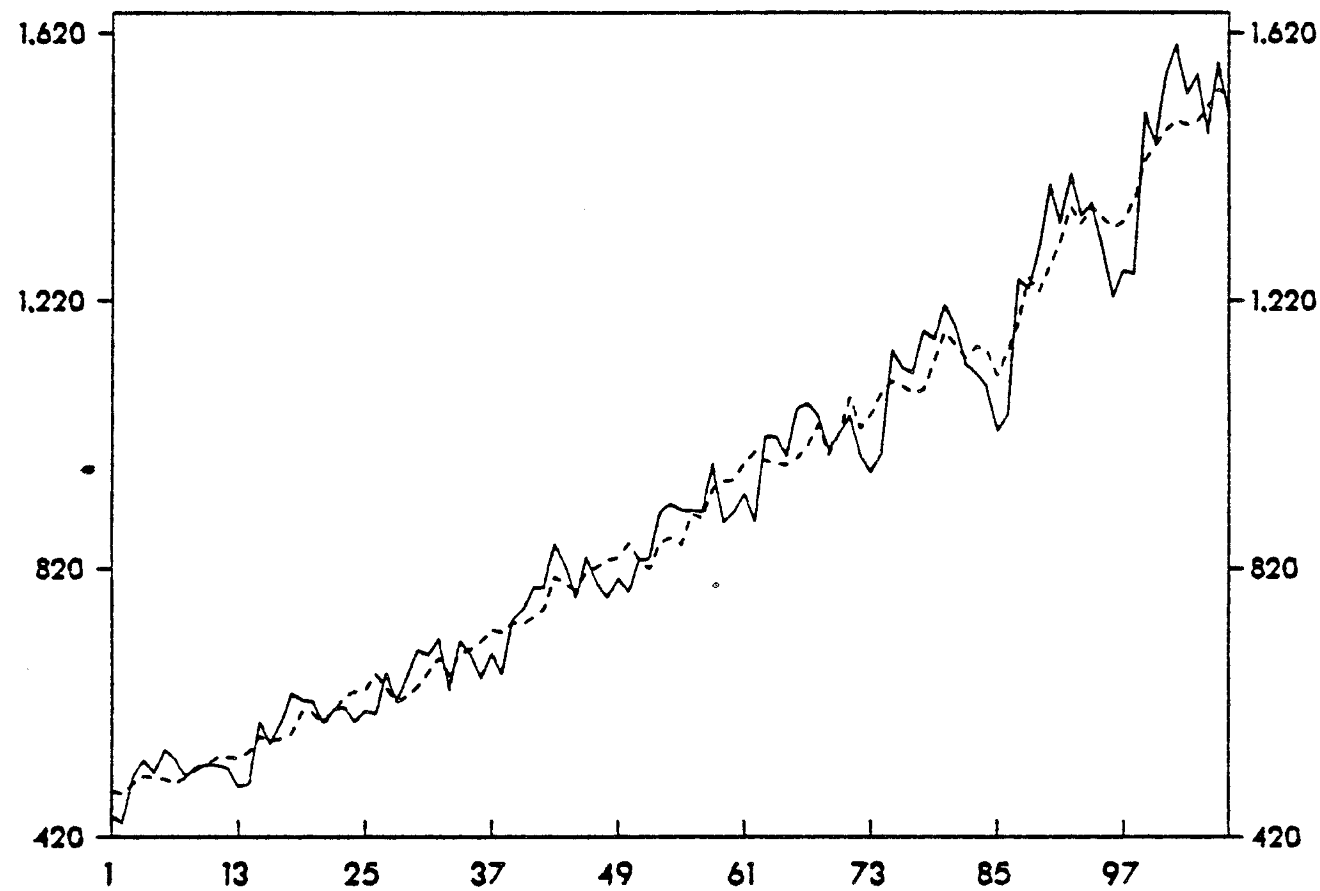


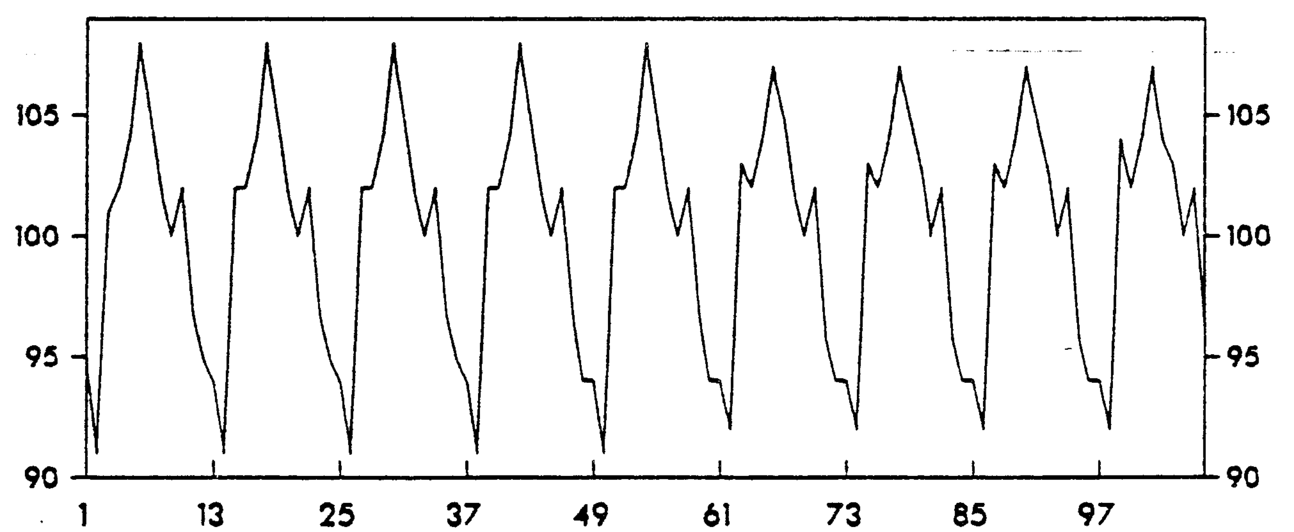
RETAIL 750000 MODEL BASED



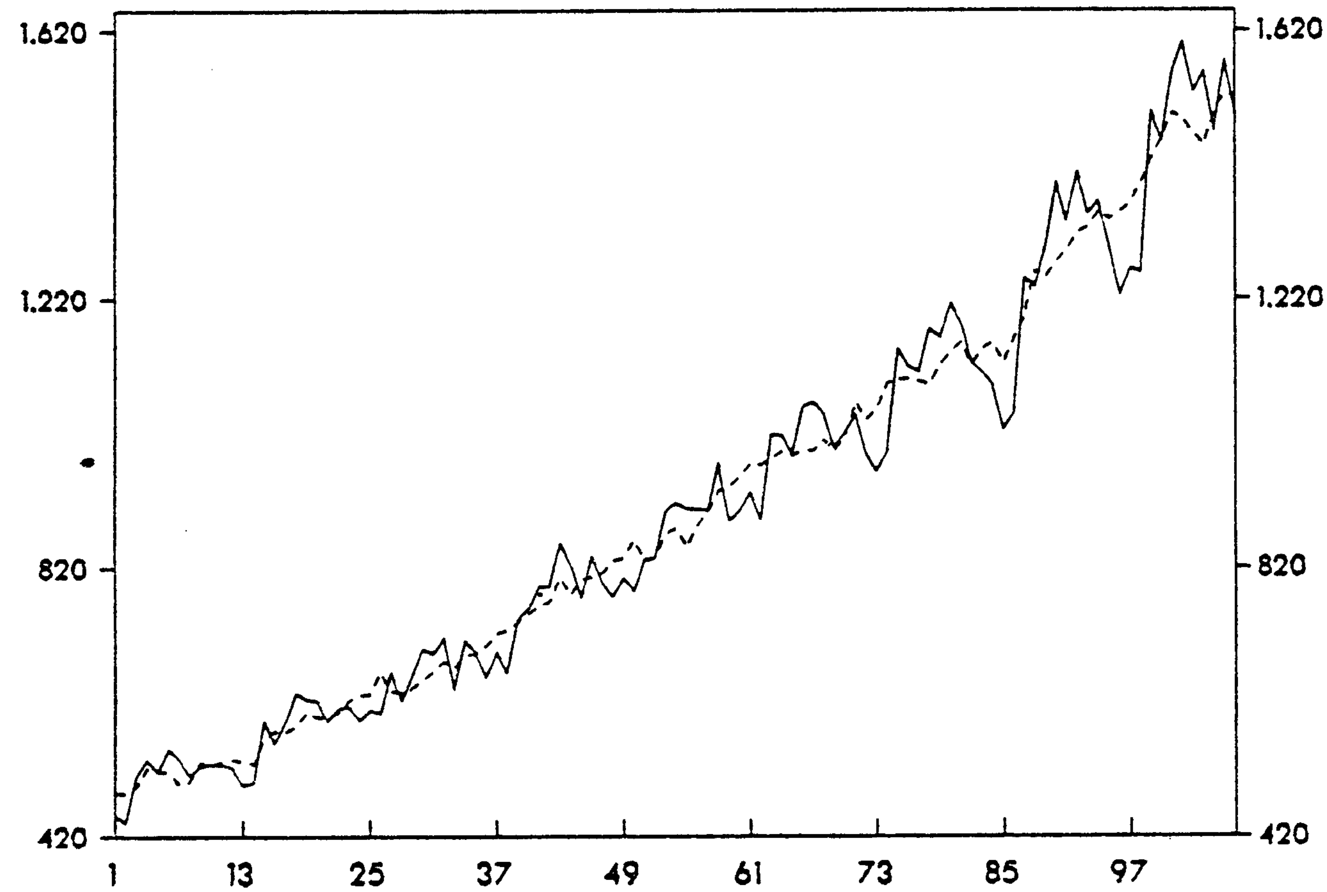


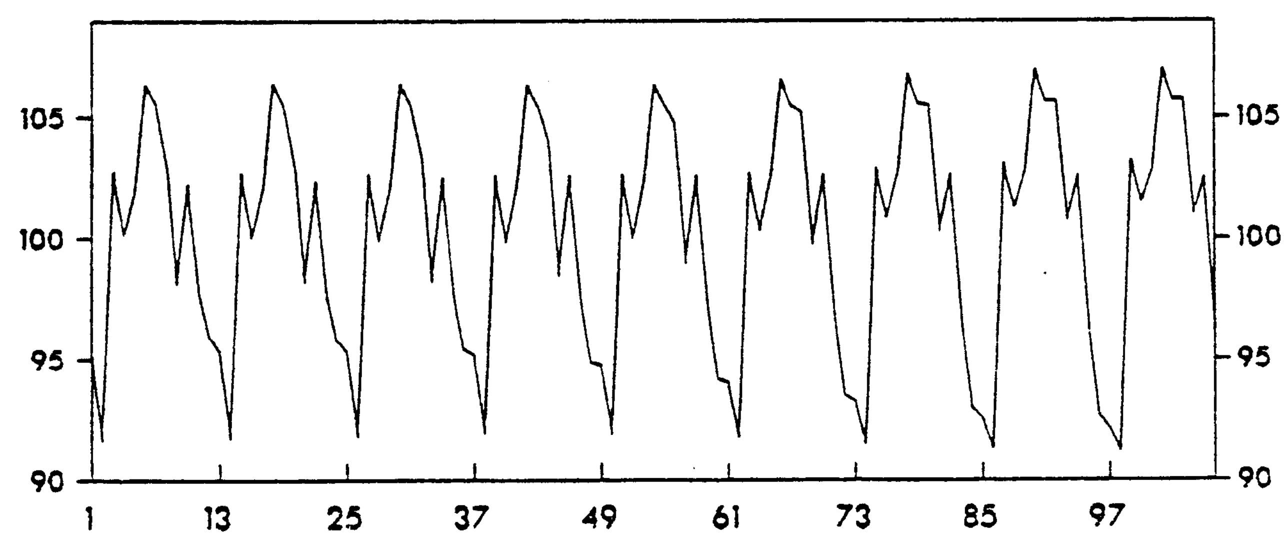
RETAIL 753000 MODEL BASED



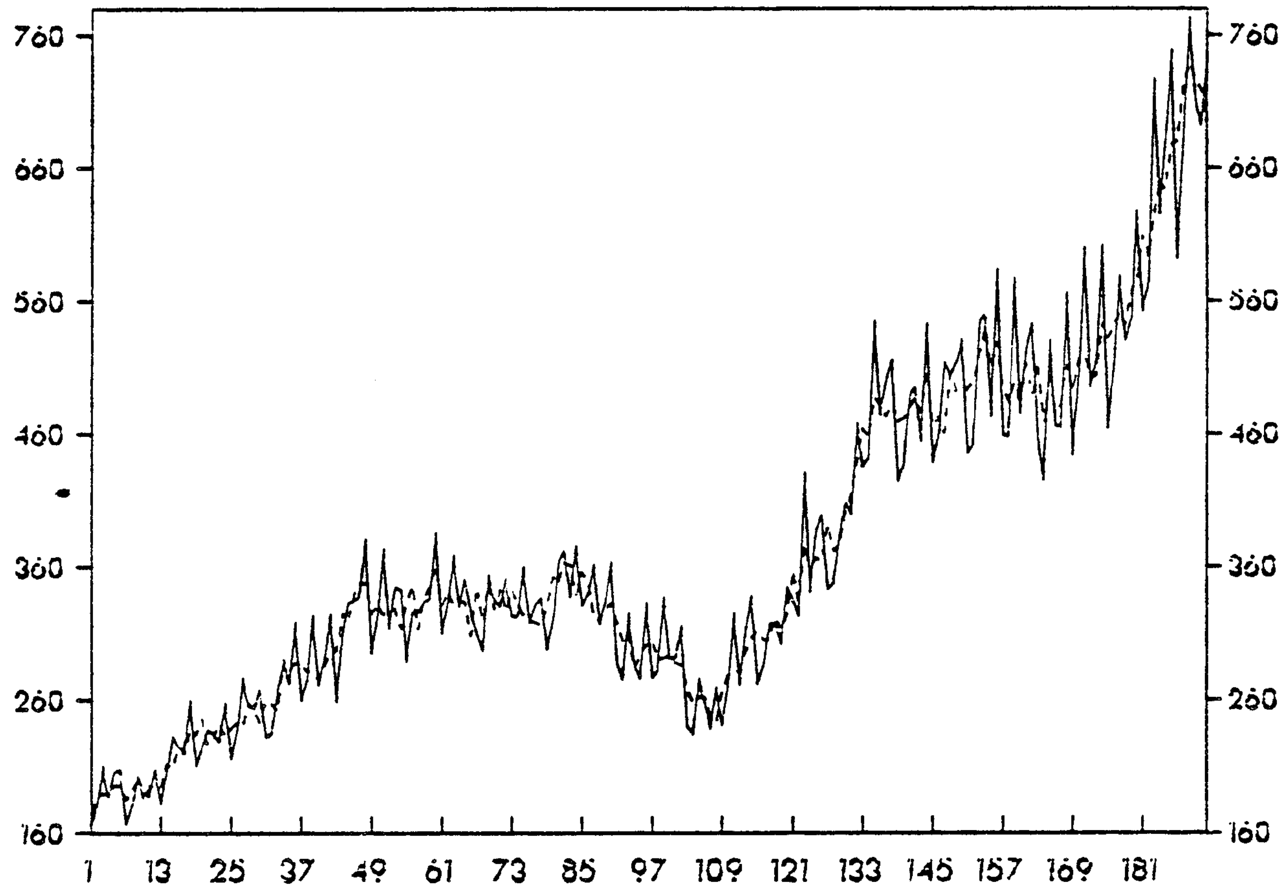


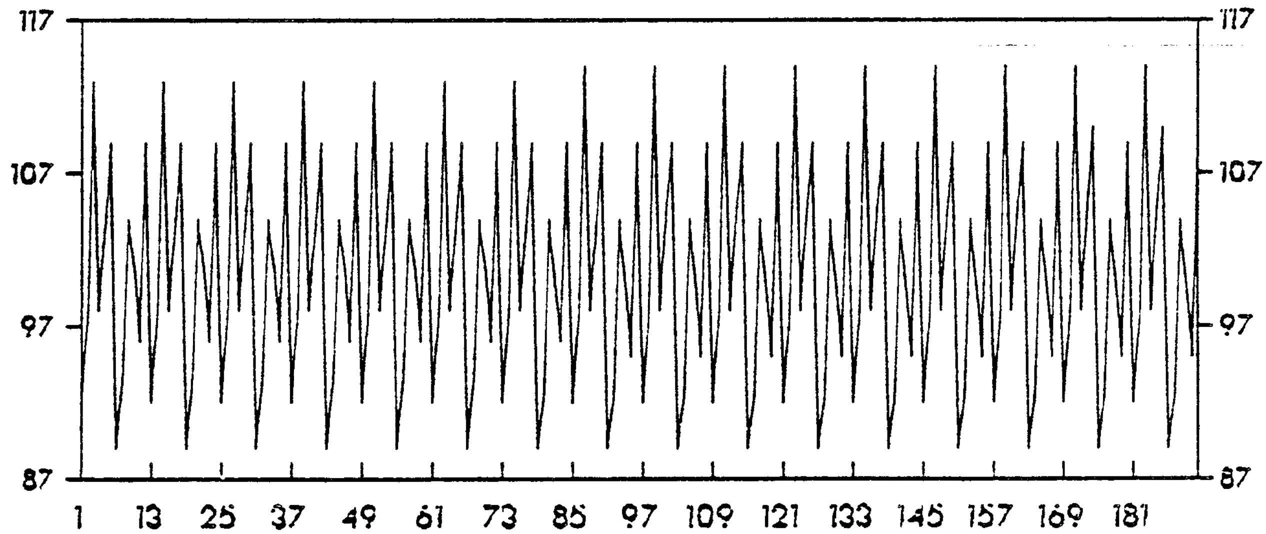
RETAIL 753000 X-11



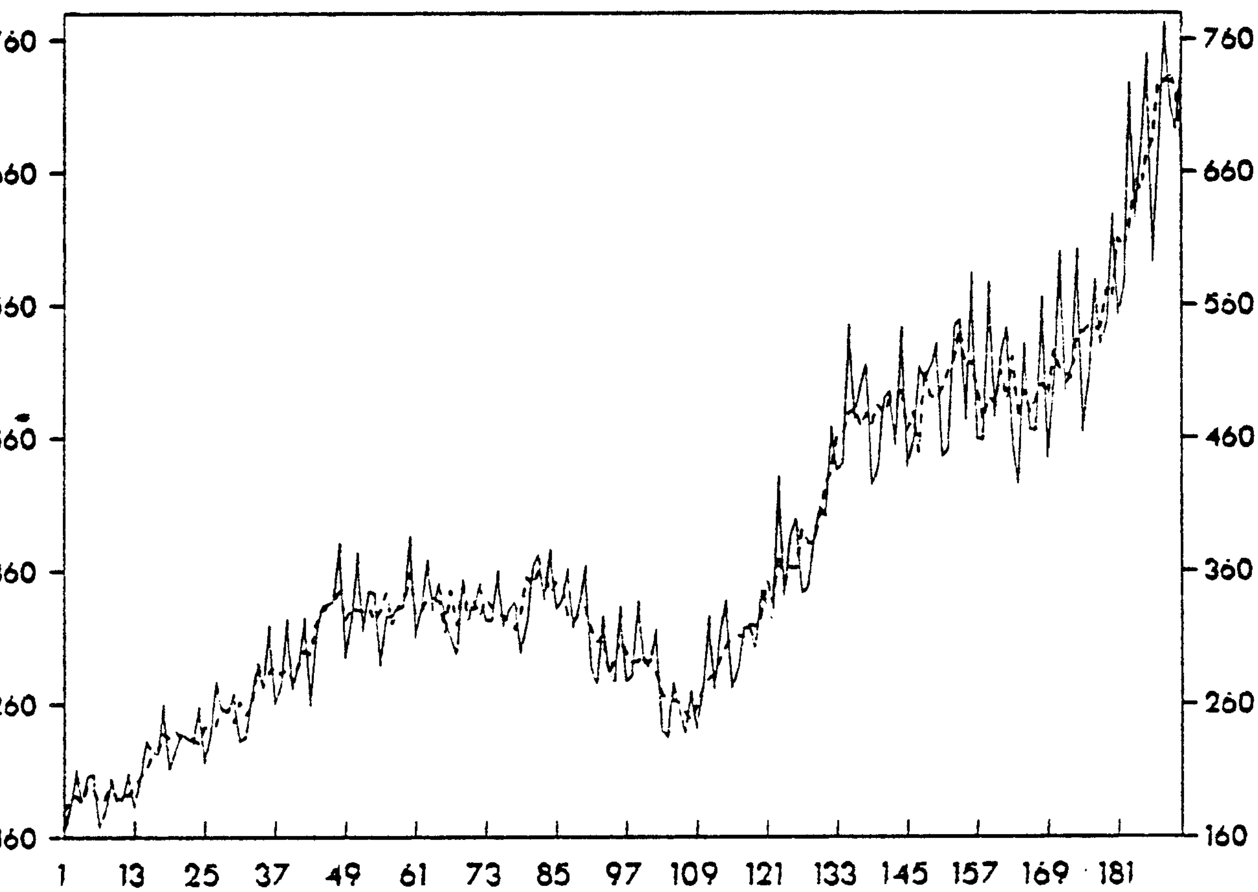


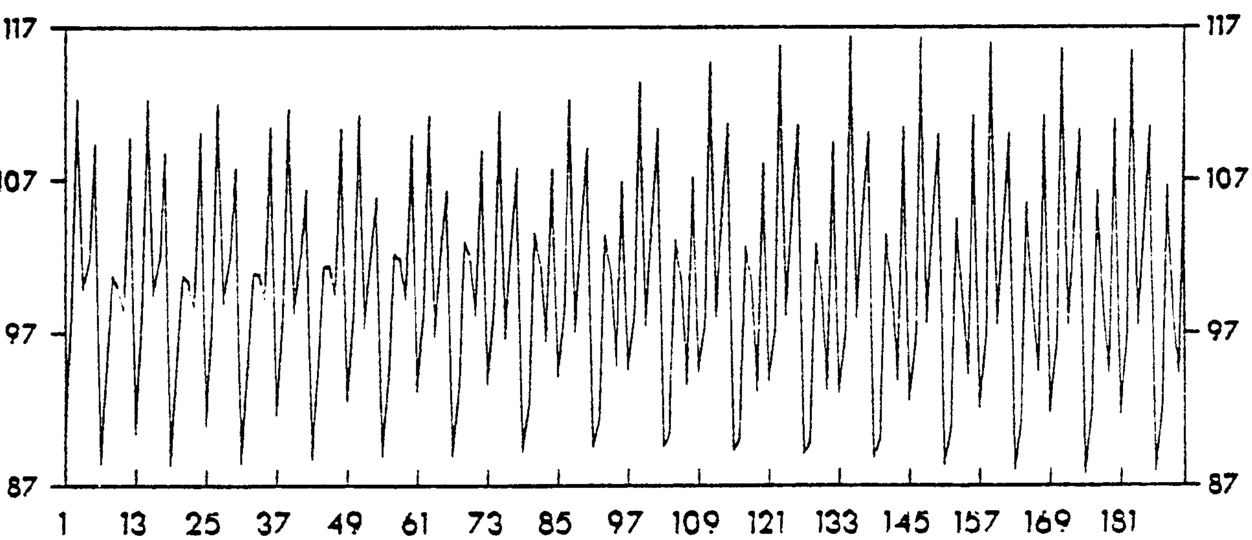
INDUSTRY S25YSU MODEL BASED



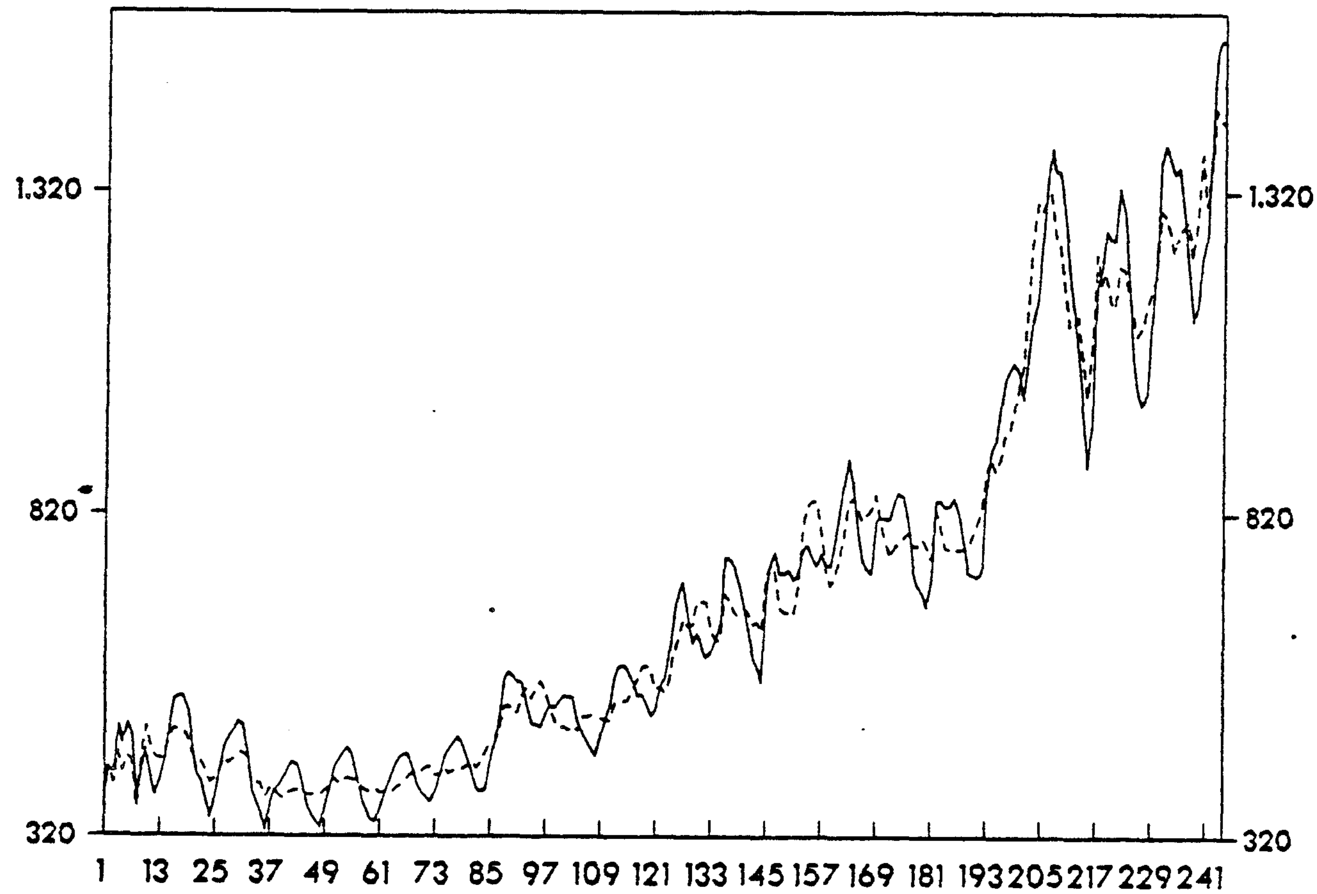


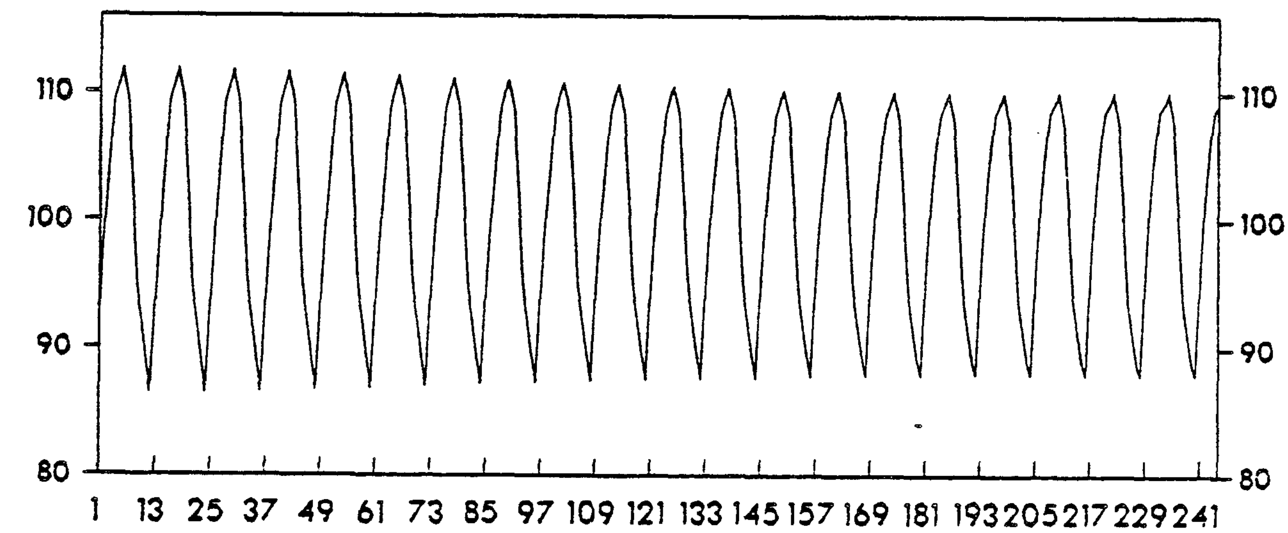
INDUSTRY S25VSU X-11 LOG



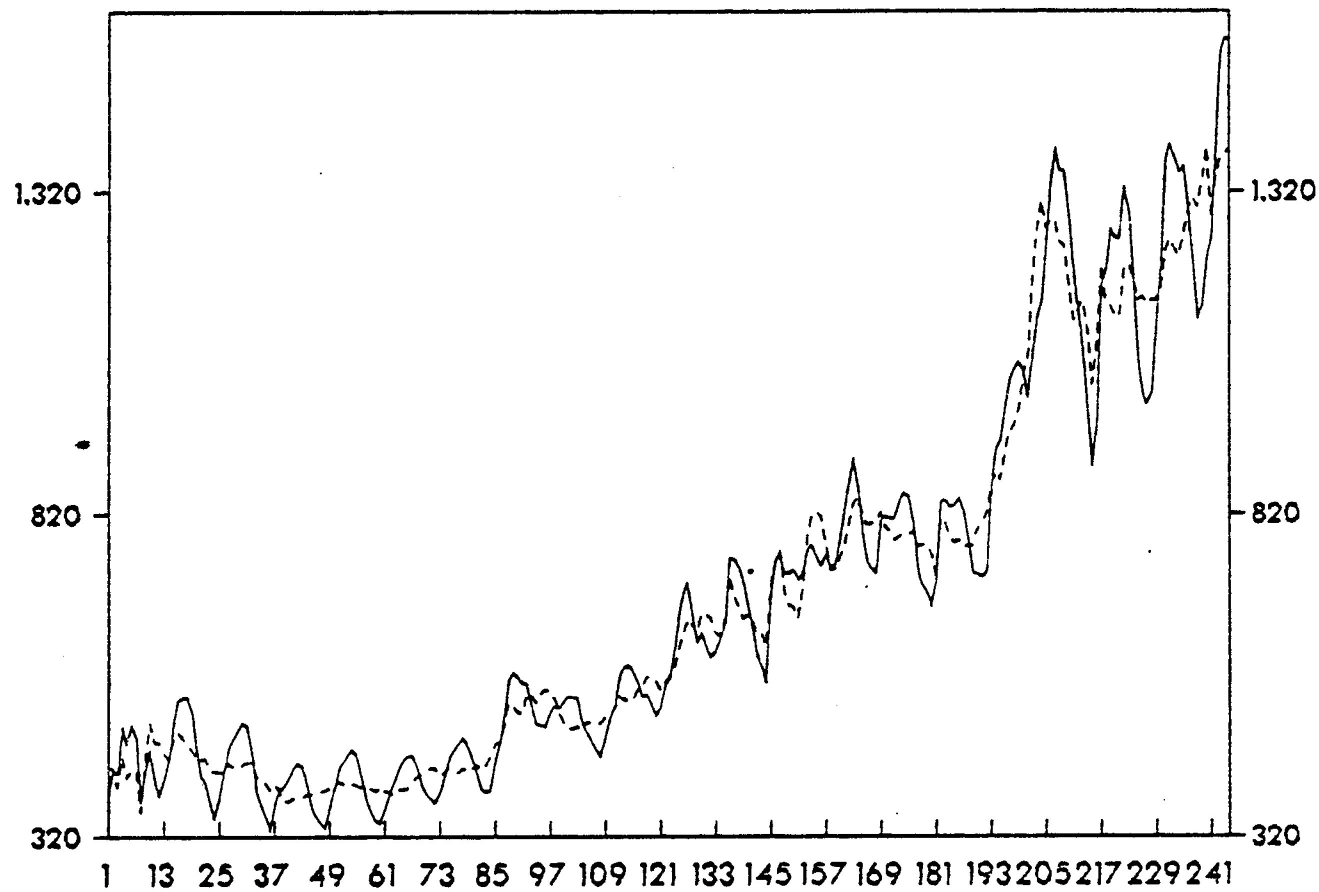


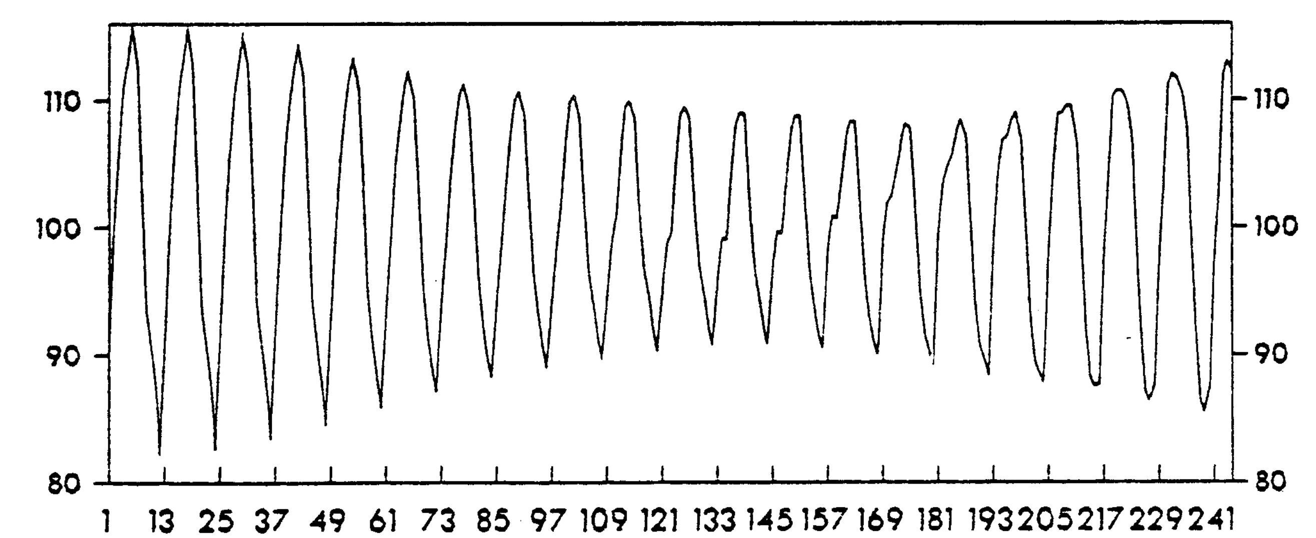
INDUSTRY S16TI MODEL BASED



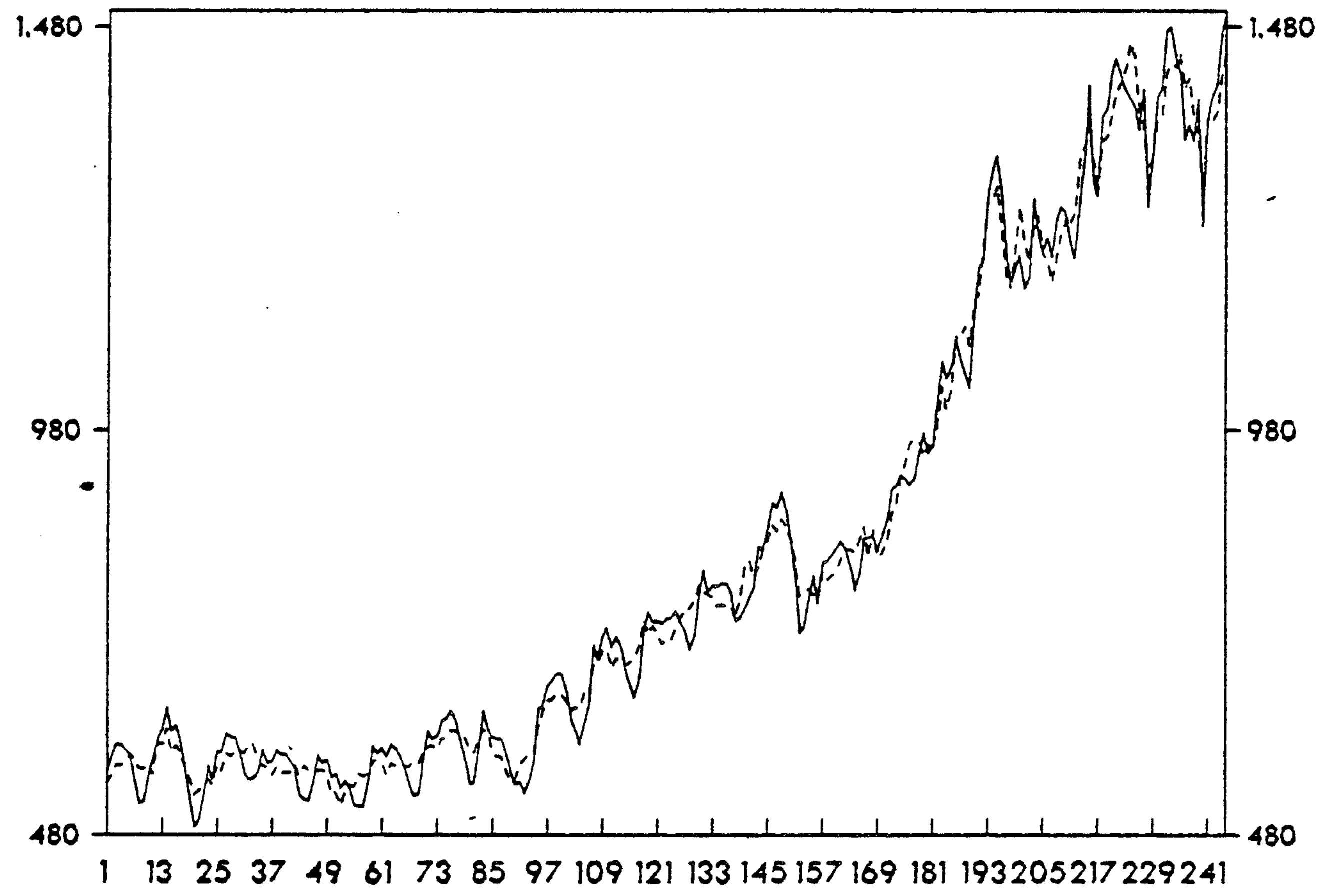


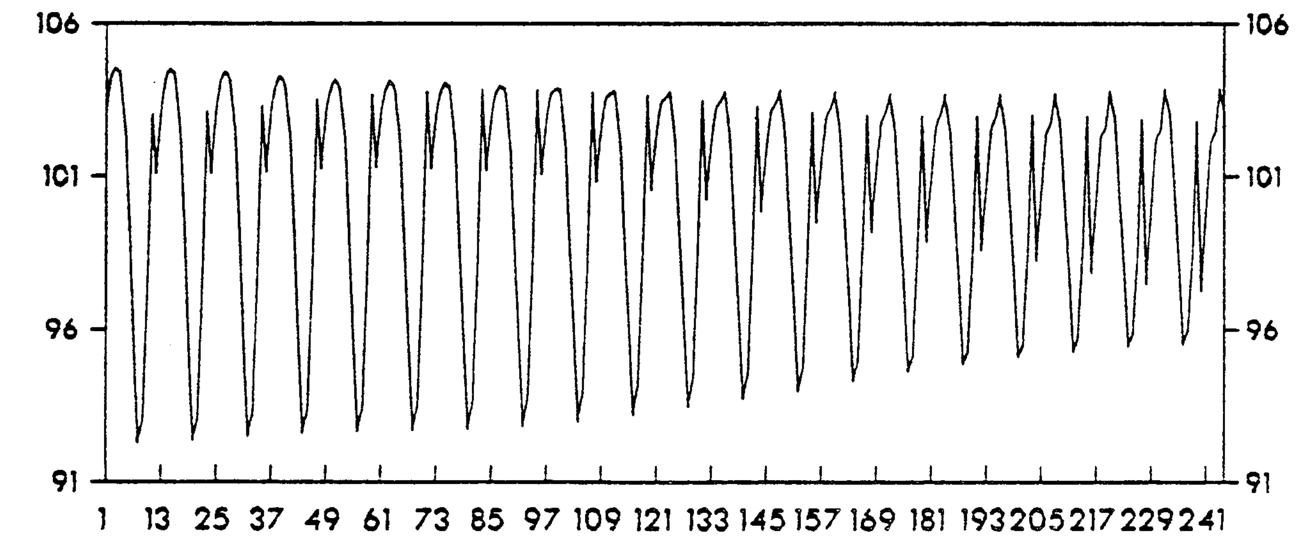
INDUSTRY S16TI X-11 LOG



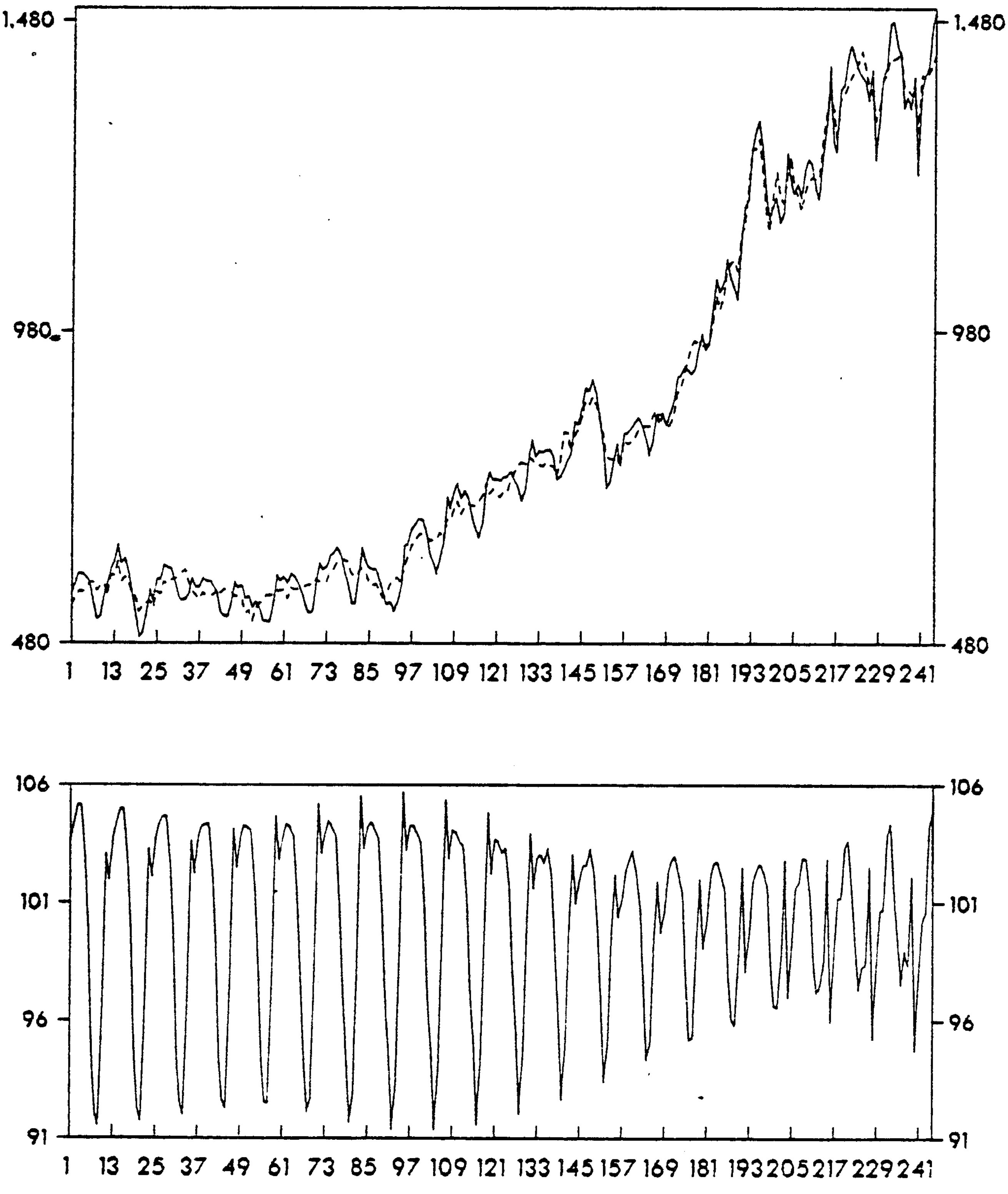


INDUSTRY S60TI MODEL BASED

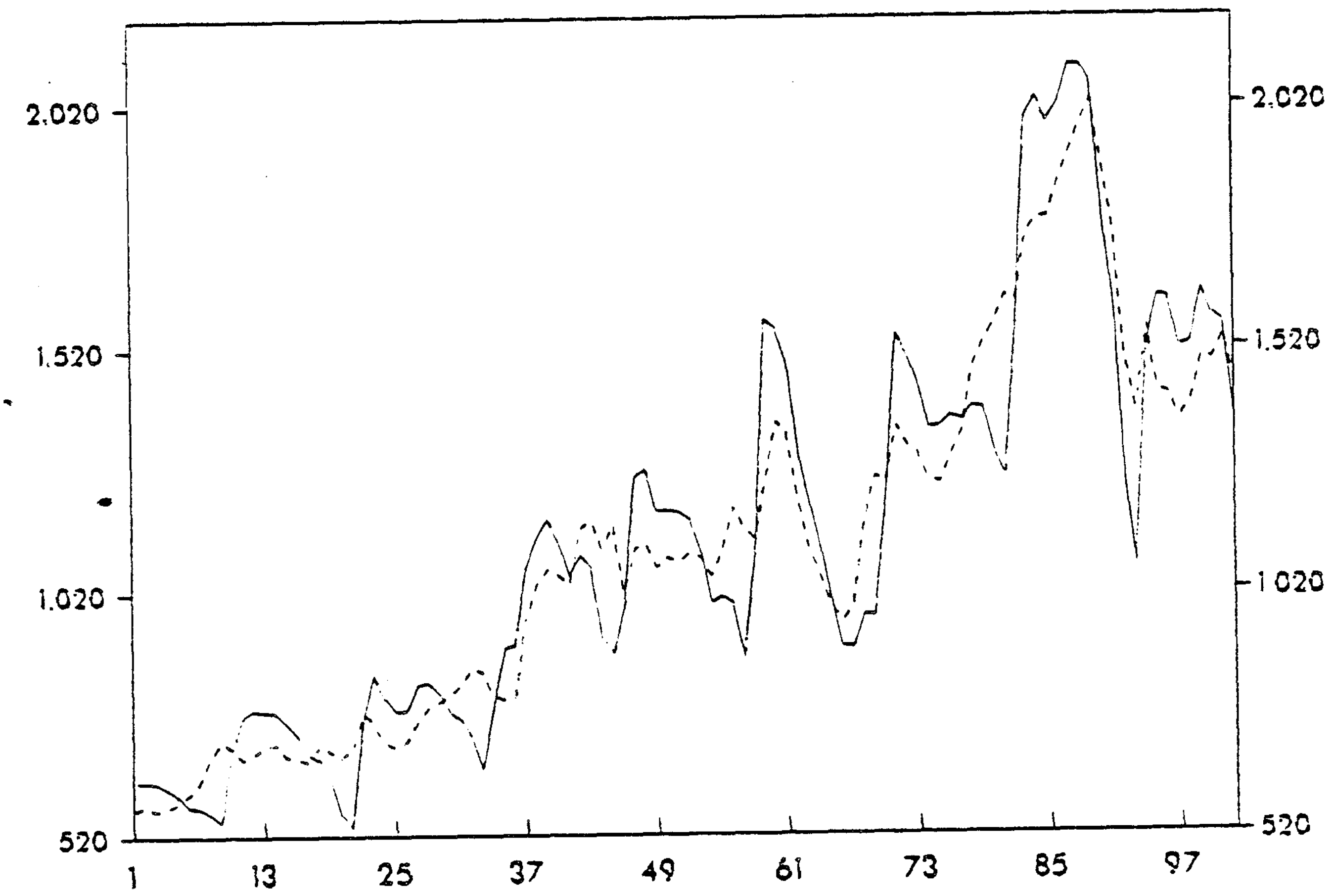


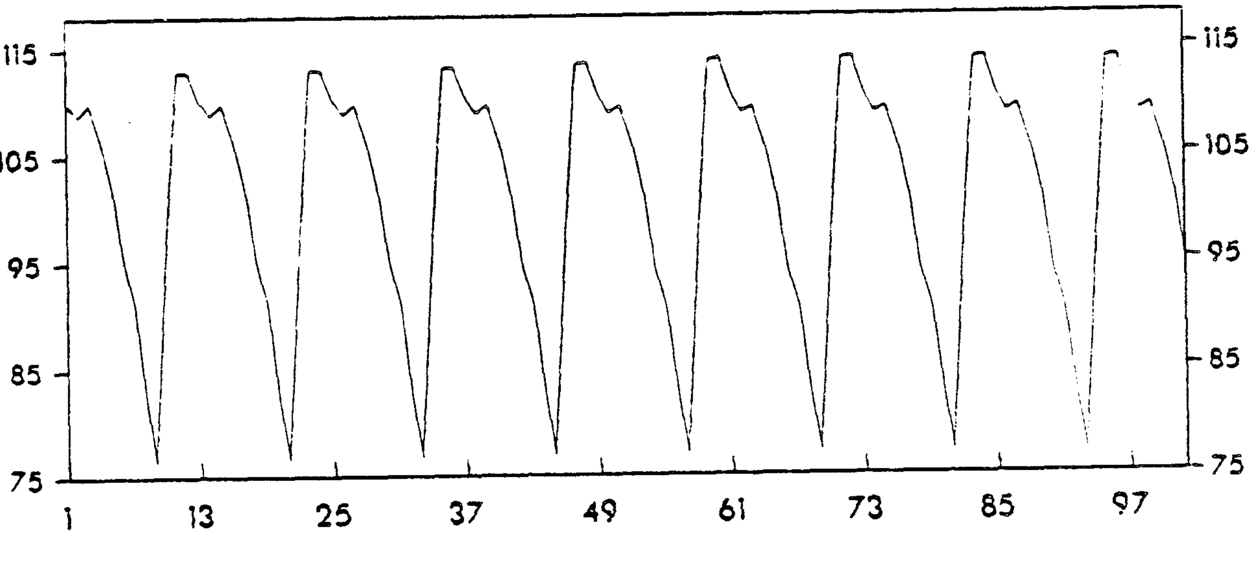


INDUSTRY S60TI X-11 LOG

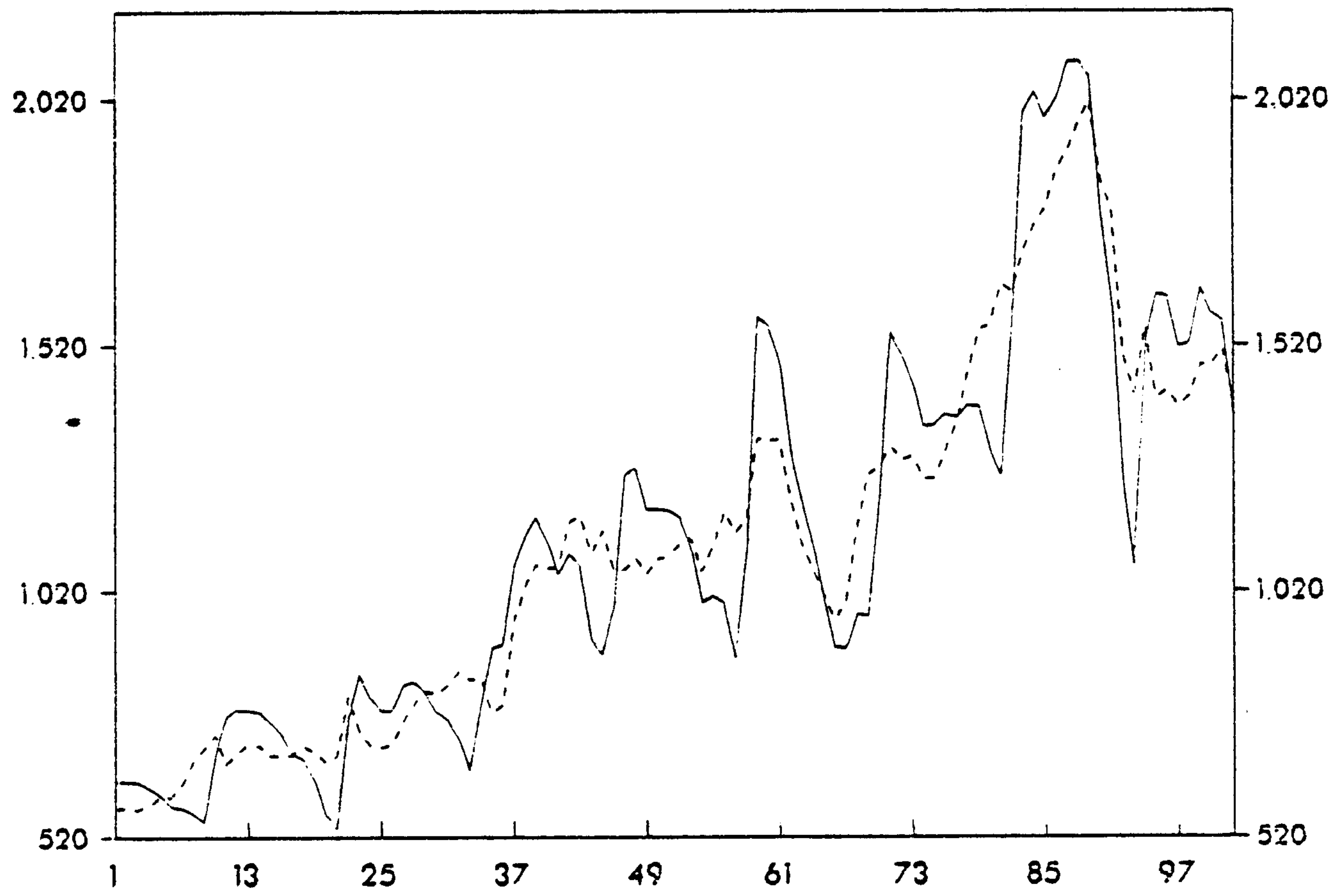


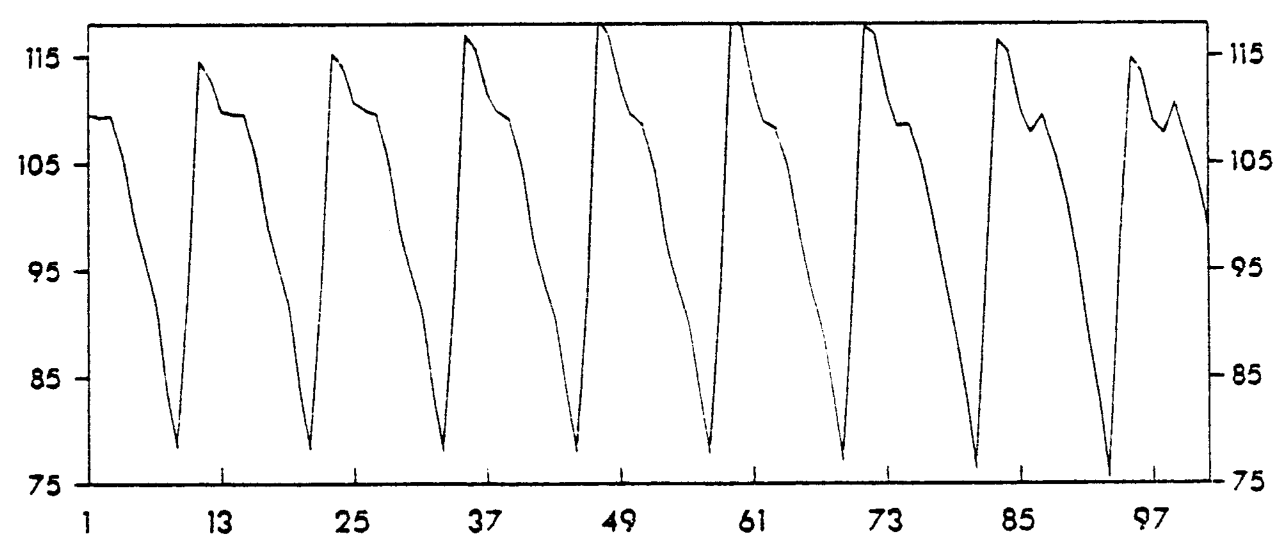
S63TI MODEL BASED



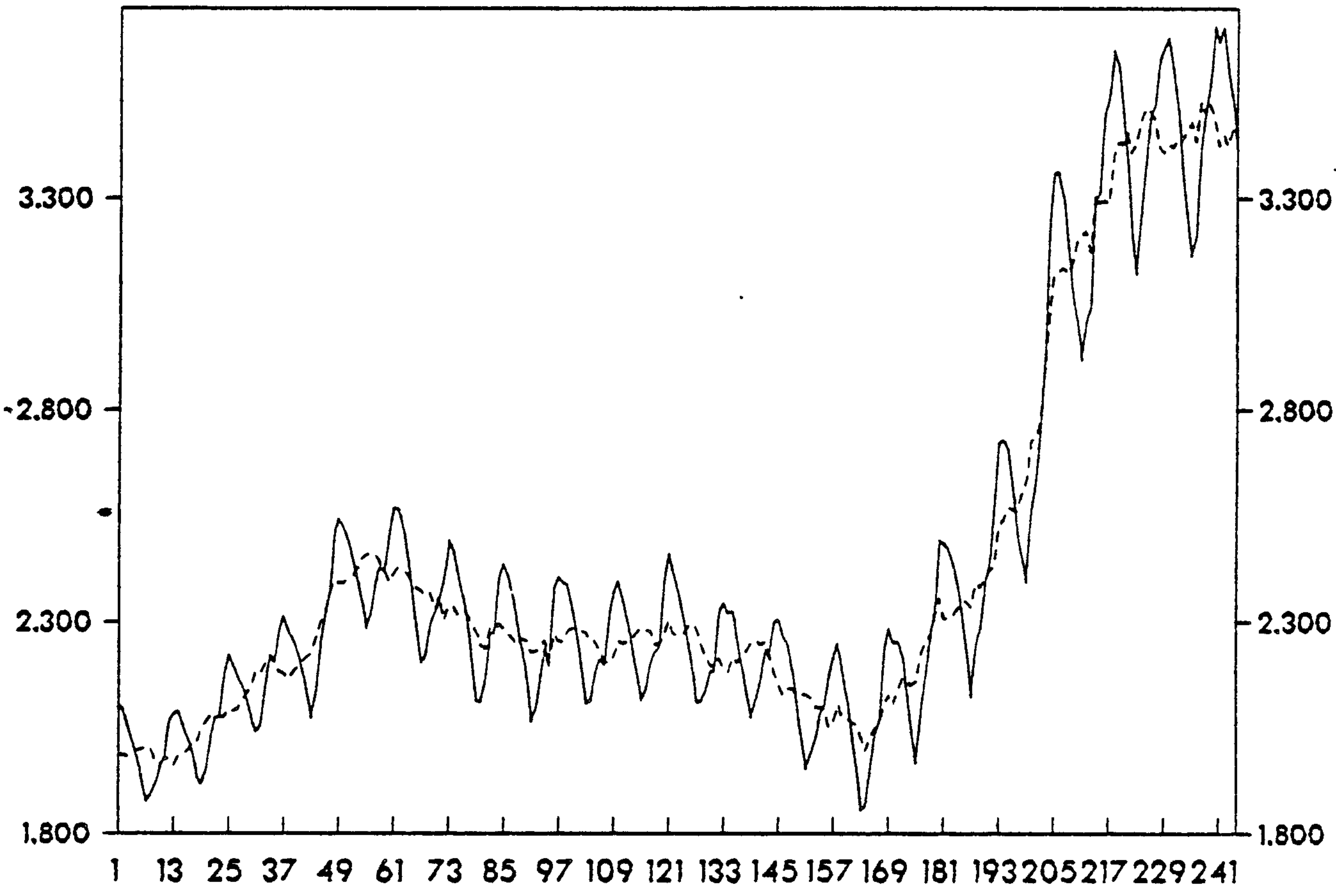


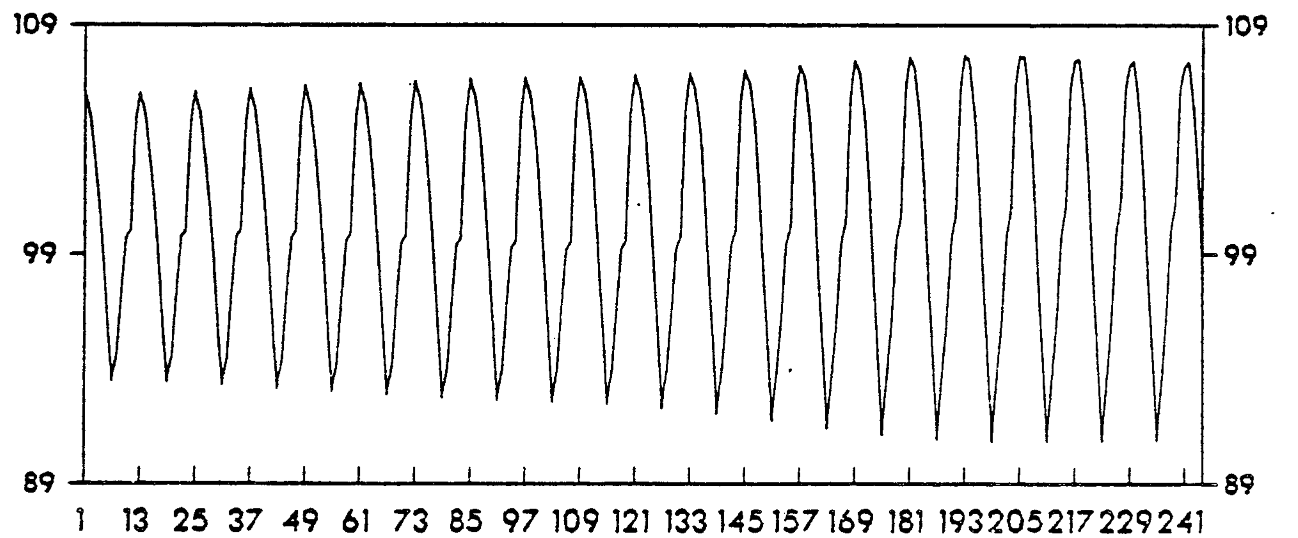
S63TI X-11 LOG



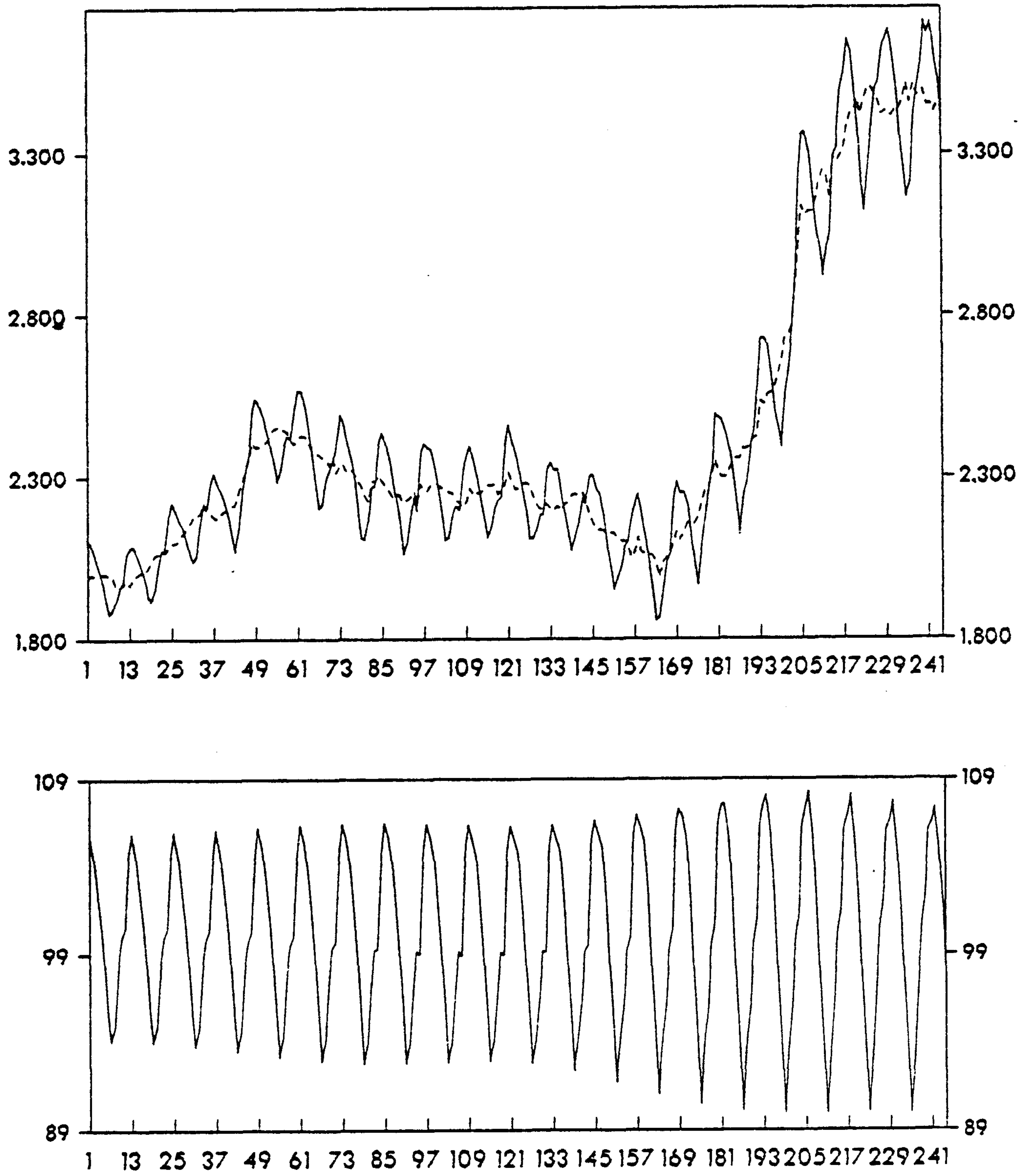


INDUSTRY S65TI MODEL BASED

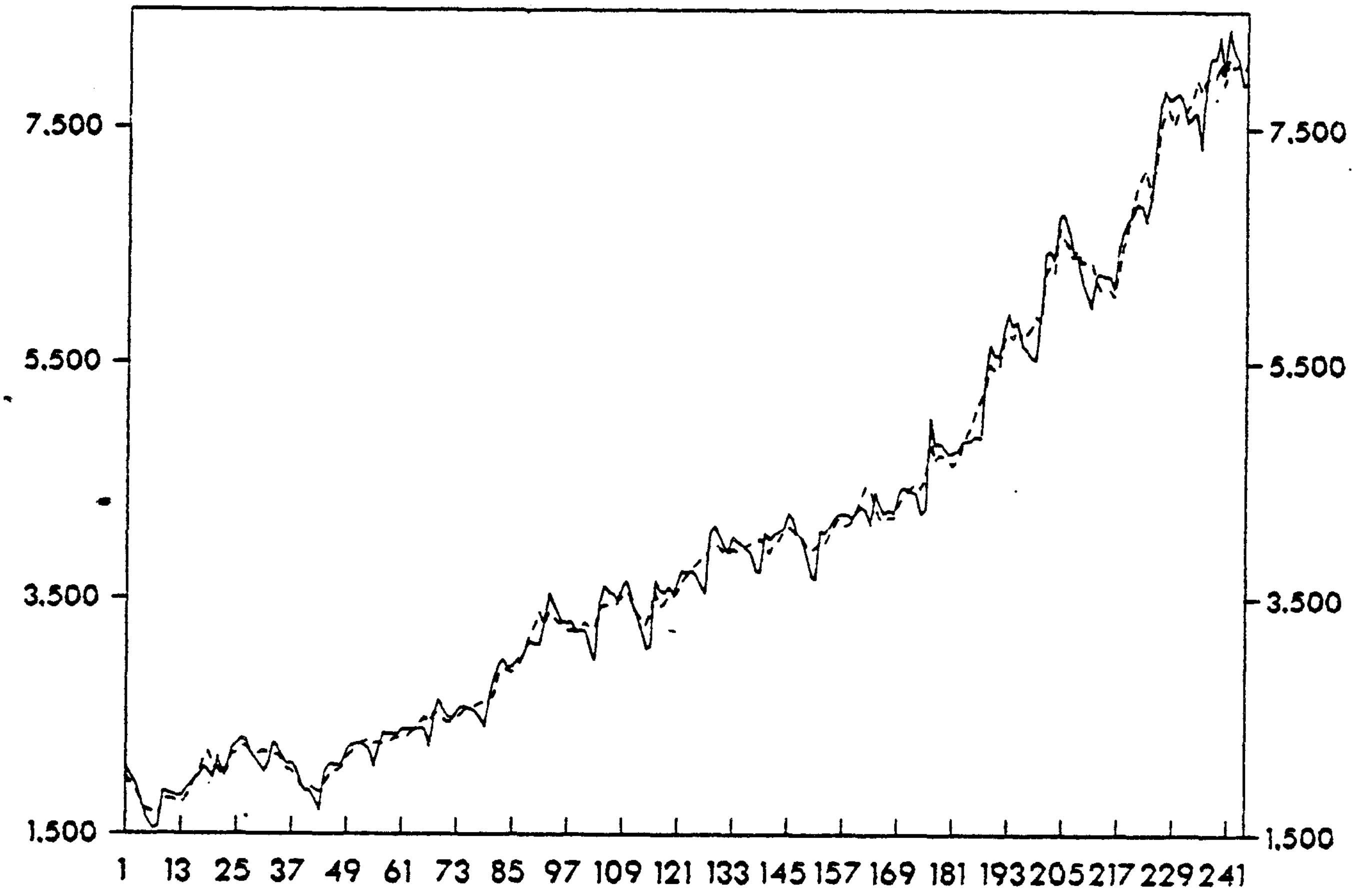


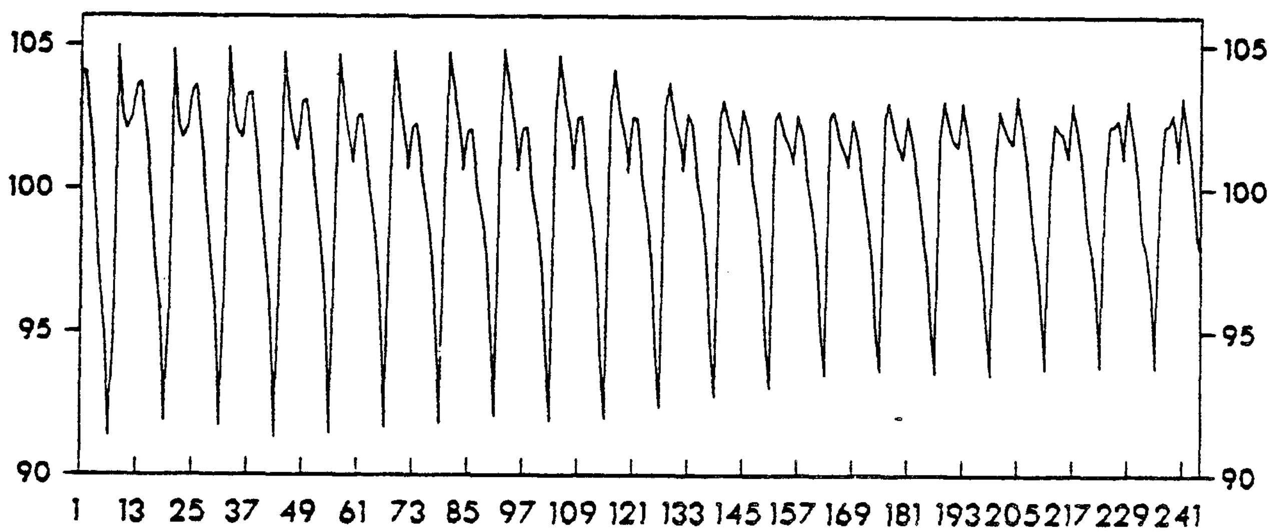


INDUSTRY S65T! X-11 LOG

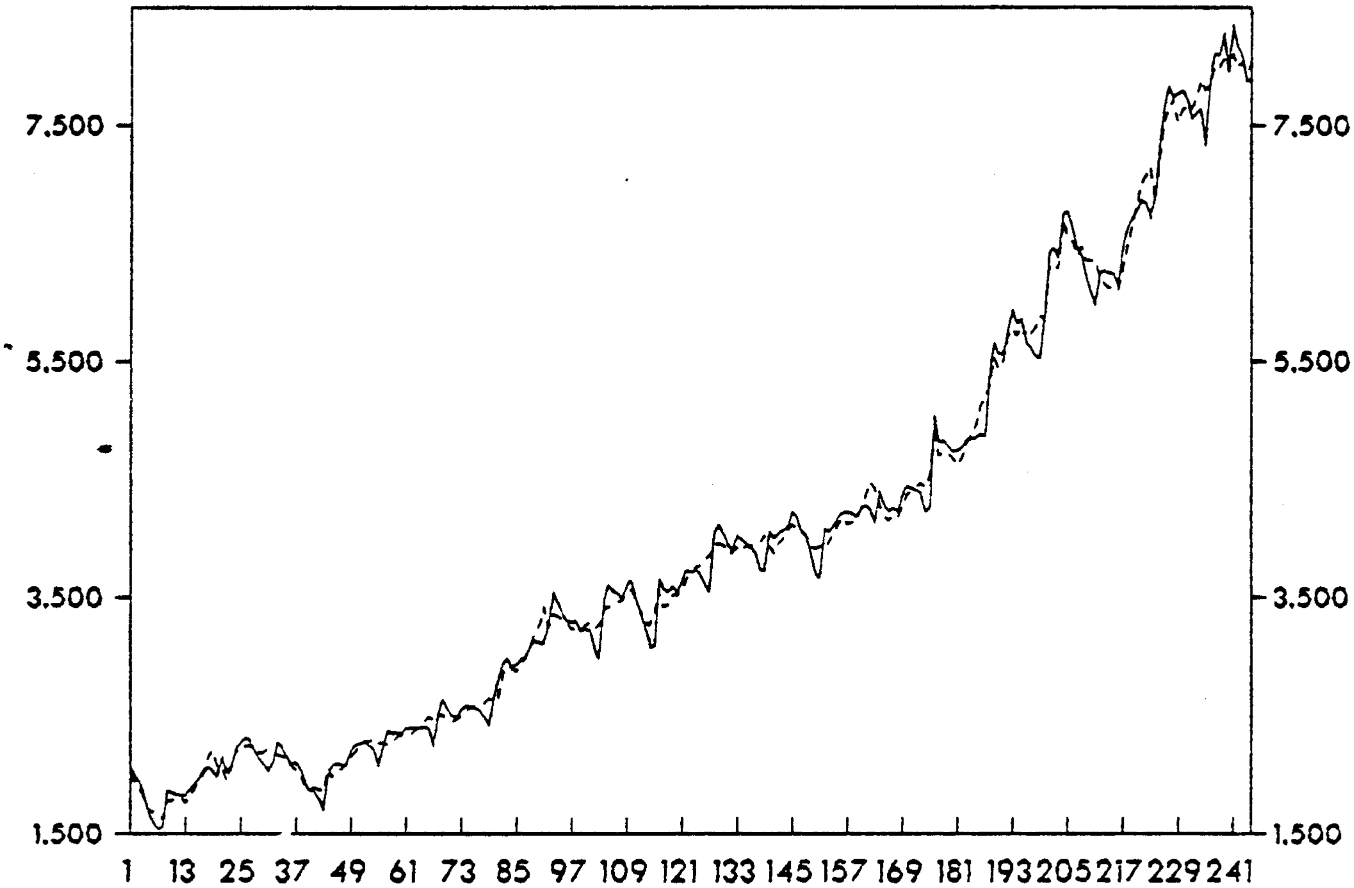


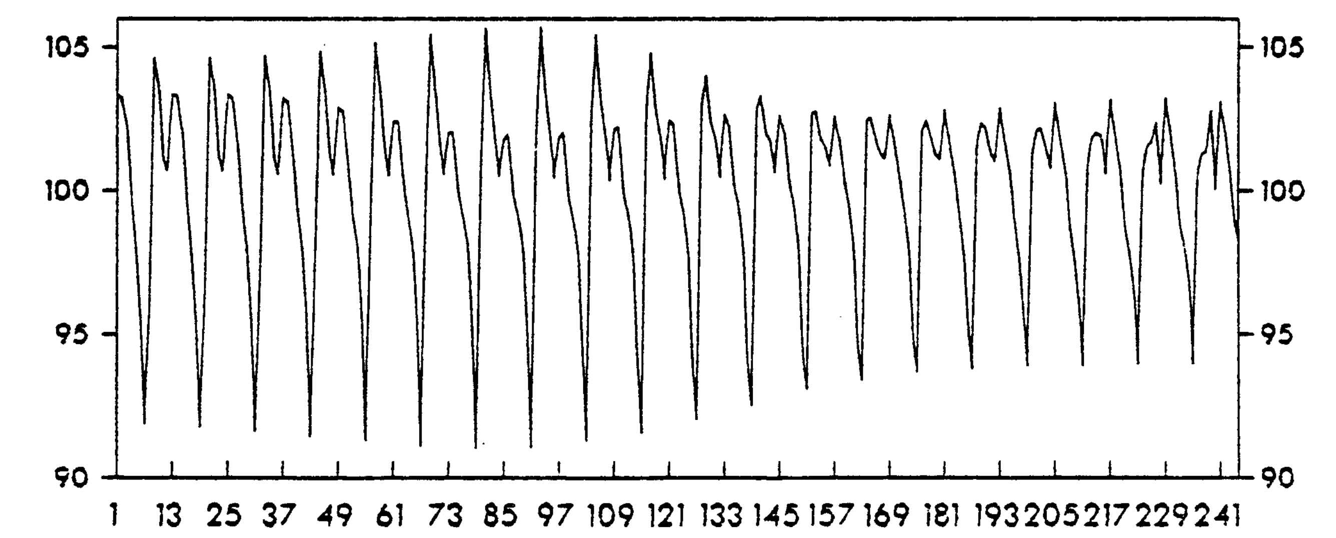
INDUSTRY SX4TI MODEL BASED



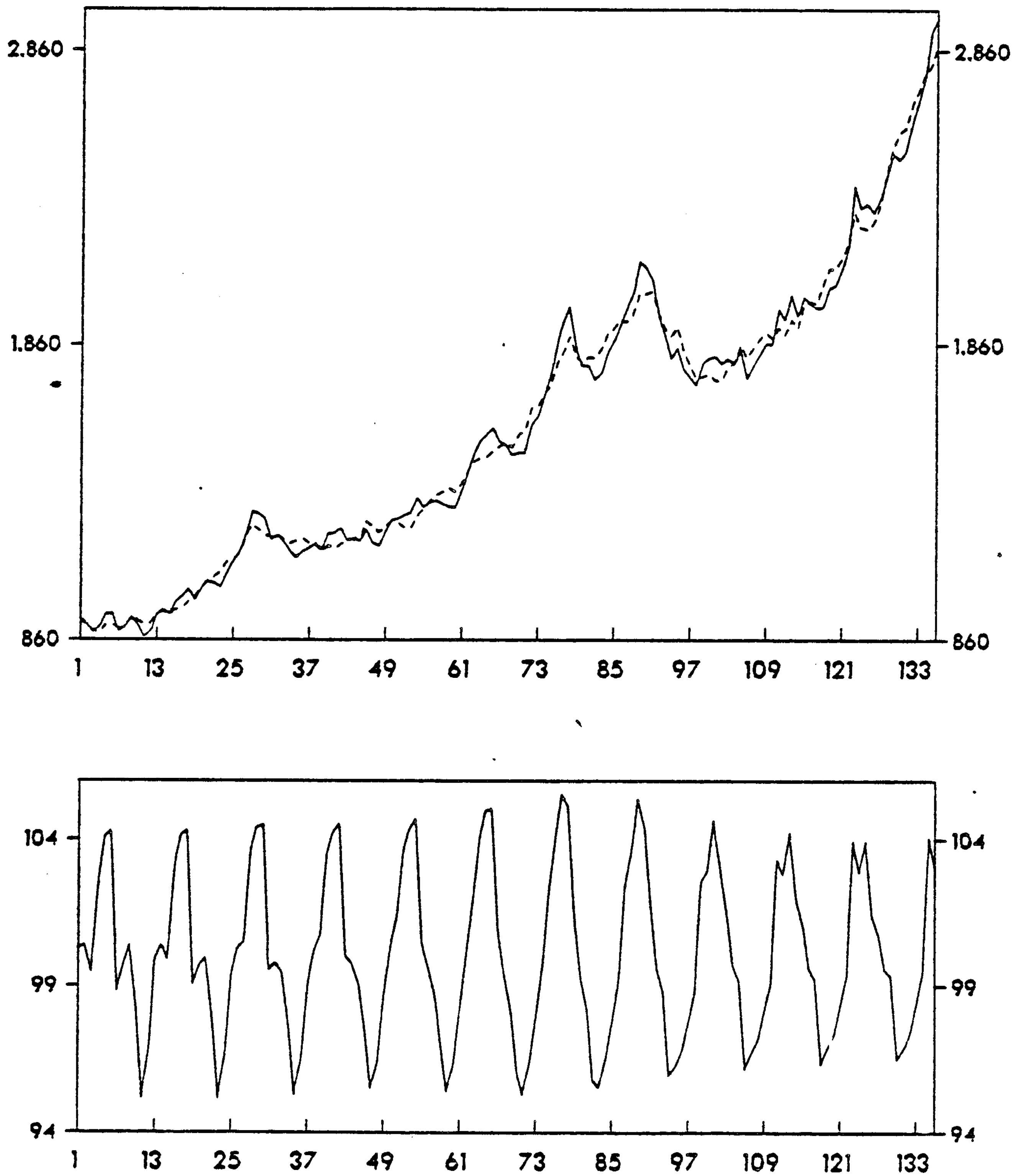


INDUSTRY SX4TI X-11 LOG

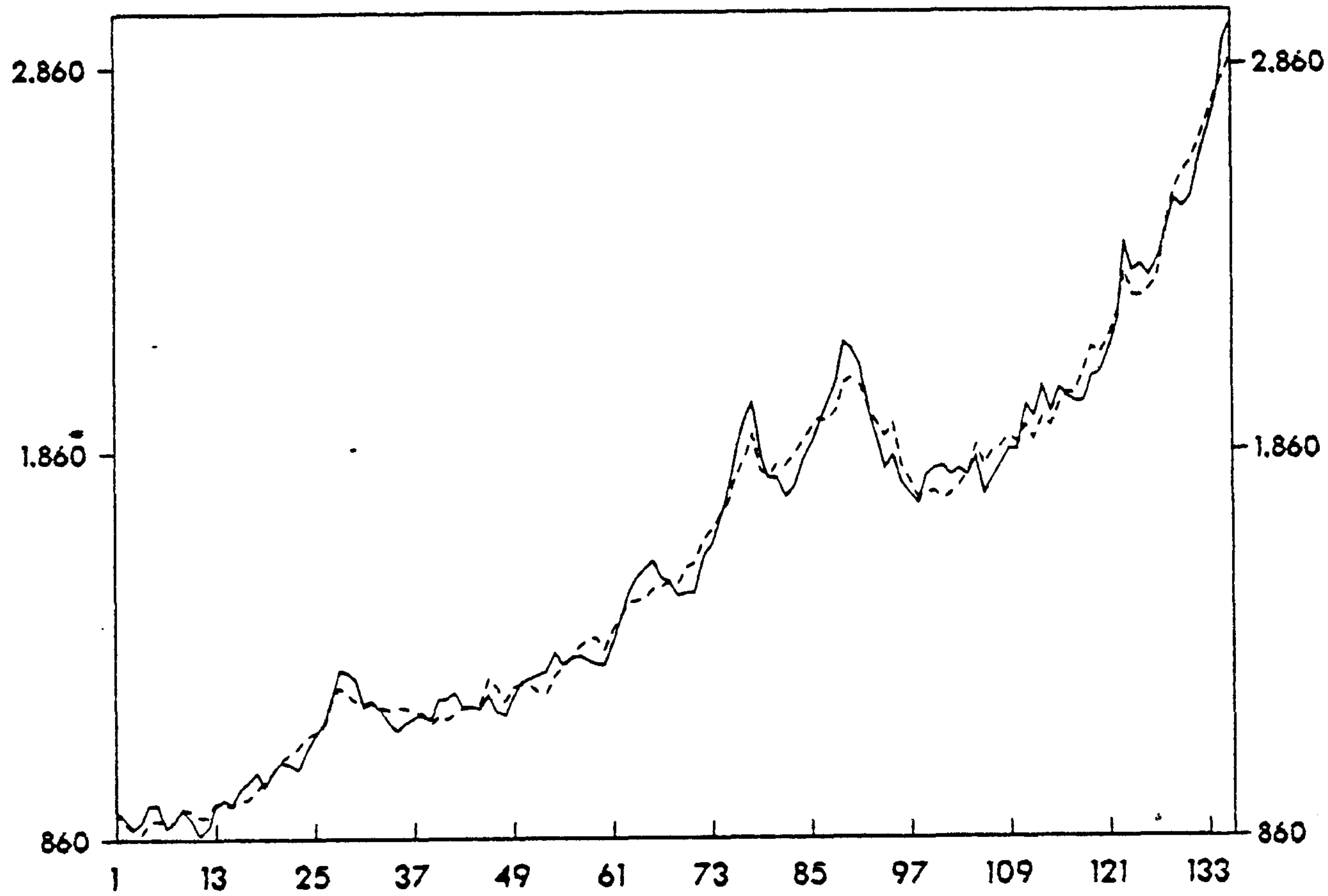


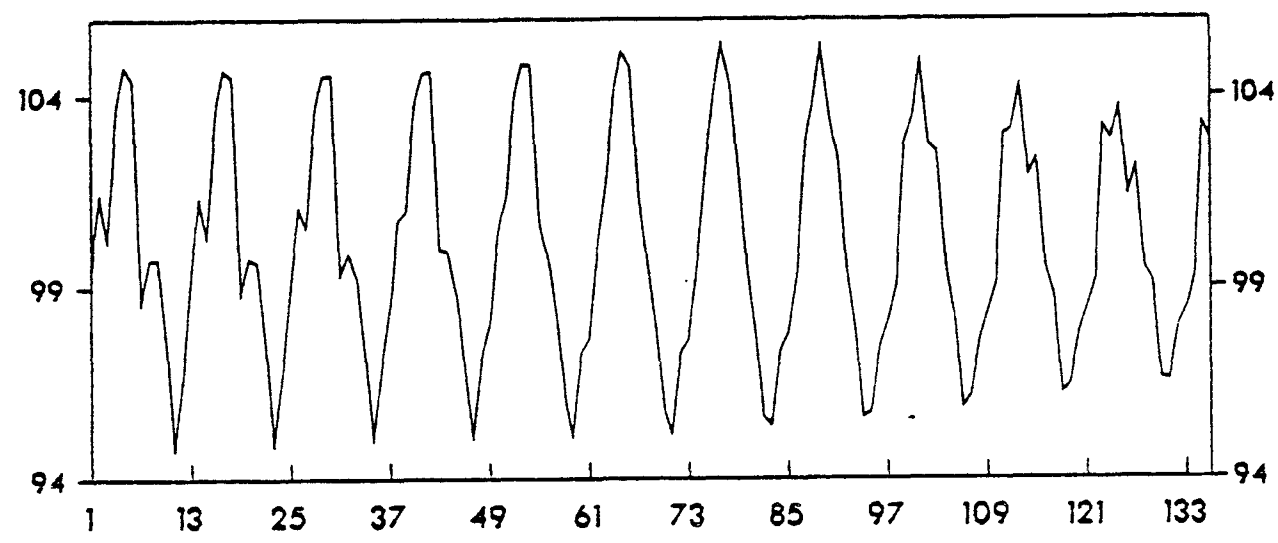


WHOLESALE T1503 MODEL BASED

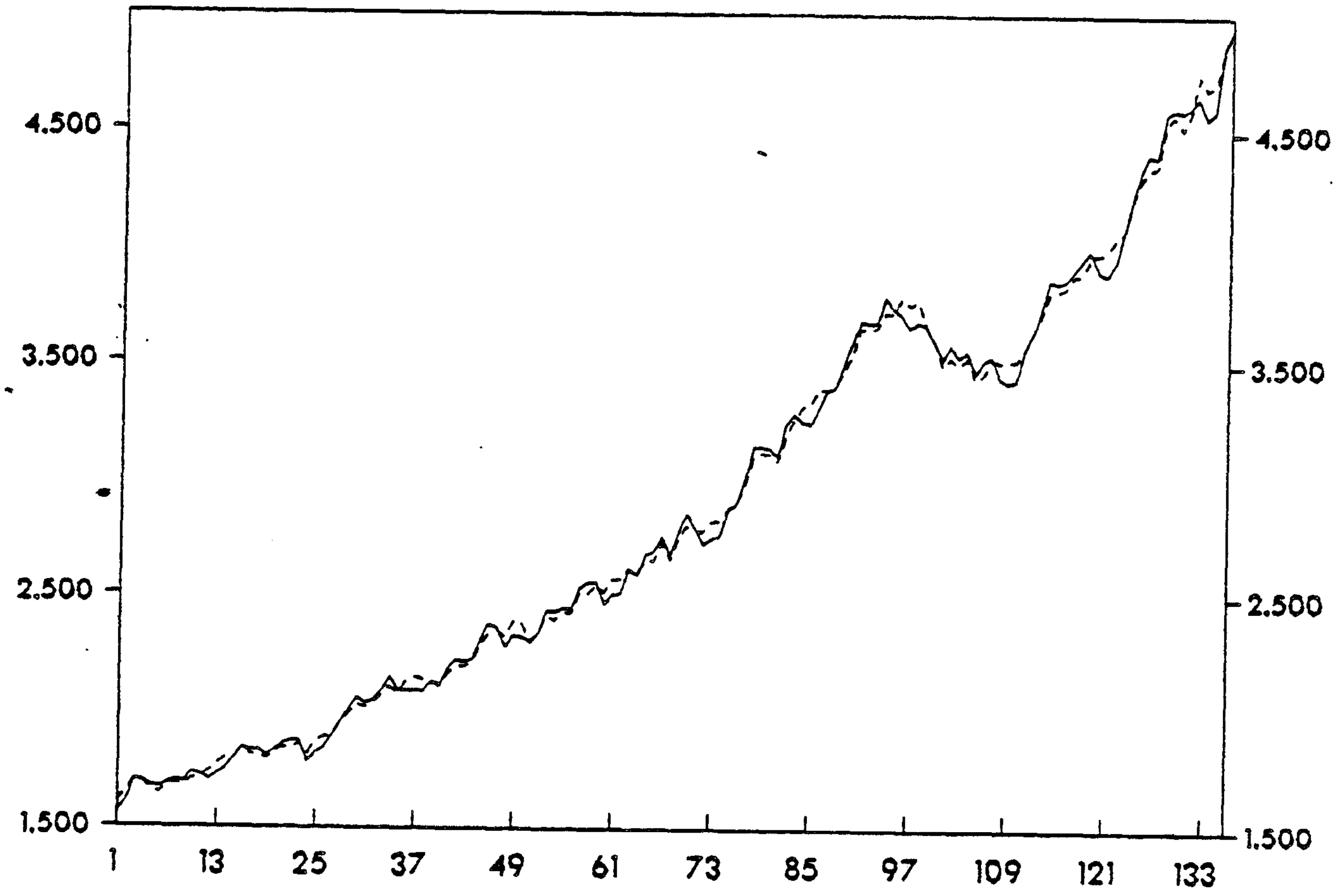


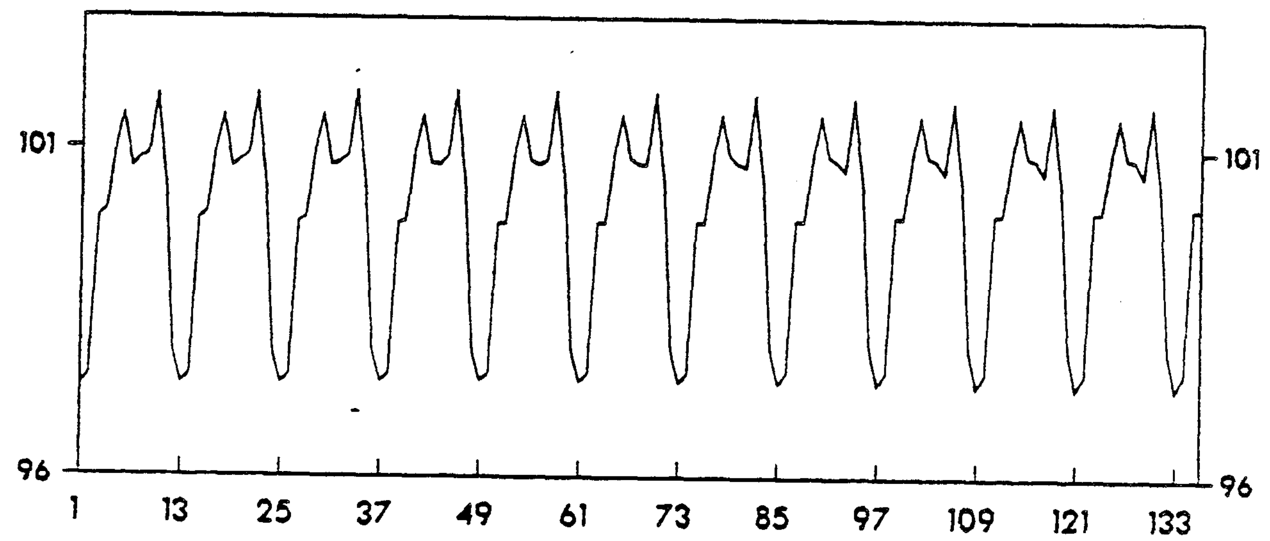
WHOLESALE T1503 X-11 LOG



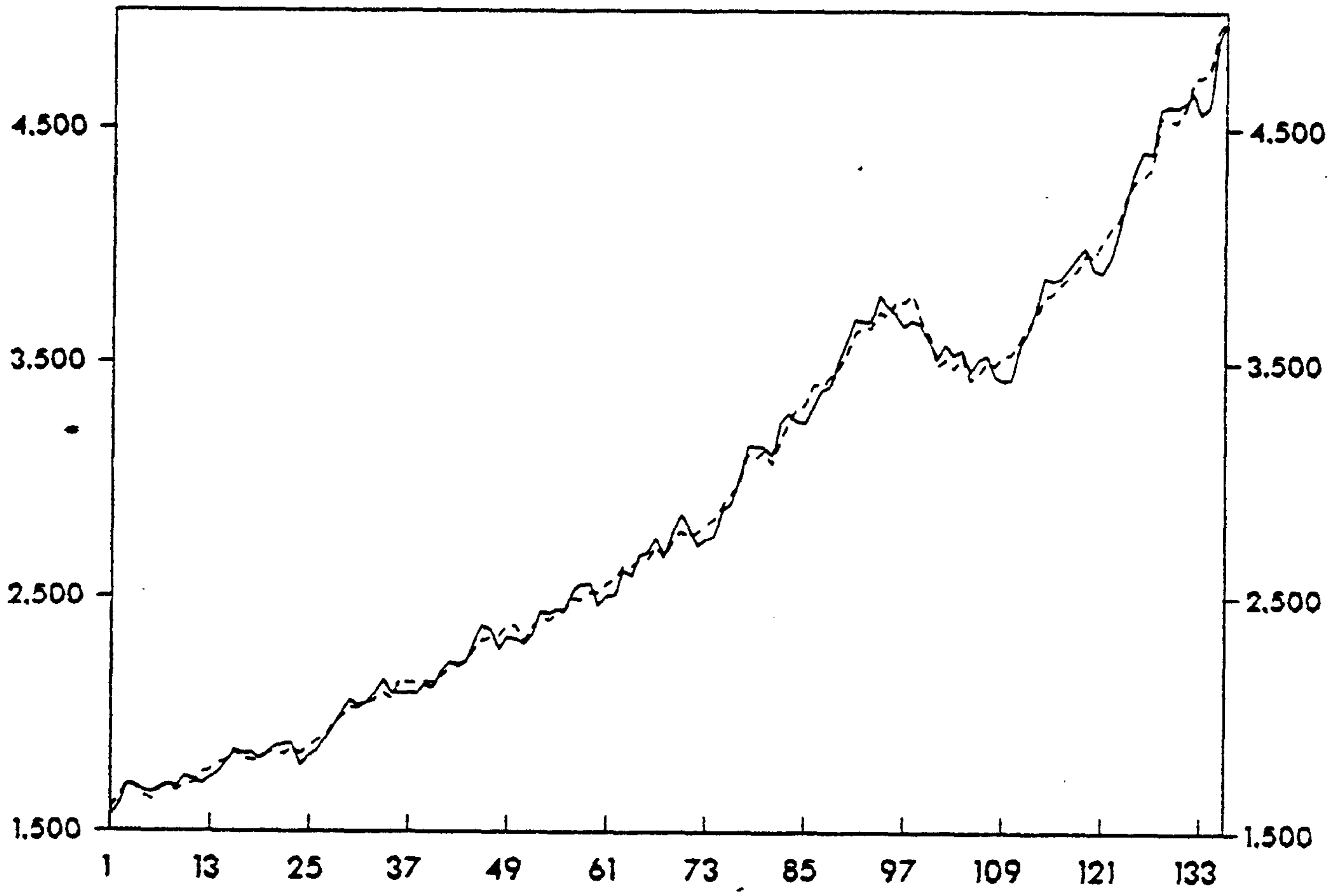


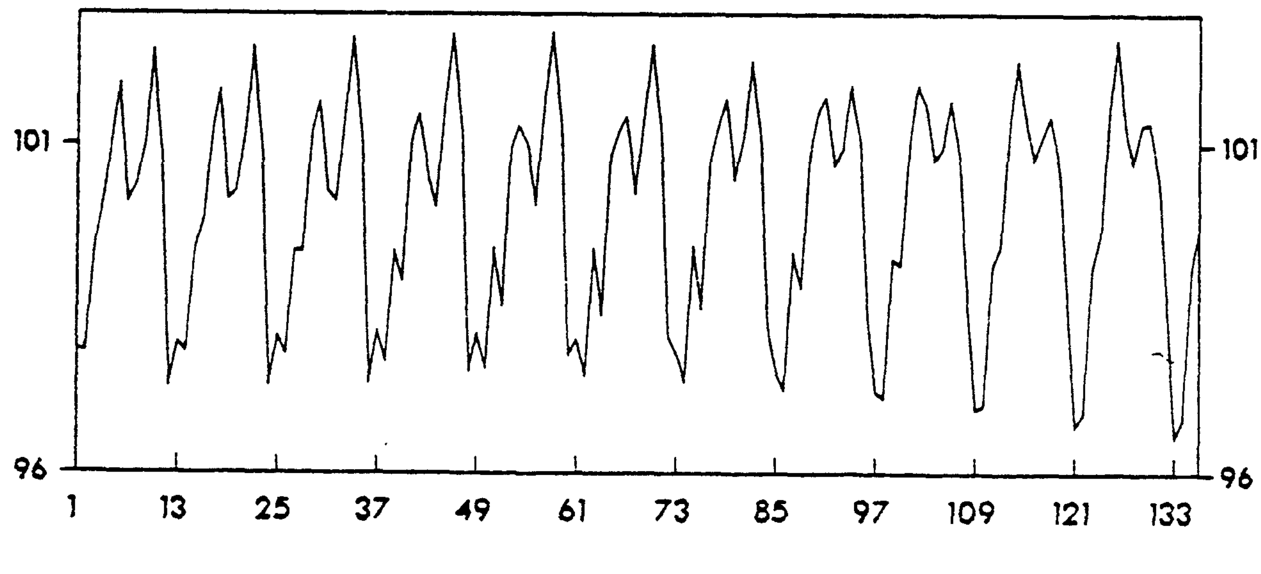
WHOLESALE T1506 MODEL BASED



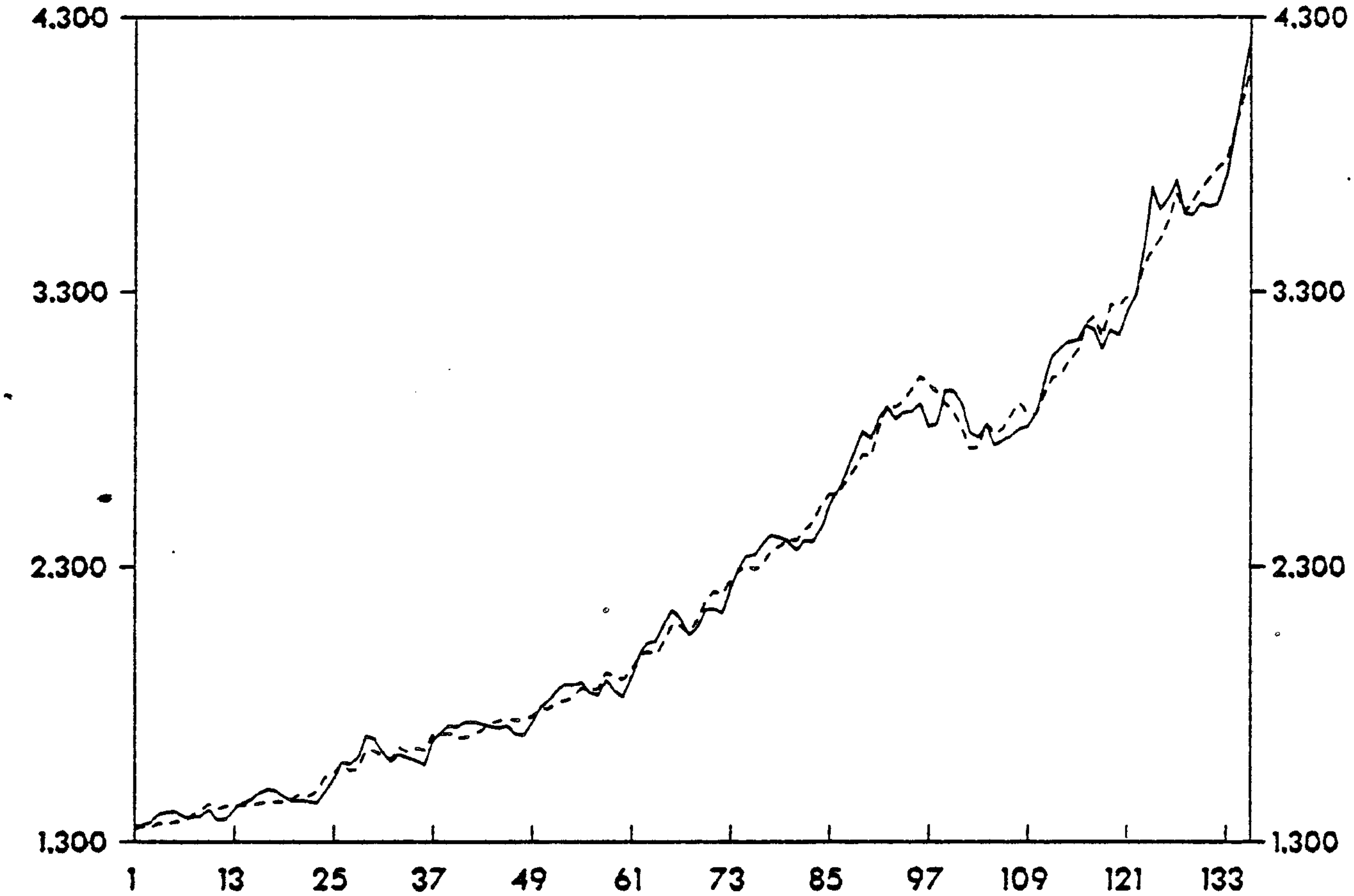


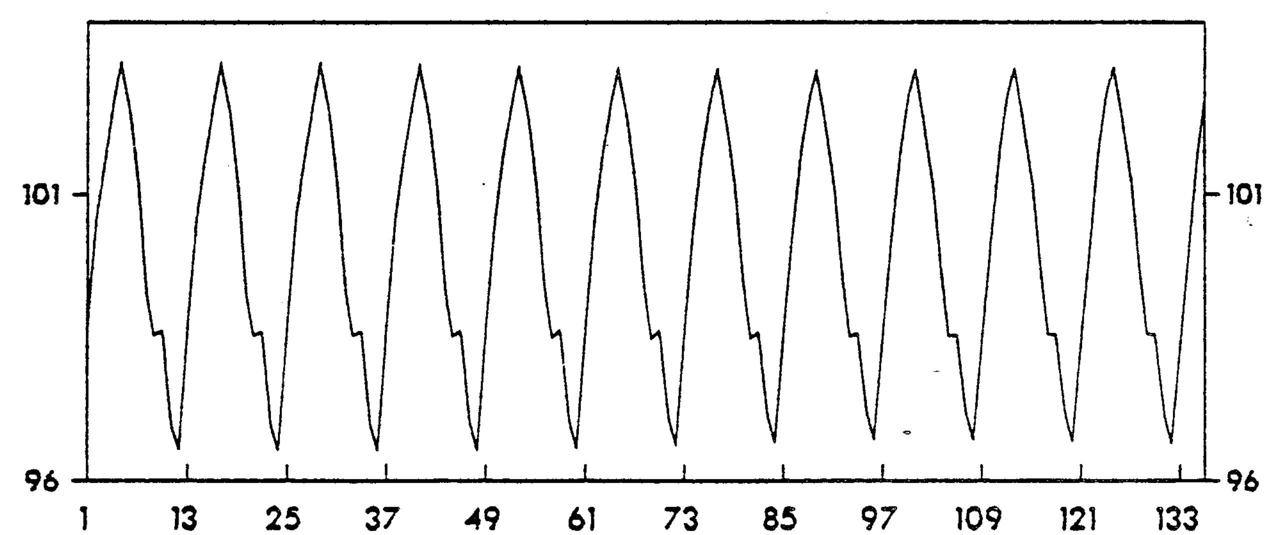
WHOLESALE T1506 X-11 LOG



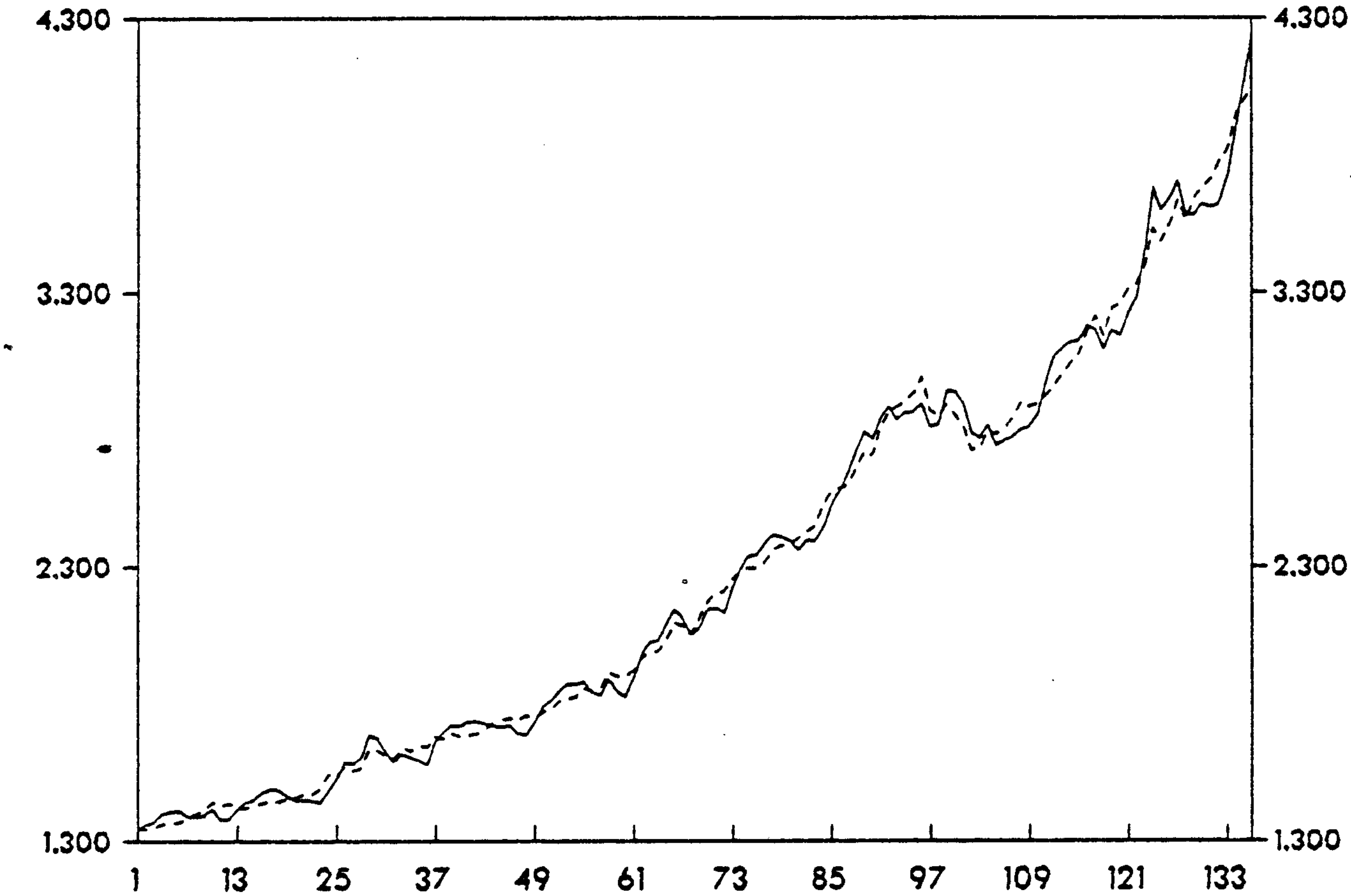


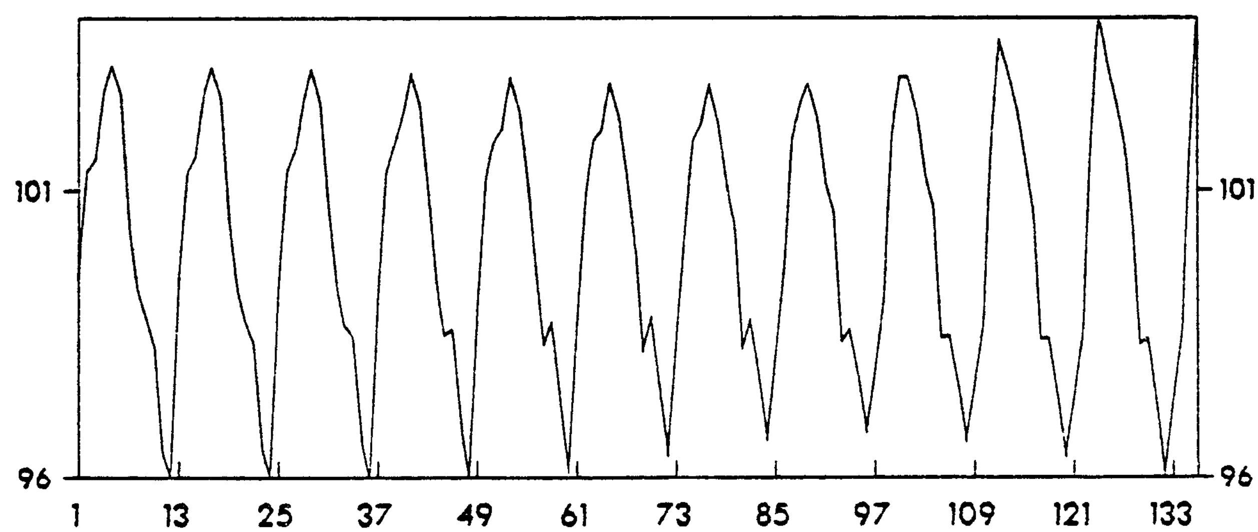
WHOLESALE T1507 MODEL BASED



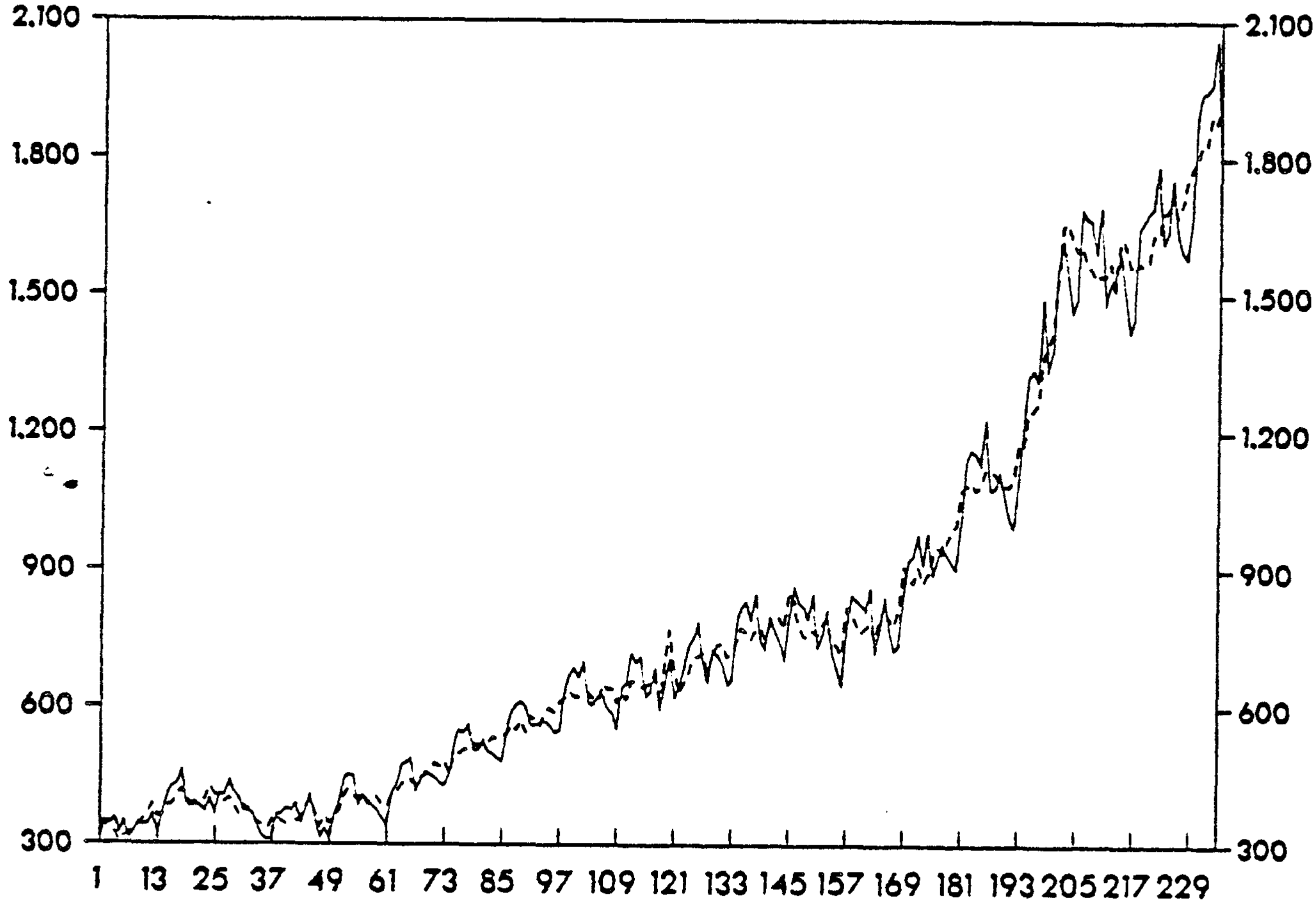


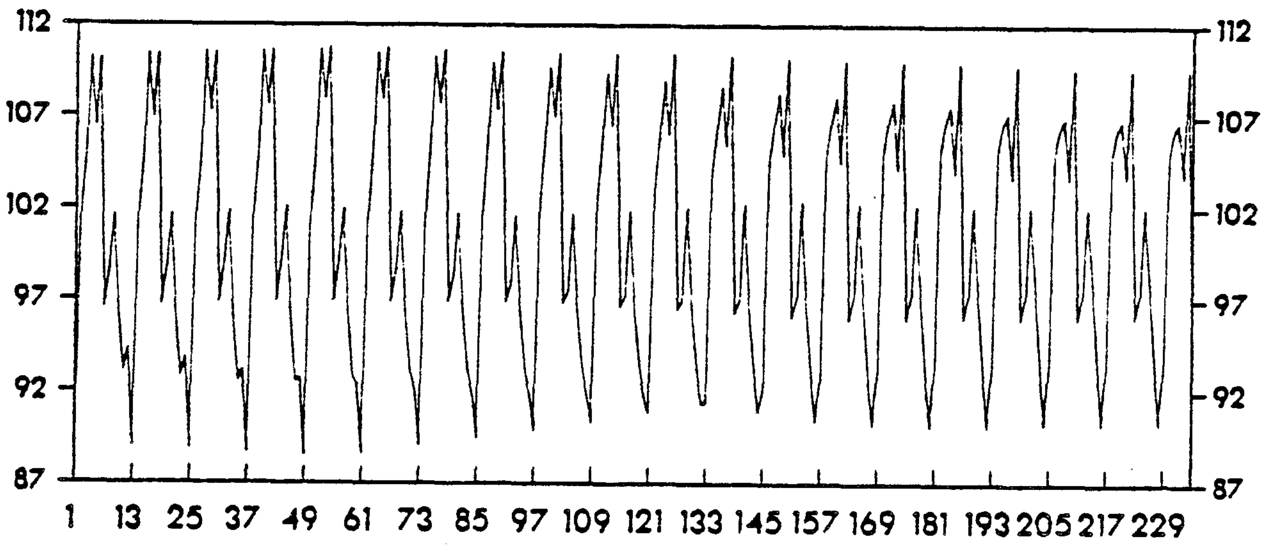
WHOLESALE T1507 X-11 LOG



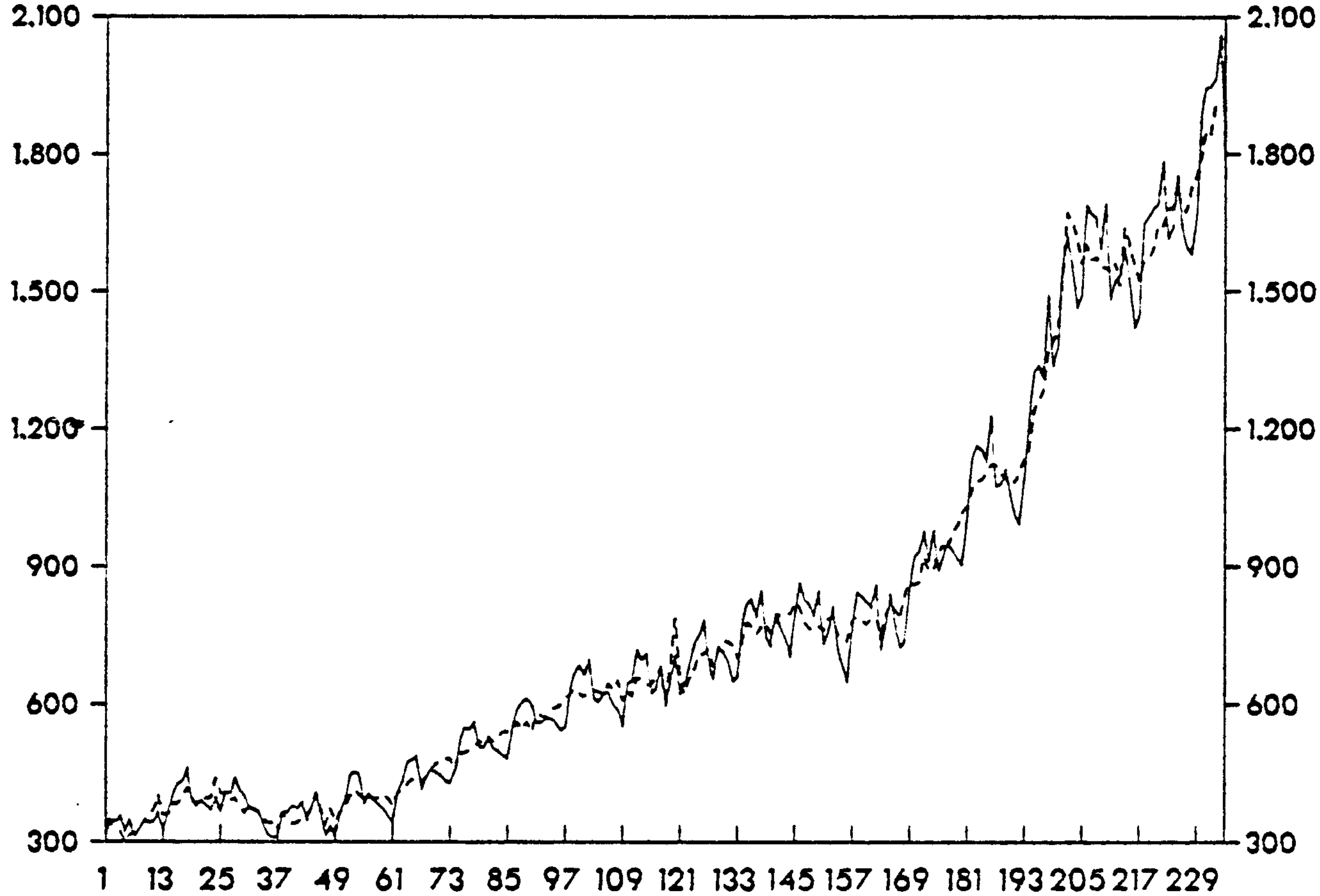


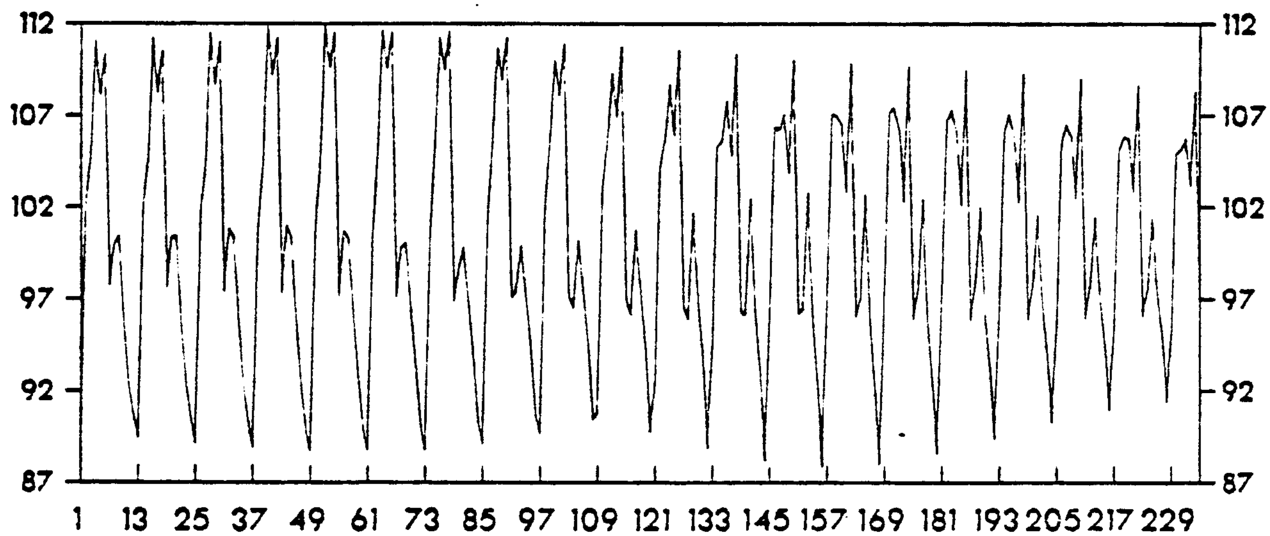
SHIPMENTS MODEL BASED



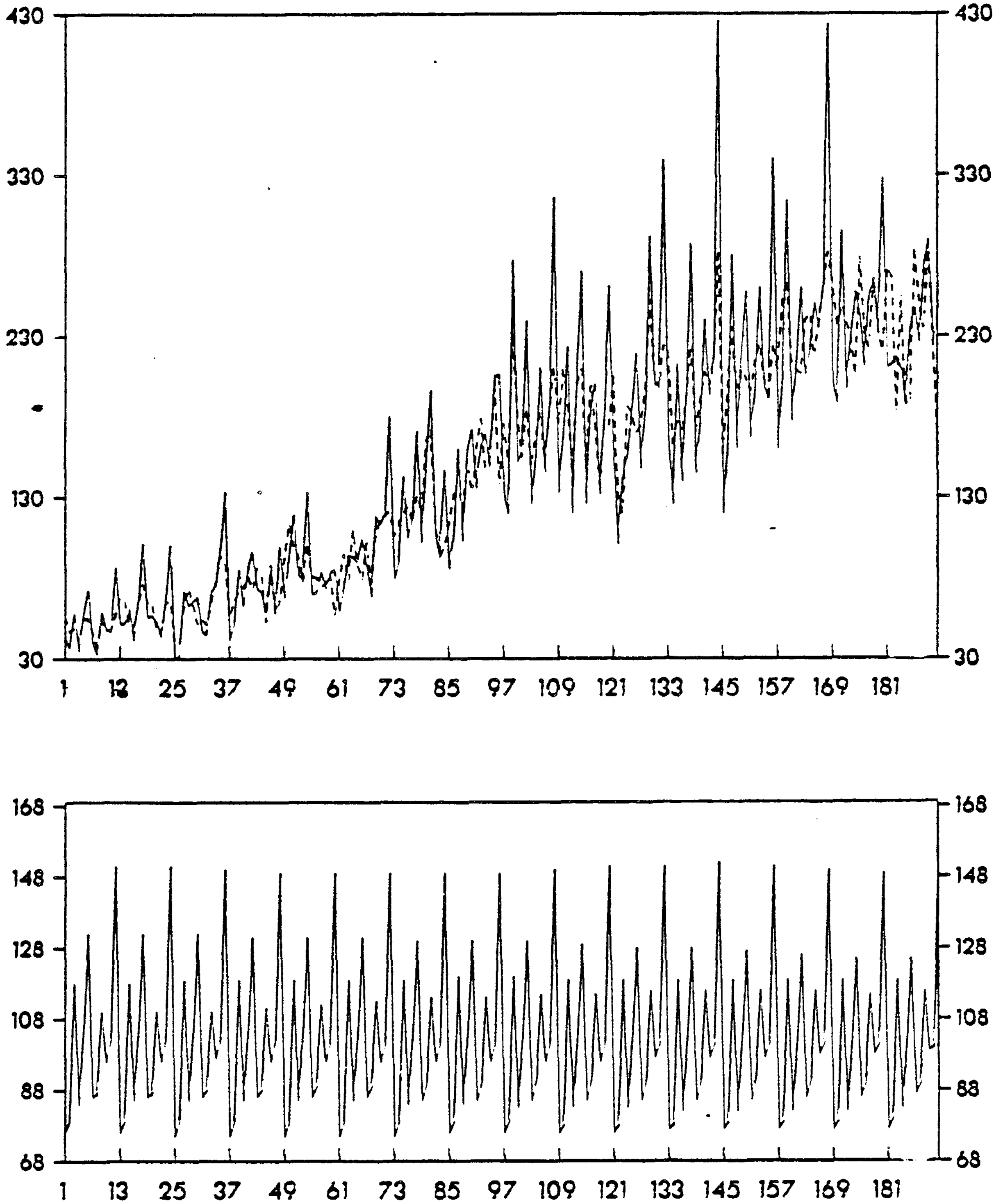


SHIPMENTS X-11 LOG

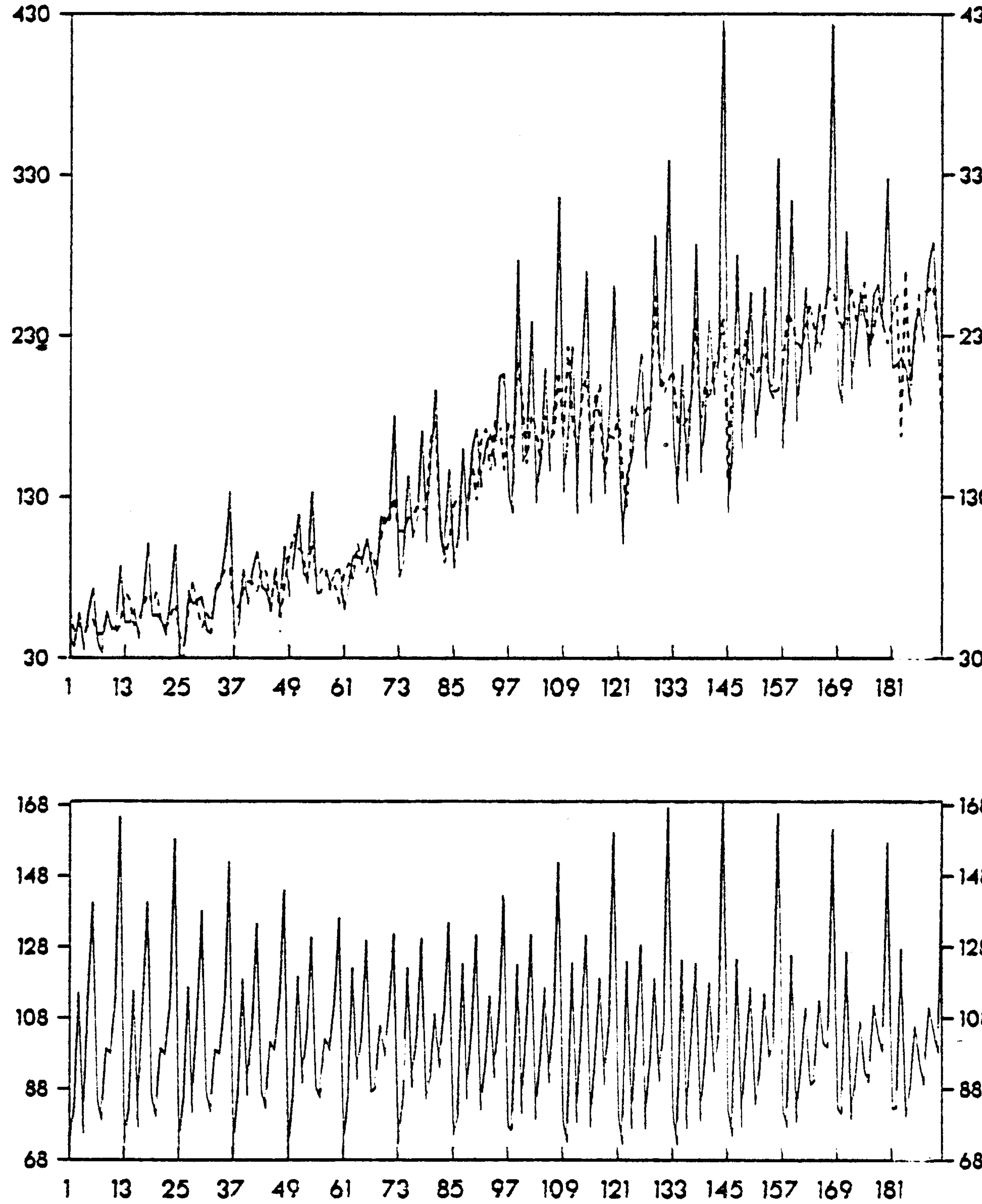




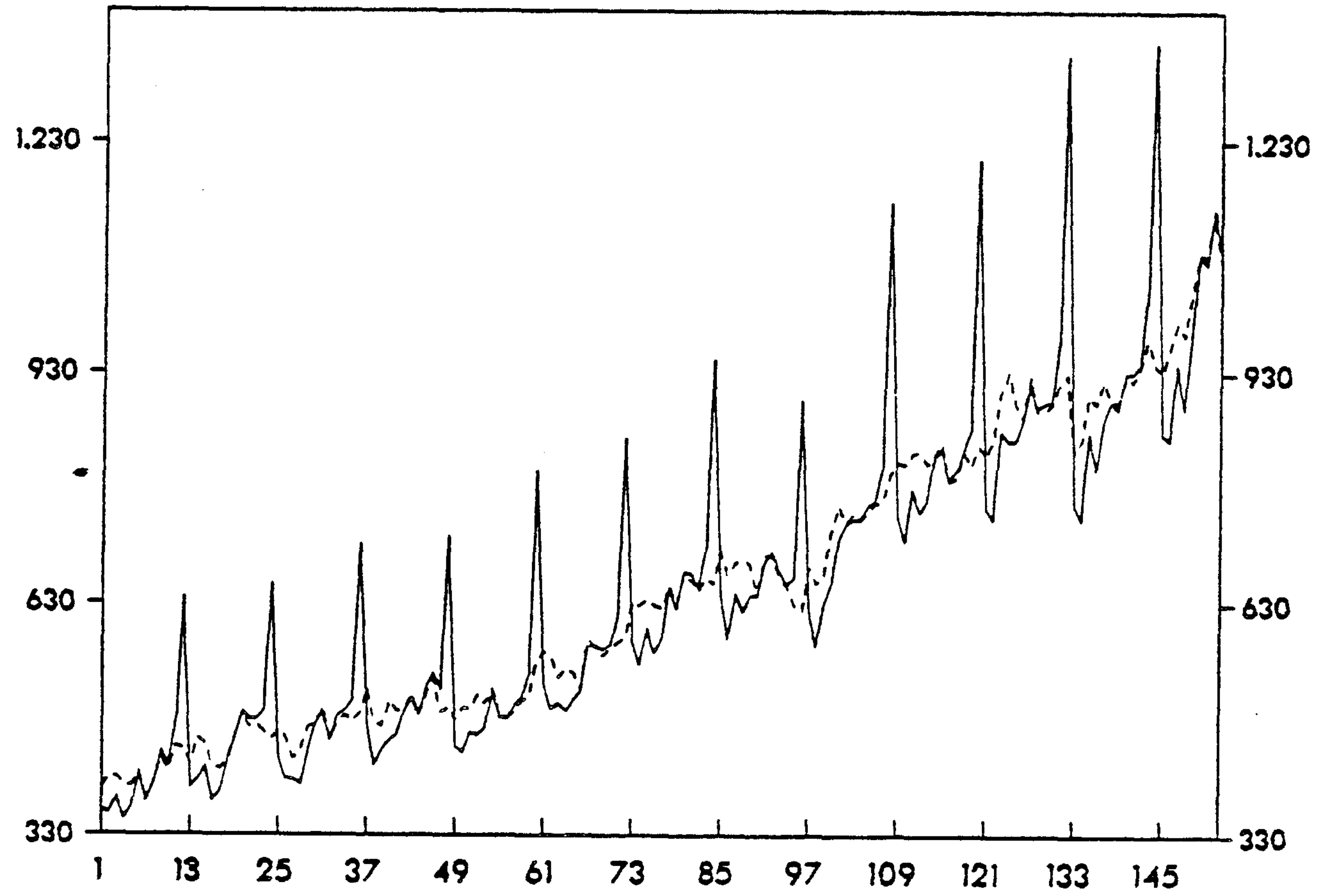
INDUSTRY S21VSU MODEL BASED

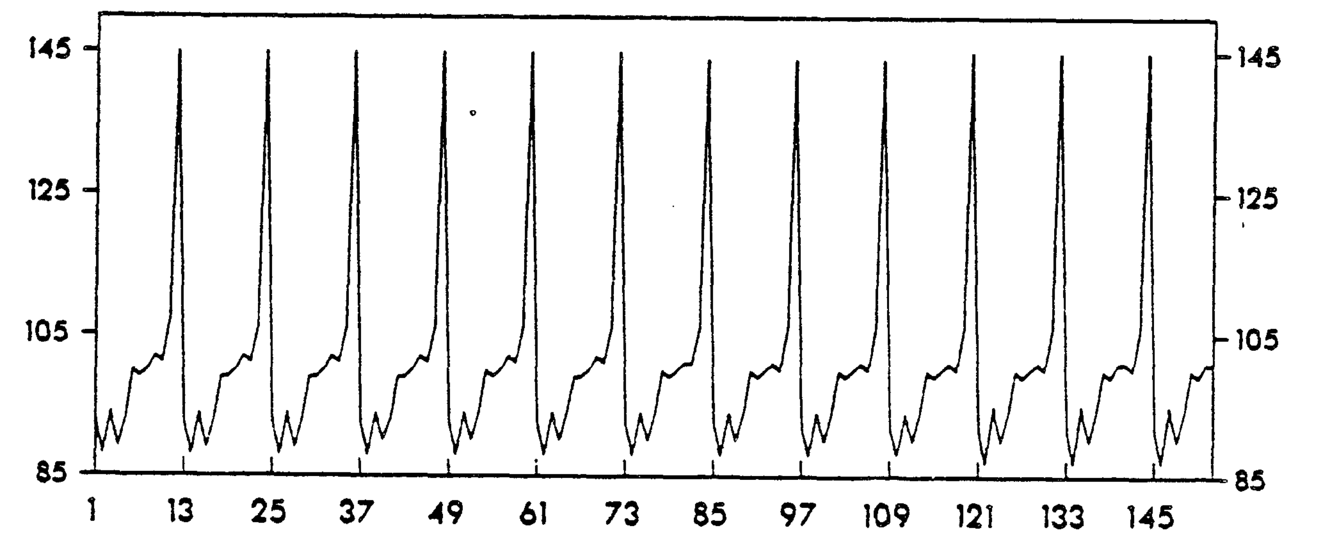


INDUSTRY S21VSU X-11 LOG

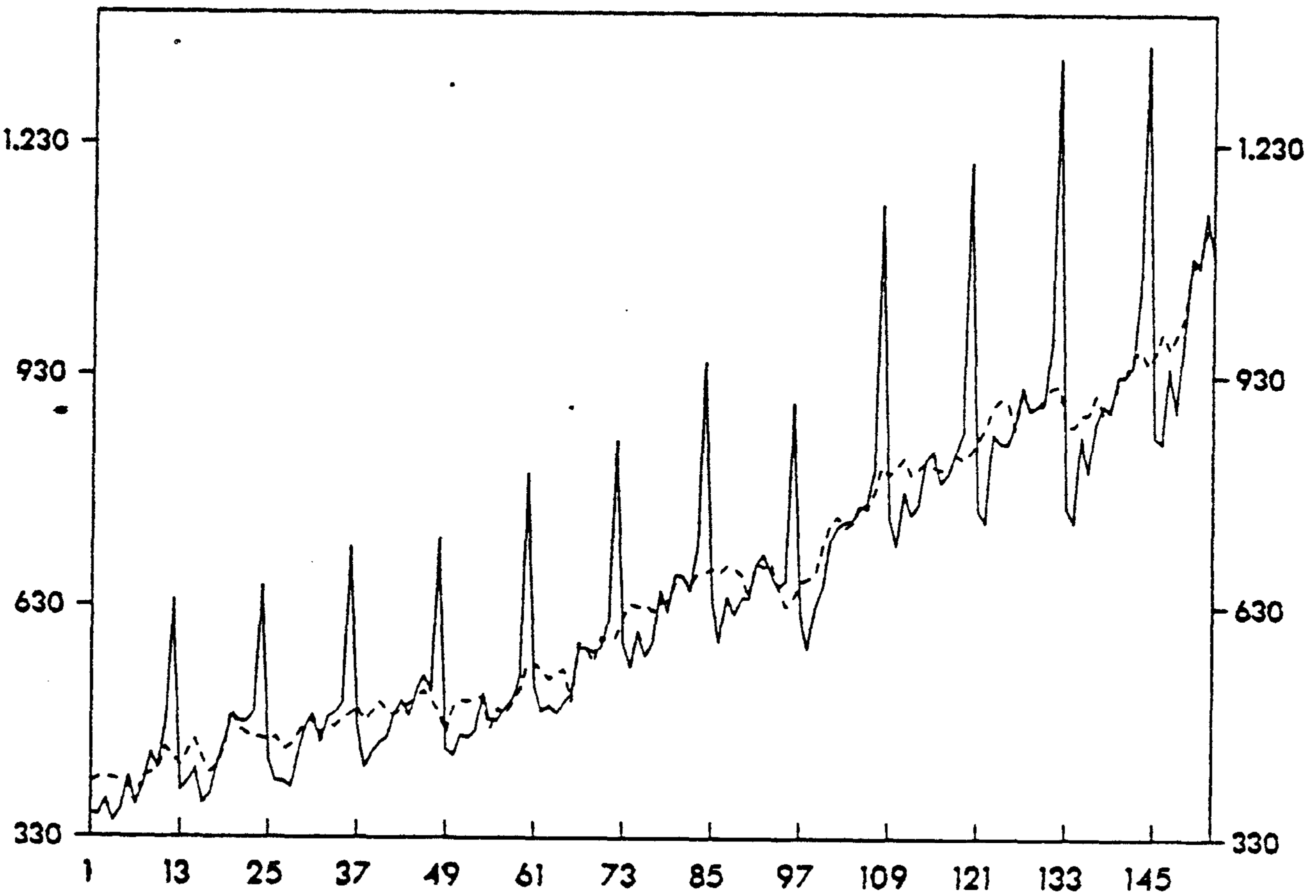


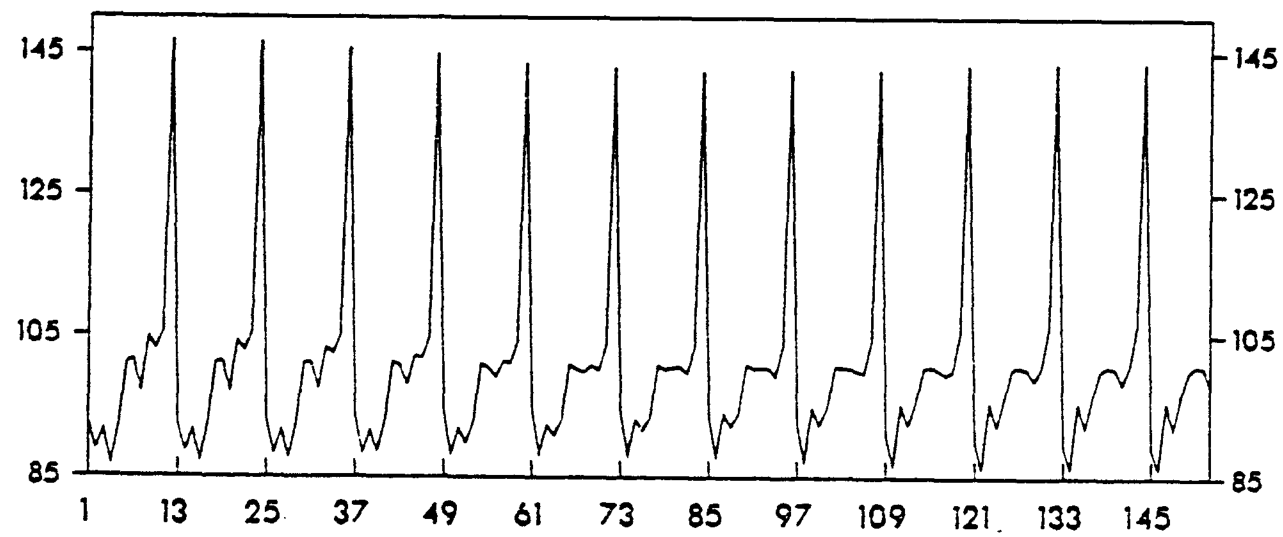
RETAIL 570002 MODEL BASED



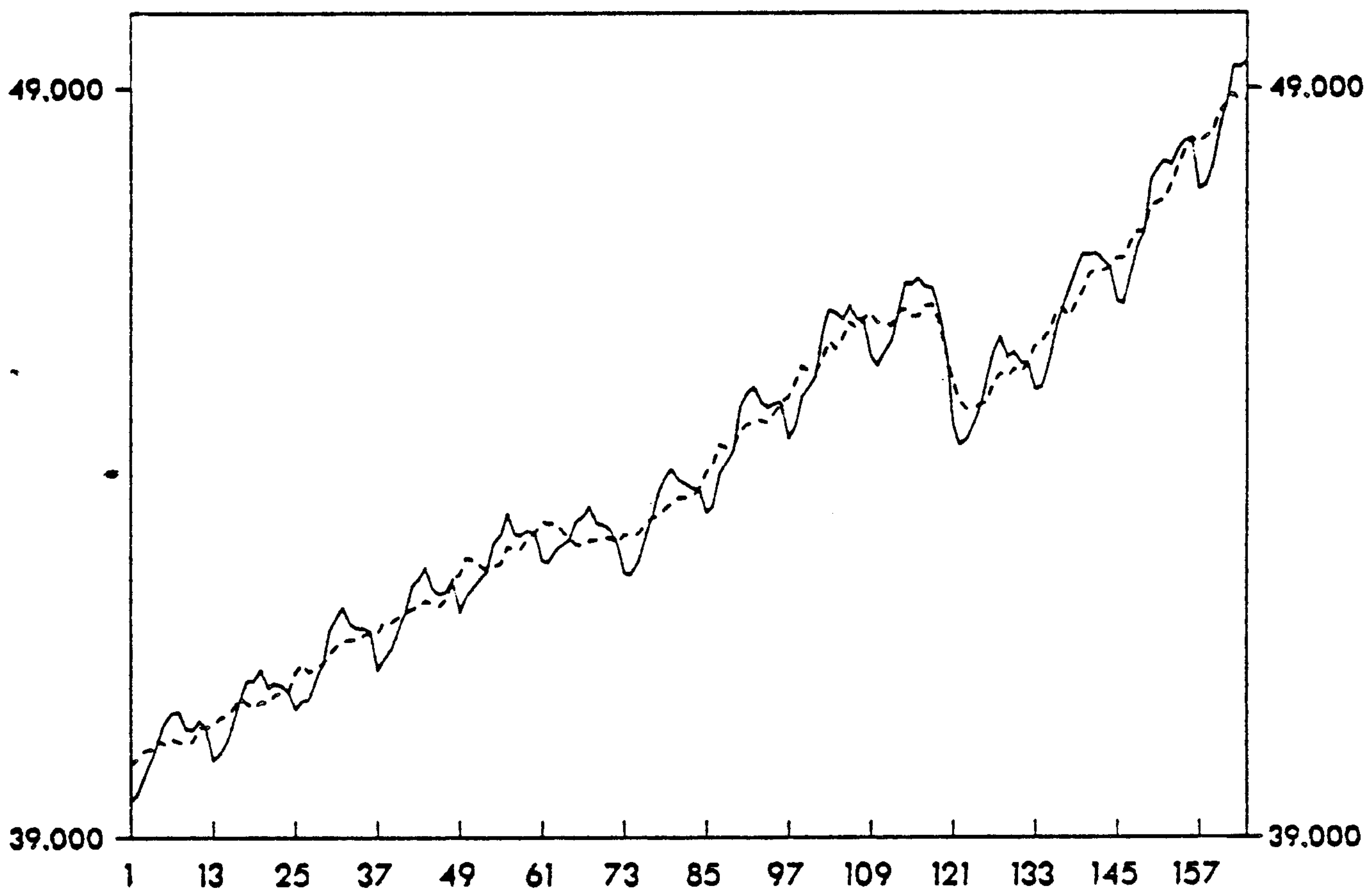


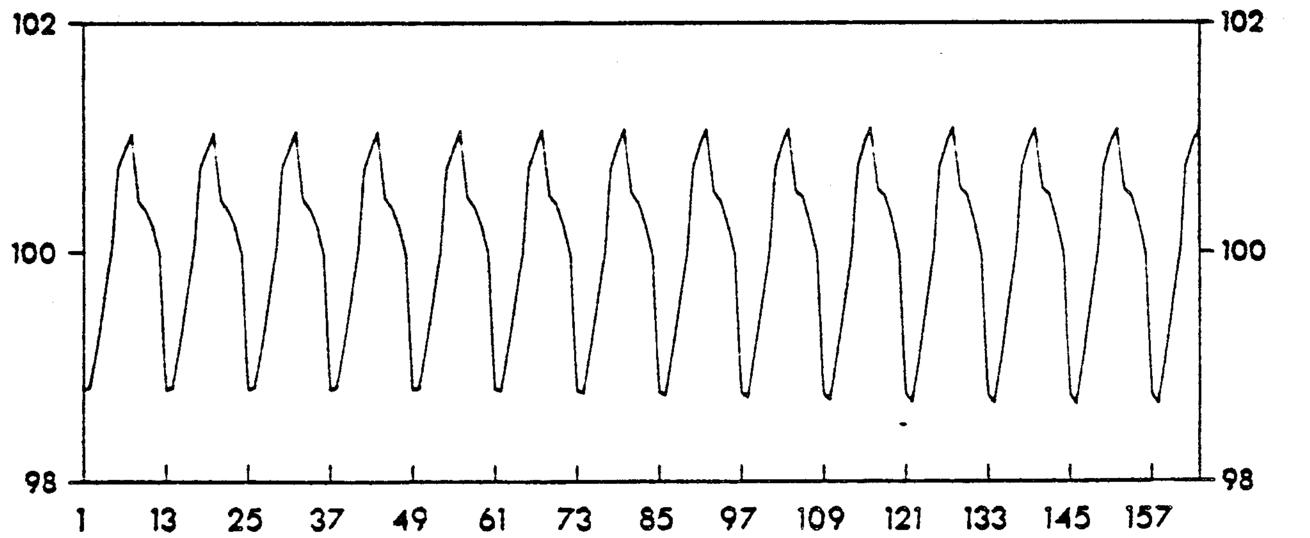
RETAIL 570002 X-11



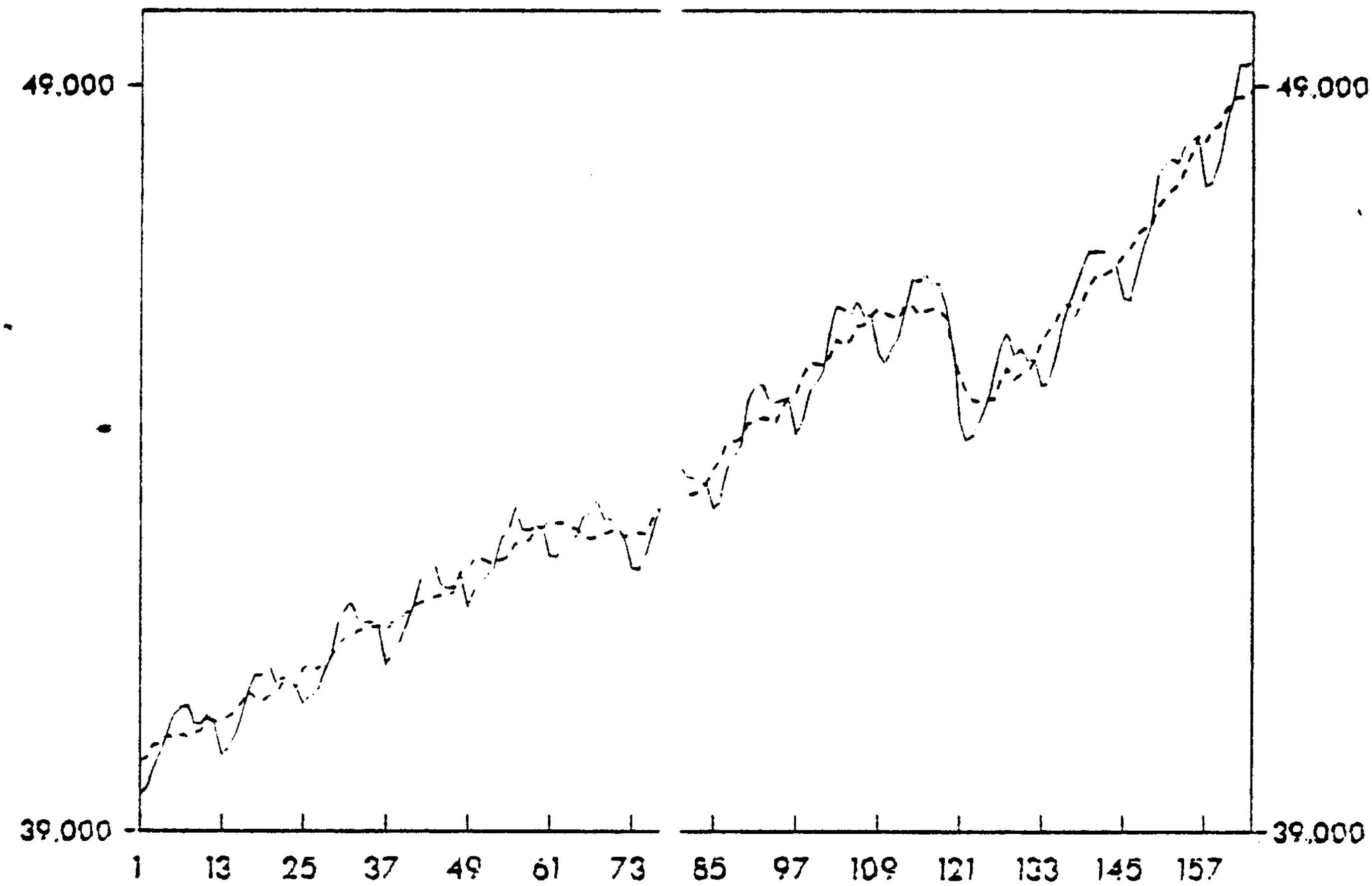


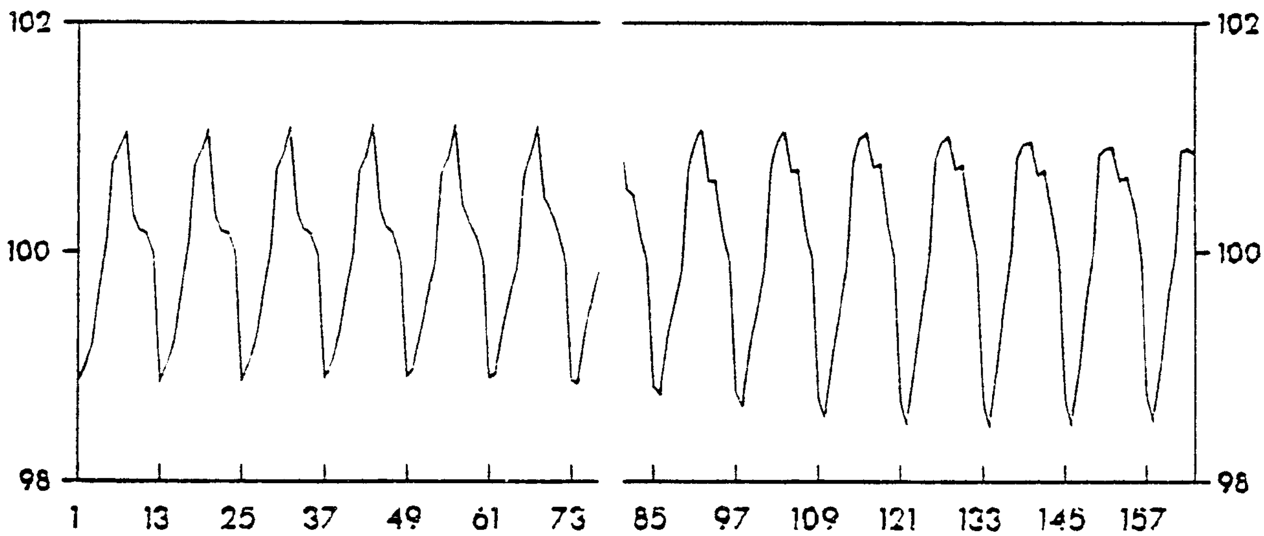
ENM20 MODEL BASED



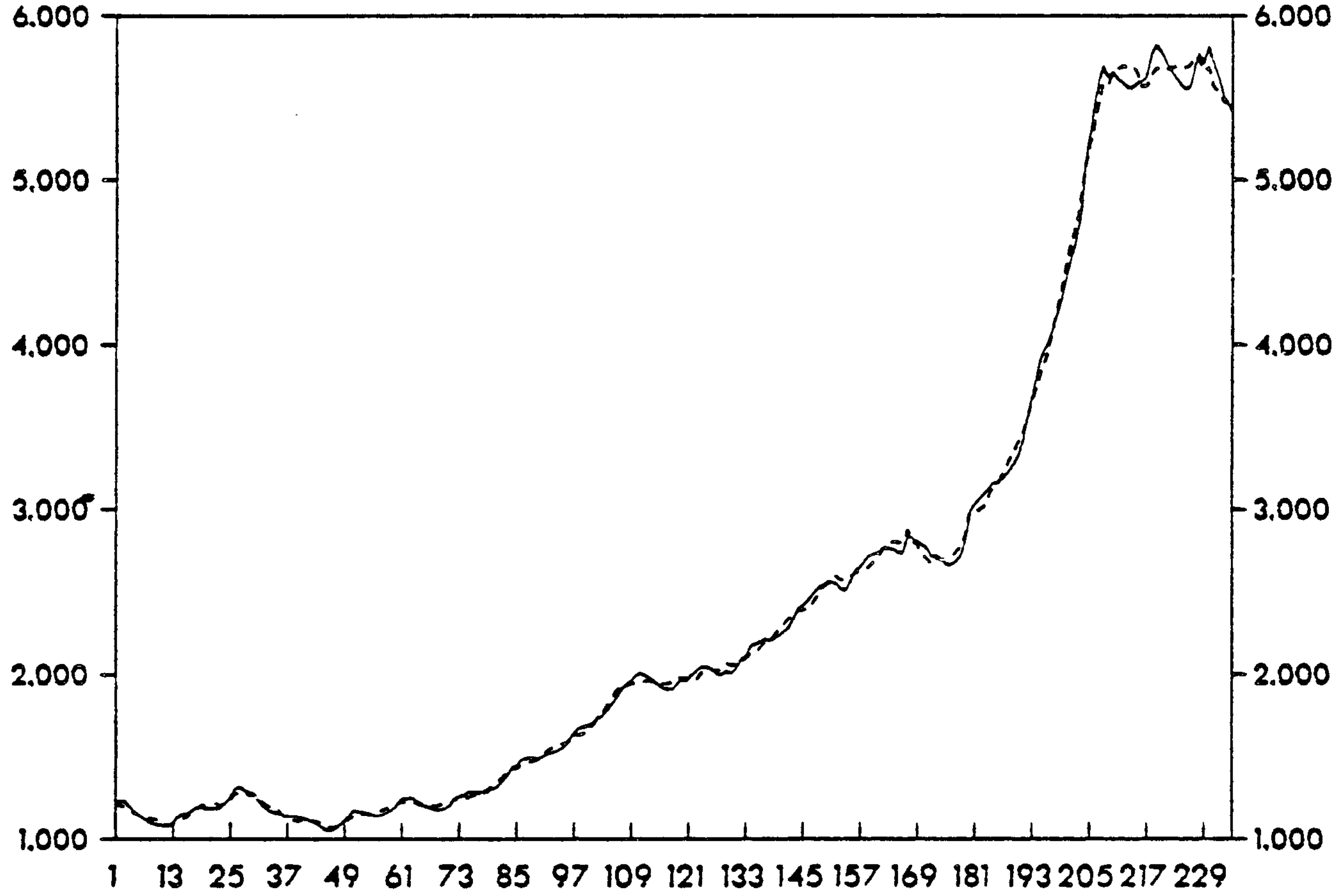


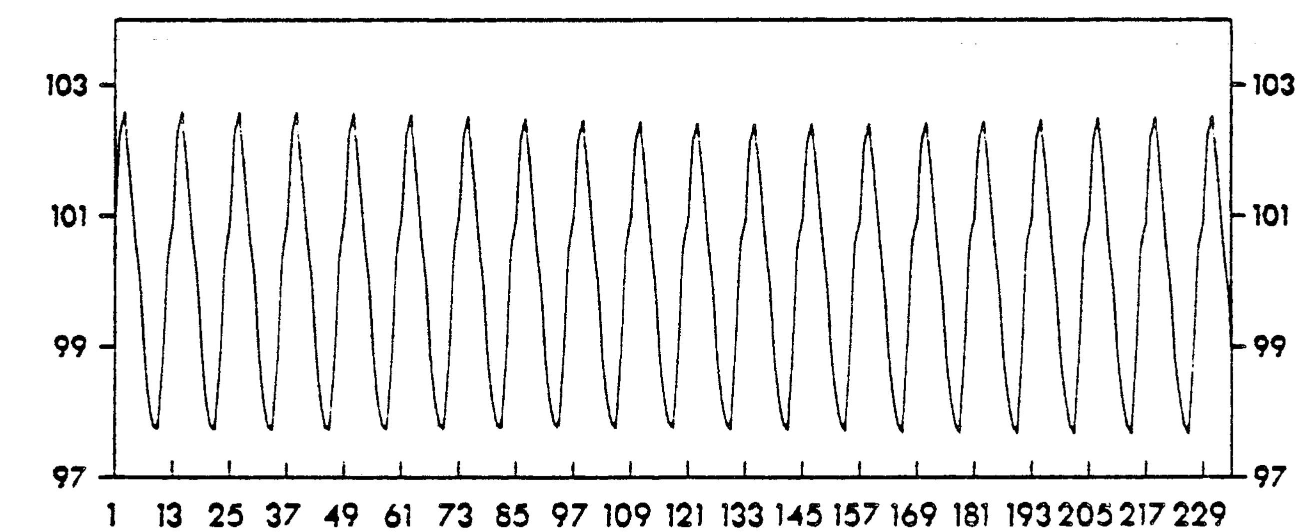
ENM20 X-11 LOG



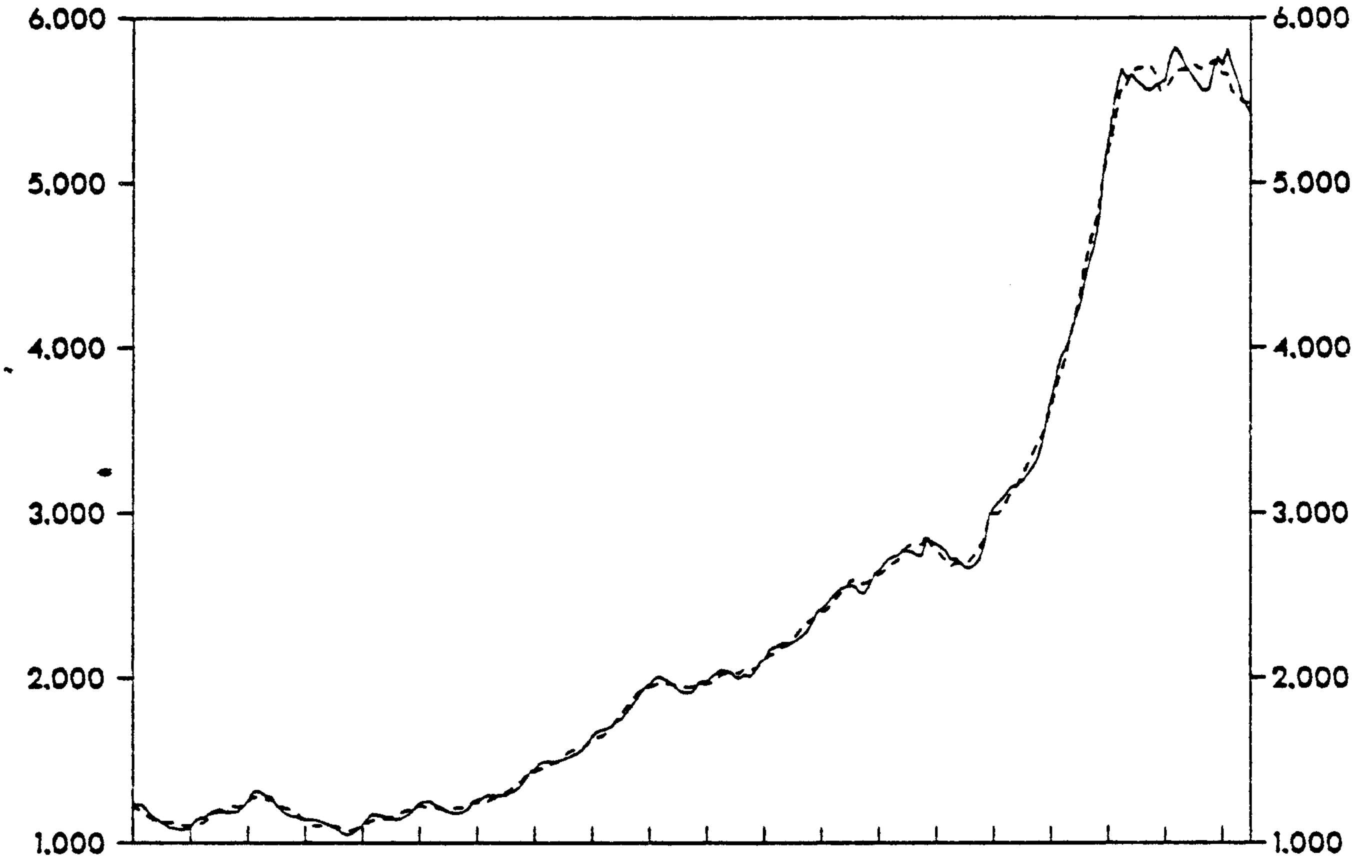


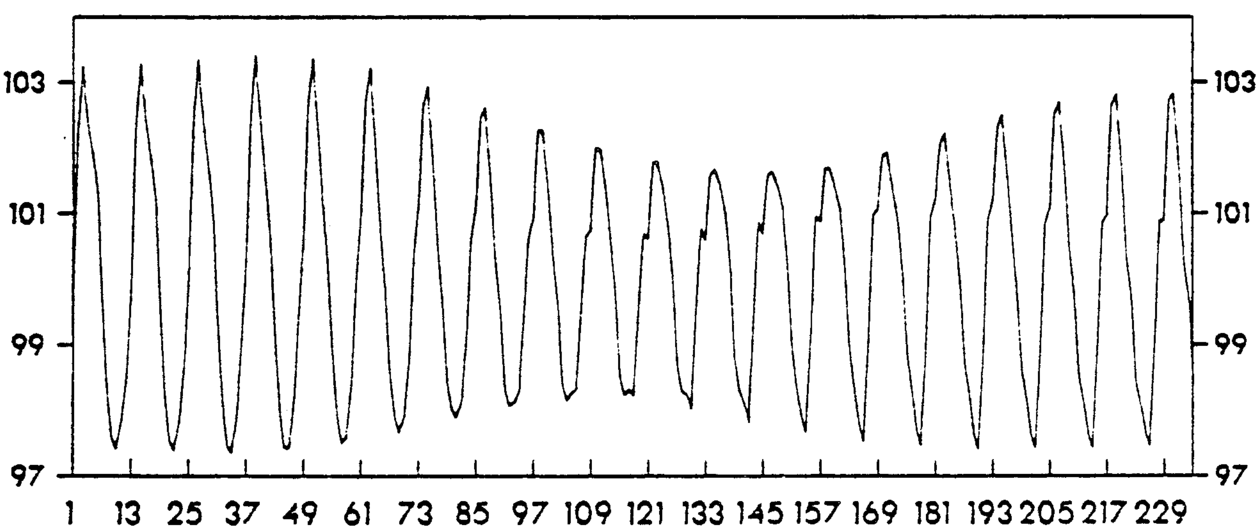
INVENTORIES MODEL BASED



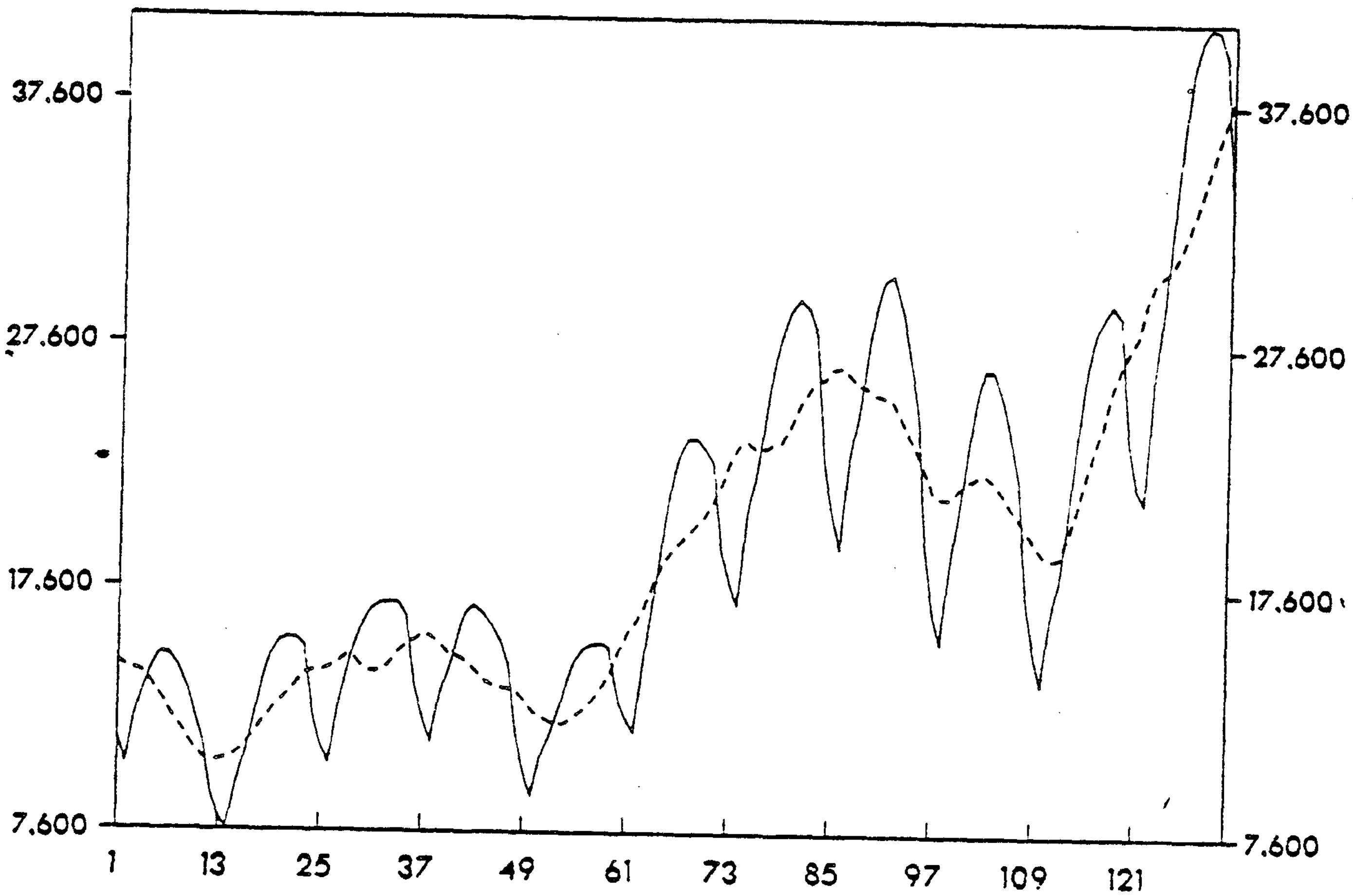


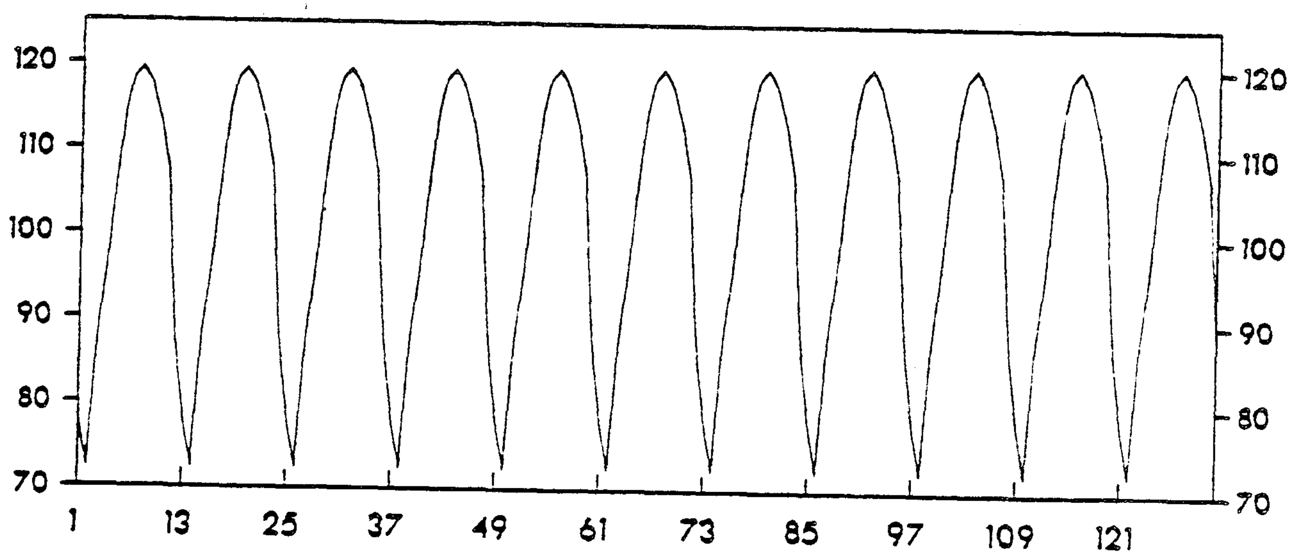
INVENTORIES X-11 LOG



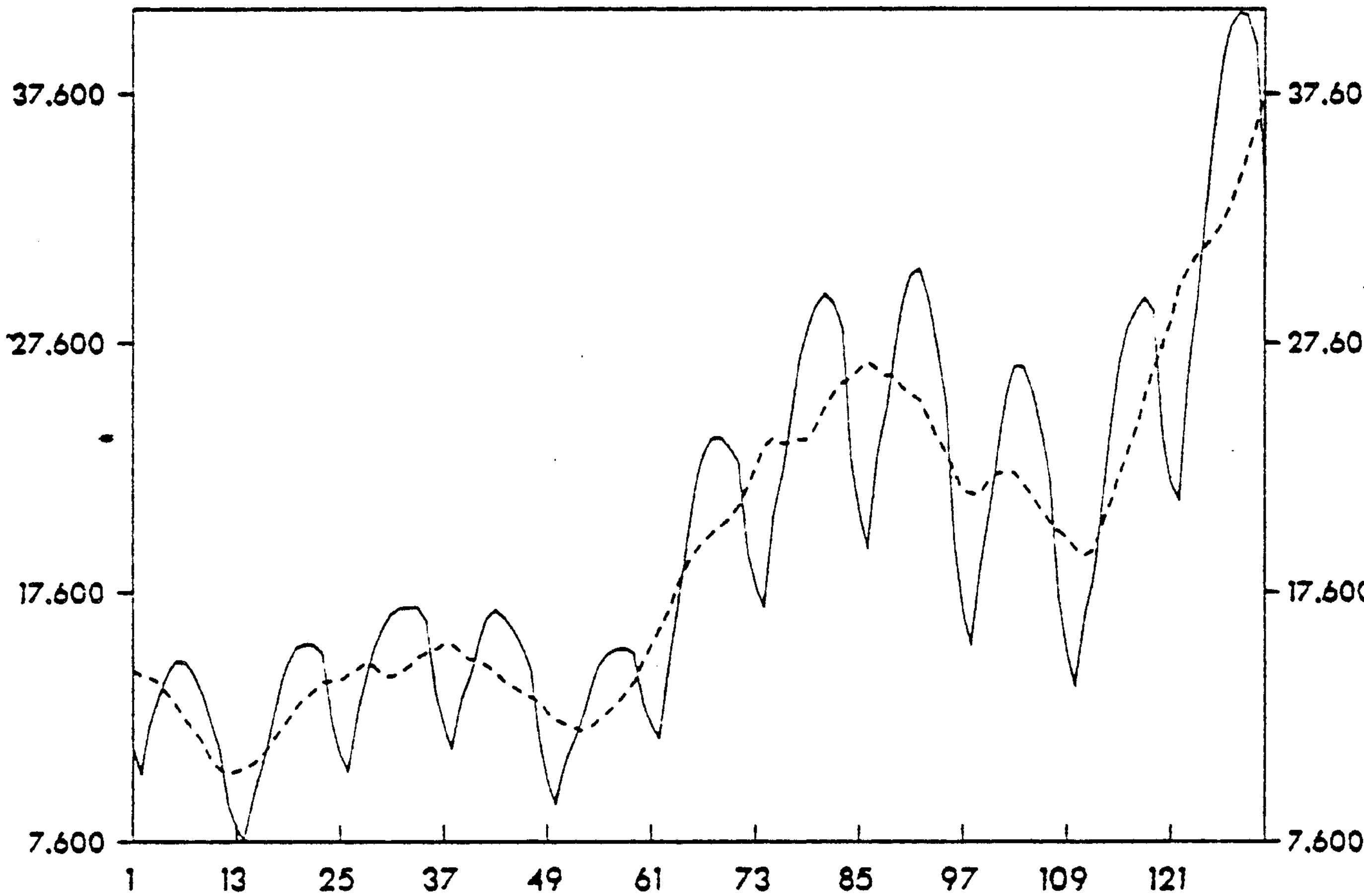


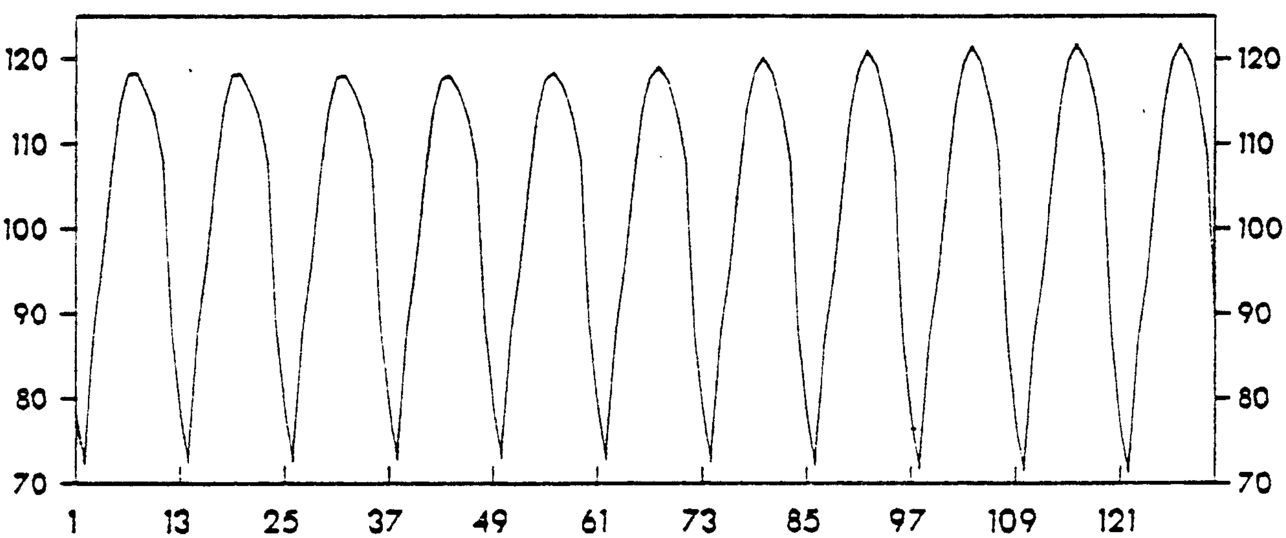
VALUE PUT IN PLACE MODEL BASED



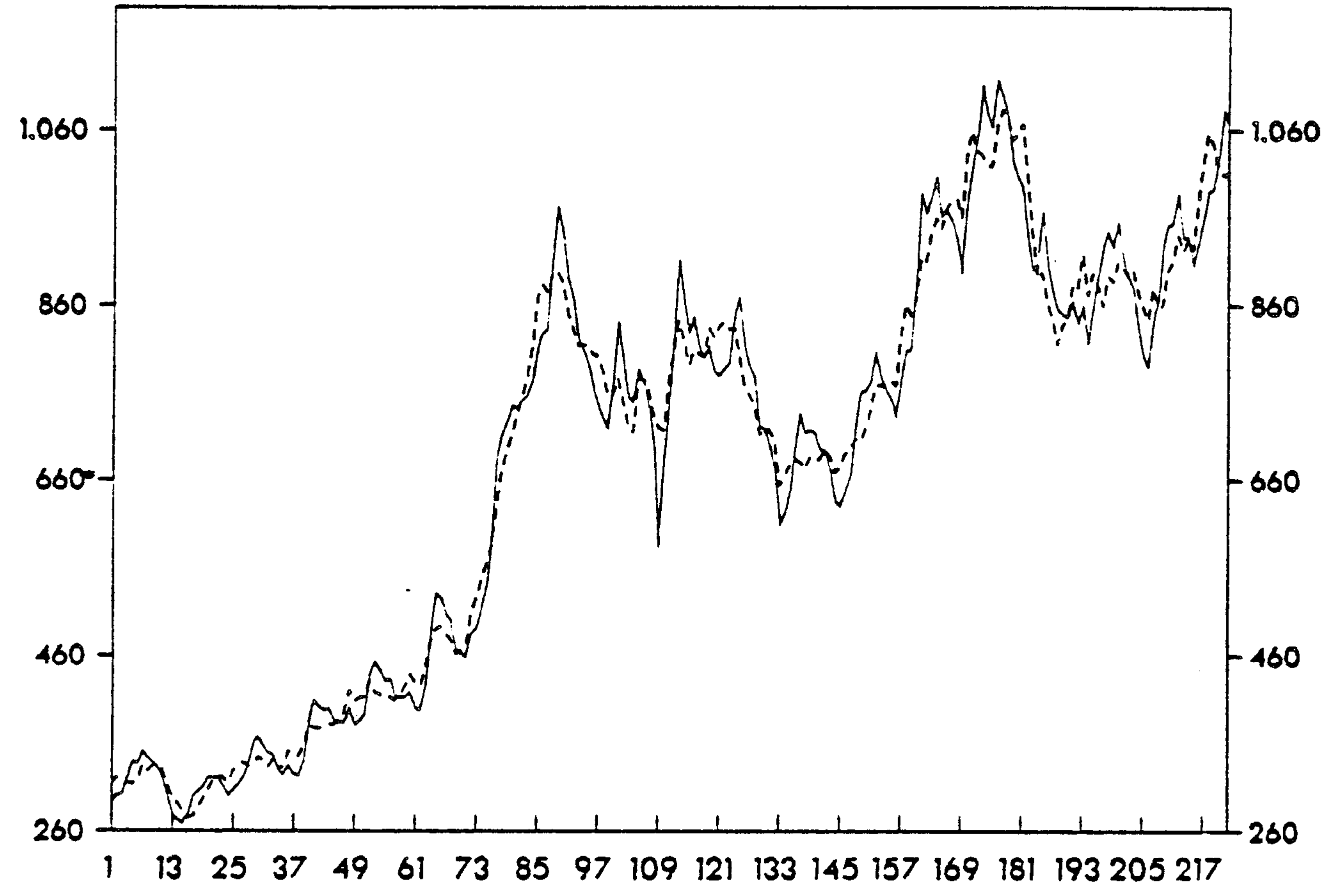


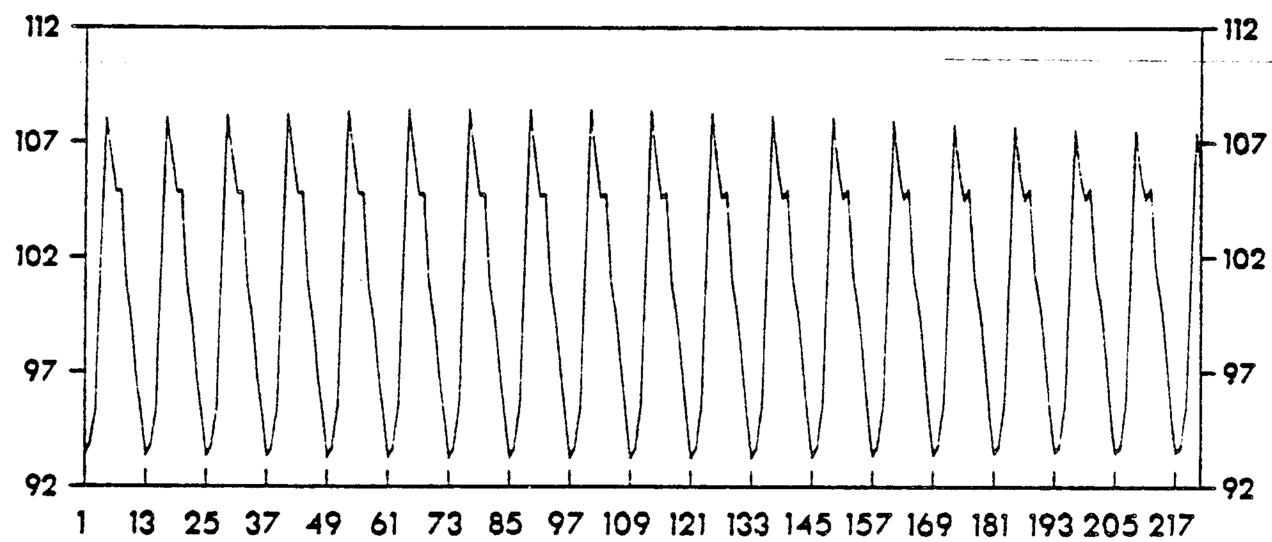
VALUE PUT IN PLACE X-11 LOG



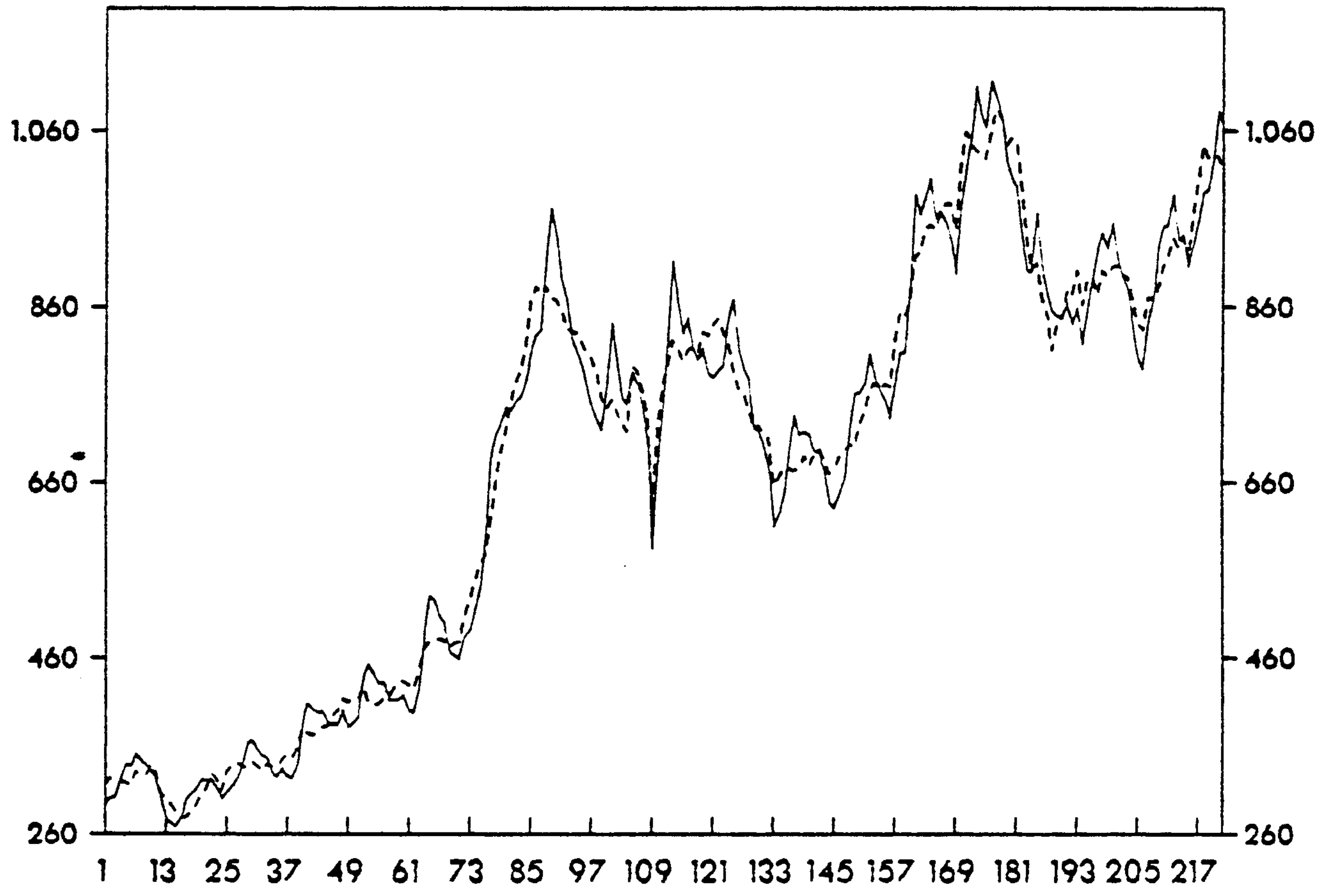


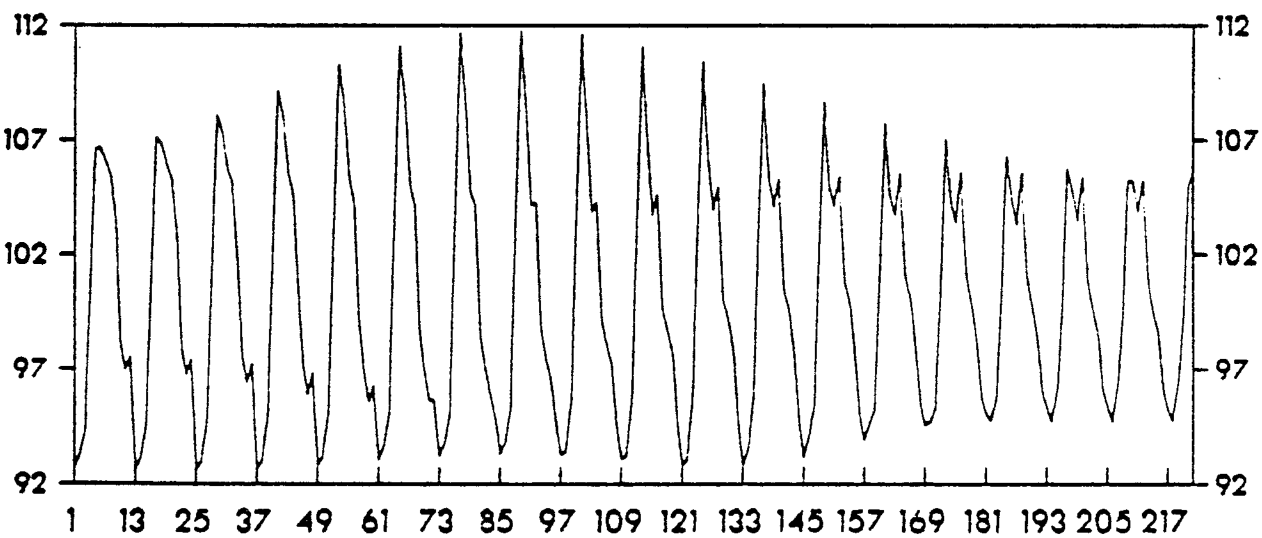
S36TI MODEL BASED



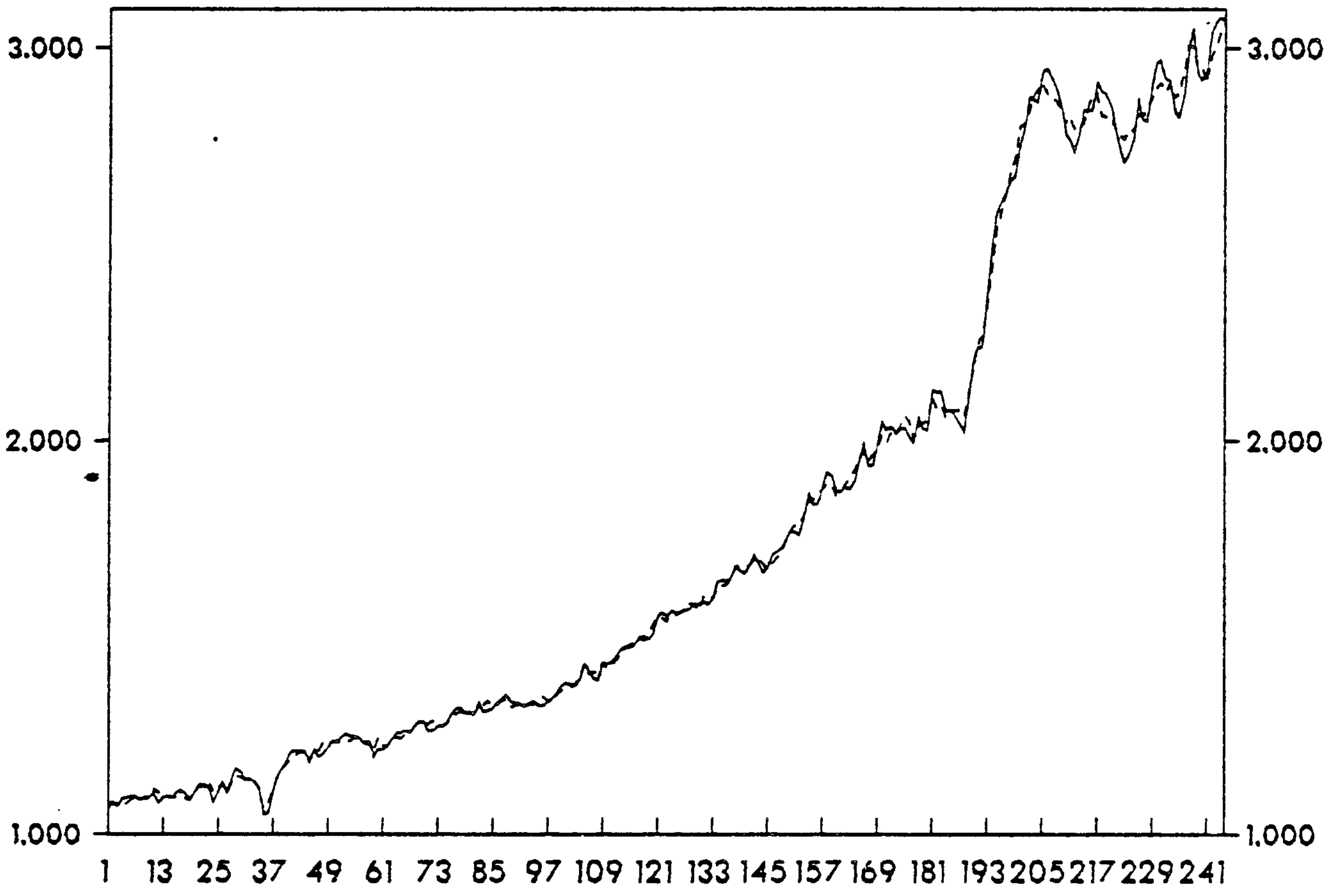


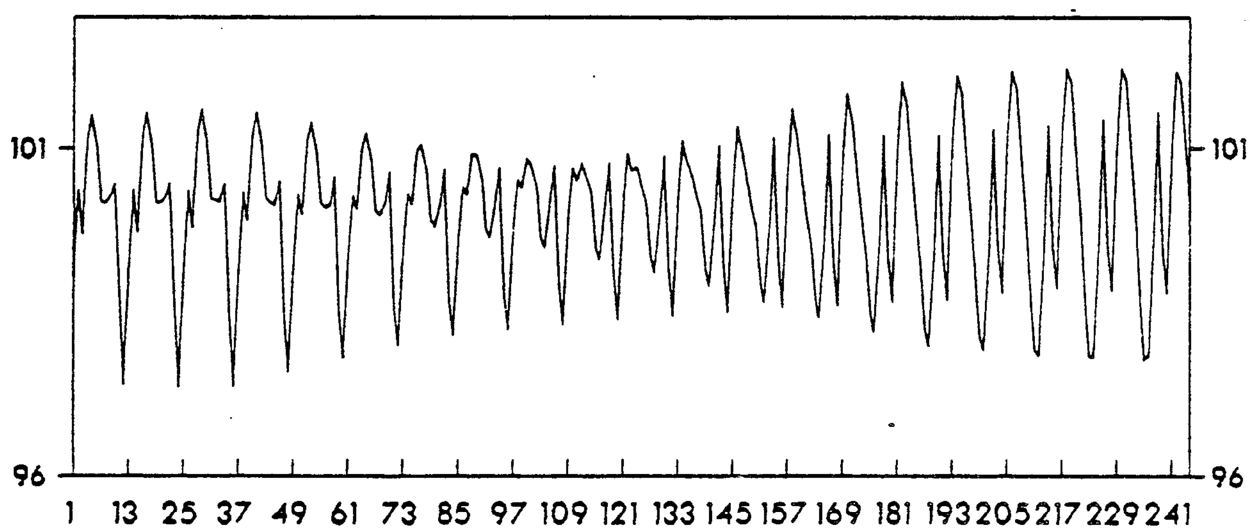
S36TI X-11 LOG



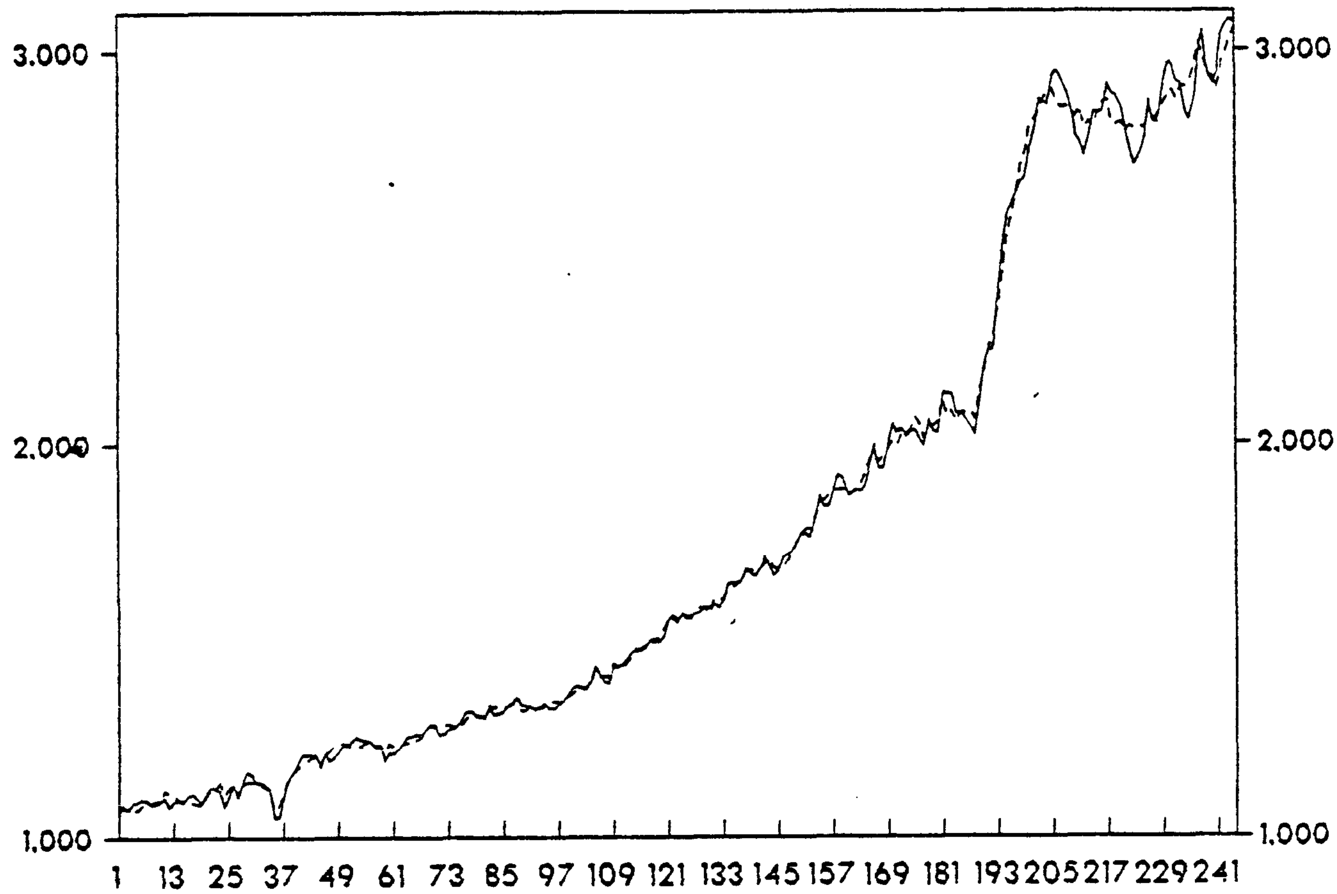


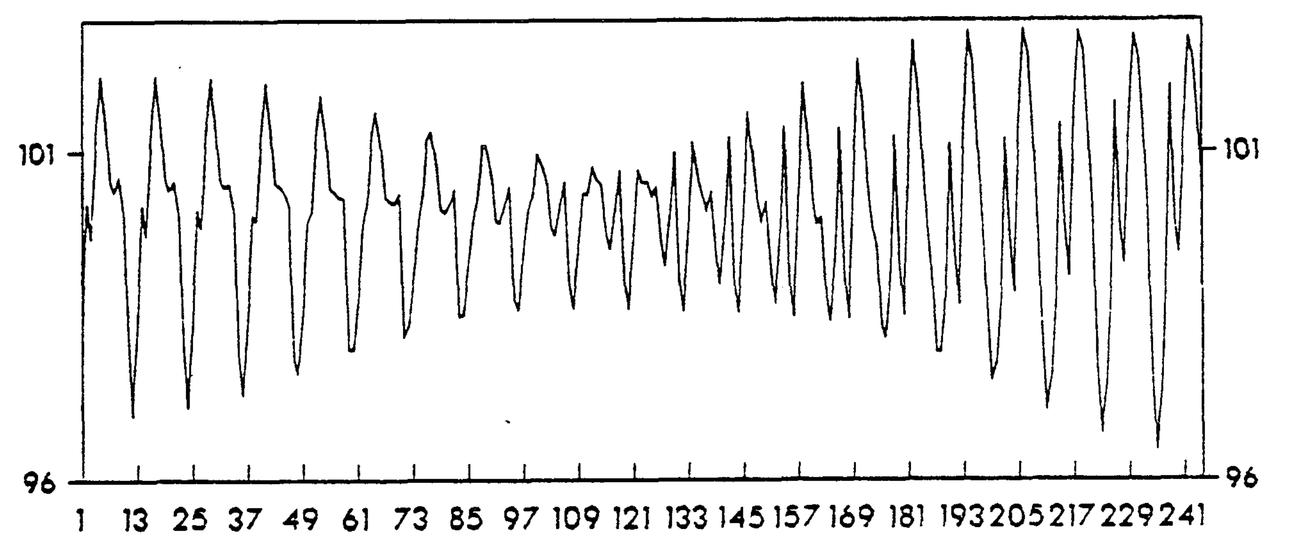
INDUSTRY S62TI MODEL BASED



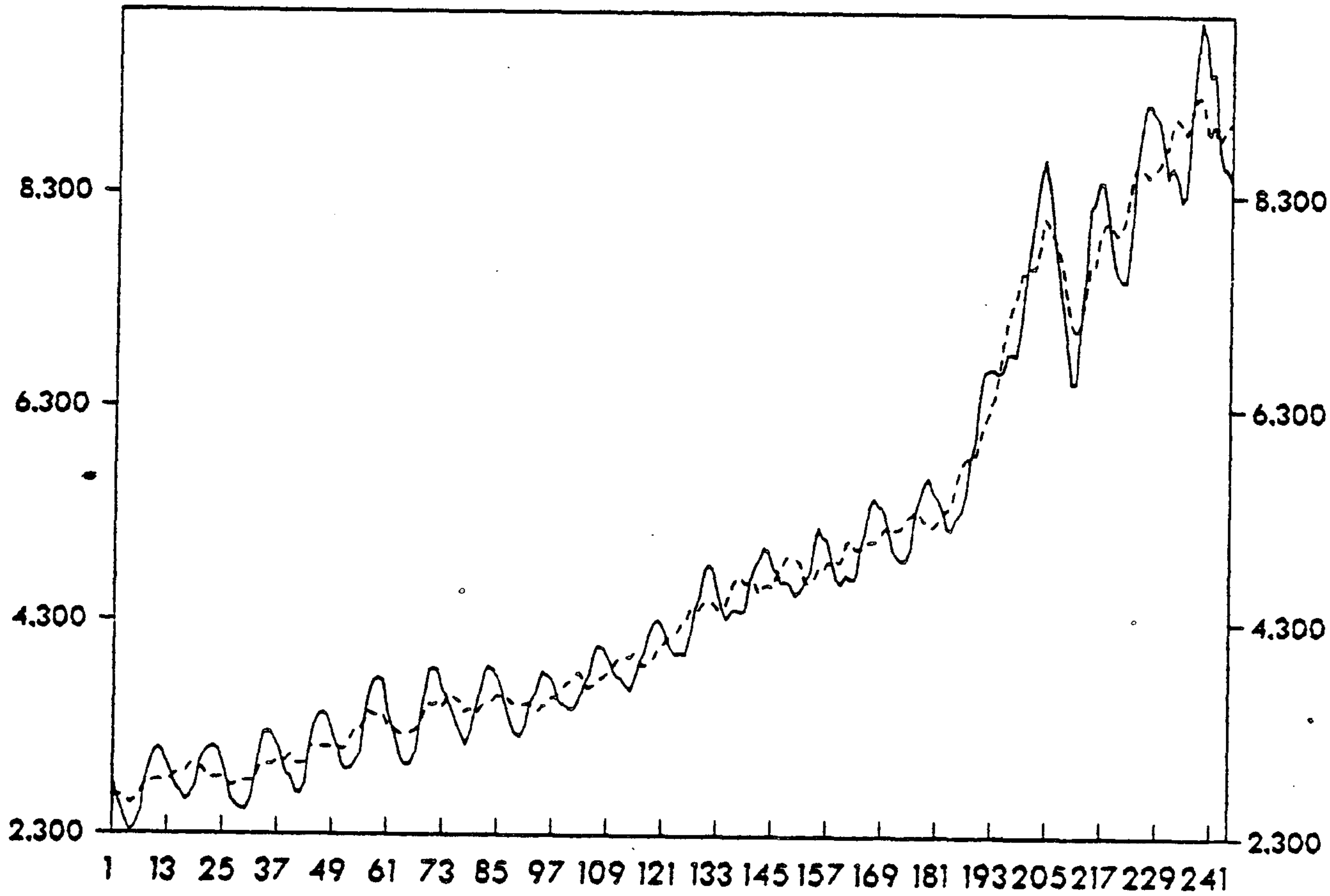


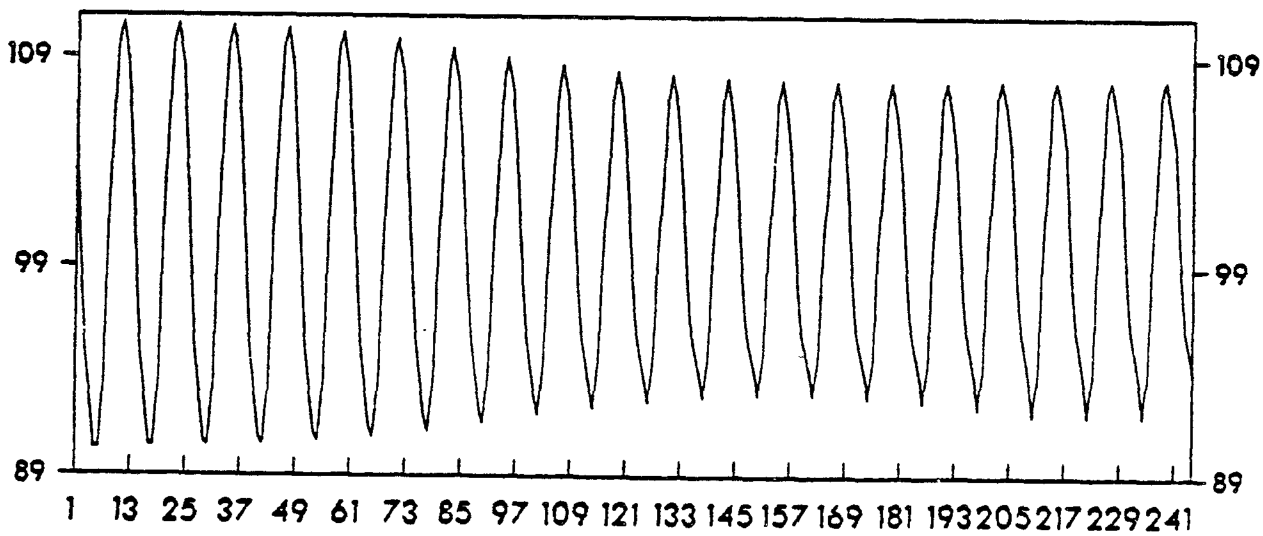
INDUSTRY S62TI X-11 LOG



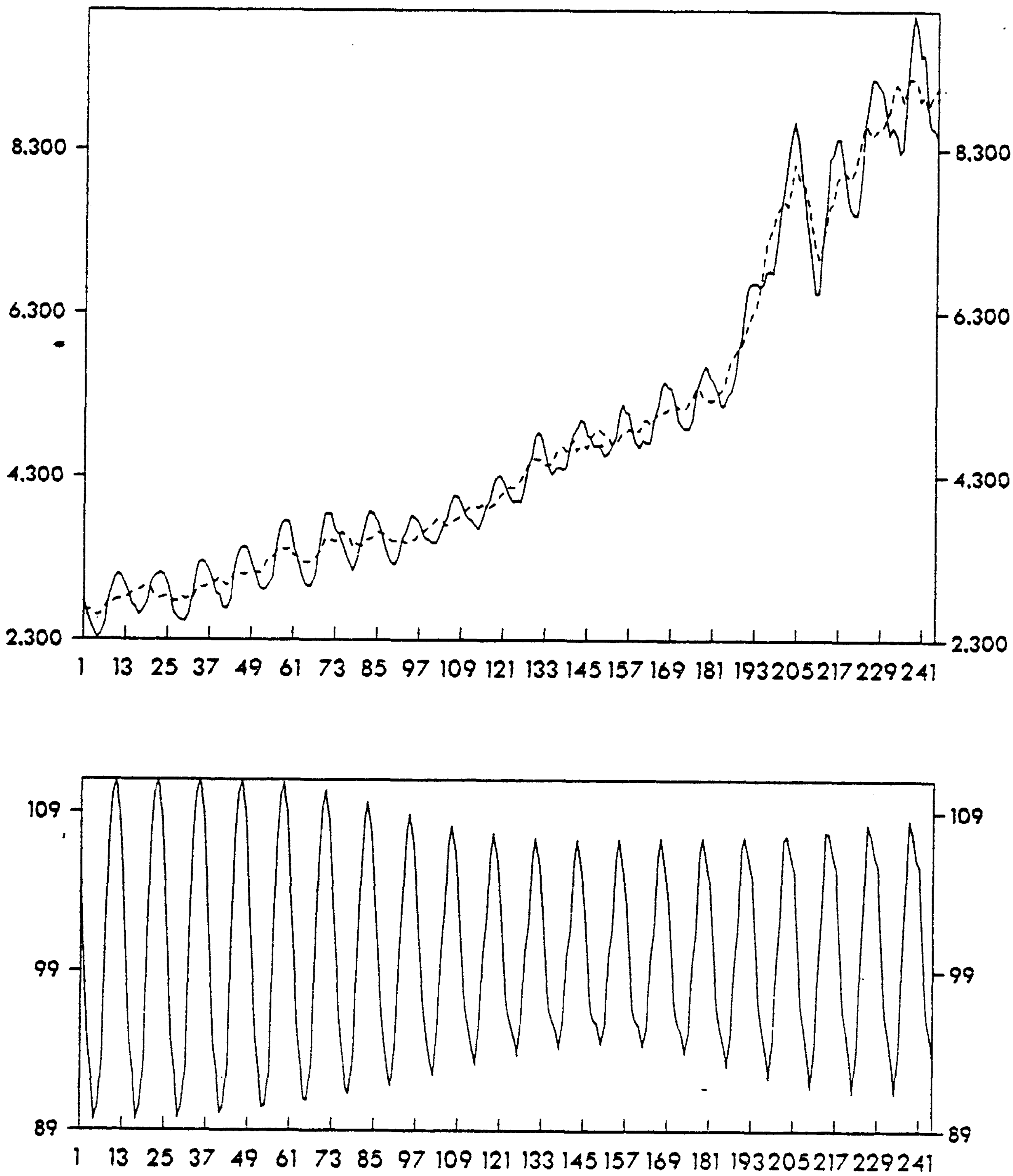


INDUSTRY S64TI MODEL BASED

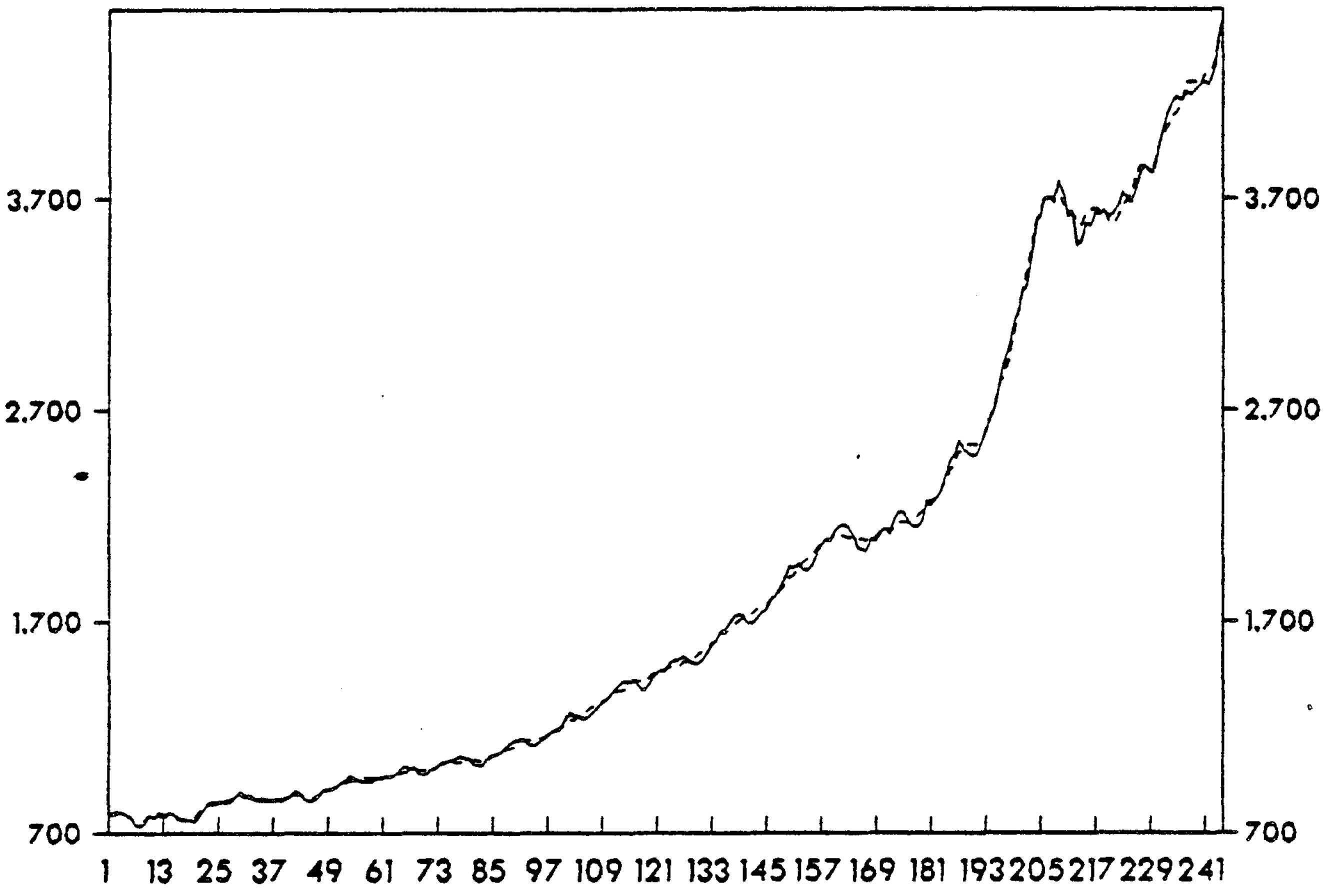


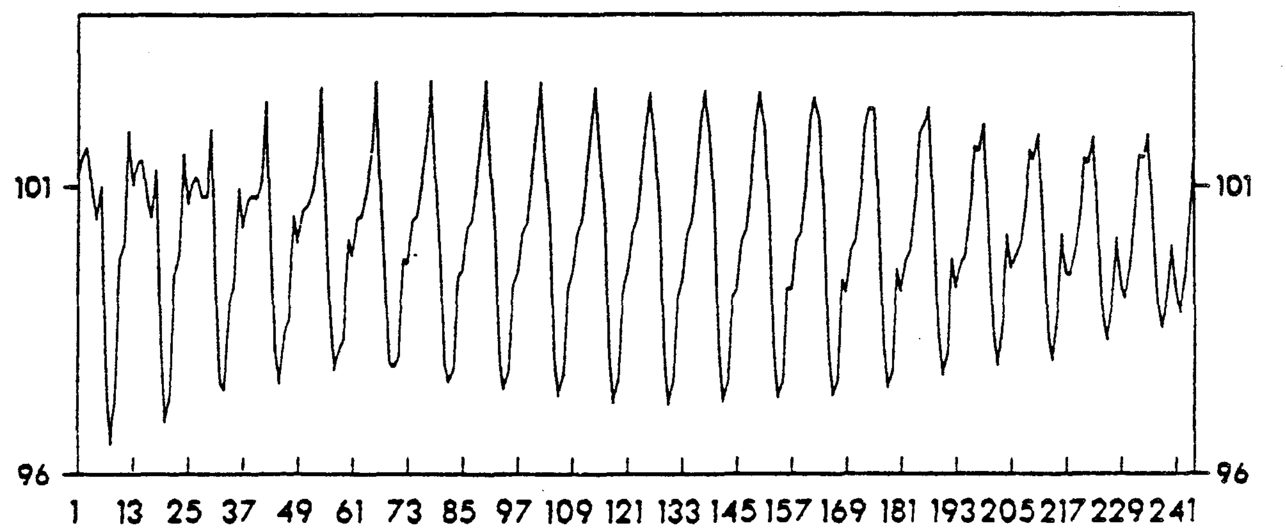


INDUSTRY S64TI X-11 LOG

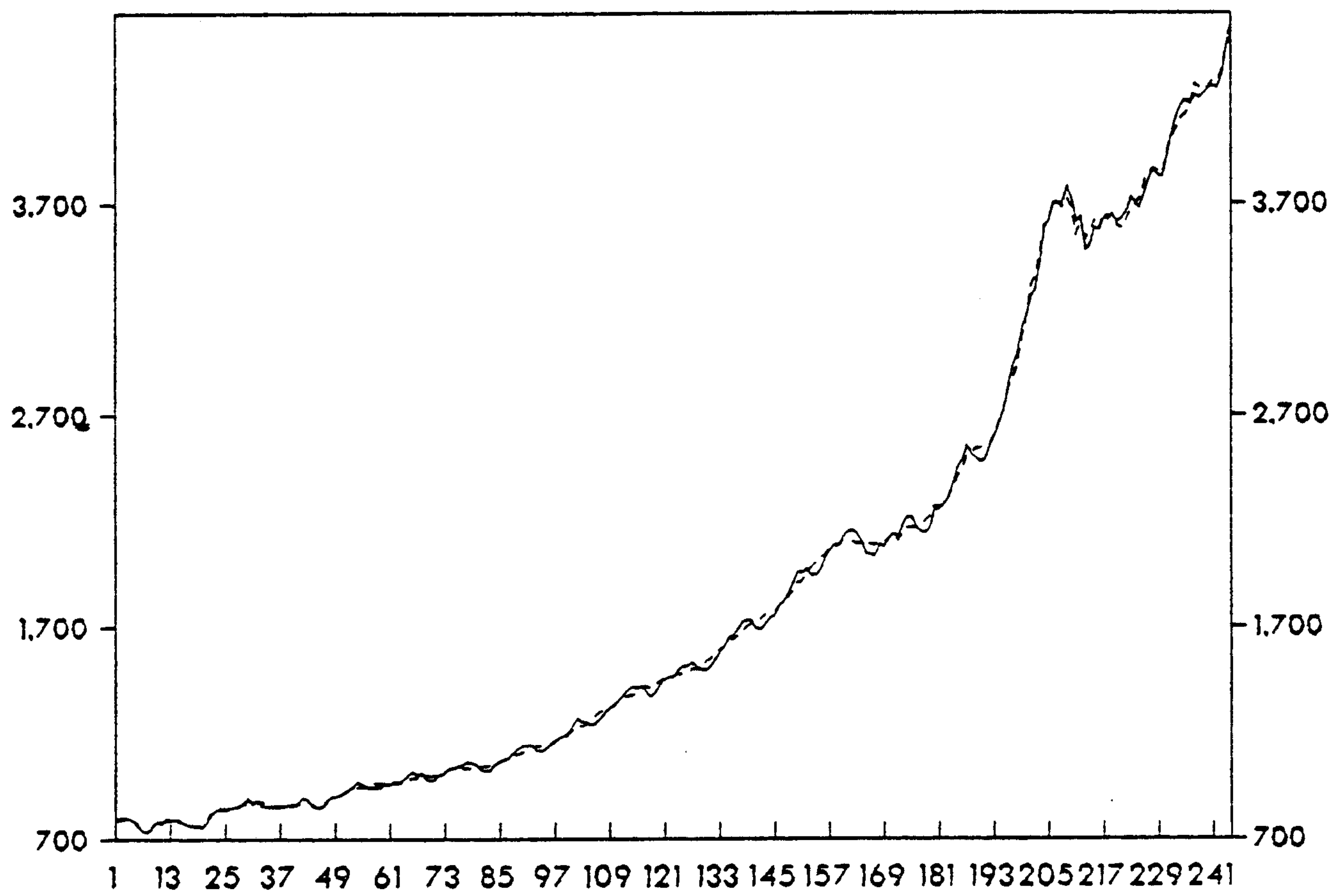


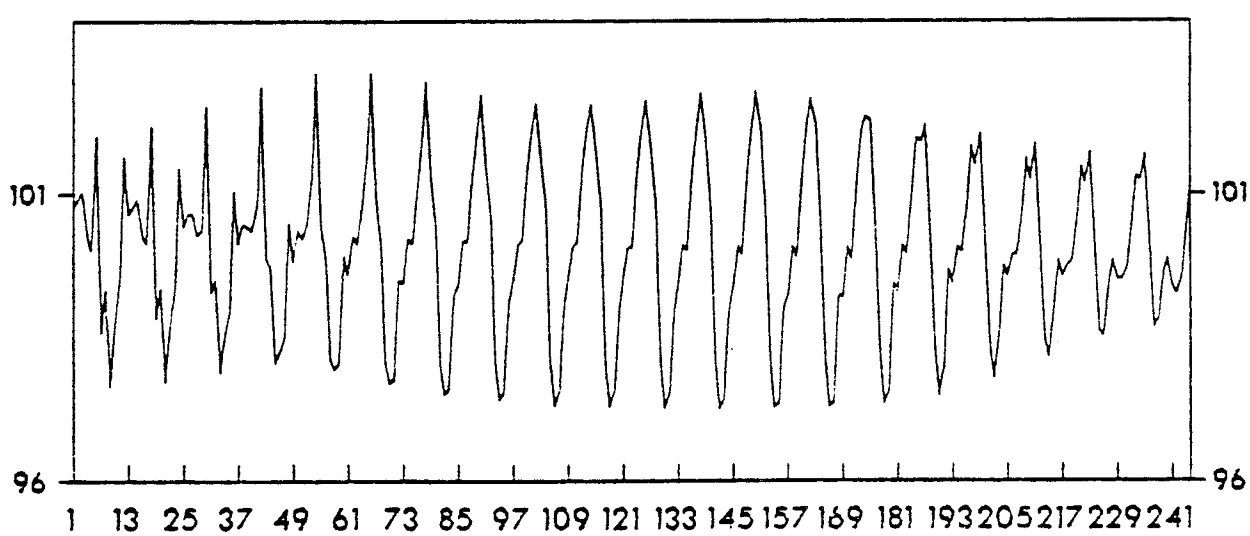
INDUSTRY S85TI MODEL BASED



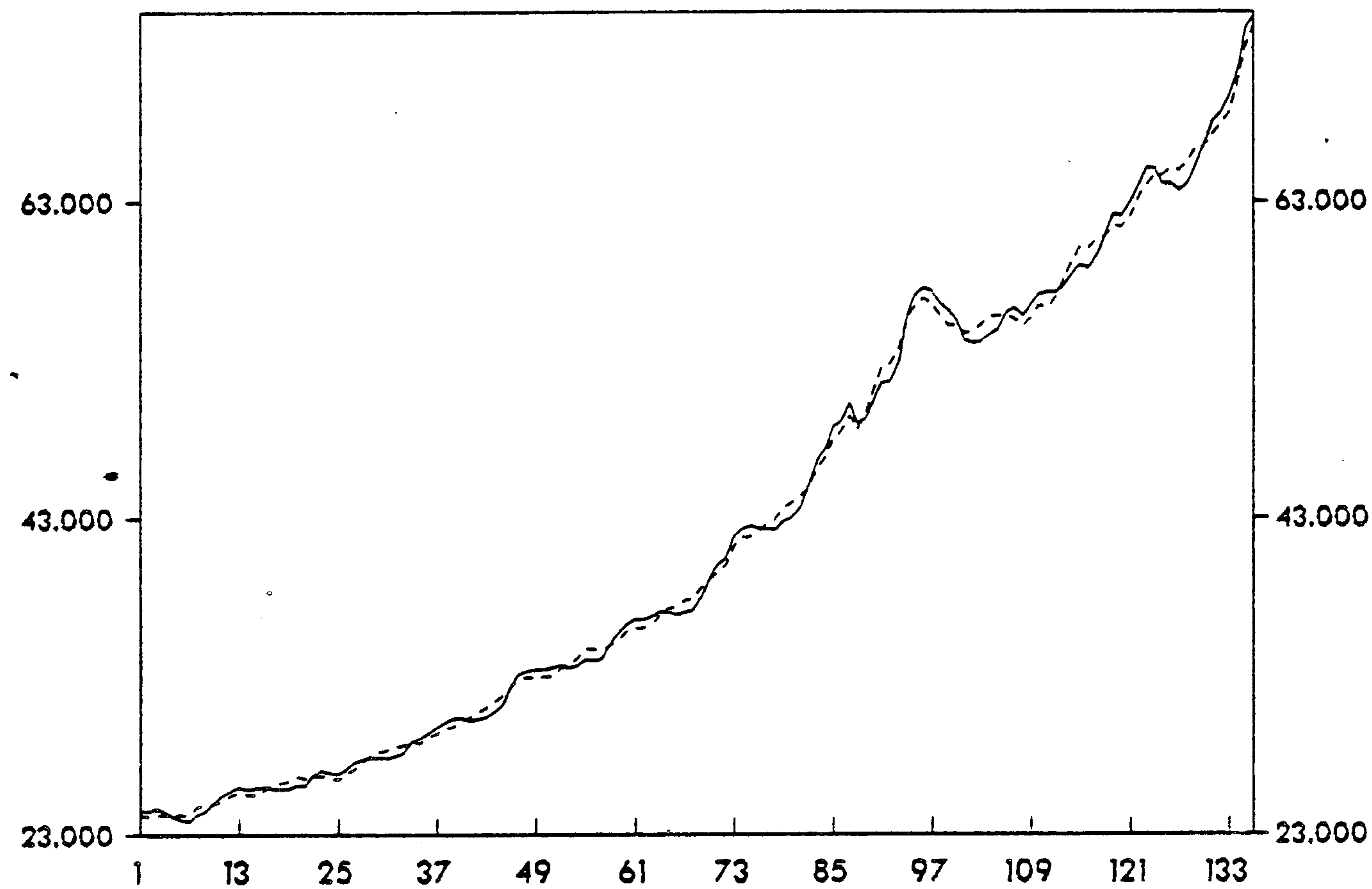


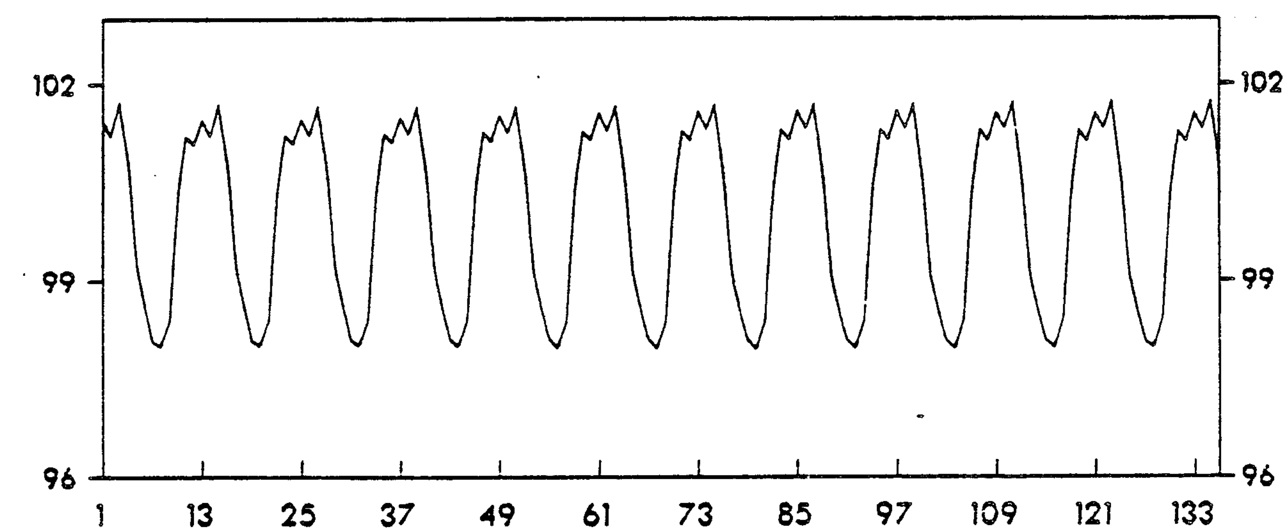
INDUSTRY S85T1 X-11 LOG



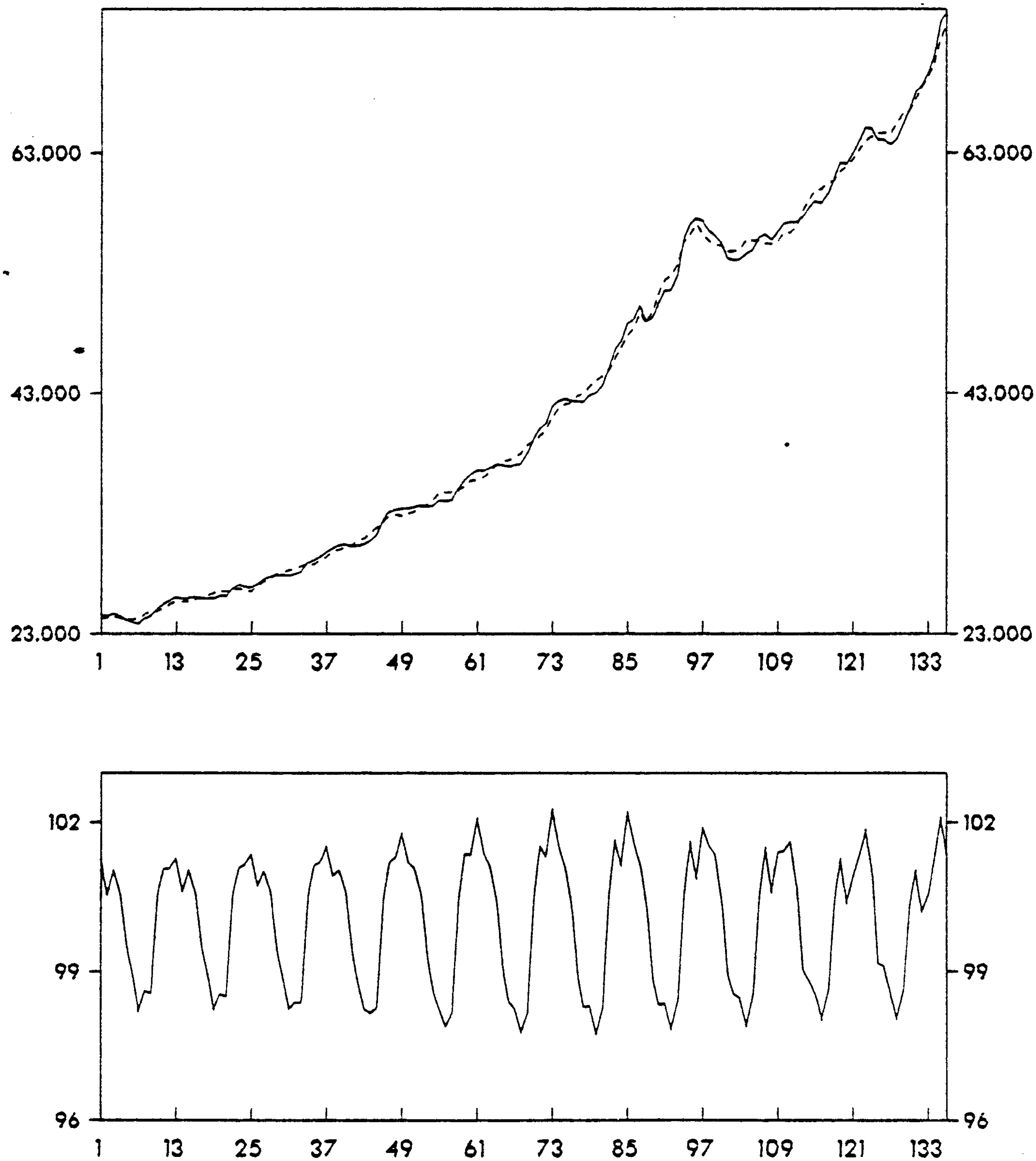


WHOLESALE T1500 MODEL BASED

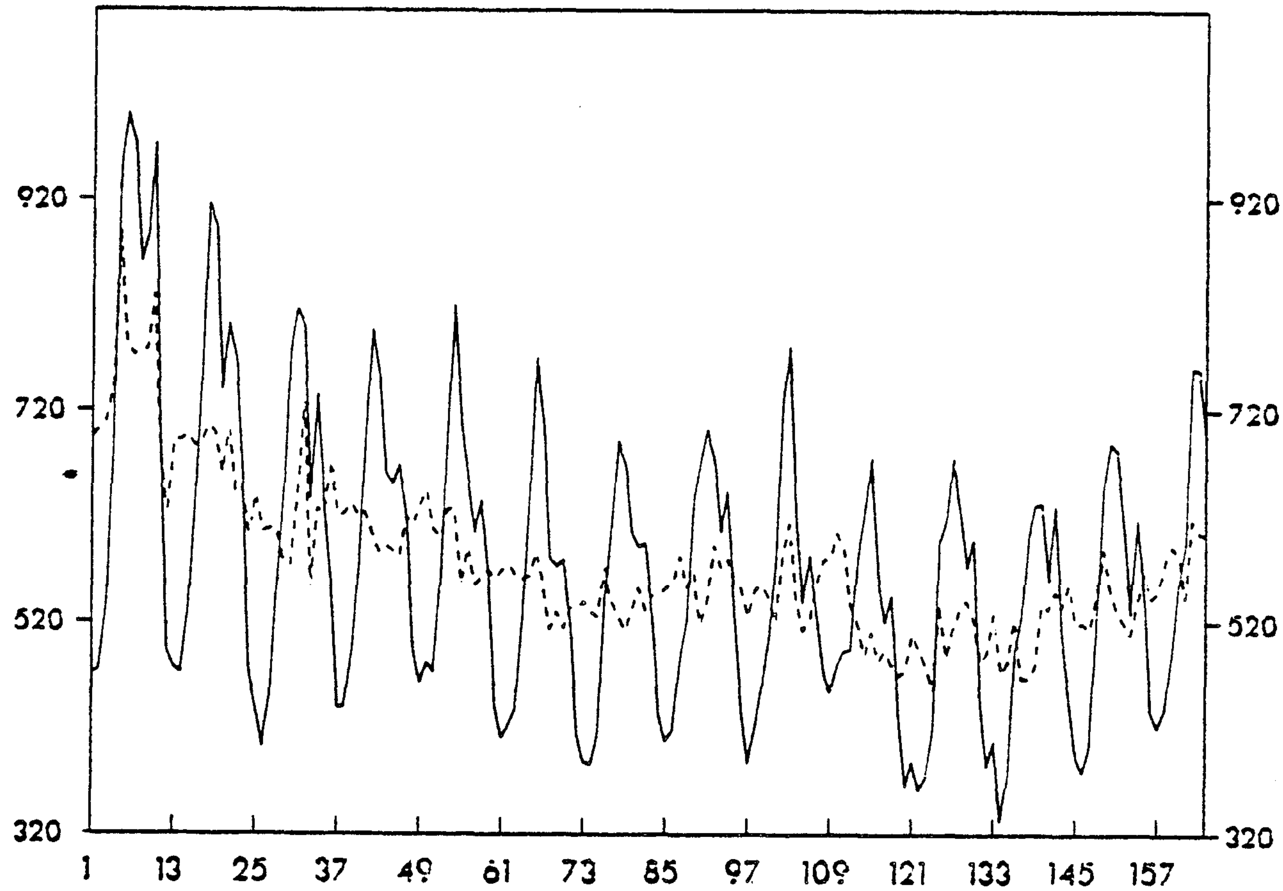


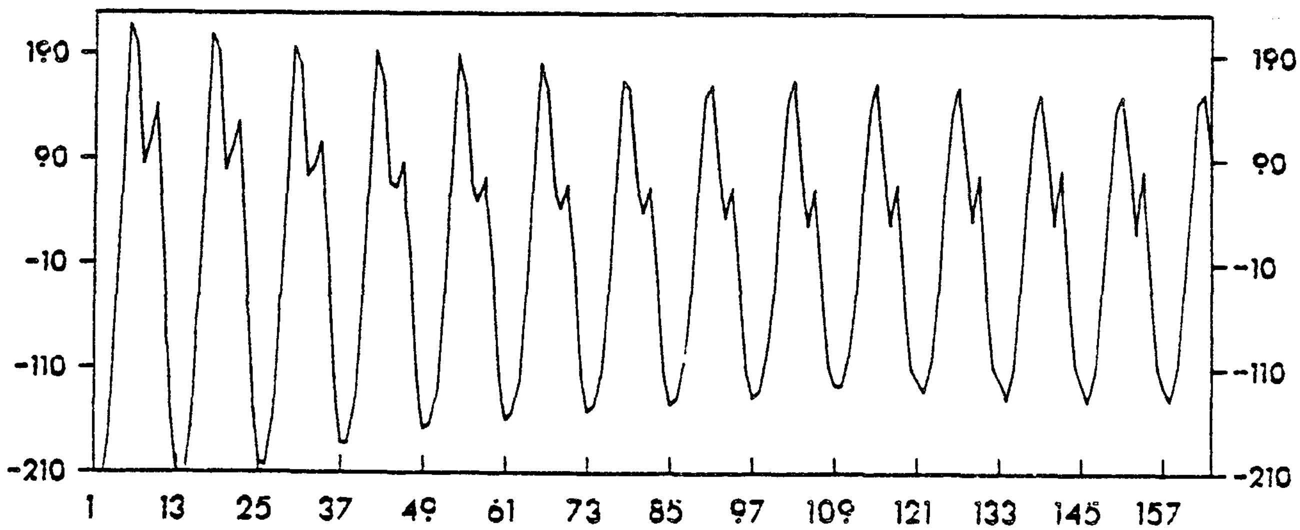


WHOLESALE T1500 X-11 LOG

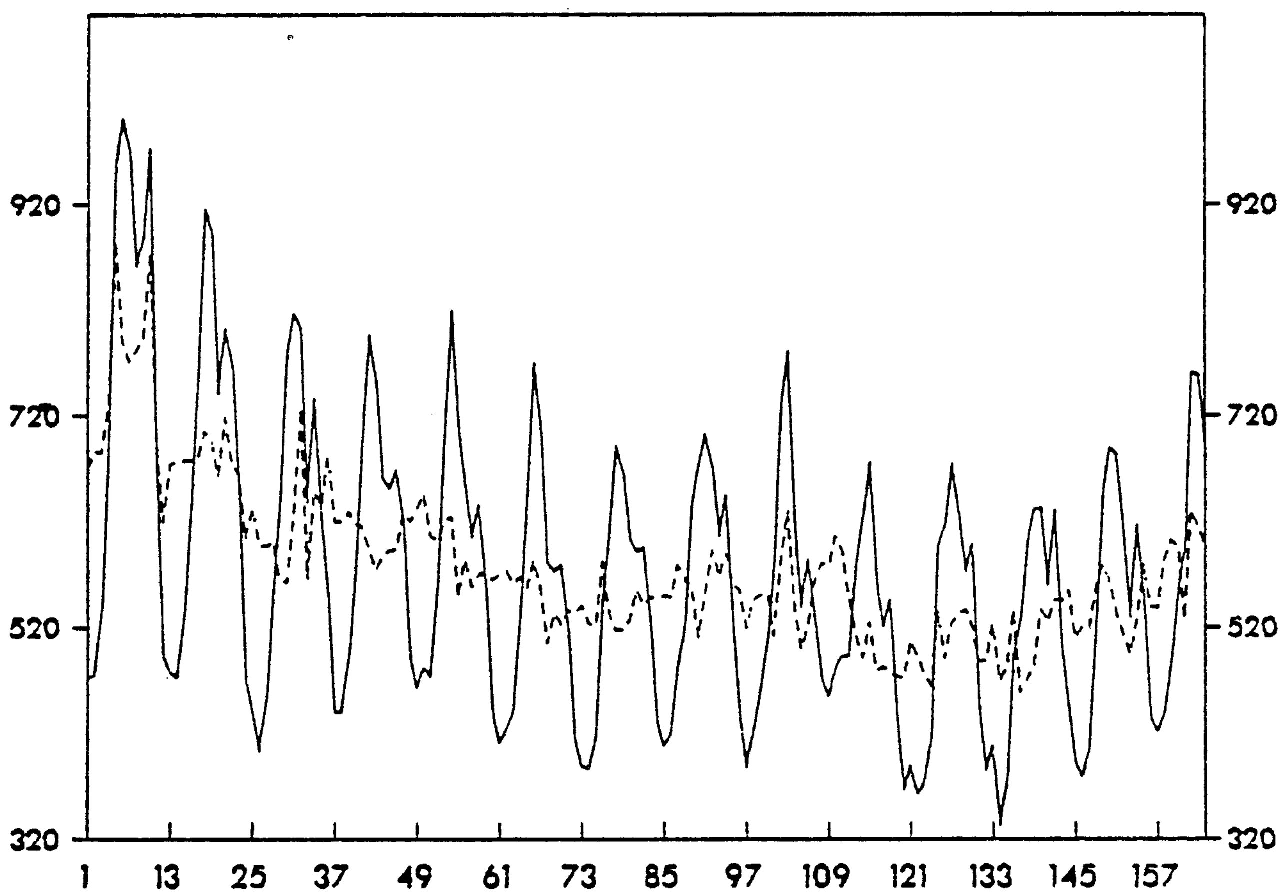


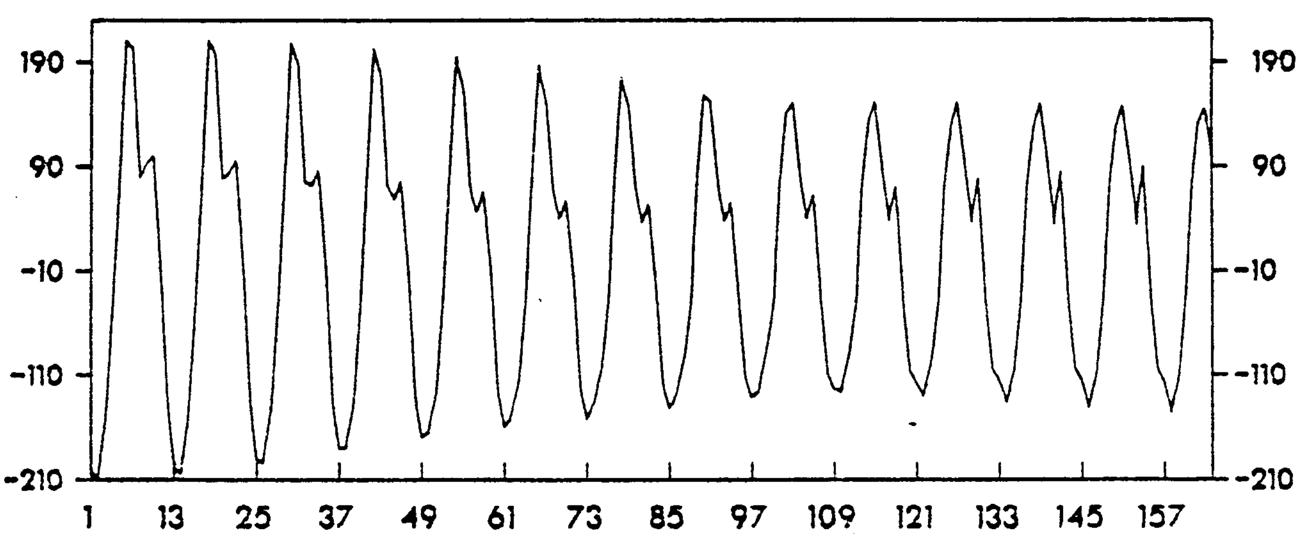
EAF20 MODEL BASED



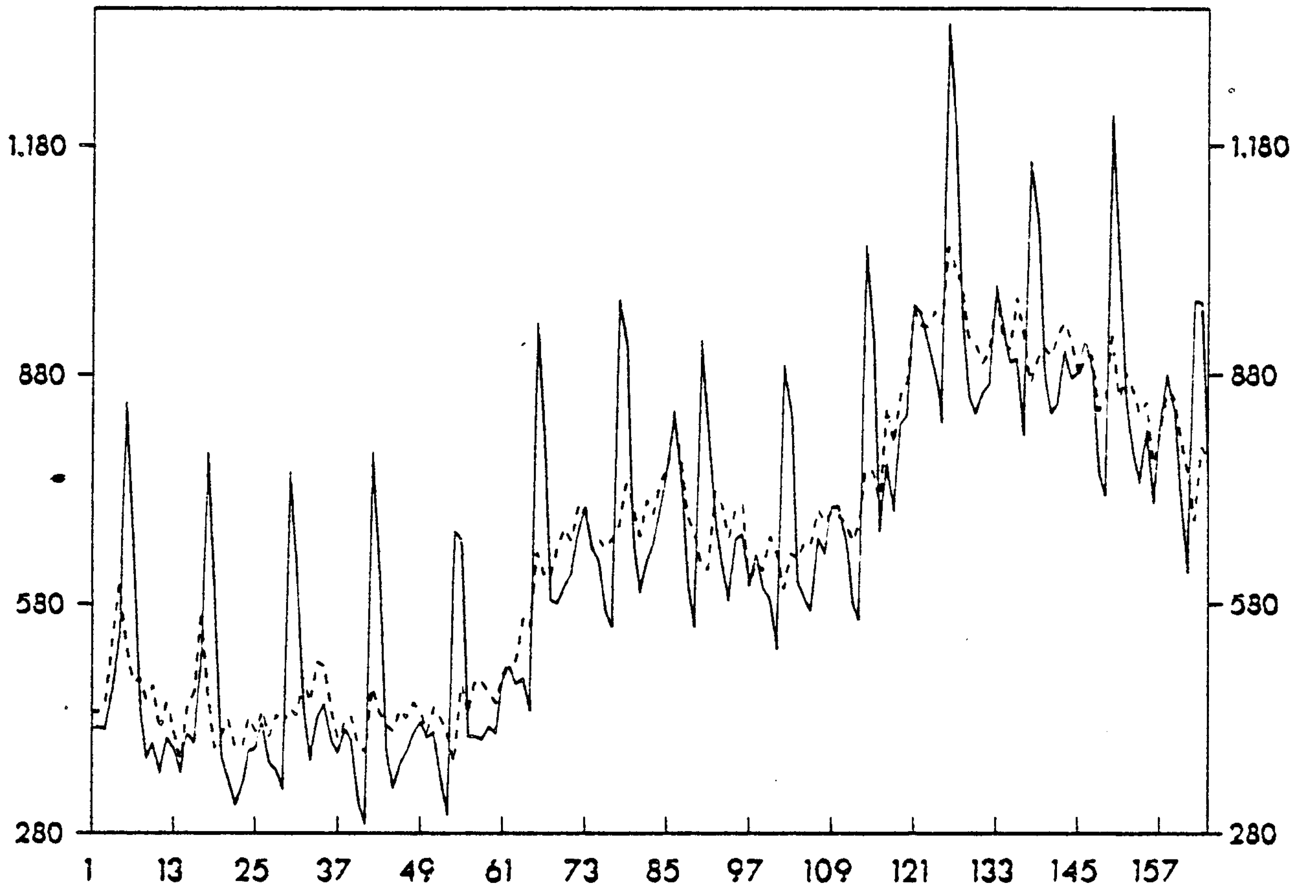


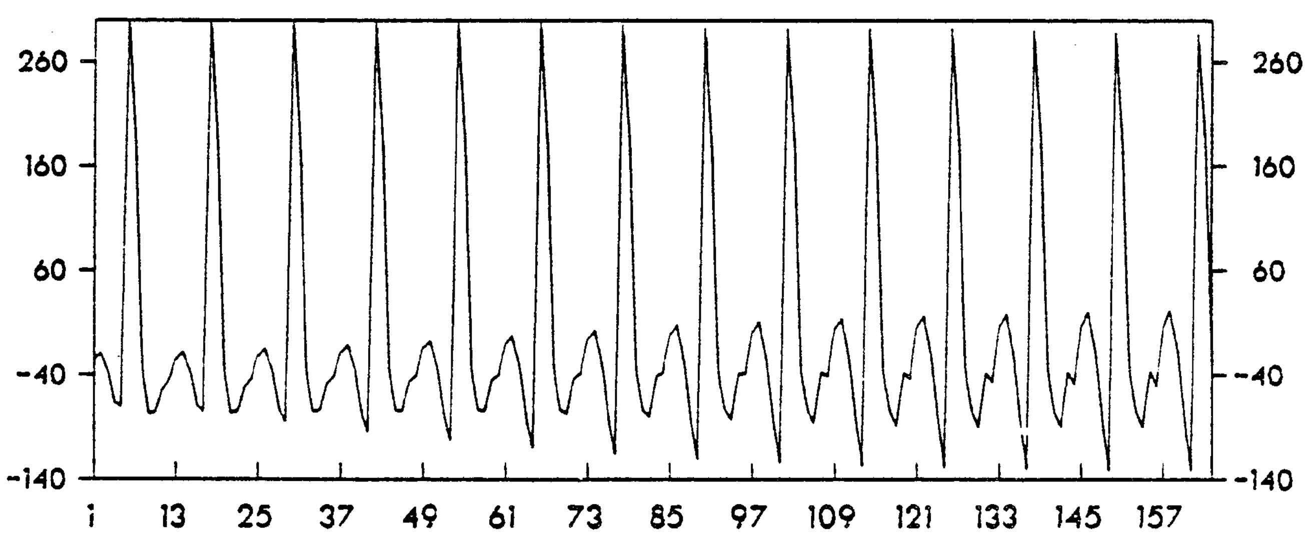
EAF20 X-11 LOG



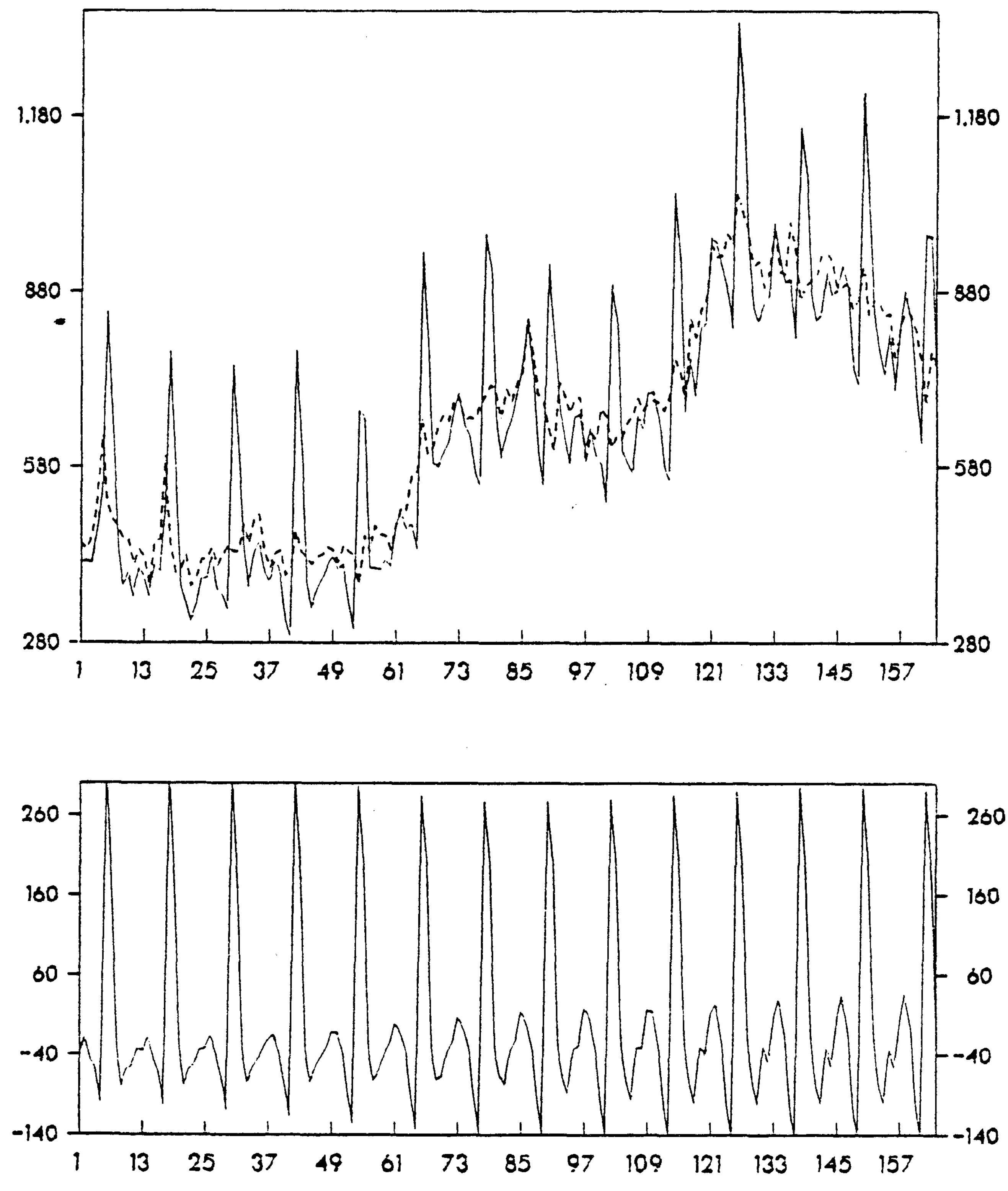


UM16 MODEL BASED

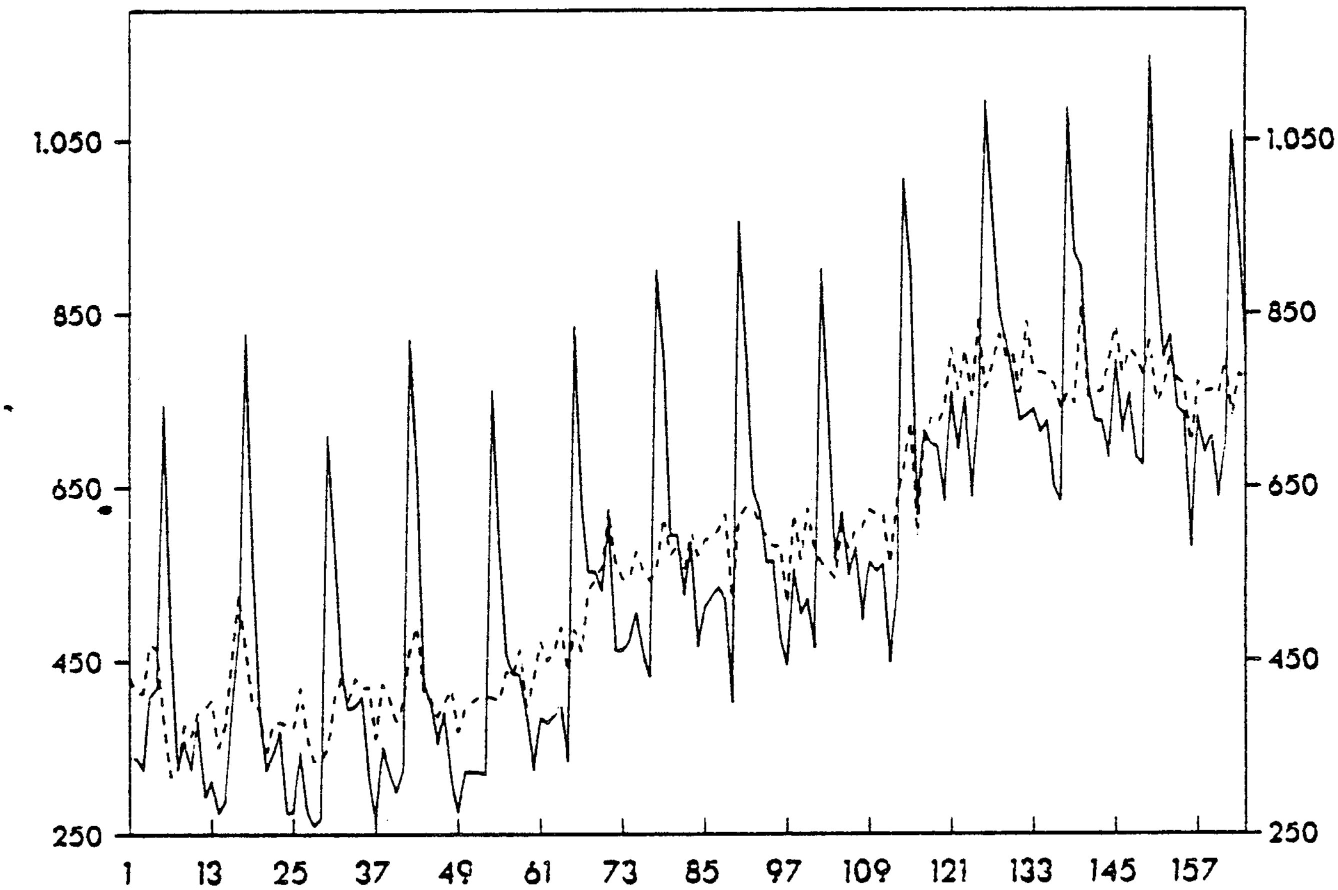


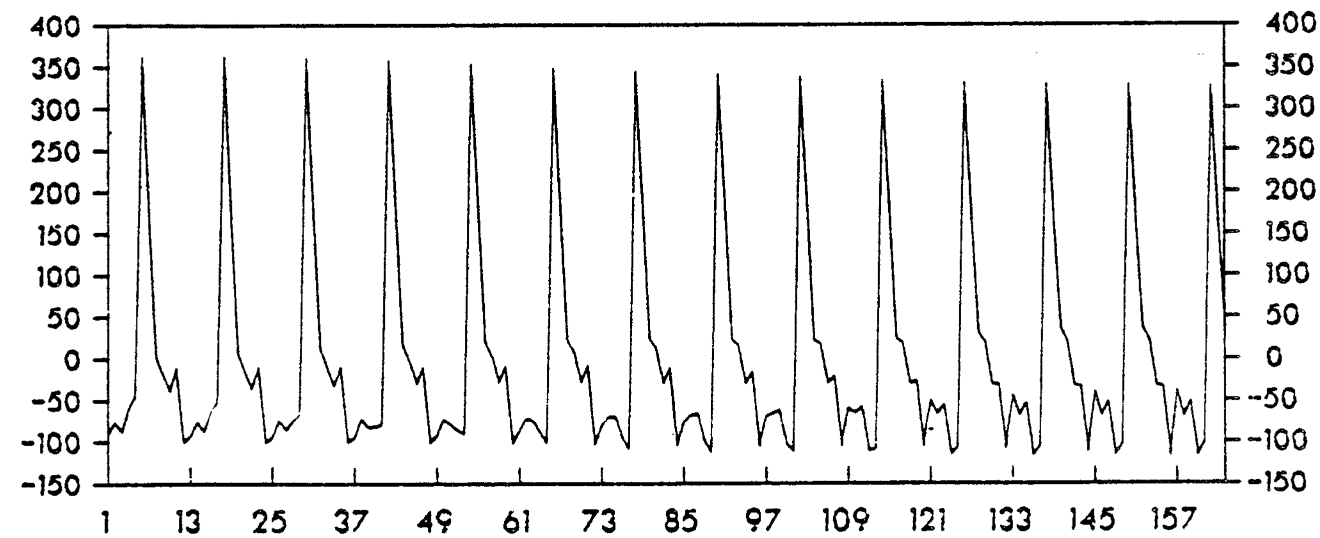


UM16 X-11 LOG

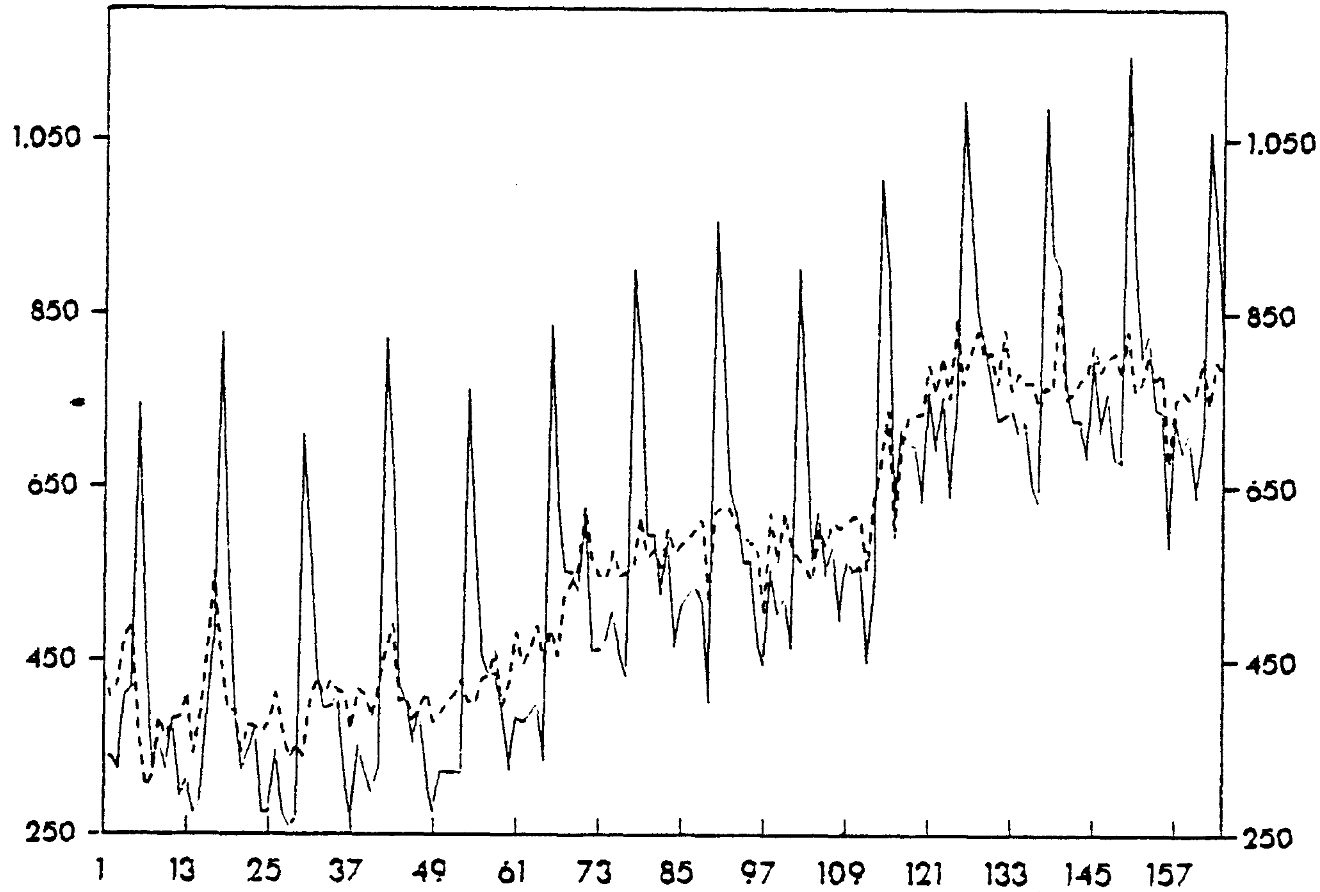

1,180
880
580
280
1
13
25
37
49
61
73
85
97
109
121
133
145
157
260
160
60
-40
-140

UF16 MODEL BASED

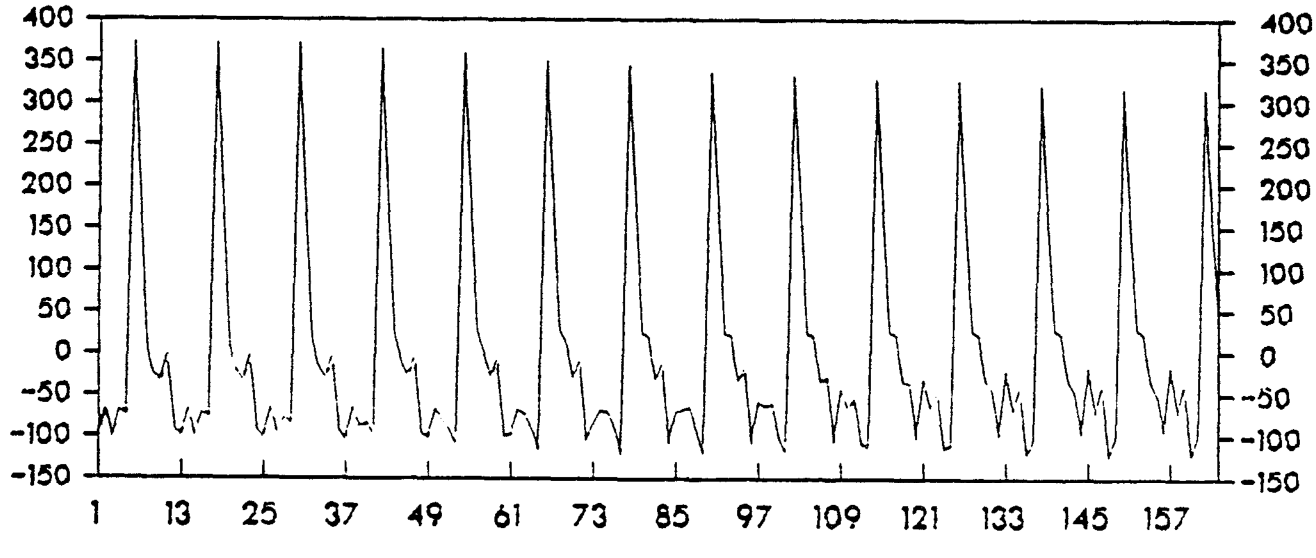




UF16 X-11 LOG

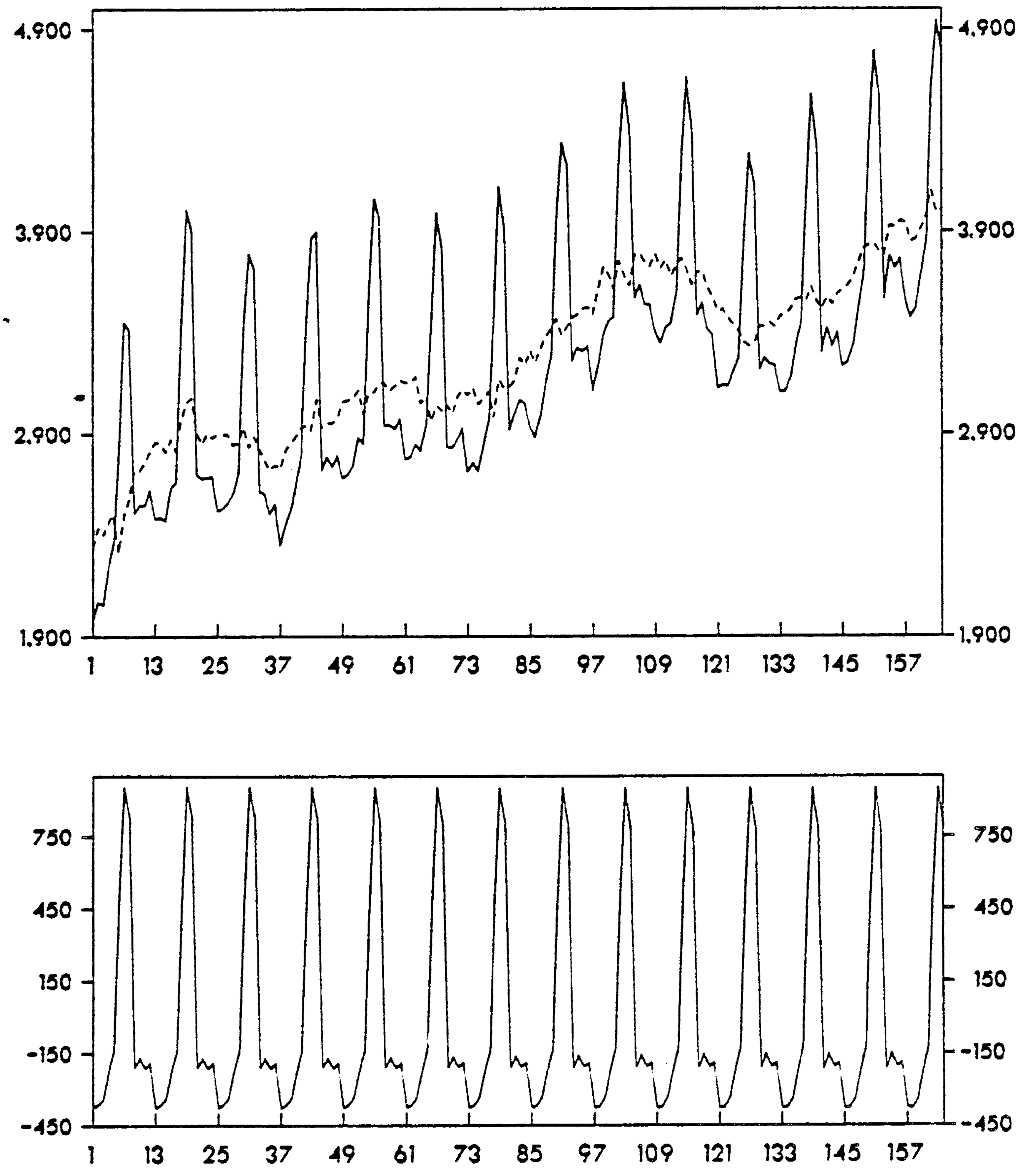


1,050
850
650
450
250
1
13
25
37
49
61
73
85
97
109
121
133
145
157

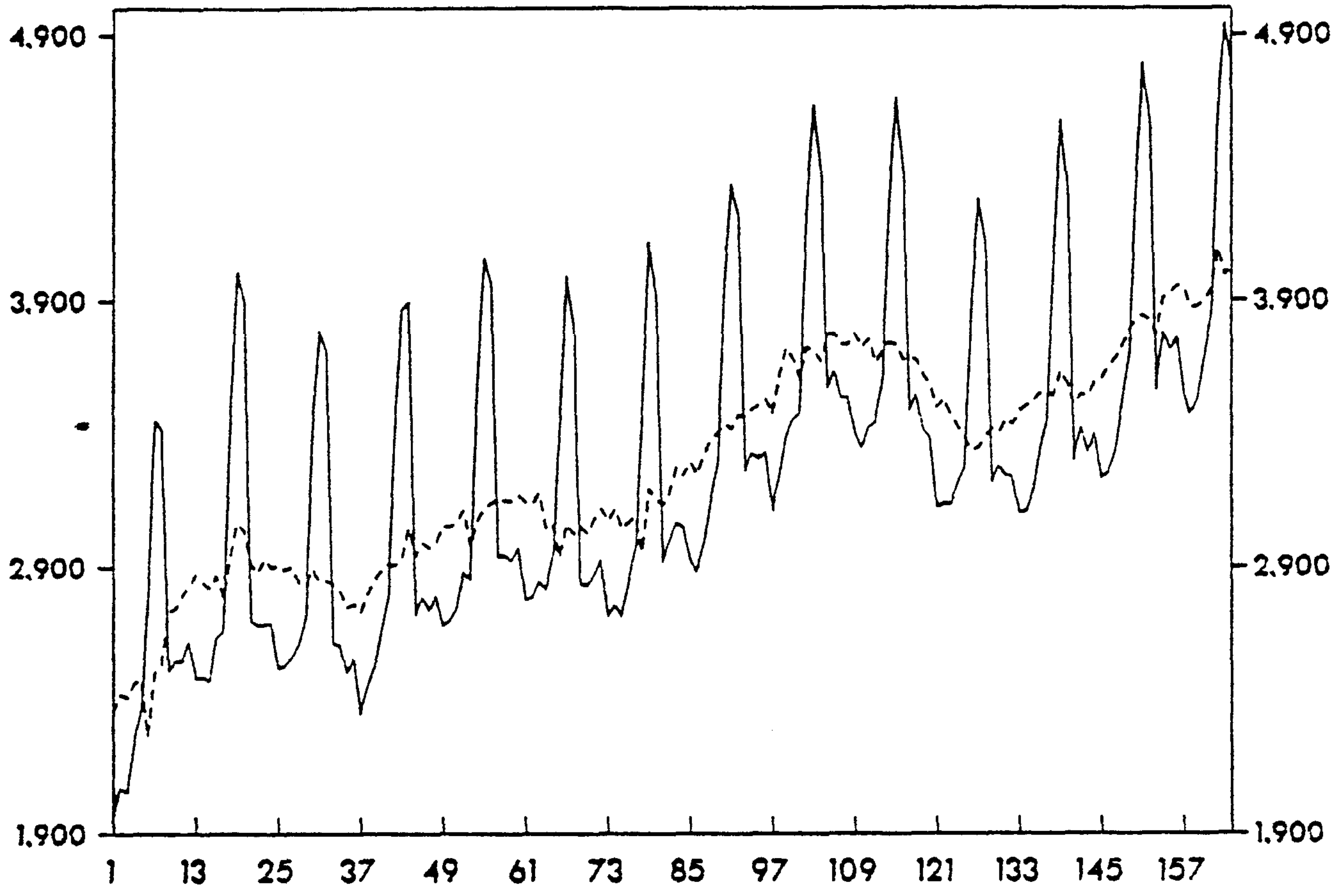


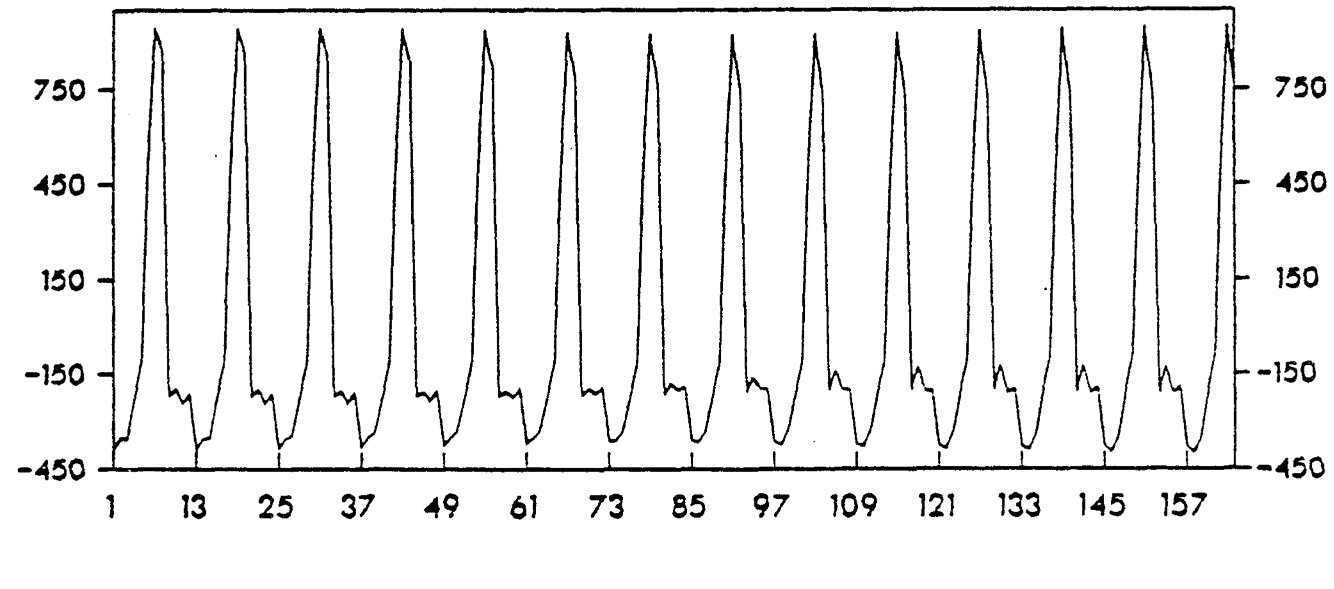
400
350
300
250
200
150
100
50
0
-50
-100
-150
1
13
25
37
49
61
73
85
97
109
121
133
145
157

ENM16 MODEL BASED

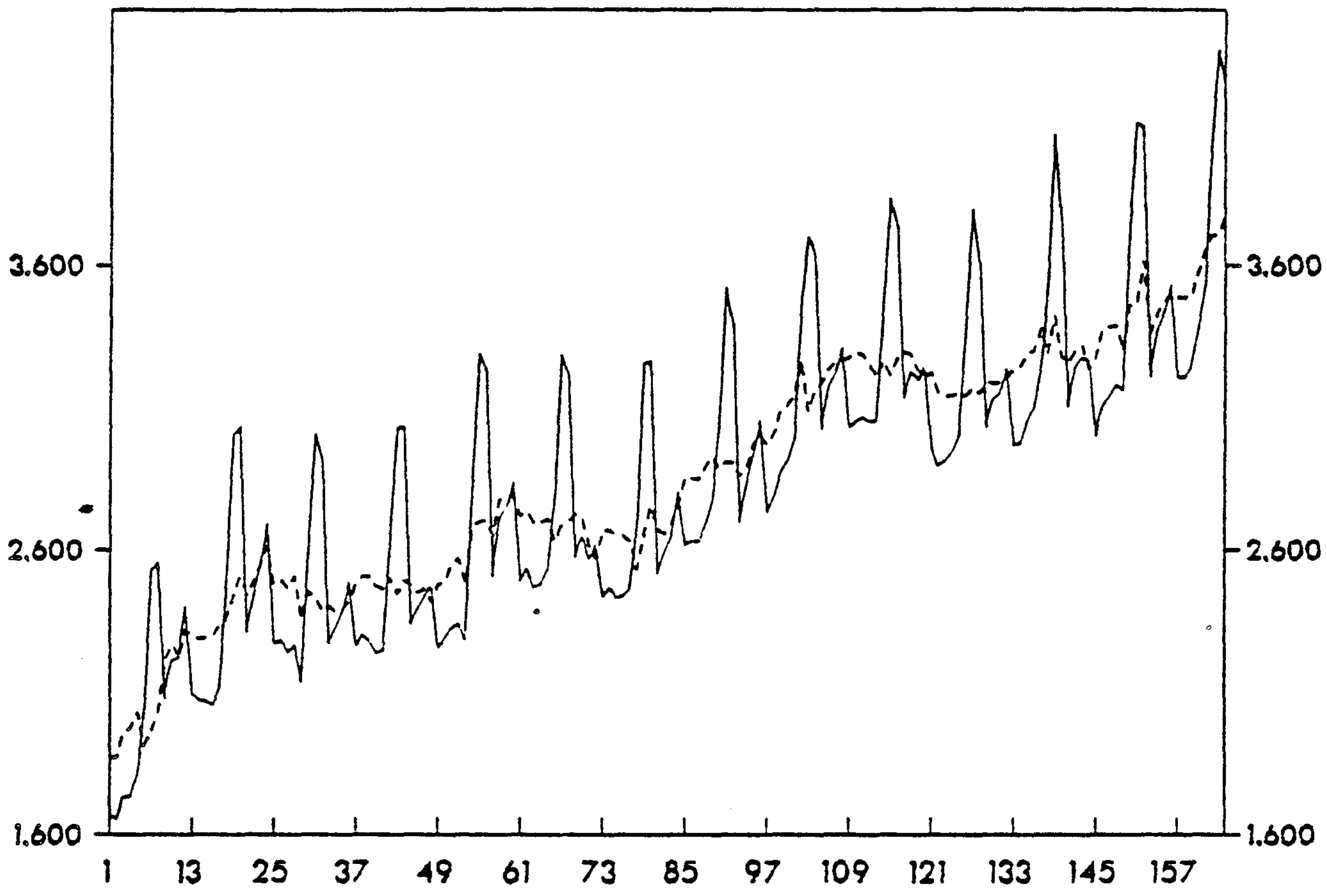


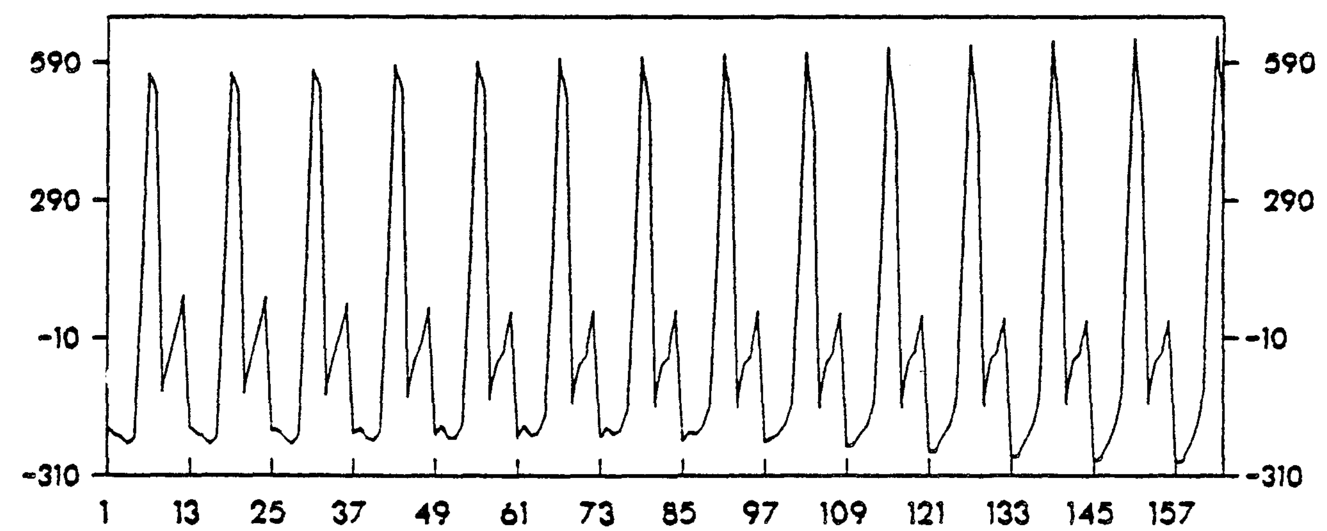
ENM16 X-11 LOG



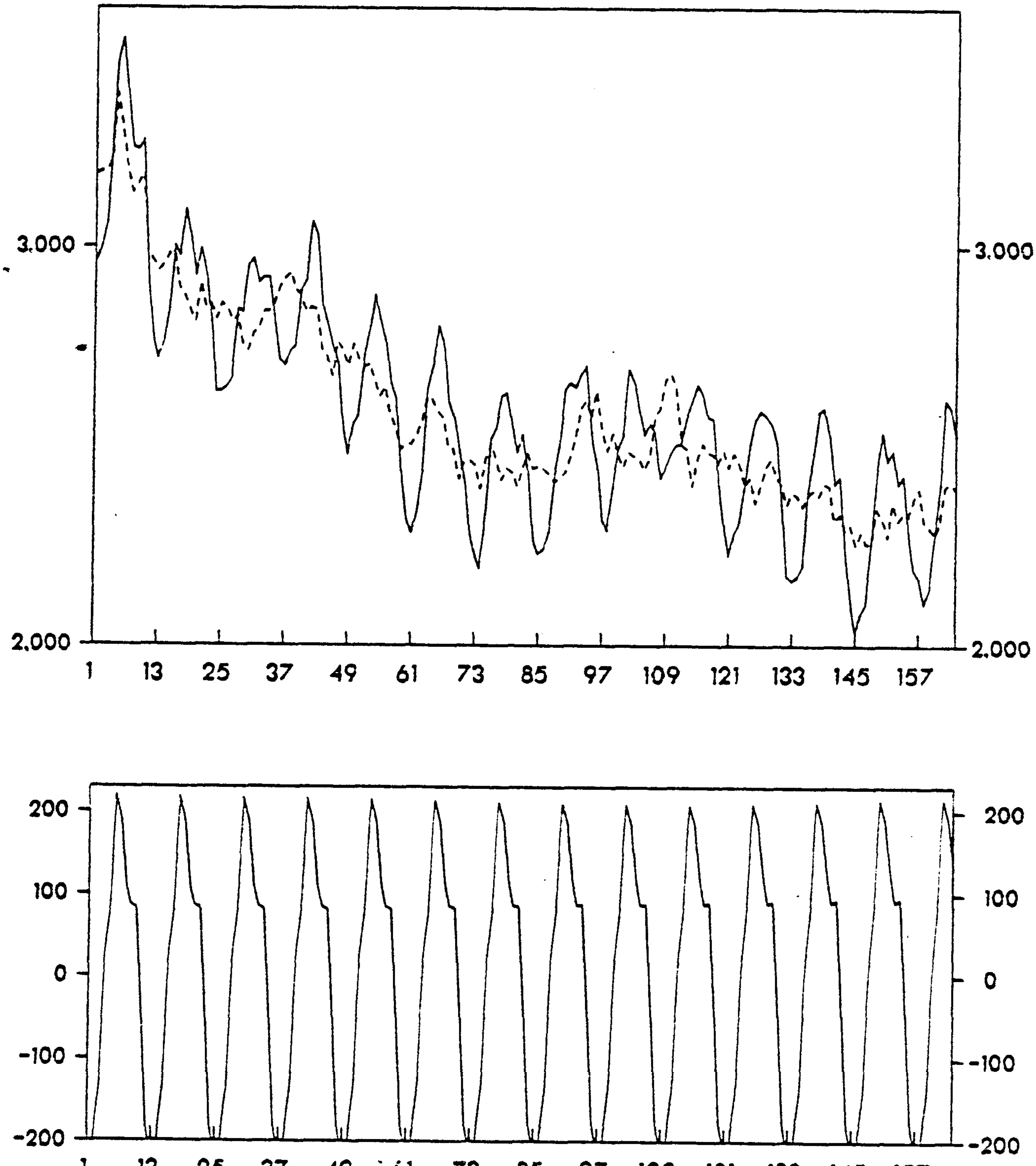


ENF16 X-11 LOG

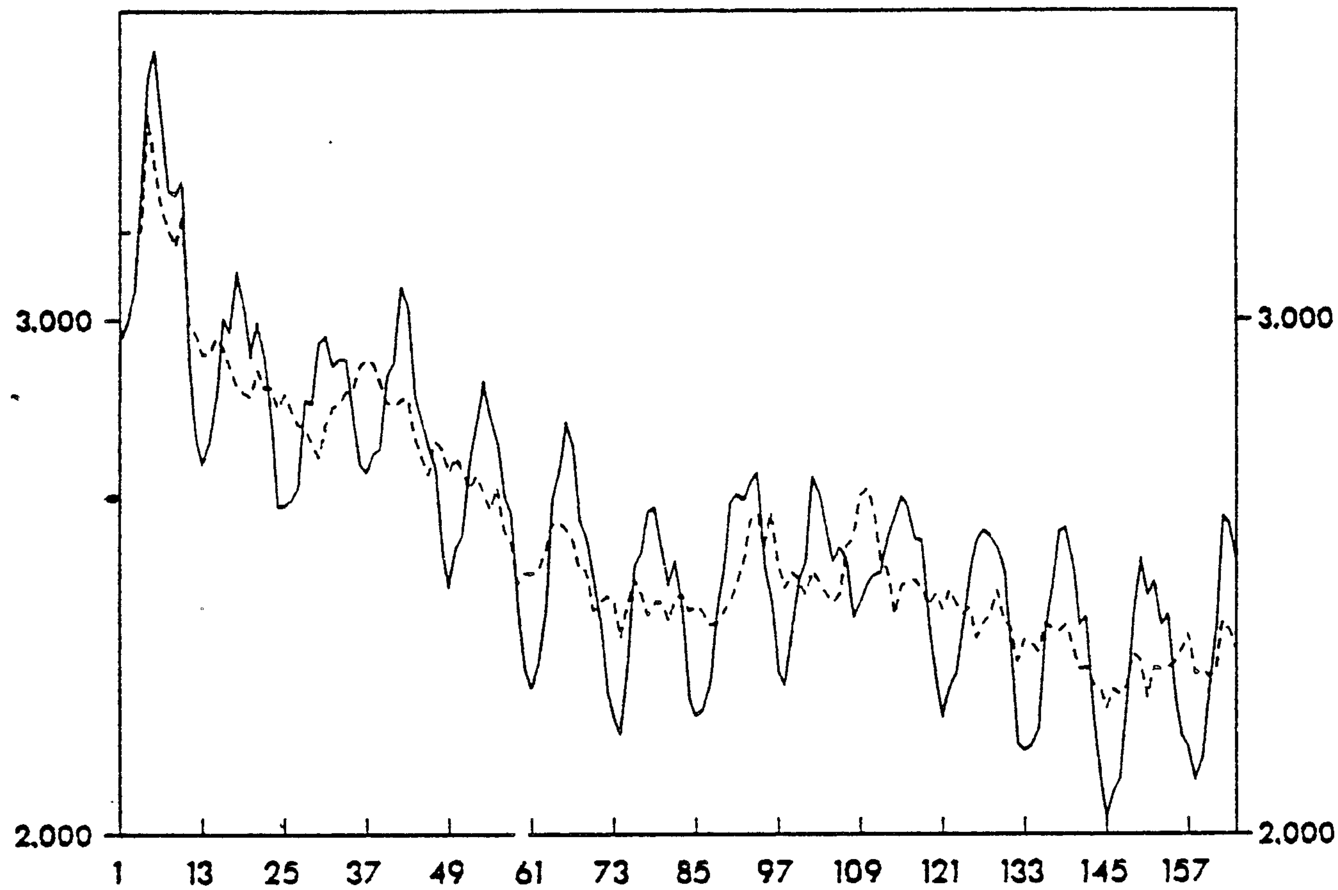


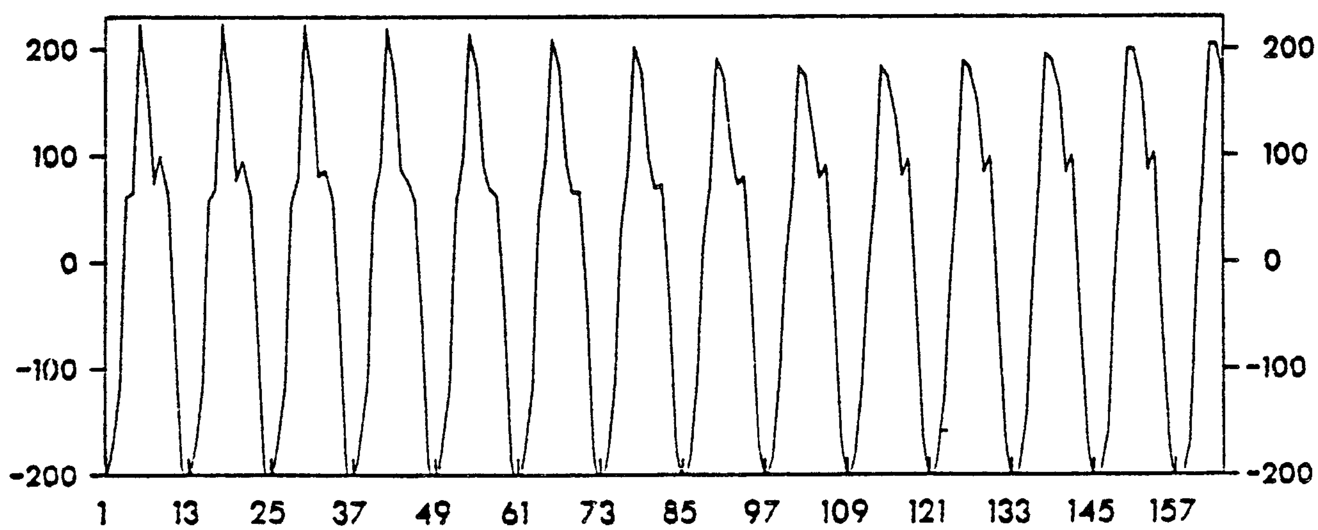


EAM20 MODEL BASED

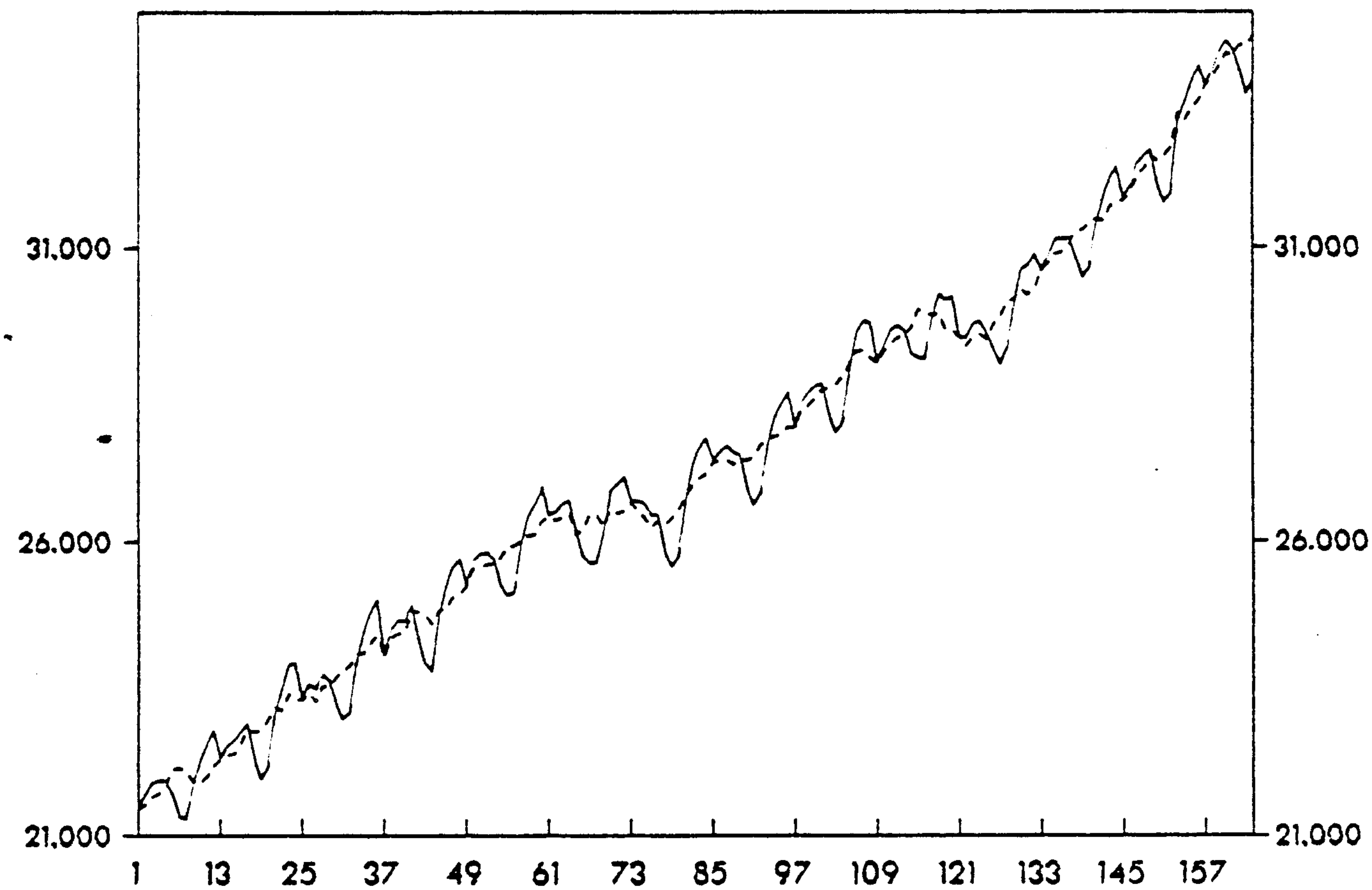


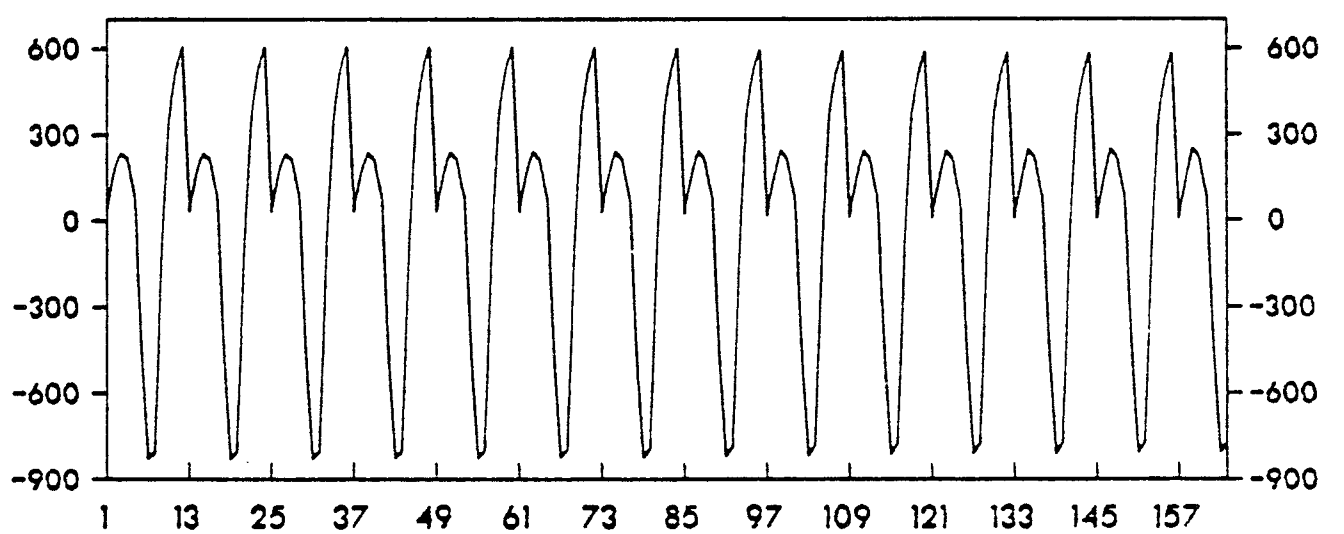
EAM20 X-11 LOG



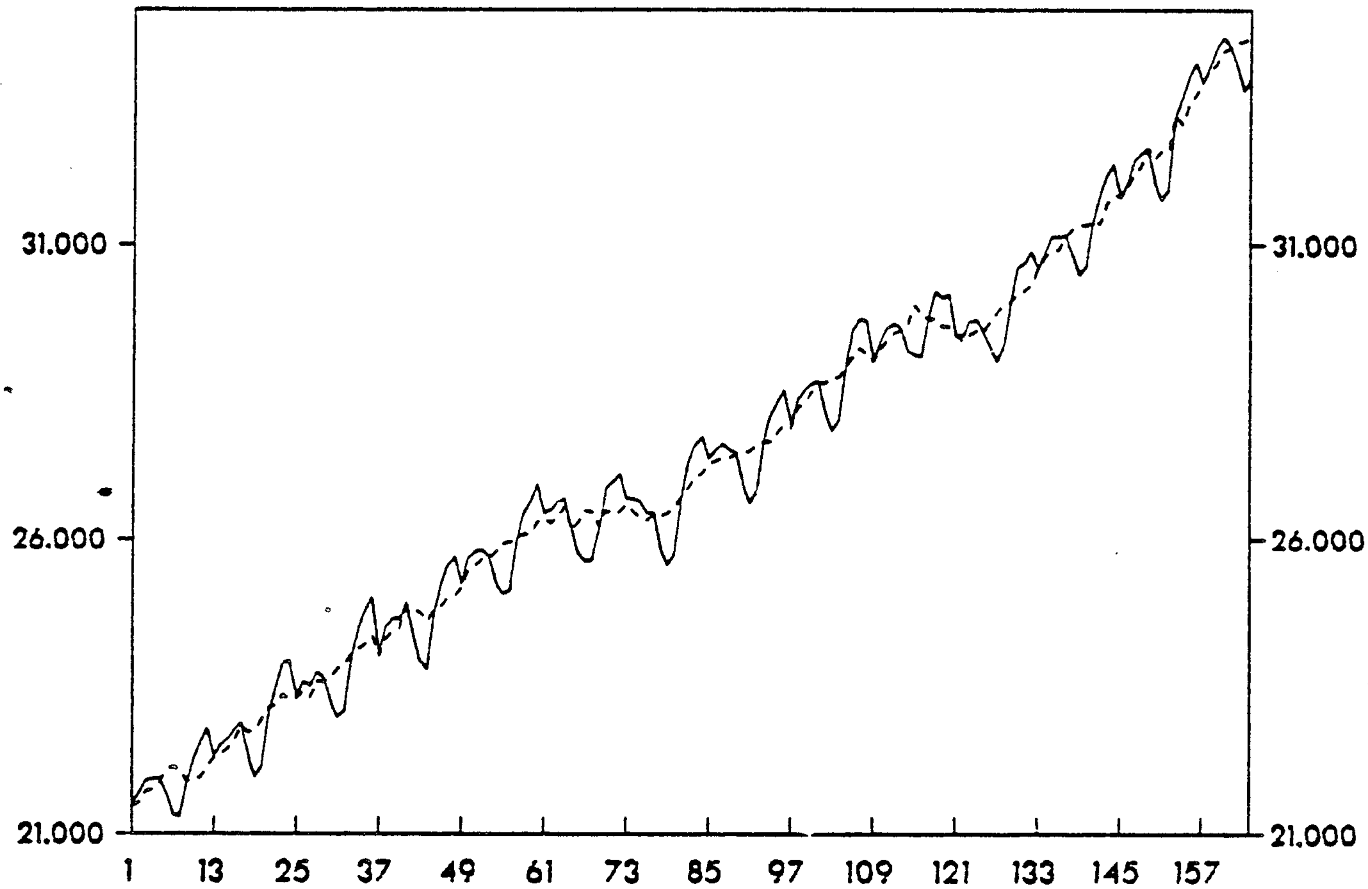


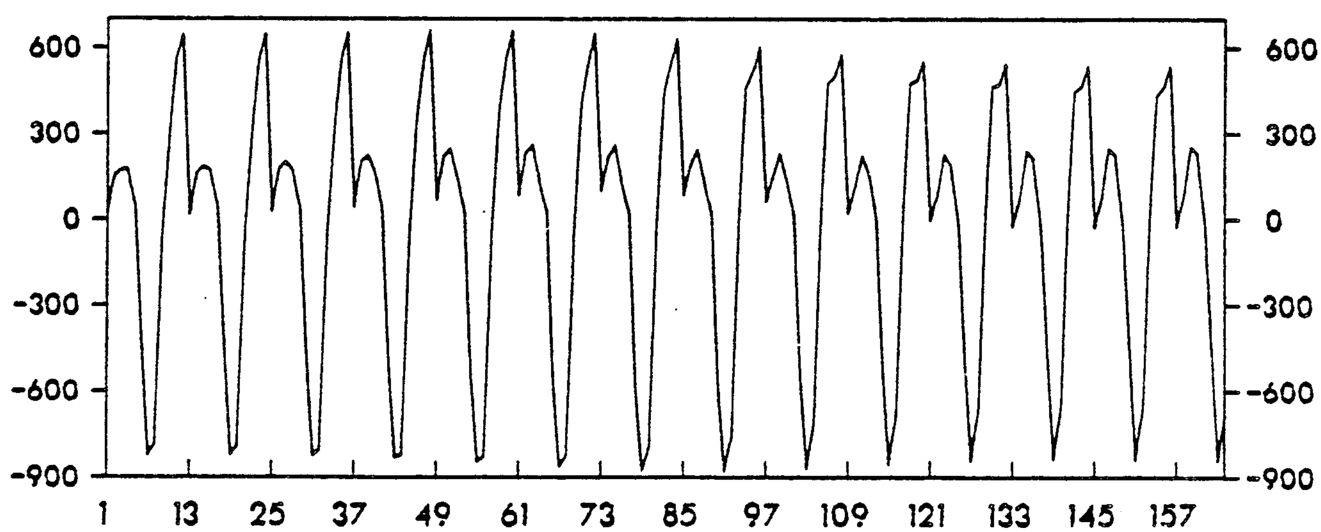
ENF20 MODEL BASED



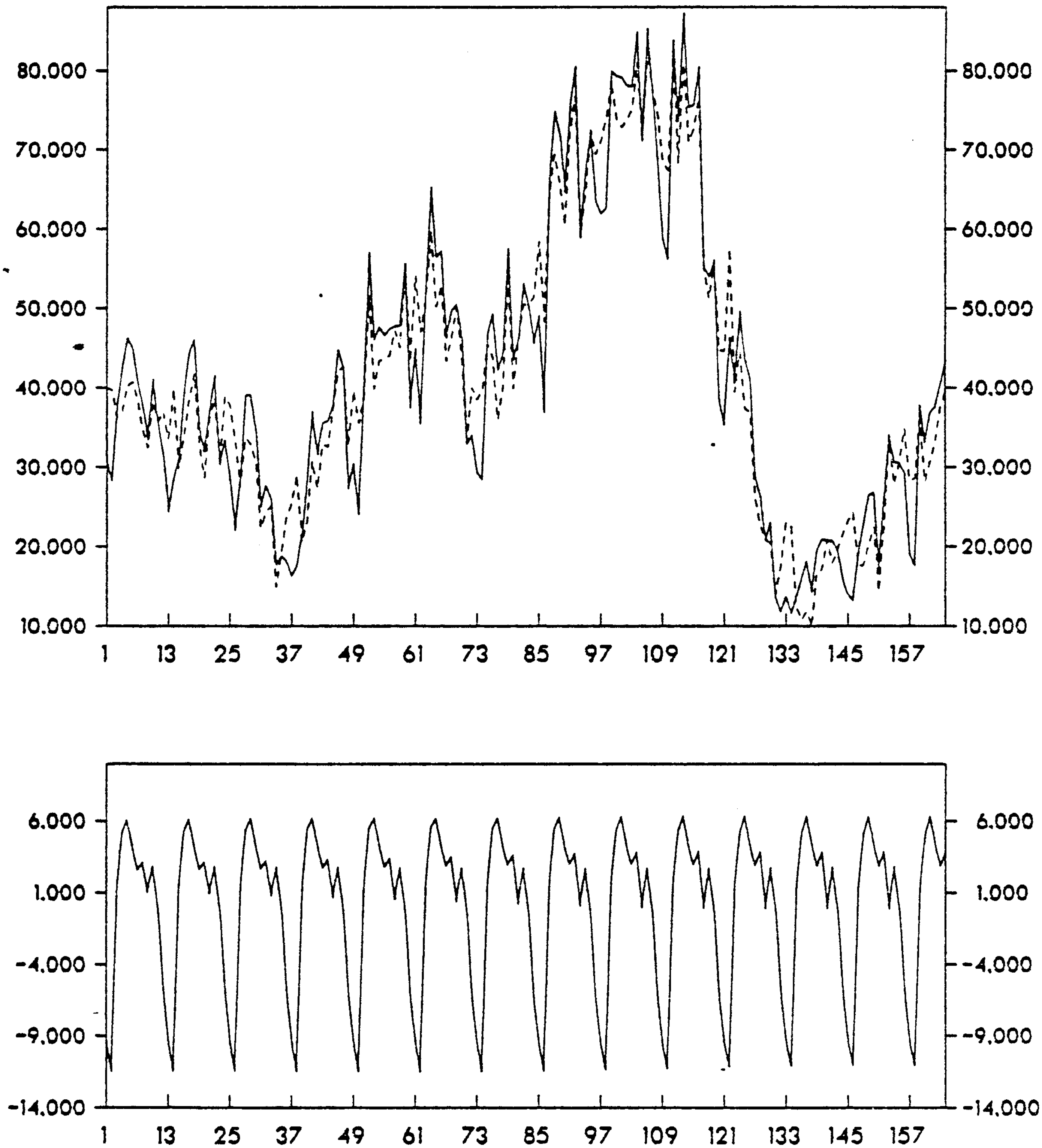


ENF20 X-11 LOG

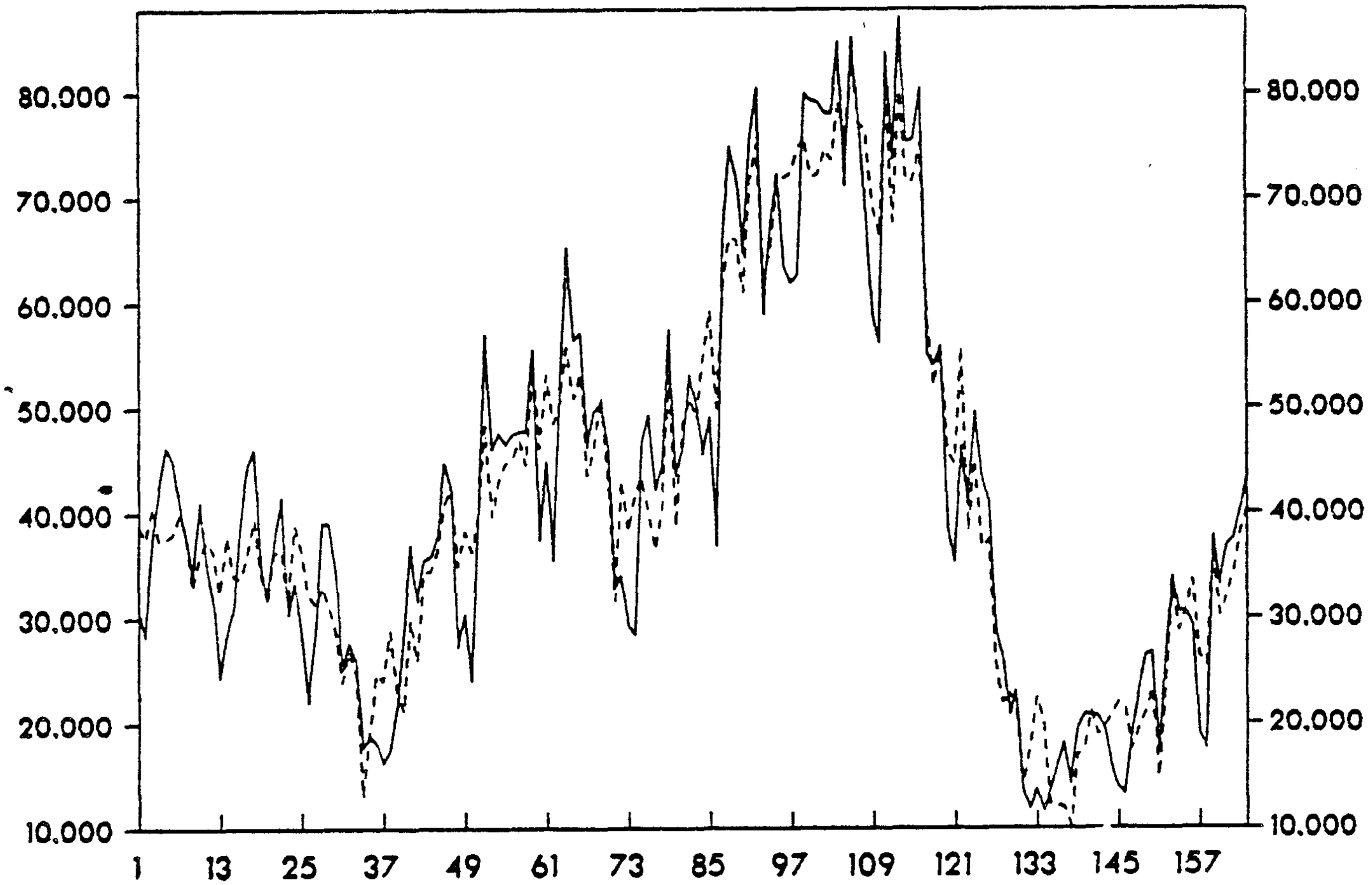


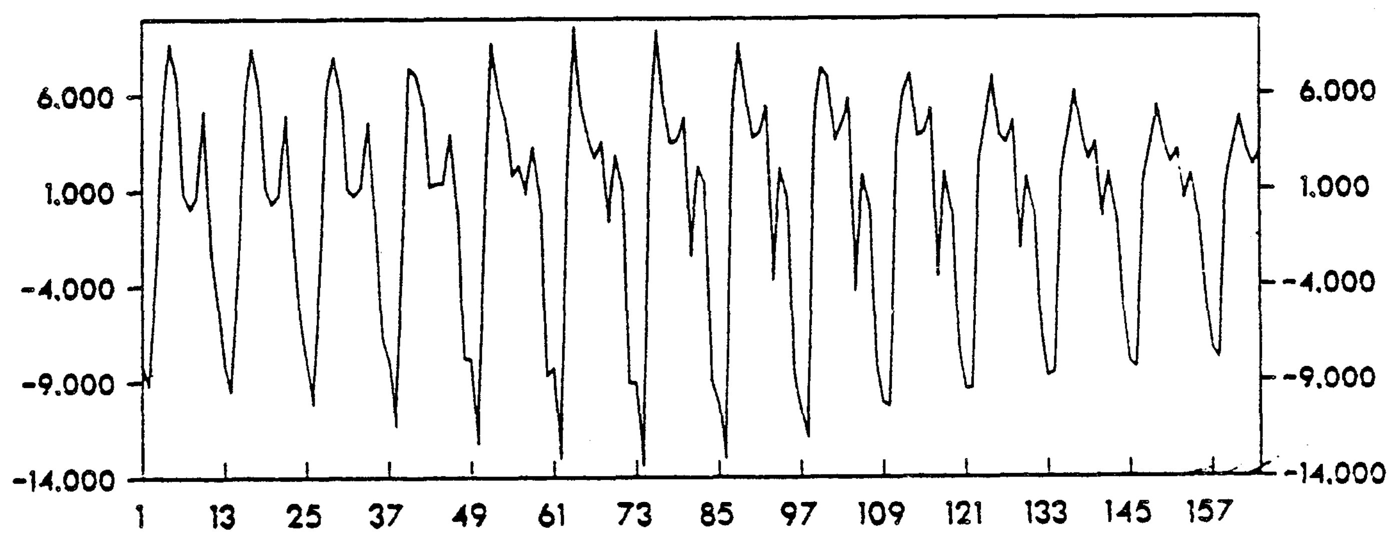


TOTAL 5 OR MORE UNIT HOUSING STARTS MODEL BASED

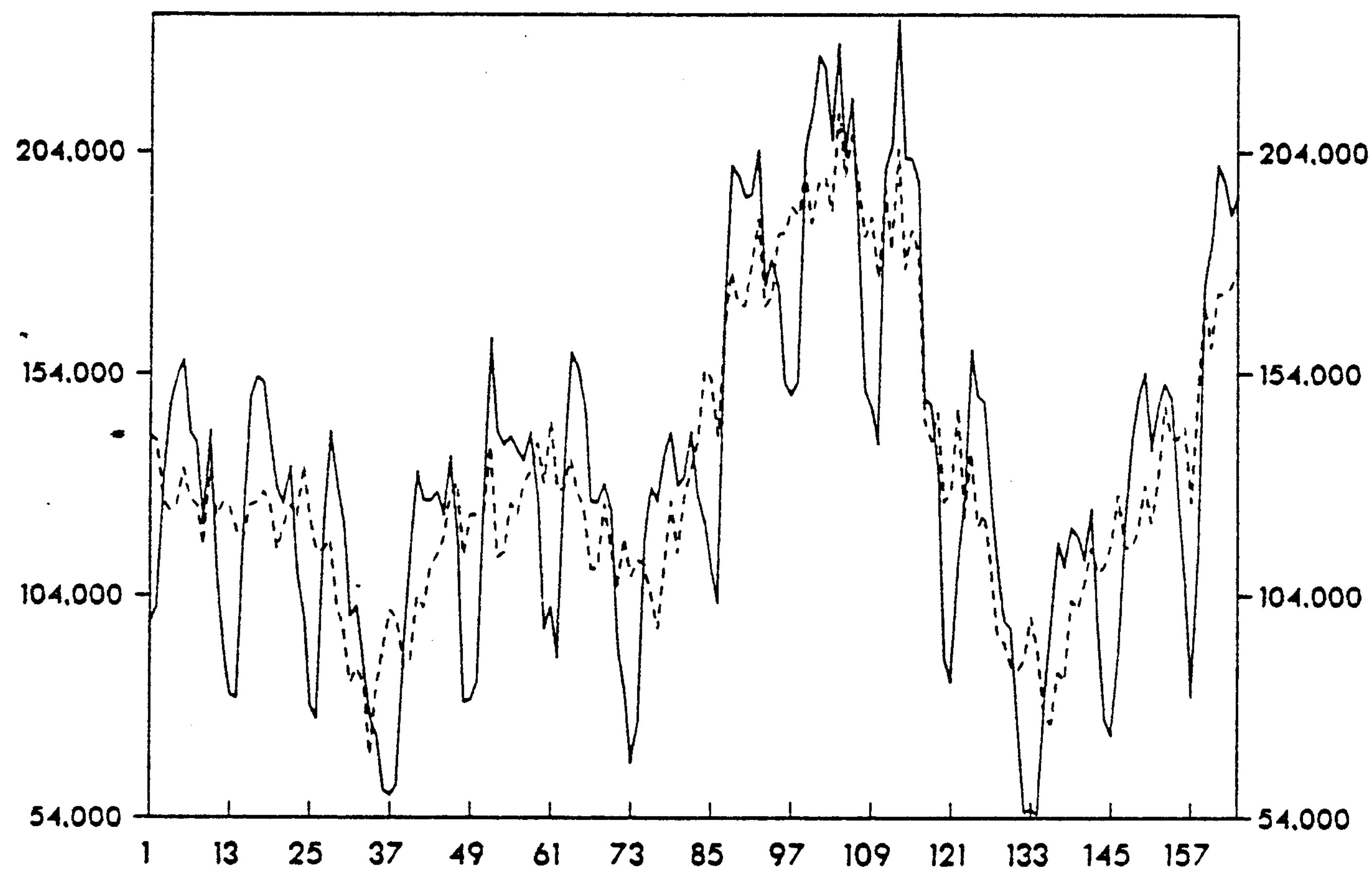


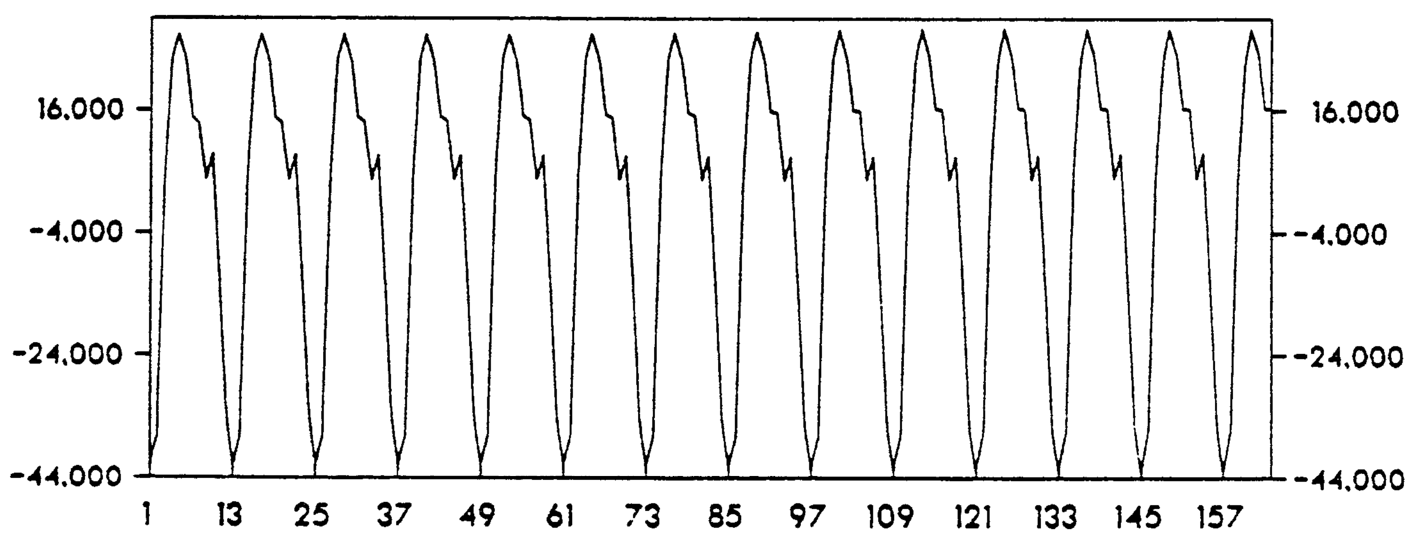
TOTAL 5 OR MORE UNIT HOUSING STARTS X-11



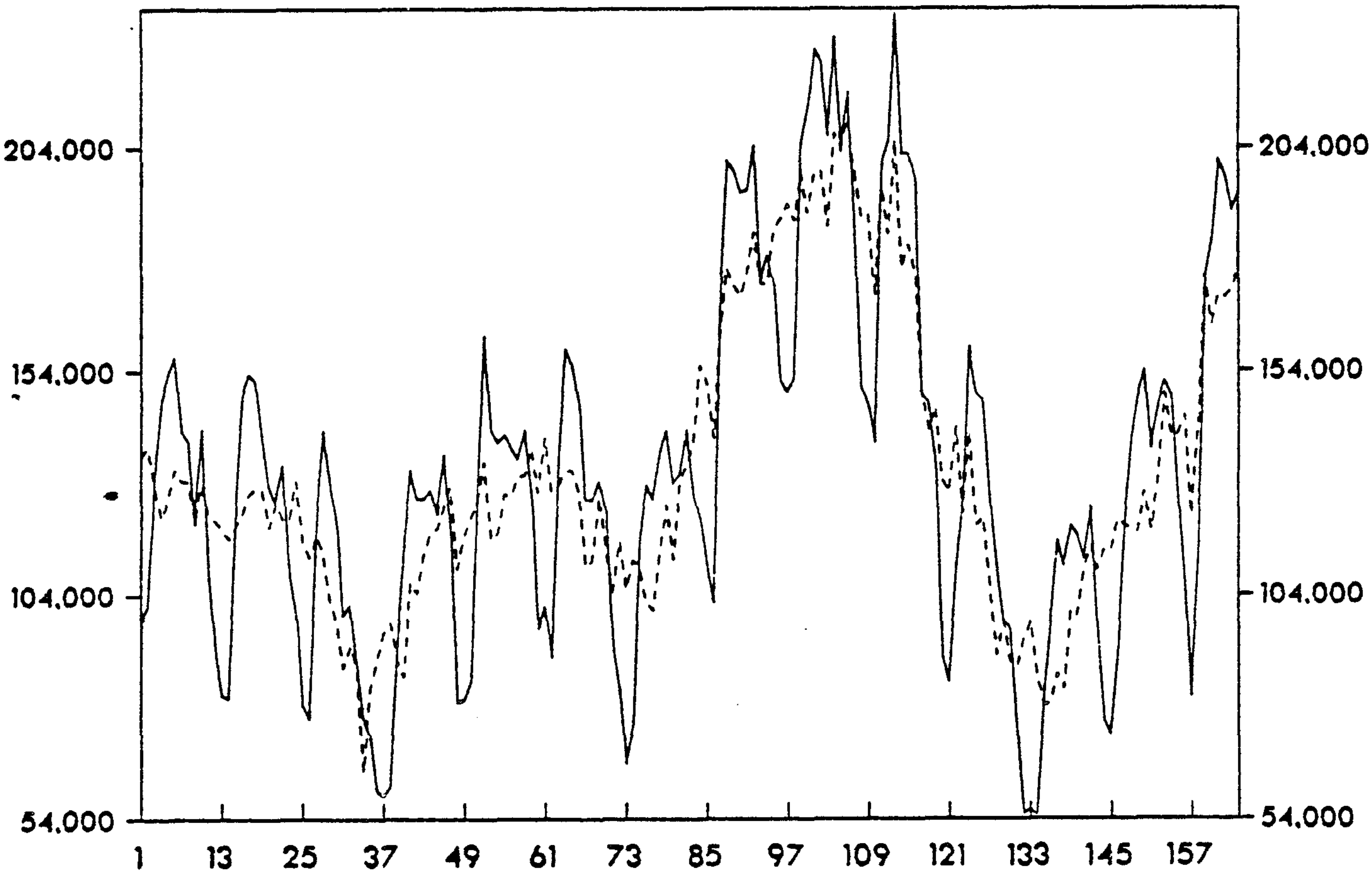


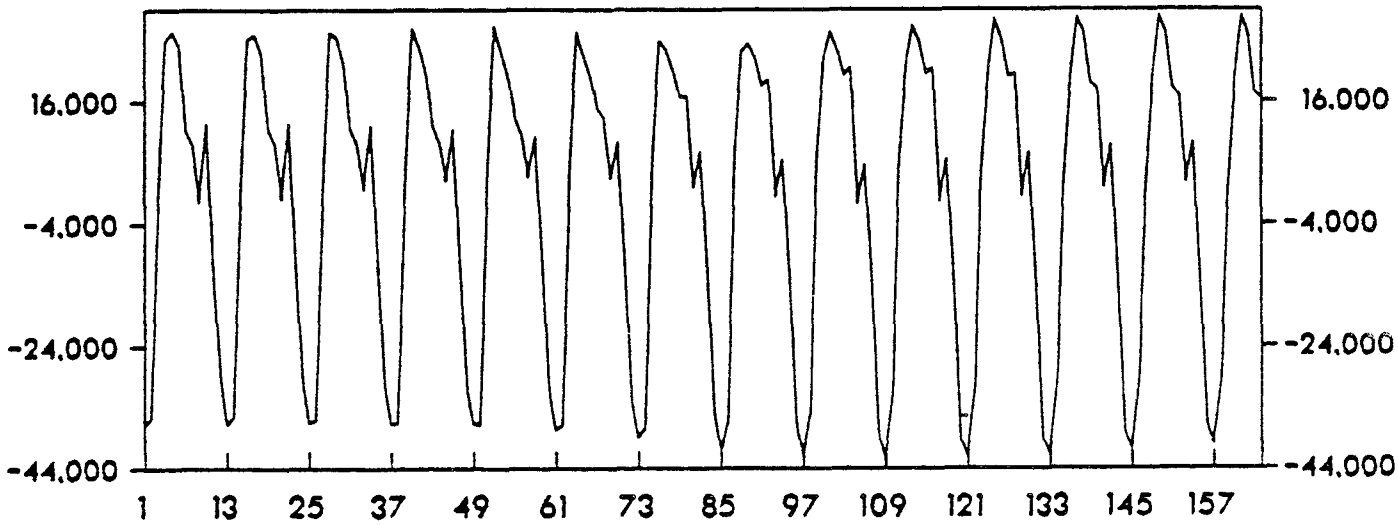
TOTAL HOUSING STARTS MODEL BASED



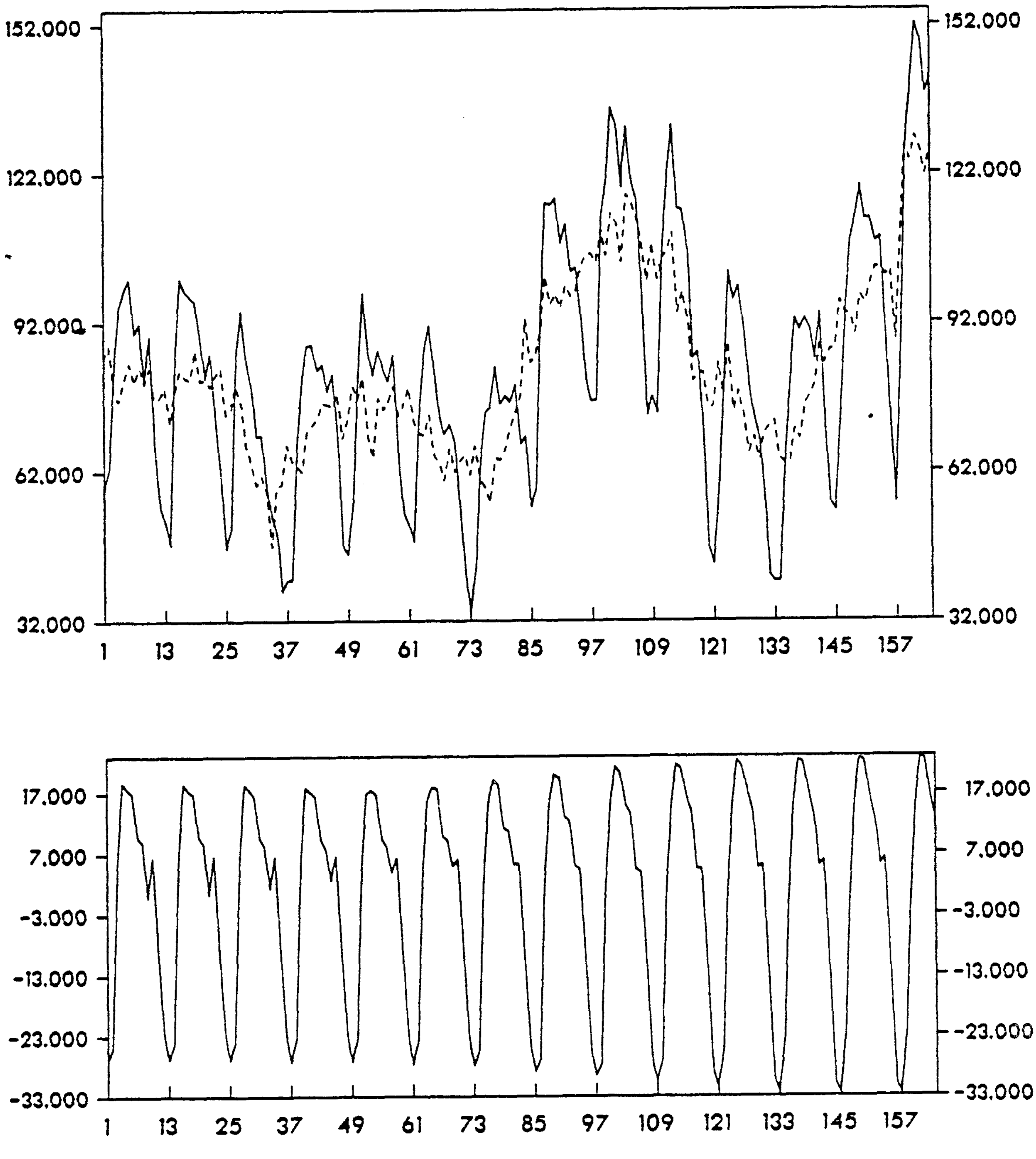


TOTAL HOUSING STARTS X-11

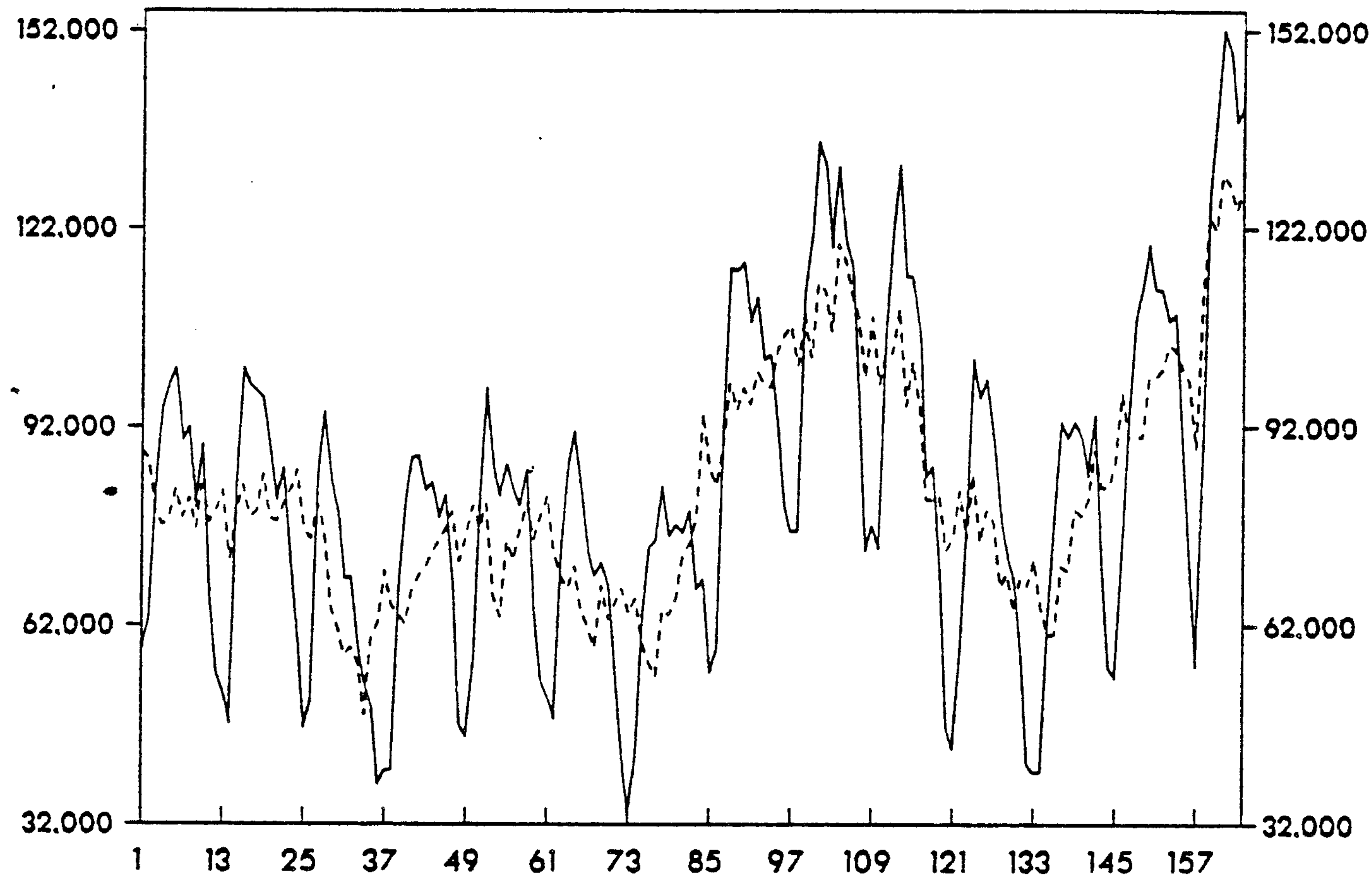


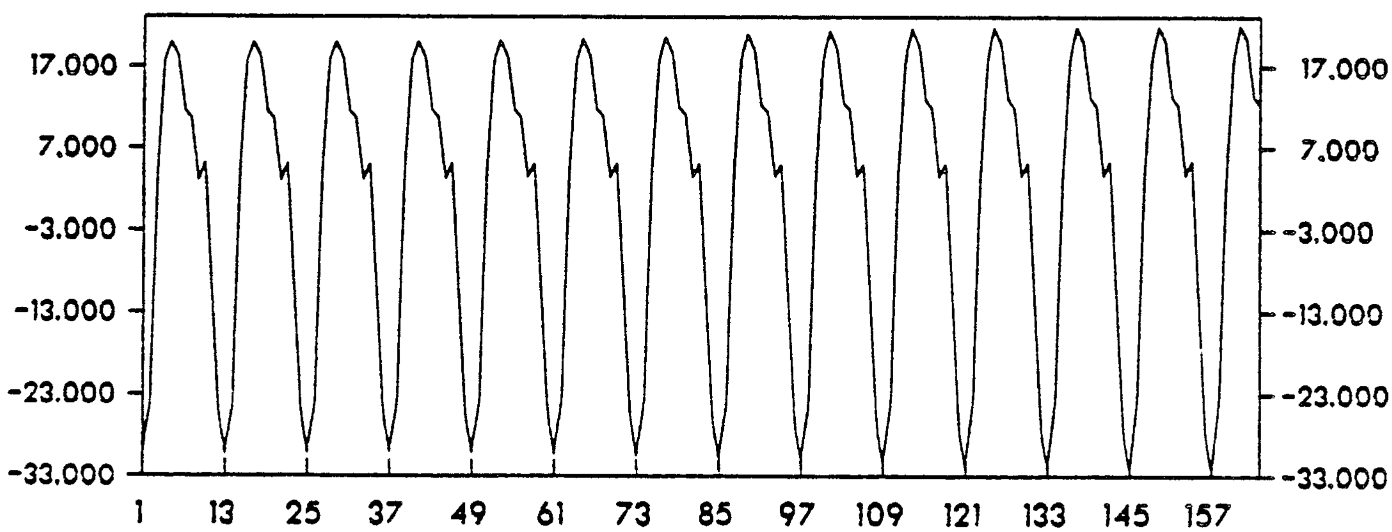


TOTAL SINGLE FAMILY HOUSING STARTS X-11

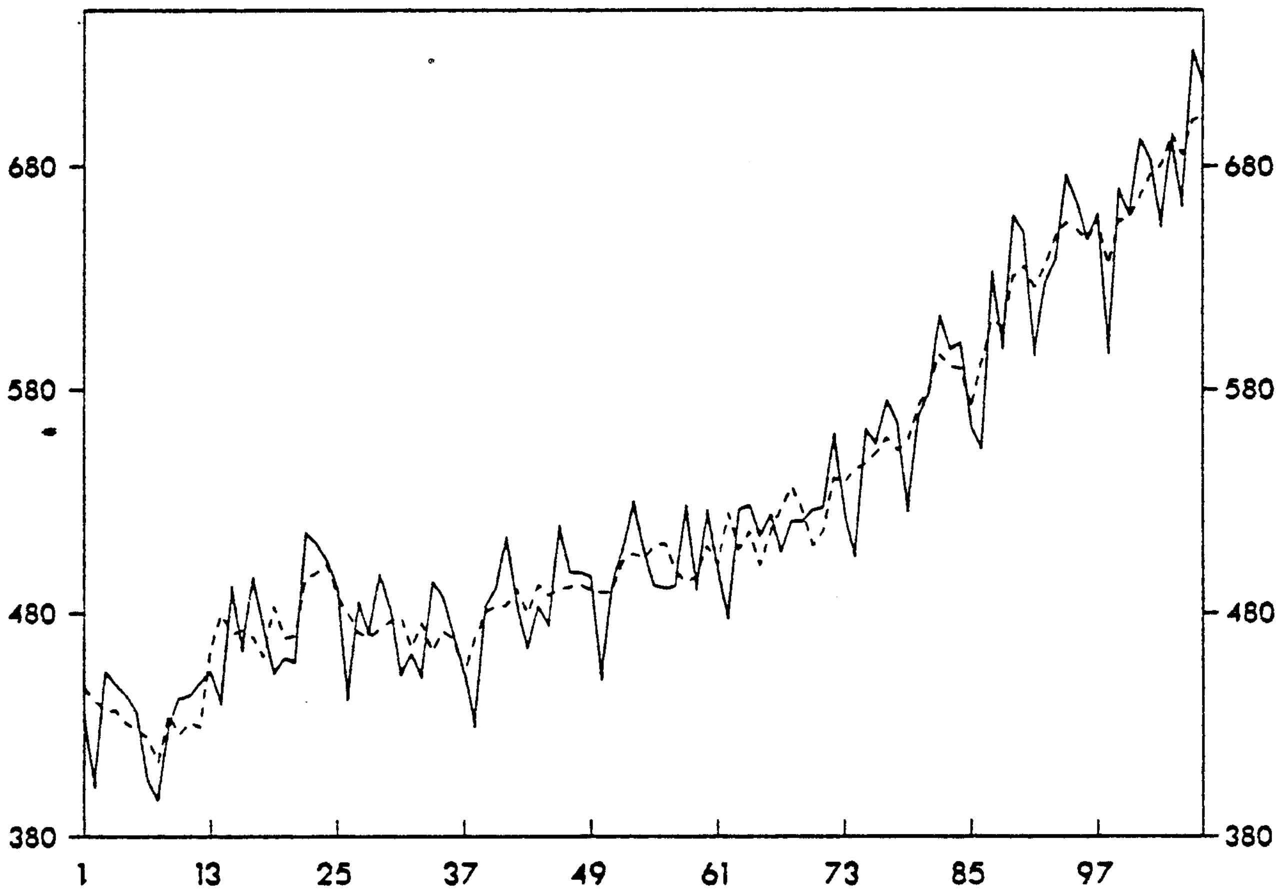


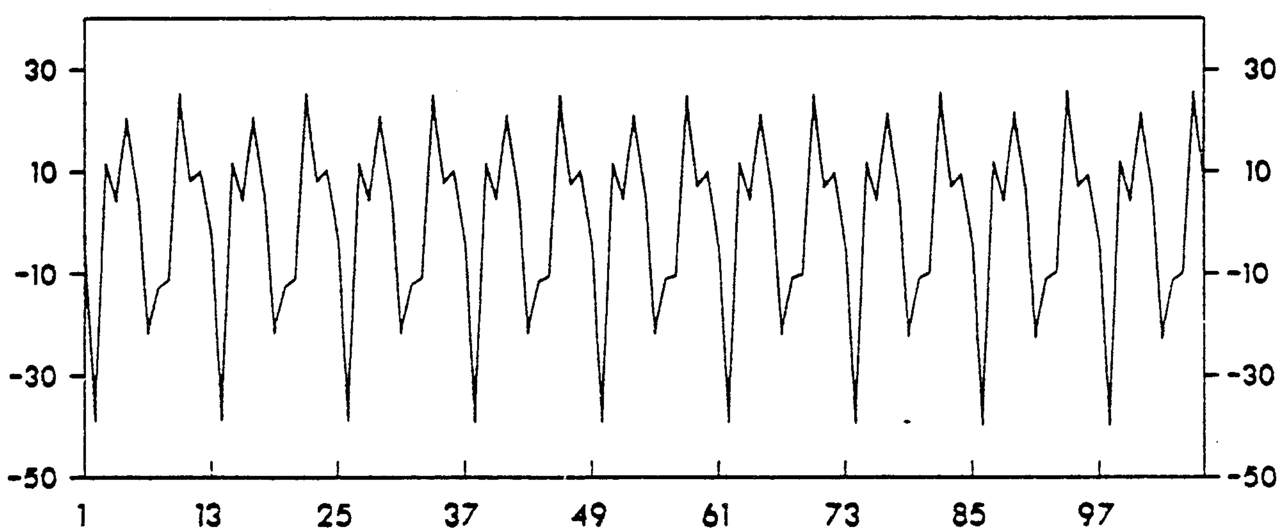
TOTAL SINGLE FAMILY HOUSING STARTS MODEL BASED.



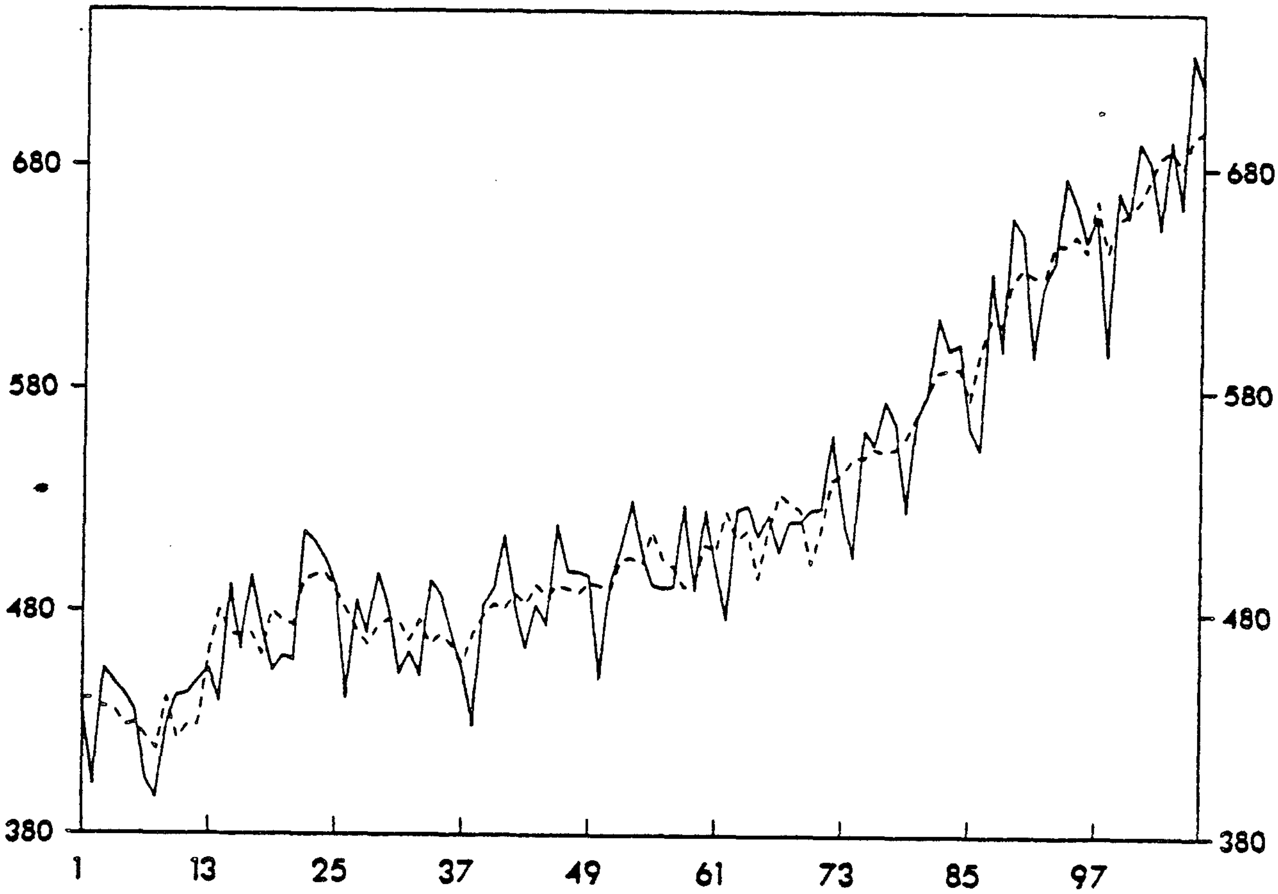


721000 MODEL BASED

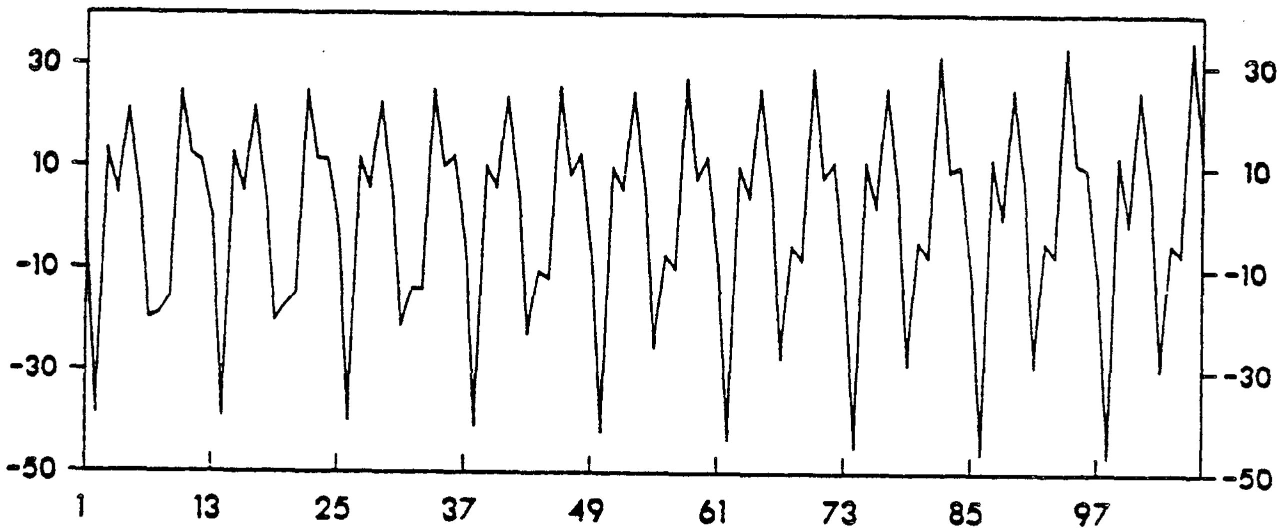




721000 X-11

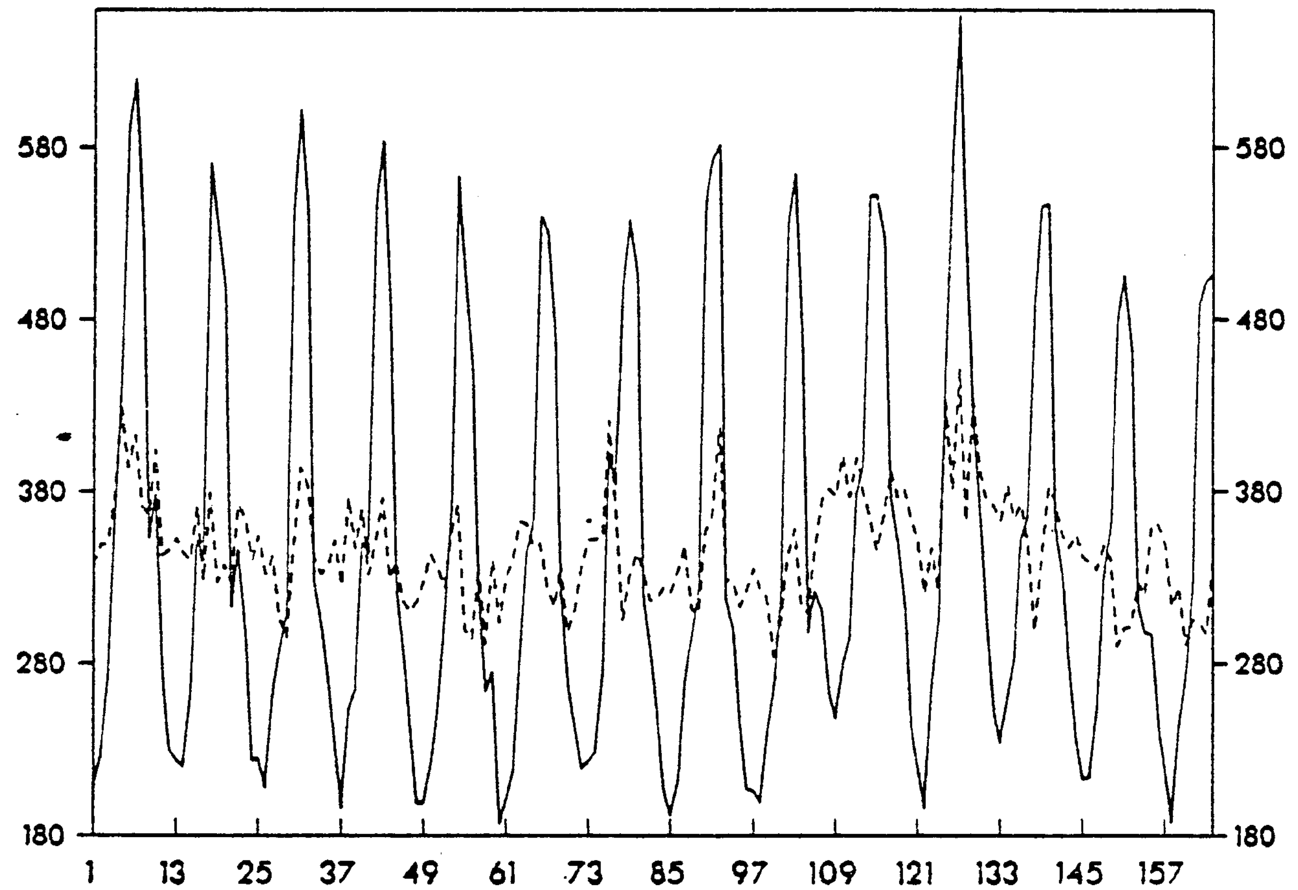


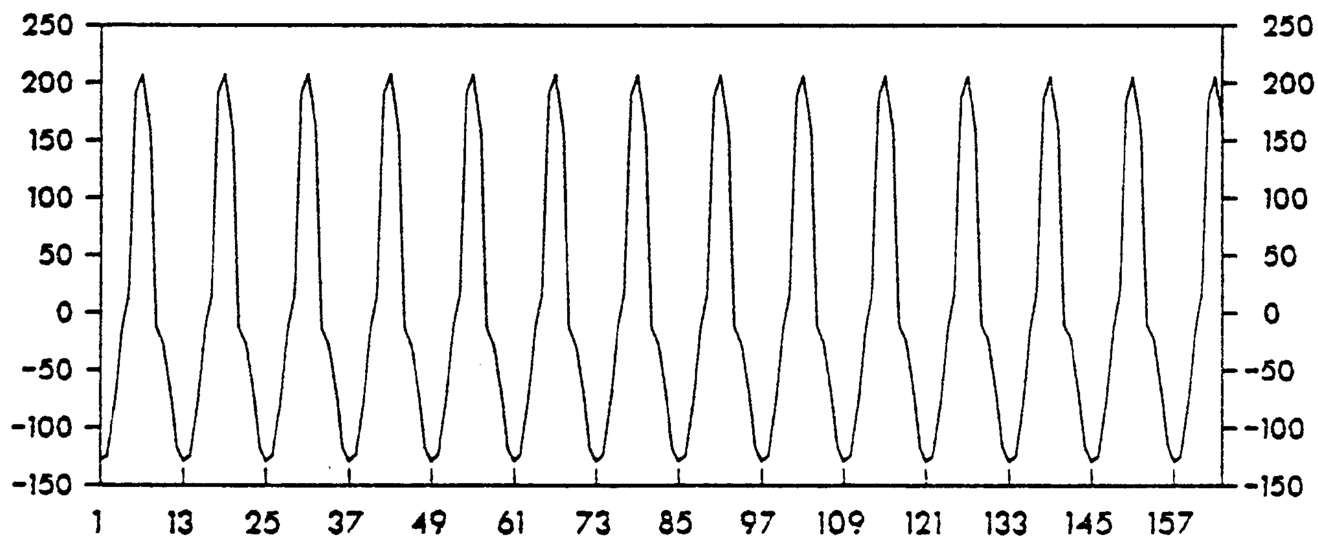
680
580
480
380
1
13
25
37
49
61
73
85
97



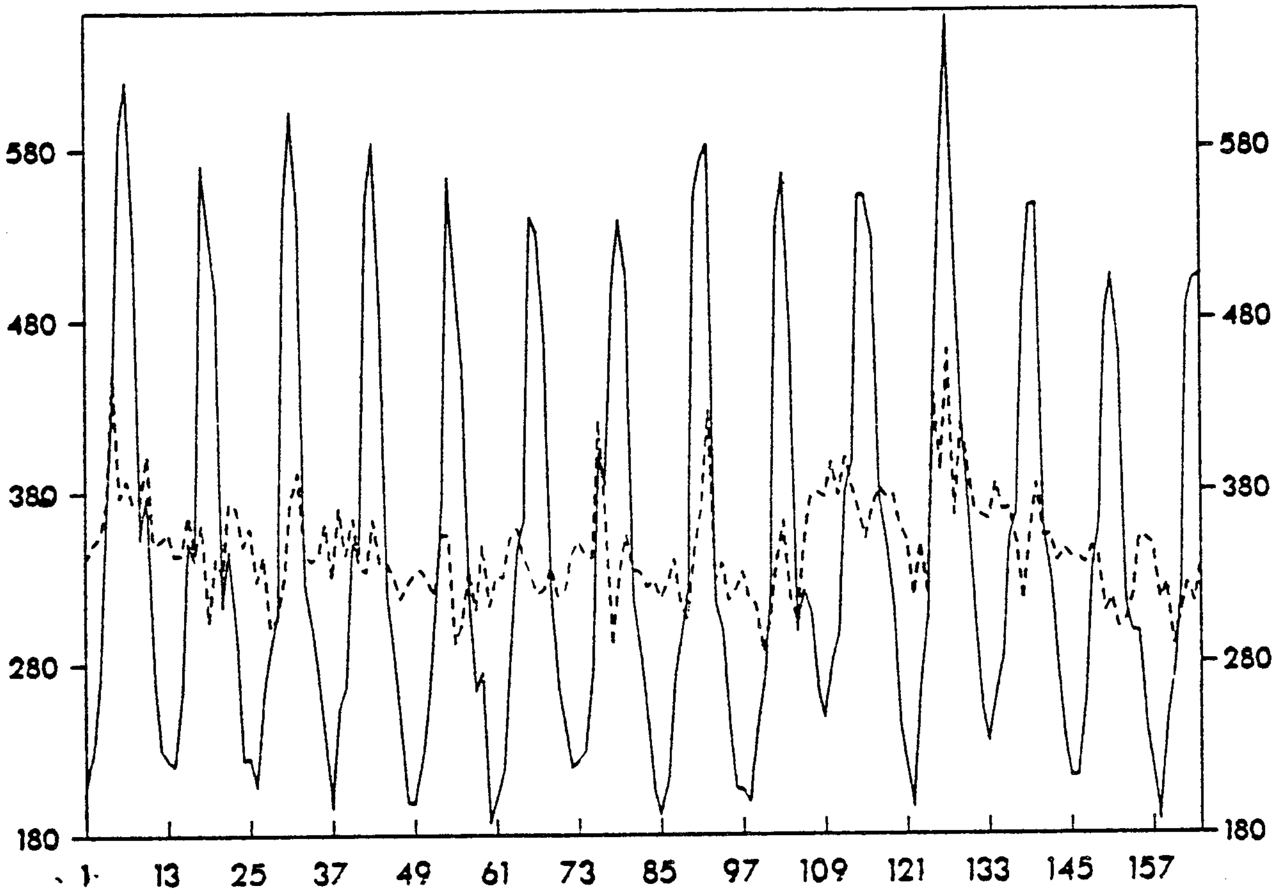
30
10
-10
-30
-50
1
13
25
37
49
61
73
85
97

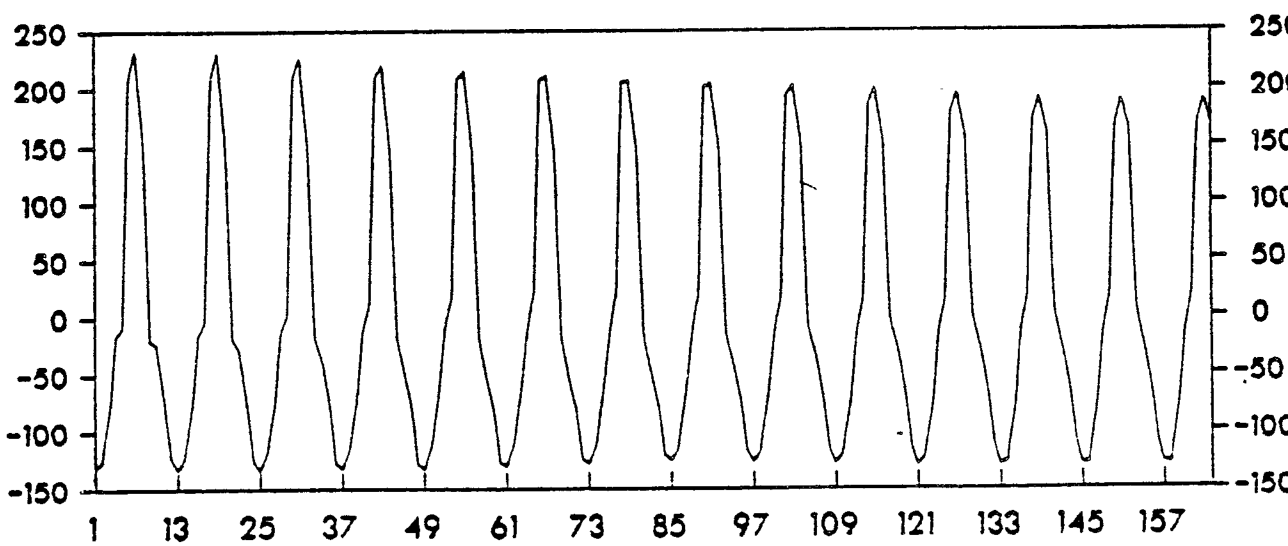
EAM16 MODEL BASED



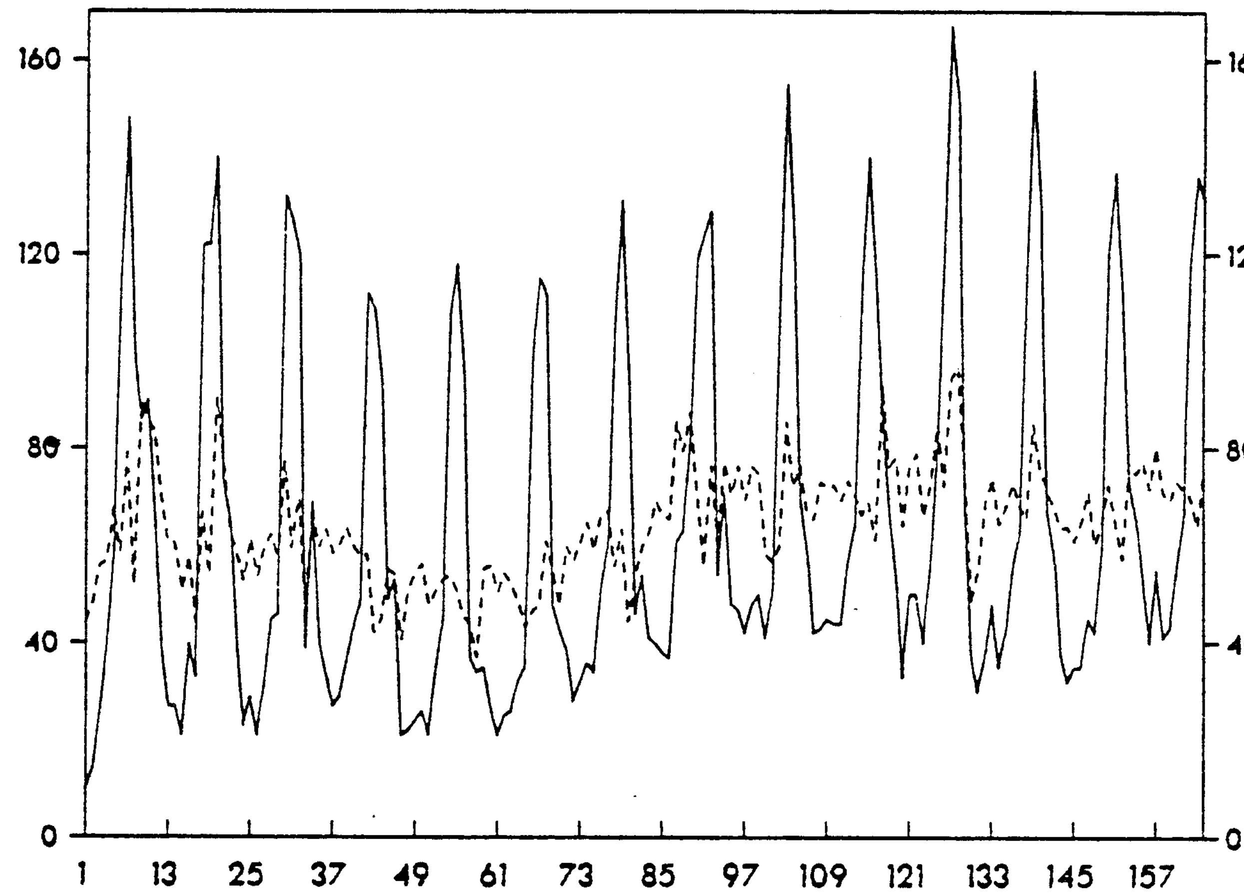


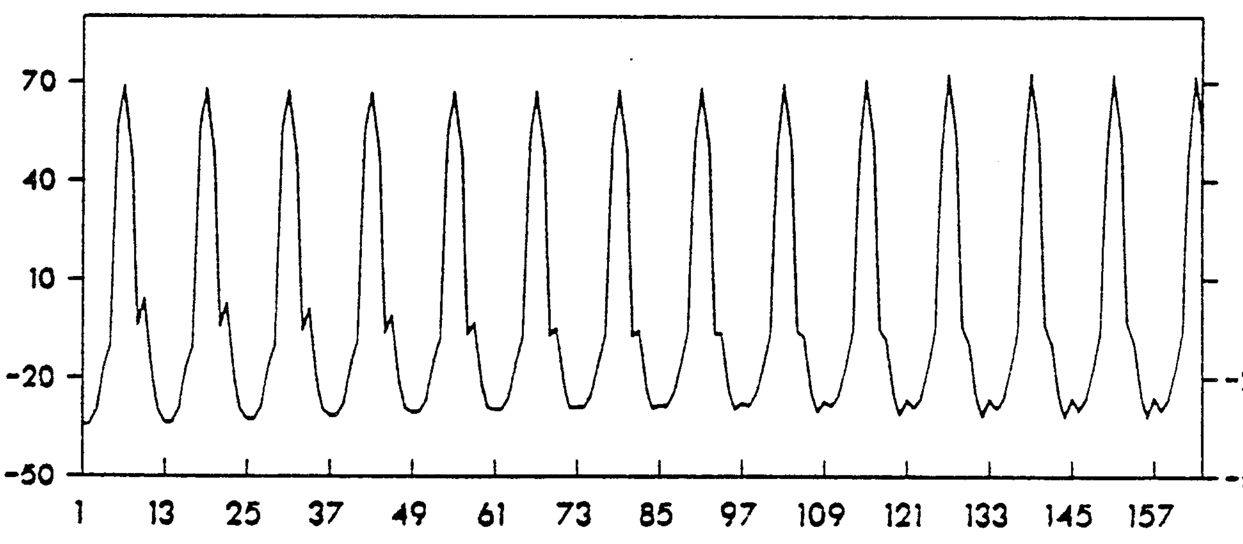
EAM16 X-11 LOG



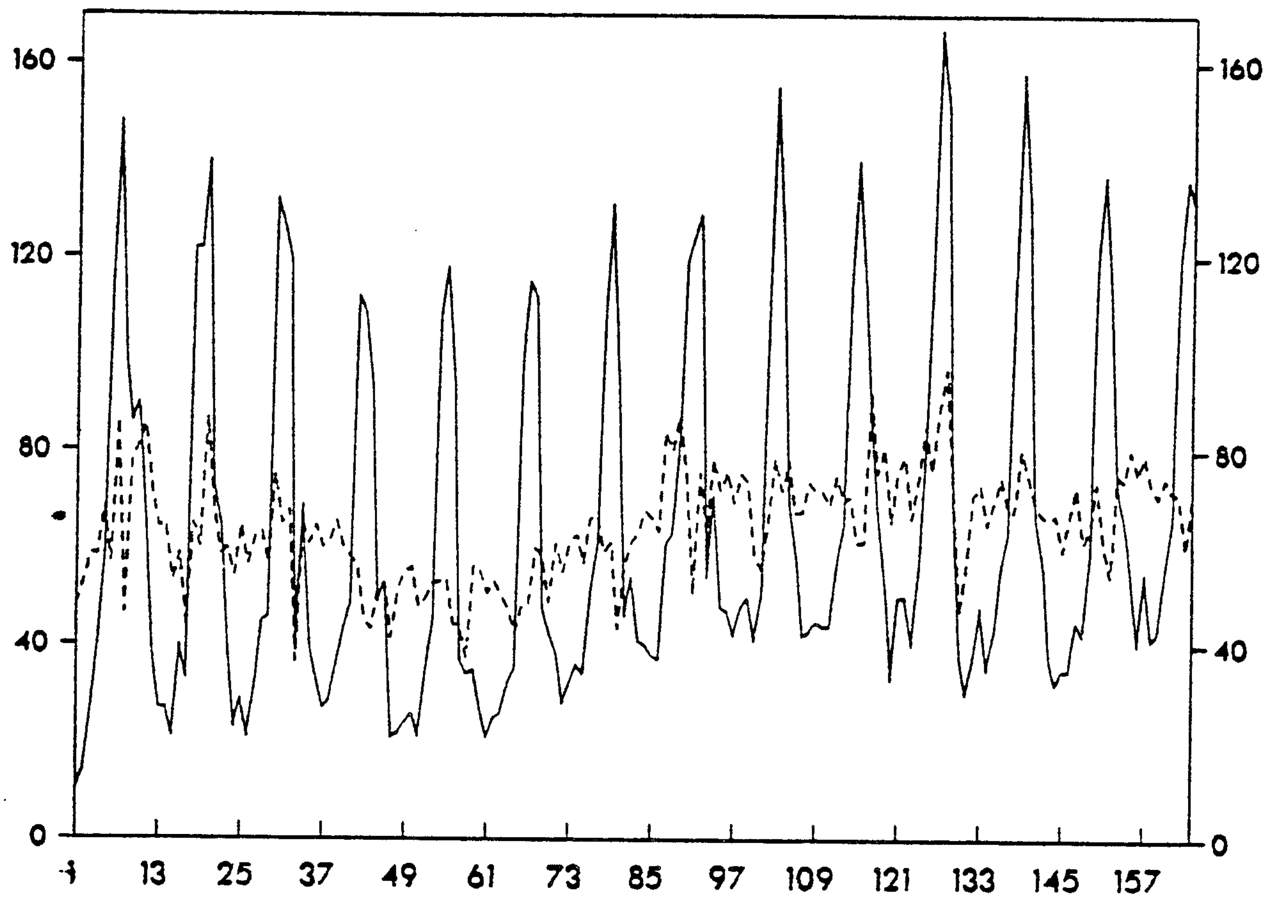


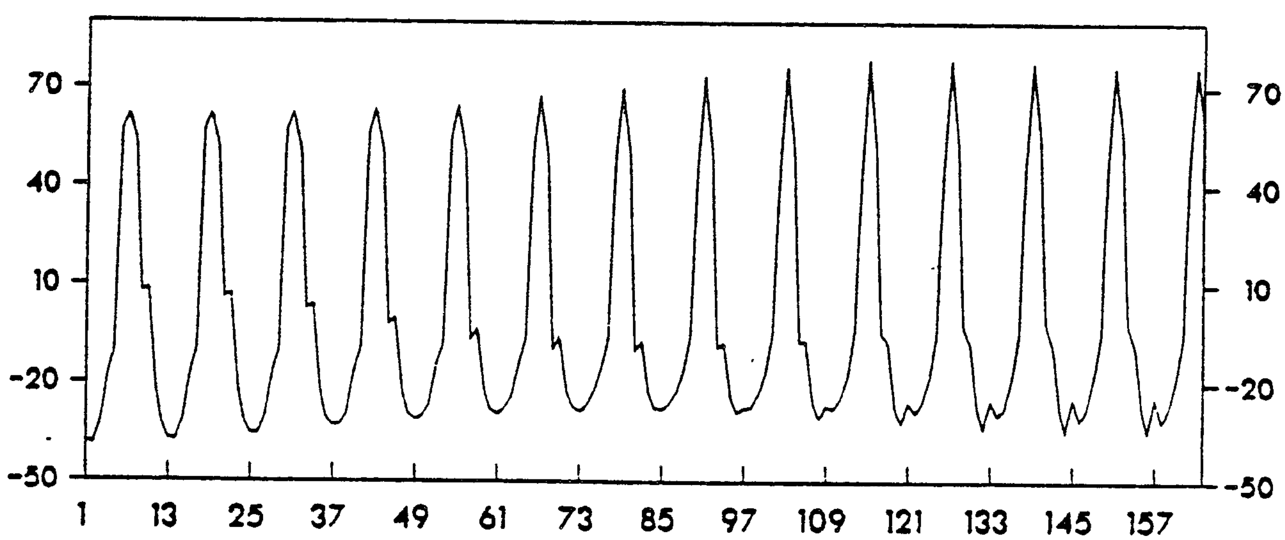
EAF16 MODEL BASED



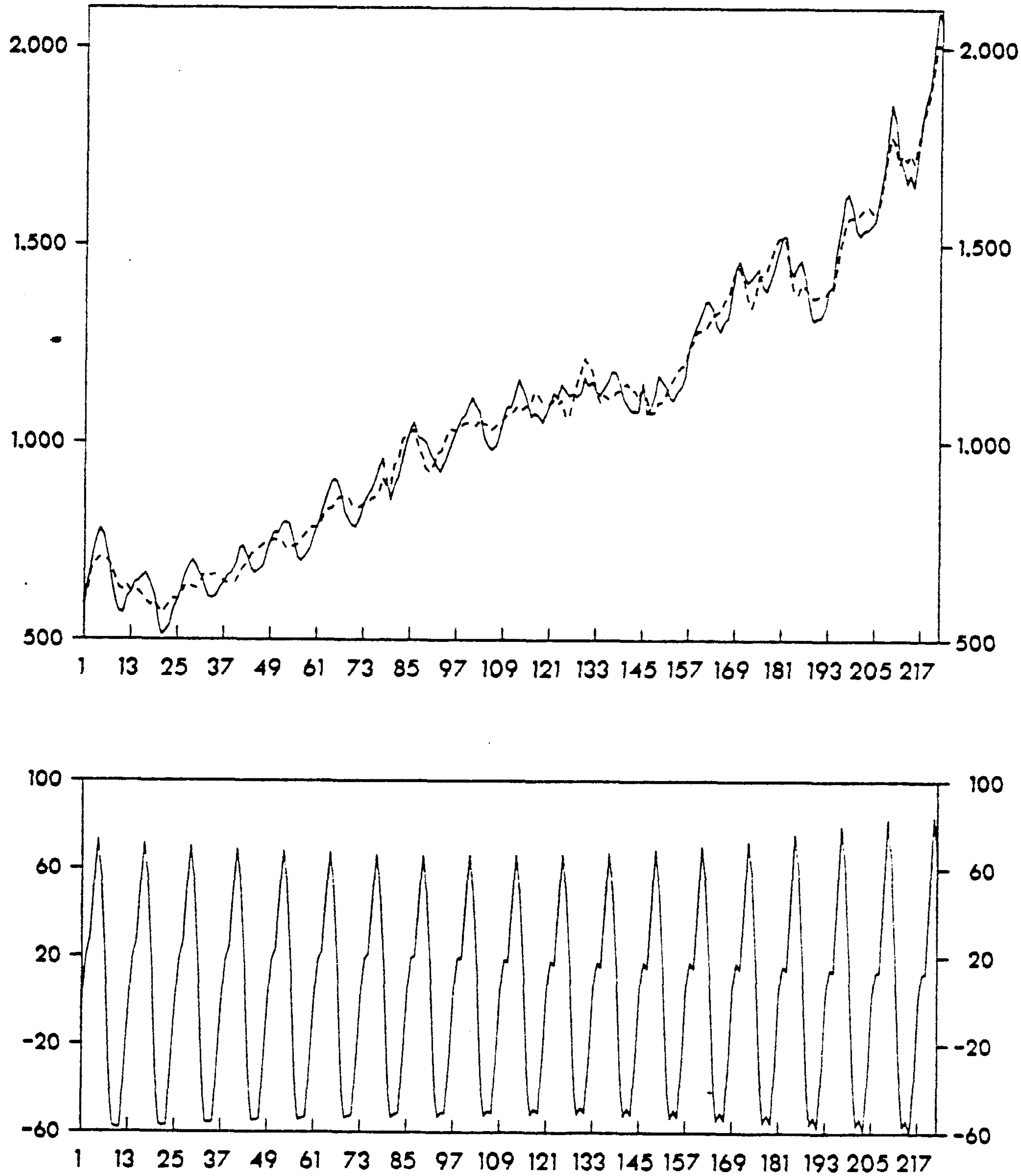


EAF16 X-11 LOG

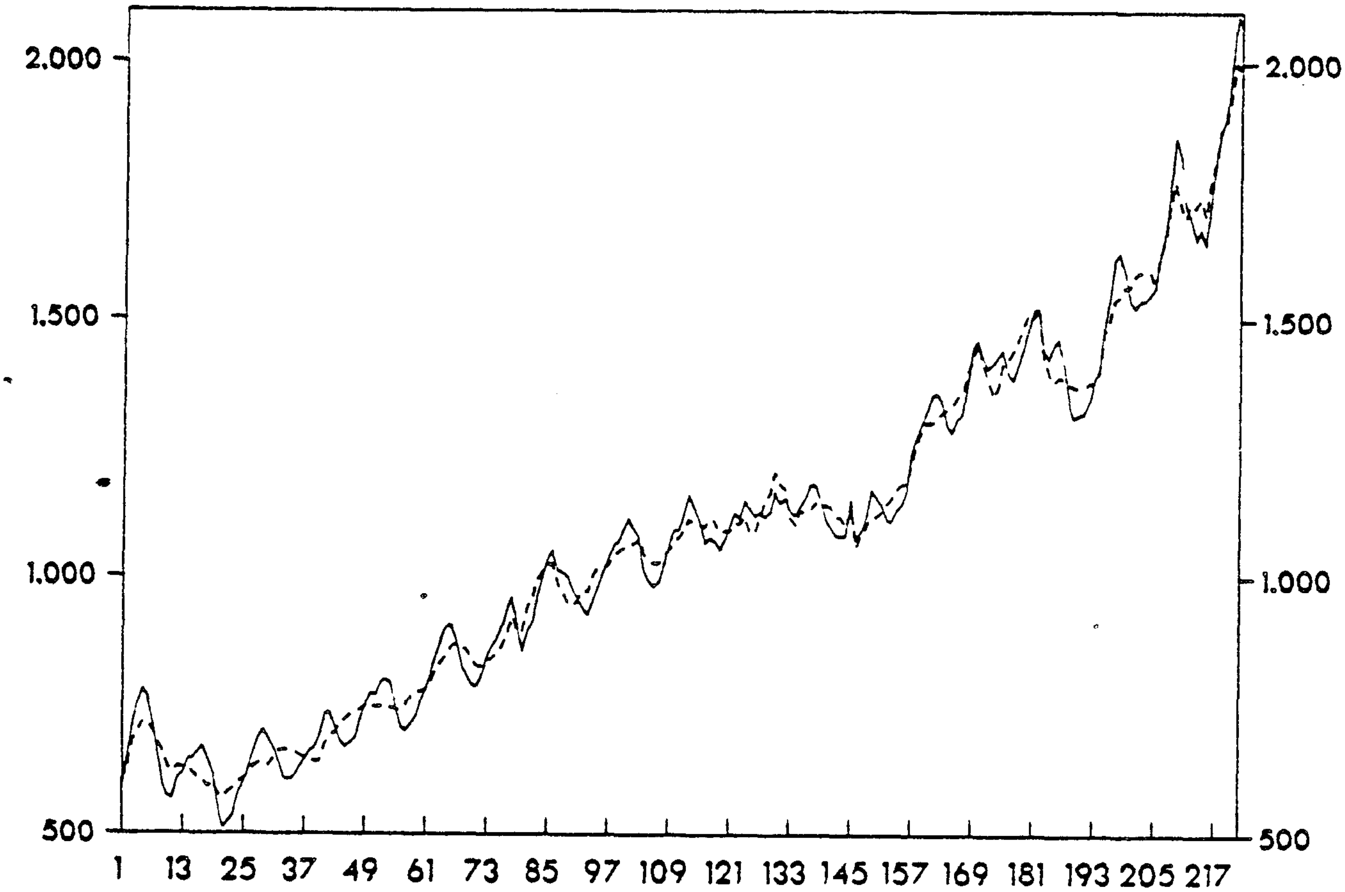


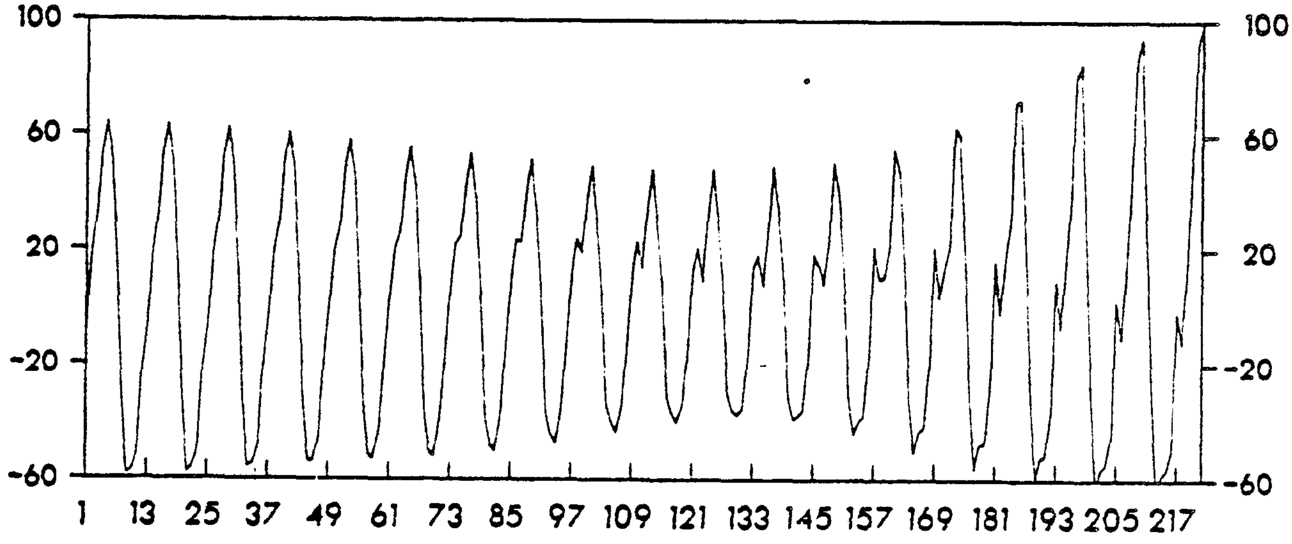


S35T1 MODEL BASED

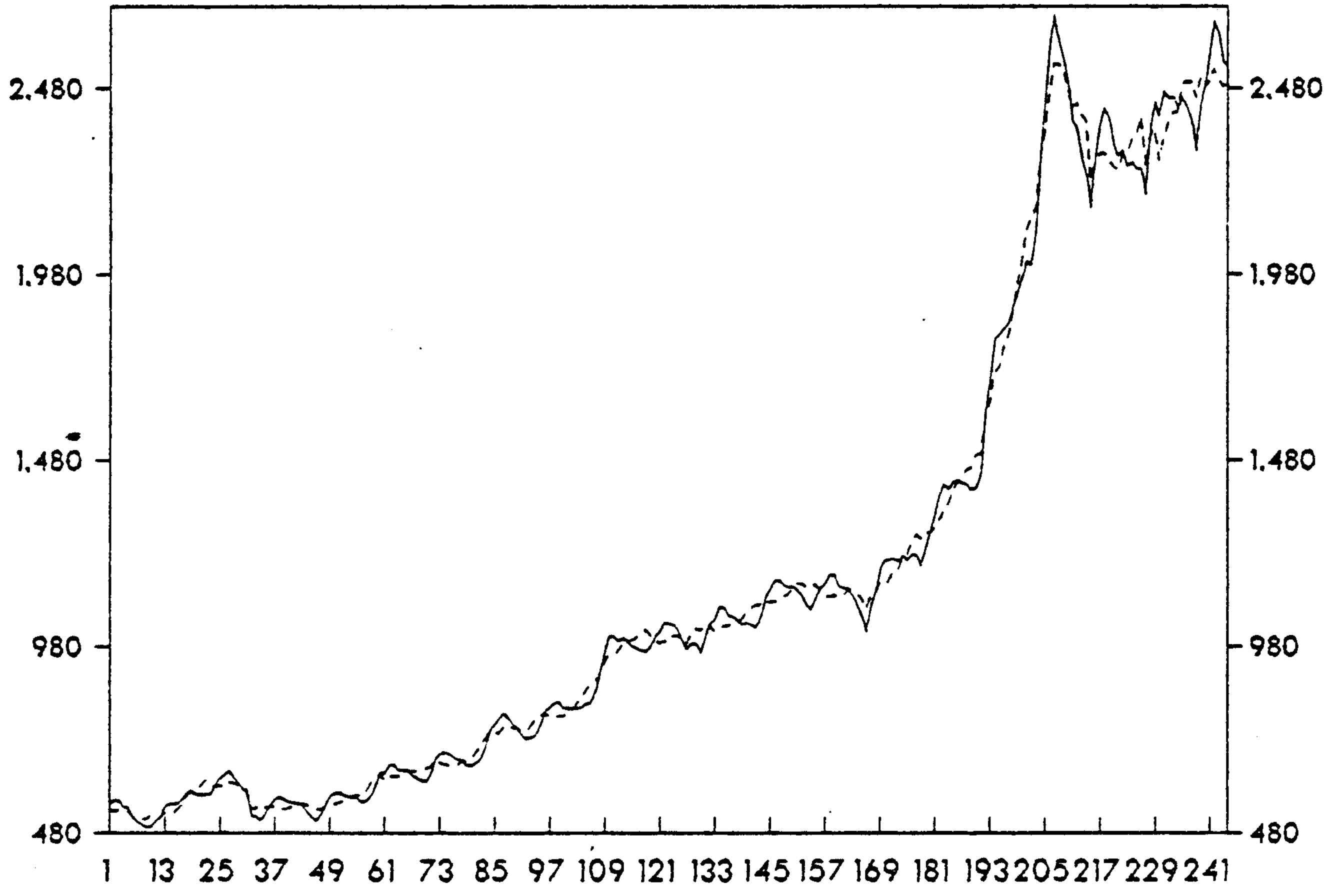


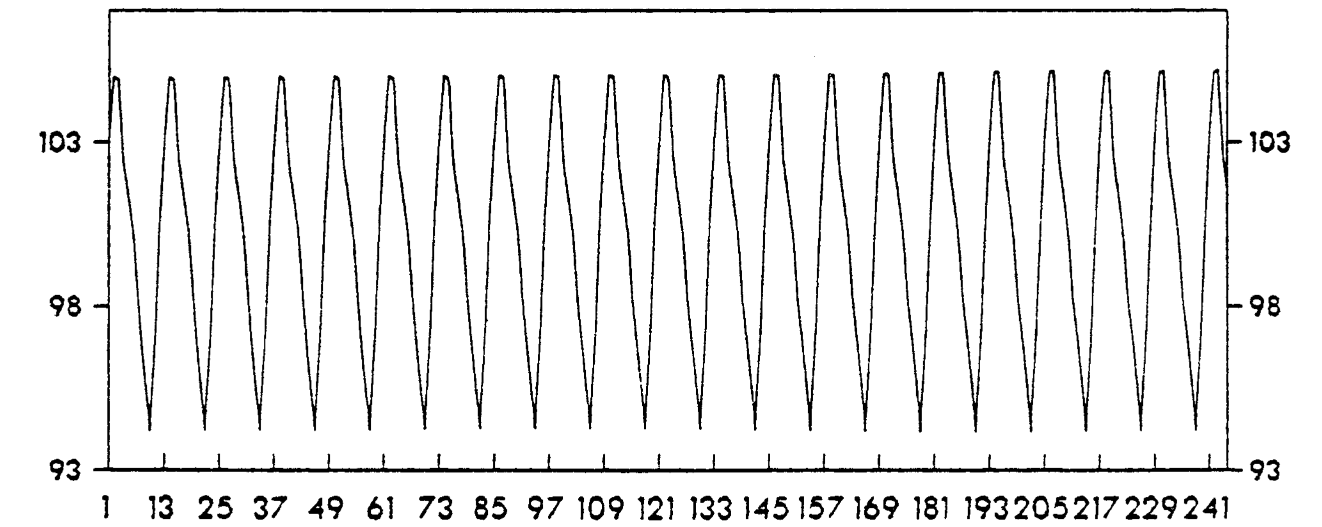
S35T1 X-11


2,000
1,500
1,000
500
1 13 25 37 49 61 73 85 97 109 121 133 145 157 169 181 193 205 217

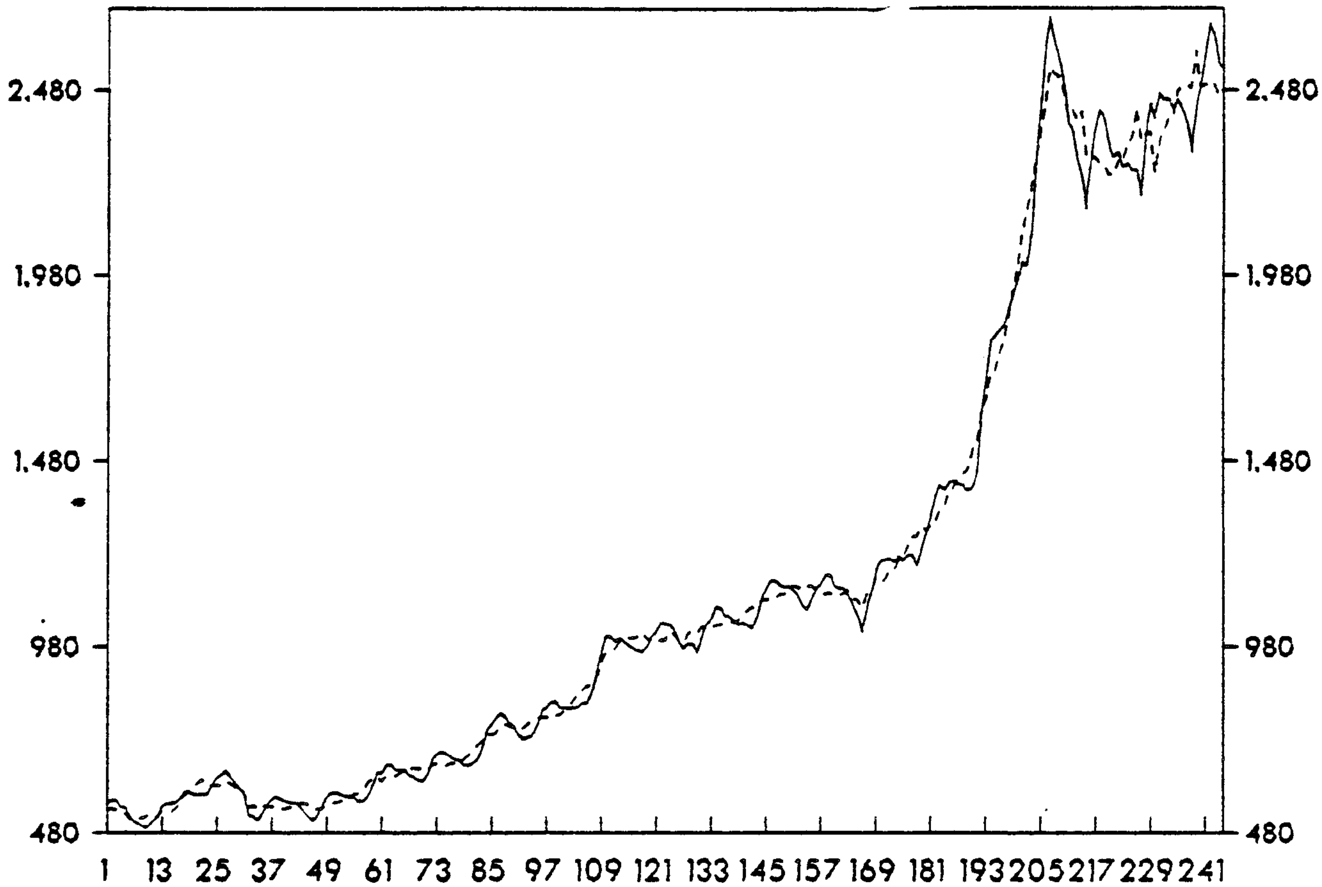

100
60
20
-20
-60
1 13 25 37 49 61 73 85 97 109 121 133 145 157 169 181 193 205 217

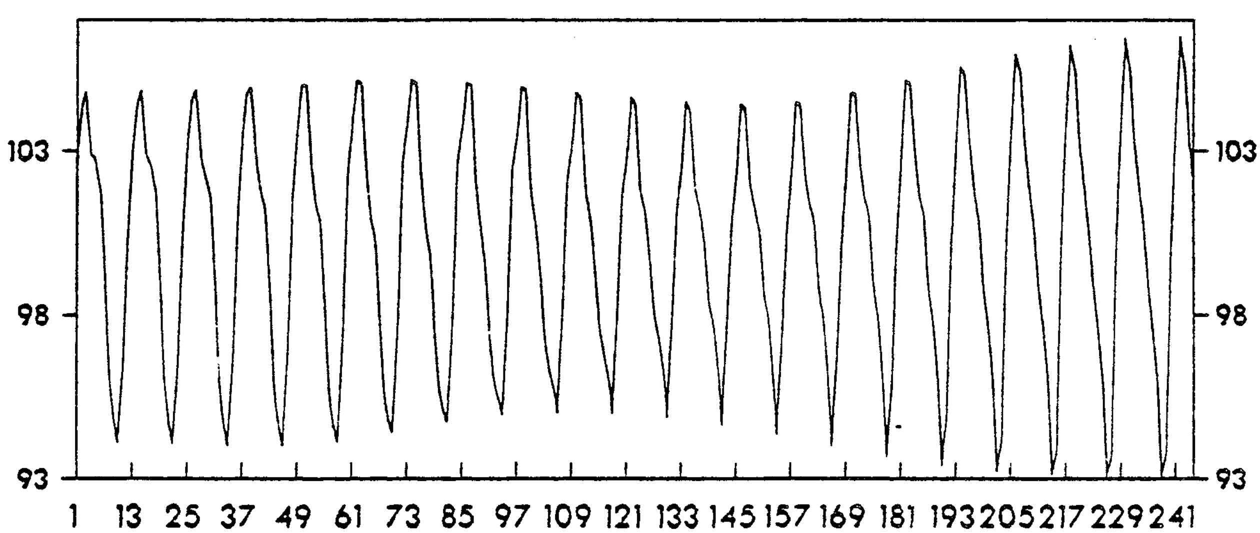
INDUSTRY S23TI MODEL BASED

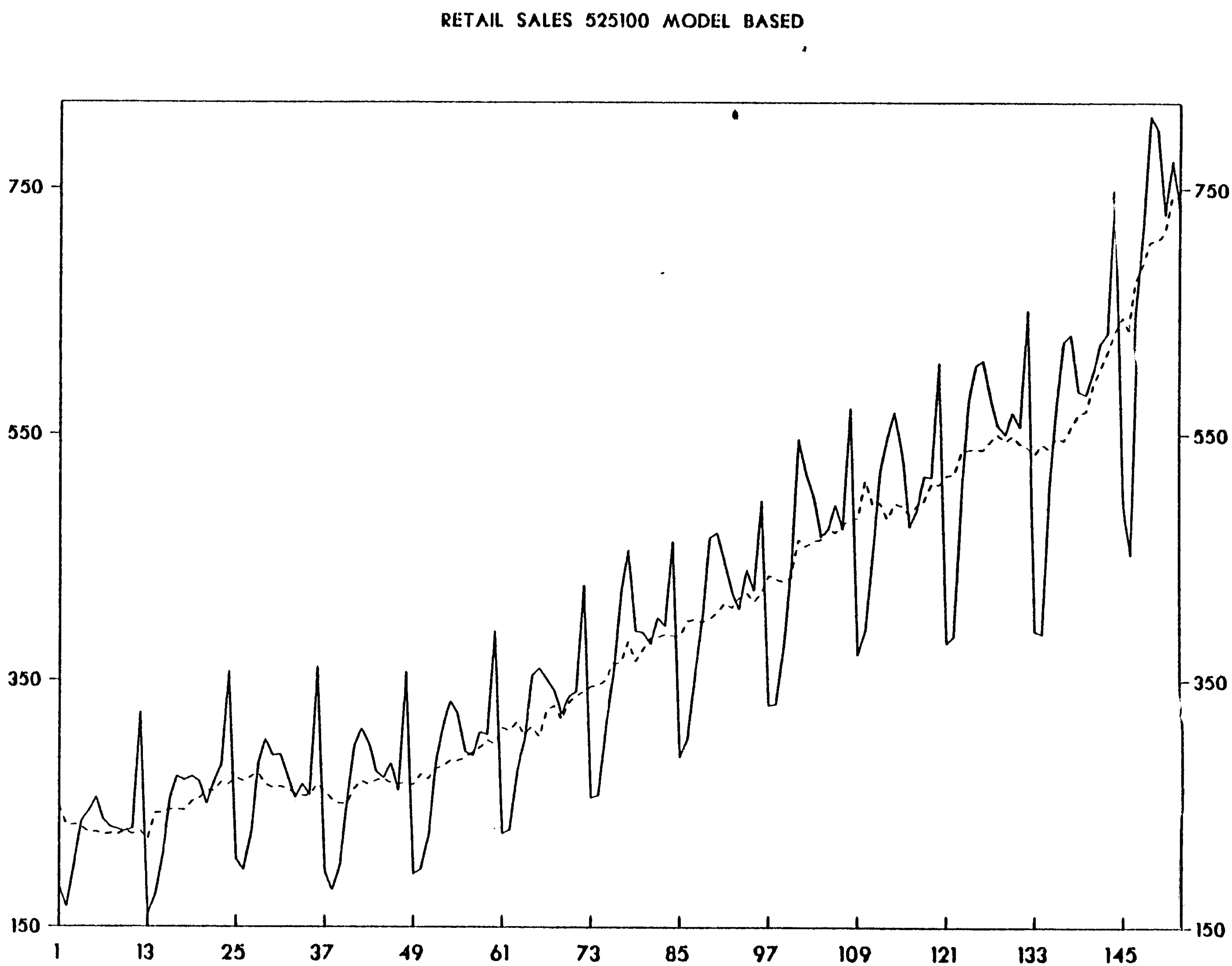




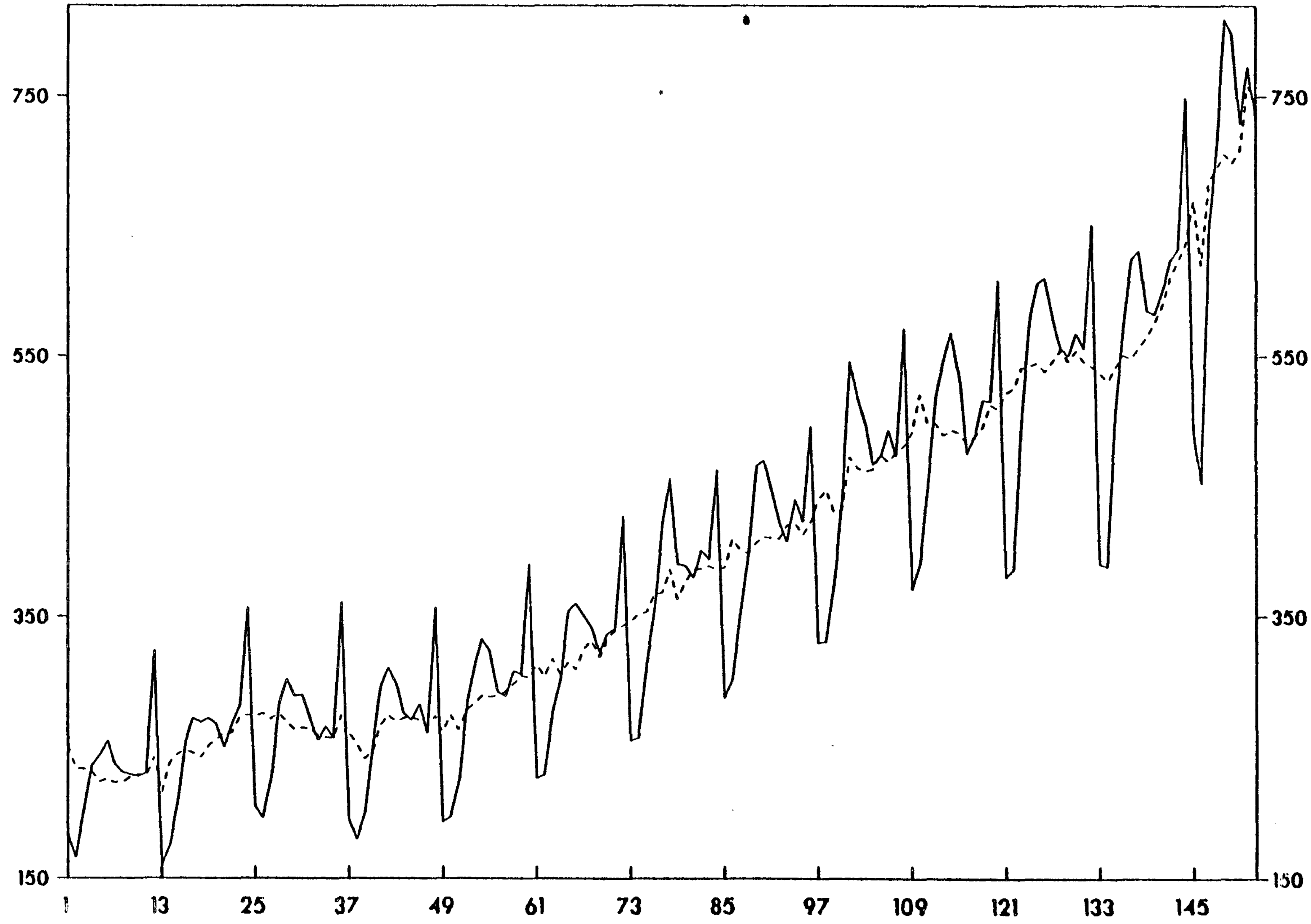
INDUSTRY 32[illegible] X-11 LOG



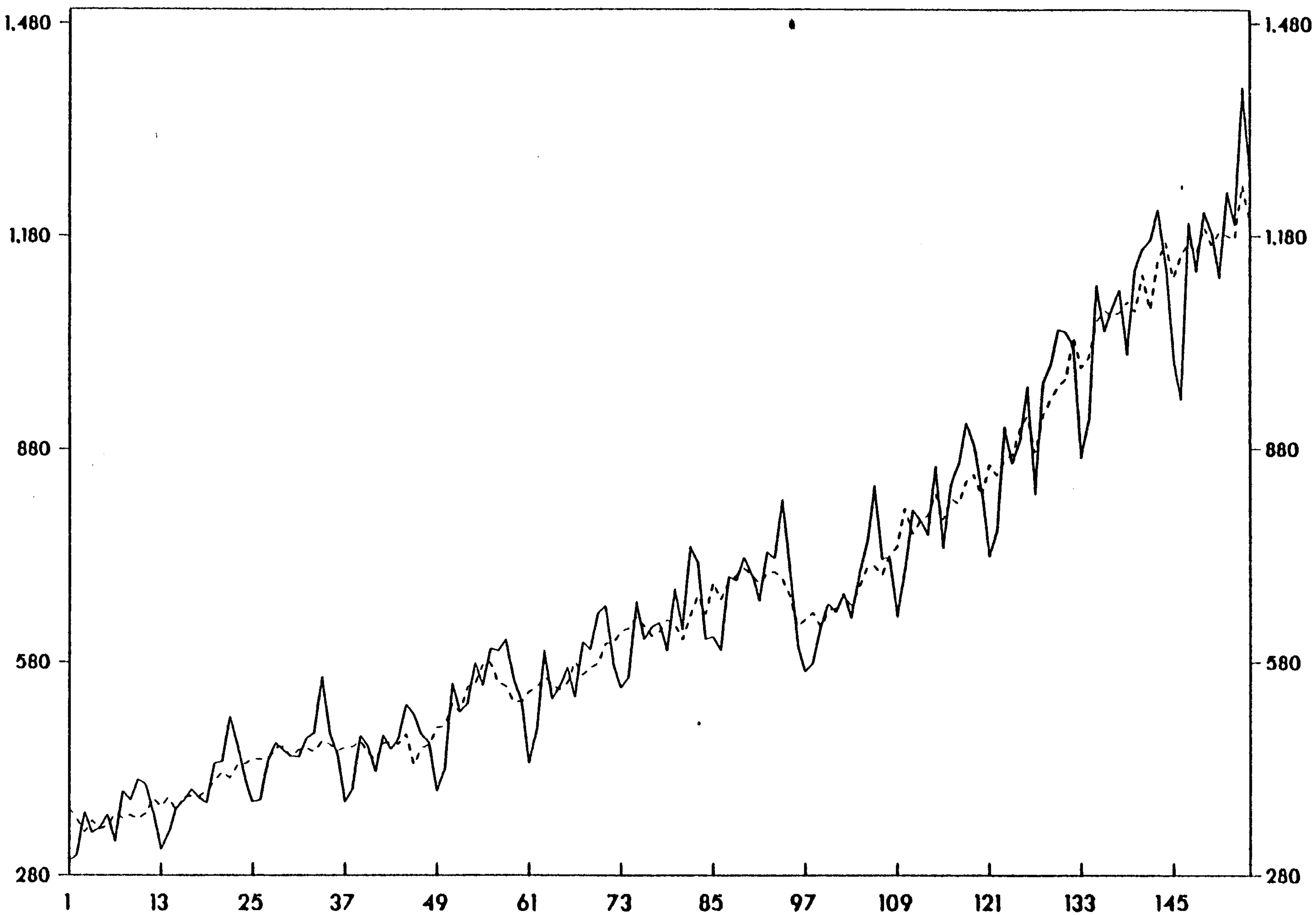


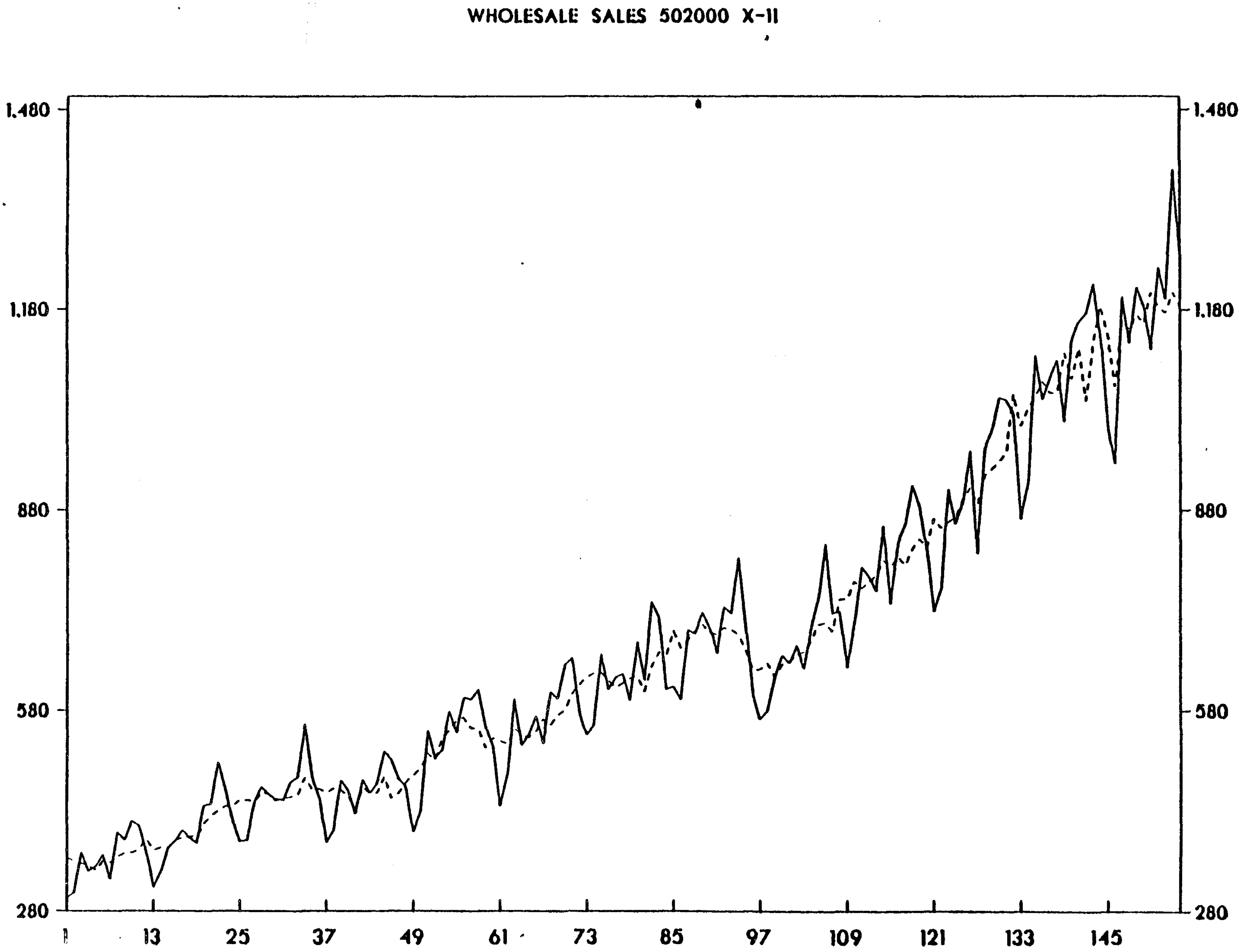

RETAIL SALES 525100 MODEL BASED
750
550
350
150
1
13
25
37
49
61
73
85
97
109
121
133
145

RETAIL SALES 525100 X-11



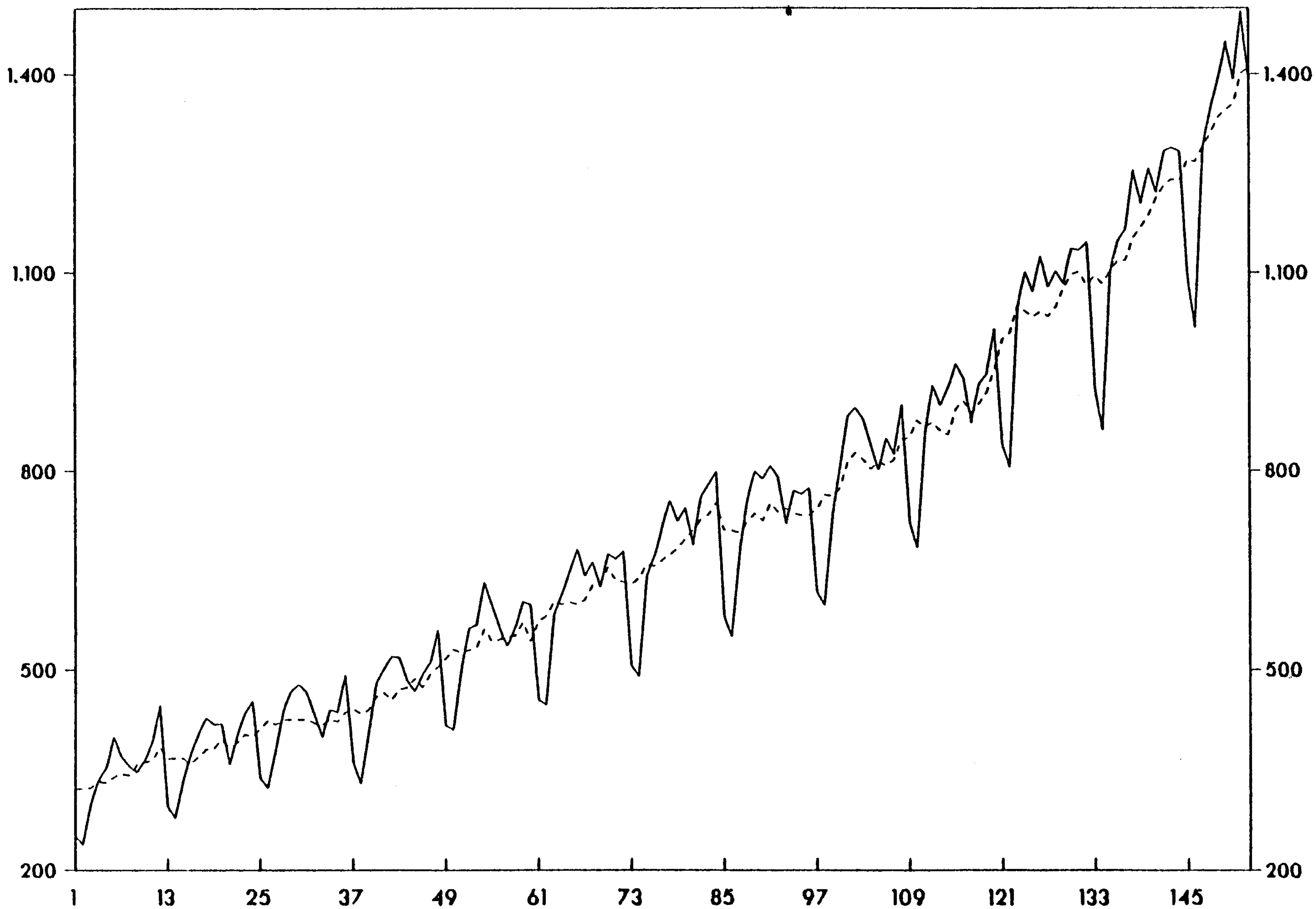
WHOLESALE SALES 502000 MODEL BASED





WHOLESALE SALES 502000 X-11
1,480
1,180
880
580
280
1
13
25
37
49
61
73
85
97
109
121
133
145

RETAIL SALES 553100 MODEL BASED



RETAIL SALES 553100 X-11