BUREAU OF THE CENSUS

STATISTICAL RESEARCH DIVISION REPORT SERIES

SRD Research Report Number: CENSUS/SRD/RR-82/03

THE PREDICTION OF TIME SERIES WITH TRENDS AND SEASONALITIES*

by

Will Gersch and Genshiro Kitagawa

Will Gersch
Department of Information and Computer Sciences
University of Hawaii
Honolulu, Hawaii 96822

Genshiro Kitagawa
The Institute of Statistical Mathematics
4-6-7 Minami-Azabu,
Minato Ku, Tokyo Japan-106

*This paper was written when both authors were American Statistical Association Fellows in Time Series at the U.S. Bureau of the Census, 1981-1982.

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: Myron J. Katzoff

Report completed: June 30, 1982

Report issued: November 23, 1982

SUBMITTED TO
ASA JOUR
BUS & FIN

THE PREDICTION OF TIME SERIES WITH TRENDS AND SEASONALITIES*

Will Gersch
Department of Information and Computer Sciences
University of Hawaii
Honolulu, Hawaii 96822

and

Genshiro Kitagawa
The Institute of Statistical Mathematics
4-6-7 Minami-Azabu,
Minato Ku, Tokyo Japan - 106

ABSTRACT: A maximization of the expected entropy of the predictive distribution interpretation of Akaike's minimum AIC procedure is exploited for the modeling and prediction of time series with trend and seasonal mean value functions and stationary covariances. The AIC criterion best one-step-ahead and best twelve-step-ahead prediction models are different. They exhibit the relative optimality properties for which they were designed. The results are related to open questions on optimal trend estimation and optimal seasonal adjustment of time series.

*This paper was written when both authors were American Statistical Association Fellows in Time Series at the U.S. Bureau of the Census, 1981-1982.

1. INTRODUCTION

In this paper we consider the optimal smoothing and forecasting of nonstationary time series with trend and seasonal mean value components with stationary covariance. Two classes of smoothness priors trend models are considered. In one the trend is modeled as a stochastically perturbed local polynomial function of time. In the other, the trend is assumed to consist of both the stochastically perturbed local polynomial component plus a "global" stationary time series component. A predictive likelihood interpretation of Akaike's AIC is exploited to determine the best of the models from the alternative trend model classes, for best one-step-ahead and best-twelve-step-ahead prediction criteria. The smoothness priors approach to time series modeling was developed earlier in papers by Akaike and by us. The innovative step in this paper is the maximization of the expected entropy of the predictive distribution interpretation of the minimum AIC procedure. The modeling and smoothing of time series is done using a Kalman predictor/smoother-Akaike AIC criterion methodology. The modeling is applied to econometric time series data that typically are seasonally adjusted by Census X-11 and by ARIMA type models. The treatment in the paper is largely phenomenological.

In detail, we consider two alternative decompositions of the observed time series data,

$$M_1: \ y(n) = t(n) + s(n) + \epsilon(n); \quad n = 1,\dots,N \tag{1a}$$

$$M_2: \ y(n) = t(n) + s(n) + v(n) + \epsilon(n); \quad n = 1,\dots,N \ . \tag{1b}$$

In (1a) and (1b), $t(n)$ is a local polynomial component, $s(n)$ is a seasonal component, $v(n)$ is a globally stationary autoregressive time series component and $\varepsilon(n)$ is an i.i.d. $N(0, \sigma^2)$ observation noise component of the observed time series $y(n)$, $n = 1,\dots,N$. The term $t(n) + s(n)$, $n = 1,\dots,N$ is thus the unknown mean value function of a nonstationary in the mean time series with the stationary covariance sequence $\epsilon(n)$ in model M_1 and $v(n) + \epsilon(n)$ in model M_2.

Following our earlier work, Brotherton and Gersch (1981), Kitagawa (1981), and Kitagawa and Gerscsh (1982), the trend t(n), the seasonal s(n) and the stationary time series v(n), are expressed in stochastically perturbed contending model order - dynamic state-space constraint models with unknown process noise variance. The Kalman filter facilitates computation of the likelihood for the unknown. (In a Bayesian framework, the process noise variances are hyperparameters.) Then, the Akaike's minimum AIC procedure is used to determine the best of the alternative trend models fitted to the observed data. A smoothing algorithm is subsequently applied to the AIC criterion best modeled data. The final results thus obtained are a "smoothness prior" or Bayesian smooth decomposition of the t(n), s(n) and possibly v(n) components of the observed time series y(n), n = 1,...,N. Particular time series that are of interest in the Census Bureau and Bureau of Labor Statistics for seasonal adjustment are analyzed. The one-step-ahead, increasing horizon and twelve-step-ahead forecasts are computed and shown for both the best fitted M_1 and M_2 models separately under optimal one-step-ahead and optimal twelve-step-ahead prediction error performance criteria.

Our frame of reference for the treatment of time series with trends and seasonalities is the smoothing problem as defined by Whittaker (1923), and the explicit smoothness priors solution to that problem by Akaike (1979). The Kalman filter/smoother is a Bayesian computationally efficient device that simplifies the task of inverting the generalized covariance for the computation of likelihoods of particular models. The richness of that technology beckons us to implement new models when the existing models do not appear to be satisfactory. When that event occurs, we compute with both the original and the newly invented models. Then we resort to the Akaike AIC criterion to determine the best of the alternative parametric models fitted to the data. In fact, consideration of the M_2 model class only evolved after extensive experience was obtained with the M_1 model class. Very simply, it appeared that in a considerable number of M_1 type models fitted to data, the local polynomial trend wiggled with a suggestive global stationary time series fluctuation. The M_2 model allows for both the local polynomial and globally stationary time series trend components.

A most comprehensive treatment of the smoothing problem approach to the modeling of time series with trend and seasonalities is in Kitagawa and Gersch (1982). Here we emphasize the prediction of such time series. The Analysis is in Section 2. State-space representations of the M_1 and M_2 models are described in Section 2.1. The minimum AIC procedure, including the maximization of the entropy of the predictive distribution interpretation of that procedure is in Section 2.2. The Kalman predictor and smoother are described in Sections 2.3 and 2.4 respectively. Examples of M_1 and M_2 modeled time series according to one-step-ahead and twelve-step-ahead prediction criteria are shown in Section 4. In the Summary and Discussion, Section 5, among other things we compare our methodology with the Box-Jenkins-Tiao methodology.

2. ANALYSIS

In this section the state space representation of the local polynomial trend or M_1 model and the local polynomial plus globally stationary time series component or M_2 model are shown. Seasonal and observation noise components are also included in both the M_1 and M_2 models. The minimum AIC procedure for determining the best of alternative models is discussed next. The maximization of the expected entropy of a predictive distribution interpretation of the minimum AIC proceduure is exploited to determine the AIC criterion best one-step-ahead and twelve-step-ahead predictor models. Following that, the Kalman predictor is discussed. The discussion includes formulas for the appropriate likelihoods and predictors. The Kalman smoother is applied to the AIC best prediction modeled data. The conventional smoother formulas are shown.

2.1 THE MODELS M_1 AND M_2

Consider the two alternative models M_1 and M_2 for the observed data $y(n)$, $n = 1,\ldots,N$,

$$M_1: \; y(n) = t(n) + s(n) + \epsilon(n), \quad n = 1,\ldots,N \tag{2.1a}$$

$$M_2: \; y(n) = t(n) + s(n) + v(n) + \epsilon(n), \; n = 1,\ldots,N \tag{2.1b}$$

In (2.1a,b) $t(n)$ is the local polynomial trend, $s(n)$ is the seasonal, $v(n)$ is the globally by stationary stochastic component of the observed time series $y(n)$, $n=1,\ldots,N$ and $\varepsilon(n)$, $n=1,\ldots,N$ is an i.i.d. "observation error" sequence assumed for convenience, to be $\mathcal{N}(0,\sigma^2)$, σ^2 unknown.

The trend and seasonal components are assumed to be represented by stochastically perturbed difference equation constraints

$$\nabla^k t(n) = w_1(n) \quad \text{for } k = 1,2,3 \tag{2.2a}$$

$$\sum_{i=o}^{L} s(n-i) = w_2(n) \quad . \tag{2.2b}$$

The stationary process $v(n)$ is assumed to be in the autoregressive (AR) model form

$$v(n) = \alpha_1 \, v(n-1) + \ldots + \alpha_p v(n-p) + w_3(n) \tag{2.2c}$$

In (2.2a) $\nabla t(n) = t(n) - t(n-1)$, $\nabla^2 t(n) = t(n) - 2t(n-1) + t(n-2)$ etc. Also, in (2.2a)-(2.2c) the "process noise" components $w_j(n)$, $j = 1,2,3$ and observation noise components $\varepsilon(n)$ are assumed to be zero-mean independent Gaussian distributed with

$$\begin{bmatrix} w_1(n) \\ w_2(n) \\ w_3(n) \\ \epsilon(n) \end{bmatrix} \sim N\left[\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_1^2 & 0 & 0 & 0 \\ 0 & \tau_2^2 & 0 & 0 \\ 0 & 0 & \tau_3^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}\right] \tag{2.3a}$$

In a compatible vector-matrix notation (2.3a) is

$$\begin{pmatrix} w(n) \\ \epsilon(n) \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix}\right) . \tag{2.3b}$$

The constraints in equations (2.2a,b,c) and the observation equations in (2.1a,b) are imbedded into the dynamical state space model for the observations,

$$\begin{aligned} x(n+1) &= F\,x(n) + G\,w(n) \\ y(n) &= Hx(n) + \epsilon(n) \end{aligned} \tag{2.4}$$

Under models M_1 and M_2 the matrices F,G, and H and an interpretation of the state vector (using an obvious subset notation) become

$$x(n) = \left[\begin{array}{c} t(n) \\ \vdots \\ t(n-k+1) \\ \hline s(n) \\ \vdots \\ s(n-L+2) \end{array}\right] ; \quad F_1 = \left[\begin{array}{cccc|cccc} c_1 & \cdots & & c_k & & & & \\ 1 & & & & & & 0 & \\ & \ddots & & & & & & \\ & & 1 & 0 & & & & \\ \hline & & & & -1 & \cdots & & -1 \\ & & & & 1 & & & \vdots \\ & & 0 & & \vdots & \ddots & & \vdots \\ & & & & 0 & \cdots & 1 & 0 \end{array}\right] ; \quad G_1 = \left[\begin{array}{cc} 1 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ \hline 0 & 1 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{array}\right] , \quad H_1' = \left[\begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \\ \hline 1 \\ 0 \\ \vdots \\ 0 \end{array}\right] \tag{2.5a}$$

$$x(n) = \begin{bmatrix} t(n) \\ \vdots \\ t(n-k+1) \\ \hline v(n) \\ \vdots \\ v(n-p+1) \\ \hline s(n) \\ \vdots \\ s(n-L+2) \end{bmatrix} \quad F_2 = \left[\begin{array}{c|c|c} \begin{matrix} c_1 \cdots & c_k \\ 1 & \\ & 1 \;\; 0 \end{matrix} & 0 & 0 \\ \hline 0 & \begin{matrix} \alpha_1 \cdots & \alpha_p \\ 1 & \\ & 1 \;\; 0 \end{matrix} & 0 \\ \hline 0 & 0 & \begin{matrix} -1 \; \cdots & -1 \\ 1 & \\ & 1 \;\; 0 \end{matrix} \end{array}\right] ; \; G_2 = \left[\begin{array}{c|c|c} \begin{matrix} 1 \\ \vdots \\ 0 \end{matrix} & 0 & \\ \hline 0 & \begin{matrix} 1 \\ \vdots \\ 0 \end{matrix} & 0 \\ \hline 0 & 0 & \begin{matrix} 1 \\ \vdots \\ 0 \end{matrix} \end{array}\right] ; \; H_2' = \begin{bmatrix} 1 \\ \vdots \\ 0 \\ \hline 1 \\ \vdots \\ 0 \\ \hline 1 \\ \vdots \\ 0 \end{bmatrix} . \quad (2.5b)$$

In (2.5a) the matrices F_1, G_1, H_1 are respectively $(k+L^{\perp} x k+L^{\perp})$, $(k+L^{\perp} x 2)$, $(1 x k+L^{\perp})$ where k is the order of the difference equation constraint in (2.2a) and $L^{\perp}$ = L-1 is the seasonal period duration minus one, with L = 4, L = 12 for quarterly and monthly data respectively. The vector $(c_1,\ldots,c_k)$ in (2.5a), (2.5b) reflects the trend constraint in (2.2a). It is respectively: (1) for k=1; (2,-1) for k=2; and (3,-3,1) for k = 3.

In model M_1 only the process noise parameters τ_1^2, τ_2^2 and the observation noise σ^2 are unknown. In model M_2 the process noise parameters τ_1^2, τ_2^2, τ_3^2 the AR parameters $\alpha_1,\ldots,\alpha_p$ and the observation noise variance σ^2 are unknown. From smoothing problem point of view, the parameters τ_1^2, τ_2^2, τ_3^2 are tradeoff parameters or hyperparameters. The ratio of parameters τ_j^2/σ^2 is a "signal to noise" measure. It expresses the relative uncertainty of the constraints (2.2a), (2.2b) and (2.2c) assumed for the model. Larger values of τ_j^2 imply stricter adherence to the j^{th} difference equation constraint.

2.2 THE MINIMUM AIC PROCEDURE

Akaike's minimum AIC procedure is interpreted here from a maximization of the entropy of the predictive distribution point of view. The treatment is phenomenological. The most relevant paper by Akaike on this topic is in Akaike (1980).

Let the true distribution be g and the fitted distribution be f, then the entropy of g with respect to f is

$$B(g,f) = \int g(y) \log \left\{ \frac{f(y)}{g(y)} \right\} dy \qquad (2.2.1)$$

$$= E_Y \log \left\{ \frac{f(Y)}{g(Y)} \right\}$$

$$= E_Y \log f(Y) - E_Y \log g(Y) .$$

The true distribution g is unknown. It is known that $B(g,f) \le 0$ and $B(g,f) = 0$ if and only if $f = g$ almost everywhere. The closer $B(g,f)$ is to zero, the closer we regard f as being to g . Hence the closeness of alternative f's to the unknown true g can be ordered if the quantity $E_Y \log f(Y)$ can be estimated from the observed data.

Let $y_1,\ldots,y_N$ be the observed data that occurs under the true distribution g and let $f(y|\theta)$ be an assumed distributional model of the data with parameter. Let

$$\ell(\theta) = \sum_{i=1}^{N} \log f(y_i|\theta) \qquad (2.2.2)$$

be the likelihood of the parameter given θ the data $y_1,\ldots,y_N$. Then, from the law of large numbers $\frac{1}{N}\ell(\theta)$ forms a natural estimate of $E_Y \log f(Y|\theta)$. Akaike's AIC criterion is a bias corrected estimate of $-2E_Y \log f(Y|\hat{\theta})$ for the practical situation in which the value of θ must be estimated from the data, Akaike (1973, 1974).

The AIC statistic, a bias corrected estimate of $E_Y \log f(Y|\hat{\theta})$ is

$$AIC(\hat{\theta}, y_1, \ldots, y_N) = -2\,(\hat{\theta}) + 2k \qquad (2.2.3)$$

where $\hat{\theta}$ is the MLE of θ, $\ell(\hat{\theta})$ is the maximized log likelihood and k is the number of parameters fitted to the model. The minimum AIC procedure preferred model is the one for which the value of the AIC statistic is smallest.

In the case of a one-step-ahead prediction criterion and Gaussian data, we consider the estimation of

$$E_Y \log f(y(n)|y(n-1),\ldots,y(1)) \qquad \text{by} \qquad (2.2.4)$$

$$\frac{1}{N}\sum_{n=1}^{N-1} \log f(y(n+1)|y(n),\ldots,y(1))$$

$$= -\frac{2}{2(N-1)}\sum_{n=1}^{N-1} \log 2\pi\, u^2(n+1|n) - \sum_{n=1}^{N-1} \frac{(y(n+1) - Hx(n+1|n))^2}{2u^2(n+1|n)}$$

with

$$u^2(n+1|n) = HV(n+1|n)H' + \sigma^2 \quad . \qquad (2.2.5)$$

In (2.2.4-5) the notation $x(n|n-1)$ and $V(n|n-1)$ are respectively the conditional mean and conditional covariance of the state vector $x(n)$ given the past data $y(n-1), y(n-2), \ldots, y(1)$ and $\bar{x}(0)$. Also in (2.2.5) $\nu(n+1) = (y(n+1) - Hx(n+1|n))$, $n=1,\ldots,N$ are the innovations. From the preceeding formula, they represent the difference between the observed data and the conditional mean of the data given the past. The innovations are a normally distributed zero mean independent process and are uniquely determinable from the observed process, $y(n), n = 1,\ldots, N$ (Anderson and Moore, 1979).

In the case of twelve-step-ahead prediction, consider the estimation of $E_Y \log f(y(n+12)|y(n),\ldots,y(1))$ by

$$\frac{1}{N-12} \sum_{n=1}^{N-12} \log f(y(n+12)|y(n),\ldots,y(1)) \quad (2.2.6)$$

$$- \frac{1}{2(N-12)} \sum_{n=1}^{N-12} \log 2 \quad u^2(n+12|n) - \sum_{n=1}^{N-12} \frac{(y(n+12) - Hx(n+12|n))^2}{2u^2(n+12|n)}$$

with

$$u^2(n+12|n) = HV(n+12|n)H^{\perp} + \sigma^2 \quad . \quad (2.2.7)$$

The last lines in (2.2.4) and (2.2.6), the approximations for the one-step-ahead and twelve-step-ahead maximized predictive likelihoods, are computed for particular values of τ_1^2, τ_2^2 in the M_1 model and particular values of τ_1^2, τ_2^2, τ_3^2 in the M_2 model. Formulas for the computation of the relevant condition state mean value and variance terms in the predictive likelihoods are in the section immediately following. (In the M_2 model the AR parameter estimates α_1, $\ldots,\alpha_p$ are estimated by a quasi Newton-Raphson type procedure for particular values of τ_1^2, τ_2^2 and τ_3^2.)

Under the AIC procedure, an exact maximum likelihood computation is assumed. The number of unknown parameters in the state space model is the dimension of the state, for the implicit or explicit estimation of x(0), plus the number of hyperparameters and the number of AR parameters fitted. Thus in the M_1 model, the number of parameters fitted is (k + (L-1) + 2), the order of the different equation constraint, the period of the seasonal duration minus one plus two for the hyperparameters τ_1^2, τ_2^2. Similarly, under the M_2 model the number of parameters fitted is (k + (L-1) + 2p + 3) where the dimensionality of the state is (k + (L-1) + p), p is the number of AR parameters fitted and there are 3 noise process terms or hyperparameters τ_1^2, τ_2^2 and τ_3^2.

2.3 THE KALMAN PREDICTOR FORMULAS, (Anderson and Moore, 1979).

Let the mean and covariance of the Gaussian density function of the state x(n+1) given the observations y(n),y(n-1),...,y(1) be denoted by x(n+1|n) and V(n+1|n). Starting with the initial conditions x(0|0) = x(0), V(0|0) = V_0 :

One-step-ahead prediction equations are computed recursively from:

$$\begin{aligned} x(n+1|n) &= F\, x(n|n) \\ V(n+1|n) &= F\, V(n|n)\, F^{\perp} + GQG^{\perp} . \end{aligned} \tag{2.3.1}$$

Time update equations are computed from:

$$\begin{aligned} x(n+1|n+1) &= x(n+1|n) + K(n+1)\, \nu(n+1) \\ K(n+1) &= V(n+1|n)\, H^{\perp}[HV(n+1|n)H^{\perp} + R]^{-1} \\ \nu(n+1) &= y(n+1) - H\, x(n+1|n) \\ u^2(n+1|n) &= H\, V(n+1|n)\, H^{\perp} + R \\ V(n+1|n+1) &= (I-K(n+1)\, H)V(n+1|n) . \end{aligned} \tag{2.3.2}$$

In (2.3.2) K(n+1) is the Kalman filter gain at time n+1, ν(n+1) is the innovations at time n+1 and u^2(n+1|n) is the conditional variance of y(n+1|n), the observation process, at time n+1 given the past data y(n),...,y(1).

The k-step-ahead predictions formulas for k = 1,2,... are

$$\begin{aligned} x(n+k|n) &= F^{k-1}\, x(n+1|n) \\ V(n+k|n) &= F^{k-1}\, V(n+1|n)F^{k-1^{\perp}} + \sum_{j=0}^{k-1} F^j GQG^{\perp} F^{\perp j} \\ y(n+k|n) &= Hx(n+k|n) \\ u^2(k+n|n) &= HV(n+k|n)H^{\perp} + R \end{aligned} \tag{2.3.3}$$

These quantities are used for the calculation of the likelihood.

For the examples worked in the next section, models are fitted to the y(1), ...,y(N) data and several types of predictors are computed for the future data y(N+1),...,y(N+M) from that model. For convenience the formulas for those predictors are:

One step-ahead-prediction (2.3.4a)

$$y(n+1|n) \qquad n = N,N+1,\ldots,N+M-1$$

Increasing horizon prediction (2.3.4b)

$$y(N+i|N) \qquad i = 1,\ldots,M$$

Twelve step-ahead prediction (2.3.4c)

First: $y(N+i|n) \qquad i = 1,\ldots,12$

Then: $y(N+12+j|N+j), \; j = 1,\ldots,(M-12)$

2.4 THE BACKWARD SMOOTHING ALGORITHM

The smooth of the state and of the observation at time n given all of the data y(1),...,y(N) are denoted respectively x(n|N) and y(n|N). The smoothed estimates are derived from the forward state estimates by the backward smoothing algorithm for n = N-1,...,1, (Anderson and Moore, 1979). It is the smoothed estimates of the trend, the seasonal, and when appropriate, the AR component of the series that are used in the final estimates.

$$\begin{aligned} x(n|N) &= x(n|n) + A(n)(x(n+1|N) - Fx(n|n) \\ V(n|N) &= V(n|n) + A(n)(V(n+1|N) - V(n+1|n))A(n)^{\perp} \\ y(n|N) &= Hx(n|N) \end{aligned} \qquad (2.3.5)$$

where

$$A(n) = V(n|n)\, F^{\perp} V(n+1|n)^{-1} \quad .$$

3. EXAMPLES

M_1 and M_2 type models, models with local polynomial trends and local polynomial plus globally stochastic stationary trend components were fitted to the BLSALLFOOD data, Jan. 66 - Dec. 79, N = 156.

The original data, and data decompositions including trend, seasonal, and AR components and the innovations are exhibited for each of the AIC best M_1 and M_2 models under both the best one-step-ahead prediction and twelve-step-ahead prediction criteria. The models are fitted to the observed $y(1),\ldots,y(N)$ data. One-step-ahead $(y(n+1|n),\ n=N,N+1,\ldots,N+M-1)$ increasing horizon $(y(N+j|N),\ j=1,\ldots,M)$ and twelve-step-ahead predictions $(y(N+j|N),\ j=1,\ldots,12,\ y(N+j+12|N+j),\ j=1,\ldots,M-12)$, are shown as are the true data $y(N+1),\ldots,y(N+M)$, and plus and minus one sigma confidence intervals. From a likelihood interpretation of the AIC, Akaike (1979), the AIC best of the best M_1 and M_2 model classes is that for which the AIC is minimum. The legends with each illustration are rather complete. The information in those legends is not repeated in this section. Instead, here we concentrate on interpretations of the illustrations.

EXAMPLE 1:

BLSALLFOOD data Figures 1A, $1B_1$, and $1C_1$ show the trend components on the BLSALLFOOD data computed by the Census X-11, default option and models M_1 and M_2 respectively. The data points are connected together in Figure $1B_1$ for easier interpretation. A comparison of Figures 1A and $1B_1$ illustrates that the M_1 modeled trend is very similar to that obtained by the Census X-11 program. The X-11 trend is computed by an ad-hoc two sided filtering method that was developed to achieve acceptable or pleasing results for the knowledgeable consumer. The local polynomial trend plus AR component computed by the M_2 model is in Figure $1C_2$. The trend computed by the M_2 model is much smoother than that computed by the M_1 model. The trend plus AR model of order 2 component of the M_2

model are very similar to the trend of the M_1 model. The seasonal components computed in the M_1 and M_2 models, Figures $1B_2$ and $1C_3$, are very similar. The appearance of the innovations of the M_1 model, Figure $1B_3$, does not suggest that a stationary time series component could be profitably removed from the M_1 model. Nevertheless, the The M_2 model is the AIC criterion best model.

The minimum AIC criterion model has an optimal mean square one-step-ahead prediction property. Therefore, the suggestion is that, for this data set, the decomposition of the trend into a local polynomial plus a globally stationary AR part will yield superior one-stepahead prediction performance than that obtainable by modeling the trend by a local polynomial model.

In Figures 2A, 2B the statistical performance of the one-step-ahead predictions of the M_1 and M_2 models have similar appearances. A careful examination of Figures 2A, 2B suggests that the M_2 model prediction performance is slightly superior to that achieved by the M_1 model predictions. This comparative one-step-ahead prediction performance is consistent with a likelihood interpretation of the AIC, Akaike (1979). The increasing horizon performance predictions achieved by the M_1 model, Figure 2C, show the increasing divergence between true and predicted data and the increasing with horizon forbiddingly large plus and minus one sigma confidence intervals. The increasing horizon predictions achieved by the M_2 model, Figure 2D, appears to be quite satisfactory. It is important to note that the one-step-ahead best prediction performance does not have any necessary implications about increasing horizon prediction performance.

EXAMPLE 2. BLSALLFOOD date re-examined, 12-month horizon best M_1 and M_2 models.

Figures 3 and 4 respectively illustrate the original data, the additive component decomposition of the BLSALLFOOD data for the 12-month ahead prediction criterion best M_1 and M_2 models and the one-step-ahead, increasing horizon and twelve-months-ahead prediction performance. For this performance criterion, the

12-month ahead prediction criterion best local polynomial trend model M_1, is slightly superior to the M_2 model. The 12-month forecast criterion M_1 trend, Fig. 3A, is somewhat similar to the one-month forecast criterion M_2 trend, Fig. 1C. The seasonal components of the 12-month forecast M_1 and M_2 models are very similar to each other. The superiority of the M_1 model over the M_2 model for 12 months prediction performance is supported by comparisons of Figs. $3B_3$ and Fig. $4B_3$ for the M_1 and M_2 models respectively. Also, the increasing horizon prediction performance of M_1 is superior to that achieved by M_2. As before, the one-step-ahead prediction performance of the M_2 model is superior to that achieved by the M_1 model, Figures $4B_1$ and $3B_1$ respectively. Figure $3A_4$ shows the the 12-month-ahead prediction performance achieved by the one-step-ahead best M_1 model. The excessively large one sigma confidence intervals make this model unsatisfactory for 12 month prediction for the given data.

4. SUMMARY AND DISCUSSION

A maximization of the expected entropy of the predictive distribution interpretation of Akaike's minimum AIC procedure was exhibited and exploited here in the modeling and prediction of time series with trends and seasonalities. The AIC criterion best one-step-ahead and best twelve-step-ahead prediction models are different and individually, they exhibit the relative optimality properties for which they were designed. These results relate to the trend estimation and seasonal adjustment procedures in the Census X-11, Shiskin et al (1978), 1967), and the ARIMA based seasonal adjustment methods, Cleveland and Tiao (1970), Hillmer, Bell and Tiao (1981), and Hillmer and Tiao (1982).

An emphasis in the employment of the Census X-11 is in achieving an appraisal of the current status or current trend of an econometric time series. The X-11 procedures are subject to certain practical public data reporting constraints which influence the determination of that trend, Shiskin and Plenes

(1967). We do note that the X-11 seasonal adjustment procedures are implicitly prediction motivated procedures in that seasonalities one year in advance are computed to facilitate deseasonalization of current data.

The ARIMA model is an innovations type model. Thus, it has optimal one-step-ahead prediction properties under the class of signal and noise model constraints with which it is designed. Other horizon prediction performance ARIMA models are not known to have been sought for with the Box-Jenkins-Tiao modeling procedure.

The exhibited statistical performance of the AIC maximized predictive distribution performance procedure suggests new inquiries as to what is really the problem in the seasonal adjustment of time series. Our evidence suggests that rather reliable one-step-ahead and twelve-step-ahead predictions can be obtained by our methodology. The models, and hence the estimate of trend, differ according to whether the desired optimal prediction performance is one-step-ahead or twelve-steps-ahead. Trend estimation and seasonal adjustments might well be considered as procedures whose results depend upon the purpose for which the data is modeled.

5. REFERENCES

Akaike, H., (1973). Information theory and an extension of the maximum likelihood principle, in 2nd international symposium on information theory, B.N. Petrov and F. Csaki, eds., (Akademiai Kiado, Budapest), 267-281.

Akaike, H., (1974). A new look at the statistical model identification, IEEE Transactions on Automatic Control, AC-19, 716-723.

Akaike, H., (1978). On the likelihood of a time series model, The Statistician, 27, 217-235.

Akaike, H., (1980). Likelihood and the Bayes procedure, in Bayesian Statistics, J.M. Bernado, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds., University Press, Valencia, Spain, 141-166.

Akaike, H., (1981). Likelihood of a model and information criteria. Journal of Econometrics, 16, to appear.

Anderson, B.D.O. and Moore, J.B., (1979). Optimal Filtering, Prentice Hall, New Jersey.

Brotherton, T. and Gersch, W., (1981). A data analytic approach to the smoothing problem and some of its variations, in Proceedings of the 20th IEEE Conference on Decision and Control, 1061-1069.

Cleveland, W.P. and Tiao, G.C., (1976). Decomposition of seasonal time series: A model for the Census X-11 program, Journal of the American Statistical Assoc. 71, 581-587.

Hillmer, S.C., Bell, W.R., and Tiao, G.C., (1981). Modeling considerations in the seasonal adjustment of economic time series, in Applied time Series Analysis of Economic Data, ASA-CENSUS-NBER Conference Proceedings.

Hillmer, S.C. and Tiao, G.C., (1982). An ARIMA based approach to seasonal adjustment, Journal of the American Statistical Association, 77, 63-70.

Kitagawa, G., (1981). A nonstationary time series model and its fitting by a recursive technique, Journal of Time Series Analysis, 2, 103-116.

Kitagawa, G. and Gersch, W., (1982). A smoothing problem approach to the modeling of time series with trends and seasonalities, preliminary draft Technical paper, Bureau of the Census, U.S. Department of Commerce.

Shiskin, J. and Plewes, T.J., (1978). Seasonal Adjustment of the U.S. employment rate, The Statistician, 27, 181-202.

Shiskin, J., Young, A.H., and Musgrave, J.C., (1967). The X-11 variant of Census method II seasonal adjustment program Technical Paper, 15, Bureau of the Census, U.S. Department of Commerce.

Whittaker, E.T., (1923). On a new method of graduation, Proceedings Edinborough Math. Society 41, 63-75.

LEGENDS

Figure 1: BLSALLFOOD data and trends, Jan. 66 - Dec. 79, N = 156.

A. Census X-11, default option

B. Model M_1: B1 Original data plus trend, AIC = 1342.49, B2 Seasonal component, B3 Innovations

C. Model M_2: C1 Original data plus trend, AIC = 1309.82, C2 Original data and trend plus AR component, C3 Seasonal component, C4 AR component, C5 Innovations

Figure 2: BLSALLFOOD data predictions, the actual data, and plus and minus one sigma confidence intervals

A. M_1: One step ahead predictions

B. M_2: One step ahead predictions

C. M_1: Increasing horizon predictions

D. M_2: Increasing horizon predictions

Figure 3: BLSALLFOOD data, Model M_1: 12-months-ahead predictions criterion, AIC = -725.46.

Original data, component decomposition and predictions, true values and plus and minus one sigma confidence intervals.

A_1: Model M_1: Original plus trend, A_2 seasonal component,

B_1 One step-ahead prediction, B_2 increasing horizon prediction,

B_3 12-step-ahead prediction

C Model M_1: One step ahead criterion model M_1 twelve months ahead predictions.

Figure 4: BLSALLFOOD data, Model M_2: 12-months-ahead prediction criterion, AIC = -715.45.

Original data and component decompositions, and predictions, true values and plus and minus one sigma confidence intervals.

A_1 M_1: Original data and trend, A_2 M_2: B_1 Original data and trend plus AR components

B_3 AR component, B_3 Seasonal component,.

C_1 M_1: One-step-ahead predictions, C_2 M_5: Increasing horizon predictions,

C_3 M_2: 12-months-ahead predictions.

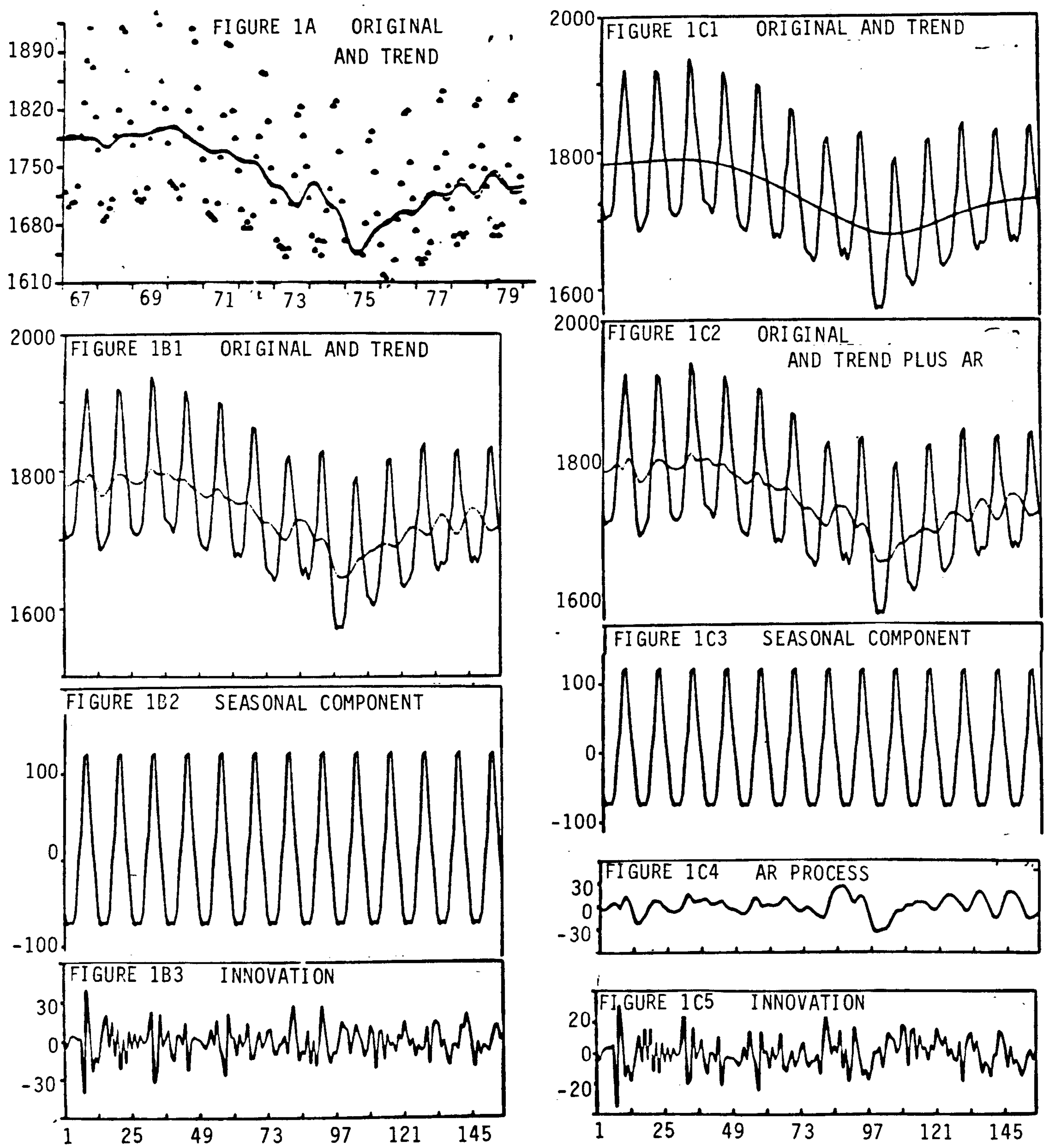
FIGURE 1A ORIGINAL AND TREND
1890
1820
1750
1680
1610
67
69
71
73
75
77
79
FIGURE 1B1 ORIGINAL AND TREND
2000
1800
1600
FIGURE 1B2 SEASONAL COMPONENT
100
0
-100
FIGURE 1B3 INNOVATION
30
0
-30
1
25
49
73
97
121
145
FIGURE 1C1 ORIGINAL AND TREND
2000
1800
1600
FIGURE 1C2 ORIGINAL AND TREND PLUS AR
2000
1800
1600
FIGURE 1C3 SEASONAL COMPONENT
100
0
-100
FIGURE 1C4 AR PROCESS
30
0
-30
FIGURE 1C5 INNOVATION
20
0
-20
1
25
49
73
97
121
145

FIGURE 2

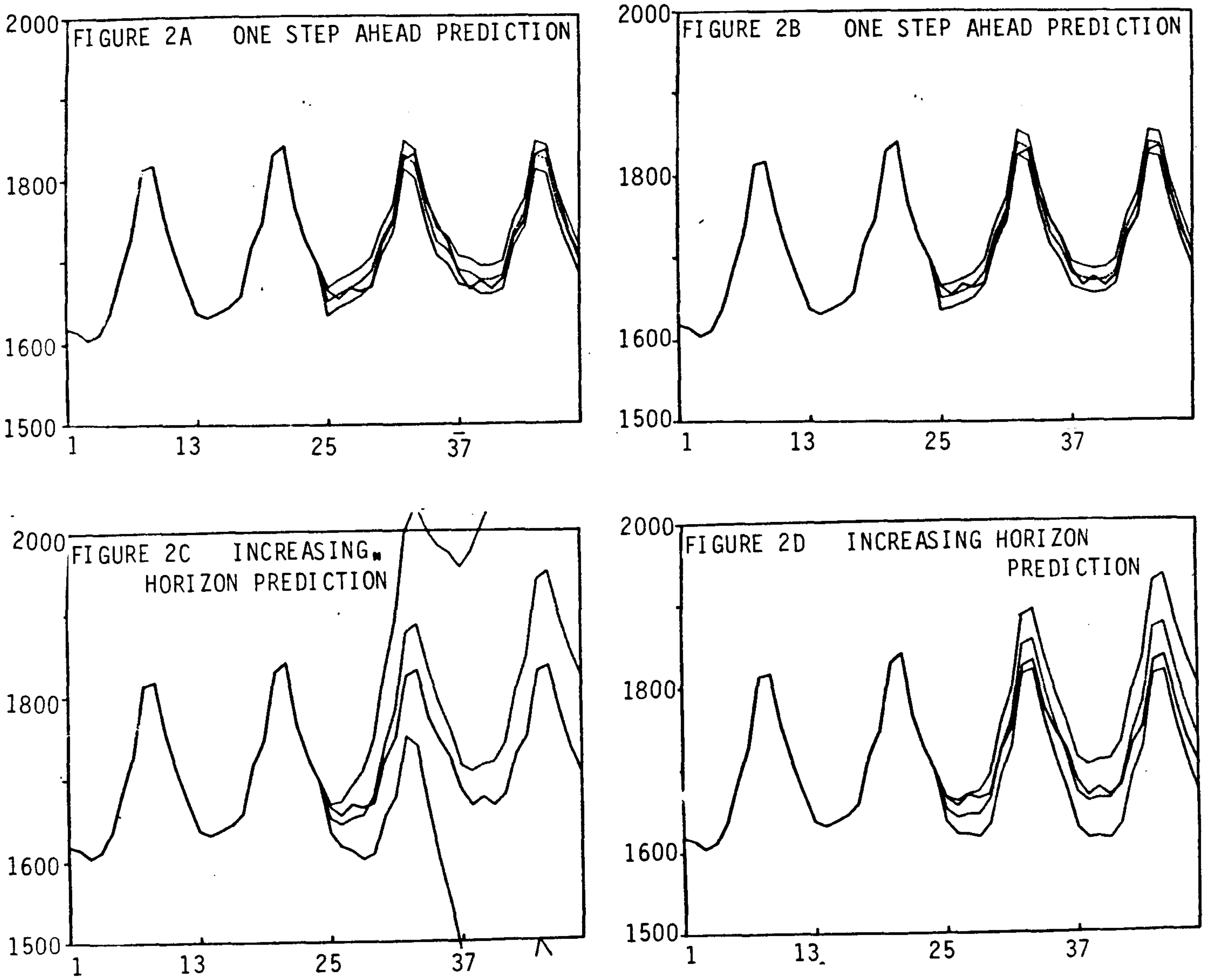
FIGURE 2A ONE STEP AHEAD PREDICTION
FIGURE 2B ONE STEP AHEAD PREDICTION
FIGURE 2C INCREASING HORIZON PREDICTION
FIGURE 2D INCREASING HORIZON PREDICTION
2000
1800
1600
1500
1
13
25
37

FIGURE 3

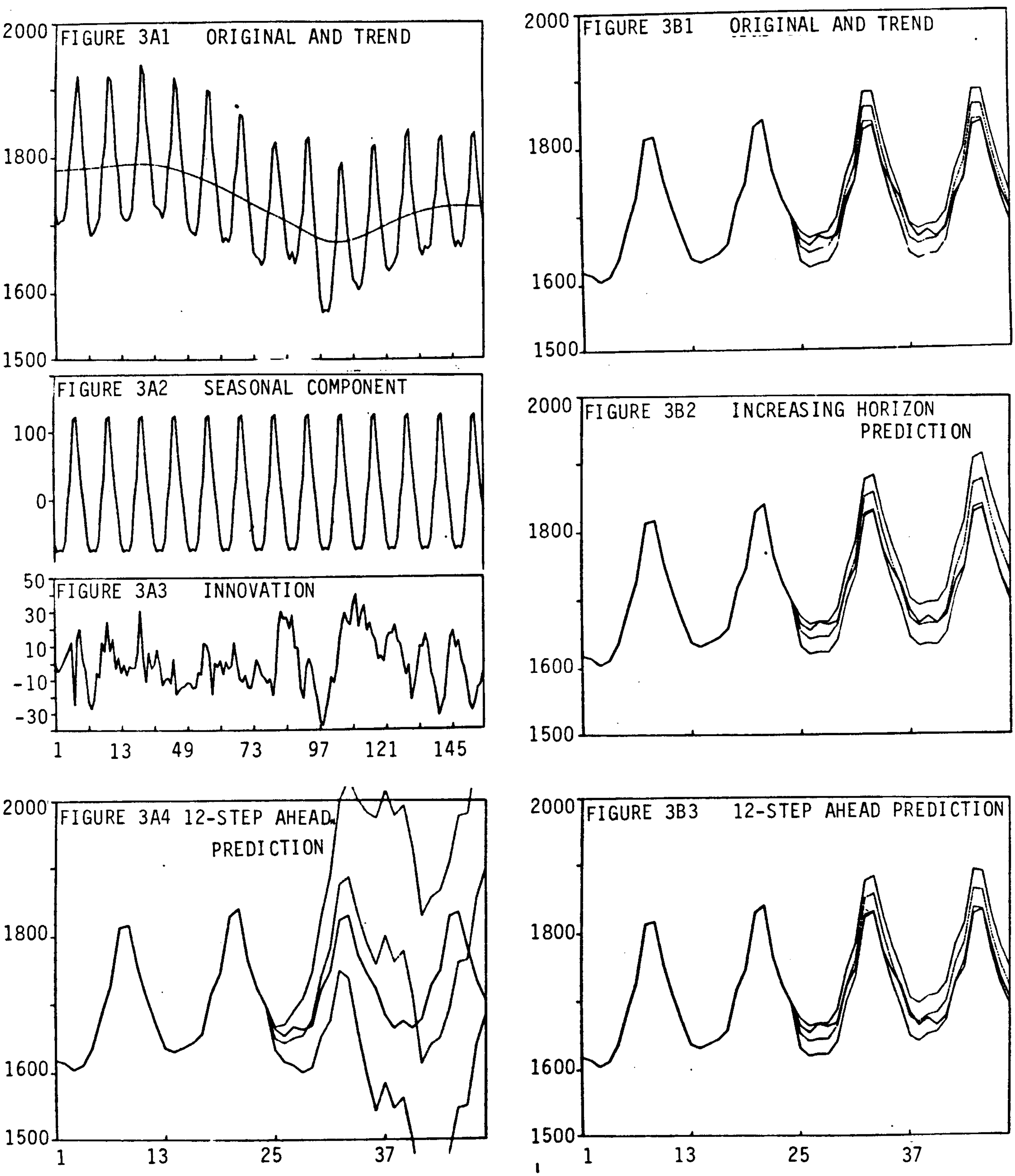
FIGURE 3A1 ORIGINAL AND TREND
2000
1800
1600
1500
FIGURE 3A2 SEASONAL COMPONENT
100
0
FIGURE 3A3 INNOVATION
50
30
10
-10
-30
1
13
49
73
97
121
145
FIGURE 3A4 12-STEP AHEAD PREDICTION
2000
1800
1600
1500
1
13
25
37
FIGURE 3B1 ORIGINAL AND TREND
2000
1800
1600
1500
FIGURE 3B2 INCREASING HORIZON PREDICTION
2000
1800
1600
1500
FIGURE 3B3 12-STEP AHEAD PREDICTION
2000
1800
1600
1500
1
13
25
37

FIGURE 4

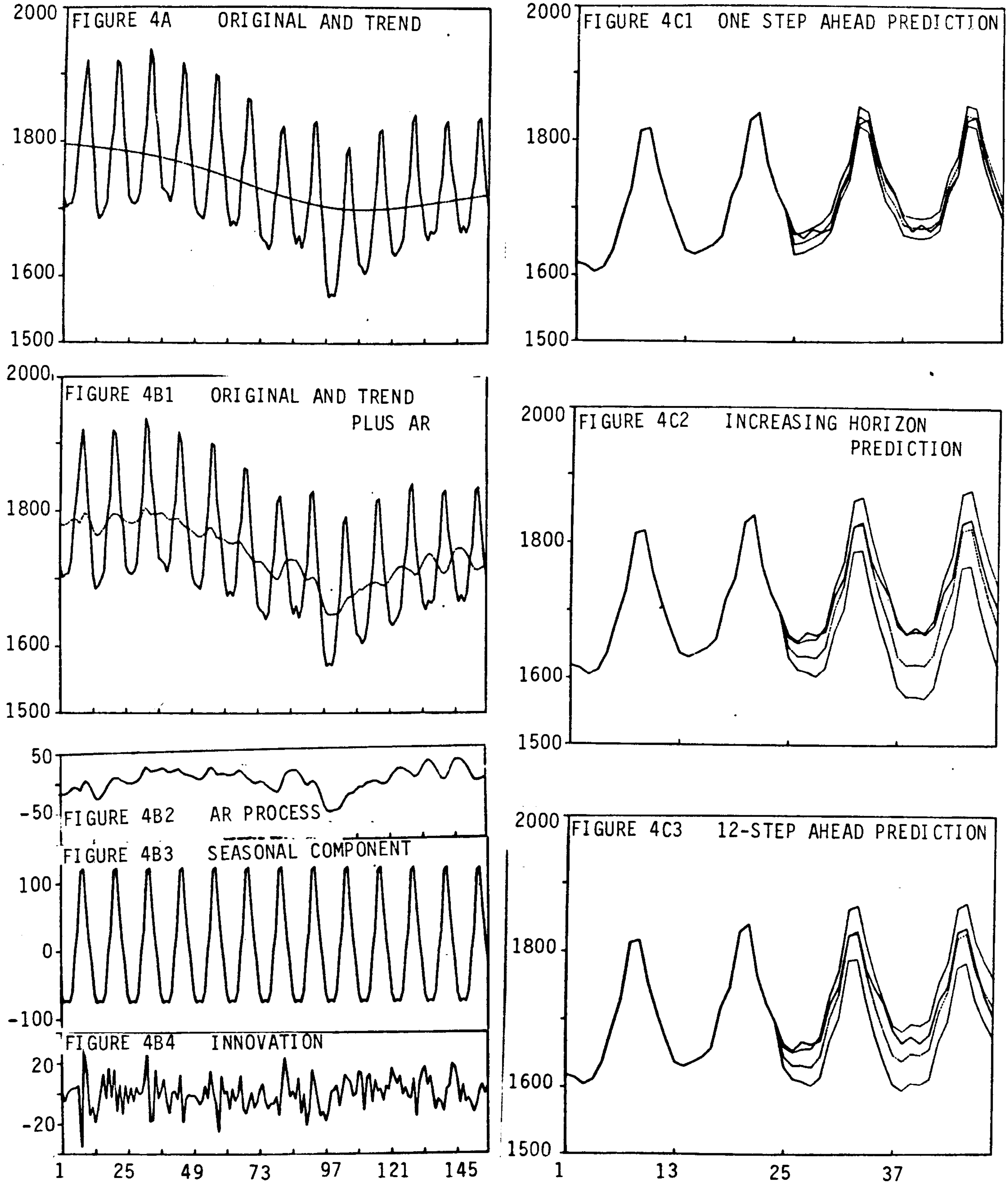

FIGURE 4A ORIGINAL AND TREND
2000
1800
1600
1500
FIGURE 4B1 ORIGINAL AND TREND PLUS AR
2000
1800
1600
1500
50
-50
FIGURE 4B2 AR PROCESS
FIGURE 4B3 SEASONAL COMPONENT
100
0
-100
FIGURE 4B4 INNOVATION
20
-20
1
25
49
73
97
121
145
FIGURE 4C1 ONE STEP AHEAD PREDICTION
2000
1800
1600
1500
FIGURE 4C2 INCREASING HORIZON PREDICTION
2000
1800
1600
1500
FIGURE 4C3 12-STEP AHEAD PREDICTION
2000
1800
1600
1500
1
13
25
37