Article Information

Journal Title: Conference papers from the winter meeting, New York, N.Y., January 26-31, 1975 /

Volume: **Issue:**
Month/Year: 1975 **Pages:** No. C75 027-8

Article Author: IEEE Power Engineering Society.

Article Title: Keyhani, A., and A. El-Abiad,; One-Step-Ahead Load Forecasting for On-Line Applications

Loan Information

Loan Title:

Loan Author:

Publisher:
Place:
Date:
Imprint:

Borrower Information

Library - ILL
Ohio State University
1858 Neil Avenue Mall
Columbus, OH 43210-1286

Odyssey: 206.107.43.112
Ariel: ariel.lib.ohio-state.edu
Email: osuill@osu.edu

NOTES

/ LANTER

C 75 027-8

ONE-STEP-AHEAD LOAD FORECASTING FOR ON-LINE APPLICATIONS

Ali Keyhani Ahmed El-Abiad

Purdue University
West Lafayette, Indiana

ABSTRACT

The problem of one-step-ahead forecasting of the load demand of an interconnected power system for on-line application is studied. It is proposed that at any time instant K, a stochastic model with unknown order, ARMA (n,m), be postulated. By using a recursive identification algorithm, the order of the model is identified on-line from the immediate past observations of load data. Having determined the number of parameters to be used (i.e., the order, n and m) at any instant K, these parameters are to be estimated in real time using the past few observations. This is done by applying a recursive estimation algorithm [1].

INTRODUCTION

In this paper the basic concepts underlying the very general stochastic models based on time series and identification of such models are studied. Identification and adaptive estimation techniques are used to choose a model from the general class of autoregressive, moving average, and mixed-autoregressive-moving-average models. The model then has been used for one-step-ahead load forecasting; and under the assumption that the resulting errors are normal, the corresponding probability limits are calculated.

As far as possible, the mathematical proofs are omitted; but the general concepts of identification and estimation techniques which have been employed are presented. Data provided by Public Service Indiana for one-minute intervals, five-minute intervals and hourly load are used for one-step-ahead forecasting to demonstrate the application of the proposed method.

Finally, it has been the task of this investigation to develop a computer software package which is recursive and completely automatic in identifying a non-seasonal time series and estimating the parameters of the identified model for real time application. Also, it should be pointed out that the method is applicable for any one-step-ahead load prediction such as 24 hours, one-week, and one-month-ahead for off-line application.

NOMENCLATURE

K - Instant of time (sec., min., hour, ...)

$Y(K)$ - The load demand as measured at time instant K, $K = 0,1,2, \ldots$

$\Phi_j(K)$ - An arbitrary function chosen by user, $j = 1, \ldots \ell$

$a^T(K) = (a_1(K), \ldots a_{n+m+\ell}(K))$ - The vector of unknown parameters

$\hat{a}^T(K) = (\hat{a}_1(K), \ldots \hat{a}_{n+m+\ell}(K))$ - The estimate of $a^T(K)$ based on the observations $\{y(j), j \leq K\}$

$\overline{W}(K)$ - The residual at instant K

$\hat{Y}(K)$ - The one-step-ahead predictor of $Y(K)$ based on observations $\{Y(K-j), j \geq 1\}$

$e(K)$ - The prediction error at instant K
$= Y(K) - \hat{Y}(K)$

$E[\cdot]$ - The operation of taking the expected value of $[\cdot]$

$C_{YW}(\cdot)$ - Crosscovariance function between Y and W

$C(\cdot)$ - Autocovariance function of a random variable

$\rho(\cdot)$ - Autocorrelation of a process

$MA(m)$ - The moving average process of order m

$AR(n)$ - The autoregressive process of order n

$ARMA(n,m)$ - The autoregressive moving average process of order n and m

$E_{0,N}$ - Mean value of the error for N observations

$E_{1,N}$ - Absolute mean value of the error for N observations

$E_{2,N}$ - Mean square value of the error for N observations

$Pr[\cdot]$ - The probability of $[\cdot]$

σ - Standard deviation

P.U.E. - Per Unit Error

$$P.U.E.(K) = \frac{Y(K) - \hat{Y}(K)}{Y(K)}$$

LOAD FORECASTING PROBLEM BY MEANS OF LINEAR STOCHASTIC DIFFERENCE EQUATION

In the present study, we are interested in predicting the one-step-ahead power demand of the entire system as it is measured at the central dispatching office. For any given power system, the power demand is available at discrete intervals of time. Let $Y(.)$ represent the power demand, as measured at the dispatching office. The problem of one-step-ahead load forecasting, given a set of past observations $Y(1), Y(2), \ldots Y(K-1), Y(K)$ (where K is the instant of time), is to find the best estimate of the load at K+1; that is, to predict $\hat{Y}(K+1)$. Naturally, if we are interested in

CONFERENCE PAPER

C 75 027-8. A paper recommended by the IEEE Power System Engineering Committee of the IEEE Power Engineering Society for presentation at the IEEE PES Winter Meeting, New York, N.Y., January 26-31, 1975. Manuscript submitted August 28, 1974; made available for printing November 4, 1974.

Price: Members $1.50
Nonmembers $2.00
At Meeting: $1.00

ℓ-step-ahead prediction, we must first determine one-step-ahead prediction; that is, $\hat{Y}(K+1)$, then two-step-ahead prediction, $\hat{Y}(K+2)$, ... $\hat{Y}(K+\ell-1)$, and finally $\hat{Y}(K+\ell)$ which is ℓ-step-ahead prediction. In this study, we are only interested in one-step-ahead prediction for real time application.

Let $\{Y(.)\}$ represent the observations of the power demand, that is

$$Y(1),\ Y(2),\ \ldots\ Y(K-n),\ \ldots Y(K-2),\ Y(K-1),\ Y(K)$$

Let the given sequence $\{Y(.)\}$ obey the following stochastic difference equation as expressed by equation (1):

$$Y(K) = a_1(K)Y(K-1) + a_2(K)Y(K-2) + \ldots + a_n(K)Y(K-n) + W(K) \quad (1)$$

in which the disturbance $W(k)$ cannot be directly observed. Alternatively, we can rewrite (1) in compact form as given by (2):

$$Y(K) = \sum_{j=1}^{n} a_j(K)Y(K-j) + W(K) \quad (2)$$

A process which can be expressed as equation (2) is called an autoregressive process of order n or, in short, AR(n).

It can be shown that one-step-ahead prediction of the stochastic process $\{Y(.)\}$, that is, $\hat{Y}(K+1)$, given the past observations $Y(1)$, $Y(2)$, ..., $Y(K-1)$, $Y(K)$, is given by equation (3):

$$\hat{Y}(K+1) = \hat{a}_1(K)Y(K) + \hat{a}_2(K)Y(K-1) + \hat{a}_3(K)Y(K-2) + \ldots \quad (3)$$

This can be written in compact form as

$$\hat{Y}(K+1) = \sum_{j=1}^{n} \hat{a}_j(K)Y(K-j+1)$$

where we have assumed that $E[W(K)] = 0$ and "$\hat{a}(K)$" is the estimate of "$a(K)$" based on the observations $[Y(j),\ j \leq K]$.

Once the observation $Y(K+1)$ is available, the residue at instant K+1 can be calculated by equation (4):

$$W(K+1) = Y(K+1) - \hat{Y}(K+1) \quad (4)$$

which is also the error at K+1 instant.

Intuitively, it is clear that we should be able to use the past errors to improve our prediction of the future. Let us assume that the observations $\{Y(.)\}$ and its past residues under certain estimation techniques are available to us. Then a better choice of a model for our process may be written as (5):

$$Y(K) = \sum_{j=1}^{n} a_j(K)Y(K-j) + \sum_{j=1}^{m} a_{n+j}(K)W(K-j) + W(K) \quad (5)$$

This is called an autoregressive moving average of order n and m, or in short ARMA(n,m). The unknown parameters in equation (5) are a_j, $j=1, \ldots n+m$.

If there is reason to believe that our process has deterministic terms and/or harmonic components with period T, or observable inputs, we may choose the general model as given by (6):

$$Y(K) = \sum_{j=1}^{n} a_j(K)Y(K-j) + \sum_{j=1}^{m} a_{n+j}(K)W(k-j) + \sum_{j=1}^{\ell} a_{n+m+j}(K)\Phi_j(K) + W(K) \quad (6)$$

where the integer ℓ is equal to all possible functions (deterministic or observable inputs) which we believe, for one reason or another, will improve the ability of our model to predict. For example, we may choose the following functions:

$\Phi_1 = 1$ $\qquad$ Φ_4 = some function of temperature

$\Phi_2 = \sin(\frac{2\pi k}{T})$ $\qquad$ $\Phi_5 = Y_K^2$

$\Phi_3 = \cos(\frac{2\pi k}{T})$ $\qquad$ $\Phi_6 = Y_K^3$, etc.

Finally, we can rewrite our general model in compact form as (7):

$$Y(K) = a^T(K)Z(K-1) + W(K) \quad (7)$$

where

$$a^T(K) = (a_1(K), \ldots, a_{n+m}(K), \ldots, a_{n+m+\ell}(K))$$

$$Z^T(K) = [Z_1(K), \ldots Z_n(K); Z_{n+1}(K), \ldots Z_{n+m}(K); Z_{n+m+1}(K), \ldots Z_{n+m+\ell}(K)]$$

$$= [Y(K-1)\ Y(K-n);\ W(K), \ldots W(K-m+1);\ \Phi_1, \ldots, \Phi_\ell]$$

The stochastic processes are divided into two classes: stationary and nonstationary. A stochastic process $\{Y(K)\}$ is stationary (in the strict sense) if its statistics are not affected by a shift in the time origin. Physically this means that the process is in a particular state of statistical equilibrium. A less restrictive requirement would be to assume that a process is weakly stationary [2]. This requires the existence of a time independent mean and autocovariance matrix for the process. Finally, a nonstationary process is a process which has a time-varying mean and variance.

A nonstationary stochastic process can be represented by equation (6) or (7).

Clearly, the model described by (6) or (7) is in the most general form. In the present study, we are only interested in one-step-ahead load prediction from the immediate past observations for on-line applications. To this end, we assume that a sample size of our process is locally stationary (weakly), and the input disturbance could be a discrete white noise. Hence, we only consider ARMA (n,m) model with a constant trend term which is given by (8):

$$Y(K) = a_0(K) + \sum_{j=1}^{n} a_j(K)Y(K-j) + \sum_{j=1}^{m} a_{j+n}(K)W(K-j) + W(K) \quad (8)$$

The time-varying coefficients $a_j(K)$ are introduced to account for the nonstationary behavior of one sample size to the next sample of the stochastic process describing the load demand of a power system.

THE ALGORITHM FOR ESTIMATION OF PARAMETERS OF PREDICTOR

We will use the algorithm developed by Kashyap in reference [1] for estimating the coefficients of the predictor.

In this method, the coefficients $a_j(K)$; $j = 0, \ldots n+m$ found in a recursive manner by minimizing the square of prediction error.

The algorithm for the coefficient of the predictor is given by (9):

$$S(K+1) = S(K) - S(K)Z(K)Z^T(K)S(K)/(1+Z^T(K)S(K)Z(K))$$

$$\hat{a}(K+1) = \hat{a}(K) + S(K+1)Z(K)(Y(K+1) - \hat{a}^T(K)Z(K))$$

$$\overline{W}(K+1) = Y(K+1) - \hat{a}^T(K+1)Z(K) \qquad (9)$$

The prediction error and predictor are given by (10):

$$\hat{Y}(K+1) = \hat{a}^T(K)Z(K)$$

$$e(K+1) = Y(K+1) - \hat{Y}(K+1) \qquad (10)$$

The initial values for $a(0)$ and $S(0)$ are arbitrary as long as $S(0)$ is a positive definite matrix. It is important to note the difference between $\overline{W}(K)$ and $e(K)$, as it is defined by (9) and (10); i.e., $\overline{W}(K)$ is only a residual and $e(K)$ is the prediction error.

MULTIPLICATIVE PREDICTOR [3]

If the observations are strictly positive, as in the case of the load demand of a power system, one can choose the multiplicative model as given by (11):

$$Y(K) = Y(K-1)^{a_1}\ Y(K-2)^{a_2} \ldots Y(K-n)^{a_n}\ W(K-1)^{a_{n+1}}$$
$$W(K-m)^{a_{n+m}}\ \Phi_1(K)^{a_{n+m+1}} \ldots \Phi\ell(K)^{a_{n+m+\ell}}\ W(K) \qquad (11)$$

If we let $y(K) = \ell n\,(Y(K))$ where $Y(K)$ is strictly positive, then, for the log transformed process $\{y(\cdot)\}$ we will have a one-step-ahead predictor of the form given by (12):

$$\hat{y}(K+1) = \hat{a}^T(K)Z(K) \qquad (12)$$

which is in the same form as (10). It should be clear that by replacing $Y(K)$ by $y(K)$ everywhere, we can apply algorithm (9) for finding the coefficients of the multiplicative predictor. It also should be pointed out that the effect of the log transformation is that of smoothing the fluctuations in the original process $\{Y(\cdot)\}$. Hence, the multiplicative predictor has a better chance of following the smoothed process $\{y(\cdot)\}$.

IDENTIFICATION

Basic Concepts

Before we proceed with a discussion of the method, the term identification should be defined. Zadeh [4] has given a definition that appears to be more or less generally accepted:

> "Identification is the determination, on the basis of input and output, of a process* (model), within a specified class of processes (models), to which the process under test is equivalent."

In order to use this definition, we must clarify what is meant by "equivalent process" and the "specific class of models." Clearly, the specific class of models to be considered here is linear systems with finite dimensions. This class contains many of the systems of practical importance. What is meant by an equivalent model is a little harder to describe. In a practical sense, the equivalent model is a model which has captured the underlying mechanism which has generated the original process. It should be pointed out that identification is necessarily inexact. It is inexact because the question of what types of processes (models) occur in practice cannot be determined by purely mathematical formulation.

Nevertheless, we accept an equivalent model from a class of finite dimensional stochastic models as a true model if it results in the smallest variance of error, or the smallest mean square error, or any other possible criteria over a finite past history of the process.

From the preceding discussion, it should be clear that we would like to identify the underlying structure which generates the process, that is, to determine the integers n and m.

Kashyap & Rao, in reference [5], tried different integers n and m and checked the accuracy of predictions of corresponding models, then chose the model among them which had the best prediction capability. This method determines n and m, that is, the order of the stochastic model which represents the structure of the finite past history of the process. It is clear that this method requires extensive simulation.

If the mechanism which will generate the entire future observations remains the same as the mechanism which generated the finite past history, then the model chosen this way would be entirely adequate. This is indeed the case in process control. For example, the output of a chemical plant can be modeled in this fashion. But, in the case of one-step-ahead load prediction, this is far from being adequate. What would happen a few minutes from now has very little or nothing to do with what has happened in the last few hours. Furthermore, unforeseen events like strikes, a sudden storm, or a fuel crisis, only to name a few, cannot be adequately represented in an equivalent model which represents the finite past history of the process. The above discussions clearly point out the need for frequent identification.

We chose to identify the integers n and m by analyzing a short sample size of the past history of the process at instant K in real time operation. Naturally, we would expect the model identified by this method to have a better or equal (if the process remains the same over finite history) prediction capability as the one identified by the previous method [5].

Before proceeding with on-line identification, the survey of literature in determining the order of the stochastic linear equation (n,m) from the autocorrelation and partial autocorrelation is in order.

As a first step in the modeling procedure [6], the model identification is done by examining the pattern of sample autocorrelations as well as partial autocorrelation functions. In practice, a set of such correlation plots is made for identifying possible underlying behavior. This method is not very effective when the process is ARMA (n,m). Furthermore, the success of identifying a mixed model relies on the ability of the individual analyst. Consequently, this method does not

*Zadeh originally used the word system here but the word process (model) is more appropriate. Our models are actually systems.

provide definite criteria for determining the exact order of a mixed model.

The problem of determining n and m was further discussed by Astrŏm and Bohlin [7] and was also analyzed by Trether and Steiglitz [8], and Jenkins and Wattes [9]. Later, Chow [10] gave a preliminary algorithm for determining the order of a linear stochastic dynamic system solely from the observation of the system output.

In the next sections, we further revise Chow's method and develop a recursive procedure which is designed to be simple, systematic and easily computerized.

ON-LINE IDENTIFICATION

In order that we might be able to use the identification algorithm on-line in real time operations, we must obey two restrictions:

1. At any instant K, only a finite, pre-specified number of real numbers can be stored in the computer, ruling out the possibility of storing a large sample size of the past history {y(j), $j \leq K$} at that instant.

2. The amount of computations such as the number of additions, multiplications, etc., needed to obtain the n and m shall be pre-specified and independent of K.

To overcome the above restrictions, a predetermined highest autoregressive order n is specified before searching for a model within (n,m) class of models. This requirement may seem to be restrictive, but hardly would present any problem in application because it is unlikely that we would need a model with AR and MA both of infinitely large orders.

THE RECURSIVE IDENTIFICATION ALGORITHM

Consider the linear difference equation for a mixed ARMA (n,m) process as given by (5) which is repeated below:

$$Y(K) = \sum_{j=1}^{n} a_j(K)Y(K-j) + \sum_{j=1}^{m} a_{n+j}(K)W(K-j) + W(K) \quad (5)$$

where we shall assume W(K) is a white excitation and is independent of Y(K-j) occurring prior to the time instant K. The autocovariance function for (5) can be written as (13):

$$C(j) = E[Y(K) - \bar{Y}][Y(K-j) - \bar{Y}] \quad (13)$$

Without loss of generality, we set the mean of Y(.), $E(Y) = \bar{Y} = 0$. Then we will have

$$C(j) = a_1C(j-1) + a_2C(j-2) + \ldots a_nC(j-n) + C_{yw}(j) + a_{n+1}C_{yw}(j-1) + \ldots + a_{n+m}C_{yw}(j-m) \quad (14)$$

where $C_{yw}(j) = E[Y(k-j)W(K)]$ is a cross-covariance function of lag j between y and w. The Y(K-j) depends on noises which have occurred up to the instant K-j. Therefore,

$$C_{yw}(i) = 0 \qquad i > 0$$

$$C_{yw}(i) = \Psi(i)\sigma_w^2 \qquad i \leq 0$$

which implies that for $j \geq m+1$ we will have

$$C(j) = a_1C(j-1) + a_2C(j-2) + \ldots + a_nC(j-n) \quad (15)$$

If the process is stationary, then variance $\sigma_y^2 = C_0$ is constant. From (15) we can write the autocorrelation for the process as (16):

$$\rho(j) = \frac{C(j)}{C_0} \text{ or } \rho(K) = \frac{C(K)}{C_0} \quad (16)$$

where we have changed j to K, finally.

$$\rho(K) = a_1\rho(K-1) + a_2\rho(K-2) + \ldots + a_n\rho(K-n) \quad (17)$$

$$\text{for } K \geq m+1$$

Equation (17) says that in ARMA(n,m) models, the autocorrelation function after lag m follows the Yule-Walker equation exactly the same way as does a pure AR process. This is the key to the method. Using (16), we can write (14) in compact notation as (18):

$$\rho(K) = \sum_{i=1}^{n} a_1\rho(K-i) + Q(K) \quad K = 1,2 \ldots \quad (18)$$

where

$$Q(K) = \frac{\sigma_w^2}{C_0} \sum_{i=0}^{m} \beta_i\Psi(i-k) \text{ for } K \leq m$$

$$= 0 \qquad \text{for } K > m$$

For a given $n=n^*$, the modified Yule-Walker equation in matrix form is given by (19):

$$\begin{bmatrix} \rho(m+1) \\ \rho(m+2) \\ \vdots \\ \rho(m+n^*) \end{bmatrix} = \begin{bmatrix} \rho(m) & \rho(m-1) \ldots & \rho(m-n^*+1) \\ \rho(m+1) & \rho(m) \ldots & \rho(m-n^*+2) \\ \vdots & \vdots & \vdots \\ \rho(m+n^*-1) & \rho(m+n^*-2) \ldots & \rho(m) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{n^*} \end{bmatrix} \quad (19)$$

Hence

$$a_i = \frac{V^m(i)}{D^m} = \frac{\begin{vmatrix} \rho(m) & \ldots & \rho(m+1) & \ldots & \rho(m+n^*-1) \\ \vdots & & \vdots & & \vdots \\ \rho(m+n^*-1) & \cdot & \rho(m+n^*) & \ldots & \rho(m) \end{vmatrix}}{\begin{vmatrix} \rho(m) & \ldots & \rho(m+1-i) & \cdot & \rho(m+n^*-1) \\ \vdots & & \vdots & & \vdots \\ \rho(m+n^*-1) & \cdot & \rho(m+n^*-i) & \ldots & \rho(m) \end{vmatrix}} \quad i=1,2\ldots n^* \quad (20)$$

For identification purposes, we only need to compute the $V^m(i)$ which are the determinants in the numerator. Equation (20) will determine the first n^* of V's.

For $i > n^*$, the lag V's can be calculated recursively by (21):

$$V^m(i) = D^m\rho(m+i) - V^m(1)\rho(m+i-1) - \ldots - V^m(n^*)\rho(m+i-n^*) \quad (21)$$

The test for determining the order of ARMA (n^*,m) consists of the following steps:

1. Fix the order of MA at m = 0.

2. Calculate $V^m(i)$ using (20) and (21) where i = 1, 2, ...

3. Check the significance of $\hat{V}^{(m)}(i)$. If $\hat{V}^{(m)}(i) \neq 0$ for $i > n^*$, then it indicates that the model is not ARMA (n^*,m). On the other hand, if $\hat{V}^{(m)}(i)$ is insignificant for $i > n_1$, where $n_1 \leq n^*$, then the model is ARMA (n_1,m).

4. Otherwise, fix $m=m+1$ and repeat steps 2 and 3.

If a model cannot be found when $m = m^*$, where m^* is a preselected MA order, then this indicates the need for fitting a model higher than ARMA (n^*,m^*). This problem can be avoided if high enough values of n^* and m^* are used for a given type of process. In the Appendix the identification of ARMA(2,m) is discussed.

RESULTS OF PARTICULAR STUDIES

From the identification and estimation algorithms proposed in this paper, a computer software package was developed [11]. The load data of Public Service Indiana for one-minute intervals (125 points), five-minute intervals (425 points) and hourly load (550 points) were used for one-step-ahead forecasting. The graphs of each case with associated errors vs. time are given in Figures 1, 2, and 3 respectively. The simulations of each time series showed that the one-minute load is ARMA (1,0) process, and the five-minute load is mostly ARMA (1,1) and occasionally ARMA (2,1) process. The hourly load is ARMA (2,0) process.

After identifying the order of the stochastic process (i.e. n,m) representing each time series, other models of the order $n\pm1$, $m\pm1$ were also simulated. In each case, the changing of the order of the identified model resulted in increased errors, which also confirms that the identified model adequately represents the underlying mechanism of the process [11]. The results for the final coefficients and errors for the three cases are given in Table 1.

CONCLUSIONS

The proposed identification and estimation algorithms can be successfully used for on(off)-line application of one-step-ahead prediction of load demand of a power system.

In addition, the advantages of the proposed method are that the method is completely automatic in identifying a non-seasonal time series, and that it is recursive in estimating the parameters of the identified model, in contrast to the usual approach which depends on the capability of the individual analyst.

As far as the applications of one-step-ahead load forecasting are concerned, the following can be mentioned:

1) 1-10 minute intervals for economic dispatch

2) 10-30 minute intervals for reserve evaluations and security calculations

3) 30-60 minute intervals for contingency assessments.

Clearly, the method can be used for forecasting total system load, a region load, or a bus load.

ACKNOWLEDGMENTS

The authors wish to thank Public Service Indiana and Mr. M. E. Kuchefski for providing the load data used in the analysis. The financial support by NSF/RANN grant GI-38315 is gratefully acknowledged.

Table 1: The Final Coefficients and Errors for the Various Models

Model Multiplicative	Model Coefficients			Error (MW)			Confidence Region (MW)	
	$\hat{a}_0(N)$	$\hat{a}_1(N)$	$\hat{a}_2(N)$	$E_{0,N}$	$E_{1,N}$	$E_{2,N}$	67% $\Pr\{-\hat{\sigma} \leq e \leq \hat{\sigma}\} \simeq .67$	97% $\Pr\{-2\hat{\sigma} \leq e \leq 2\hat{\sigma}\} \simeq .97$
AR(1,0) (1-min. load) N = 125	.0006	.978		- .899	7.2	.164	-13.7, 11.9	-26.4, 24.6
ARMA(1,1) (5-min. load) N = 425	.036	.993	.112	1.41	8.8	.134	-10.0, 12.0	-21.0, 24.0
AR(2,0) (hourly load) N = 550	.037	1.61	-.70	-1.93	30.9	1.62	-42.0, 38.8	-82.2, 78.5

where $E_{0,N} = \sum_{j=1}^{N} e(j)$ $\quad E_{1,N} = \frac{1}{N}\sum_{j=1}^{N} |e(j)|$ $\quad E_{2,N} = \frac{1}{N}\sum_{j=1}^{N} e^2(j)$

Assuming that the errors are normal, the upper and lower probability limits are

$$Y(K+1)(\text{lower/upper}) = \hat{Y}(K+1) + \text{confidence region (67\% or 97\%)}$$

That is $P_r\{Y(K+1)\ \text{lower} \leq Y(K+1) \leq Y(K+1)\text{upper}\} \simeq .67$ or .97 respectively.

REFERENCES

[1] Kashyap, R. L., "Estimation of Parameters in a Partially Specified System," Proceedings of the Ninth Allerton Conference on Circuits and Systems, University of Illinois, (Aho) Technical Report, Electrical Engineering Department, Purdue University (1971).

[2] Papoulis, A., Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 1965.

[3] Teekens, R., Prediction Methods in Multiplicative Models. Rotterdam University Press, 1972.

[4] Zadeh, L. A., "From Circuit Theory to System Theory," IRE, 50, pages 856-865, 1962.

[5] Kashyap, R. L. and Rao, A. R., "Real Time Recursive Prediction of River Flows for Optimal Operation of Water Resources Systems," Automatica, Vol. 9, pp. 175-183. Pergammon Press, 1973.

[6] Box, G. P. and Jenkins, G. M., Time Series Analysis Forecasting and Control, Holden-Day, 1971.

[7] Aström, J. J. and Bohlin, T., "Numerical Identification of Linear Dynamic Systems from Normal Operating Record," Proc. 2nd IFAC Symposium on Self-Adaptive Control System. Ed. by Hammond, Plenum Press, N. J.

[8] Tretter, S. A., and Steiglitz, K., "Power Spectrum Identification in Terms of Rational Models." IEEE, Automatic Control, April 1967, pp. 185-188.

[9] Jenkins, G. M., and Watts, P. G., Spectral Analysis and Its Application, Ch. 5, pp. 197-204. Holden Day, San Francisco, 1968.

[10] Chow, J. C., "A Preliminary Algorithm for Determining the Order of a Linear Stochastic Dynamic System," Proc. IEEE Decision Control Conference; 1971, pp. 518-520.

[11] Keyhani, A., and El-Abiad, A., "One-Step-Ahead Load Forecasting Using Recursive Identification and Estimation Techniques, TR-EE-74-32, Purdue University, 1974.

APPENDIX

IDENTIFICATION OF ARMA (2,m) MODELS

For a special case when $n^* = 2$, equation (19) can be written as

$$\begin{bmatrix} \rho(m+1) \\ \rho(m+2) \end{bmatrix} = \begin{bmatrix} \rho(m) & \rho(m-1) \\ \rho(m+1) & \rho(m) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \qquad (22)$$

From (22) a_1 and a_2 can be written as (23):

$$a_1 = \frac{V^m(1)}{D^m}, \quad a_2 = \frac{V^m(2)}{D^m} \qquad (23)$$

where $V^m(1)$ and $V^m(2)$ are given by (24):

$$V^m(1) = \rho(m)\rho(m+1) - \rho(m-1)\rho(m+2)$$
$$V^m(2) = \rho(m)\rho(m+2) - \rho(m+1)\rho(m+1) \qquad (24)$$

Using equation (21), the lag V's can be calculated recursively for $i > 2$ by (25):

$$V^m(i) = D^m\rho(m+i) - V^m(1)\rho(m+i-1) - V^m(2)\rho(m+i-2) \qquad (25)$$

where

$$D^m = -V^{m-1}(2) \text{ for } m > 0$$

The equations (24) and (25) can be used for identifying any ARMA(2,m) models. In order to see the significance of V^m's clearly, we compare the ratio of $|V^m(i)|/\sigma_{V^m}$ where $\sigma_{V^m} \simeq \frac{1}{\sqrt{N}}|D^m|$. In actual identification $\rho(K)$'s are estimated by sample autocorrelations r(K)'s. Next, we demonstrate the method by means of two examples:

Example 1: 100 points of the hourly load data were used as a sample size for recursive identification of the process. The first ten lags of V's with m=0 are calculated, and the results are given in Table 2.

Table 2

i	r(i)	$V^0(i)$	$\lvert V(i)\rvert/\sigma_{V^0}$
1	.958	.1396	21.39
2	.854	-.064	9.83
3	.711	.0002	.03
4	.550	.0008	.12
5	.404	.0011	.17
6	.272	.0014	.21
7	.169	.0018	.27
8	.103	.0022	.33
9	.072	.0024	.36
10	.063	.0017	.27

Clearly, the $V^0(i)$'s are insignificant after lag 3. Therefore, the sample process is identified as AR(2,0).

Example 2: In this example, 100 points of five-minute load were used as a sample size for recursive identification. The first ten lags of V's with m=0 and m=1 are calculated, and the results are given in Table 3.

Table 3

i	r(i)	$V^0(i)$	$\lvert V^0(i)\rvert/\sigma_{V^0}$	$V^1(i)$	$\lvert V^1(i)\rvert/\sigma_{V^1}$
1	.127	.118	1.51	-.097	13.77
2	.068	.052	.67	.0038	.89
3	.066	.0507	.64	-.0027	.64
4	.119	.106	1.35	.0026	.63
5	.076	.057	.73	-.0039	.94
6	.150	.133	1.69	-.0011	.26
7	.181	.156	1.99	.0082	1.96
8	.032	.002	.03	-.0074	1.76
9	.163	.147	1.87	.0019	1.87
10	.141	.118	1.50	.0076	1.50

Clearly, with m=0, all the lags of V's, except for V(8), have about the same significance. Therefore, no decision can be made; and we proceed to check the significance of V^1's. It is clear that only $V^1(1)$ is significant. Thus the sample process of five-minute load is identified as ARMA (1,1).

In reference [11], the identification of ARMA (4,m) is discussed.

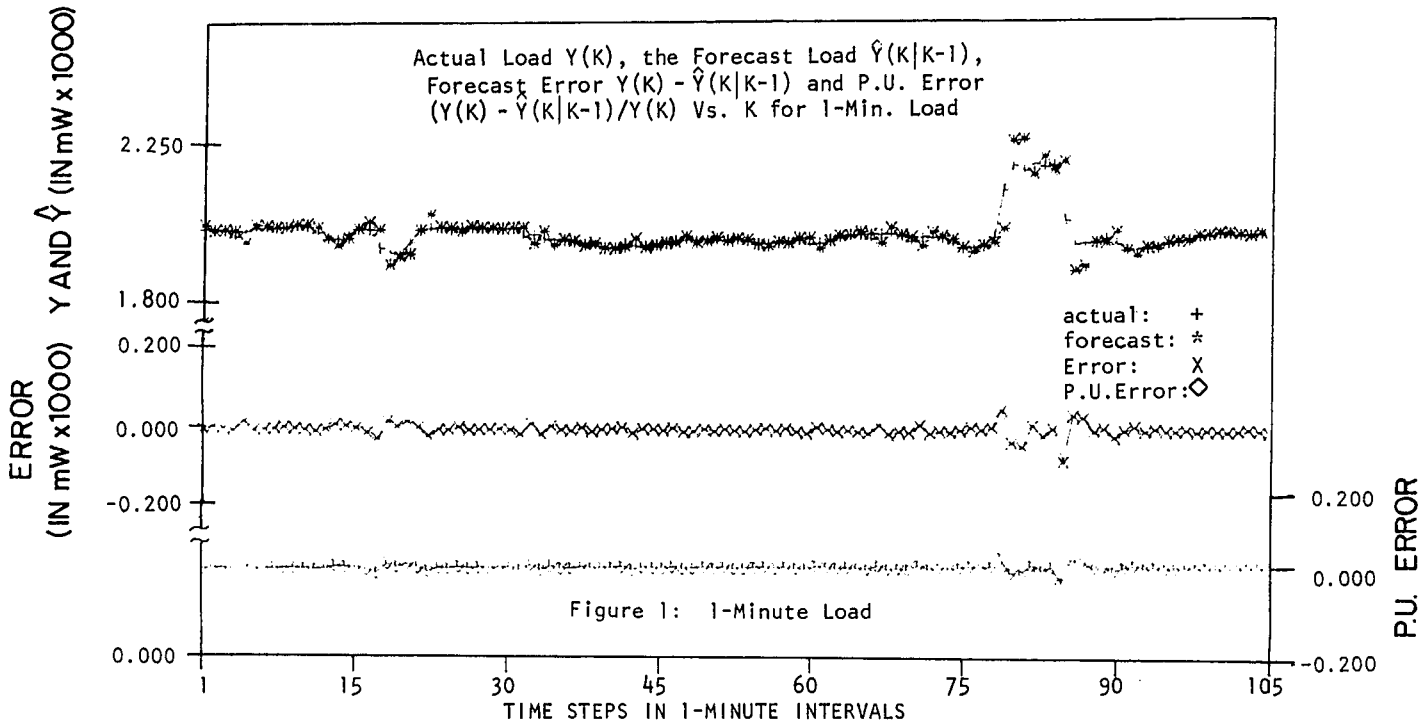

Figure 1: 1-Minute Load

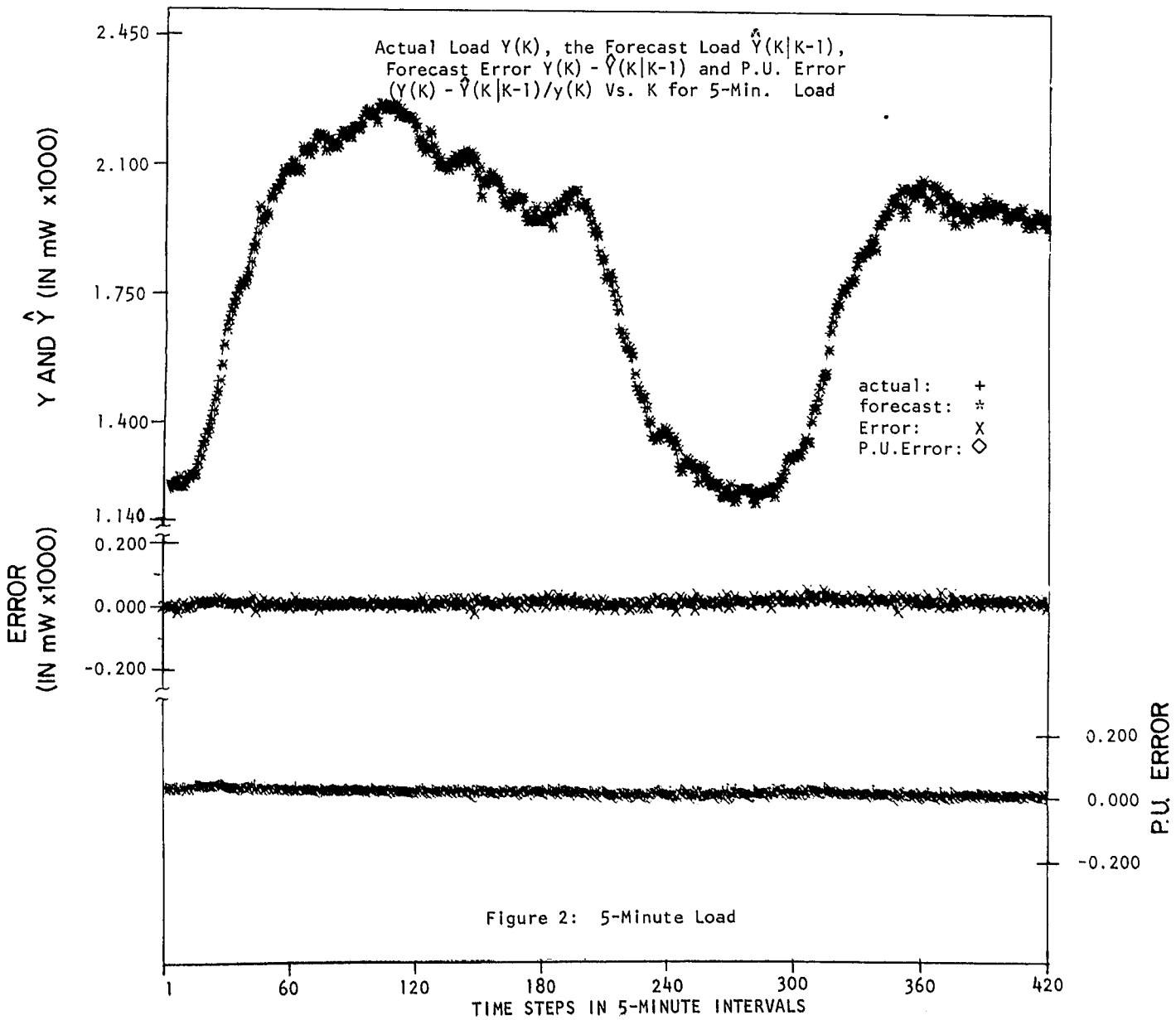

Figure 2: 5-Minute Load

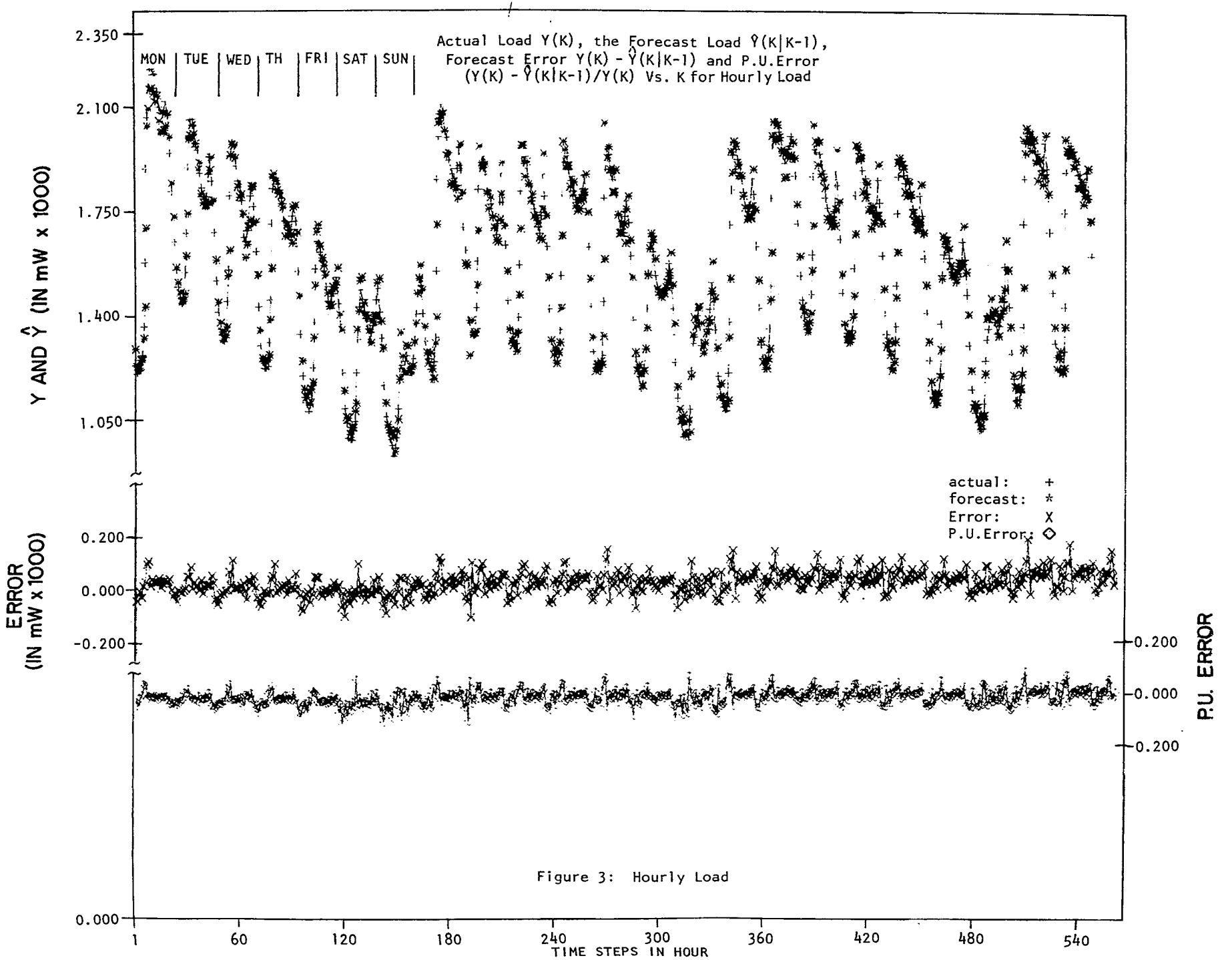

Figure 3: Hourly Load