論 文
32～6～1

칼만 필터와 시계열을 이용한 순환 단기 부하예측

Recursive Short-Term Load Forecasting Using Kalman Filter and Time Series

朴 永 文* · 鄭 正 周**
(Young-Moon Park · Chung-Choo Chung)

요 약

本 論文은 電力系統負荷의 短期豫測에 適用될 수 있는 다른 형태 모형의 應用에 관하여 記述하였다.

칼만 필터와 시계열을 利用한 온-라인 앨고리즘으로 한 시간 後의 電力負荷를 豫測하도록 구성하였다.

보흘린이 提示한 모델을 利用하여 24時間 단위로 負荷를 먼저 豫測하고, 박스-젠킨스 모형의 시계열을 利用하여 한 시간 후의 잔차를 매 시간 豫測하여 보정하였다.

칼만 필터에서 시스템 방정식 및 측정 방정식은 고정시켰으며 시계열의 패러미터들은 한 주를 주기로 변화시켰다.

한국 전력 공사의 1981년 4월 부터 8월까지 한 시간 간격의 데이타를 갖고서 시뮬레이션한 결과 1.2% rms 오차가 있었다.

Abstract

This paper describes the application of different model which can be used for short-term load prediction. The model is based on Bohlin's approach to first develop a load profile model representing the nominal load component and the Box-Jenkins approach is used to predict residuals.

An on-line algorithm using Kalman Filter and Time Series is implemented for an hour-ahead prediction.

In the Kalman Filter system equation and measurement equation were fixed and parameters of Time Series were varied week after week.

A set of data for Korea Electric Power Corporation from April to June 1981 was used for the evaluation of the model. As the result of this simulation 1.2% rms error was acquired.

1. Introduction

One of the main objects of electric power system control is to provide the load as economically as possible to a certain security level. Therefore on-line short-term forecasts are useful for the commitment of equipment and spinning reserve allocation, for reliability calculations related to uncertainty in load forecast for economic allocation of generation.

* 正 會 員：서울大 工大 電氣工學科 教授・工博
** 正 會 員：(株) 金星社 中央研究所 研究員
接受日字：1983年 1 月31日

Literature in short-term load forecasting dates as far as 1966. At present, techniques and algorithms in this area include at least five fundamental methods which are state estimation techniques, method utilizing load weather linear regression, spectral expansion techniques, exponential something techniques, application of Box-Jenkins modeling concepts.[1)]

In the area of load forecasting we have to consider the accuracy, the simplicity and the apaptivity of the models of the nominal load. Moreover the computation efficiency is required.[1),2),2)]

In the reference G.D. Irisarri suggested the BGLS algorithm and as he commented, model identification and paramenter tuning of the BGLS model is not a trivial task. Unfortunely as with most practical problems where modelling is of prime importance, a certain amount of educated quessing coupled with trial and error experimentation is necessary. (1), page 75) Without regard to this fact it is generally a large task to obtain the weather sensitive term.

By reason of the above fact, we proposed the different model with is robust and can be used without artificial model changes for a some duration. And due to its adaptive nature the model parameter are continuously changed weekly to track seasonal (dialy, weekly) load variations as well as large range trends.

In this paper the nominal load characteristic is modeled by 24 hour load profiles with enough type to characterize the weekly seasonality by the method which was suggested by Bohin's approach.[1),3)]

For any given data type the discrepancy between the actual load and load profile model exists because 24 hour load prediction can not describe the component affected by the exogeneous input and random noise component, e.g., the weather, so that such residuals is modeled by the ARMA model.

2. Development of forecasting load model

2.1 Nominal load model using Kalman Filter

(a) Modeling of state and measurement equation

For $k=1,2,\ldots,9$ let $Y_k(t)$ the niminal load on day k at hour t and let $X_i(t)$ $i=1,2,\ldots,9$ be the the load profile state variables. (i means the special day type of a week.)

The load profile model for day k, at hour t is given as [1], [3]

$$Y_k(t) = x_1(t) + \delta_2(k)X_2(t) + \ldots + \delta_9(k)X_9(t) + r(t) \qquad (1)$$

$$\text{where } \delta_i(k) = \begin{cases} 1, & \text{for } i=k,\ i=1,2,\ldots,9 \\ 0, & \text{otherwise} \end{cases}$$

The term $r(t)$ represents a random noise, which accounts for the modelling erors and it will be assumed that $r(t)$ also includes measurement errors.

And if we let

$$\underline{H}_k = [\,1\ \delta_2(k)\ \delta_3(k) \cdots\cdots \delta_9(k)\,] \qquad (2)$$

The Eq. (1) can be written as follow.

$$\underline{Y}_k(t) = \underline{H}_k\,\underline{X}(t) + r(t) \qquad (3)$$

where $X_1(t)$ = the reference nominal base load for hour t

$X_2(t)$ = The Monday incremental nominal load from $X_1(t)$ at hourt t

⋮

$X_8(t)$ = the Sunday incremental nominal load from $X_1(t)$ at hour t

$X_9(t)$ = the odd day incremental nominal load from $X_1(t)$ at hour t

Where $\{r(t),\ t=1,2,\ldots\ldots\}$ is white Gaussian sequence with mean zero and covariance $R_y(t)$ Assume for any t the dynamics of the load profile state variable are given by

$$X_i(t+T) = X_i(t) + W_i(t) \qquad (4)$$

where $i=1,2,\ldots,9$ and $t=1,2,\ldots,T$

The term $W_i(t)$ was used to compomensate the change of load after time T, which is white Gaussian sequence with mean $E\{W_i(t)\}$ zero and covariance $E\{W_i(t)\,W_j(s)\} = \delta_i(t)\,\delta_j(l)\,\delta_s(t)$.

For an example if $i=1$ (This means the reference nominal base load.), the tomorrow reference nominal base load will be sum of today reference nominal load $X_1(t)$ and today random noise $W_1(t)$. For another example if $i=2$, the next week Monday increment nominal load will be the sum of this week Monday reference nominal load $X_2(t)$ and this week Monday random noise $W_2(t)$.

So far we made the state equation which describes the nominal load. The state equation can be used by kalman Filter to predict 24 hour nominal load. In order to use the Kalman Filter it is needs to re-writer the state equation, measurement equation.

state equation : $\underline{X}(t+T) = A\,\underline{X}(t) + \underline{W}(t)$ (5)

measurement equation : $Y_k(t) = \underline{H}_k\,\underline{X}(t) + r(t)$ (6)

where $X(t) = [X_1(t) \quad X_2(t) \;\ldots\ldots X_9(t)\;]$

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$E\{W(t)\} = \underline{0},\; E\{\underline{W}(t)\,\underline{W}^T(t)\} = R_x(t)$$

$$\underline{W}(t) = [W_1(t) \quad W_2(t)\ldots W_9(t)]$$

(b) A description for the filtering equation

Thanks to the assumption that a point on a profile is coupled only to that of the previous day, the model is seperated into T independent ninth order model, all defined by (5) and (6), and so does the Kalman Filter. At this point if it is assumed that $R_x(t)$ and $R_y(t)$ are independent of time, the solutions of Riccati equations (filtering equation) will be equal.[3]

Let $t=kT+\tau$, where k indicates the day, t the time of the day. For $\tau=1$ and all k compute[3)4),10)]

$$P_{yy}(k\mid k\text{-}1) = \underline{H}_k\,P_{xx}(k\mid k\text{-}1)\,\underline{H}_k^T + R_y \quad (7)$$

$$\underline{K}(k) = P_{xx}(k\mid k\text{-}1)\;\underline{H}_k^T\;P_{yy}^{-1}(k\mid k\text{-}1) \quad (8)$$

$$P_{xx}(k\mid k\text{-}1) = R_x + P_{xx}(k\mid k\text{-}1) - \underline{K}(k)\;P_{yy}^{-1}(k\mid k\text{-}1)\,\underline{K}(k)^T \quad (9)$$

At a given day k, for each $t=1,2,\ldots, T$ the residual load component is computed as each obserbation becomes available by using

$$r(t) = \widetilde{Y}_k(t\mid t\text{-}T) = Y_k(t) - \underline{H}_k\,\hat{\underline{X}}(t\mid t\text{-}T) \quad (10)$$

the t-th row of the state matrix X is updated by

$$\underline{X}(t+T\mid t) = \underline{X}(t\mid t\text{-}T) + \underline{K}(k)\;\widetilde{Y}_k(t\mid t\text{-}T) \quad (11)$$

2.2. Redisual load model using Time Series

In the nominal load model described before no recent load values of the last hours are used for the prediction of the load, and it may be expected that an improvement of the prediction will be obtained when more recent data is used.

A general class of models which are useful for this purpose are the so-called Box-Jenkins model. As daily and weekly periodicity in the load curve are eliminated by the Kalman Filter, the Box-Jenkins model is apparently simplified.[1]

(a) A brief description for AR, MA and ARMA model

If the residual process $r(t)$ is stationary, $r(t)$ can be represtnted as one of autoregressive or moving average or mixed autoregressive moving average. (from now we note r_t stead of $r(t)$)[6),7)]

(a-1) Autoregressive model

This model is based on the fact that the current value r_t is written as a finite linear combination of previous values $r_{t-1}, r_{t-2}, \ldots, r_{t-p}$

$$r_t = \theta_0 + \sum_{j=1}^{p} \phi_j\,r_{t-j} + a_t \quad (12)$$

or $(1-\phi_1 B - \ldots\ldots - \phi_p B^p)\,r_t = \phi(B)\,r_t = \theta_0 + a_t$ (13)

where $\mu = E\{r_t\},\; E\{a_t\} = 0$

$$E\{a_t a_s\} = \delta_a^2\;\delta_t(s)$$

$$\theta_0 = \mu\,(1 - \sum_{j=1}^{p}\phi_j)$$

ϕ_j = weight of AR parameters

For this process becomes the stationary process all roots of characteristic equation $\phi(B)$ exist outside of the unit circle.

(a-1) Moving average model

Another model is based on fact that the current value r_t can be represented by a linear combination of a finite number of forecasted errors.

Let $a_t = r_t - \hat{r}_t$, then

$$r_t = \mu + \sum_{i=1}^{q} \theta_i\,a_{t-i} + a_t \quad (14)$$

where θ_i as weights of moving average parameters

(a-3) Autoregressive moving average model

Often a more parsimonious model can be achieved with mixed autoregressive moving average model of (p,q)

$$r_t = \theta_0 + \sum_{j=1}^{p} \phi_j r_{t-j} + \sum_{i=1}^{q} \theta_i a_{t-i} + a_t \qquad (15)$$

or $\phi(B)\, r_t = \theta(B)\, a_t$

It may be thought of in two ways. Namely

(a-3-1) as a p-th order autoregressive process

$\phi(B)\, r_t = e_t$

with e_t following the *q-th* order moving average process $e_t = \theta(B)\, a_t$

(a-3-2) as a *q-th* order moving average process

$r_t = \theta(B)\, b_t$

with b_t following the *p-th* order autoregressive process

$\phi(B)\, b_t = a_t$

so that

$\phi(B)\, r_t = \theta(B)\, \phi(B)\, b_t = \theta(B)\, a_t$

$\theta(B)\, r_t = \theta(B)\, a_t$ will be stationary process, provided that the characteristic equation $\theta(B) = 0$ has all its roots lying outside the unit circle.

Similarly, the roots of $\theta(B) = 0$ must lie outside unit circle if the process to be invertible.

(b) Interative stages in the selection of a model

The identification of the model and estimation of its parameters consist of a three-step iterative procedure. (refer to figure 1). The initial estimate of the model parameters $\theta_0, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$ is based on the principle that an ARMA *(p,q)* process has a unique covariance structure. The autoregressive parameter $\phi_1, \phi_2, \ldots, \phi_p$ are estimated from the calculated autocovariances by solving p linear equations written as (16)

The linear equation which we are solving is as following.

$$A\Phi = X \qquad (16)$$

where $A_{ij} = C_{|q+i-j|}$, $i,j, = 1,2, \ldots, p$

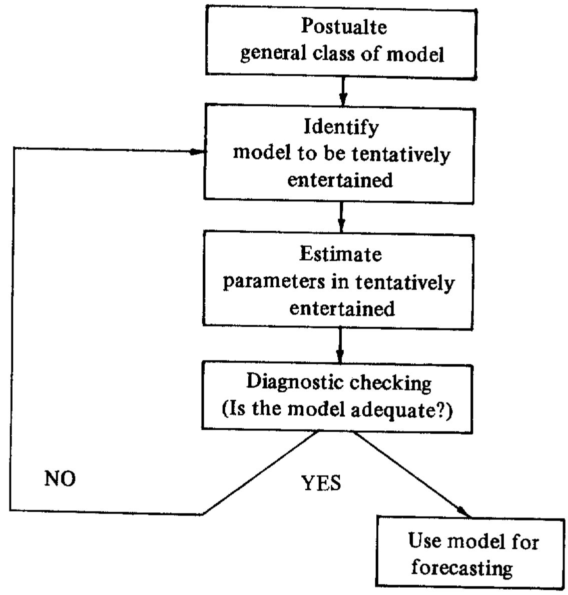

Fig. 1. Selection of appropriate model.

$$X_i = C_{q+i}$$

$$C_k = \frac{1}{N} \sum_{t=1}^{N-k} (r_t - \mu)(r_{t+k} - \mu), \qquad k=1,2,\ldots,K$$

N: number of observation

K: maximum lag of autocorelation function

or

$$\begin{bmatrix} C_q & C_{q+1} & C_{q+2} \cdots\cdots & C_{q+p-1} \\ C_{q+1} & C_q & C_{q+1} \cdots\cdots & C_{q+p-2} \\ C_{q+2} & C_{q+1} & C_q \cdots\cdots & C_{q+p-3} \\ \vdots & \vdots & \vdots & \vdots \\ C_{q+p-1} & C_{q+p-2} & C_{q+p-3} \cdots\cdots & C_q \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} C_{q+1} \\ C_{q+2} \\ C_{q+3} \\ \\ C_{q+p} \end{bmatrix}$$

After solving this linear equation, let $\phi_1, \phi_2, \ldots, \phi_p$ be estimated AR parameters $\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_p$. And the moving average parameters $\theta_1, \theta_2, \ldots, \theta_q$ should be estimated. For this purpose a time series $\{z_t\}$ is defined according to

$$z_t = r_t - \sum_{j=1}^{p} \phi_j r_{t-j} - \phi_0 \qquad (17)$$

It is supposed that $\{z_t\}$ can be described by an MA (q) model. The autocovariance at lag k of the MA (q) process can be written as

$$C'_k = \begin{cases} (-\theta_k + \theta_1 \theta_{k+1} + \cdots\cdots + \theta_{q-k} \theta_q) \sigma_a^2 \\ \quad , \text{for } k=1,2,\ldots,q \\ 0, \text{for } k>q \end{cases} \tag{18}$$

and $C'_0 = (1+\theta_1^2 + \ldots\ldots + \theta_q^2)\sigma_a^2$

The autocovariances of the MA (q) process $\{z_t\}$ can be expressed in the estimated autoregressive parameters ϕ_i and autocovariances of the process $\{z_t\}$ according to

$$C'_k = \begin{cases} \sum_{i=0}^{p} \sum_{j=0}^{p} \hat{\phi}_i \hat{\phi}_j C_{|j+i-k|}, \; p>0 \; (\phi_0 = -1) \\ C_j, \; p=0 \end{cases} \tag{19}$$

From (18) and (19) the parameters $\phi_1, \ldots\ldots, \phi_1$ and σ_a can be solved. As the equations are nonlinear, the Nowton-Raphson algorithm is used to solve these equations.

So far we have obtained only the initial estimated parameter, but under these parameters density function usually does not be maximized.[2),6),7)]

For most situations the maximum likelihood estimates are closely approximated by the least square estimate which make

$$S(\underline{\phi},\underline{\theta}) = \sum_{t=-\infty}^{N} [a_t \mid \underline{\phi}, \underline{\theta}, r_t]^2 = \sum_{t=-\infty}^{N} [a_t]^2 \tag{20}$$

a minimum, and in practice, the infinite sum can be replaced by a manageable finite sum $\sum_{t=1-Q}^{N} [a_t]^2$

In the case of AR process the solution can be obtained easily by applying least square method, but in the case of process we have to solve nonlinear equation. To overcome this problem there are several methods such that steepest descent method and Marquardt algorithm.[6),11),12)]

In this paper with the initial estimated parameters the steepest descent method is used to optimize the sum of residuals.

If parameters are optimized then we should test if this model is adequate. The test is based on the assumption that the autocorrelation $\rho(k)$ of the residuals a_t of the best ARMA (p,q) model will have a zero mean and a variance reciprocal to the length of the Time Series. Consequently[7),8)]

$$Q = n \sum_{i=1}^{K} \rho(i) \tag{21}$$

where $n=N-p-q$

can be approximated by a Chi-square distribution with $v=k-p-q$ degree of freedom. Here K is the maximum lag of autocorrleation function and N is the number of obserbations. K should be efficiently large so that the weight Ψ_j in the model, written in the form

$$r_t = \phi^{-1}(B)\,\theta(B)\,a_t = \Psi(B)\,a_t$$

will be negligibly small after $j=k$. From the Chi-square test if $Q<x^2_{v,\alpha}$ the model is accepted. (α is confidence limit. Usually the value of α is 0.025 or 0.01)

The identification process is started with an underfitted low order ARMA (p,q) model, and if the model is rejected by the Chi-square test, alternatively q are increased until the test is accepted.

3. Forecasting algorithm

Until now we have described nominal load model and residual load model. Nominal load is modeled by state equation in order to forecast 24 hour ahead load with Kalman Filter and residual load by ARMA so as to predict one hour ahead load with forecasted residuals by Kalman Filter.

Once the statistics of the state equation is determined, that is not changed in the forecasting routine as long as we don't change. But the parameters of ARMA process will be varied by one week period because after one week ARMA is remodeled by past two week residuals of Kalman Filter.

In this paper Kalman Fliter equation form, Eq. (22), proposed by Joseph[4),5),10)] is used in stead of Eq.(9) to make $P_{xx}(k+1 \mid k)$ insensitive to errorsin the filter gain and obtain more faithfully the positive definiteness and symmetry of $P_{xx}(k+1 \mid k)$.

$$P_{xx}(k+1 \mid k) = R_x + [I - \underline{K}(k)\underline{H}_k]\, P_{xx}(k \mid k-1)$$

$$\times [I-\underline{K}(k)\underline{H}_k]^T + \underline{K}(k)R_y\underline{K}(k)^T \quad (22)$$

Forecasting algorithm

Step 0.

Initialize the time t and the day type k and model Kalman Filter parameters and time series parameters.

Step 1

Compute the Kalman Filter parameters $P_{yy}(k\mid k-1)$, $\underline{K}(k)$ and $P_{xx}(k+1\mid k)$ in the order specified.

$$P_{yy}(k\mid k-1) = \underline{H}_k P_{xx}(k\mid k-1)\underline{H}_k^T + R_y \quad (7)$$

$$\underline{K}(k) = P_{xx}(k\mid k-1)\underline{H}_k^T P_{yy}^{-1}(k\mid k-1) \quad (8)$$

$$P_{xx}(k\mid k+1) = R_x + [I-\underline{K}(k)\underline{H}_k P_{xx}(k\mid k-1)$$

$$\times [I-\underline{K}(k)\underline{H}_k]^T + \underline{K}(k)\underline{R}_y K(k)^T \quad (22)$$

Step 2

Forecast the nominal load for day type k

$$Y_k(t+T\mid t) = \underline{H}_k \underline{X}(t+T\mid t) \quad t=1,2,\ldots,T \quad (23)$$

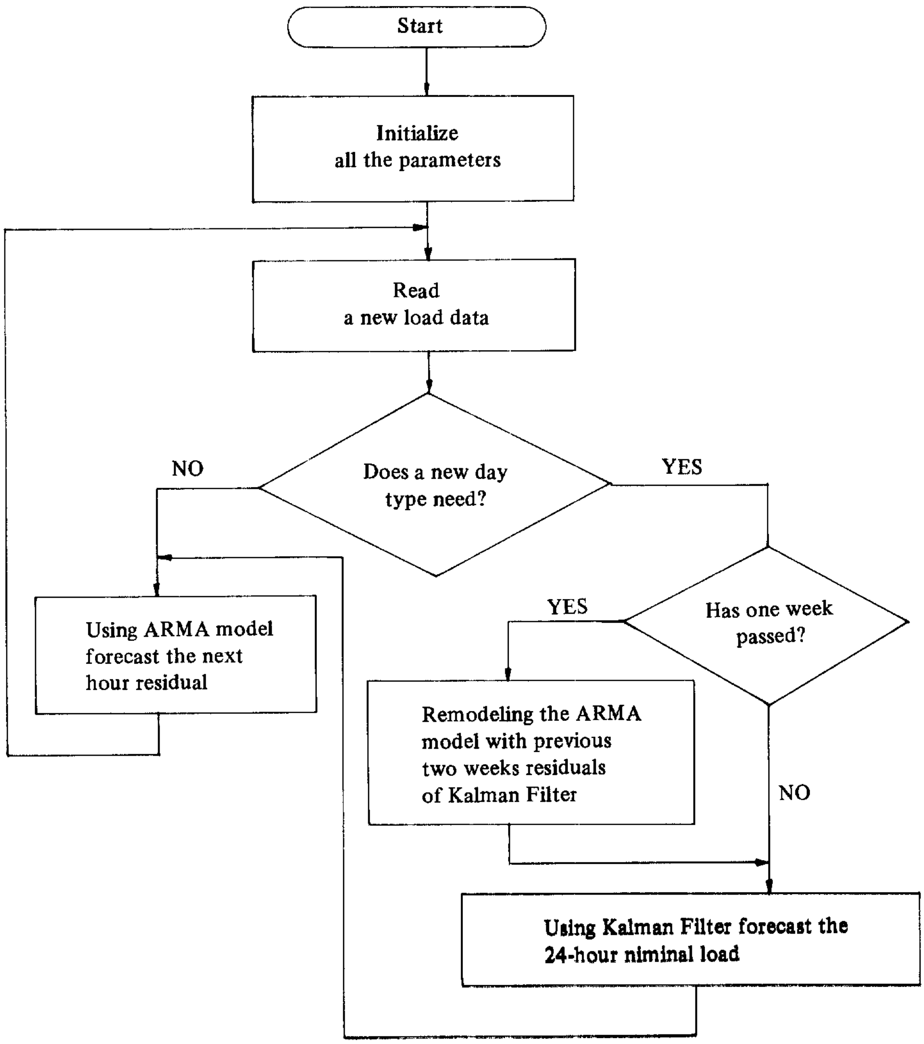

Fig. 2. Simulation flowchart.

Step 3

When a new data is available predict the next hour residual load until a new day type appears.

$$r_t = \mu + \sum_{i=1}^{p} \phi_i r_{t-1} + \sum_{j=i}^{q} \theta_j a_{t-j} + a_t \qquad (15)$$

where $a_t = \begin{cases} r_t - \hat{r}_t, & t<0 \\ 0, & t \geq 0 \end{cases}$

Step 4

If a new day type appears, test if one week has passed . In the case that one week has passed then go to step 5 otherwise go to step 1.

Step 5

Remodeling the ARMA model using the past two weeks residuals due to Kalman Filtering.

The flow for this algorithm is given at Figure 2

4. Numerical Example

A set of data for Korea Electric Power Corporation from April to June 1981 was used for the evaluation of the model. Four different day types were chosen to specify the nominal load profile model. With regard to this choice the state variables for the nominal load are defined as follows:

$X_1(t)$ = the reference nominal base load for hour t

$X_2(t)$ = the Saturday nominal load increment from $X_1(t)$ at hour t

$X_3(t)$ = the Sunday nominal load increment from $X_1(t)$ at hour t

$X_4(t)$ = the Monday nominal load increment from $X_1(t)$ at hour t

Other parameters chosen for the nominal load profile model are $P_{xx}(0 \mid 0) = 3.6 \times 10^7 I$, R_x=diag (500 40 80 80) and the initial X is the zero 0 with dimension 4x24. (NOTE : unit 0.1 MVA) As a means of compensating for model errors adaptive noise estimation is performed as Eq. (24).[4),5)]

$$R_y = \frac{1}{48} \left(\sum_{t=1}^{24} r_1^2(t) + \sum_{t=1}^{24} r_2^2(t) \right) \qquad (24)$$

where $r_1(t)$ = yesterday residuals $r(t)$

$r_2(t)$ = residuals $r(t)$ one week ago

Results of the evaluation are tabulated in Table 1. The rms forecast error as a percentage of peak load are plotted in Figure 3 for the data from April June 1981.

No attempt was made to alter the data or distinguish holidays from the day of the week on

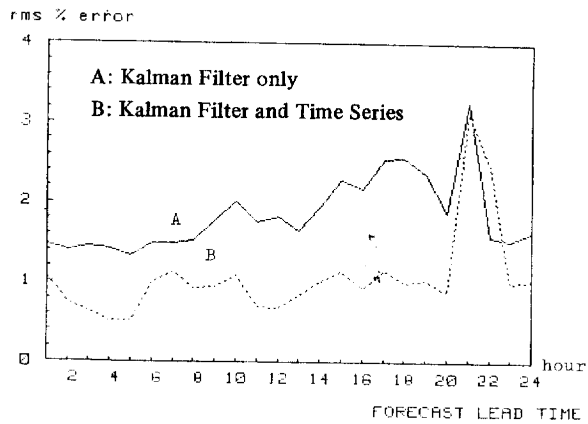

Fig. 3. Rms error as a percentage of peak load.

Table 1. Simulation result.

lead time	1	2	3	4	5	6	7	8	9	10	11	12
A	1.46	1.40	1.45	1.41	1.33	1.49	1.49	1.53	1.77	2.01	1.75	1.82
B	1.27	0.75	0.63	0.52	0.51	0.99	1.14	0.92	0.93	1.08	0.70	0.67
lead time	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
A	1.65	1.95	2.28	2.16	2.56	2.57	2.36	1.86	3.26	1.57	1.52	1.62
B	0.83	1.00	1.12	0.83	1.14	0.99	1.04	0.88	3.14	2.46	0.89	1.02

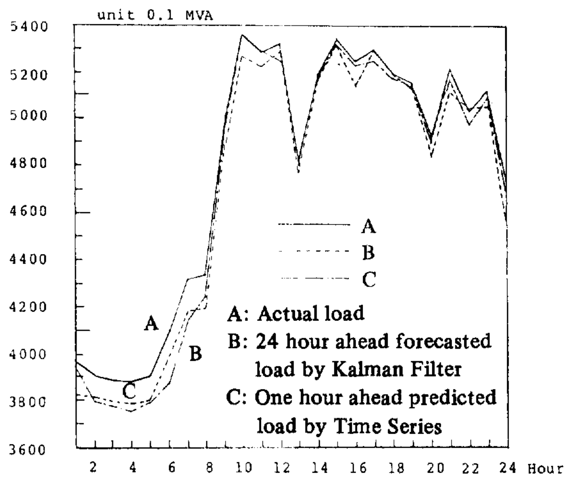

Fig. 4. The comparison of simulation results.

which they fell. All the Chi-square tests for adequacy of residual model of Time Series were satisfied when ARMA (1,1) model was adopted. As an example of result one typical load is plotted in Figure 4 with forecasted loads by Kalman Filter and one hour ahead predicted load by Time Series. In this case study 1.2% rms error was acquired.

5. Conclusions

The general algorithm presented permits the modeling of various day types, and this algorithm can be easily and efficiently implemented. Its on-line adaptive nature makes it highly effective in control center environment.

In comparison with BGLS suggested by Irisarri[1)] the fact that BGLS needs a number of data to identify its parameter is adisadvantage, and, more over, the appropriate order of resulting residual sequence and exponential decay factors δ and ρ should be found by trial and error. In the algorithm presented such an effort is not required but it only is necessary to decide statistics of noise.

And computer simulation results with the varied noise statistics have shown no significant discrepancy in forecasting among them.

Reference

1) G.D. Irisarri; "On-line load forecasting for energy control center application," IEEE PAS, vol. 101, Jan. 1982, pp. 71-78.
2) H.P. Van Meeteren; "Short-term load prediction with a combination of different model," 1979, Power Industry Applications Conference.
3) T. Bohlin; "Four cases of identification of changing systems," in Mehrea, Lainoiotis, System Identification, Academic Press, 1976.
4) A.H. Jazwinski; "Stocastic process and filtering theory," Academic Press, 1970.
5) B.D.O. Anderson, J.R. Moore; "Optimal filtering," Prentice-Hall, 1977.
6) G.E.P. Box, G.M. Jenkins; "Time seies analysis for forecasting and control," Holden-Day, 1970.
8) S.J. Park, C.W. Ahn; "Box-jenkins time seies analysis," KAIST Software Guidance Series No. 1, Korea, 1981.
8) G.E.P. Box, David A. Pierce; "Distriubtion of residual autocorrelations in autoregressive integrated moving average time series models," Journal of the American Statistical Association, vol. 65 (Dec. 1970).
9) K.L.S. Sharma, B.E., M. Tech; "Recursive short-term load forecasting algorithm," PROC. IEE, vol. 121, No. 1, Jan. 1974.
10) Peter S. Maybeck; "Stochastic models, estimation, and control volume 1," Academic Press, 1979.
11) N.R. Draper, H. Smith; Applied rregression analysis," John Wiley & Sons, 1966.
12) D.C. Montgomery, L.A. Johnson; "Forecasting and time series analysis," McGraw-Hill, 1976.