

SOME DISCUSSIONS ON POWER LOAD PREDICTION BASED ON FARMER'S METHOD

MORIKI TOYAMA and KEIGO YAMADA

Mitsubishi Electric Corporation

(Received February 6, 1967)

1. Introduction

The prediction of electric load takes an important part in determining operation schedules of power stations and sub-stations as well as economical dispatch and load frequency control. As the techniques for this purpose they are roughly classified into two kinds. One is to estimate statistically by means of past load data, the other is to estimate functionally by seeking any functional relation between electric load and measurable main factors which seem to be influential to the electric load. The Exponential Smoothing Method [1] [2] and the Farmer's Method [3] which will be described in this article belong to the former and a Regression Model Method [4] belongs to the latter. They are to be employed according to the objects, for example, the former is better to predict the range from several minutes to several hours but the latter seems more appropriate to predict the peak load of the next day.

In this paper we are trying to report what we contemplate on prediction errors by means of experiments with actual data and on some problems which arise in its application. The matters we discuss here are as follows:

(1) The relation between prediction errors and number of the characteristic mode functions and other parameters which the Farmer's Method uses.

(2) Automatic detection of the load pattern change which is necessary

to put the Farmer's Method to practical use, especially as a part of automatic power feeding systems.

(3) Applicable method to the load series which have a certain trend in order to put this method to practical use.

2. Farmer's Method

For later discussions the abstract of the Farmer's Method will be described below.

The Farmer's Method is essentially suitable for a short time prediction. In such a forecasting, it is assumed that electric load doesn't have a constructive change and its pattern is almost constant. Now daily electric load is represented with $x(t)$ ($0 \leq t \leq T$); where t means any daily time and T is its final time.

In the Farmer's Method it is assumed that such electric load $x(t)$ is represented by a linear combination of some characteristic mode functions.

The characteristic mode functions are introduced as follows:

Suppose that $x_m(t)$, $m=1, 2, \dots, M$, are M sample functions of the stochastic process of electric load and that it is required to define, in some sense, the characteristic modes of the process. A simple way of specifying the first mode is to seek a function $\phi_1(t)$, a scaling factor $\lambda_1^{1/2}$ and a set of coefficients a_{m_1} , such that $\lambda_1^{1/2} a_{m_1} \phi_1(t)$ approximates, in a least square sense, to the sample functions $x_m(t)$ over a specified time interval $(0, T)$. The m th error takes the form:

$$(1) \quad e_m(t) = x_m(t) - \lambda_1^{1/2} a_{m_1} \phi_1(t)$$

The mean square error averaged over time and over the group is:

$$(2) \quad E_1 = \frac{1}{MT} \sum_{m=1}^M \int_0^T \{x_m(t) - \lambda_1^{1/2} a_{m_1} \phi_1(t)\}^2 dt$$

The function $\phi_1(t)$ and the vector a_{m_1} may, without loss of generality, be normalized such that

$$(3) \quad \left\{ \begin{array}{l} \int_0^T \phi_1^2(t) dt = 1, \\ \sum_{m=1}^M a_m^2 = M. \end{array} \right.$$

The minimum conditions then take the form

$$(4) \quad \left\{ \begin{array}{l} a_m = \frac{1}{\lambda_1^{1/2}} \int_0^T \phi_1(t) x_m(t) dt, \\ \lambda_1^{1/2} \phi_1(t) = \frac{1}{M} \sum_{m=1}^M a_m x_m(t). \end{array} \right.$$

On the other hand, sample correlation function $R(t, \tau)$ is described with eqn (5).

$$(5) \quad R(t, \tau) = \frac{1}{M} \sum_{m=1}^M x_m(t) x_m(\tau).$$

From eqns (4) & (5), the following eigen function eqn (6) is introduced.

$$(6) \quad \int_0^T R(t, \tau) \phi_1(\tau) d\tau = \lambda_1 \phi_1(t) \\ 0 \leq t \leq T.$$

Combing eqns (2) and (4) gives for the minimum mean square error

$$(7) \quad E = \frac{1}{T} \left\{ \int_0^T R(t, \tau) dt - \lambda_1 \right\},$$

and it is necessary to take the greatest eigen value as λ_1 to minimize eqn (7).

Therefore, $\phi_1(t)$ is determined as the eigen function corresponding to the dominant eigen value λ_1 .

Furthermore the 2nd characteristic mode function $\phi_2(t)$ is calculated by treating $e_m(t)$ as sample functions in place of $x_m(t)$, and it is proved that this function $\phi_2(t)$ is corresponding to the 2nd greatest eigen value λ_2 of eqn (6).

The k th characteristic mode function is determined from eqn (6) in

the same manner.

Sample function takes the form:

$$(8) \quad x_m(t) = \lambda_1^{1/2} a_{m_1} \phi_1(t) \lambda_2^{1/2} a_{m_2} \phi_2(t) + \dots$$

When the series in eqn (8) is terminated after K terms, the least mean square error E_K is as follows:

$$(9) \quad E_K = \frac{1}{T} \left[\int_0^T R(t, t) dt - \sum_{i=1}^K \lambda_i \right]$$

Eqn (9) shows that it is possible to select K for the specified accuracy.

Now prediction mechanism is expressed as follows, using the characteristic mode functions which were described above:

$$(10) \quad \hat{x}(t) = \sum_{i=1}^K C_i \phi_i(t).$$

Where $\hat{x}(t)$ is estimated load for actual load $x(t)$

Here it is necessary to decide the value of $C_i (i=1, 2, \dots, K)$. A simple method is to choose them such that eqn (10) is the best means square approximation to past data.

Namely that is to minimize eqn (11) when load data $x(t)$ over a specified time interval $(0, T_0)$ are given. Then it is possible to estimate the load at a time ahead of T_0 with eqn (10).

$$(11) \quad I = \int_0^{T_0} \left\{ x(t) - \sum_{i=1}^K C_i \phi_i(t) \right\}^2 dt.$$

In reference (3) another method is described using Volterra series expansion. But that is omitted in this paper because no experiments about it are done.

Note: On the estimation of prediction

Eqn (9) gives the evaluation of prediction error as Farmer describes.

But it is apparent that this gives lower evaluation than the actual

one, due to the definition itself of E_k in eqn (9).

For example, assume the following process described by eqn (12):

$$(12) \quad \begin{aligned} x(t) &= \sin t + \varepsilon(t) \\ (0 \leq t \leq 2\pi) \end{aligned}$$

where

$\varepsilon(t)$ is a random variable and independent with respect to time and, average of $\varepsilon(t)$, $E[\varepsilon(t)] = 0$ variance of $\varepsilon(t)$, $\text{Var}[\varepsilon(t)] = \sigma^2$.

Then, mean square prediction error is σ^2 , but that of eqn (9) is $(2\pi - 1)\sigma^2/2\pi$ and lower than true one.

Therefore, it is better that the following eqn (13) is used to make more precise evaluation about mean square prediction error.

$$(13) \quad E_{K^1} = \frac{1}{T} \left[\int_0^T R(t, t) dt - \sum_{i=1}^K \lambda_i + \sum_{i=1}^K \int_0^T \int_0^T \phi_i(t) \phi_i(\tau) \cdot a(t, \tau) d\tau dt \right],$$

where

$$a(t, \tau) = R(t, \tau) - \bar{x}(t) \bar{x}(\tau)$$

and $\bar{x}(t)$ expresses the mean of sample functions at time t .

Eqn (13) is introduced as follows:

Assuming that the prediction mechanism is given by eqn (14):

$$(14) \quad \sum_{i=1}^K C_i \phi_i(t),$$

mean square prediction error at time t , $D(t)$ is defined by eqn (15).

$$(15) \quad D(t) = E \left[x(t) - \sum_{i=1}^K C_i \phi_i(t) \right]^2$$

Here, $C_i (i=1, 2, \dots, K)$ which minimize the following equation:

$$E \left[\int_0^T \left\{ x(t) - \sum_{i=1}^K C_i \phi_i(t) \right\}^2 dt \right] \text{ are adopted as } C_i \text{ in eqn (14)}$$

Then

$$(16) \quad C_i = \int_0^T \phi_i(t) E[x(t)] dt$$

$$(i=1, 2, \dots, K)$$

According to the fact that eigenfunctions $\phi_i(t)$ are orthonormal, eqn (15) is modified as follows:

$$D(t) = E[x^2(t)] - \sum_{i=1}^K \lambda_i \phi_i^2(t) + 2 \sum_{i=1}^K \phi_i(t) \int_0^T \phi_i(\tau) a(t, \tau) d\tau$$

$$- \sum_i^K \sum_j^K \phi_i(t) \phi_j(t) \int_0^T \phi_i(\tau) \int_0^T \phi_j(\mu) a(\tau, \mu) d\mu d\tau$$

where

$$a(t, \tau) = R(t, \tau) - E[x(t)]E[x(\tau)].$$

Therefore, the mean square prediction error averaged over time interval $(0, T)$ becomes as follows:

$$E_K^1 = \frac{1}{T} \int_0^T D(t) dt = \frac{1}{T} \left[\int_0^T R(t, t) dt \right. \\ \left. - \sum_{i=1}^K \lambda_i + \sum_{i=1}^K \int_0^T \int_0^T \phi_i(t) \phi_i(\tau) a(t, \tau) d\tau dt \right].$$

Eqn (13) obtained by substituting $\bar{a}(t, \tau)$, which is calculated from sample data, in place of $a(t, \tau)$ in eqn (17).

In case of eqn (12),

$$a(t, \tau) = \sigma^2 \delta(t - \tau)$$

where

$$\delta(t - \tau) = \begin{cases} 1 & t = \tau \\ 0 & t \neq \tau \end{cases}$$

and there is only one eigen value.

$$\lambda_1 = \pi + \sigma^2.$$

Therefore,

$$E_K^1 = \frac{1}{2\pi} [(\pi + 2\pi\sigma^2) - (\pi + \sigma^2) + \sigma^2].$$

3. Experiments Based on the Actual Date and Discussions on Them

Some discussions on experimental results with the actual data are described below by means of the Farmer's Method which was described in chapter 2. We would like point out several problems which are to be considered when this method is applied.

(a) How many samples are necessary to make the characteristic mode functions? In connection to this, too much samples are not proper as Farmer points out. This is because they may be affected by the past data of different pattern. According to Farmer's report twenty samples are employed, but our experience indicates that such large samples are apt to be affected by the past data of different pattern so that the prediction errors are inclined to increase. We have favorable results from the experiment with less number samples (about ten samples). In addition it is to be noted that weekdays have different patterns, respectively. For example, Saturday and Sunday are obviously different from the rest of the week and Monday is also somewhat different. In this experiment these Saturday, Sunday Monday and holiday are excluded from the data.

(b) What is the proper number of the characteristic mode functions, K ? According to the eqn (9), it seems that the more the number K , the better the results become. The experiments indicate, however, that the vicinity of the number from 4 to 5 is most appropriate from the favorable results. The details are to be discussed later in connection with (C).

(c) To what extent are the past data to be employed in case of determining C_i of eqn (10) by means of the least square method? Although C_i are to be determined by (11), the data of the previous day must be employed if T_0 is small. Therefore, C_i are determined by the following formula in this experiments. Using $x(t)$, ($T_0 - L \leq t < T_0$), C_i is determined so as to minimize eqn (18).

$$(18) \quad I = \int_{T_0-L}^{T_0} \left\{ x(t) - \sum_{i=0}^K C_i \phi_i(t) \right\}^2 dt.$$

In this case there arises a problem how to determine the length of the past data, L . If L is small, the so called degree of freedom becomes small so that reliability of C_i is reduced to a large extent because eqn (18) is generally used in discrete form. On the other hand, if L is large, the so called adaptability of the prediction goes worse.

Although the above three points are necessary to be discussed, the two of them, (b) and (c), will be examined in this article. Before entering into their details, we would like to show some of the experimental results to assure the general prediction accuracy in Fig. 1, 2 and 3. The data employed are daytime hourly mean load of a certain power company. We tried to predict the hourly load at one hour in advance. We em-

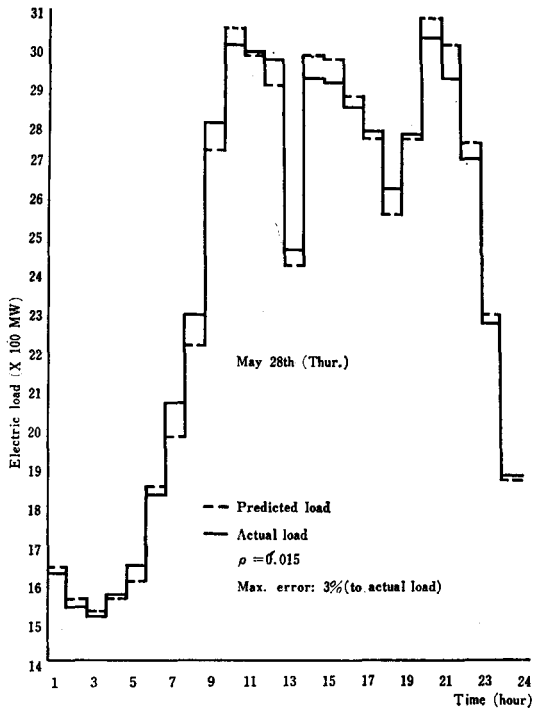


Fig. 1.

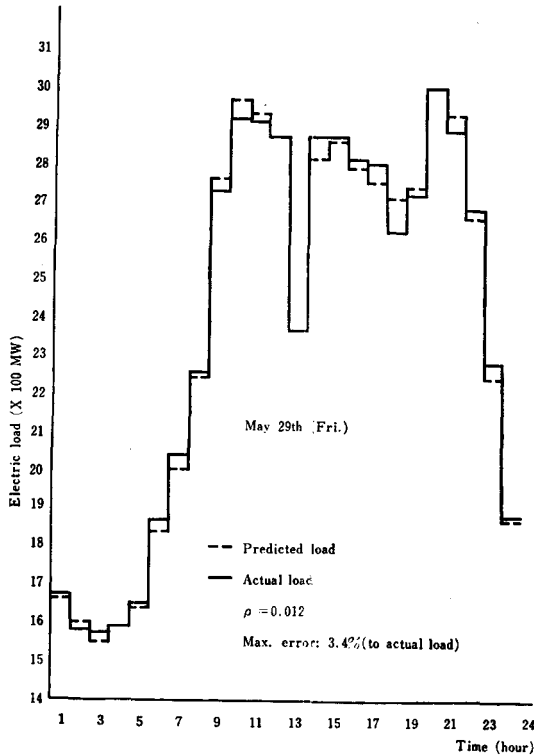


Fig. 2.

ployed 8 samples from May 14 th to May 27 th excluding Saturday, Sunday and Monday for the calculation of the characteristic mode functions. We do not use such data before May 14 th. This is because the pattern prior to May 14 th is, we think, different from that after the date and it will lower prediction accuracy.

ρ in the figures means coefficient of variation and is the standard deviation of the prediction error divided by the average of actual values. The number K of the characteristic mode functions, and L in (C) are 5 and 12, respectively.

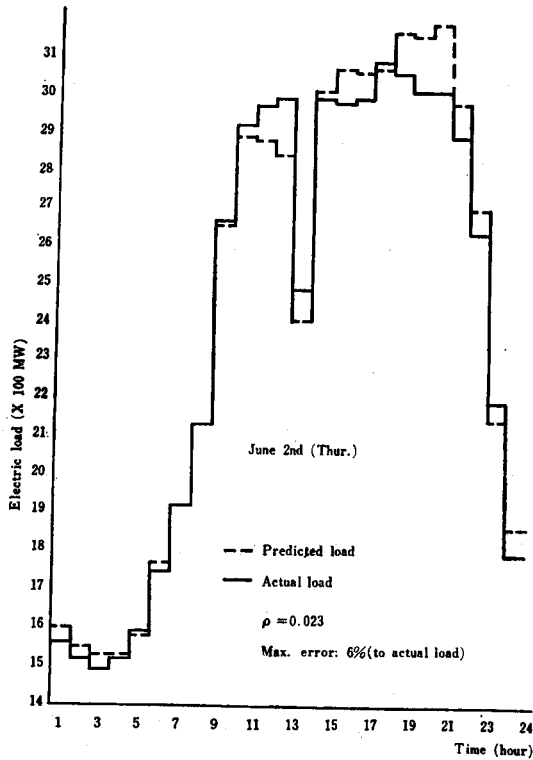


Fig. 3.

After carrying experiments on the samples of the same company for about 30 days, we could have obtained the similar accuracy.

Now, we would like to discuss the problems of (b) and (c). In other word, we will examine how the number K of the characteristic mode functions and the length L of (18) influence to the prediction accuracy. Firstly, setting K constant, we will examine the change of prediction accuracy due to L . Let us consider the variance of the prediction errors by means of the following fictitious model.

Assume that the load, $x(t)$ is realized as follows:

$$(19) \quad x(t) = \sum_{i=1}^K C_i \phi_i(t) + \varepsilon_t,$$

where ε_t is random variable and its mean and variance are zero and constant, respectively and;

$$\text{Var}(\varepsilon_t) = \sigma^2.$$

If the present time is T_0 and prediction for load at time t is carried by means of eqn (10) and C_i are determined by means of (18), then the variance $\text{Var}(e_t)$ of the prediction error e_t .

(e_t =(actual load—predicted value)) is shown as follows:

$$\text{Var}(e_t) = [1 + \phi'(t)V(T_0)\phi(t)] \cdot \sigma^2$$

where

$$\begin{aligned} \phi(t) &= (\phi_1(t), \phi_2(t), \dots, \phi_K(t)) \\ V(T_0) &= (X' X)^{-1} \\ X &= \begin{pmatrix} \phi_1(T_0-L+1) & \dots & \phi_K(T_0-L+1) \\ \phi_1(T_0-L+2) & \dots & \phi_K(T_0-L+2) \\ \phi_1(T_0) & \dots & \phi_K(T_0) \end{pmatrix} \end{aligned}$$

where prime (') means the transposed matrix.

Table 1 illustrates the calculations of ν and $\nu^{1/2}$ by applying various L , using $\phi_i(t)$ which are employed in case of the predictions which are shown in Fig. 1~3.

$$\nu = (1 + \phi'(T_0+1)V(T_0)\phi(T_0+1)).$$

Table 1. The values of ν and $\sqrt{\nu}$ for various L 's.

L	6	8	10	12	15	20	25	30	48	70
ν	7.66	3.71	2.95	2.74	2.47	2.46	1.54	1.54	1.54	1.54
$\sqrt{\nu}$	2.77	1.93	1.72	1.65	1.57	1.57	1.24	1.24	1.24	1.24

(Note) $T_0=19$.

As shown in the above, it is found that the variance of the prediction errors becomes larger as L become smaller and it becomes smaller as L

becomes larger. This is, however, based upon the assumption that $X(t)$ depends upon the model of the eqn (19). In fact, the value, C_i changes daily, therefore, the prediction accuracy becomes worse if L is too large. This is illustrated in Fig. 4.

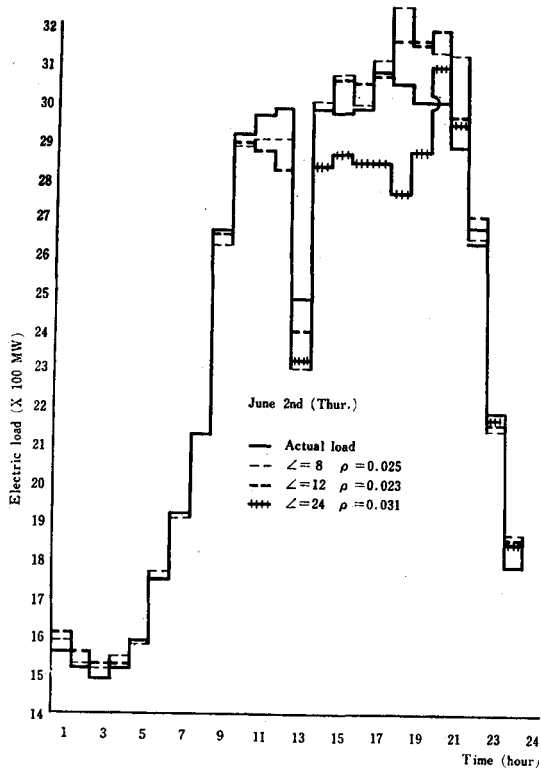


Fig. 4.

Fig. 4 shows that in case different pattern from the ordinarily one has occurred, the prediction loses adaptability when L is large. In this figure it is found that the most favorable result is obtained from $L=12$.

Next, we will examine by experiments what is appropriate for L against other K 's. A part of results is shown in Table 2.

Table 2. Comparison of prediction accuracies (ρ 's) for (K, L).

$L \backslash K$	1	3	5	10	15
6	0.023				
8	0.021	0.019	0.019		
12	0.022	0.021	0.018		
15				0.023	
20				0.029	0.031
24			0.028		0.033

(Note) Blanks show no experiment.

In this the load series are three days, May 29 th, June 1 st and 2 nd and ρ is calculated for the entire period of prediction. As noted in this, $K=5$ and $L=12$, are most favorable. The reason why the results become worse with increased K may be that L is necessary to become large in turn so that adaptability is lowered. In another examples, too, $K=5$ or its vicinity afford such a good result.

In the foregoing our experiments show that $K=4\sim 5$ and $L=12\sim 15$ are optimum. Furthermore, it is verified that the Farmer's Method falls into within allowable errors against the load series.

In general, the most favorable value of K and L depends upon the load series and it may be decided theoretically if a proper model can be constructed for load. But in practice, it will be realistic to decide them by means of the experiments as described above.

4. Discussions to put this Method the Practical use (1)

— Automatic check of load pattern change —

As described above, the Farmer's Method uses the characteristic mode function to predict electric load.

It is, therefore, necessary that load pattern at the time when the prediction is done is not so much different from that of the interval during which characteristic mode functions are produced.

If that pattern changes, it is required to recalculate characteristic

mode functions. It should be noticed here that the change of a load pattern doesn't mean one or two unusual daily-patterns.

Then it is difficult and complicated for a system operator to find the change of daily load pattern because data are random.

Therefore, it is important to detect the change automatically. Especially in the case that this prediction method is used in the on-line control system, automatic detection of pattern change is more necessary. One method to detect the pattern change automatically is described below.

Sample average value of $x_m(t)$ at the time t , $\bar{x}(t)$, is expressed as eqn (20).

$$(20) \quad \bar{x}(t) = \frac{1}{M} \sum_{m=1}^M x_m(t),$$

where $x_m(t)$ ($m=1, 2, \dots, M$) are sample functions which are used to calculate the characteristic mode functions.

Let $\bar{x}(t)$ be normalized so that an integral summation of normalized value over interval $(0, T)$ becomes a constant A . And let $y(t)$ be this normalized value, then

$$(21) \quad y(t) = \left(A / \int_0^T x(t) dt \right) x(t).$$

Now, when new daily load data $x(t)$, ($0 \leq t \leq T$) are given they are normalized as in eqn (22)

$$(22) \quad \bar{x}(t) = \left(A / \int_0^T x(t) dt \right) x(t) \\ (0 \leq t \leq T).$$

Then square of the distance between $y(t)$ and $\bar{x}(t)$, D , is defined in eqn (23) and is considered as a index of pattern difference.

$$(23) \quad D = \int_0^T (y(t) - \bar{x}(t))^2 dt.$$

If D is large, it would be considered that $x(t)$ is different from $y(t)$. Namely each pattern is different from another. On the contrary it can

be considered that each pattern is same, if D is small.

But, using the above D to check pattern change, it is considered that there is a pattern change even if $x(t)$ is unusual only for one day.

Therefore, to avoid such phenomena it should be considered to compare $y(t)$ with a normalized value which is obtained from weighted averaged value of past data involving $x(t)$.

According to the exponentially smoothing method, we define weighted average value $x'(t)$ as follows :

$$(24) \quad \tilde{x}(t) = \alpha x(t) + (1 - \alpha)x''(t).$$

α is a positive number and smaller than 1, and $x''(t)$ is a weighted average value obtained by eqn (24) using data of the previous day.

Furthermore, as initial value of $x''(t)$ in eqn (24), we use $\tilde{x}(t)$ in eqn (20)

Normalized form of $x'(t)$ is :

$$(25) \quad \tilde{x}'(t) = \left(A / \int_0^T x'(t) dt \right) x'(t) \quad (0 \leq t \leq T)$$

The new distance between $y(t)$ and $\tilde{x}'(t)$ is defined as follows :

$$(26) \quad D = \int_0^T \{y(t) - \tilde{x}'(t)\}^2 dt.$$

Then our method is as follows :

For some constant value w , we consider that the pattern changes if $D \geq w$ and it doesn't change if $D < w$.

There arises a problem to decide α and w . For this problem it is more practical to decide them experimentally. Namely, using the series data of load of which a pattern changes on the way, daily calculation of D in eqn (26) is carried. Considering D 's be time series, D must become large on the way if the pattern changes at that time.

But if α is very small, response to the pattern change is slow, and D would change fairly later than the actual pattern changes. If α is large, D would response quickly to the unusual load pattern as described above.

As an example, a result of experiments which are made using load data during April and May is shown in Table 3. It must be noticed that the peak load time changes from the middle of April due to the length of day time.

From the above table, the following fact is apparent.

In the case of $\alpha=0.2$ there is a large change of D on April 9 because

Table 3. Sequencies of D values α 's.

date α	8/4	9/4	10/4	14/4	15/4	16/4	17/4	21/4	22/4	23/4
0.05	0	0	0	0	1	1	1	1	1	1
0.1	0	1	1	1	2	2	2	2	2	2
0.2	1	4	3	3	6	3	4	4	5	6

date α	24/4	28/4	30/4	7/5	8/5	12/5	13/5	14/5	15/5
0.05	1	1	1	2	2	2	3	3	3
0.1	3	3	4	5	5	5	7	8	9
0.2	7	5	7	10	10	9	13	18	14

(Note 1) Sat., Sun., Mon. and holidays are excluded.

(Note 2) $A=24$. Figures show $D \times 1000$.

this day has an unusual pattern and it seems too sensitive.

Next in the case of $\alpha=0.05$, D begins to change on May 9th and this response is too slow.

In the case of $\alpha=0.1$, it is most suitable case and when w is set to 0.003, D becomes larger than w on the appropriate day.

Several other experiments showed that it is best when α and w are respectively set to about 0.1 and 0.003.

It is necessary to calculate the characteristic mode function again when $D > w$.

5. Discussions to put this Method to Practical use (2)

— Dealing with the case containing trend factor —

The case without trend factor is discussed above. But there may be a system whose load has trend factor even in a short term.

In such a case, it is not suitable to apply Farmer's Method directly to load prediction.

So it is necessary to modify the prediction system, considering trend factor.

A method is proposed, in which the Farmer's Method is involved.

Briefly speaking, trend prediction is made by the Exponentially Smoothing Method, and the Farmer's Method is applied to trend-eliminated data, and then load prediction is obtained by combing these two predicted values.

And so, we assume that the Farmer's Method can be applied to predict trend-eliminated part.

First we define load pattern factor which is calculated from sample functions $x_m(t)$, ($m=1, M$) which are used to seek characteristic mode functions. That procedure is as follows:

For convenience later, we treat $x_m(t)$ as a discrete function, and let x_{mn} denote $x_m(t)$ at $t=n$, and n ranges from 1 to N , here N is the number of sampling instants.

Now we estimate trend factor at n th time of m th day by the following moving average method:

$$\begin{aligned}
 m_{mn} &= \left\{ y_{t-\frac{N}{2}} + 2y_{t-\frac{N}{2}+1} + \dots + 2y_{t+\frac{N}{2}-1} + y_{t+\frac{N}{2}} \right\} \\
 (27) \quad & n = \frac{N}{2}, \frac{N}{2}+1, \dots, N \quad (m=1) \\
 & n = 1, 2, \dots, \frac{N}{2} \quad (m=M) \\
 & n = 1, 2, \dots, N \quad (m \neq 1, M)
 \end{aligned}$$

where

$$\begin{aligned}
 m_{mn} &: \text{estimated trend factor} \\
 y_{(m-1)N+n} &= x_{mn} \\
 t &= (m-1)N+n,
 \end{aligned}$$

here we assume N is even (e.g. $N=24$).

Next deviding x_{mn} by the estimated trend m_{mn} , we get s_{mn} . Namely

$$(28) \quad s_{mn} = x_{mn} / m_{mn}.$$

Using s_{mn} , load pattern factor, s_i are defined as follows:

$$(29) \quad s_i = \frac{1}{M-1} \sum_{j=2}^M s_{ji} \left(i=1, 2, \dots, \frac{N}{2} \right) = \frac{1}{M-1} \sum_{j=1}^{M-1} s_{jt} \left(i = \frac{N}{2} + 1, \frac{N}{2}, \dots, N \right)$$

Using load pattern factor $s_i (i=1, N)$ described above, prediction is carried as follows:

First, we get characteristic mode functions from data s_{mn} .

Let the calculated characteristic mode functions be

$$\phi_k(t), (k=1, 2, \dots). \text{ Here } t \text{ takes discrete value:} \\ t=1, 2, \dots, N$$

Next, prediction of trend factor is done. When data up to some time $\tau: y_\tau, y_{\tau-1}, \dots$ are given, trend factor is predicted by exponentially smoothing method. Namely, using estimated trend factor $m_{\tau-1}$ at time $(\tau-1)$ and the latest data y_τ , we estimate trend factor at time τ by the following eqn (30):

$$(30) \quad m_\tau = \alpha \left(\frac{y_\tau}{s_\tau} \right) + (1-\alpha)(m_{\tau-1} + R_{\tau-1}).$$

Where s_τ is load pattern factor at time τ and α is a positive number which is smaller than 1. $R_{\tau-1}$ is estimated value of trend difference at time $(\tau-1)$, and R_τ is obtained by eqn (31).

$$(31) \quad R_\tau = \beta(m_\tau - m_{\tau-1}) + (1-\beta)R_{\tau-1},$$

where β is a positive number which is smaller than 1. And predicted value $\hat{m}_{\tau+L}$ of trend factor at time $(\tau+L)$ is obtained as follows:

$$(32) \quad \hat{m}_{\tau+L} = m_\tau + L \cdot R_\tau.$$

While prediction of trend-eliminated part is done by the Farmer's Method using characteristic mode functions calculated above as follows:

We determine C_i by eqn (18) using data

$$(33) \quad x'(\mu) = y(\mu) / m\mu \\ (\mu = \tau, \tau - 1, \dots).$$

Thus prediction value $\hat{y}_{\tau+L}$ of trend-eliminated part at time $\tau+L$ is obtained by eqn (34)

$$(34) \quad \hat{y}_{\tau+L} = \sum_{i=1}^K C_i \phi_i(\tau+L).$$

Final prediction value $x_{\tau, L}$ of load at time $\tau+L$ is obtained by eqn (35)

$$(35) \quad x_{\tau, L} = \hat{m}_{\tau+L} \cdot \hat{y}_{\tau+L}.$$

It is important to select initial values of m_r and R_r and smoothing constant α and β optimally.

About initial values, we determined them as follows:

(i) Using sample data x_{mn} ($m=1, M; n=1, N$) with which characteristic mode functions are calculated, average load per day V_m for each m is calculated as follows:

$$(36) \quad v_m = \frac{1}{N} \sum_{n=1}^N x_{mn} (m=1, 2, \dots, M)$$

(ii) Initial value R° of R_r is:

$$R^\circ = (V_M - V_1) / (MN - N).$$

(iii) Initial value m° of m_r is:

$$m^\circ = V_1 + R^\circ (MN - N/2).$$

About smoothing constants, some methods can be thought, for example, gradient method, simulation and son on. In our experiments, when they are about 0.1, good results were obtained.

Although experiments should be done using actual data which contain trend factor, we could not get such data. So fictitious data were produced by adding various trend factor to data we used in the experiments in chapter 3 and some experiments were done using them. Prediction was possible with almost same error as in chapter 3. But, to test our method with actual data remains as our subject.

6. Conclusion

The matters discussed in this article can be summarized as follows:

(1) We showed that the Farmer's Method can predict within the allowable accuracy using actual data.

(2) We proposed an automatic calculation method which shows when the recalculation of characteristic mode functions is to be carried, and experiments showed its practical usefulness.

(3) A method for predicting load containing trend was proposed. It combines the Farmer's Method and the Exponentially Smoothing Method and experiments showed good results.

On the other hand there are still many problems to be resolved in order to make the Farmer's Method more practical. Some of them are:

(1) to improve prediction accuracy against unusual patterns. It is necessary to find out factors such as meteorological components which affect the electric load and make rapid correspondense to such unusual conditions.

(2) to dispose against the loss of data.

(3) to speed up the calculation by the method of least squares.

Lastly, we would like to express our hearty thanks to Mr. H. Ogawa, Central Research Institute of Electric Power Industry, Dr. H. Kimura, and Dr. J. Baba, Mitsubishi Electric Corporation who kindly extended assistance.

REFERENCES

- [1] R.G. Brown, "Smoothing, Forecasting and Prediction of Discrete Time Series", Prentice-Hall, 1962.
- [2] P.R. Winters, "Forecasting Sales by Exponentially Weighted Moving Averages" *Management Science*, 6, 3 (1960).
- [3] E.D. Farmer, "A Method of Prediction for Non-stationary Process and its Application to the Problem of Load Estimation," *Trans. I.F.A.C.*, (1963).
- [4] G.T. Heinemann *et al.*, "The Relationship between Summer Weather and Loads—A Regression Analysis," *PICA* 1965 (Session IV).
- [5] Reference [1], p. 238.