Asian Journal of Information Technology 5 (2) : 160-163, 2006

The Automatic Forming of Typical Daily Load Curve Based on the Electric Energy Measurement System

[1,2]Li Xiao-ming, [1]Yin Xiang-gen, [1]Zhang Lian-mei, [1]Sun Jing
Department of Electrical Engineering, Huazhon University of Science and Technology
[2]College of Electrical Engineering, Wuhan University, Wuhan 43

Abstract: In this study a method of automatically forming typical daily load curves based on the real time data is presented, which is got from the electric energy measurement system. We classify daily load data for nearly a month according to workday and festival or holiday and calculate them, and use regression model and estimation of least squares to decide its regression coefficient, later proceed to significance testing of regression result. So the typical daily load curve of workday and holiday or festival can be obtained This kind of typical daily load curve can be applied to dynamic daily load forecast of the next latest week and distribute and dispatch electric power. This method has been tested and proved by data of electric energy measuring and analyzing system in Xiangfan power supply bureau. The error from the result is smaller, and the result is proved to be reasonable and practicable.

Key words: Daily load curve, electric energy measurement, load forecast

INTRODUCTION

The conventional drawing of daily load curve is generally based on using man-made collected historical data, and the accuracy the objectivity and the reliability of these data themselves are not ensured, so the error is very big to draw daily load curve in common use by using of these data.

In this study a method of automatically forming typical daily load curves based on the real time data is presented. This method adopts the real time data of the electric energy measurement system. The data are collected at the master station by communication gathering workstation, which dials to register on the Web by several demodulators, through public telephone net connects indirectly to remote main user gathering implement, then gathers real time data in the connected electric energy meter from gathering implement. We add up all the powers of the same kind of main users from integer time 0~23 h. to form the this day's synthetic daily load curve, then take every day's synthetic daily load curve for a month as the sample of counting. We divide these samples into two different classes: one is the work-day and the other is the holiday or festival, and finally form the typical daily load curves of this kind of load by means of statistics and count.

ESTABLISH MODEL FOR SYSTEM

The category of daily load curve: Generally speaking, the characteristic of load about electric user from Monday to Friday is basically identical, with only a little random interference, and Saturday, Sunday and other significant festivals have great difference from other days, so the mode of daily load curve is fundamentally classified into two type: workday and festival or holiday.

Identify and eliminate false data: In the primitive load records, some random disturbance, such as the electromagnetic disturbance, communication breakdown results in bad data, so we must discern these data according to regularity that the data should be obeyed and logic relationship about these data, and clear them in the correspondent load array. In order to guarantee the data's integrity, we then substitute these data with an average value of the same type data of a forward day and a backward day at the integer time.

Establish regression model: Because the synthetic daily load characteristic of workday and festival or holiday is basically similar, so their typical daily load curve is similar to their corresponding synthetic daily load. Thus we can adopt single-variable linear regression model, such as formula (1) shows:

Corresponding Author: Yin Xiang-gen, Department of Electrical Engineering, Huazhon University of Science and Technology, Wuhan,430074,P.R.China

$$\begin{cases} y_{dt} = a_{0t} + a_{1t} x_{dt} + \varepsilon_t \\ \varepsilon_t \sim (0, \sigma^2) \end{cases} \qquad (1)$$

In the formula, ε_t is represent for random variable of every user's overall powers at *t* o'clock every day, and it obeys normal distribution of which the average value is zero and the variance is σ^2, and a_{0t}, a_{1t} are unknown parameters to be resolved, t= 0, 1, 2 **…** 22, 23. With regard to the workday, the argument x_{dt} is defined as the d workday of the t hour (t=0~23 h), which is representative for all days sorted in turn except holidays or special days, the value of x_{dt} takes one number of 1~20 or 1~22, the function y_{dt} is on behalf of the same kind of main user's power in some integer time and someday; with regard to the holiday or some special day, the argument x_{dt} is defined as the different holiday or special day of some integer time (0~23 h), which is representative for all days sorting in turn of holidays or special days, the value of x_{dt} takes one number of 1~8 or 1~10, the function y_{dt} is on behalf of the same kind of main user's power in some integer time and some holiday or special day. Every variable x_{dt} (i=1,2…d; j=1,2…t) is independent each other.

ESTIMATE MODEL OF UNKNOWN PARAMETER AND EXAMINE NOTABILITY OF REGRESSION RESULT

Here we adopt estimation of least squares to get regression coefficients: α_{0t}, α_{1t} at t o'clock, then by using F distribution to examine their notability, and obtain typical power forecast expression formulas of different day at *t* o'clock. These power forecast expression formulas make up this kind of main user's typical daily load curves. The specific algorithm refers to the following instance. This instance uses real time data collected from electric energy measuring and analyzing system in some power supply bureau. The typical daily load curves can form by using these data. We take all data of one month, 30days, in 2002 to analyze and calculate.

APPLICATIONS

Primitive data:

Parameter estimation with method of least squares

Calculate parameter estimation at t o'clock:

$$y_{dt} = \hat{\alpha}_{0t} + \hat{\alpha}_{1t} x_{dt} \qquad (2)$$

$$\hat{\alpha}_{1t} = l_{xyt} / l_{xxt} \qquad (3)$$

$$\hat{a}_{0t} = \bar{y}_t \, \hat{a}_{1t} \, \bar{x}_t$$

$$(4) \begin{cases} \bar{x}_t = \frac{1}{n}\sum_{d=1}^{n} x_{dt} \ldots\ldots\ldots\ldots\ldots\ldots(5) \\ \bar{y}_t = \frac{1}{n}\sum_{d=1}^{n} y_{dt} \ldots\ldots\ldots\ldots\ldots\ldots(6) \\ l_{xyt} = \sum_{d=1}^{n} x_{dt} y_{dt} - n\bar{x}_t \bar{y}_t \ldots\ldots\ldots(7) \\ l_{xxt} = \sum_{d=1}^{n} x_{dt}^2 - n\bar{x}_t^2 \ldots\ldots\ldots\ldots(8) \end{cases}$$

Table 1: Workday of one month: x_{dt}, correspondent user power: y_{dt}(unit: kw)

功率 (小时) (天数)	0	1	2	3	t	23
1	682.5	577.5	525	525	…	945
2	630	577.5	472.5	472.5	…	787.5
3	577.5	472.5	525	472.5	…	787.5
4	577.5	472.5	525	525	…	892.5
5	682.5	682.5	630	525	…	840
6	630	577.5	577.5	577.5	…	787.5
7	630	577.5	525	577.5	…	892.5
8	682.5	630	630	682.5	…	945
9	735	682.5	682.5	682.5	…	840
10	735	630	630	577.5	…	840
11	682.5	630	630	577.5	…	997.5
12	630	630	682.5	577.5	…	945
13	787.5	735	735	787.5	…	892.5
14	840	892.5	787.5	682.5	…	892.5
15	840	787.5	682.5	630	…	735
16	735	630	630	630	…	892.5
17	682.5	682.5	682.5	735	…	997.5
18	630	525	525	525	…	787.5
19	525	472.5	577.5	525	…	682.5
20	577.5	472.5	630	525	…	735
21	682.5	630	682.5	682.5	…	630
22	577.5	577.5	525	630	…	997.5

Table 2: Holidays of one month: x_{dt}, correspondent user power: y_{dt}(unit: kw)

功率 (小时) (天数)	0	1	2	3	t	23
1	735	630	787.5	630	…	840
2	630	630	630	630	…	840
3	787.5	735	630	735	…	997.5
4	840	840	735	787.5	…	840
5	787.5	735	682.5	682.5	…	945
6	840	787.5	682.5	787.5	…	682.5
7	840	682.5	682.5	525	…	997.5
8	892.5	735	682.5	577.5	…	787.5

Table 3: The results of workdays parameter estimation and calculation (22 days)

parameter hour	$\bar{x}_t$	$\bar{y}_t$	l_{xxt}	l_{xyt}	$\hat{a}_{0t}$	$\hat{a}_{1t}$
0	11.5	670.6	885.5	280.7	667	0.32
1	11.5	615.7	885.5	1517.9	596	1.71
2	11.5	613.3	885.5	4277.6	557.7	4.83
3	11.5	596.6	885.5	4775.2	534.6	5.39
…	…	…	…	…	…	…
23	11.5	851.9	885.5	-2695.7	886.9	-3.0

n is represent for the number of workday or festival or holiday.

Calculate parameter estimation of 0~23 h: Put the data in Table 1 and 2 into formulas (5)~(8) and calculate, by

Table 4: The result of holiday parameter estimation and calculation(8days):

parameter hour	$\bar{x}_t$	$\bar{y}_t$	l_{XXt}	l_{xyt}	$\hat{a}_{0t}$	$\hat{a}_{1t}$
0	4. 5	713.5	42	1812	521.2	43.1
1	4. 5	662.8	42	2127	434.9	50.6
2	4.5	649.7	42	866	556.9	20.6
3	4.5	649.7	42	1916	444.4	45.6
…	…	…	…	…	…	...
23	4.5	833.4	42	290	802.3	6.9

formulas (3), (4), later get user load power forecast expression formulas of the *d* day at *t* o'clock by $\bar{y}_t$ formula (2), the calculation result is shown in Table 3 and 4.

Put data in Table 3 and 4 into Eq. 2, then we can obtain y_t the load forecast value of 0~23 h of the d day.

The typical daily load forecast value of the d workday is decided by the following forecast formulas:

$y_{d0} = 667+0.32 x_{d0}$ ¦ $y_{d1}=596$£«$1.71 x_{d1}$ ¦

$y_{d2} =557.7+4.83 x_{d2}$ ¦ $y_{d3} =534.6+5.39 x_{d3}$, …,

$y_{d23}=886.9-3 x_{d2}$ …. ¦ 9 ¦

$x_{d0} = x_{d1}= x_{d2}=\ldots = x_{d23}=1,2,\ldots,22$ (workday)

The typical daily load forecast value of the d holiday is decided by the following forecast formulas:

$y_{d0}=521.2+43.1 x_{d0}$ ¦ $y_{d1}=434.9+50.6 x_{d1}$ ¦

$y_{d2}= 556.9+20.6 x_{d2}$ ¦ $y_{d3}=444.4+45.6 x_{d3}$ ¦ …,

y_{d23} ¦½$802.3+6.9 x_{d4}$ …. ¦ 10 ¦

$x_{d0} = x_{d1}= x_{d2}=\ldots = x_{d23}=1,2\ldots 8$ (holiday)

Through those two groups of power forecast formulas, we can put the value of ✷ into them and acquire typical daily load curves of workday and holiday or festival.

Significance testing of regression effect

Significance testing of regression effect of the *t* o'clock:

Hypothesis: the level α is 0.05 and

H_0 ¦ $\hat{\alpha}_{1t} =0$ ¦ $H_1 \hat{\alpha}_{1t} \neq 0$

For workday ¦ $F_1 (1,n-2) = F_{0.95} (1,20)=4.35$

$S_{Rt} = \hat{a}_{1t}^2$ ¦ $lxx_t = S_{Tt} = lyy_t \sum_{d=1}^{22} y_{dt}^2 - n\, \bar{y}_t^2$

For holiday ¦ $F_{1-a} (1,n-2) = F_{0.95} (1,6)=5.99$

$S_{Rt} =\hat{\alpha}_{1t}^2 l_{xxt}$ ¦ $S_{Tt} = lyy_t = \sum_{d=1}^{8} y_{dt}^2 - n\, \bar{y}_t^2$

Significance testing of regression effect from 0~23 hours: For workday ¦ $F_0 = S_{Ro}*(n-2) / (S_{ro}- S_{T0})=0.45$

$F_1 =1.02$, $F_2=0.12$, $F_3=0.38,\ldots, F_{23}=2.12$

$F_0 \sim F_{23} < F_{0.95} (1,18)=4.35$

For holiday ¦ $F_0 =S_{Ro}*(n-2) / (S_{R0}-S_{To})=5.447$,

$F_1 =5.326$, $F_2=5.631$, $F_3=5.1, \ldots, F_3=5.336$

$F_0 \sim F_{23} < F_{0.95} (1,8)=5.99$

Hypothesis H_0 is true, so we make $\hat{\alpha}_{1t}=0$. According to that significance testing, the linear part of the known power forecast formula isn't notable, so we can ignore it. We can use formula (11) directly to forecast near power of the same type.

$$\hat{y}_{dt} = \hat{\alpha}_{0t} \quad (t=0,1,2,3,\ldots,23) \qquad (11)$$

Table 5: The table of comparing result(Y1represents regression value of least square estimation,Y2 represents mean value of load power ,Y3 represents real value)

Time	Y1	Y2	y3
0	667	670.6	735
1	596	615.7	735
2	557.7	613.3	630
3	534.6	596.6	735
4	692.4	622.8	735
5	737.6	682.5	787.5
6	922	899.7	1260
7	1215	1341	1523
8	1678	1527.3	1942.5
9	1696	1658.5	1732.5
10	1789.3	1677.6	1890
11	1832	1734.9	1942.5
12	1555	1458.1	1575
13	1562	1410.3	1575
14	1417	1479.5	1575
15	1788	1429.4	1785
16	1665	1441.4	1575
17	1580	1513	1785
18	1818.2	1527.3	1942.5
19	1689	1684	1680
20	1588	1577	1628
21	1447	1420	1470
22	1156	1138	1103
23	886.9	851.9	840

Comparing power forecast between least square estimation and the maximum of greatest similar estimation : The forecast power formulas of least square estimation and the maximum similar estimation are as follows (11), (12):

$$Y1= \hat{y}_1 = \hat{\alpha}_{0t}, \quad t=0,1,2,\ldots 23 \text{¡–¡–¡–¡–¡–¡–£¨11£©}$$

$$Y2= \hat{y}_2 = \bar{y}_t, \quad t=0,1,2,\ldots 23 \text{¡–¡–¡–¡–¡–¡–£¨12£©}$$

In order to analyze the accuracy of those two methods, we take $\hat{\alpha}_{0t}$ and $\bar{y}_t$ obtained by the known data from February 21to March 24 to forecast the daily load on March 25, then compare them with the real data y3 on March 25, and observe error square sum of these two methods from 0~23 h, the comparing result is as Table 5.

The error square sum of taking $Y_{dt}= \bar{y}_t$ to express daily load of March 25 ($nS^2= \sum_{d=0}^{23}(y_t - \bar{y}_t)^2$ ¦ y_t is the real value of March 25 ¦ is the mean value of every hour) is1950685.8, but the error square sum of taking $y_{dt}=\hat{\alpha}_{0t}$($nS^2 = \sum_{t=0}^{23}(y_t - \hat{a}_{0t})^2$ ¦ y_t is the real value of March 25) is1743807.18. According to those two data, we know that the error square sum led by using formula $y_{dt}=$ is much smaller than using formula $y_{dt}=\hat{\alpha}_{0t}$. According to least square estimation principle, we can also get that using $y_{dt}=$ can make error square sum least, but $y_{dt}=\hat{\alpha}_{0t}$(using load mean value) can't be restricted by making error

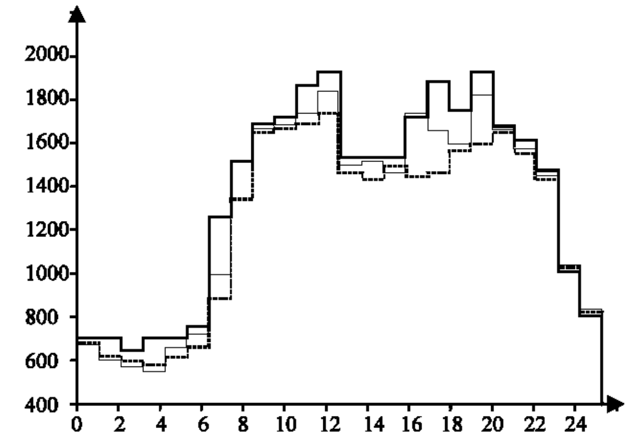

Fig. 1: Typical daily load curve and load forecist

square sum least. So it is more accuracy using $y_{dt}=\hat{\alpha}_{0t}$ than using $y_{dt}=\bar{y}_t$.

Deciding confidence interval:
α=0.05, workday| n=22£»holiday|n=8
workday | confidence interval is

$[\hat{\alpha}_{0t}-S_t/\sqrt{n-1}, \hat{\alpha}_{0t}+t_{1-\frac{\alpha}{2}}(n-1)\,S_t/\sqrt{n-1}]$ =[649.2,761.6], [604.1,783.1], [604.8,671], [589.4,765.4], ..., [814.6,915.09]
holiday| confidence interval is
[726.2,862], [661.2,782.6], [645.8,732.4], [589.3,749.5], ...,[776.25,956.25] |

Typical daily load curve and load forecast: Because typical daily load curve consists of mathematics characteristic of statistics law, we can make exterior extent appropriate, and forecast the just day's daily load curve with the near forward typical daily load curve of the same kind as shown in Fig.1. In Fig.1 forecast daily load on March 25. That is: express y_{dt} with $\hat{\alpha}_{0t}$(t=0,1,2,3…,23) to get the typical daily load curve on March 25. The wide solid line represents the real load value on March 25, the thin solid line represents forecast value with $y_{dt}=\hat{\alpha}_{0t}$, the dotted line represents forecast value with

$$y_{dt}=\bar{y}_t$$

Unit kw, horizontal axis represents hours, vertical axis represents load value.

CONCLUSIONS

The automatic forming typical daily load curve based on real time data obtained by the electric energy measurement system has been applied in Xiangfan power supply bureau. The error between the real measured data and the data got by calculating from typical daily load curve forecast formulas is small, and the result is acceptable, the actual application shows that this kind of method is practicable and reasonable. The method can be used for dispatch automation and distribution management in power systems.

REFERENCES

1. Yonghua, Z., T. Shucai and S. Yunyu, 2000. application statistics, Wuhan university publisher.
2. Dongxiao, N. and Z. 2000. Lei, electric load forecast technique and its application, China electricity publisher.
3. Xuheng, T., 1996. superficial speaking about the method of drawing two typical daily load curves about village load plan of water and electricity basic electrification, village electrification.
5. Murata, J. K. Kimura, S. Sagara and T. Kamakura, 1991. Daily load curve forecasting via automatic selection of models, IEE Conference Publication n 336 Publ by IEE 272: 0537-9987.
6. Cheng, Xu,Kang, Chongqing, Xia, Qing ,Shen, Yu , 2000. Integrated Model of Short Term load Forecasting,Dianli XitongZidonghue/Automation of Electric Power Systems 24 9 2000 Autom Electr Power Syst Press, 44: 1000-1026.