UNIVERSITI TEKNIKAL MALAYSIA MELAKA

MODELLING DAILY LOAD PROFILES IN THE UTILITY OF MALAYSIA

ARFAH BINTI AHMAD
SUZIANA BINTI AHMAD
INTAN AZMIRA BINTI WAN ABDUL RAZAK
NUR ILYANA BINTI ANWAR APANDI

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

MODELLING DAILY LOAD PROFILES IN THE UTILITY OF MALAYSIA

ARFAH BINTI AHMAD
SUZIANA BINTI AHMAD
INTAN AZMIRA BINTI WAN ABDUL RAZAK
NUR ILYANA BINTI ANWAR APANDI

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

Modelling Daily Load Profiles in the Utility of Malaysia

Arfah Ahmad[1], Suziana Ahmad[2], Intan Azmira Wan Abdul Razak[3] , Nur Ilyana Anwar Apandi[4]

[1,3,4]Faculty of Electrical Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.
[2]Faculty of Engineering Technology, Universiti Teknikal Malaysia Melaka (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.
arfah@utem.edu.my[1], suziana@utem.edu.my[2] ,intan.azmira@utem.edu.my[3] , ilyana@utem.edu.my[4]

Abstract: This paper presents a statistical framework for the modelling of daily load profiles. Data of short term load profiles in the utility of Malaysia with the reading recorded every half an hour was used for this study. The daily plot and analysis of average maximum demand and minimum demand was done in order to choose the best statistical distribution to be fit this data. Analysis shows that the average maximum load, minimum load and average load consumption are vary for every seven days. The frequency distribution plots of the data reveal different pattern for different days. Based on these, the normal distribution, log-normal distribution and Weibull distribution was fit to the data. Maximum likelihood method was used to estimate parameters in these distributions. The assessments of the best distribution are by using numerical method, namely the mean square error (MSE) and normal absolute error (NAE). Result shows that the load consumption for each seven day can be model using different statistical distributions. Hence results from this study can be used to forecast the daily load consumption using confidence interval method.

Keywords: load profiling, statistical distribution, maximum likelihood estimation, goodness of fit.

1. Introduction

Nowadays, most country around the world generated their electrical energy by burning huge fossil fuels that will lead to the end of source for a number of years ahead. In order to overcome the source shortage, a lot of countries search for other source such as by using wind power, wave power, hydro power, solar power, geothermal energy and also nuclear energy. Thus the high cost of electrical energy generation requires an accurate forecasting of energy usage by the customer. As discussed in [1], short term load forecasting has long been an issue of major interest for electricity industry. Traditionally, hourly forecast with a lead time of between one hour and seven days are required for the scheduling and control of power systems.

Generally the purpose of load profiling data is to produce load profile and it is derived when load curve is produced for a group of consumers which have common characteristics over a given period. Load profiling can be used to monitor load shape of the customers which can be operating based on past or present day data. It could act as reference to derived data or can also be refer to demand and consumption data types such as regression and profile coefficients. The common thing with all these data types is they presents the pattern of electricity usage of a segment on supply market customers. In distribution transformer loss of life and distribution transformer modelling, load profiles have been used as assessor and evaluator. Beside, load profiling is the best way for maintenance and to find out equipment demands especially to install the transformers traditionally and to maintain it in a good condition. Power utilities have been used load profiling for such a long time which operated to facilitate the formulation of retail tariffs and also to give all the needed information in forecasting, financial planning, rate design, system planning, site planning's demanding and regulatory reporting [2].

There are several factors that influencing the load profiling such as customer, time and other electrical load correlated to the load required. The customer factor due influencing load profiling as the electricity consumption is depend particularly by the number, the type and the size of customer electrical requirements. Apart of that, human and economic activity make the electric load varies with time as there are more loads produce during day time but less during the night. Further, the load also varies during weekdays and weekend [2]. As the production capacity is limited and the price of the incremental power to maximum load is sometimes very high, time factor is very important of Peninsular Malaysia power system.

Electric load forecasting can be divided into three categories that are short term load forecasting, medium term load forecasting and long term load forecasting. The short term load forecasting predicts the load demand in time interval from one day to several weeks. It can help to estimate load flows and to make decisions that can prevent overloading. The medium term load forecasting predicts the load demand from a month to several years. It provides information for power system planning and operations. The long term load forecasting predicts the load demand in time interval from a year up to twenty years. It is mainly for power system planning [3].

Several different methods and techniques of analyzing load demands have been developed with the focus on improving the prediction accuracy. The approach using statistical time series analysis and neural network (NN) are the most widely interest with rich research effort from all around the world [4]. In time series analysis, the approach by using ARMA models, ARIMA models, Box-Jenkin's methods and double seasonal ARIMA model have been applied as is [5-8]. On the other hand, studies in [9-11] present the analysis of short term load forecasting using NN models such as multilayer feed-forward, combined artificial NN and hybrid NN.

Interestingly, a study in [12] discusses the analysis and forecasting of short term load profile using statistical distribution approach. In this study, the author make used of normal and log-normal distribution to forecast customer load confidence interval in Finland. The selection of normal and log-normal distribution is due to the loads that distributed around the mean and in many cases; the distribution has a bell shape. While a study by G.W. Irwin et.al. [13] discussed about statistical electricity demand modeling based on customer billing data. The histograms of the data are highly-skew and in some cases are normally-skew, the model selected are exponential distribution, Rayleigh distribution and Weibull distribution. Result from this study shows that two parameters Weibull distribution can be employed to model demand patterns in the province of Northern Ireland [13].

This paper described the analysis of load profiling data in the utility of Malaysia. The analysis process begin by observe the trend of the load profiling; that are the weekly and daily trend. Based on the trend and descriptive statistics of the data, statistical distribution was fitted to the data. Parameter estimates for each distributions was determined by using the method of maximum likelihood estimation. The choosing of the best statistical distribution is based on the plot of histogram with density line and the smallest value of error distribution. Hence the best distribution can further be used for forecasting purposes.

2. Load Profiling Data

In the general context of daily time series of economic activity, the main features of load series have been extensively reported as having the trend, superimposed levels of seasonality, short-term dynamics, special days, nonlinear effects of meteorological variables, possible nonlinear time dependence, etc. [1]. The load profiling data for this study is based on the utility of Malaysia for a period of eight weeks. The reading was recorded in Megawatt and taken for every 30 minutes.

Table 1 shows the average maximum load, minimum load and average load consumption within a week. From Table 1, the average maximum load is different among the days with the highest demand in a day is about 10,000MW to 12,500MW. The lowest maximum load occurred in the weekend due to the off-day from works and school in Malaysia is on Saturday and Sunday. For the weekday load consumption, the average maximum load on Friday is slightly lower than Monday, Tuesday, Wednesday and Thursday. This is due to the long break on Friday, for Friday prayers and it is also an off-day for three states in Malaysia that are Kelantan, Terengganu and Kedah. The average minimum load ranges from 7,500MW to 8,500MW with the lowest minimum load consumption occurs in Sunday. The average minimum load on Monday is a bit lower compare to other days in the weekdays. Overall, the average daily load rages from 8,800MW to 10,700MW with the lowest average on the weekend while the highest average load on Wednesday. As the average maximum, minimum and average daily load for the days is different; hence it is suggested to fit the statistical distribution based on daily load data.

Table 1 Average maximum load, minimum load and average load consumption.

Day	Average Maximum Load (MW)	Average Minimum Load (MW)	Average Daily Load (MW)
Monday	12 032	7738.25	10110
Tuesday	12 456	8377.50	10568
Wednesday	12 491	8511.88	10659
Thursday	12 435	8509.88	10604
Friday	11 945	8470.63	10299
Saturday	10 767	8150.38	9540
Sunday	10 122	7576.00	8801

3. Distribution Functions

Load researches have common understanding that the electric load does not follow any known distribution function [12]. The frequency distribution plots of load data (Fig. 1 to Fig. 7) reveal that the pattern of load consumption is different for each day in the week. Based on the shape of the plots, three statistical distributions are suggested to be fit to the data. The statistical distribution chosen are Normal distribution, log-normal distribution and Weibull distribution. The distribution functions for all the test distribution is given as below.

For Normal distribution, the parameter $\overline{x}$ is the mean of the load profile data, s^2 is the variance and x is the load profile data. This distribution is chosen as the loads are distributed around the mean and in some cases the histogram has a bell shaped.

Normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi s^2}} \exp\left[-\frac{(x-\overline{x})^2}{2s^2} \right] \quad (1)$$

Log-normal distribution is a continuous distribution in which the natural logarithm transformation of the variables will results in normal distribution. The parameters involve in this distribution are the mean, x^2, the variance, s^2 of daily load consumption and the load data which represent by x. The form of log-normal distribution is given in equation (2).

Log-normal distribution

$$g(x)=\frac{1}{x\sqrt{2\pi s^2}}\exp\left[-\frac{[\ln(x)-\overline{x}]^2}{2s^2}\right] \qquad (2)$$

The Weibull distribution is characterized by two parameters that are the shape parameter, β and the scale parameter, α. In Weibull distribution, β also known as the slope that measures the rate behavior of load consumption with respect to time. Different value of β will affects the behavior of the distribution. Equation (3) below present the form of Weibull distribution with x is the load profiling data.

Weibull distribution

$$h(x)=\frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \qquad (3)$$

The estimation for the parameters is by using the method of maximum likelihood estimation. Let $x_1, x_2, \ldots, x_n$ be the load profiling data drawn from a selected probability density function with unknown parameter. The likelihood function of the data is the joint density of the random variables and it is a function of unknown parameter, θ. Thus the likelihood function is $L=\prod_{i=1}^{n} f_{x_i}(x_i,\theta)$. Taking the logarithm of the likelihood functions, differentiate with the respective parameters and equate to zero, the procedure will gives the estimations for each parameters in the distribution.

For normal distribution, the estimation of the mean parameter, and varians are given by $\hat{\mu}=\overline{x}$ and $\hat{\sigma}^2=\sum_{i=1}^{n}(x_i-\mu)^2/n$ respectively. For log-normal distribution, the estimation of the mean is given by $\hat{\mu}=\sum_{i=1}^{n}\ln(x_i)/n$ while the estimation for variance is

$$\hat{\sigma}^2=\frac{\sum_{i=1}^{n}\left[\ln x_i-\left(\sum_{i=1}^{n}\ln x_i/n\right)\right]}{n} \qquad (4)$$

For Weibull distribution, the maximum likelihood procedure with Newton Raphson iteration was use to get the estimation of the parameters. Equation (5) and (6) display the estimates of the parameters α and β.

$$\hat{\alpha}=\left[\left(\frac{1}{n}\right)\sum_{i=1}^{n}x_i^{\hat{\beta}}\right]^{\frac{1}{\hat{\beta}}} \qquad (5)$$

$$\hat{\beta}=\left[\left(\sum_{i=1}^{n}x_i^{\hat{\beta}}\ln x_i\right)\left(\sum_{i=1}^{n}x_i^{\hat{\beta}}\right)^{-1}-\frac{1}{n}\sum_{i=1}^{n}\ln x_i\right]^{-1} \qquad (6)$$

4. Goodness of Fit Test

To assess the suitability of these statistical distributions, three goodness of fit test use in this study are mean square error (MSE) and normal absolute error (NAE). The calculation of MSE and NAE is given as in equation (7) and (8) where P_i is referring to the predicted value while O_i is the observe value. A smaller value of MSE and NAE, the better fitted the distribution on load profiling data.

Mean square error

$$\sum_{i=1}^{n}(P_i-O_i)^2 \Big/ n-1 \qquad (7)$$

Normal absolute error

$$\sum_{i=1}^{n}Abs(P_i-O_i) \Big/ \sum_{i=1}^{n}O_i \qquad (8)$$

5. Results

The plots of frequency distribution of load profiling data for each day in the week are shown in Fig. 1 to Fig. 7. A study on these plots suggested that the plots for Monday, Tuesday, Wednesday and Thursday are skewed to the left. A unimodal plot is detected for Monday and Tuesday (Fig.1 and Fig.2) while the plot for Wednesday, Thursday and Saturday display a bimodal pattern (Fig.3, Fig.4 and Fig.6). In addition, the Friday and Sunday plots show a multimodal pattern. The different mode patterns on the plots are due to the different peak hours of load consumption for different days. Visually, it is apparent that the plots of the load consumptions for each day can be modeled by the selected three statistical distributions.

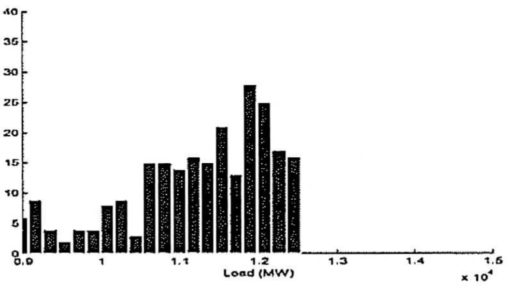

Fig. 1 Frequency distribution for Monday.

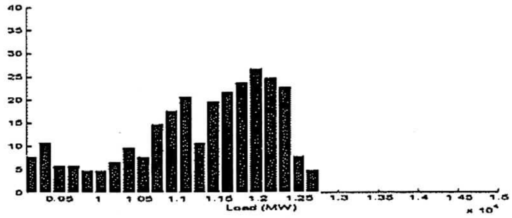

Fig. 2 Frequency distribution for Tuesday.

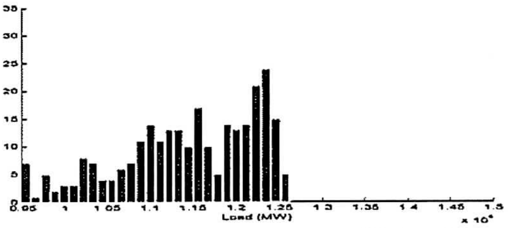

Fig. 3 Frequency distribution for Wednesday.

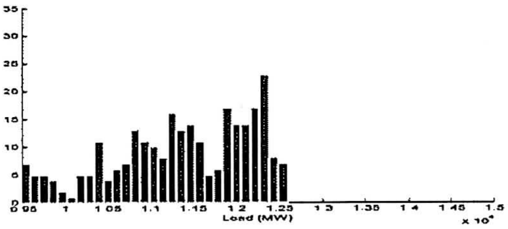

Fig. 4 Frequency distribution for Thursday.

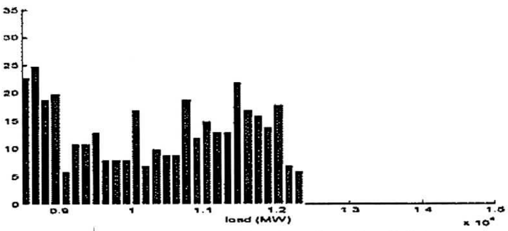

Fig. 5 Frequency distribution for Friday.

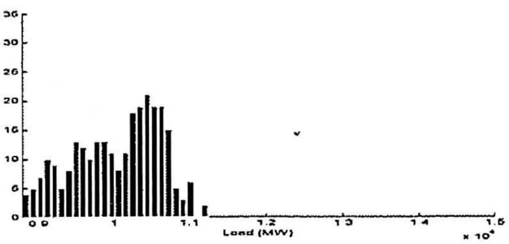

Fig. 6 Frequency distribution for Saturday.

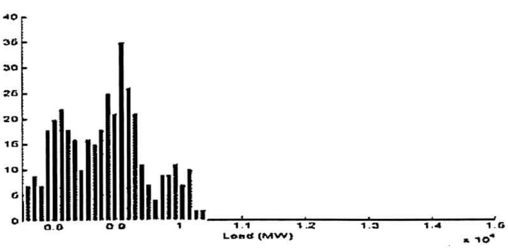

Fig. 7 Frequency distribution for Sunday.

Table 2 below shows the value of parameter estimates for normal, log-normal and Weibull distributions while equation (9), (10) and (11) present and example of fitted probability distribution function for Monday.

Table 2 Parameter estimates for selected distributions.

Day	Normal distribution		Log-normal distribution		Weibull distribution	
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\alpha}$	$\hat{\beta}$
Mon	10110	1652.1	9.208	0.167	6.953	10819
Tues	10568	1431.5	9.256	0.141	8.563	11183
Wed	10659	1373.0	9.266	0.132	8.927	11261
Thurs	10604	1345.8	9.261	0.130	9.068	11195
Friday	10299	1214.5	9.233	0.119	9.900	10829
Sat	9540.2	911.3	9.157	0.097	12.41	9938.8
Sun	8800.6	697.68	9.080	0.079	15.37	9101.1

Normal distribution function for Monday

$$f(x)=\frac{1}{\sqrt{2\pi(1652.1)^2}}\exp\left[-\frac{(x-10110)^2}{2(1652.1)^2}\right] \tag{9}$$

Log-normal distribution function for Monday

$$g(x)=\frac{1}{x\sqrt{2\pi(0.167)^2}}\exp\left[-\frac{[\ln(x)-9.208]^2}{2(0.167)^2}\right] \tag{10}$$

Weibull distribution function for Monday

$$h(x)=\frac{10819}{6.953}\left(\frac{x}{6.953}\right)^{10818}\exp\left[-\left(\frac{x}{6.953}\right)^{10819}\right] \tag{11}$$

Table 3 and Table 4 respectively presents the value of MSE and NAE for each distributions on each day of the week. From these value, the selection of best model is based on the smallest value of MSE and NAE. From the tables, it can be conclude that the best statistical distribution for Monday, Tuesday, Wednesday, Thursday and Saturday is Weibull distribution. Meanwhile the best distribution for Friday and Sunday is log-normal distribution.

Table 3 Mean square error (MSE)

Day/ MSE	Normal distribution	Log-normal distribution	Weibull distribution
Monday	10.05	11.20	4.56
Tuesday	6.54	7.30	4.41
Wednesday	10.16	7.22	4.46
Thursday	6.73	7.44	4.49
Friday	10.38	4.45	4.62
Saturday	10.54	5.07	4.56
Sunday	1.80	1.77	4.66

Table 4 Normal absolute error (NAE)

Day/ MSE	Normal distribution	Log-normal distribution	Weibull distribution
Monday	0.6225	0.6701	0.5004
Tuesday	0.5123	0.5441	0.4933
Wednesday	0.8831	0.5714	0.4965
Thursday	0.5634	0.5820	0.4979
Friday	0.8930	0.4772	0.5049
Saturday	0.8936	0.5177	0.5005
Sunday	0.3003	0.3000	0.5162

6. Discussion & Further Study

Electricity demand forecasting has significant impact on the costs and security of the energy supply. To produce a reliable energy system operation, accurate forecasting models are highly important. This paper presents statistical distributions model for Malaysia's load profiling data. Results from this study shows that the daily load profiling data in the utility of Malaysia can be modeled using different statistical distribution. Among the chosen distributions, the Weibull distribution emerged as the best distribution for 5 days of the week while log-normal distribution serve the best distribution for Friday and Sunday.

The analysis has shown that some of the plots display a unimodal and multimodal pattern. The existence of this pattern permits an examination to be made on the data sample. It is also suggested that other statistical distribution such as beta distribution, gamma distribution or three parameters Weibull distribution can be test to suit the pattern of load profiling data. It is also propose that another method of parameters estimates such as least square method can be employ to test the robustness of the estimated parameter.

This study just presents the best statistical model for daily load profiling data in the utility of Malaysia. As a proposal for future work, the result from this study can be used to estimate the electricity demand using the approach of confidence interval. The estimation method can be verified with the daily load measurement data and this will be looked for future research.

References

[1] Cancelo, J.R., Espala, A. and Grafe, R. Forecasting the electricity load from one day to one week ahead for Spanish system operator. *International Journal of Forecasting 24*, (2008), pp. 588-602.

[2] Gross, G. and Galiana, F.D. Short-term load forecasting. *Proceedings of the IEEE*, Volume 25, No. 12, (1987), pp. 1558-1573.

[3] Intan Azmira, W.A.R., Md. Shah, M., Hasimah, A.R. and Mohammad, Y.H. Short term load forecasting using data mining technique. *2nd IEEE International Conference on Power and Energy Conference*, (2008), pp.139-142.

[4] Jianguang Deng and Panida Jirutitijaroen. Short-term load forecasting using time series analysis: A case study for Singapore. *IEEE Conference on Cybernetics and Intelligent Systems*. (2010), pp. 231-236.

[5] Salim, N.A., Abdul Rahman, T.K., Jamaludin, M.F. and Musa, M.F. Case study of short term load forecasting for weekends. *Proceedings of IEEE Student Conference on Research and Development*, (2009), pp. 332-335.

[6] Fadhilah, A.R., Shitan, M., Amir, H.H. and Izham, Z.A. Forecasting the maximum demand of electricity in Malaysia. *International Conference on Energy and Environment*, (2006), pp. 6-11.

[7] Fadhilah, A.R., Shitan, M., Amir, H.H. and Izham, Z.A. Load forecasting using time series models. *Jurnal Kejuruteraan ke-21 UKM*, (2009), pp. 53-62.

[8] Norizan, M., Maizah, H.A. and Suhartono. Short term load forecasting using double seasonal ARIMA model. *Proceedings of the Regional Conference on Statistical Sciences*, (2010), pp. 57-73.

[9] Norizan, M., Maizah, H.A., Zuhaimy, I. and Suhartono. Shor term load forecasting using multilayer feed-forward (MLFF) neural network model. *UMT 11th International Annual Symposium on Sustainability Science and Management*. (2011), pp. 40-46.

[10] Subbaraj, P. and Rajasekaran, V. Short term hourly load forecasting using combined artificial neural networks. *International Conference on Computational Intelligence and Multimedia Applications*, (2007), pp. 155-163.

[11] Pang Qingle and Zhang Min. Very short-term load forecasting based on neural network and rough set. *International Conference on Intelligent Computation Technology and Automation*, (2010), pp. 1132-1135.

[12] Anssi Seppala. Statistical distribution of customer load profiles. *IEEE Catalogue*, (1995), pp.696-701.

[13] Irwin, G.W., Monteith, W. and Beattie, W.C. Statistical electricity demand modeling from consumer billing data. *IEEE Proceedings, Volume 133*, (1986), pp. 328-335.