Short-Term Load Forecasting for Campus Building with Small-Scale Loads by Types Using Artificial Neural Network

Sung-Jun Baek
Department of Electrical Engineering
Soongsil University
Seoul, Republic of Korea
sjbaek@soongsil.ac.kr

Sung-Guk Yoon
Department of Electrical Engineering
Soongsil University
Seoul, Republic of Korea
sgyoon@ssu.ac.kr

***Abstract*—Nowadays, the large portion of total energy is consumed by buildings, so it is essential to use the energy efficiency in buildings. With the deployment of a smart meter in the building, a small-scale load data are now available such as lighting, general and the other loads. We first analyze the small-scale loads through the fixed k-means clustering algorithm. Based on the analysis of load characteristics, we propose five artificial neural network-based forecasting models for campus buildings. The five models use total load, temperature and small-scale loads as input data with different input combinations. The case studies show that using small-scale loads improves the forecasting performance.**

Keywords: artificial neural network, campus building, small-scale load, smart meter, k-means clustering algorithm, short-term load forecasting.

I. Introduction

The energy consumption in buildings accounts for about 35% of total energy consumption [1]. The building sector represents 39% and 40% of the energy consumption in the U.S. and Europe [2]. In Korea, about 24% of the total energy consumption is used in buildings and the portion of electric energy used in buildings is 66.4% in 2015 [3]. Therefore, the electrical load forecasting for buildings becomes important to use the electrical energy efficiently.

Several studies have been investigated building load forecasting recently. Lim et al. have proposed a load forecasting model for campus building by using regional temperature data [4]. In [5], the accuracy of SVR and ANN based forecasting scheme has been analyzed in campus load. Ke et al. have analyzed several factors which highly affect the building load pattern [6]. In [7], load forecasting performances for the entire campus and one building on campus have been analyzed. In [8], the ANN trained by two different input combinations and compared performance of both cases.

When the load size becomes smaller, the forecasting performance is generally degraded compared to the performance of a large scale load [7] [9]. Therefore, to improve the forecasting accuracy, different approaches should be considered. Nowadays, many intelligent meters are installed in diverse sectors such as the substation, feeder and even inside a building. They can measure and send the power usage data in real time. Since the total load consists of the small-scale loads in a building, it is expected that small-scale data help to reveal unique characteristics of buildings and improve the forecasting accuracy.

In [10], to group small-scale loads, the k-means clustering algorithm has been used. The loads in the same cluster are summed and the load forecasting is performed in each cluster. In [11], the similar load and the other loads were called the regular and irregular nodes, respectively. Differently than [10] and [11], we use the fixed centroid k-means clustering algorithm to group small-scale loads based on pattern similarity with a total building load.

In this paper, we propose five artificial neural network (ANN)-based forecasting models for campus buildings. The proposed forecasting model uses total load, temperature and small-scale loads as input data with five different combinations. The small-scale load groups are obtained by the fixed centroid k-means clustering algorithm. Through simulations, it is confirmed that the small-scale data improve forecasting performance than the previous model which does not use the small-scale data.

The rest of the paper is organized as follows. We analyze the load data of campus building in Section II, and Section III describes the proposed load forecasting model. After evaluating the proposed model in Section IV, we conclude our paper in Section V.

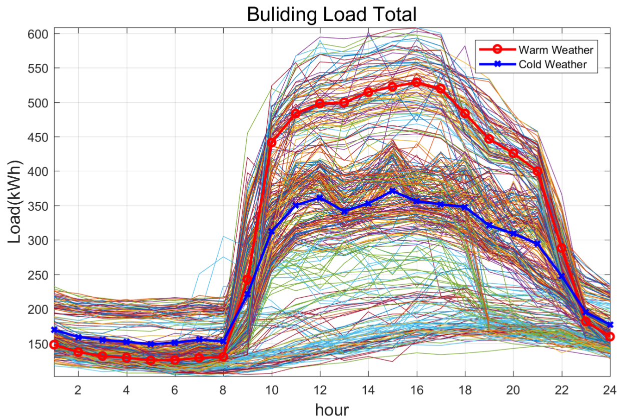
II. Analysis of Load Data for a Campus Building

We have been gathering load data of an engineering building in Soongsil University, Seoul, Korea since August 2015. This section analyzes the building load data.

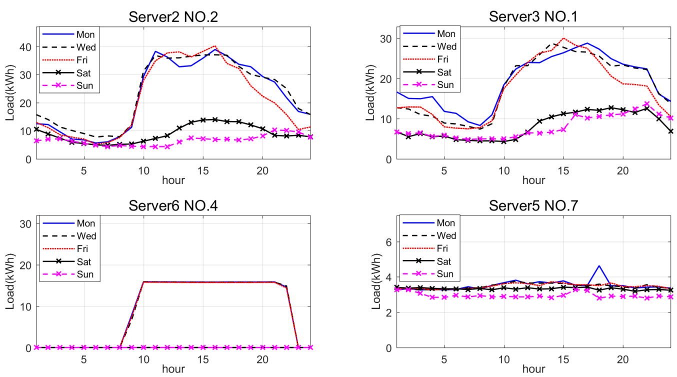
A. Configuration of the Campus Building Load

In the underground electrical room in the engineering building, smart meters have been installed to measure the electrical

This work was supported by Korea Electric Power Corporation (Grant number: R17XA05-62 and R18XA04).



(a) Total load of building from August 2015 to July 2016



(b) Small-scale loads in the building for one week

Fig. 1. Daily Load Curves for the Campus Building

loads which consist of different types of loads[1]. The small-scale loads consist of general loads (office or classrooms), lighting load, rectifier load, motor control sensors, ventilation load, pump load, and elevator load. There are several spare loads that are currently used as cafe and laboratories. The total number of the small-scale loads in the building is 56.

B. Characteristics of the Campus Load

Fig. 1(a) shows the total load of the building from August 2015 to July 2016. The thick red line with circles and the thick blue line with x show the average power consumptions for summer and winter, respectively. The daily load curves show different patterns depending on seasons. On the weekday, the load rapidly increases from 7:00 to 9:00, which is usually the beginning of the lectures. The peak value appears approximately from 14:00 to 16:00. After most of the lectures are over, i.e., at 6:00 pm, power consumption gradually decreases, and the load at night is small and stable.

Fig. 1(b) shows the load curves for small-scale loads in the building for one week. As shown in the figure SV2-2 and SV3-1 show a similar pattern to the total building load since they are general purpose load. On the other hand, SV6-4 and SV5-7 show a different pattern. The load shown in SV6-4 is only used on certain days, and that in SV5-7 shows a random pattern that cannot easily define its characteristic.

[1]In this paper, these loads are called small-scale loads.

C. Clustering Analysis of Campus Load Data

Since the patterns for small-scale loads in the building are very different, it is important to analyze load patterns for selecting important features from the small-scale load data. We use the k-means clustering algorithm which makes k groups, i.e., clusters, from the input data set [12]. It gathers close distant small-scale loads into the same cluster.

In this work, the input data is daily load curves of small-scale loads that are 24-hour load data. Since the absolute values of the total load and small-scale load are very different, each data is normalized before the k-means clustering algorithm. We use a similar normalization method as shown in [11], which is

$$L_i^N(t,d) = \frac{x_i(t,d) - \min_{t\in T} x_i(t,d)}{\max_{t\in T} x_i(t,d) - \min_{t\in T} x_i(t,d)} \tag{1}$$

where $L_i^N(t,d)$ is normalized load i at time t on day d and N stands for normalization. x_i and T are the original power consumption of load i (kWh) and the set of time for one day. Since following analysis performs on the same day, we drop the day index d to improve readability.

We use the Euclidean distance to measure the distance between the total building load and a small-scale load i, i.e., $D(L_i, C)$ which can be obtained by

$$D(L_i, C) = \sqrt{\sum_{t\in T} (L_i^N(t) - C^N(t))^2} \tag{2}$$

where $C^N(t)$ is a centroid, which is normalized total building load at t.

The number of clusters is an important parameter for the k-means clustering algorithm. Although the more clusters give better grouping result, it does not guarantee that the solution is better. To get a proper number of clusters, we use the Silhouette coefficient [13], which is given by

$$s(j) = \frac{b(j) - a(j)}{\max(a(j), b(j))} \tag{3}$$

where $a(j)$ is the average distance of j with all other data within the same cluster (cohesion), and $b(j)$ is the average distance of j to other cluster which not containing the j (separation). We use the average value of total data, i.e., $S_k = \frac{1}{m}\sum_{j=1}^{m} s(j)$, where m and k denote the numbers of sample data and clusters, respectively. The value of the Silhouette coefficient lies between -1 and 1. When S_k is close to 1, it means a good clustering result. To determine the optimal number of clusters k, S_k values are obtained by changing k from 2 to 10. As a result, we choose $k = 3$ which maximizes the Silhouette coefficient, i.e., $S_3 = 0.635$. Note that the proper criterion of clustering is $S_k > 0.6$. The three clusters are named as on-off, regular and irregular as shown in Fig. 2.

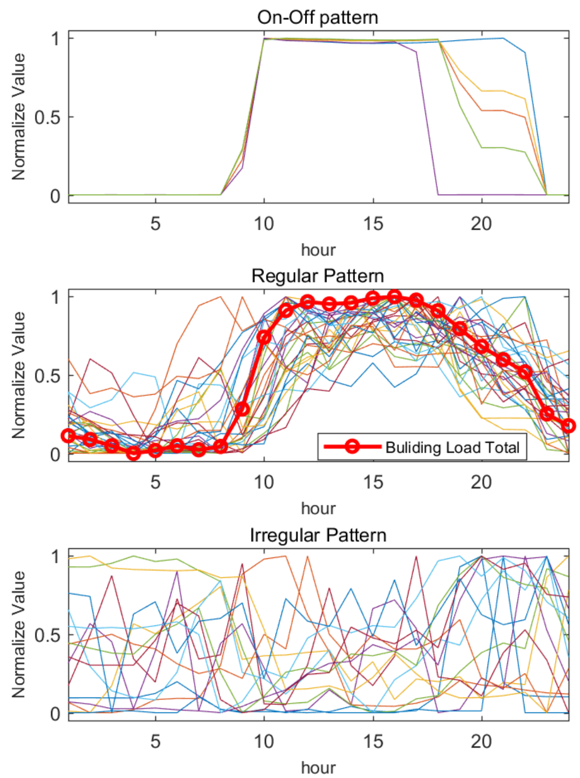


Fig. 2. The clustering results of a day.

Among the three clusters, the regular pattern cluster shows a similar pattern to the total building load by setting a fixed centroid as the total building load. There are 31 small-scale loads out of 56 in this cluster. In a regular pattern group, the thick red line with a circle is the normalized total building load. The other two groups show quite different patterns. The on-off pattern group has six small-scale loads. Load curves in this group have on- and off-periods. In off-period, the load consumes almost zero electricity. On the other hand, it consumes high electricity in on-period. The last cluster is the irregular pattern group which is not easy to find any important characteristics. However, since the loads in this pattern have relative small peak, such as less than tens of kW, we do not use this cluster as the input data of the proposed ANN.

III. Forecasting Model

A. Artificial Neural Network

The ANN is an algorithm designed to imitate the human brain and to train the machine with various data. In ANN, a perceptron receives inputs from its child perceptron, and then the perceptron also sends its data to the next perceptron by using the activation function.

Multi-layer perceptron (MLP) has several hidden layers to solve nonlinear problems. To train the MLP, a back-propagation (BP) algorithm is used which adjusts weights until the error reached a threshold. The error is defined as the difference between the output of MLP and the target value. The output value obtained by the feed-forward algorithm is calculated as

$$y(\boldsymbol{x}) = \sum_{k=1}^{s} w_{y,k} h_k(\boldsymbol{x}) + b \tag{4}$$

$$h_k(\boldsymbol{x}) = \phi(\boldsymbol{w_k} \cdot \boldsymbol{x} + b_k) \tag{5}$$

where $w_{y,k}$, $h_k(\boldsymbol{x})$, b, and s denote the weight for the output neuron, the output of a hidden neuron k, the bias and the number of hidden neurons[2], respectively. There is no concrete way of determining the number of neurons in the hidden layer, so they must be determined empirically by the network designer [14]. In this paper, we choose one hidden layer with seven neurons, which has the best performance.

In (5), $\boldsymbol{x}$, $\boldsymbol{w_k}$, and b_x denote the input vector, the weight vector for the hidden neuron k, and the bias, respectively. In this work, we use the ReLU function[3] $\phi(\cdot)$ as the activation function. The ReLU function is defined as

$$\phi(x) = \max(0, x). \tag{6}$$

The gradient of this function is always 1 when $x \geq 0$ and 0 otherwise. The ReLU function is a good solution to solve the vanishing gradient problem since the error at each layer is effectively delivered to the previous layer. Therefore, an ANN can be trained well with the BP algorithm without the vanishing gradient problem [15].

The error is defined by the difference between the output value y and the target value l, it is given by

$$E(\boldsymbol{w}) = |y(\boldsymbol{x}) - l|^2. \tag{7}$$

Training the ANN means that adjusting the weights to minimize $E(\boldsymbol{w})$. To this end, we use the gradient descent method as a BP algorithm. That is,

$$\boldsymbol{w(t+1)} = \boldsymbol{w(t)} - \eta \frac{dE}{dw}, \tag{8}$$

where η is the learning rate, and it is 0.01 in this work.

B. Input Data Selection

To get a good performance in ANN, it is very important to decide input data combination [8]. In this work, to forecast the total building load, the one day before total building load for 24 hours, the temperature, and the small-scale load are considered as the input data. A detailed explanation of the selected input data is as follows.

- **Total building load**: The load pattern of today is similar to that for the previous day in general, so the one day before (D-1) total building loads for 24-hour is selected as input data. [24 inputs]
- **Temperature**: Among the weather data, temperature shows the highest correlation to the power load, especially in summer and winter seasons. We use six temperature data that are average, minimum, and maximum

[2] We choose seven hidden neurons in this work.

[3] Typical activation functions include step function, sigmoid function and ReLU function.

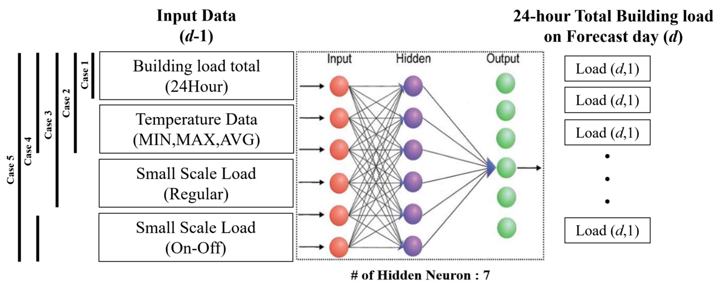


Fig. 3. The proposed ANN-based short-term load forecasting model for campus building

temperatures for D-1 and D day. Note that temperature data for D day is forecasted one. [6 inputs]

- **Small-scale loads (regular)**: The 31 small-scale loads in the regular pattern cluster are highly correlated to the total load. Therefore, we use the summation of the 31 regular small-scale loads as input data. [24 inputs]
- **Small-scale loads (on-off)**: The six small-scale loads in the on-off pattern cluster are also used as input data. To reflect the impact of the on-off load on the total load, we use a binary value per a small-scale load which has 0 and 1 when the load is off and on for the day, respectively. [6 inputs]

C. Normalization

The ranges of input data are very different. For instance, temperature changes about from -10 to 35 ^{o}C while the building load changes from 100,000 Wh to 600,000 Wh. Therefore, we use normalized data for training. There are several normalization methods such as min-max and Z-score methods. We use the Z-score normalization method since this method shows better forecasting result than the min-max method. The Z-score method for the input data x is formally defined as

$$Z_x = \frac{x - \mu_x}{\sigma_x}, \tag{9}$$

where μ_x and σ_x denote the average and variance of the whole input data during the training period, respectively.

D. Load Forecasting Model

The proposed ANN-based short-term load forecasting model predicts day-ahead hourly load. That is, the output of the ANN is 24 hours load for D-day. The structure of the proposed forecasting model is shown in Fig. 3. There is one hidden layer and the number of hidden nodes in the hidden layer is seven [5]. We propose five forecasting model to investigate the effect of input data selection.

- **Case 1**: 24-hour total building load of D-1 day (24 input nodes)
- **Case 2**: Case 1 + three temperature data (average, minimum, and maximum temperatures) of D-1 day and D day (30 input nodes).
- **Case 3**: Case 2 + 24-hour regular loads of D-1 day (54 input nodes).

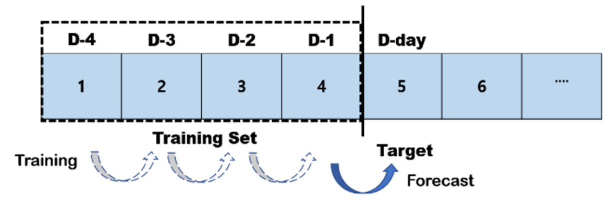


Fig. 4. Training and forecasting procedure.

TABLE I
FORECASTING PERFORMANCE FOR FIVE MONTHS

		Jan	Apr	Jun	Aug	Oct
Case 1	MAPE	8.3544	6.1975	4.347	4.225	7.6316
	CV	9.7287	8.6003	5.2824	5.2902	12.408
	RMSE	27.112	20.328	17.347	15.96	28.79
Case 2	MAPE	8.9894	**6.0506**	4.4717	4.2201	7.2298
	CV	10.314	**8.1480**	5.2707	5.4217	11.719
	RMSE	28.596	**19.315**	17.319	16.374	27.149
Case 3	MAPE	**8.2917**	6.1923	**4.3120**	4.3369	**7.0632**
	CV	**9.6368**	8.3959	**5.1213**	5.4371	**11.485**
	RMSE	**26.855**	19.883	**16.852**	16.419	**26.635**
Case 4	MAPE	8.4942	6.5405	4.4186	**4.1029**	7.9225
	CV	9.6228	8.9241	5.2061	**4.9694**	13.128
	RMSE	26.857	21.1	17.112	**15.013**	30.409
Case 5	MAPE	8.6472	6.0708	4.4305	4.1961	7.4703
	CV	9.9515	8.1679	5.3455	5.2701	12.022
	RMSE	27.754	19.345	17.523	15.915	27.9

- **Case 4**: Case 2 + On/Off loads of D-1 day (36 input nodes).
- **Case 5**: Case 2 + 24-hour regular loads of D-1 day + On/Off loads of D-1 day (60 input nodes).

Fig. 4 shows the training procedure for the proposed model. To train the ANN, we use four days which consist of one to four days prior to the forecasted day (D-day), and the training data set is used sequentially. Among the four training days, three days and one day are used for training and test, respectively. There are two reasons to use the four days as training data. First, it is because the most recent data have the most similar pattern to D day. The other reason is that we only have the building data set for one year. It is not easy to train well and to get a good forecasting performance for the ANN with one year data.

IV. RESULTS AND DISCUSSION

In this section, we evaluate the performance of the proposed ANN-based short-term load forecasting models in terms of the mean absolute percentage error (MAPE), root mean square error (RMSE) and the coefficient of variation (CV)[4].

Table I shows the forecasting performances of five months for the five ANN models. The five months are selected as representative months for the season, semester and vacation. Among the five ANN models, Case 3 generally shows good performance. The other four cases show back and forth result. For the best performing model in each month, those without

[4]The most commonly-used three evaluation indices [2].

TABLE II
AVERAGE PERFORMANCE AND STANDARD DEVIATION FOR THE WHOLE PERIOD

		MAPE	CV	RMSE
Case 1	mean	7.2535	9.0713	24.6602
	std	5.2652	6.3523	15.9862
Case 2	mean	7.2058	8.9359	24.3279
	std	5.2076	6.0008	15.2044
Case 3	mean	**7.0527**	**8.7372**	**23.7705**
	std	**4.9579**	**5.7092**	**14.3469**
Case 4	mean	7.2758	9.0633	24.6032
	std	5.1260	6.3525	15.8154
Case 5	mean	7.1443	8.8724	24.1833
	std	5.0089	6.0121	15.2607

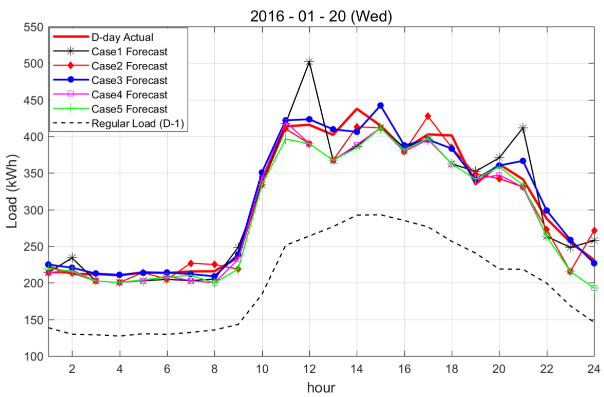


Fig. 5. Forecasting results of January 20, 2016.

small-scale loads (Cases 1 and 2) and those with small-scale loads (Cases 3, 4, and 5) get three times and nine times, respectively. Therefore, it is reasonable to infer that the small-scale loads contribute to the performance improvement.

Table II shows the yearly mean and standard deviation performances for all cases. As shown in Table. I, although Case 3 does not always show the best performance throughout a year, it shows the best and the most stable performance. Since Case 5 shows the second-best performance, it is confirmed that the use of the regular small-scale load data improves forecasting performance. On the other hand, on-off small scale load, i.e., Case 4, fails to improve the forecasting performance than those without small-scale loads because capturing the exact timing of an on-off pattern is not easy. Comparing the performances of Cases 1 and 2, it is also confirmed that temperature data contribute to performance improvement for building loads.

To explain the performance improvement of Case 3, Fig. 5 shows the forecasting results and the real load on January 20, 2016. In the figure, the summation of regular small-scale loads for D-1 day, i.e, input data for Cases 3 and 5, is also plotted as a dotted line. The dotted line shows a smooth curve since it is a summation of regular loads, and this data makes the forecasting output smooth as shown in Fig. 5. Therefore, the general forecasting performance improves for the building load.

V. CONCLUSION

In this paper, we propose ANN-based building load forecasting models using small-scale load data in a building. The small-scale loads are classified by the fixed k-means clustering algorithm resulting in three clusters: regular, on-off, and irregular loads. The inputs for the proposed forecasting model are the one day before total building loads, temperature, and regular and on-off small-scale loads. To evaluate the performance of each input data selection, we build five ANN models which have different input data combinations. The results of the case study show that ANN with the regular small-scale loads performs the best because the regular small-scale loads make the forecasting output smooth. For the future work, the proposed forecasting model is expanded to other buildings or distribution feeders. Deep learning algorithm can be developed with more data of the building.

REFERENCES

[1] IEA, Energy Technology Perspectives, Transition to Sustainable Buildings, “Final Energy Consumption By sector and Building Energy Mix,” 2010. Available: http://www.iea.org/etp/buildings/

[2] K. Amasyali, N. M. El-Gohary, “A review of data-driven building energy consumption prediction studies,” *Renewable and Sustainable Energy Reviews*, vol. 81, part 1, pp.1192-1205, January, 2018.

[3] Korea Energy Agency,“Energy consumption statistics,” 2015. Available: http://www.energy.or.kr/

[4] J. H. Lim, O.S. Kwon. K. B. Song and J.D. Park, “Short-term load forecasting for educational building with temperature correlation,” in *Proc. IEEE POWERENG*, Istanbul, Turkey, May, 2013.

[5] J. H. Moon, S. H. Jun, J. W. Park, Y. H. Choi and E. J. Hwang, “An Electric load forecasting scheme for university campus buildings using artificial neural network and support vector regression,” *KIPS Trans. Comput. Commun. Syst.*, vol. 5, no. 10, pp. 293-302, October, 2016.

[6] X. Ke, A. Jiang and N. Lu, “Load profile analysis and short-term building load forecast for a university campus,” in *Proc. IEEE/PES General Meeting*, Boston, MA, USA, July, 2016.

[7] S.-J. Baek and S.-G. Yoon, “Short-term load forecasting for university campus building using artificial neural network,” in *Proc. KIEE summer annual conference*, Busan, Korea, July, 2017 (Korean).

[8] M. B. Tasre, V. N. Ghate, “Input vector comparison for hourly load forecast of small load area using artificail neural network,” in *Proc. International conference on Communication Systems and Network Technologies*, Rajkot, India, May, 2012.

[9] L. Hernandez, C. Baladron, J. M.Aguiar, B. Carro, A. E. Sanchez, J. Lloret. “Short-term load forecasting for microgrids based on artificial neural networks,” *Energies*, vol. 6, no. 3, pp. 1385-1408, March, 2013.

[10] F. L. Quilumba, W. J. Lee, H. Huang, D. Y. Wang and R. L. Szabados, “Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities,” *IEEE Trans. Smart Grid*, vol. 6, no. 2, pp.911-918, March, 2015.

[11] X. Sun, P. B. Luh, K. W. Cheung, W. Guan, L. D. Michel, S.S. Venkata and M. T. Miller, “An Efficient approach to short-term load forecasting at distribution level,” *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2526-2537, July 2016.

[12] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” in *Proc. Berkeley Symp. Math. Statist. Prob.*, vol. 1, no. 14, pp. 281-297, 1967.

[13] P. J. Rousseeuw, “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis,” *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53-65, November, 1987.

[14] M. U. Usman, J. Ospina and Md. Faruque, “Fault Classification and location identification in a smart distribution network using ANN,” in *Proc. IEEE/PES General Meeting*, Portland, OR, USA, July, 2018.

[15] S. H. Ryu, J. K. Noh and H. S. Kim, “Deep Neural Network Based Demand Side Short Term Load Forecasting,” *Energies*, vol. 10, no. 3, December, 2016.

[16] B. Yuce, M. Mourshed and Y. Rezgui, “A smart forecasting approach to district energy management,” *Energies*, vol. 10, no. 8, July, 2017.