26th Iranian Conference on Electrical Engineering (ICEE2018)

Loss Estimation in Power Distribution Networks with Limited Information Using Clustering Techniques

Hamed Sarjan
Faculty of Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
hamed.sarjan@mail.um.ac.ir

Habib Rajabi Mashhadi
Faculty of Engineering
Ferdowsi University of Mashhad
Mashhad, Iran
h_mashhadi@um.ac.ir

***Abstract*—An accurate loss estimation in distribution lines dramatically depends on load estimation in distribution transformers. The lack of network databases in large-scale distribution networks leads to use estimation methods. In this work, we introduce a new method to estimate the loss of distribution lines based on transformer load. This method contains two parts. The first one introduces a new load estimation method using Fuzzy C-means clustering. In this part, extraction of two load patterns (residential and commercial) for each cluster subject to a small fraction of network data is the main step followed in this paper. In the second part, loss estimation is performed based on load flow. IEEE-33 bus test system was applied to demonstrate the accuracy of the proposed method.**

***Index Terms*—distribution transformer, loss estimation, load estimation, residential customers, commercial customers, limited information, Fuzzy C-means clustering**

I. Introduction

Energy losses are inescapable in power system. Energy losses which occur in both transmission and distribution networks exist at all levels from generation to loads. The major part of the energy is lost in distribution networks, especially in low voltage level. Energy losses in low voltage distribution networks are almost 70% of energy transport technical losses and can climb to 75% in some power systems [1]. An accurate energy loss estimation in distribution network highly depends on the precision of determining the load in distribution transformers.

Estimation of load, which is the most important requirements for efficient operation, real-time monitoring and controlling of power distribution systems, plays a significant role in loss estimation [2]. The expansion of the distribution networks needs an accurate examination. It is obvious that loss and load estimation in a large distribution system is not as easy as small one and relates to the observability of distribution systems significantly. Moreover, observability relates to the number of installed smart meters in distribution systems nodes and their locations but the installation of smart meters in each node is technically and economically impossible [3]. Load estimation methods need the limited number of installed smart meters to estimate. A daily load curve is representative of Consumer behavior and corresponds to the electric power consumption. Availability of these data depends on the type of consumer [4].

The smart meters are not installed in each costumer of large-scale distribution networks because of economic and technical reasons and these smart meters are used executively for customers with high consumption. On the other hands, smart meters are not installed on the consumers which are connected to LV distribution networks. It is obvious that the load characteristics of all customers required to evaluate for improving the observability of distribution systems.

A novel approach for the loss estimation of the power distribution networks is proposed by the author in [5] and it is based on the typical load profiles of the nodes with the minimum information. The clustering methods are applied to categorize the nodes in power distribution system. In this paper, the optimal number of smart meters is use.

Technical energy loss estimation in power distribution systems is discussed in the paper [6]. In this paper, energy losses were estimated using information about the mean and variance components of the load curve. In this study, the link between the average loss and the time necessary to achieve the equivalent loss at the average load was obtained. Additionally, it is said the information about the energy consumption is more meaningful and usually more reliable than the maximum demand in loss estimation. Furthermore, the information about losses during the peak loads is obtained from the energy loss distribution through statistical methods.

A method for estimation and evaluation of the loss estimation on a distribution feeder is introduced in [7] which can produce useful outputs, by network segments, obtained from a limited database. Load curve of the studied feeder or substation and some complimentary data are required for this methodology. This method was applied to two real Brazilian distribution systems.

978-1-5386-4916-9/18/$31.00

Most of the existing loss estimation studies concentrate on distribution system clustering, estimation function or artificial intelligence based approaches and the majority of these are aimed at off-line studies at the distribution level and utilize mainly the historical data similarity [8]. It should be mentioned some of those studies also applied a limited amount of smart meters [9]. Because of the similarity in energy consumption of the customers, all of the approaches use the fact that most of the loads follow a very similar daily load profile. Learning methods, such as artificial neural networks, fuzzy logic, have also been proposed [10].

Herein, an improved approach for the loss estimation of the electric distribution system, based on load estimation in distribution transformer which is used the residential and commercial load profiles obtained from a limited number of smart meters and heuristic formula is proposed. We employed Fuzzy C-means classification method for load estimation. Due to lack of information in distribution networks, the main aim of this study is loss estimation based on minimum information.

II. Optimal Smart Meters Placement

The consumer features such as customer bills and the number of commercial and residential customers of each transformer are determinant for clustering and organization of the distribution transformers which is followed by the determination of the candidate points for smart meter installation. It is because of installation of measurement devices on all network transformers does not consider wisely. Therefore, it is important to place existing measurement devices on the appropriate buses of the network to gain more accurate information from their group transformers. If the equipment install randomly, the result will not be trustworthy. Hence, it should be necessary to consider the criterion to choose the perfect locations. The selected locations should be proper representative of transformer groups.

It is possible that the characteristics of the cluster center and transformers are not similar. Therefore, it is not correct to place the measurement devices on it center. Whit regard to differences that are between characteristics of the cluster center and transformers the similarity of data related to the transformer of a cluster within center should be expressed by using a criterion. The degree of membership is considered suitable criterion for determining appropriate location of measurement devices.

The degree of membership for each transformer is evaluated in a cluster that is gained from FCM and one of which that has the highest degree of membership in the cluster is introduced as the cluster center. We introduce two cluster center namely the math cluster center and physical one. The math cluster center is gained from FCM, and the physical cluster center is the transformer that has selected based on the most degree of membership.

III. Load Pattern Extraction

In distribution networks operation studies, it is necessary to know the consumer consumption pattern. There is not a wide disparity in consumer's behavior of urban distribution networks (including residential and commercial customers). The network operator can create the consumption pattern which can be recognized using a group of consumers [11]. The network operator adopt a method to extract and rebuild the load profile of the distribution transformer. There are many methods such as, clustering, using a mean feeder, etc. were used in different investigations that had different accuracy. It is an appropriate approximation for all customer's load profiles. However, it is not precise load pattern for all customers and transformers.

In this study it is crucial to discover the load pattern of distribution transformer. Each kind of customers has its executive pattern. Therefore, it is important to introduce a method which includes two types of load patterns: residential and commercial users. The load pattern estimation of residential and commercial users gets high errors without network transformers categorize. Therefore, clustering technique divides network into smaller networks. For each small network, a couple of load profiles can be gained from residential and commercial models. Dividing the network into several regions causes the pattern extraction error be reduced.

In the following, we introduce a relationship between transformer load and commercial and residential load pattern. This relationship is run as an equation for each cluster exclusively. The number of unknowns of this relationship is more than the number of clusters. Because the unknowns are the residential and commercial consumer patterns.

$$r_{kn}u_{k1}\times P^t_{r1}+c_{kn}u_{k1}\times P^t_{c1}+r_{kn}u_{k2}\times P^t_{r2}+c_{kn}u_{k2}\times P^t_{c2}$$
$$+\ldots+r_{kn}u_{kn}\times P^t_{rn}+c_{kn}u_{kn}\times P^t_{cn}=P^t_{kn}. \quad (1)$$

The above equation is written for each cluster, where:

r, c and n: Number of residential and commercial customers and number of clusters, respectively

r_{kn}: The number of residential customers of the transformer that has K smart meter in the cluster number n

c_{kn}: The number of commercial customers of the transformer that has K smart meter in the cluster number n

P^t_{kn}: The registered hourly power of the transformer that has K smart meter in the cluster number n

P^t_{rn}: Hourly power consumption pattern of residential customers in cluster number n

P^t_{cn}: Hourly power consumption pattern of commercial customers in cluster number n

u_{kn}: The degree of transformer membership corresponds to the k smart meter of the cluster number n

If the equation is run for each cluster separately, the number of unknowns is twice the number of equations. To solve these mentioned equations, we need an ancillary equation for each cluster, which can be generalized for the sum of transformers of each cluster:

$$(\sum r_{in}u_{i1})\times P_{r1}^{t}+(\sum c_{in}u_{i1})\times P_{c1}^{t}+(\sum r_{in}u_{i2})\times P_{r2}^{t}+(\sum c_{in}u_{i2})\times P_{c2}^{t} +\ldots+(\sum r_{in}u_{in})\times P_{rn}^{t}+\sum(c_{in}u_{in})\times P_{cn}^{t}=LAF_n\times P_{kn}^{t}. \quad (2)$$

In the above equation, LAF_n is the coefficient which is obtained from the following equation:

$$LAF_n=\frac{E_{TCn}}{EKn} \quad (3)$$

Where:

E_{TCn}: Total consumption of transformers of cluster number n

E_{Kn}: The total consumption of transformer that has smart meter k in the cluster number n

This method is based on the segregation of residential and commercial transformer load patterns. It is obvious that transformer supply customers follow a specific load curve. In addition, the information of transformers within a cluster, such as the number of customers and the total customer bills, are almost same. Furthermore, transformers that is selected as physical cluster center within a cluster follow a residential pattern and a commercial pattern. The load profile of each transformers inside each cluster approximately equivalent to cluster candidate.

There is a restriction in the number of measurement equipment. Therefore, using one pattern curves for all transformers in a cluster will be contained some errors in estimating the load. There are two ways to reduce the estimation errors. The increase in the number of clusters cause to increase in amount of measurement equipment is the first way. The second way is increase in the number of measurement equipment in each cluster especially on distant cluster center transformers. It is clear the transformers on which were installed the smart meter has the minimum errors and also the measurement equipment error was assumed to be zero so it can be assumed the error of estimation is considered zero for these transformers.

IV. Load Estimation

The load pattern extraction of residential and commercial customers can be easily performed from the suggested method. The load of the transformers was estimated by using the load pattern of residential and commercial customers and proposed equation. The pattern load curves were gained from (1) and (2) were specific to the study period. In other words, we estimated the load of other cluster transformers by referring to physical cluster center load pattern at 15-minute intervals in all days. The load of distribution transformer can be estimated using (1). In the following, they were aggregated and the curves were fitted for transformers.

V. Loss Estimation

With respect to the instantaneous load of each transformer, it is easy to calculate the average network losses by load flow at the distribution level. In this study, values of the active power of each transformer have been estimated. Therefore, the reactive power values of each transformer must be determined in order to carry out the load flow problem. The Power Factor (PF) value in each transformer is assumed to be constant, equal to 0.85. As a result, it is easy to obtain instantaneous reactive power values (every 15 minutes). As soon as knowing the active and reactive power values for each node, the load flow problem easily runs and the power losses for each line from the network are obtained. It should be noted that due to the selection of the IEEE 33 BUS network, the resistance and reactance values of the 20 KV lines are specified.

VI. Case Study

In this work, IEEE 33-bus radial distribution system test is used. There is a distribution transformer (20kv/0.4kv) at each node of the network which is connected to the specific amount of customers. Two types of consumers (residential and commercial) were considered in this research.

Fig. 1 shows the topology of IEEE 33-bus system test with a substation at the first node and loads at the remaining nodes. Each transformer consists of residential and commercial customers.

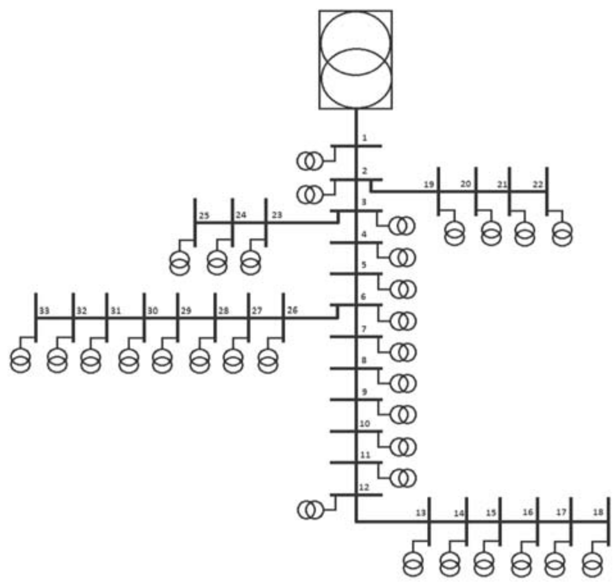

Figure 1. IEEE-33 bus system

VII. Distribution Transformer Clustering

The accurate features of distribution transformer are necessary to reach acceptable results of clustering. But, the number of residential and commercial customers and periodic bills customers is the only available information of transformers. Therefore, these features can be selected to apply in FCM as input information.

After collecting information, we needed to choose the number of clusters. We used FCM for 2 to 25 clusters. Dispersion of data is shown in Fig. 2. The data is shown in two dimensions to be more obvious. The data was normalized because of the difference of dimension of data. Energy consumption of consumers was registered with measurement equipment placed in each house. Aggregation of energy consumption of all customers for a specific period (2 months) was operated and the energy consumption of each transformer was finally calculated.

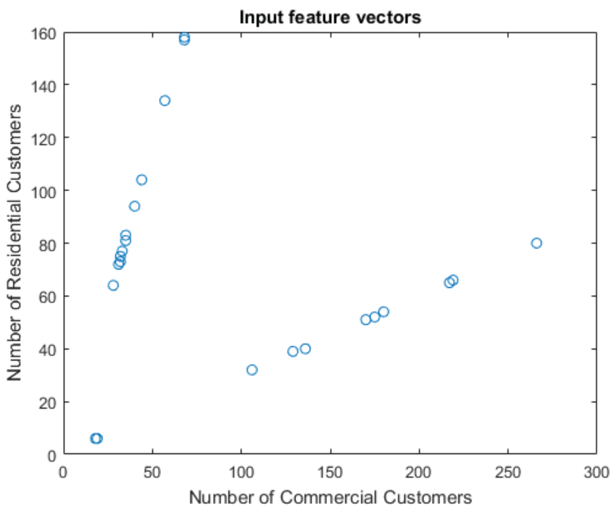

Figure 2. Dispersion of transformer data

In next stage, the program was run for a different number of clusters. For example, the clustered data and center of clusters is shown for 7 clusters in Fig. 3.

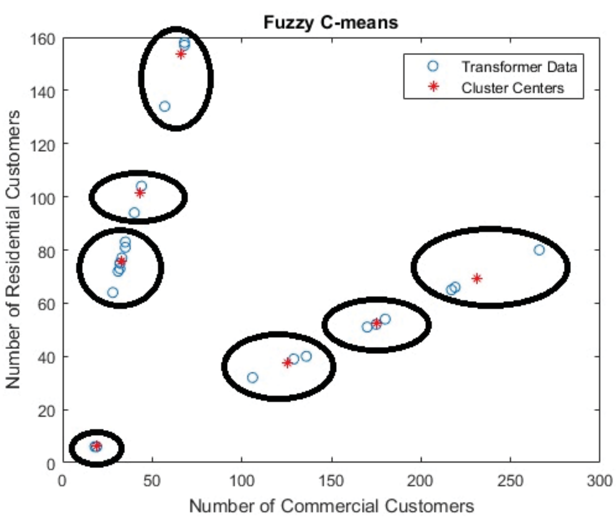

Figure 3. clustered transformers and center of clusters for 7 clusters

A transformer that had the most degree of membership is selected as the cluster center.

In table I, the difference of physical centers and math centers has been shown for 7 clusters.

TABLE I THE DIFFERENT OF PHYSICAL CENTERS AND MATH CENTERS

Number of clusters	**Math cluster centers**		**Physical cluster centers**	
	Residential customer number	***commercial customer number***	***Residential customer number***	***commercial customer number***
1	101.4877	43.2571	104	44
2	6.0911	18.7934	6	19
3	75.6893	32.6950	75	32
4	52.4467	175.3310	52	175
5	37.6093	125.6064	39	129
6	69.4717	231.0968	66	219
7	153.4780	66.1637	157	68

The investigation duration was 7 days and the sampling interval was every 15 minutes. By installation of smart meters on the representative transformers, after the 7-day study period, the active inject power samples had been registered. Extracted samples from the measurement devices are shown in Fig. 4.

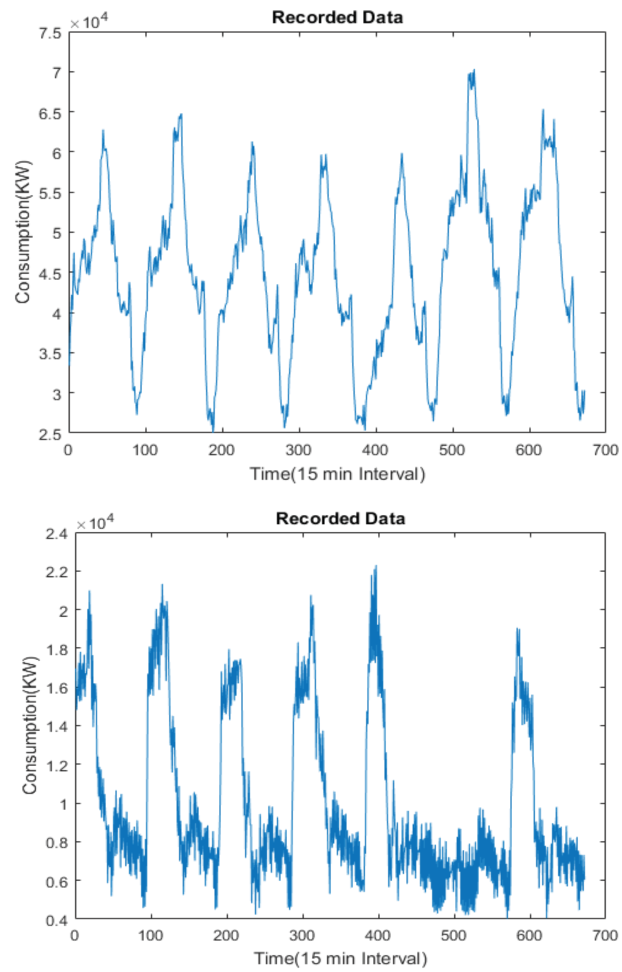

Figure 4. samples extracted from the smart meters

Using (1), transformer load estimation is gained easily. The estimation program was run each 15 minutes for distribution transformer then was aggregated. In order to demonstrate the accuracy of load estimation, it is necessary to compare the estimated load with their actual values. For this purpose, the estimated load profile and actual load profile was drawn in one graph.

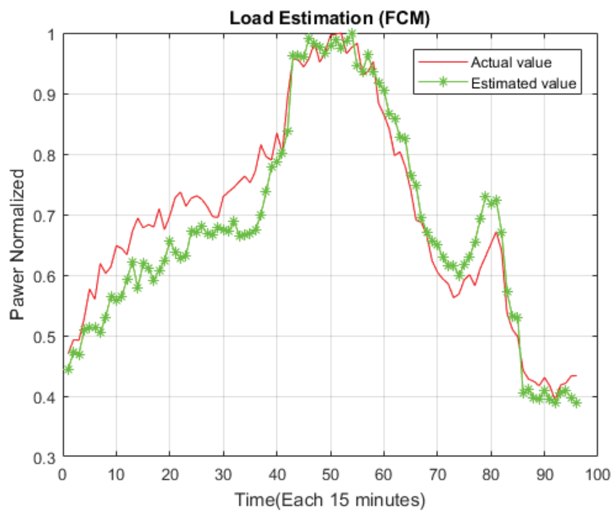

Figure 5. estimation of transformer load (BUS 33)

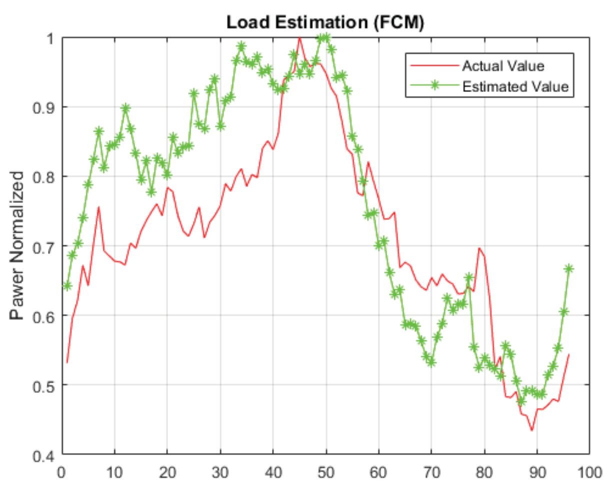

Figure 6. estimation of transformer load (BUS 4)

As can be obvious in Fig. 5, the amount of errors is low due to the low destination of the transformer to the cluster center, and there is not a high difference between estimated values and their actual values at different moments. But in Fig. 6, the estimated error is slightly greater in comparison with Fig. 5. The distance of the transformer from the center of the cluster plays the main rule in rate of error. This problem can be solved by increasing in the number of clusters and the number of measurement devices. The amount of error of representative transformers equals to zero because the transformer placement is in the cluster center and as a result, two curves which show the actual loads and estimated load values are exactly similar.

As can be seen in Fig. 7, the estimated power losses on the lines are not significantly different from the actual value. However, in some lines containing a transformer with a high load estimation error, the amount of loss estimation error is different. In this study, the estimation error was 8.176%, which, in comparison with other estimation methods, is a valid amount. It should be noted that the estimation error significantly depends on the correlation of consumption of subscribers. The consumption of subscribers with a greater correlation has less error rate. Consequently, the more correlation of data reduces the average distance between the center of the cluster and transformer. Finally, the purpose of this research is to calculate the percentage of technical losses that was 4.14%. The percentage of technical losses depends on network length. The error in rural areas is more significant than urban areas.

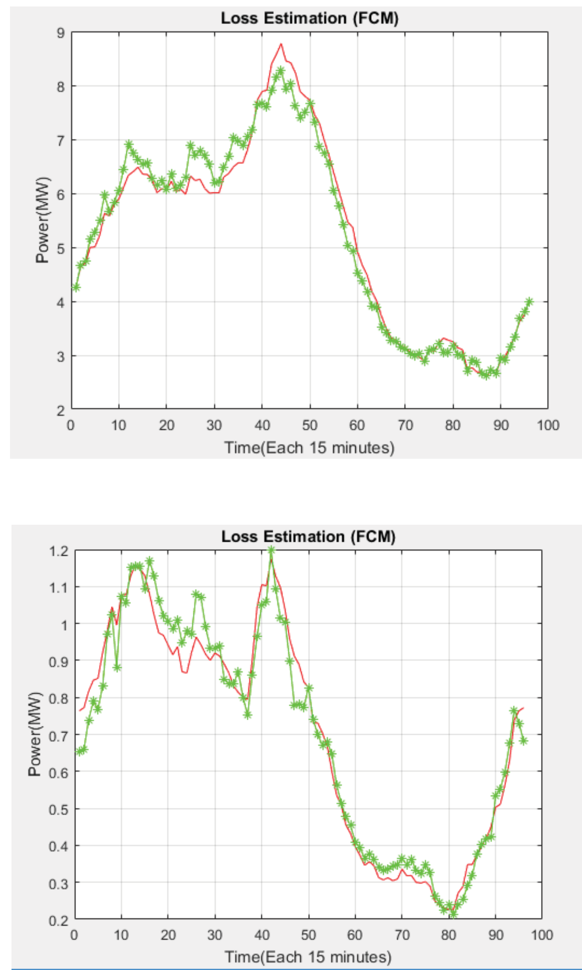

Figure 7. Estimation of line loss

VIII. CONCLUSIONS

The loss estimation in different parts of the distribution system especially distribution lines is one of the most significant requirements for efficient operation of electric distribution systems.

In this study, a new approach which estimates the distribution lines loss is proposed based on distribution transformers load. The load profiles of the distribution transformer were obtained from the heuristic approach and through the FCM method. The validation of method was done by the actual transformers data in IEEE 33-bus radial distribution system test. The loss of lines successfully estimated and they were compared with the actual one. Although, in this paper the loss estimation method is used on the IEEE 33-bus system, it can be applied to all types of power distribution systems, meshed or radial.

REFERENCES

[1] H. Schau and A. Novitskiy, "Analysis and prediction of power and energy losses in distribution networks," in *Universities Power Engineering Conference 2008,* Italy, 2008, pp. 1-5.

[2] G. Grigoraş, F. Scarlatache, and G. Cârţină, "Load estimation for distribution systems using clustering techniques," in *13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), 2012*, Romania, 2012, pp. 301-306.

[3] J. A. Massignan, C. A. Fantin, J. B. London, and M. H. Camillo, "Real-time load estimation for distribution feeders," in *PowerTech, 2015 IEEE Eindhoven*, Netherland, 2015, pp. 1-6.

[4] U. Singh, V. Zamani, and M. Baran, "On-line load estimation for distribution automation using AMI data," in *Power and Energy Society General Meeting (PESGM)*, 2016, pp. 1-5.

[5] A. K. Dashtaki, M. R. Haghifam, "A new loss estimation method in limited data electric distribution networks," *IEEE Transactions on Power Delivery*, vol. 28, pp. 2194-2200, 2013.

[6] L. M. Queiroz, M. A. Roselli, C. Cavellucci, and C. Lyra, "Energy losses estimation in power distribution systems," *IEEE Transactions on Power Systems*, vol. 27, pp. 1879-1887, 2012.

[7] M. Oliveira, A. Padilha-Feltrin, "A top-down approach for distribution loss evaluation," *IEEE Transactions on Power Delivery*, vol. 24, pp. 2117-2124, 2009.

[8] X. Tang, K. N. Hasan, J. V. Milanovic, K. Bailey, and S. J. Sgtott, "Estimation and Validation of Characteristic Load Profile through Smart Grid Trials in a Medium Voltage Distribution Network," *IEEE Transactions on Power Systems*, vol. 33, pp. 1848 - 1859, 2017.

[9] H. Wang and N. N. Schulz, "Using AMR data for load estimation for distribution system analysis," *Electric Power Systems Research*, vol. 76, pp. 336-342, 2006.

[10] Y. Chen, P. B. Luh, C. Guan, Y. Zhao, L. D. Michel, M. A. Coolbeth, P. B. Friedland, and S. J. Rourke, "Short-term load forecasting: similar day-based wavelet neural networks," *IEEE Transactions on Power Systems*, vol. 25, pp. 322-330, 2010.

[11] A. Khodabakhshian, R. Hooshmand, and Y. Raisee-Gahrooyi, "A New Pseudo Load Profile Determination Approach in low Voltage Distribution Networks," *IEEE Transactions on Power Systems*, vol. 33, pp. 463- 472, 2017.