Research Article

Assessment of Electrical Load in Water Distribution Systems Using Representative Load Profiles-Based Method

Gheorghe Grigoras

Power System Department, Electrical Engineering Faculty, “Gheorghe Asachi” Technical University of Iasi, Boulevard Dimitrie Mangeron, No. 21-23, 700050 Iasi, Romania

Correspondence should be addressed to Gheorghe Grigoras; ghgrigoras@yahoo.com

Received 17 April 2014; Accepted 28 June 2014; Published 21 July 2014

Academic Editor: Mamun B. Ibne Reaz

Copyright © 2014 Gheorghe Grigoras. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The problem of optimal management of a water distribution system includes the determination of the operation regime for each hydrophore station. The optimal operation of a water distribution system means a maximum attention to assess the demands of the water, with minimum electrical energy consumption. The analysis of load profiles corresponding to a water distribution system can be the first step that water companies must make to assess the electrical energy consumption. This paper presents a new method to assess the electrical load in water distribution systems, taking into account the time-dependent evolution of loads from the hydrophore stations. The proposed method is tested on a real urban water distribution system, showing its effectiveness in obtaining the electrical energy consumption with a relatively low computational burden.

1. Introduction

Water and energy are critical resources that affect virtually all aspects of daily life. A huge amount of electrical energy is necessary for the transportation, treatment, and distribution of water for drinking and industrial consumption and for different internal technological processes of water distribution systems. Water distribution systems are massive consumers of energy, which is consumed in each of the stages of the water production and supply chain: starting from pumping the water to the water treatment plant, followed by the treatment process while distributing the water via the network. In the Report Watergy by Alliance to Save Energy, it has been asserted that 2-3% of the world's electrical energy consumption is used to pump and treat water for civil and industrial supply [1]. Energy costs constitute the largest expenditure for nearly all water utilities worldwide and can consume up to 65 percent of a water utility's annual operating budget [2]. The energy requirements vary significantly from city to city, depending on local factors such as topography, location and quality of water sources, pipe dimensions and configurations, treatment standards required, and the types and numbers of consumers [1-8]. Water industry decisions on operational strategies and technology selection can also significantly influence electrical energy consumption [5]. A

high electrical energy consumption may be due to various reasons: inefficient pump stations, poor design, installation or maintenance, old pipes with high head loss, bottlenecks in the supply networks, excessive supply pressure, or inefficient operation strategies of various supply facilities [2-4, 9-16].

Energy-saving measures in water supply systems can be realized in many ways, from decreasing the volume of water pumps (e.g., adjusting pressure zone boundaries) to reducing the price of energy (e.g., avoiding peak hour pumping and making effective use of storage tanks) or increasing the efficiency of pumps (e.g., ensuring that pumps are operating near their best efficiency point). These energy-saving measures often pay for themselves in months, most do so within a year, and almost all recover their costs within three years. Prolonging this enactment period would increase the investment required for long-term [11-16].

Utilities can further reduce energy costs by implementing on-line telemetry and control systems (SCADA) and by managing their energy consumption more effectively and improving overall operations from water supply systems [2]. The motivation for introducing such systems is due to the following factors [2, 9-15]:

- (i) operation of water supply systems which is in many cases becoming more complex, with rising demands,

incorporating water from a variety of sources and aging systems;

- (ii) high operating costs, which justify investments to improve efficiency;
- (iii) control and computer hardware and software which are available and more reliable;
- (iv) the fact that as more computer control systems are installed, there is more experience from which to learn.

Therefore, a permanent policy for the reduction of electrical energy consumption not only involves the technical improvement in the water distribution system, but also requires the use of software tools to facilitate the operation process. Based on this concept, a new approach is proposed in the paper based on similarities that exist between daily load profiles of each hydrophore station from water distribution system and their grouping into representative clusters. By knowing the load profile, water companies can simplify the demand determination for their supply zone. Thus, they can provide better and improved efficiency marketing strategies.

Different techniques have been used in literature for the classification and load profiling, but most of them were implemented to solve the problems from power systems. Table 1 presents a synthesis of the solutions proposed in literature depending on the type of technique.

It can be seen that most techniques belong to Artificial Intelligence (clustering, data mining, self-organizing maps, neural networks, and fuzzy logic) and the number of references is higher after the year 2000. This aspect can be explained by the fact that more and more information is becoming available, faster than ever before. The impact is felt in the control and operation of the distribution and transport networks (electricity, steam, gas, and water). Under these circumstances of momentous changes, Artificial Intelligence has the potential to play a more important role.

This paper proposes an extension of profiling techniques in the area of water distribution systems for the assessment of electrical using on a hybrid algorithm. The algorithm uses the K -means clustering method for obtaining of representative load profiles and a statistical approach for the assessment of the electrical load in the water distribution systems. The remainder of this paper is organized as follows. In Section 2, the K -means clustering method is presented. Section 3 presents the steps of process for obtaining of representative load profiles. Section 4 presents the statistical method for assessment of the electrical load in a water distribution system. Section 5 presents the steps of proposed algorithm. Section 6 shows the results of testing the proposed method on a real system. Finally, Section 7 contains the concluding remarks.

2. K -Means Clustering Method

The K -means clustering is an algorithm to classify or to group the objects based on attributes/features into K number of groups (K is positive integer number). The grouping is done

by minimizing the sum of squares of distances between data and the corresponding cluster centroid [14, 17–20]:

$$\min(E) = \min \left(\sum_{i=1}^K \sum_{x \in C_i} d(x, z_i)^2 \right), \quad (1)$$

where z_i is the center of cluster C_i , while $d(x, z_i)$ is the Euclidean distance between a point x and z_i .

Thus, the criterion function E attempts to minimize the distance of each point from the center of the cluster to which the point belongs. More specifically, the algorithm begins by initializing a set of K cluster centers. Then, it assigns each object of the dataset to the cluster whose center is the nearest and recomputed the centers. The process continues until the centers of the clusters stop changing.

The steps of the algorithm are the following [14, 19–22].

Step 1. Choose K initial clusters centres $z_1^{(0)}, z_2^{(0)}, \dots, z_K^{(0)}$.

Step 2. At the k th iterative step, distribute the samples $\{x\}$ among the K clusters using the relation

$$x \in C_i^{(k)} \quad \text{if } d(x, z_i^{(k)}) < d(x, z_j^{(k)}), \quad i = 1, 2, \dots, K; \quad i \neq j, \quad (2)$$

where $C_i^{(k)}$ denotes the set of samples whose cluster centre is $z_i^{(k)}$.

Step 3. Compute the new cluster centres $z_i^{(k+1)}$, $i = 1, 2, \dots, K$. The new cluster centre is given by

$$z_i^{(k+1)} = \frac{1}{n_i} \sum_{x \in C_i^{(k)}} x, \quad i = 1, 2, \dots, K, \quad (3)$$

where n_i is the number of objects in cluster $C_i^{(k)}$.

Step 4. Repeat Steps 2 and 3 until convergence is achieved, that is, until a pass through the training sample causes no new assignments.

It is obvious in this algorithm that the final clusters will depend on the initial cluster centers chosen and on the values of K .

The optimal number of clusters K_{opt} can be determined using the following algorithm [19–21].

- (1) Determination of the maximum of clusters K_{max} : the maximum of clusters K_{max} should be set to satisfy condition $2 \leq K_{\text{max}} \leq \sqrt{N}$, where N is the clustered objects from database.
- (2) Use the K -means clustering method with given K ($2 \leq K \leq K_{\text{max}}$) for the set of objects from database.
- (3) According to the obtained clusters structure, determine partition quality is evaluated.
- (4) Increase the number of clusters to the K_{max} to see if K -means clustering method finds a better grouping of the data (to repeat Steps 2 and 3).

TABLE 1: Synthesis of the literature techniques for classification and load profiling.

Technique	References	
	Before year 2000	After year 2000
Statistical	[29–33]	[34, 35]
Clustering		[14, 17, 36–43]
Data mining	[44]	[45–47]
Self-organizing maps		[17, 36, 37, 48, 49]
Neural networks	[33]	[50–54]
Fuzzy logic		[11, 34, 39, 50, 51]

- (5) Show number of clusters (K_{opt}) that have obtained the optimal value of the silhouette global coefficient.

Assessing the results of the K -means algorithm represents the main subject of cluster validity. In the process of cluster analysis, the following properties of clusters are being examined: density, sizes and form of cluster, separability of clusters, robustness of classification. There are many approaches to cluster validation [17, 20, 23–27], but internal cluster validation tests are more popular in practice of cluster analysis. The test based on the Silhouette Global Index calculation is one of the most used [20, 23, 24, 28]. This calculates the silhouette width for each sample, average silhouette width for each cluster, and overall average silhouette width for data set. Using this approach each cluster could be represented by a so-called silhouette, which is based on the comparison of its tightness and separation. Cluster validity checking is one of the most important issues in cluster analysis related to the inherent features of the data set under concern. It aims at evaluating clustering results and the selection of the scheme that best fits the underlying data.

3. Determination of Representative Load Profiles

The load profiling represents an alternative to the settlement based on energy meters because in many countries, in water distribution systems there is a lack of necessary metering and monitoring systems to collect data. The load diagram of the hydrophore stations is reconstructed using the normalized load profile and their daily (monthly, yearly, depending on the case) electrical energy consumption. The time interval of sampling load curve data can be one hour. In this situation, a load profile is represented by 24 load values throughout the day. The shape of load profiles is influenced by the type of hydrophore station and, on the other hand, by the type of day or season of the year. Because a large number of load profiles regarding various water hydrophore stations create unnecessary problems in handling them, they could be grouped into coherent groups, seeing that some similarities exist between load profiles. For this purpose, the K -means clustering method is applied to classify profiles of the hydrophore stations into coherent groups—representative load profiles (RLPs).

Each RLP is represented by a vector $x_j = \{x_{jh}, h = 1, \dots, T\}$ for $j = 1, \dots, N$, and the comprehensive set of

RLPs is contained in the set $P = \{x_j, j = 1, \dots, N\}$. The time scale along the day is partitioned into T time intervals of duration Δt_h , for $h = 1, \dots, T$. Hourly values are used in this paper to exemplify the application. The variables used in the calculations are assumed to be represented as constant (average) values within each time interval. The clustering process forms K clusters corresponding to the hydrophore stations. Further, a RLP is assigned to each station.

The algorithm is based on the load profiling process. The major steps are as follows.

(1) *Measurements*. In this step a representative sample of the set of load profiles is identified, the most relevant attributes are to be measured, and the cadence for data collection is defined. Finally the collected data is gathered in a large database.

(2) *Data Cleaning and Preprocessing*. In real problems, like this, involving a large number of measurements, spread over a large geographic area, and collecting data during a considerable period of time different kind of problems will affect the quality of the database. The most relevant and frequent are communication problems, outages, failure of meters, and irregular atypical behavior of some consumers. The result will be a very large database with problems like noise, missing values and outliers. These data (after being cleaned, preprocessed, and reduced) are used in clustering process [43, 55].

(3) *Classification*. For realization of this classification, the K -means clustering method is used. The normalized load profile for each hydrophore station from the water distribution system is determined using a suitable normalizing factor (e.g., energy over the surveyed period):

$$p_{jh} = \frac{P_{jh}}{\sum_h P_{jh}}, \quad j = 1, \dots, N, \quad h = 1, \dots, T, \quad (4)$$

where p_{jh} is normalized value [p.u.]; P_{jh} is actual value [kW]; $\sum P_{jh}$ is the energy over the surveyed period [kWh]; and N is the total number of vectors x_j from database.

(4) *Determination of Representative Load Profiles*. Using the K -means clustering method, the normalized load profiles are refined so as to desist at the unrepresentative profiles. The representative load profile for each cluster is obtained by averaging the normalized values for each hour. These values (called the load factors) lead us to the representative load profiles corresponding to the active powers. The load factors for each cluster are calculated with relation

$$m_h^i = \frac{\sum_{j=1}^{n_i} P_{jh}^i}{n^i}; \quad h = 1, \dots, 24; \quad i = 1, \dots, K, \quad (5)$$

where n^i represents the number of hydrophore stations from cluster i .

The signification of load factors m_h^i , $h = 1, \dots, 24$, $i = 1, \dots, K$ from relation (5) is the following: these factors

transform the electrical energy consumed by the medium member of cluster i in average power demanded by it.

The representative load profiles can characterize very well the operation mode of the water hydrophore stations (identified by the clusters obtained), related to the electrical energy consumption.

(5) *Assignment*. Finally, to each water hydrophore station is made the assignation of a representative load profile.

4. The Assessment of Electrical Load in Water Distribution Systems

The assessment of electrical load in water distribution systems can be made using an improved simulation method based on the representative load profiles of hydrophore stations. The method is based on the following hypotheses [30, 31, 56].

- (i) The mean loads corresponding to a cluster of hydrophore stations from the water distribution system in any hour during the analyzed period are approximately proportional to the electrical energy consumption of those stations.
- (ii) The loads for any hour during the analyzed period have a statistical distribution that can be regarded as normal.

Using these hypotheses, the load estimation of a water distribution system, at any hour $h = 1, \dots, 24$, is given by the following formula:

$$P_h^S = \sum_{i=1}^K N_{HS}^i \cdot W_{med}^i \cdot m_h^i + \sqrt{\sum_{i=1}^K N_{HS}^i \cdot (W_{med}^i \cdot \sigma_h^i)^2}, [\text{kW}],$$

$$h = 1, \dots, T, \quad (6)$$

where P_h^S is the load of the water distribution system at the hour h ; K is the number of clusters corresponding to the hydrophore stations from water distribution system. N_{HS}^i is the number of the hydrophore stations from cluster i ; W_{med}^i is the average energy consumption of the hydrophore stations from cluster i ; m_h^i is the average load factor of the hydrophore stations from cluster i , at the hour h ; σ_h^i is the standard deviation for load factor from hydrophore stations corresponding to cluster i .

The standard deviations for each cluster are calculated with relation

$$\sigma_h^i = \sqrt{\frac{\sum_j^{N_{HS}^i} (p_{jh}^i - m_h^i)^2}{N_{HS}^i}}; \quad h = 1, \dots, T, \quad i = 1, \dots, K. \quad (7)$$

5. Algorithm for Assessment of Electrical Load in Water Distribution Systems

The algorithm adopts a procedure composed of two calculation stages:

FIGURE 1: Flow chart of the proposed method.

- (i) the use of a *clustering* technique (K -means method) in the first stage to determine the patterns of electrical load and determine a subset of representative load profiles to be processed by the second stage: at each iteration, the clustering outcomes simplify the process of selecting a relatively small number of load profiles corresponding to hydrophore stations;
- (ii) the use of *load simulation* in the second stage: this approach is based on a statistical method that considers the loads for any hour during the analyzed period have a statistical distribution that can be regarded as normal.

The flow chart of the proposed method is shown in Figure 1.

For a water distribution system, the input information refers to the technical characteristics (rated power of the force pumps, rated water flow) and the hourly load patterns of hydrophore stations.

FIGURE 2: The silhouette global coefficient for different values of the number of clusters.

FIGURE 3: The silhouette plot for $K_{opt} = 3$.

5.1. Determination of Representative Load Profiles. The load profile of each hydrophore station is normalized relatively to the daily electrical energy consumption. The normalization is made in relation to daily electrical energy consumption because it is always known. This consumption is recorded with meters placed in each hydrophore station. The optimal number of clusters is obtained using the K -means clustering algorithm presented in Section 2. Finally, the silhouette global coefficient is calculated to assess the partition quality. After aggregation of the normalized load profiles of each cluster the representative profiles are determined and a RLP is assigned to each hydrophore station.

5.2. The Assessment of Electrical Loads. In the second step of the study, using information from clustering process (hourly average load factors, average energy consumption, and standard deviation of load factors for each hydrophore station from cluster i , $i = 1, \dots, K$) and relation (6), the hourly loads of the analyzed water distribution system will be obtained.

The accuracy of the estimates is expressed, depending on the data available. Thus, if the actual value of the estimated quantity is available (such as during method development and testing), the following quantity can be useful to verify the method:

$$MAPE = \frac{\sum_{h=1}^T (|P_{h \text{ real}}^S - P_{h \text{ est}}^S| / P_{h \text{ real}}^S)}{T} \cdot 100, \quad (8)$$

TABLE 2: The technical characteristics of hydrophore stations from water distribution system.

Type	Number of stations	Rated power [kW]	Rated water flow [m ³ /h]
I	3	3×2.2	24
II	6	4×2.2	32
III	11	4×4	64
IV	13	3×5.5	96
V	33	4×5.5	128
VI	7	4×7.5	128

where $P_{h \text{ real}}$ and $P_{h \text{ est}}$ represent the real and estimated values for the load of water distribution system at the hour h .

The mean absolute percentage error (MAPE) from (8) is dimensionless and thus it can be used to compare the accuracy of the model on different data sets.

6. Case Study

In order to show the characteristics of the proposed method for assessment of electrical energy consumption, a real water distribution system with $N = 73$ hydrophore stations is considered. For this system, the input information refers to the technical characteristics (Table 2) and the hourly load patterns of the hydrophore stations.

Thus, in all stations, three or four similar force pumps having the rated power between 2.2 and 7.5 kW are installed. The required water flow is delivered at a constant pressure, by changing the frequency of the source supplying the electrical engines of the force pumps. The load patterns are represented by load profiles of the water hydrophore stations. The measurements of individual load profiles were performed using an electronic meter. A sensor and an electronic device for pulse counting and data storage compose this meter. These profiles were processed for the day when it registered the maximum load in the water distribution system. The time interval is defined by taking hourly steps within a day ($T = 24$ and $\Delta t_h = 1$ hour).

The normalization of the load profiles was made in relation to daily electrical energy consumption. Further, the optimal number of clusters was determined using the algorithm described in Section 2. Getting started, the maximum of clusters $K_{\max}$ was calculated ($K_{\max} = \sqrt{N} \approx 9$). Then, for the set of normalized active power profiles, the K -means clustering method with a given K ($2 \leq K \leq K_{\max}$) is used. Finally, the silhouette global coefficient is calculated for the assessment of partition. Because the silhouette global coefficient has the highest value for $K = 3$, this represents the optimal solution for clustering process, Figure 2. For this solution, the silhouette plot is presented in Figure 3.

The characteristics of clusters (p_h^{Ci} and σ_h^{Ci} , $h = 1, \dots, 24$, $i = 1, \dots, 3$) are presented in Table 3. After aggregation of normalized load profiles of each cluster, Figure 4, the representative profiles were determined. Representative load profile for each cluster is obtained by averaging the values for each hour represented by load factors p_h^{Ci} , $i = 1, \dots, 3$.

FIGURE 4: Clustering results with 3 clusters ((a) cluster C1, (b) cluster C2, (c) cluster C3).

TABLE 3: Hourly coefficients of representative load profiles.

Hour	p^{C1} [p.u.]	p^{C2} [p.u.]	p^{C3} [p.u.]	σ^{C1} [p.u.]	σ^{C2} [p.u.]	σ^{C3} [p.u.]
1	0.026	0.022	0.033	0.0028	0.0053	0.0041
2	0.023	0.020	0.031	0.0036	0.0043	0.0039
3	0.021	0.019	0.030	0.0032	0.0040	0.0040
4	0.022	0.019	0.030	0.0033	0.0036	0.0036
5	0.022	0.020	0.031	0.0029	0.0042	0.0042
6	0.026	0.025	0.035	0.0025	0.0044	0.0039
7	0.036	0.040	0.041	0.0047	0.0040	0.0028
8	0.050	0.054	0.047	0.0089	0.0072	0.0048
9	0.054	0.048	0.045	0.0045	0.0055	0.0048
10	0.055	0.048	0.046	0.0062	0.0058	0.0045
11	0.053	0.046	0.044	0.0100	0.0036	0.0033
12	0.048	0.043	0.043	0.0040	0.0040	0.0027
13	0.051	0.041	0.042	0.0056	0.0028	0.0037
14	0.054	0.041	0.041	0.0064	0.0034	0.0030
15	0.051	0.040	0.041	0.0059	0.0028	0.0037
16	0.047	0.044	0.043	0.0047	0.0028	0.0026
17	0.046	0.048	0.045	0.0039	0.0045	0.0030
18	0.047	0.055	0.047	0.0072	0.0041	0.0049
19	0.046	0.060	0.050	0.0061	0.0058	0.0057
20	0.050	0.063	0.051	0.0067	0.0066	0.0059
21	0.056	0.070	0.056	0.0072	0.0077	0.0082
22	0.047	0.058	0.048	0.0076	0.0073	0.0058
23	0.039	0.048	0.043	0.0039	0.0049	0.0037
24	0.032	0.030	0.037	0.0030	0.0043	0.0037

TABLE 4: Results of clustering process.

Cluster	Number of stations	[%]	Type of stations						W_{med} [kWh]
			I	II	III	IV	V	VI	
C1	24	32.88			4	2	11	7	41.42
C2	8	12.33	3	5	1				28.58
C3	41	54.79		1	6	11	22		32.35
Total	73	100	3	6	11	13	33	7	34.11

FIGURE 5: Representative load profile for cluster C1.

FIGURE 7: Representative load profile for cluster C3.

FIGURE 6: Representative load profile for cluster C2.

Graphical representation of the representative load profiles corresponding to the three clusters obtained (C1, C2, and C3) is given in Figures 5, 6, and 7.

One hourly value from the representative load profiles denotes the load of a water hydrophore station in per unit of the total average daily load of this station. The hourly load pattern can be employed to approximate the load pattern of any water hydrophore station within the same cluster.

The results of calculations carried out during the first stage of the clustering procedure are presented in Table 4.

From Table 4 it can be seen that the most consistent clusters are C1 and C3, which together accounted for about 85% of the total load profiles of the water hydrophore stations. In terms of technical characteristics, the water hydrophore stations from cluster C2 belong to types I and II (installed rated power is less than 8.8 kW); the stations from clusters C1 and C3 belong to types III–VI (installed rated power is

TABLE 5: Real and estimated hourly load of water distribution system.

Hour	P_{real}^S [kW]	P_{est}^S [kW]	Err. [%]
1	71.05	72.95	2.68
2	64.49	66.95	3.81
3	63.50	65.34	2.90
4	63.23	65.19	3.10
5	65.14	67.27	3.26
6	75.85	78.61	3.64
7	107.44	105.16	2.12
8	127.37	130.96	2.82
9	118.80	122.11	2.79
10	119.89	123.57	3.07
11	115.29	119.09	3.29
12	109.36	113.26	3.57
13	106.22	109.95	3.51
14	107.96	110.44	2.30
15	110.15	108.64	1.37
16	109.78	113.11	3.04
17	114.28	119.76	4.79
18	125.28	130.13	3.87
19	137.02	140.12	2.26
20	139.93	145.00	3.62
21	165.31	160.36	3.00
22	131.87	136.24	3.31
23	111.54	115.69	3.72
24	85.50	87.94	2.85

between 16 and 30 kW). But, in cluster C3 approximately 50 percent of the total stations have an installed rated power by

FIGURE 8: Real and estimated load profiles of the analyzed water distribution system.

FIGURE 9: Estimation errors obtained for hourly load of analyzed water distribution system.

22 kW. In terms of the electrical energy consumption, it can be observed that the highest medium value is registered for cluster C1, and the lowest value is registered for cluster C2.

In the second step of the study, using information from Tables 3 and 4 and relation (6), the hourly loads of the analyzed water distribution system were obtained. Thus, in Table 5, Figures 8 and 9, the real loads, estimated loads, and estimated errors are presented. The estimation maximum error is 4.79%, and minimum error is 1.37%. Global results show that the value for MAPE is 3.11%.

This may be considered a very good result, especially if we take into account the arbitrariness inherent to loads behavior and not all hydrophore stations have monitoring system.

7. Conclusions

The investigation of combined actions, which account for connected water and electrical aspects, allows improving the global efficiency of the water distribution systems and supplying the consumers using the least possible amount of water and energy.

In water distribution companies, the system operator predicts the consumption for today and tomorrow based on

recent consumption trends, weather forecasting, day of the week, knowledge of future events, and historical knowledge of utility system performance. This approach conducts at a large percentage error (more than 5%) for assessment of load.

The proposed method is based on two stages in which the *K*-means clustering method and a load simulation technique are exploited for estimation of the electrical energy consumption. The method based on the *K*-means clustering algorithm was used for determination of the representative load profiles of the hydrophore stations. A comparison of obtained results using the proposed method with the real registered data indicates an error by 3.11%, which is very close to the expected one by water companies (2.5–3%) [57].

Results obtained demonstrate the ability of the proposed method to become the first step in an efficient management of the water distribution systems. On the one hand, the representative load profiles can be used to model overall water company demands in a way that show how changes in use by one category affect the hourly load profiles for the system as a whole and on the other hand these profiles can contribute to a better understanding of the opportunities for linking water-efficiency and energy-efficiency programs.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- [1] Alliance to Save Energy, "What watery involves," 2014, <http://www.watery.net>.
- [2] P. Boulous, Z. Wu, C. Orr, M. Moore, P. Hsiung, and D. Thomas, *Optimal Pump Operation of Water Distribution Systems Using Genetic Algorithms*, H₂ONET-Users Guide, MW Software Consulting, Dallas County, Tex, USA, 2000.
- [3] European Commission, Directorate General for Environment, 2013, http://ec.europa.eu/environment/water/index_en.htm.
- [4] A. Anton and S. Perju, "Monitoring the main parameters of a water supply pumping station over ten years," in *Proceedings of the 6th International Conference on Hydraulic Machinery and Hydrodynamics*, Timisoara, Romania, 2004.
- [5] S. J. Kenway, A. Priestley, S. Cook et al., "Energy use in the provision and consumption of urban water in Australia and New Zealand," Water for a Healthy Country National Research Flagship Report, CSIRO, 2008.
- [6] L. House, *Water Supply Related Electricity Demand in California*, Report for California Energy Commission, 2006.
- [7] R. Goldstein and W. Smith, "Water and sustainability: U.S. electricity consumption for water supply & treatment—the next half century," EPRI Report, Electric Power Research Institute, Palo Alto, Calif, USA, 2000.
- [8] I. Pulido-Calvo and J. C. Gutiérrez-Estrada, "Selection and operation of pumping stations of water distribution systems," *Environmental Research Journal*, vol. 5, no. 3, pp. 1–20, 2011.
- [9] K. Feldman, "Aspects of energy efficiency in water distribution systems," in *Proceedings of the International Conference on Water Loss*, pp. 26–30, IWA, Cape Town, South Africa, April 2009.

- [10] G. C. Ionescu and D. S. Ionescu, "The optimization of energy consumption in water supply systems," *Acta Electrotehnica*, vol. 46, pp. 191–194, 2005.
- [11] M. Istrate and G. Grigoras, "Energy consumption estimation in water distribution systems using fuzzy techniques," *Environmental Engineering and Management Journal*, vol. 9, no. 2, pp. 249–256, 2010.
- [12] I. Sârbu, *Energetic Optimization of the Water Distribution Systems*, Publishing House of the Romanian Academy, Bucharest, Romania, 1997.
- [13] L. Szychta, "Energy consumption of water pumping for selected control systems," *Electrical Power Quality and Utilization Journal*, vol. 12, pp. 21–27, 2006.
- [14] G. Grigoras and M. Istrate, "An efficient clustering method in evaluation of the electric energy consumption from water hydrophore stations," *International Review on Modelling and Simulations*, vol. 4, no. 2, pp. 813–818, 2011.
- [15] U. Zessler and U. Shamir, "Optimal operation of water distribution systems," *Journal of Water Resources Planning and Management*, vol. 115, no. 6, pp. 735–752, 1989.
- [16] P. Cutore, A. Campisano, Z. Kapelan, C. Modica, and D. Savic, "Probabilistic prediction of urban water consumption using the SCEM-UA algorithm," *Urban Water Journal*, vol. 5, no. 2, pp. 125–132, 2008.
- [17] G. Chicco, "Overview and performance assessment of the clustering methods for electrical load pattern grouping," *Energy*, vol. 42, no. 1, pp. 68–80, 2012.
- [18] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," <http://dataclustering.cse.msu.edu/papers/MSU-CSE-00-16.pdf>.
- [19] R. Siddheswar and R. Turi, "Determination of number of clusters in K-Means clustering and application in color image segmentation," in *Proceedings of the 4th International Conference on Advanced in Pattern Recognition and Digital Techniques*, Calcutta, India, 1999.
- [20] I. Yatskiv and L. Gusarova, "The methods of cluster analysis results validation," in *Proceedings of International Conference RelStat'04*, Riga, Latvia, 2004.
- [21] J. Yu and Q. Cheng, "The upper bound of the optimal number of clusters in fuzzy clustering," *Science in China Series F*, vol. 44, pp. 119–124, 2001.
- [22] "A tutorial on clustering algorithms," <http://home.dei.polimi.it/matteucc/Clustering/tutorial.html>.
- [23] 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, 1987.
- [24] M. Halkidi, Y. Batistakis, and M. Vazirgiannis, "On clustering validation techniques," *Journal of Intelligent Information Systems*, vol. 17, no. 2–3, pp. 107–145, 2001.
- [25] M. R. Rezaee, B. P. F. Lelieveldt, and J. H. C. Reiber, "A new cluster validity index for the fuzzy C-mean," *Pattern Recognition Letters*, vol. 19, no. 3–4, pp. 237–246, 1998.
- [26] P. Berkhin, "Survey of clustering data mining techniques," Tech. Rep., Accrue Software, San Jose, Calif, USA, 2002.
- [27] M. Hølgersson, "The limited value of cophenetic correlation as a clustering criterion," *Pattern Recognition*, vol. 10, no. 4, pp. 287–295, 1978.
- [28] A. D. Gordon, *Classification*, Chapman & Hall, New York, NY, USA, 2nd edition, 1999.
- [29] C. G. Carter-Brown, "Load profile modeling for integrated energy planning," in *Proceedings of the Domestic Use of Electrical Energy Conference*, pp. 13–18, Cape Town, South Africa, April 1999.
- [30] British Electricity Boards, "Report on the design of low voltage underground networks for new housing," ACE Report No. 105, 1986.
- [31] British Electricity Boards, "Report on the computer program DEBUTE for the design of LV radial networks. Part 1—general considerations , part 2—program user guide," Tech. Rep. 115, 1988.
- [32] A. Seppälä, *Load research and load estimation in electricity distribution [Ph.D. thesis]*, Helsinki University of Technology, Esbo, Finland, 1996.
- [33] J. Nazarko and Z. A. Styczynski, "Application of statistical and neural approaches to the daily load profiles modelling in power distribution systems," in *Proceedings of the IEEE/PES Transmission and Distribution Conference*, pp. 320–325, New Orleans, La, USA, May 1999.
- [34] Z. Zakaria, M. N. Othman, and M. H. Sohod, "Consumer load profiling using fuzzy clustering and statistical approach," in *Proceedings of the 4th Student Conference on Research and Development (SCORED '06)*, 274, p. 270, IEEE, Selangor, Malaysia, June 2006.
- [35] I. H. Yu, J. K. Lee, J. M. Ko, and S. I. Kim, "A method for classification of electricity demands using load profile data," in *Proceedings of the 4th Annual ACIS International Conference on Computer and Information Science (ICIS '05)*, pp. 164–168, Jeju Island, South Korea, July 2006.
- [36] G. Chicco, R. Napoli, and F. Piglion, "Comparisons among clustering techniques for electricity customer classification," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 933–940, 2006.
- [37] G. Chicco, R. Napoli, and F. Piglion, "Application of clustering algorithms and self organising maps to classify electricity customers," in *Proceedings of the IEEE Bologna PowerTech Conference*, vol. 1, Bologna, Italy, June 2003.
- [38] D. Gerbek, S. Gasperic, and F. Gubina, "Comparison of different classification methods for the consumers' load profile determination," in *Proceedings of the 17th International Conference on Electricity Distribution*, Barcelona, Spain, 2003.
- [39] G. Grigoras, M. Istrate, and F. Scarlatache, "Electrical energy consumption estimation in water distribution systems using a clustering based method," in *Proceedings of the International Conference on Electronics, Computers and Artificial Intelligence (ECAI '13)*, pp. 1–6, Pitesti, Romania, June 2013.
- [40] N. Mahmoudi-Kohan, M. P. Moghaddam, and S. M. Bidaki, "Evaluating performance of WFA K-means and modified follow the leader methods for clustering load curves," in *Proceedings of the IEEE/PES Power Systems Conference and Exposition (PSCE '09)*, Seattle, Wash, USA, March 2009.
- [41] G. Grigoras, M. Istrate, and F. Scarlatache, "Electrical energy consumption estimation in water distribution systems using a clustering based method," in *Proceedings of the International Conference on Electronics, Computers and Artificial Intelligence (ECAI '13)*, pp. 1–6, Pitesti, Romania, June 2013.
- [42] A. Mutanen, M. Ruska, S. Repo, and P. Järventausta, "Customer classification and load profiling method for distribution systems," *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1755–1763, 2011.
- [43] G. Grigoras, C. Barbulescu, G. Cartina, and D. Comanescu, "A comparative study regarding efficiency of the hierarchical clustering techniques in typical load profiles determination," *Acta Electrotehnica*, vol. 52, no. 5, pp. 184–188, 2011.

- [44] B. D. Pitt and D. S. Kirschen, "Application of data mining techniques to load profiling," in *Proceedings of the 21st IEEE International Conference Power Industry Computer Applications*, Santa Clara, Calif, USA, 1999.
- [45] S. Ramos and Z. Vale, "Data mining techniques application in power distribution utilities," in *Proceedings of the IEEE/PES Transmission and Distribution Conference and Exposition*, pp. 1–8, Chicago, Ill, USA, April 2008.
- [46] V. Figueiredo, F. Rodrigues, Z. Vale, and J. B. Gouveia, "An electric energy consumer characterization framework based on data mining techniques," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 596–602, 2005.
- [47] A. H. Nizar, Z. Y. Dong, and J. H. Zhao, "Load profiling and data mining techniques in electricity deregulated market," in *Proceedings of the IEEE Power Engineering Society General Meeting (PES '06)*, Montreal, Canada, June 2006.
- [48] G. J. Tsekouras, N. D. Hatziaargyriou, and E. N. Dialynas, "Two-stage pattern recognition of load curves for classification of electricity customers," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1120–1128, 2007.
- [49] S. V. Verdú, M. O. García, F. J. G. Franco et al., "Characterization and identification of electrical customers through the use of self-organizing maps and daily load parameters," in *Proceedings of the IEEE PES Power System Conference and Exposition*, vol. 2, pp. 899–906, Atlanta, Ga, USA, 2004.
- [50] K. L. Lo and Z. Zakaria, "Electricity consumer classification using artificial intelligence," in *Proceedings of the 39th International Universities Power Engineering Conference*, Bristol, UK, 2004.
- [51] D. Gerbec, S. Gašperič, I. Šmon, and F. Gubina, "Determining the load profiles of consumers based on fuzzy logic and probability neural networks," *IEE Proceedings: Generation, Transmission and Distribution*, vol. 151, no. 3, pp. 395–400, 2004.
- [52] D. Gerbec, S. Gašperič, I. Šmon, and F. Gubina, "Allocation of the load profiles to consumers using probabilistic neural networks," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 548–555, 2005.
- [53] M. Sarlak, T. Ebrahimi, and S. S. Karimi Madahi, "Enhancement the accuracy of daily and hourly short time load forecasting using neural network," *Journal of Basic and Applied Scientific Research*, vol. 2, no. 1, pp. 247–255, 2012.
- [54] H. K. Alfares and M. Nazeeruddin, "Electric load forecasting: Literature survey and classification of methods," *International Journal of Systems Science*, vol. 33, no. 1, pp. 23–34, 2002.
- [55] L. A. Garcia-Escudero and A. Gordaliza, "A proposal for robust curve clustering," *Journal of Classification*, vol. 22, no. 2, pp. 185–201, 2005.
- [56] Z. Zisman and G. Cartina, "Application of fuzzy logic for distribution system estimation," in *Proceedings of the 22nd Seminar on Fundamentals of Electrotechnics and Circuit Theory*, Ustroń, Poland, May 1999.
- [57] L. Jentgen, H. Kidder, R. Hill, and S. Conrad, *Water Consumption Forecasting to Improve Energy Efficiency of Pumping Operations*, Awwa Research Foundation, 2007.