UNSUPERVISED LEARNING-BASED METHODOLOGY TO ANALYZE THE VOLTAGE QUALITY IN THE LOW-VOLTAGE ELECTRIC DISTRIBUTION NETWORKS

METODOLOGIE BAZATĂ PE ALGORITMI DE INVĂȚARE NESUPRAVEGHEATĂ PENTRU ANALIZA CALITĂȚII TENSIUNII IN REȚELELE ELECTRICE DE DISTRIBUȚIE DE JOASĂ TENSIUNE

Ștefania GĂLBĂU[1], Gheorghe GRIGORAȘ[2], Bogdan-Constantin NEAGU[3]

DOI: 10.37410/EMERG.2025.2.01

Abstract: *The paper addresses the voltage quality issue, which plays a key role in describing the technical and economic performance of low-voltage electrical distribution networks (LV-EDNs) and how they affect energy efficiency at the level of activities performed by end users. Given the challenges faced by distribution network operators in defining a global performance indicator for evaluating voltage quality due to the diverse technical characteristics of LV-EDNs, the authors proposed an efficient methodology integrating unsupervised learning algorithms based on clustering to recognise the voltage quality categories applicable to LV nodes. The methodology was tested on an LV-EDN with 37 nodes in a suburban, urban area supplying 53 end-users and 26 prosumers, with active and reactive power profiles obtained from smart meters. The steady-state calculations lead to phase voltage profiles for the analysed period (one week). The clustering process had three clusters (patterns) for the nodes from the network, and regardless of the phase, the voltage quality categories "Unsatisfactory" and "Satisfactory" were assigned for each node. As a technical measure, the On-load Tap Changer is used to improve voltage quality, and the classification of each node has been changed within the "Very Good" and "Good" patterns.*

[1] PhD Student, Dept. of Power Engineering, “Gheorghe Asachi” Technical University of Iasi, Romania, e-mail: stefania-loredana.galbau@student.tuiasi.ro
[2] Prof., Dept. of Power Engineering, “Gheorghe Asachi” Technical University Iasi, Romania, e-mail: gheorghe.grigoras@academic.tuiasi.ro, *ORCID ID (ex: 0000-0003-3811-6298)*
[3] Assoc. Prof., Dept. of Power Engineering, “Gheorghe Asachi” Technical University of Iasi, Romania, e-mail: bogdan-constantin.neagu@academic.tuiasi.ro, *ORCID ID (ex: 0000-0002-8413-0317)*

Keywords: voltage quality; electrical distribution networks; unsupervised learning; clustering.

Rezumat: *Lucrarea tratează problema calității tensiunii care are un rol semnificativ în caracterizarea performanţelor tehnico-economice a rețelelor electrice de distribuție de joasă tensiune (RED-JT), influențând eficiența energetică a activităților desfășurate la nivelul utilizatorilor finali. În contextul în care este dificil pentru operatorii de distribuție a energiei electrice să definească un indicator global de performanță pentru evaluarea calității tensiunii din cauza diferitelor caracteristici tehnice ale RED-JT, autorii au propus o metodologie eficientă ce integrează algoritmi de învățare nesupravegheată bazați pe clustering pentru a putea recunoaște categoriile de calitate a tensiunii în care se încadrează nodurile RED-JT. Metodologia a fost testată într-o RED-JT cu 37 de noduri dintr-o zonă urbană periferică care alimentează 53 de utilizatori finali, dintre care 26 sunt prosumatori, pentru care au fost atașate profiluri de putere activă și reactivă provenite de la contoarele inteligente. Profilurile tensiunii de fază au fost obținute din calculele de regim permanent efectuate pentru fiecare eșantion de timp din perioada analizată (o săptămână). Au fost identificate 3 clustere pentru care, indiferent de faza de alimentare, nodurile au fost în cadrate în categoriile de calitate a tensiunii Nesatisfăcător și Satisfăcător. Ca măsură tehnică pentru îmbunătățirea calitatea tensiunii s-a folosit un transformator prevăzut cu un comutator de reglaj sub sarcină, iar clasificarea fiecărui nod a fost schimbată în categoriile de calitate ”Foarte bine” și “Bine”.*

Cuvinte cheie: calitatea tensiunii; rețelele electrice de distribuție; învățare nesupravegheată; clustering

1. Introduction

The rapid development of low-voltage electric distribution networks (LV-EDNs) in recent years and the emergence of the microgrid concept have made power quality one of the most important objectives that Distribution Network Operators (DNOs) must consider in their management strategies and decision-making processes [1], [2].

Electrical energy quality indicators, in general, and supply voltage quality indicators, in particular, play a significant role in characterising the technical and economic performance associated with the operation and development of LV-EDNs. They considerably influence the technical and economic efficiency of activities carried out by end-users [3]. End-users include both consumers and prosumers. Prosumers are individuals or entities that consume energy from the electrical grid and choose to produce and supply surplus energy generated at their premises through renewable sources, with a maximum installed capacity of 400 kW per consumption point [4].

The technical quality of the supply voltage has become increasingly important due to the growing number of nonlinear consumers, the sensitivity of specific consumers (such as computer networks and microprocessors), and the integration of a significant number of single-phase prosumers. [5], [6].

Each DNO must monitor voltage quality using certified equipment, such as network analysers, which evaluate key indicators related to voltage amplitude, transient interruption frequency, and voltage sags. In addition to these indicators, the decision-makers can use metrics associated with flicker phenomena and current harmonics to quantify them [7].

In recent years, smart meters have been installed at each end-user's location, providing DNOs with detailed information from every installation point. This data enables the calculation of the previously mentioned specific indicators [7], [8].

Due to the multitude of specific indicators calculated for each node, it is challenging for DNOs to characterise voltage quality across LV-EDNs. Thus, defining representative voltage quality patterns/classes/models, characterised by typical indicators calculated based on information from smart meters, poses a significant challenge for EDOs, as it requires working with large-scale databases. In these conditions, unsupervised learning algorithms, such as clustering, can provide solutions for DNOs to identify the voltage quality models/patterns into which each LV node in the EDN [9] – [12].

In this context, the paper proposes an efficient methodology that integrates unsupervised learning algorithms based on clustering to recognise the voltage quality patterns in the LV nodes of EDNs. The input features for the clustering process refer to the specific indicators of slow voltage variations calculated at each LV node in the EDN for each of the three phases over the analysed period: the phase voltage deviation, the mean of phase voltage deviation, the variance of phase voltage deviation, and the mean squared value of phase voltage deviation. The advantages of the proposed approach are related to the fact that each LV node will be assigned a voltage quality pattern that integrates all three typical indicators of slow voltage variations.

The structure of the paper is as follows: Section 2 presents the typical indicators of voltage variations, Section 3 describes the evaluating procedure of voltage quality in LV-EDNs using clustering-based unsupervised learning, Section 4 includes the results obtained in a real case study, and Section 5 provides the conclusions.

2. Typical indicators of phase voltage variations

The DNOs can identify the slow voltage variations in the network elements due to operating conditions influenced by changes in active and reactive power flows resulting from the dynamic load demand of end-users and the power injections into the grid at the common boundary point of prosumers [6].

Voltage variations in the LV-EDNs, to which end-users (consumers and prosumers) are connected, can cause changes in the production of the industrial consumers, affect their energy performance, lead to rapid equipment wear, and disrupt technological processes.

The calculation relations of the typical indicators characterising slow phase voltage variations at the level of each LV node from the EDNs are as follows [6], [7], [12], [13].

- ***Phase Voltage Deviation (PVD):***

$$PVD_{m,t}^{\{p\}} = \frac{V_{m,t}^{\{p\}} - V_r}{V_r} \cdot 100, \qquad [\%] \qquad m = 1, \dots N_m;$$
$$t = 1, \dots, T; \quad \{p\} \in \{a, b, c\} \tag{1}$$

where: N_m – the number of LV nodes in the analysed EDN; T – the analysis period (day, week, year); $V^{\{p\}}_{m,t}$ – the amplitude of the phase voltage measured on the phase p, in the LV node m, at the time t within the analysis period T; V_r – the rated voltage on each phase of the analysed LV-EDN.

- ***Mean of Phase Voltage Deviation (MPVD):***

$$MPVD_m^{\{p\}} = \frac{1}{T}\int_0^T \Delta V_{m,t}^{\{p\}}\, dt = \frac{100}{T}\int_0^T \frac{V_{m,t}^{\{p\}} - V_r}{V_r} dt, [\%]$$
$$m = 1, \dots, N_m; \ \{p\} \in \{a, b, c\} \tag{2}$$

- ***Variance of Phase Voltage Deviation (VPVD):***

$$VPVD_m^{\{p\}} = \frac{1}{T}\int_0^T \left(\Delta V_{m,t}^{\{p\}} - MPVD_m^{\{p\}}\right)^2 dt, [\%] \quad m = 1, \dots, N_m;$$
$$\{p\} \in \{a, b, c\} \tag{3}$$

- ***Mean Squared Value of Phase Voltage Deviation (MSVVPD):***

$$MSVVPD_m^{\{p\}} = \frac{1}{T}\int_0^T \left(\Delta V_{m,t}^{\{p\}}\right)^2 dt, [\%] \quad m = 1, \dots, N_m; \quad \{p\} \in \{a, b, c\} \tag{4}$$

3. Methodology integrating clustering-based unsupervised learning for voltage quality analysis

Understanding the complexity of voltage quality in LV-EDNs is essential to analyse the relationships between the specific indicators that define it. In this context, unsupervised learning can define these relationships (see Fig. 1) [14]. Clustering, a subset of unsupervised learning, is a powerful tool that reveals patterns in data, predicting future or unknown data from a training dataset [10], [12], [14], [15].

Clustering refers to special techniques that group the input data represented as vectors. To study the similarity/difference between elements associated with a data set and to group them, each element corresponds to a vector whose components are associated with the representative features/ attributes. These features/attributes could be voltage magnitude, frequency, or harmonic content. Within the clustering process, the vectors are grouped based on the distance calculated between them. At the end of the clustering process, one or more clusters (groups/patterns/models/classes) will be obtained, depending on the studied problem. These clusters represent the spatial arrangement of the considered features for the elements involved in the process [10], [12].

The clustering methods, regardless of whether they belong to the hierarchical clustering or K-means category, lead to the iterative reduction of the number of elements from the database, grouping them into representative clusters (patterns/models/classes).

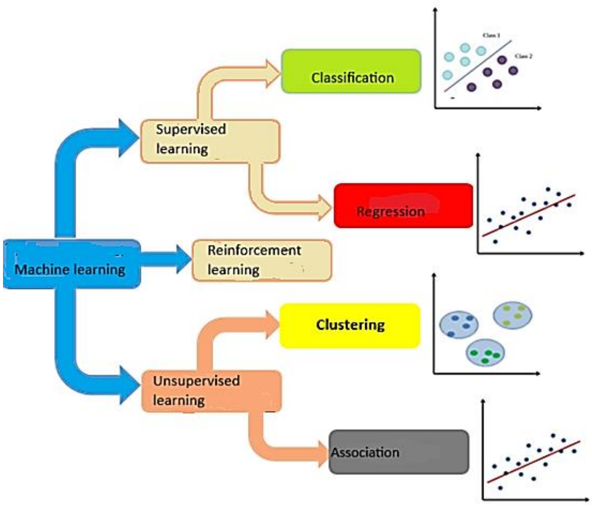

Figure 1. Classification of Machine Learning techniques, adapted from [14]

Concerning the K-means, the algorithm partitions the data into K clusters based on various features (in this case, the typical indicators of slow voltage variations presented in the previous paragraph) of the elements (LV nodes in the analysed EDN), attempting to determine the centre of the clusters. The algorithm's objective is to minimise the Euclidean distance between each element (corresponding to an LV node in the EDN), characterised by a vector from the initial database, and assign it to the closest cluster (referred to as patterns in our study) [9], [12].

$$F = min\sum_{k=1}^{K}\sum_{m=1}^{N_m^k}\left\|n_m^k - c_k\right\|^2 \tag{5}$$

where: $\left\| n_m{}^k - c_k \right\|^2$ - the distance calculated between the vectors associated with the LV nodes $n^k{}_m$ in the EDN and the centre c_k of pattern k, where $k = 1$, …, K and $m = 1$, …, $N^k{}_m$; $N_m{}^k$ - the number of LV nodes from the EDN assigned to pattern k, where $k = 1, \ldots, K$; K – the number of patterns.

The verification of the obtained partitions will be done at the end of the clustering process using a performance indicator called the silhouette coefficient [1], [9]:

$$SC = \frac{1}{K}\cdot\sum_{k=1}^{K}\left(\frac{1}{N_m^k}\cdot\sum_{m=1}^{N_m^k}\left(\frac{D(t_m^k) - h(t_m^k)}{\max\{D_m - h_m\}}\right)\right) \tag{6}$$

where: $D(t_m{}^k)$ - the average value of the distance between LV node $n_m{}^k$, $m = 1, \ldots, N_m{}^k$, and the elements of the same pattern k, $k = 1, \ldots, K$; $h(t_m{}^k)$ - represents the minimum average value associated with the distance between the node $t_m{}^k$ and the elements of the nearest pattern k.

The proposed methodology for evaluating voltage quality at the LV nodes of an EDN, which includes unsupervised learning based on clustering, consists of two stages.

Stage I – Input Data and Information Processing

The decision-makers upload the phase voltages from the Smart Metering system database. These form the basis of a variable with a matrix structure [V] of size (N_m x (3 x E)), where N_m represents the number of LV nodes in the analysed EDN, E is the number of records corresponding to the sampling step selected for the smart meter installed at the end-user connected to node *m*, and the number 3 refers to the number of phases of the LV-EDN to which the end-users are connected. Considering the short length of the branching, the used assumption assumes that the voltage at LV node *m* is identical to the voltage measured by the smart meters installed at the end-users connected to this node.

The matrix [*V*] represents the input variable for a power quality analysis software for the LV-EDNs developed in MATLAB professional software. When the end-users do not have smart meters installed, the voltage values at the LV nodes will be determined through a load-flow calculation method proposed in [16], knowing the network topology and the phases to which the end-users are connected and then stored in the variable [*V*].

The output data refers to the matrix [*ICV*] of size (N_m x 3), where the columns are associated with specific voltage variation indicators: Mean of Phase Voltage Deviation (MPVD), Variance of Phase Voltage Deviation (VPVD), and Mean Squared Value of Phase Voltage Deviation (MSVVPD).

$$[\mathrm{ICV}] = \begin{vmatrix} MPVD_1 & VPVD_1 & MSVVPD_1 \\ \vdots & \vdots & \vdots \\ MPVD_j & VPVD_j & MSVVPD_j \\ \vdots & \vdots & \vdots \\ MPVD_m & VPVD_m & MSVVPD_m \end{vmatrix} \quad (7)$$

Stage II** – **Obtaining the voltage quality patterns

All *m* LV nodes, *m* = 1, …, N_m, from the EDN will be grouped using any clustering algorithm. In the current study, the authors selected the K-means algorithm due to its advantages over other algorithms in the hierarchical clustering category. The algorithm used the matrix [*ICV*], containing the indicators calculated inside Stage I, as its input variable.

The output data will ultimately consist of *K* clusters representing voltage quality categories. These categories will serve as the basis for decisions regarding improving measures of the voltage quality in the LV-EDN.

4. Case Study

To test the proposed methodology, an EDN with 37 LV nodes from a suburban area belonging to a Romanian DNO was considered, as illustrated in Fig. 2 [2].

The electric distribution substation, which has been put in operating since the 80s of the last century, is equipped with a 20/0.4 kV power transformer (S_r = 250 kVA) with a 3-tap changer, without load adjustment, having a voltage step of ±5% from the rated voltage supplied to the LV-EDN.

All line sections have a length of approximately 0.04 km and conductors with a cross-section of 50 mm², with characteristic parameters r_0 = 0.61 [Ω/km] and x_0 = 0.298 [Ω/km], and a maximum admissible current of 205 A. The MV bus of the electric distribution substation was considered the slack bus in the steady-state calculations.

The LV-EDN has 53 end-users, of which 26 are prosumers (24 with single-phase branching and two with three-phase branching). These end-users have active and reactive power profiles with a sampling step of 1 minute, taken from the Pecan platform [17]. The pillars highlighted with red squares have the prosumers connected (see Fig. 2). All consumers have single-phase branching.

In the first stage, knowing the requested and injected powers by the end-users, the decision-maker performs the steady-state calculations to determine the voltage values on all three phases at each LV node, m = 1, …, 37 and stored in the variable [V].

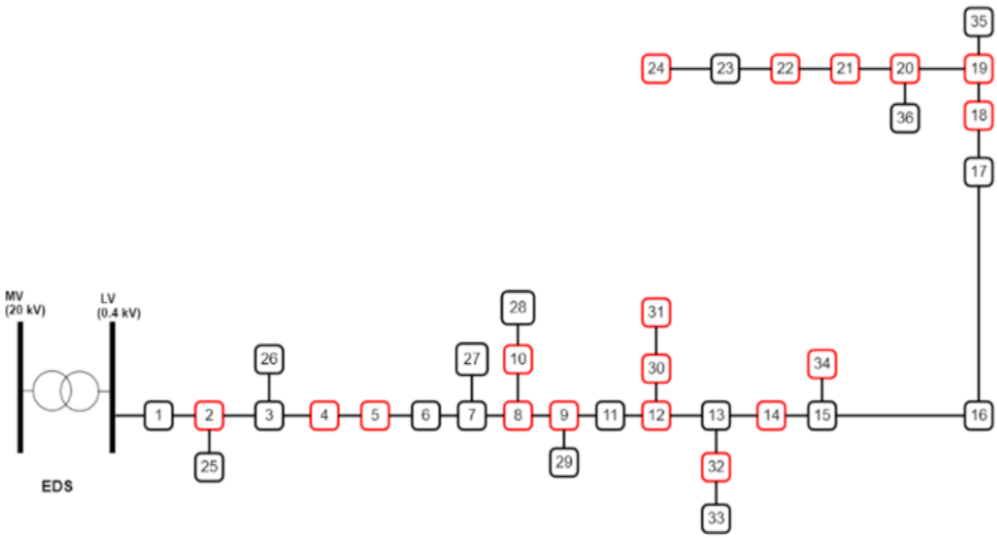

Figure 2. The test LV-EDN topology

The programming code developed in MATLAB accesses the variable [*V*] to calculate the typical voltage variation indicators recorded in the output variable [*ICV*].

In the second stage, each row in the matrix [*ICV*] represented a vector associated with each LV node *m*, *m* = 1, …, 37, included in the clustering process.

The initial maximum number of patterns is determined during the clustering process using the following relation [12]:

$$K_{max} = \sqrt{N_m} = \sqrt{37} \cong 6$$

The K-means algorithm runs for each number of patterns $k = 2, \ldots, K_{max}$. Figure 3 shows the SC values obtained for each clustering process associated with *k*, taking the values between 2 and K_{max}. The partition that led to the best results corresponds to $K_{opt} = 3$, the one for which SC = 0.8103 (as seen in Fig. 3). Figure 4 presents the values of the silhouette coefficient for the optimal partition.

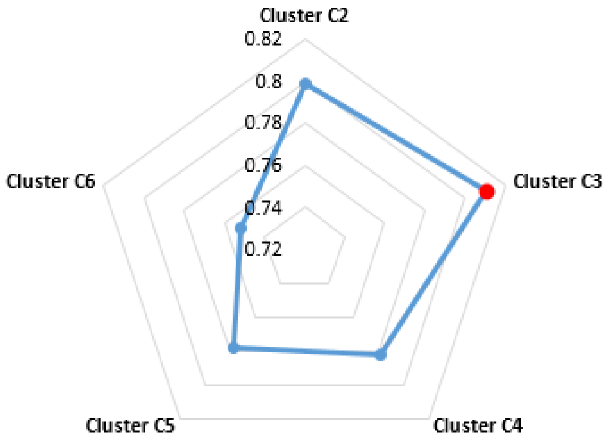

Figure 3. The values of the silhouette coefficient for each pattern $k = 2, \ldots, 6$

Each of the three patterns can be characterised by a representative element associated with a characteristic (fictive) LV node to which statistical indicators (mean and standard deviation) can be attached, calculated for each of the three quality indicators (Mean of Phase Voltage Deviation - MPVD, Variance of Phase Voltage Deviation - VPVD, and Mean Squared Value of Phase Voltage Deviation - MSVVPD, as shown in Table 1.

Table 1. **The values of statistical indicators associated with the obtained patterns**

Pattern	Statistical indicators	MPVD [%]	MPVD [%]	MPVD [%]	VPVD [%]	VPVD [%]	VPVD [%]	MSVVPD [%]	MSVVPD [%]	MSVVPD [%]
C1	Mean	5.2	5.7	4.3	55.1	65.7	36.9	41.3	52.4	29.5
	Standard Deviation	0.1	0.3	0.1	3.5	9.9	2.8	1.5	4.4	1.8
C2	Mean	4.5	4.4	3.7	34.1	29.4	22.5	31.5	35.9	22.2
	Standard Deviation	0.2	0.4	0.2	5.4	7.5	4.4	2.7	3.8	1.8
C3	Mean	3.8	4.3	3.7	8.9	7.3	6.8	19.8	24.9	18.4
	Standard Deviation	0.4	0.3	0.5	6.9	5.8	5.1	1.8	2.3	1.2

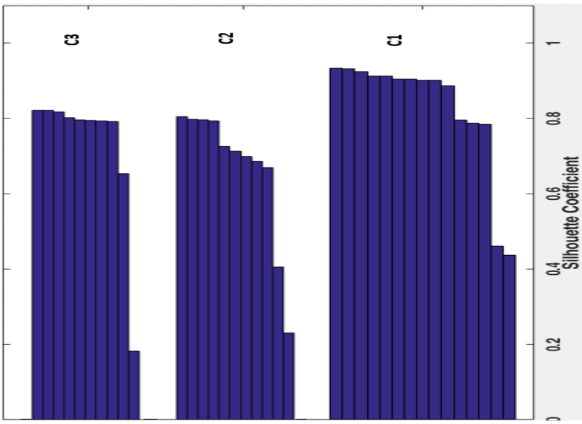

Figure 4. The values of the silhouette coefficient for the optimal partition $K_{opt} = 3$

As a result of the analysis, all three patterns are representative (C1 with 15 LV nodes, C2 with 11 LV nodes, and C3 containing the rest of the LV nodes), highlighting the excellent quality of the obtained partition. Figure 5 presents the clusters assigned to the location of the LV nodes in the EDN.

From the perspective of voltage quality across the three phases of the LV-EDN, the three patterns are associated with the following categories, as shown in Fig. 6: **Very Good** (VG), **Good** (G), **Satisfactory** (S) and **Unsatisfactory** (NS).

The findings are reinforced by the statistical representations in boxplots of the voltage deviation in all LV nodes on the three phases from Figures 7 - 9.

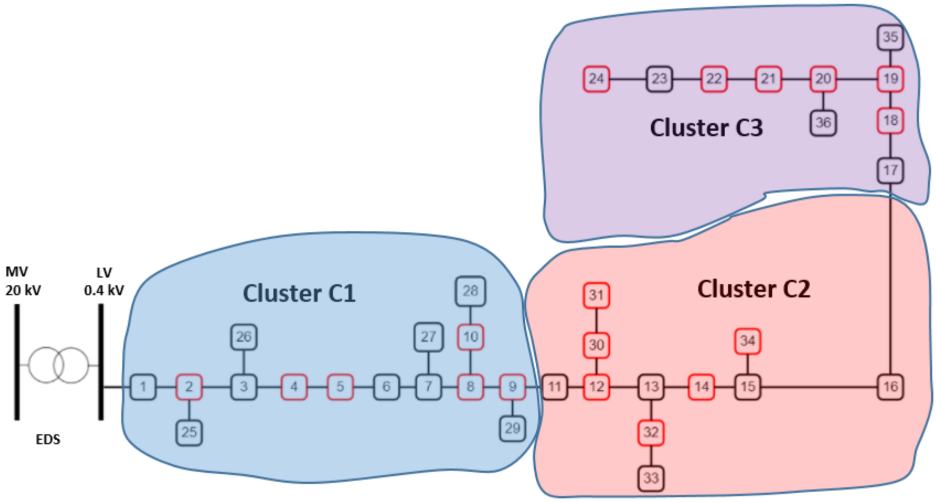

Figure 5. Association of the LV nodes from the EDN at the obtained patterns

	a	b	c
C1	S	NS	S
C2	S	S	S
C3	G	S	S

VG
G
S
NS

Figure 6. Association of the three patterns with the voltage quality categories

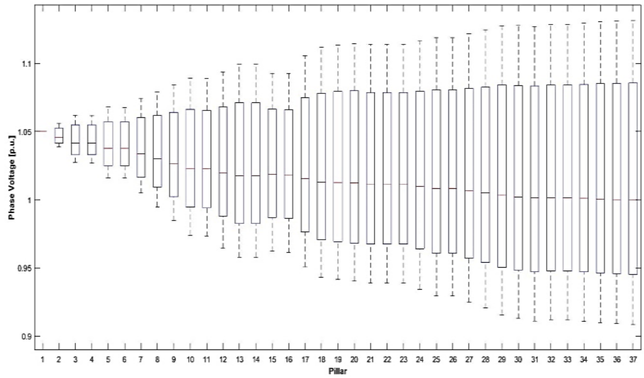

Figure 7. The boxplot representation of the voltage deviation in the analysed LV-EDN, phase a

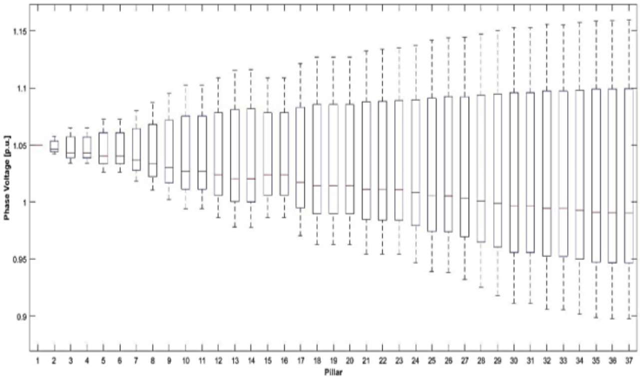

Figure 8. The boxplot representation of the voltage deviation in the analysed LV-EDN, phase b

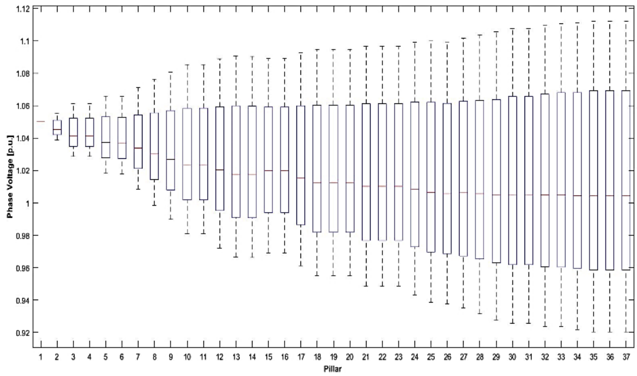

Figure 9. The boxplot representation of the voltage deviation in the analysed LV-EDN, phase c

Thus, the conclusions converge toward measures DNO must take to improve the voltage quality supplied to end-users. Some of the most effective measures include voltage control using the On Load Tap Changer (OLTC) or phase load balancing.

In the following, voltage regulation using the On Load Tap Changer (OLTC) was applied to improve the voltage quality in the LV nodes of the analysed EDN. The used assumption corresponds to replacing the old power transformer, having a low-energy efficiency standard and the lifespan has exceeded, with a new one with the same rated power (S_r = 250 kVA) having a high-energy efficiency standard following the European Eco-Design Directive [18] and On-load Tap Changer (OLTC) incorporated to maintain voltage control. The voltage tap range is between -10% and +10%, with ±8 steps. The authors used the algorithm proposed in [2] to obtain optimal voltage control according to two objectives: minimising the phase voltage deviation in all LV nodes and the power losses.

Variable [*V*] will store, in this case, the phase voltages from each LV node obtained after optimisation. Next, the decision-maker will complete the two stages of the proposed methodology. Figures 10 - 12 and Table 2 present the obtained results.

The partition that led to the best results is the same as the base case, corresponding to K_{opt} = 3, the one for which SC = 0.7689 (as seen in Fig. 10). Figure 11 presents the values of the silhouette coefficient for the optimal partition K_{opt} = 3. The analysis of Figure 12 indicates that pattern C2 contains more LV nodes than in the base case, integrating 10 LV nodes from the initial pattern C1 and one LV node from the initial pattern C3.

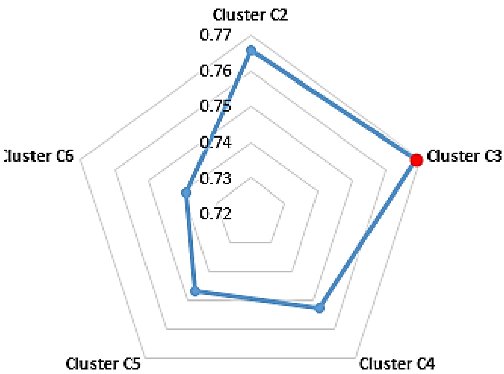

Figure 10. The values of the silhouette coefficient for each pattern *k* = 2, …, 6, case with OLTC

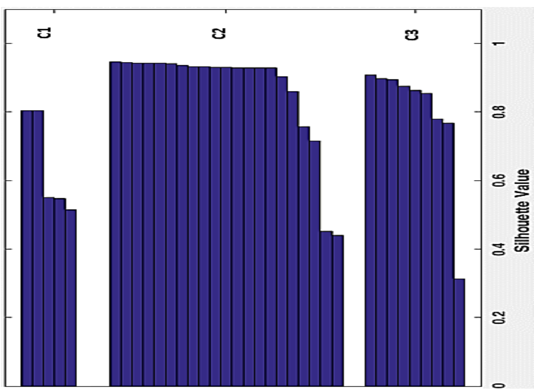

Figure 11. The values of the silhouette coefficient for the optimal partition Kopt = 3, case with OLTC

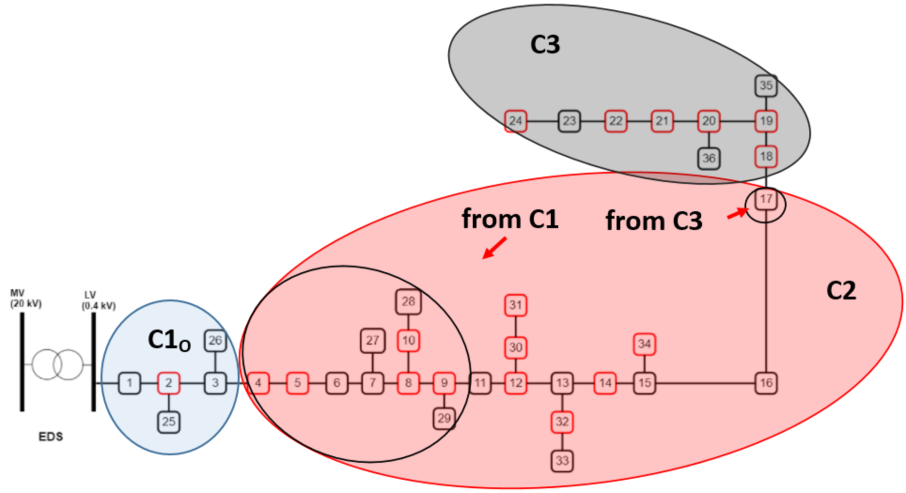

Figure 12. Association of the LV nodes from the EDN at the obtained patterns, case with OLTC

Table 2. **The values of statistical indicators associated with the obtained patterns, case with OLTC**

Pattern	Statistical indicators	$MPVD_a$ [%]	$MPVD_b$ [%]	$MPVD_c$ [%]	$VPVD_a$ [%]	$VPVD_b$ [%]	$VPVD_c$ [%]	$MSVVPD_a$ [%]	$MSVVPD_b$ [%]	$MSVVPD_c$ [%]
C1	Mean	3.0	3.3	3.1	9.8	10.4	10.9	12.8	15.4	13.6
	Standard Deviation	0.4	0.3	0.4	2.8	2.5	2.4	3.5	2.7	3.2
C2	Mean	1.3	1.8	1.2	4.3	4.5	4.1	2.9	5.5	2.9
	Standard Deviation	0.4	0.3	0.4	3.1	3.2	1.9	1.6	1.8	1.5
C3	Mean	2.4	3.1	1.7	18.1	26.2	10.9	8.1	13.8	4.9
	Standard Deviation	0.1	0.3	0.1	1.6	5.5	8.8	0.6	2.2	0.3

As a result of the analysis, all three patterns are well-defined (C1 with 4 LV nodes, C2 with 22 LV nodes, and C3 containing the rest, i.e. 9 LV nodes), indicating, once more, the high quality of the obtained partition.

The data from Table 2 for each of the three patterns indicated an improvement in voltage quality characterised by much lower values for all three indicators (Mean of Phase Voltage Deviation - MPVD, Variance of

Phase Voltage Deviation - VPVD, and Mean Squared Value of Phase Voltage Deviation – MSVVPD.

The observations concerning the analysis of the results presented in Table 2 refer to the voltage quality, which is **Good** (G) for all phases of pattern C1, **Very Good** (VG) for all phases of pattern C3, **Very Good** (G) for phases a and c and **Good** (G) for phase b of pattern C3, see Fig. 13.

The improvement of the results is reflected in the statistical results represented as boxplots where the voltage deviation on each phase in all LV nodes is reduced, see Figs 14 - 16.

	a	b	c
C1	G	G	G
C2	VG	VG	VG
C3	VG	G	VG

VG
G
S
NS

Figure 13. Association of the three patterns with the voltage quality categories, case with OLTC

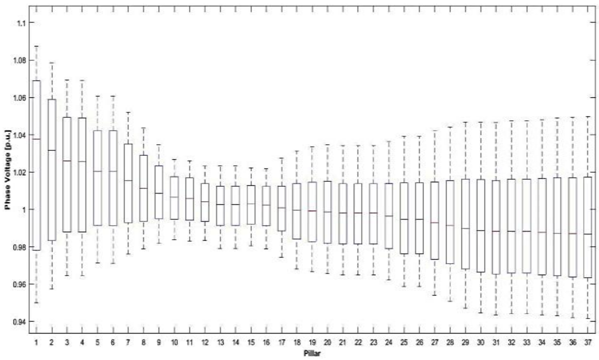

Figure 14. The boxplot representation of the voltage deviation in the analysed LV-EDN, phase a, case with OLTC

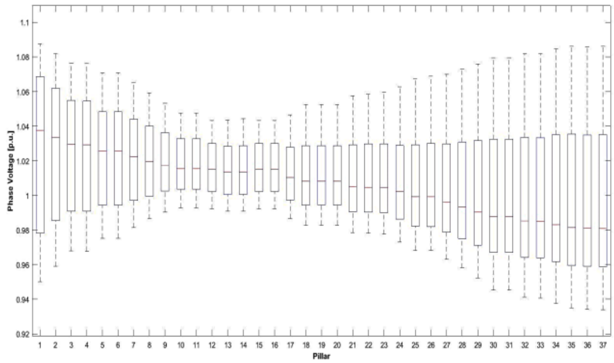

Figure 15. The boxplot representation of the voltage deviation in the analysed LV-EDN, phase b, case with OLTC

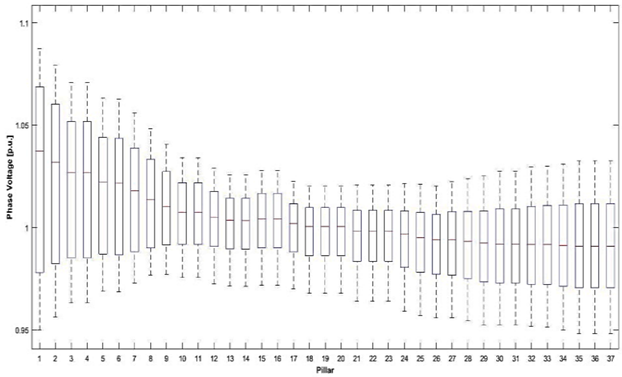

Figure 16. The boxplot representation of the voltage deviation in the analysed LV-EDN, phase c, case with OLTC

5. Conclusions

This study proposed a methodology for evaluating voltage quality in the LV-EDNs using a clustering-based unsupervised learning technique. The input features used in the clustering process referred to phase voltage indicators calculated at the level of each LV node of the EDN (the phase voltage deviation, the mean of phase voltage deviation, the variance of phase voltage deviation, and the mean squared value of phase voltage deviation).

An EDN with 37 LV nodes in a suburban area supplying 53 end-users and 26 prosumers has been used to test the methodology. Phase voltage profiles were obtained through steady-state calculations for each time interval of the analysed period (one week). Three patterns were obtained from the clustering process, classifying the LV nodes according to the **Satisfactory** and **Unsatisfactory** voltage quality categories regardless of the phase.

Using an efficient transformer equipped with an OLTC has improved the voltage quality by placing it in the **Very Good** and **Good** categories, regardless of the phase.

The proposed methodology can help the decision-makers define a global voltage quality indicator that considers typical indicators of voltage variations to characterise the LV-EDNs effectively.

R E F E R E N C E S

[1]. *I.A.Ibrahim, M.J.Hossain*, „Low Voltage Distribution Networks Modelling and Unbalanced (Optimal) Power Flow: A Comprehensive Review," IEEE Access, Vol. 9, pp. 143026-143084, 2021.

[2]. *G.Grigoras, F.Scarlatache, S.Galbau*, „An Efficient Distribution Transformer Fleet Modernization Strategy for a Rapid Transition toward the Active Electric Networks," 13th International Symposium on Advanced Topics in Electrical Engineering (ATEE), Bucharest, Romania, 2023.

[3]. V. *Ignatova, D. Villard, and J. -M. Hypolite*, "Simple Indicators for an Effective Power Quality Monitoring and Analysis," 15th International Conference on Environment and Electrical Engineering (EEEIC), Roma, Italy, 2015.

[4]. * Ordonanță de Urgență nr. 143 din 28 decembrie 2021 pentru modificarea și completarea Legii energiei electrice și a gazelor naturale nr. 13/2012, precum și pentru modificarea unor acte normative, Monitorul Oficial nr. 1259/31.12.2021 (Emergency Ordinance no. 143 of December 28, 2021 amending and supplementing the Electricity and Natural Gas Law no. 13/2012, as well as amending some normative acts, Official Gazette no. 1259/31.12.2021).

[5]. *M.Guder, O.Salor, I.Çadirci, B.Ozkan, and E.Altintas*, „Data Mining Framework for Power Quality Event Characterization of Iron and Steel Plants," IEEE Transactions on Industry Applications, Vol. 51, pp. 3521-3531. 2015.

[6]. *A. Kriukov, G. Grigoraș, M. Gavrilaș, F. Scarlatache, and B. Neagu*, „An Analyze of Slow Voltage Variations from the Electric Distribution Systems with a Clustering Based Approach," 16th International Conference on Harmonics and Quality of Power", Bucharest, Romania, 2014.

[7]. *G.Patrizi, C.Garzon Alfonso, L.Calandroni, A.Bartolini, C.Iturrino Garcia, L.Paolucci, F.Grasso, L.Ciani*, „Anomaly Detection for Power Quality Analysis Using Smart Metering Systems," Sensors, Vol. 24, 5807, 2024.

[8]. *Romanian Energy Regulatory Agency*, „Standard de performanţă pentru serviciul de distribuţie a energiei electrice," (Performance Standard for the Electricity Distribution Service)" 2021, https://www.anre.ro/ro/ energieelectrica/legislatie/norme-tehnice/ standarde-de-performanta1387201290.

[9]. *M.J. Li, M.K. Ng, Y. Cheung, and J.Z. Huang*, „Agglomerative Fuzzy K-Means Clustering Algorithm with Selection of Number of Clusters," IEEE Transactions on Knowledge and Data Engineering, Vol. 20, pp. 1519-1534, 2008.

[10]. *A.M.Ikotun, A.E.Ezugwu, L.Abualigah, B.Abuhaija, J.Heming*, „K-means Clustering Algorithms: A Comprehensive Review, Variants Analysis, and Advances in the Era of Big Data," Information Sciences, Vol. 622, pp. 178-210, 2023.
[11]. *A.Kiranmai, J.Laxmi*, „Data Mining for Classification of Power Quality Problems Using WEKA and the Effect of Attributes on Classification Accuracy," Protection and Control of Modern Power Systems, vol. 3, 29, 2018.
[12]. *S.Galbau, G.Grigoras, B.C.Neagu, F.Scarlatache, D.Lucache, V.Hustiuc*, „Characterization of the Power Quality in the Electric Distribution Networks Using Data Mining with K-Means Clustering," 10th International Conference on Systems and Control (ICSC 2022), Marseille, France, 2022.
[13]. *M. Iordache, I. Conecini*, „Calitatea energiei electrice" (Power Quality), Ed. Tehnică, Bucharest, Romania, 1997.
[14]. *D.Polzer*, „7 of the Most Used Regression Algorithms and How to Choose the Right One," Towards Data Science, 2021, https://towardsdatascience.com/7-of-the-most-commonly-used-regression-algorithms-and-how-to-choose-the-right-one-fc3c8890f9e3
[15]. *G.Grigoras, G.Cartina, M.Istrate, F.Rotaru*, „The Efficiency of the Clustering Techniques in the Energy Losses Evaluation from Distribution Networks," International Journal of Mathematical Models and Methods in Applied Sciences, Vol. 5, pp. 133-141, 2011.
[16]. *G.Grigoras, B.C.Neagu*, „Smart Meter Data-Based Three-Stage Algorithm to Calculate Power and Energy Losses in Low Voltage Distribution Networks," Energies, Vol. 12, 3008, 2019.
[17]. *Pecan Street Data Port*, www.pecanstreet.org/dataport.
[18]. *European Commission*, „Ecodesign and Energy Labelling Directive 2009/125/EC and Regulation (EU) 2017/1369," https://single-market-economy.ec.europa.eu/single-market/european-standards/harmonised-standards/ecodesign_en

Authors' biographies

Ștefania GĂLBĂU was born in Iasi, Romania, in 1994. He received the B.S. and M.S. degrees in power systems engineering from the ‘‘Gheorghe Asachi’’ Technical University of Iaşi, in 2019 and 2021, respectively. She is currently pursuing the Ph.D. degree in the Power Engineering field. Her Ph.D thesis is focused on the impact of smart technologies on power quality in electric distribution networks.

Email: stefania-loredana.galbau@student.tuiasi.ro

Gheorghe GRIGORAŞ was born in Vatra Dornei, Suceava, Romania, in 1976. He received an M.S. degree in electrical power engineering and a Ph.D. degree in power engineering from the ‘‘Gheorghe Asachi’’ Technical University of Iaşi, Romania, in 2000 and 2005, respectively. He is a professor at the Department of Power Engineering and a PhD Supervisor in the Power Engineering Field. He is the Head of the Research Laboratory “Availability of Energy Systems Supplied from Renewable Sources”, Electrical Engineering Faculty, Power System

Department. He is the author of 20 books, 14 book chapters, and more than 300 research papers. His research interests include Artificial Intelligence applications for monitoring and optimal control of power systems.
He is a Topic Editor for two journals indexed by WoS (Mathematics - ISSN: 2227-7390) and Applied Sciences - ISSN 2076-3417). Professor Grigoras was a Guest Editor of five Special Issues for peer-reviewed journals indexed by WoS. He was a General Chair of 8 international conferences and a member of over 300 technical program committees of the international symposia and conferences carried out in countries from Europe, Asia, North America, Australia, and Africa.
He is a member of the Advisory Board of the American Society for Electrical Engineering (ASEEE) and a member of the International Association of Engineers (IAENG) and the World Academy of Science, Engineering and Technology (WASET).

Email: gheorghe.grigoras@academic.tuiasi.ro

Bogdan-Constantin NEAGU (Member, IEEE) was born in Roman, Neamţ, Romania, in 1984. He received the B.S. and M.S. degrees in power engineering and the Ph.D. in electrical engineering from the ‘‘Gheorghe Asachi’’ Technical University of Iaşi in 2008 and 2014. Since 2022, he is an Associate Professor with the Department of Power Engineering and a PhD Supervisor in Power Engineering Field at ‘‘Gheorghe Asachi’’ Technical University of Iaşi. He is the author of eight books and five chapters books, and more than 150 articles. His research interests include electric energy transmission and distribution grids, microgrids, smart grids, power quality, and artificial intelligence applied in power systems.
He is a member of the International Association of Engineers (IAENG) and the World Academy of Science, Engineering and Technology (WASET).

Email: bogdan-constantin.neagu@academic.tuiasi.ro