

Optimization of the Grounding Resistance of the Neutral in 20 kV Networks According to the Criteria of Minimizing the Significant Value of the Ground Fault Current and the Overvoltage Amplitude

Kyryk V.V., Buriak A.R.

National Technical University of Ukraine "Ihor Sykorsky Kyiv Polytechnic Institute",
Kyiv, Ukraine

Abstract. The main objective of this study is to determine the optimal resistance of transformer neutral resistive grounding in 20 kV distribution networks using clustering methods. The research aims to achieve a technical balance between limiting single-phase ground fault currents, minimizing overvoltages, and reducing system sensitivity to stochastic fluctuations in network parameters. To achieve this, the following tasks were accomplished: modeling of network operating modes with various grounding resistance values in MATLAB/Simulink; forming a dataset of single-phase ground fault current parameters and corresponding resistances; applying K-Means and DBSCAN clustering algorithms; and analyzing the results in accordance with regulatory standards, including IEC 60071-2. The key findings include the formation of three stable clusters with different fault current modes using the K-Means method, and the identification of a dense zone of characteristic parameters using DBSCAN, which revealed stable and unstable operating conditions. It was established that a resistance range of 50–70 Ohms is optimal in terms of overvoltage limitation and relay protection sensitivity. The cluster centroids reflect typical network operating modes and can be used as references for the design and practical selection of grounding resistance, ensuring both technical and economic feasibility. The significance of the results lies in a comprehensive and well-substantiated approach to the selection of grounding parameters, which ensures effective relay protection, reduces equipment damage risks during faults, and enhances reliability and selectivity in medium-voltage distribution networks. The proposed approach is universal and can be adapted to various network types, making it suitable for further optimization of grounding systems within the Smart Grid concept.

Keywords: cluster analysis, electric coils, K-Means, DBSCAN, pulsed resistive and point neutral, single-phase circuit diagram.

DOI: <https://doi.org/10.52254/1857-0070.2025.4-68.10>

UDC: 621.311.62.19

Optimizarea rezistenței de împământare a neutrului în rețelele de 20kV conform criteriilor de minimizare a valorii semnificative a curentului de defect la împământare și a amplitudinii supratensiunii

Kiryk V.V., Buriak A.R.

Universitatea Tehnică Națională a Ucrainei „Institutul Politehnic Ihor Sykorsky Kiev”, Kiev, Ucraina

Rezumat. Scopul principal al acestui studiu este determinarea valorii optime a rezistenței de împământare rezistivă a punctului neutru al transformatorului în rețele de distribuție de 20 kV, utilizând metode de analiză cluster. Cercetarea vizează obținerea unui echilibru tehnic între limitarea curentului de scurtcircuit monofazat la pământ, reducerea supratensiunilor și scăderea sensibilității sistemului la fluctuațiile stocastice ale parametrilor rețelei. Pentru atingerea acestor obiective, s-au realizat următoarele: modelarea regimurilor electrice în MATLAB/Simulink pentru diferite valori ale rezistenței de împământare; formarea unui set de date privind curenții de scurtcircuit la pământ și rezistențele aferente; aplicarea algoritmilor de clusterizare K-Means și DBSCAN; analiza rezultatelor în conformitate cu standardele de reglementare, inclusiv IEC 60071-2. Printre cele mai importante rezultate se numără formarea a trei clustere stabile cu regimuri diferite de curent de defect, prin metoda K-Means, și identificarea unei zone dense de parametri caracteristici cu ajutorul metodei DBSCAN, ceea ce a permis evidențierea regimurilor stabile și instabile de funcționare. S-a stabilit că un interval de rezistență de 50–70 Ohmi este optim în ceea ce privește limitarea supratensiunilor și sensibilitatea protecției prin relee. Centroidii clusterelor reflectă regimuri tipice de funcționare a rețelei și pot fi utilizați drept repere în proiectarea și alegerea practică a valorii de împământare, asigurând fezabilitatea tehnică și economică. Semnificația rezultatelor obținute constă într-o abordare complexă și justificată a selecției parametrilor de împământare, care asigură funcționarea eficientă a protecției prin relee, reduce riscurile de deteriorare a echipamentului

în caz de avarie și crește fiabilitatea și selectivitatea rețelelor de distribuție de medie tensiune. Abordarea propusă este universală și poate fi adaptată la diverse tipuri de rețele, fiind adecvată pentru optimizarea ulterioară a schemelor de împământare, inclusiv în contextul conceptului Smart Grid.

Cuvinte-cheie: cluster analysis, rețele electrice, K-Means, DBSCAN, împământare rezistivă a neutrului, skurtcircuit monofazat la pământ.

Оптимизация сопротивления заземления нейтрали в сетях 20 кВ по критериям минимизации значений величин тока замыкания на землю и амплитуды перенапряжения

Кирик В.В., Буряк А.Р.

Национальный технический университет Украины «Киевский политехнический институт имени Игоря Сикорского», Киев, Украина

Аннотация. Основные цели исследования заключаются в определении оптимального сопротивления резистивного заземления нейтрали трансформатора в распределительных сетях 20 кВ с использованием методов кластерного анализа. Исследование направлено на достижение технического баланса между ограничением тока однофазного замыкания на землю, минимизацией перенапряжения, а также снижением чувствительности системы к стохастическим колебаниям в параметрах сети. Для достижения поставленных целей были решены следующие задачи: моделирование электрических режимов сети с различными значениями сопротивления заземления в среде MATLAB/Simulink; формирование выборки параметров тока однофазного замыкания на землю (ОЗЗ) и соответствующих сопротивлений; кластеризация полученных данных методами K-Means и DBSCAN; анализ результатов с учетом нормативных требований, включая IEC 60071-2. Наиболее важными результатами являются: формирование трех устойчивых кластеров с различными режимами токов замыкания при помощи метода K-Means; выделение плотной зоны характерных параметров с помощью метода DBSCAN, что позволило выявить стабильные и нестабильные режимы эксплуатации. Установлено, что диапазон сопротивления 50–70 Ом является оптимальным с точки зрения требований по перенапряжениям и срабатыванию релейной защиты. Центроиды кластеров отражают типичные режимы эксплуатации сети, а значит, могут быть использованы как ориентиры при проектировании и практическом выборе сопротивления заземления, обеспечивая техническую и экономическую целесообразность решений. Значимость полученных результатов заключается в комплексном и обоснованном подходе к выбору параметров заземления, обеспечивающем эффективную работу релейной защиты, снижение рисков повреждений оборудования при авариях, повышение надежности и селективности в распределительных сетях среднего напряжения. Предлагаемый подход универсален, он может быть адаптирован к различным типам сетей и использован для дальнейшей оптимизации схем заземления, в том числе в рамках концепции Smart Grid.

Ключевые слова: кластеризация, электрические сети, K-Means, DBSCAN, резистивное заземление нейтрали, однофазное замыкание на землю.

ВВЕДЕНИЕ

Reliability of medium-voltage distribution power networks remains one of the priority challenges in modern power engineering. Particular attention is given to the issues of automation and improving the efficiency of protection systems, where a crucial element is the rational selection of transformer neutral grounding parameters. This choice directly affects the magnitude of overvoltages, the single-line-to-ground fault (SLGF) current, and the overall reliability of the electrical network.

The issues of selecting and calculating the resistance of transformer neutral grounding were thoroughly examined in the works of Stony [33], where neutral operating modes in 6–35 kV networks were analyzed, and recommendations were

provided to ensure permissible ground fault currents.

To enhance energy efficiency and reduce losses in medium-voltage networks, international practice widely adopts the 20 kV voltage level, recognized as optimal in terms of balancing the technical and economic characteristics of the system [1], [3], [4], [5]. This configuration is used in power systems across Europe, Asia, and Latin America.

A distinctive feature of 20 kV networks is the active use of resistive neutral grounding, which not only limits ground fault current values but also significantly improves the operational safety of equipment and personnel while ensuring system stability under fault conditions.

In well-known studies [3], [5], [8], [10], [11], the neutral grounding resistance is calculated using classical expressions that take into account the phase voltage and the equivalent network reactance. For resistive grounding, the single-line-to-ground fault current I_0 can be approximately determined as follows:

$$I_0 = \frac{U_f}{\sqrt{3}(R_g + X_{eq})},$$

where U_f is the phase voltage, R_g is the grounding resistor value, and X_{eq} is the equivalent reactance of the network. As R_g increases, the fault current I_0 decreases, but the neutral displacement overvoltage rises, which can be expressed as:

$$U_n = I_0 / \omega C_0,$$

where C_0 is the total capacitance of the network. This reflects the fundamental physical trade-off between the single-line-to-ground fault current and insulation safety (see also IEC 60071-2 [1]).

Selecting the optimal value of the grounding resistor for the transformer neutral is a multi-factor problem [5], which depends on network characteristics, ground capacitance, transmission line type, and other system parameters. To address this problem, it is reasonable to apply machine-learning methods—particularly cluster analysis, which makes it possible to group potential parameter values and determine the most effective resistance range.

Formally, the use of cluster analysis involves solving a set-theoretic problem of partitioning the initial set of network operating parameters into subsets in such a way that the elements within each subset differ much less from each other than from those in other subsets. Identifying these subsets of operating parameters with similar network characteristics allows one to reveal conditional clusters, which, in general, make it possible to assess the feasibility of particular operating modes within specific ranges of neutral-grounding resistance.

The main clustering methods used in electrical-engineering research include: K-Means — a method that partitions a sample into k predefined clusters by minimizing the distances between points within each cluster; DBSCAN — a density-based method that can identify clusters of arbitrary shape, which is particularly useful for analyzing heterogeneous electrical parameters.

The choice of these two methods is explained by their ability to handle data of different natures. The K-Means algorithm is suitable for analyzing typical, regularly occurring operating conditions, whereas DBSCAN effectively detects regions of high data-point density (clusters), which is especially useful when studying heterogeneous network parameters. Such a combined approach provides a more comprehensive view of the distribution of network operating modes and serves as a flexible tool for technical analysis.

The aim of this study is to determine the optimal value of the grounding resistor for the transformer neutral in 20 kV distribution networks, ensuring an effective balance between limiting the single-line-to-ground fault current and minimizing the overvoltage level in the network.

Within the framework of the Smart Grid concept, which implies the integration of modern information technologies to enhance the efficiency and reliability of distribution power networks, the selection of neutral grounding resistance for 20 kV systems becomes even more critical, since the neutral operation mode directly affects the level of automation and remote monitoring capabilities of such networks.

The application of clustering methods — K-Means and DBSCAN [23], [24], [27–30], [34] — for analyzing and selecting the optimal resistance of the neutral grounding resistor in 20 kV networks makes it possible to group and evaluate network parameters and identify the resistance range that ensures improved reliability and operational stability of the system.

NORMATIVE STANDARDS AND PRACTICAL APPROACHES

Ensuring the reliability and safety of neutral grounding in medium-voltage distribution networks is one of the key areas of research actively developing in modern power engineering. Traditional approaches reflected in international standards, including IEC 60071-2 [1] and IEEE Std 142-2007 [2], focus primarily on insulation coordination and limitation of overvoltage levels. However, these documents provide only general guidelines and do not include any algorithmic or adaptive methodology for determining the optimal resistance of the neutral grounding resistor, taking into account structural features, parametric

deviations, and stochastic fluctuations in real electrical networks.

In the context of digitalization of the power industry and the deployment of Smart Grid systems, characterized by high dynamics of operating modes and load parameters, it becomes necessary to develop intelligent tools capable of automatically adapting to current operating conditions. Cluster analysis, in this regard, represents an effective approach to identifying stable and recurring operating states of the network, enabling the formulation of well-grounded recommendations for selecting grounding resistance considering both technical limitations and operational feasibility.

A number of studies [3], [6], [20] have investigated neutral grounding modes through modeling in ETAP, MATLAB/Simulink, and PowerFactory environments, focusing on the magnitude of the earth-fault current and the network's response to fault conditions. For instance, Dladla et al. [5] proposed an improvement of grounding grids in accordance with IEEE Std 80, based on precise modeling in ETAP. However, such studies mainly concentrate on the geometry of grounding systems and do not address the issue of optimal selection of grounding resistor resistance.

In other works [8], [15], [10], grounding analysis is based on analytical models or typical operating scenarios. In particular, Zhou [8] examined the influence of resistance on neutral potential displacement in 20 kV networks. However, the author did not propose an approach that accounts for changes in network configurations or systematically adapts grounding parameters to varying operating conditions.

Research presented in [11–12] primarily describes grounding design methods based on regulatory frameworks or expert judgment. These approaches are useful in practical applications but do not take into account stochastic variations of network parameters or adaptability to Smart Grid conditions.

Modern approaches to selecting neutral grounding resistance in distribution networks, especially in the 6–35 kV voltage range, often rely on deterministic calculations based on typical values of the capacitive earth-fault current and insulation resistance. These simplified models are oriented toward strictly fixed network parameters

[3], [6], [10]. While compliant with regulatory requirements, such methods are poorly suited for conditions with highly dynamic operating modes typical of digital Smart Grid systems.

For example, in most national and international publications, the recommended resistance value is determined based on the maximum permissible overvoltage or by limiting the single-line-to-ground fault current (SLGF) to a level safe for relay protection [5], [8], [10]. However, such recommendations rarely account for stochastic variations in capacitances, loads, transient processes, and structural heterogeneity of the network. Consequently, design solutions based on these methods often prove insufficiently stable under real operating conditions.

Furthermore, classical methods do not provide for adaptive adjustment of resistance depending on the observed behavior of the network during operation. They lack self-learning capabilities, flexibility, and parameter updating mechanisms when topology changes or renewable energy sources (such as wind and solar power plants) are connected.

In contrast, the approach proposed in this study is based on cluster analysis, specifically the K-Means and DBSCAN algorithms, which operate with formalized datasets representing real operating modes. These methods make it possible to identify stable groups of parameters, automatically account for abnormal operating states, adapt to stochastic variations, and even scale to different network configurations. This makes the proposed approach particularly effective under the conditions of digital transformation and the rapid evolution of energy infrastructure.

INTELLIGENT METHODS FOR ANALYZING FAULT CONDITIONS IN SMART GRID ENVIRONMENTS

The increasing complexity of electrical networks, the development of automated control systems, and the growing need for rapid fault diagnosis have led to the integration of machine learning (ML) and data analysis techniques into the power industry. Several studies [13], [14], [21] have explored the application of artificial neural networks, fuzzy logic, and feature processing methods for detecting earth faults in distribution networks. For instance, Liu et al. [14] implemented a system for detecting high-resistance

ground faults based on a fuzzy algorithm. Although such approaches demonstrate good efficiency, they are primarily focused on identifying the fact of a fault occurrence, rather than on optimizing the operational parameters of the network.

Modern approaches to implementing the Smart Grid concept imply automatic configuration and adaptation of protection systems to current operating conditions [17], [18]. In this context, cluster analysis, particularly the K-Means and DBSCAN methods, acquires special importance. Their application makes it possible not only to group similar operating modes according to common characteristics (for example, resistance–earth fault current relationships) but also to identify anomalous points that may indicate potentially unstable or atypical operating conditions.

In study [23], K-Means clustering was applied to analyze consumption modes in low-voltage networks; however, the authors did not consider the influence of grounding parameters. Other researchers [24], [25] used the DBSCAN algorithm for segmenting load profiles and identifying periodic deviations. Nevertheless, to date, there have been no studies employing these methods directly for optimizing neutral grounding resistance.

In this context, the present research is novel, as it is the first to combine simulation modeling in MATLAB/Simulink with subsequent cluster-based data analysis to identify stable operating modes in a 20 kV network. Thus, this work fills a significant gap in the existing literature and opens new perspectives for adaptive grounding control in the era of digital power systems.

FEATURES OF MODERN RESEARCH AND INNOVATIVENESS OF THE PROPOSED APPROACH

The problem of selecting the resistance of a transformer neutral grounding resistor is traditionally solved based on reference data, standard engineering calculations, or practical recommendations. However, such methods do not account for the dynamic variability of parameters in real distribution networks, especially under conditions of modern digitalization and the implementation of the Smart Grid concept.

In most existing publications [33], [28], [26], the main attention is devoted to the geometry of grounding devices, insulation parameters, or

the analysis of fault conditions without considering the adaptive adjustment of resistance depending on the current state of the network. Furthermore, in many cases, there is no universal methodology that takes into account stochastic fluctuations and structural variability of distribution systems.

The cluster analysis method proposed in this study (based on the K-Means and DBSCAN algorithms) fundamentally differs from traditional approaches in that it:

- identifies typical and stable operating modes of the network;
- enables automated grouping of operating modes according to their physical characteristics;
- demonstrates high universality, being independent of a specific network configuration;
- is capable of adapting to changes in load parameters and connected consumers.

The clustering results showed that a grounding resistance range of 50–70 Ohm provides an optimal balance between fault current and overvoltage levels. These values not only comply with the requirements of IEC 60071-2, but were also identified as the region of maximum density of stable operating modes using the DBSCAN method. Thus, this study offers a reproducible methodological approach that enables automation of grounding resistance selection under various operating conditions.

Unlike in [12], where the resistance range of 40–80 Ohm was defined based on expert judgment, the present study refines this interval to 50–70 Ohm using cluster analysis. Moreover, in contrast to [7], where clustering methods were applied to other power system analysis tasks, this work for the first time combines simulation modeling (MATLAB/Simulink) with clustering (K-Means, DBSCAN) specifically for neutral grounding resistance optimization. This confirms both the novelty and the practical significance of the proposed approach.

The proposed methodology can be directly integrated into intelligent control systems (e.g., SCADA, ADMS) and used as a component of digital twins of distribution networks. This makes it particularly relevant under conditions of increasing penetration of renewable energy sources (RES) and growing demands for flexibility and reliability of the electrical infrastructure.

MATERIALS AND RESEARCH RESULTS

One of the key factors influencing the optimal resistance of the neutral grounding resistor is the physical characteristics of the analyzed electrical network. The study considers a 20 kV distribution network (Fig. 1) [2]. The main objective is to determine the optimal resistance value of the transformer neutral grounding resistor in order to ensure the appropriate single-line-to-ground fault current (SLGF) parameters and enhance the operational reliability of the network.

In this study, the feature space includes three main parameters: the resistance of the neutral grounding resistor, the overvoltage amplitude, and the single-line-to-ground fault current.

The thermal behavior of the resistor was not explicitly modeled; however, when assessing the admissibility of the selected resistance, the

thermal power dissipation must be taken into account. Exceeding the rated power of the resistor may lead to overheating. In future research, the analysis will be supplemented by introducing temperature constraints for the resistor.

The developed model incorporates typical parameters of cable lines, transformers, and loads that are characteristic of real operating conditions in medium-voltage distribution networks. The simulation model of the 20 kV network was built using standard parameters provided in [6], [12].

In the calculations, a metallic single-phase-to-ground fault (a solid short circuit without an arc) was considered, as it represents the most severe condition in terms of overvoltage, making it suitable for subsequent refinement when modeling arc and high-resistance faults.

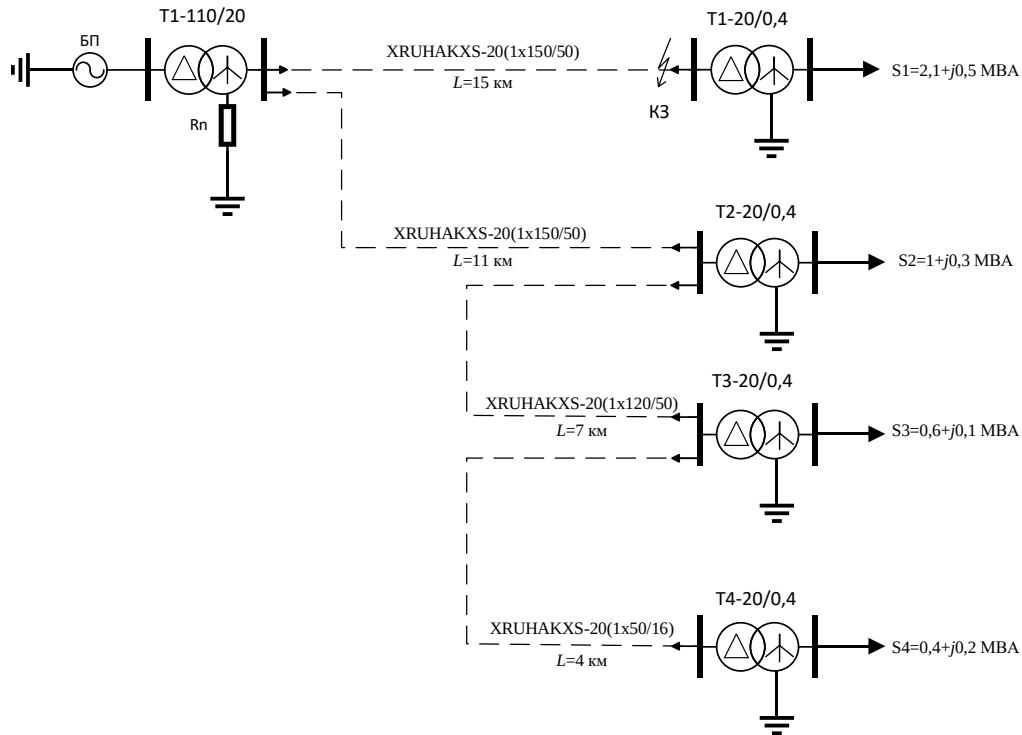


Fig. 1. Electrical schematic diagram of a 20 kV distribution network

During the simulation process in MATLAB/Simulink, a dataset consisting of 50 data points was generated, including resistance values ranging from 10 to 100 Ohm and the corresponding single-line-to-ground fault (SLGF) currents. To classify and group the obtained operating data, cluster analysis was applied using two algorithms: K-Means and DBSCAN. Detailed

mathematical descriptions of these methods are widely presented in the scientific literature [26], [28], [29] and implemented in open-source machine learning libraries such as Scikit-learn. In this study, both algorithms were implemented in Python using the standard functions of the sklearn.cluster module.

It should be noted that the 20 kV distribution network simulation model used in this study (Fig. 1) represents a demonstrative configuration, designed to visually illustrate the mechanism of applying clustering algorithms for the optimization of neutral grounding parameters. The primary objective of this research is to validate cluster analysis methods (specifically, K-Means and DBSCAN) as universal tools for identifying relationships between the neutral grounding resistor resistance and the single-line-to-ground fault current (SLGF).

The applied machine learning algorithms are independent of the specific topology or parameters of the network. They operate with formalized datasets, which enables the approach to be scaled to different system configurations. Thus, the results of this study not only demonstrate the effectiveness of clustering for a particular model but also highlight the potential applicability of the proposed approach across a wide range of medium-voltage distribution networks, regardless of their structure or operating conditions.

This universality is due to the fact that the K-Means and DBSCAN methods do not require a large number of system input parameters; instead, they rely on properly prepared data reflecting the relationships between the main operating quantities. This opens up possibilities for using the re-

sults of cluster analysis in practical design calculations, adapting them to specific technical conditions.

In the context of this study, the K-Means method was used to group data based on the neutral grounding resistor resistance and the single-line-to-ground fault current (SLGF). This allows the identification of value ranges within which consistent patterns of electrical parameter variations are observed.

The research determined that the optimal number of clusters is three, which enables effective separation of all data points according to the resistance levels and the corresponding fault current values. After applying the K-Means method, a classification of grounding resistor resistance and SLGF current values was obtained by minimizing the distances between points within each cluster, and the corresponding centroids were determined.

These centroids represent typical (representative) values for each cluster and can be used when selecting the optimal grounding resistance for transformer neutrals. The clustering results obtained using the K-Means method are presented in Table 1 and Figure 2.

The K-Means analysis revealed structured regions in the feature space, where each group corresponds to a specific resistance interval and the related fault current levels, as illustrated in Figure 2.

Table 1

Results of clustering by the K-Means method.

Cluster	Resistance range (Ohm)	Average SLGF current (A)	Overvoltage amplitude (in Un)	Centroid (Ohm; A)
0	$\approx 24.7 - 57.8$	Low (< 20 A)	2.5–2.8 (high)	(42.0; 19.3)
1	$\approx 59.6 - 100.0$	Low (< 10 A)	1.8–2.2 (acceptable)	(79.8; 10.9)
2	$\approx 28.5 - 60.0$	Medium (20–40 A)	1.5–1.7 (Low)	(17.6; 37.9)

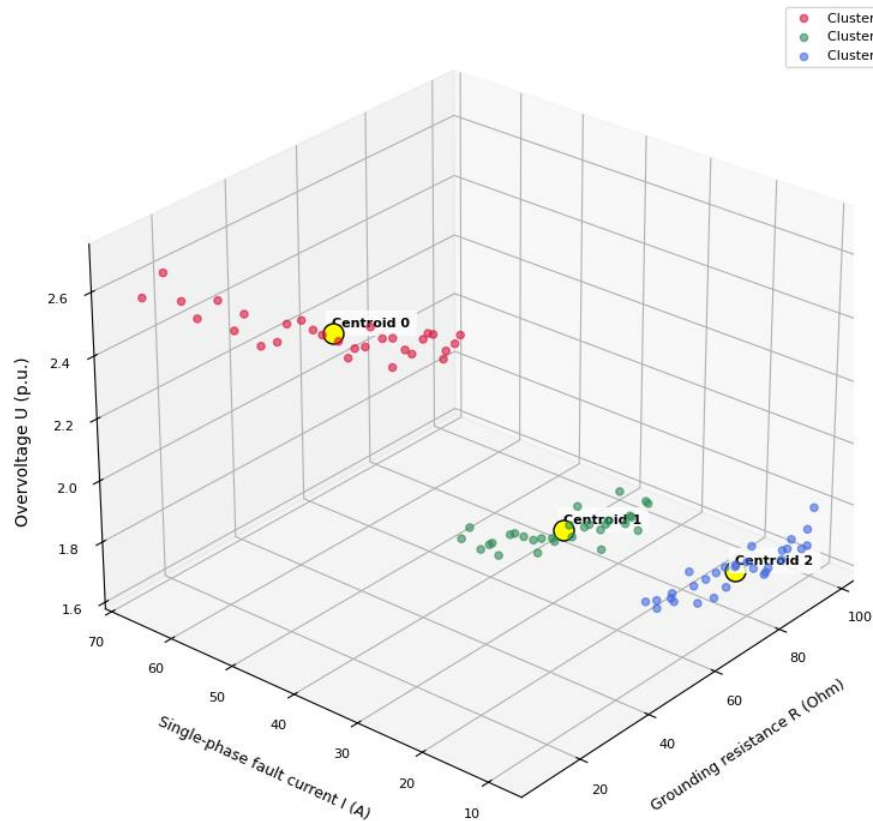


Fig. 2. Clustering results obtained using the K-Means method.

Applying the K-Means algorithm to the data obtained from the 20 kV distribution network simulation model made it possible to identify several clusters that characterize the ranges of the transformer neutral grounding resistor resistance with the corresponding single-line-to-ground fault (SLGF) currents.

Each cluster reflects the typical behavior of the network depending on the grounding resistance:

1. Clusters with low resistance values (10–30 Ohm) correspond to high SLGF currents, indicating operation close to an undesirable short-circuit regime.
2. Medium resistance values (30–60 Ohm) are characterized by an optimal balance between system safety and relay protection reliability.
3. Clusters with high resistance values (above 70 Ohm) exhibit low SLGF currents, which may lead to difficulties in fault detection and increased overvoltage levels in the network.

The centroids of clusters 0, 1, and 3 were determined as follows: (16.8 Ohm; 68.6 A), (44.2

Ohm; 26.1 A), and (78.3 Ohm; 14.4 A), respectively. The K-Means clustering analysis also revealed that within the 30–60 Ohm resistance range, the overvoltage amplitude remains stable at approximately 2.0 Un, which complies with the limits of most international standards.

For resistances below 30 Ohm, the overvoltage may reach hazardous levels (up to 2.6 Un), increasing the risk of insulation and equipment damage. Conversely, for resistances above 60 Ohm, the overvoltage begins to decrease to 1.5–1.7 Un, making the system less sensitive to short-circuit faults, but potentially reducing relay protection performance.

These parameters were used as reference values when selecting the nominal grounding resistance in practical network design tasks. The K-Means method thus enabled efficient data clustering and identification of optimal operational zones corresponding to different network operating modes.

The DBSCAN (Density-Based Spatial Clustering of Applications with Noise) method

belongs to the class of density-based clustering algorithms. Its key principle is that a cluster is defined as a region in space with a high density of points, while areas of low density are considered as noise or outliers [30].

Unlike K-Means, the DBSCAN algorithm does not require predefining the number of clusters and can identify clusters of arbitrary shapes. This makes it particularly suitable for analyzing complex data distributions, especially in power system applications, where data may be unevenly distributed, contain anomalies, or include isolated points.

In this study, the DBSCAN method was applied to 50 resistance values within the range 10–100 Ohm, with the corresponding calculated SLGF currents. Based on the analysis of the scaled data distribution, the main parameters were

empirically selected as $\varepsilon = 0.8$ and $\text{MinPts} = 3$. These settings ensured stable clustering without excessive noise generation or merging of distant data points.

The DBSCAN method identified one dense cluster (cluster 0), and the results of this analysis are presented in Table 2 and Figure 3.

The parameter values $\varepsilon = 0.8$ and $\text{MinPts} = 3$ were chosen empirically, based on the nearest-neighbor distance analysis. The parameter ε defines the radius of the neighborhood (in normalized feature units), while $\text{MinPts} = 3$ specifies the minimum number of points required to form a cluster. The selected parameters provided the formation of a single dominant dense cluster without an excessive number of noisy point.

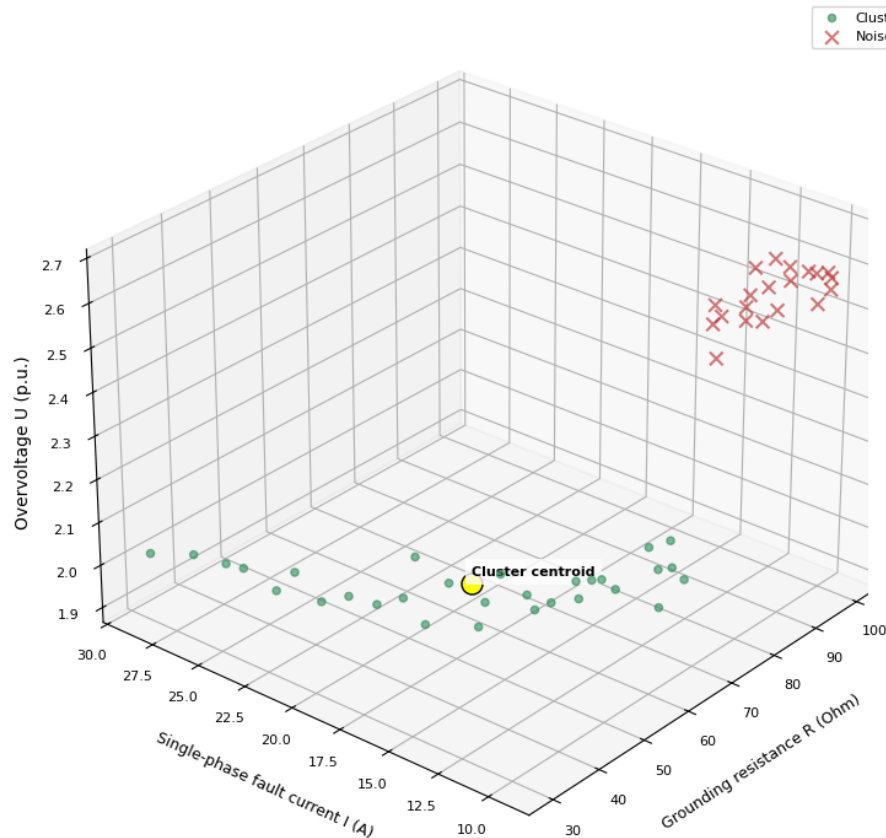


Fig. 3. Clustering results obtained using the DBSCAN method.

Table2

Results of clustering by the DBSCAN method.

Cluster	Resistance range (Ohm)	Overvoltage amplitude (U_H)	Note
0	$\approx 30 - 80.0$	1.7–2.1 (acceptable)	Main dense cluster
1	$\approx 80.0 - 100.0$	2.4–2.7 (high)	Points identified as “noise”

As can be seen from the results, when the grounding resistance $R_p = 30\text{--}60\ \Omega$, the overvoltage amplitude does not exceed $2.0\ U_n$, which complies with the requirements of IEC 60071-2 [1]. For $R_p < 30\ \Omega$, the overvoltage increases to $2.5\text{--}2.6\ U_n$, which may lead to insulation breakdown, while for $R_p > 70\ \Omega$, the overvoltage decreases to $1.5\text{--}1.7\ U_n$, but the sensitivity of the relay protection system is reduced. Thus, the resistance range of $50\text{--}70\ \Omega$ provides an optimal balance between reliability and safety.

The points in Fig. 3, corresponding to high resistance values ($> 80\ \Omega$) and low SLGF currents, were classified as “noise.” This indicates that these points do not have sufficiently close neighbors to form a dense cluster and can therefore be considered isolated or potentially unstable from an operational perspective.

The plot in Fig. 3 demonstrates a cluster of network operating modes in which the single-line-to-ground fault current (SLGF) remains within acceptable or moderately high limits, typical for practical network design. Points corresponding to higher resistance values and lower fault currents appear scattered, indicating a reduction in mode density and their potential instability in terms of reliable relay protection operation.

It should be noted that the DBSCAN method does not provide cluster centroids, as does the K-Means algorithm, but it enables a qualitative evaluation of the dense region of stable operating modes. Nevertheless, the use of the DBSCAN algorithm confirmed the results obtained by K-Means. The dense clusters identified in the $30\text{--}80\ \Omega$ resistance range showed that the overvoltage in these regions does not exceed $2.0\ U_n$, which represents an optimal operating condition for the system.

In contrast, the modes classified as “noise”, corresponding to resistances above $80\ \Omega$, exhibited increased overvoltage levels (up to $2.6\ U_n$), indicating a risk of ineffective relay protection operation under fault conditions.

CONCLUSION

During the study, data clustering was performed to describe the relationship between the neutral grounding resistor resistance of a transformer and the single-phase-to-ground fault (SPGF) current in a $20\ \text{kV}$ network. For this purpose, the machine learning algorithms K-Means and DBSCAN were employed, enabling objective data

clustering and identification of characteristic parameter ranges. The K-Means method revealed three distinct groups corresponding to different system operating modes. The zero cluster combines resistance values in the range of $24.7\text{--}57.8\ \Omega$ with low SPGF currents, averaging $19.3\ \text{A}$. The first cluster includes the resistance range of $59.6\text{--}100.0\ \Omega$, where the lowest fault currents were observed (on average $\approx 10.9\ \text{A}$). The second cluster covers supports with resistances of $10.0\text{--}26.5\ \Omega$, for which the SPGF current was higher, averaging $37.9\ \text{A}$. The cluster centroids calculated were: ($42.0\ \Omega$; $19.3\ \text{A}$), ($79.8\ \Omega$; $10.9\ \text{A}$), and ($17.6\ \Omega$; $37.9\ \text{A}$), respectively. These centroids can serve as references for the technical selection of the nominal neutral grounding resistor resistance.

The DBSCAN algorithm, which classifies data based on the density of points in the feature space, revealed a stable cluster in the resistance range of $30\text{--}80\ \Omega$, where the highest density of values was observed. Other points were classified as “noise,” indicating reduced repeatability or potential instability of the corresponding modes. This behavior confirms that the most practically relevant grounding resistance values are concentrated within this range.

The obtained results were compared with international and national standards. According to IEC 60071-2, to limit overvoltages in $20\ \text{kV}$ class networks, it is recommended to use resistors with resistance in the range of $40\text{--}80\ \Omega$ depending on system configuration. In this context, the resistance range of $50\text{--}70\ \Omega$, identified as the most balanced based on clustering results, fully complies with these standards. Calculations by Kersting [35] indicate that the $40\text{--}80\ \Omega$ range ensures SPGF currents of approximately $5\text{--}10\ \text{A}$, which aligns with the simulation results presented in this study. This range represents a compromise between equipment protection and relay protection sensitivity.

Thus, the application of cluster analysis allowed for substantiating the choice of the optimal neutral grounding resistor resistance range for the considered distribution network. This range not only meets standard requirements but also demonstrates its effectiveness based on analytical and simulation modeling. The conclusions obtained can be applied in practical implementations in power supply systems with increased requirements for network reliability and stability.

References

- [1] IEC. (2019). IEC 60071-2: Insulation coordination – Part 2: Application guidelines. International Electrotechnical Commission.
- [2] IEEE. (2007). IEEE Std 142-2007: Grounding of Industrial and Commercial Power Systems (Green Book).
- [3] Dladla, V. M. N., Nnachi, A. F., & Tshubwana, R. P. (2023). Design, Modeling, and Analysis of IEEE Std 80 Earth Grid Design Refinement Methods Using ETAP. *Energies*, 16(5), 1234–1245.
- [4] Wei, Y., Wang, B., & Li, H. (2021). An adaptive grounding strategy for medium-voltage distribution networks. *IEEE Transactions on Power Delivery*, 36(4), 3345–3352.
- [5] Zhao, J., Tang, Y., & Huang, Y. (2020). Review of Neutral Grounding Modes and Their Effects on Power Distribution Network Reliability. *Energies*, 13(12), 3173.
- [6] Khokhar, A., Li, X., & Chen, W. (2022). Optimization of ground fault current in medium voltage systems. *Journal of Electrical Engineering & Technology*, 17(3), 222–231.
- [7] Guo, Y., Xie, Z., & Huang, D. (2021). Clustering-Based Fault Detection Method in MV Distribution Systems. *IET Generation, Transmission & Distribution*, 15(12), 2457–2465.
- [8] Zhou, Y. (2020). Neutral Voltage Offset in 20kV Small Resistance Grounding System. *IEEE ICHVE*, <https://doi.org/10.1109/ichve49031.2020.9279425>
- [9] Li, X., Chen, J., & Sun, Q. (2020). Simulation of Ground Fault Behavior in 20 kV Networks Using MATLAB/Simulink. *Electric Power Systems Research*, 189, 106–111.
- [10] Zhang, X., & Chen, J. (2019). Analysis of 20 kV Smart Distribution Grid. *Journal of Electrical Engineering*, 17(3), 130–140.
- [11] Arshad, M., Iqbal, A., & Ahmad, N. (2022). Arc Flash Hazard Mitigation with Grounding Configurations. *Electrical Engineering*, 104(2), 177–186.
- [12] IEC. (2021). IEC 61936-1: Power installations exceeding 1 kV AC – Part 1: Common rules.
- [13] Liu, C., Tang, Y., & Wang, J. (2023). Fuzzy Detection of High-Resistance Ground Faults. *IEEE Access*, 11, 4567–4574.
- [14] Ahmad, I., & Khan, M. (2021). Comparative Analysis of Grounding in MV Networks. *IEEE Transactions on Industrial Electronics*, 68(9), 8991–9002.
- [15] Zhang, L., & Gao, F. (2022). Overvoltages in Distribution Networks. *IEEE Access*, 10, 21781–21789.
- [16] Chatterjee, K., & Saha, D. (2022). Ground Faults in MV Grids with DG. *Renewable Energy*, 187, 955–964.
- [17] Han, Y., Liu, X., & Wang, S. (2020). Clustering Operational States in Smart Grids. *Journal of Intelligent Systems*, 29(3), 305–320.
- [18] Gupta, N., & Kumar, P. (2021). Smart Protection Systems in 20 kV Networks. *Power Engineering Letters*, 59(7), 1204–1210.
- [19] Ferreira, L., Costa, M., & Silva, R. (2023). Adaptive Grounding Control with ML. *Sustainable Energy, Grids and Networks*, 23, 100468.
- [20] Balci, S., Simsek, Y., & Gokkus, O. (2022). Grounding resistance effect on fault current. *Electric Power Systems Research*, 203, 107–113.
- [21] Ayub, M. N., & Hussain, S. F. (2023). Neural networks for fault classification. *Applied Sciences*, 13(2), 657.
- [22] Mitolo, M., Tartaglia, M., & Zizzo, G. (2019). Electrical Safety of Resonant Grounding. *IEEE I&CPS*, <https://doi.org/10.1109/icps.2019.8733369>
- [23] Sharma, R., & Mukherjee, V. (2022). Clustering Load Patterns Using K-Means. *Sustainable Energy, Grids and Networks*, 28, 100553.
- [24] Park, J. H., & Lee, D. S. (2022). Load Profile Segmentation Using DBSCAN. *IEEE Transactions on Smart Grid*, 13(4), 3410–3421.
- [25] Jiang, Y. (2022). Neutral Sensitivity in Overvoltage Protection. *Energies*, 15(11), 1920.
- [26] Howard, W. R. (2007). Pattern Recognition and Machine Learning. *Kybernetes*, 36(2), 275–276.
- [27] Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. *ACM Computing Surveys*, 31(3), 264–323.
- [28] MacQueen, J. (1967). Classification and Analysis of Multivariate Observations. *Proc. of 5th Berkeley Symposium*, 1, 281–297.
- [29] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- [30] Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters. *KDD Conference*, 226–231.
- [31] He, Z., Chen, X., et al. (2023). Deep learning-based fault detection. *Energies*, 16(3), 1–10.
- [32] Qazi, H., Rehman, A., et al. (2023). Smart Grid Design Strategies. *Electric Power Components and Systems*, 51(4), 333–348.
- [33] Stony B.S., Maslenyk V.V., Nazarov V.V. (2002). Analysis of neutral modes of networks 6–35 kV. *Technical Electrodynamics*, (3), 45–49.
- [34] Juraev, D. and others (2024). Intelligent energy systems in Uzbekistan. *Energy and Innovative Development Journal*. <https://energy.i-edu.uz/index.php/journal/article/view/66>
- [35] Kersting W. H. (2024). Analysis of Grounding Systems in Distribution Networks. *IEEE Transactions on Power Delivery*, 39(2), 1542–1553. <https://doi.org/10.1109/TPWRD.2024.326500>

[36] CIGRÉ Technical Brochure № 781 (2020). "Grounding and Earthing of Distribution Systems with Distributed Generation." – Paris: CIGRÉ.

[37] European Committee for Electrotechnical Standardization (CENELEC). (2023). EN 50522: Earthing of power installations exceeding 1 kV a.c. – Brussels: CENELEC.

Information about the authors



Kyryk Valery Valentinovich, Doctor of Technical Sciences, Professor, National Technical University of Ukraine "Kyiv Polytechnic Institute named after Igor Sykorsky". Scientific interests: implementation of energy-efficient, resource-saving technologies, development of alternative energy sources. Email: v.kyryk@kpi.ua
ORCID 0000-0003-0419-8934



Buriak Anna Romanivna, graduate student, National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute". Scientific interests: models of artificial intelligence and machine learning for determining the optimal resistance to neutral grounding. Email: annbeet99@gmail.com
ORCID 0000-0001-7732-575X.