Enhanced Prediction of Insulator Flashover Voltage Using Artificial Neural Networks Optimized with Particle Swarm Optimization

Lazreg Taibaoui
Laboratoire d'Etudes et Developpement des Materiaux Semi-Conducteurs et Dielectriques, Amar Telidji University of Laghouat, Algeria
l.taibaoui@lagh-univ.dz (corresponding author)

Abdelhalim Mahdjoubi
Laboratoire d'Etudes et Developpement des Materiaux Semi-Conducteurs et Dielectriques, Amar Telidji University of Laghouat, Algeria
ah.mahdjoubi@lagh-univ.dz

Boubakeur Zegnini
Laboratoire d'Etudes et Developpement des Materiaux Semi-Conducteurs et Dielectriques, Amar Telidji University of Laghouat, Algeria
b.zegnini@lagh-univ.dz

Received: 25 January 2025 | Revised: 19 May 2025 and 8 June 2025 | Accepted: 14 June 2025

 DOI: https://doi.org/10.48084/etasr.10330

ABSTRACT

Outdoor insulators are critical components in power systems but are highly susceptible to environmental factors such as moisture, rain, and contaminants. These adverse conditions often lead to surface flashovers, causing insulation failures and compromising the reliability of power systems. To address this challenge, this study aimed to develop a robust predictive model for flashover voltage under polluted conditions, leveraging the capabilities of Artificial Neural Networks (ANN) optimized with Particle Swarm Optimization (PSO). The primary objective was to enhance prediction accuracy and overcome the limitations of traditional methods. The proposed ANN-PSO model was trained and validated using a comprehensive dataset comprising experimental and simulated data under various pollution conditions. Key input features included Equivalent Salt Deposit Density (ESDD), insulator height, leakage distance, diameter, form factor, and environmental conditions. PSO was employed to optimize the ANN parameters, minimizing error metrics, such as Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), while maximizing the regression coefficient (R^2). The results demonstrate that the proposed ANN-PSO model significantly outperforms traditional approaches, including standard ANN and hybrid models such as LS-SVM-PSO and LS-SVM-GWO, in terms of predictive performance. The model achieved exceptionally low values for RMSE (0.003511) and MAPE (0.842%) along with a very high R^2 value (0.997), confirming its precision, robustness, and superior capability to predict flashover voltage accurately. This study provides a practical and reliable tool for power utilities to monitor and mitigate the risk of insulator flashovers in polluted environments. Furthermore, it highlights the potential for integrating advanced hybrid AI models to address complex challenges in power system operations, paving the way for further innovation in predictive modeling for power system reliability.

Keywords-ANN; flashover voltage; polluted insulators; PSO; prediction; power transmission

I. INTRODUCTION

The electrical power transmission and distribution network relies on chains of insulators to maintain system reliability. However, contamination from industrial emissions, urban pollution, marine salt particles, and desert sand significantly impacts their performance, increasing the risk of flashover [1]. This phenomenon occurs when a conductive layer forms on the insulator surface due to moisture from dew, rain, or fog, leading to leakage currents and potential power disruptions. As a result, predicting and mitigating flashover voltage in contaminated insulators is a critical challenge [2-3].

Extensive research has been carried out to investigate the behavior of polluted insulators, employing both physical and

mathematical approaches, including the widely used Obenaus model, to estimate the critical flashover voltage [4-5]. Experimental studies have significantly enhanced our understanding of the stresses caused by pollution on insulators [6-7], while sophisticated simulation techniques have been developed to represent real-world environmental conditions more accurately [8]. Despite these substantial efforts, conventional models often fail to effectively capture the intricate and nonlinear interactions between the insulator geometry, pollution severity, and environmental influences.

To overcome these limitations, Artificial Intelligence (AI) techniques have emerged as highly effective tools for predicting flashover voltage. Techniques including Artificial Neural Networks (ANN) [9], Fuzzy Logic (FL) [10], Support Vector Machines (SVM) [11], Least Squares Support Vector Machines (LS-SVM) [12], and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) [13] have shown impressive accuracy in predicting flashover phenomena. For example, in [10], a fuzzy logic-based method was introduced to estimate the critical flashover voltage, whereas in [14], the Grey Wolf Optimizer (GWO) was successfully combined with LS-SVM, obtaining prediction errors below 10%. Furthermore, a refined LS-SVM model employed a quadratic Rényi criterion [15], achieving even greater accuracy in voltage prediction tasks. Recently, in [13], an ANFIS model achieved exceptional accuracy, with a mean absolute percentage error of 0.027011 and a determination coefficient of 0.999997049.

Among neural network architectures, the Multi-Layer Perceptron (MLP) is one of the most widely recognized and applied models, typically utilizing the Backpropagation (BP) algorithm or one of its derivatives, known as the Backpropagation Neural Network (BPNN). However, the BP algorithm's reliance on the steepest descent search technique makes it prone to convergence issues, such as getting stuck in local optima, computational overflow, or oscillations. These limitations have led researchers to explore more robust optimization techniques to enhance neural network effectiveness. A significant breakthrough in addressing these challenges is the integration of evolutionary algorithms into neural network training. Among these, Particle Swarm Optimization (PSO), inspired by the social behavior of bird flocks and fish schools, has proven highly effective. Originally designed to simulate natural, graceful, but unpredictable flock movements graphically, it was subsequently refined by eliminating unnecessary parameters, resulting in an efficient basic PSO algorithm.

This study proposes a novel PSO-optimized ANN model specifically tailored to predict flashover voltage in polluted outdoor insulators. This approach leverages the cooperative and adaptive strengths of the swarm intelligence inherent in PSO, significantly enhancing prediction accuracy and reliability compared to traditional neural network methods. Furthermore, meticulous design and fine-tuning of hyperparameters ensure that this model is precisely adapted to the dataset, maximizing its predictive performance and generalization capabilities.

The originality of this research lies in its innovative combination of ANN with PSO optimization to predict flashover voltage under polluted conditions, an area where traditional prediction models often falter. By integrating AI-driven predictive modeling, this work contributes substantially to ongoing efforts to improve the reliability and efficiency of power transmission systems, particularly under harsh environmental conditions. This advancement is critical for ensuring system stability, reducing the risk of insulation failures, and enhancing overall operational efficiency.

II. METHODOLOGY

A. Data Selection and Mathematical Model

The data used to train and test the ANN-PSO model came from a variety of sources. Some of it was sourced from experiments at the High Voltage Laboratory of the Testing, Standards and Research Center of the General Energy Company of Athens, conducted in compliance with IEC standards [16-17]. The pollutants used included 675 g/l of silica powder, 75 g/l of kaolin clay, and sodium chloride as required, all suspended in isopropyl alcohol. This contamination process lasted 30 minutes, and the insulators were left to dry for one hour. The ESDD value was then set on the insulator surface based on the contamination intensity index. In addition to these experimental data, measurements of [18, 19] were also utilized.

The Obenaus model is a theoretical framework developed to examine flashover phenomena on polluted insulators. It conceptualizes the process by considering a partial arc that traverses a dry zone on the insulator surface, in combination with the resistive properties of the contamination layer. These two elements are modeled as a series circuit [9]. The model is particularly significant in defining the critical voltage U_c, which represents the threshold voltage at which the partial arc transitions into a full flashover event. This critical voltage, expressed in volts, is determined through a specific formula that integrates the physical and electrical characteristics of the insulator and its contamination. By providing a structured approach to quantify flashover conditions, the Obenaus model plays a vital role in understanding and predicting insulator performance under polluted conditions. Furthermore, it serves as a foundational basis for enhancing insulation design and optimizing maintenance strategies in outdoor power systems exposed to environmental challenges [9].

$$U_C = \frac{A}{n+1}(L + \pi D_m F K n)(\pi D_m \sigma_S A)^{-n/(n+1)} \quad (1)$$

Additionally, the model incorporates various parameters: A and n are arc constants, L represents the creepage distance of the insulator (measured in centimeters), and D is the maximum diameter of the insulator disc (also measured in centimeters). K is the pollution layer resistance coefficient, while F refers to the form factor, which accounts for the geometric profile of the insulator. Figure 1 illustrates the equivalent circuit model based on the Obenaus representation. In the proposed model, V_{arc} represents the arcing voltage, R_p denotes the resistance of the pollution layer, and U corresponds to the stable voltage supply source. The critical voltage, expressed in volts, is represented by U_c [9].

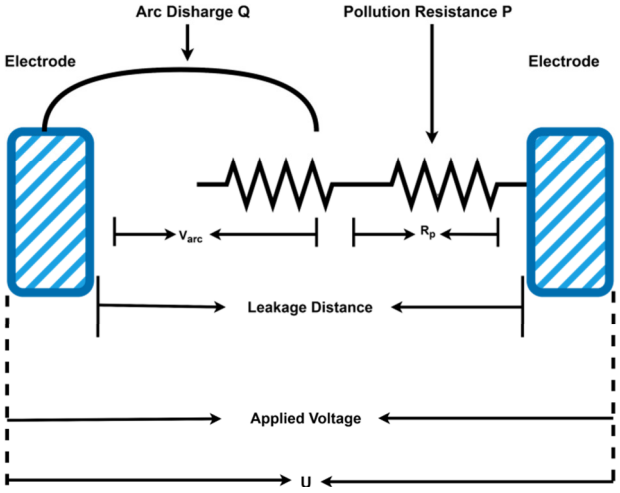

Fig. 1. Obenaus model.

A genetic optimization process was used to calculate the constants A and n, which are $n = 0.409$ and $A = 124.8$ [13]. The high-voltage insulator's form factor F is defined as:

$$F = \int_0^L \frac{dl}{p(l)} \quad (2)$$

Experimental evidence has shown that the flashover voltage of a contaminated insulator is not consistent, even when subjected to similar conditions. The surface conductivity σ_s is defined by:

$$\sigma_s = (369.05 \cdot C + 0.42) \cdot 10^{-6} \quad (3)$$

where C represents the Equivalent Salt Deposit Density (ESDD), measured in mg/cm².

For cap-and-pin insulators, the coefficient K represents the resistance of the pollution layer.

$$K = 1 + \frac{n+1}{2 \cdot \pi \cdot F \cdot n} \cdot ln\left(\frac{L}{2 \cdot \pi \cdot R \cdot F}\right) \quad (4)$$

The radius of the arc foot R, measured in centimeters, is calculated using:

$$R = 0.469(\pi A \cdot D_m \sigma_s)^{1/(2(n+l))} \quad (5)$$

A training dataset was curated with a sufficient number of representative data points. The input parameters considered include the leakage length per insulator element (L, cm), insulator height (H, cm), insulator diameter (D, cm), form factor (F), and the ESDD (C in mg/cm²). These parameters were utilized as the input vector for the model, while the output was the critical flashover voltage (V_c, kV). As a result, an input-output dataset was systematically constructed to facilitate the training process.

Table I summarizes the parameters utilized in the mathematical model for calculating the flashover voltage. Flashover voltage was calculated using (1). These calculations were based on the input parameters provided in Table I, along with specific values of ESDD (C, mg/cm²), specifically: 0.02, 0.03, 0.04, 0.05, 0.06, 0.13, 0.16, 0.23, 0.28, 0.34, 0.37, 0.49, 0.52, and 0.55 [9].

TABLE I. VALUES UTILIZED IN THE MATHEMATICAL MODEL [9]

Insulator	*H* (cm)	*D* (cm)	*F*	*L* (cm)
(T1)	15.9	26.8	0.79	33
(T2)	15.9	26.8	0.86	40.6
(T3)	16.5	25.4	0.90	43.2
(T4)	14.6	25.4	0.72	31.8
(T5)	15.9	29.2	0.92	47.0
(T6)	15.6	27.2	0.76	36.8
(T7)	17.8	32.1	0.96	54.6
(T8)	17.0	28.0	0.80	37.0
(T9)	14.5	25.4	0.74	30.5
(T10)	16.5	20.0	1.29	40.0

The experiments were conducted at an insulator test station located in the High Voltage Laboratory of the Public Power Corporation's Testing, Research, and Standards Center in Athens [8], following guidelines specified by IEC standards [12]. This station primarily comprises two chambers: a pollution chamber and a fog chamber. Tests were performed on artificially polluted insulators to determine their critical flashover voltage. Table II summarizes the experimental results obtained from this station.

TABLE II. OBSERVED EXPERIMENTAL DATA [20]

Insulator	*L* (cm)	*D* (cm)	*H* (cm)	F	*C* (mg/cm²)	U_c (KV)
Type 1	27.9	25.4	14.6	0.68	0.13	12.0
	27.9	25.4	14.6	0.68	0.16	11.1
	27.9	25.4	14.6	0.68	0.23	8.10
	27.9	25.4	14.6	0.68	0.28	9.10
	27.9	25.4	14.6	0.68	0.34	7.50
	27.9	25.4	14.6	0.68	0.37	7.80
	27.9	25.4	14.6	0.68	0.49	6.80
	27.9	25.4	14.6	0.68	0.52	6.20
	27.9	25.4	14.6	0.68	0.55	6.10
Type 2	30.5	25.4	14.6	0.70	0.02	22.0
	30.5	25.4	14.6	0.70	0.05	16.0
	30.5	25.4	14.6	0.70	0.10	13.0
	30.5	25.4	14.6	0.70	0.16	11.0
	30.5	25.4	14.6	0.70	0.22	10.0
	30.5	25.4	14.6	0.70	0.30	8.50
Type 3	43.2	25.4	14.6	0.92	0.02	26.0
	43.2	25.4	14.6	0.92	0.05	19.0
	43.2	25.4	14.6	0.92	0.10	15.0
	43.2	25.4	14.6	0.92	0.16	13.0
	43.2	25.4	14.6	0.92	0.22	12.0
	43.2	25.4	14.6	0.92	0.30	10.5
Type 4	43.2	22.9	16.6	1.38	0.02	23.5
	43.2	22.9	16.6	1.38	0.03	20.9
	43.2	22.9	16.6	1.38	0.04	19.4
	43.2	22.9	16.6	1.38	0.05	18.3
	43.2	22.9	16.6	1.38	0.06	16.9
	43.2	22.9	16.6	1.38	0.10	15.8
	43.2	22.9	16.6	1.38	0.20	13.6

B. Case Studies

To effectively evaluate the predictive accuracy of the ANN-PSO, it is necessary to use a test dataset that was not used in the training phase. In the process of training and testing machine learning models, it is crucial to divide the complete dataset appropriately for each stage. Research has indicated that allocating 70-80% of the data for training and 20-30% for testing typically yields the most effective outcomes [21-22].

The dataset utilized in this study comprises 140 estimated values derived from a mathematical model and 28 real values obtained from experimental results. This study explored several cases to evaluate the performance of the proposed model, using a combination of data from the mathematical model and real experimental values. Table III provides a comprehensive overview of various case studies, highlighting the amount of mathematical and experimental data associated with each case. Furthermore, it specifies the number of data points allocated for training and testing purposes in each scenario, offering valuable insights into the dataset distribution. In each case, the training and testing datasets consist of five input parameters (L, H, D, F, and σ) with the output representing the critical flashover voltage V_c. This systematic approach ensures a robust evaluation of the model's ability to predict flashover voltage across various data distributions. The test dataset is utilized to assess the models' predictive precision. This approach aims to ensure that the model learns from the given data with minimal bias and training error, while also maintaining the capacity to generalize new data effectively with reduced variance and test error.

TABLE III. DIFFERENT CASES STUDIED

Cases		1	2	3	4	5
Training data	Mathematical model	140	140	140	140	-
	Experimental values	-	4	8	20	24
	Total of training data	140	144	148	160	24
Testing data	Mathematical model	-	-	-	-	-
	Experimental values	28	24	20	8	4
	Total of testing data	28	24	20	8	4

C. ANN Architecture and Optimization Approach

ANNs are similar in principle to the biological systems of humans and other animals, and because of this, they have become a powerful analysis tool for complex data sets [19]. They are particularly good at finding non-trivial relations between the inputs and outputs, even when the dataset is noisy and complex. MLP is one of the most commonly used neural network architectures, and its mathematical formulation can be described as [9]:

$$y_j = F\left(\sum_{j=1}^{n} w_{kj}E_j + B_j\right) \quad (6)$$

where y_j is the neuron output in the current layer, F is the activation function, w_{kj} and B_j are the weights and biases, respectively, and E_j represents the node values from the preceding layer.

D. Particle Swarm Optimization (PSO)

The PSO algorithm [23] is inspired by the collective behavior observed in animal groups, such as flocks of birds and schools of fish. A semi-evolutionary swarm intelligence algorithm is one way to describe this particular method. The process is driven by randomly selecting and testing solutions and then using the results to find, step by step, a better one. Every solution scanned in this process is attached to a search strategy that works at the speed and in the memory of the best condition it has ever been exposed to. Three critical elements play a crucial role: position, velocity, and fitness. To address an optimization issue using PSO, the steps are as follows:

- Generate an initial population of particles with random positions and velocities within the problem space.
- Calculate the fitness value for each particle.
- Update the particle positions and velocities based on [24]:

$$V_i^{(t+1)} = wV_i^{(t)} + C_1 r_1\left(P_{best,i} - X_i^{(t)}\right) + C_2 r_2\left(G_{best} - X_i^{(t)}\right) \quad (7)$$

$$x_i^{(t+1)} = x_i^t + v_i^{(t+1)} \quad (8)$$

where $V_i^{(t)}$ is the velocity of particle i at iteration t, $X_i^{(t)}$ is the position of particle i at iteration t, $P_{best,i}$ is the best position found by particle i, G_{best} is the global best position found by the swarm, w is the inertia weight, controlling the influence of the previous velocity, C_1 and C_2 are acceleration coefficients that determine the relative importance of personal best and global best solutions, and r_1 and r_2 are random numbers uniformly distributed in the range $[0,1]$.

One of the key advantages of PSO is its simplicity in coding and its low computational cost [24]. Given its effectiveness in finding global optima, the PSO algorithm was utilized to train the MLP in this study.

E. Model Development

This study employed a hybrid modeling approach that integrates PSO with an ANN to predict the critical flashover voltage of insulators. The PSO algorithm was used to optimize the weights and biases of the neural network, thus improving its performance in terms of both the training speed and the prediction accuracy. Figure 2 presents the workflow of the proposed ANN-PSO prediction model. As outlined in Figure 2, the development phase of the PSO-based ANN prediction model involves the following procedure:

1. All inputs are normalized to $[-1, 1]$ to have standard minimum and maximum values for every data input.
2. The ANN prediction model involves training with datasets to determine optimal parameters that accurately predict flashover voltage [9]. To ensure model optimization, the neural network must be trained using mathematically processed input data, enabling it to effectively capture underlying relationships and enhance predictive accuracy.
3. Prepare the data to train and validate the best parameters for the ANN-PSO model.
4. Initialize the following parameters: PSO weights coefficient (C_1 and C_2); the number of particles (i); the size of the swarm (n); the number of hidden layers; NN weights and biases. ANN-PSO parameters should be previously selected to obtain the results.
5. The model is trained and tested based on p_{best} and g_{best} in the entire swarm.
6. Particles' speed and location are upgraded based on the best fitness values. Fitness evaluation is based on MSE

$$MSE = \left(\frac{1}{N}\right)\sum_i |t_i - o_i|^2 \quad (9)$$

where t denotes the target value, o is the produced output, and N is the total number of outputs in the network's output layer.

7. The search process to find the best position of p_{best} and g_{best} proceeds until the end condition is satisfied. Training stops when it achieves the best accepted value.
8. The results of the ANN-PSO model are compared with real data. The ANN-PSO development was investigated based on trial and error. The robustness of the ANN-PSO model was determined based on the accuracy of prediction values that provide the least percentage error and Root Mean Square Error (RMSE) values.

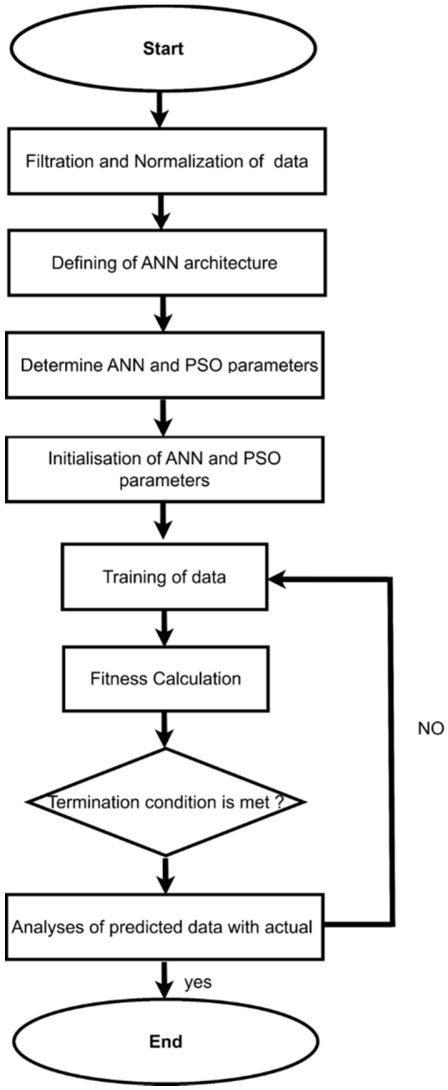

Fig. 2. ANN-PSO flow diagram.

The ANN was structured as a three-layer feedforward architecture comprising an input layer, a single hidden layer, and an output layer. The hidden layer consisted of 10 neurons, a choice justified by empirical testing and careful performance evaluation. Specifically, configurations with fewer neurons (ranging from 3 to 7) demonstrated inadequate learning capacity, resulting in underfitting, whereas configurations exceeding 12 neurons were prone to overfitting and extended training durations. Consequently, the 10-neuron configuration was found to be optimal, achieving the lowest validation error and maintaining robust generalization capabilities across multiple experimental trials. The ANN exhibited the best overall performance when trained using the Levenberg-Marquardt learning algorithm, employing a sigmoid tangent (Logsig) activation function in the hidden layer and a linear (Purelin) activation function in the output layer.

PSO was integrated into the ANN to enhance prediction performance. Careful parameter selection for PSO is essential, particularly in determining suitable coefficients within the velocity equation and the optimal swarm size. This parameter analysis involved defining the maximum iteration count, the number of neurons, and the hidden layer configuration in the feedforward ANN architecture. Initial weights were randomly assigned values within $[-1, 1]$. The ANN-PSO model was implemented and evaluated in MATLAB, aiming to identify the most effective swarm size, while parameters C_1 and C_2 were consistently maintained at a fixed value of 2.

Figure 3 illustrates the impact of particle number on the convergence behavior of the ANN-PSO by plotting the training MSE against the number of iterations for three swarm sizes: 5, 10, and 20 particles. As observed, the model with 20 particles consistently outperformed the others in terms of both convergence speed and final error level, demonstrating a steep decline in MSE and reaching values below 0.015 after 2000 iterations. In contrast, models with 5 and 10 particles converged more gradually, and their final MSE values plateaued at higher levels, around 0.02 and 0.025, respectively. These results indicate that increasing the number of particles enhances the swarm's ability to explore the search space effectively, thus improving the optimization of the ANN parameters. However, it is important to note that while larger particle numbers may yield improved accuracy, they also demand greater computational resources. Therefore, an appropriate trade-off between accuracy and efficiency must be considered when selecting the swarm size.

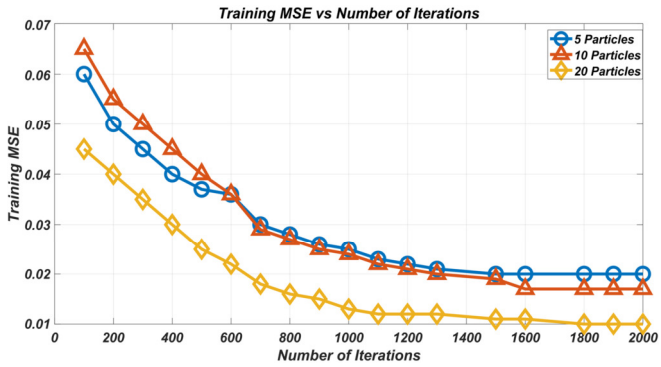

Fig. 3. Performance of ANN-PSO with different swarm sizes.

Maximum iterations and swarm size remained same for testing the other ANN-PSO parameters. Table III summarizes the combinations of acceleration coefficients used based on [23].

TABLE IV. OPTIMUM VALUES OF ACCELERATION COEFFICIENTS

Relationship	C_1	C_2	MSE
C_1=0.25C_2	0.8	3.2	0.0018
C_1=0.5C_2	1.3	2.7	0.002
C_1=0.75C_2	1.7	2.3	0.0021
C_1=0.25C_2	3.2	0.8	0.0017
C_1=0.5C_2	2.7	1.3	0.0018
C_1=0.75C_2	2.3	1.7	0.0019
C_1=C_2	2.5	2.5	0.0015
C_1=C_2	2	2	0.0021
C_1=C_2	1.75	1.75	0.0018
C_1=C_2	1.5	1.5	0.0021
C_1=C_2	1.25	1.25	0.0021
C_1=C_2	1	1	0.0025

Table III shows that the configuration where C_1 and C_2 equal 2.5 yields the lowest MSE value. Most MSE results for the testing data across different coefficient combinations were closely aligned. This suggests that the relationship between C_1 and C_2 is relatively well-defined and consistent in its impact on model performance. The parameters that give better prediction are the following: inertia weight ranges from 0.5 to 0.9, swarm size of 1000, cognitive component $C_1 = 2.5$, 100 iterations, social component $C_2 = 2.5$.

III. PERFORMANCE EVALUATION METRICS

This study selected various performance metrics to evaluate the proposed models and identify the most accurate for predicting the flashover output voltage. These metrics included the coefficient of determination (R^2), RMSE, and MAPE. The following formulas were applied to calculate these indices [9]:

$$R^2 = 1 - \frac{\sum_{k=1}^{n}\left(y_{tes,k} - y_{pre,k}\right)^2}{\sum_{k=1}^{n}\left(y_{tess,k} - \overline{y_{tes,k}}\right)^2} \quad (10)$$

$$MAPE = 100\%.\frac{\sum_{k=1}^{n}\left|y_{tes,k} - y_{pre,k}\right|/y_{tes,k}}{n} \quad (11)$$

$$RMSE = \left\{\frac{\sum_{k=1}^{n}\left(y_{tes,k} - y_{pre,k}\right)}{n}\right\}^{1/2} \quad (12)$$

IV. RESULTS AND DISCUSSION

Several cases were considered to evaluate the performance of the proposed model. For this study, Case 2 was selected to provide a detailed analysis and allow for further comparisons with the results of other studies. Figures 4 and 5 present the simulation results for the predicted flashover voltage of polluted insulators, assessed using both training and testing datasets. In Figure 4, the training results of the ANN-PSO model are compared with the values obtained from the mathematical model. This comparison, conducted using a specific subset of data from the training pattern series, highlights the accuracy and reliability of the ANN-PSO model in capturing the expected behavior. Furthermore, the results clearly demonstrate that the ANN-PSO model achieved optimal performance in the training dataset.

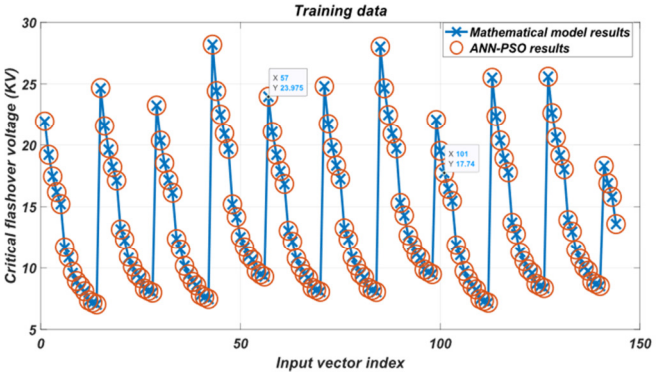

Fig. 4. Training performance of the ANN-PSO model.

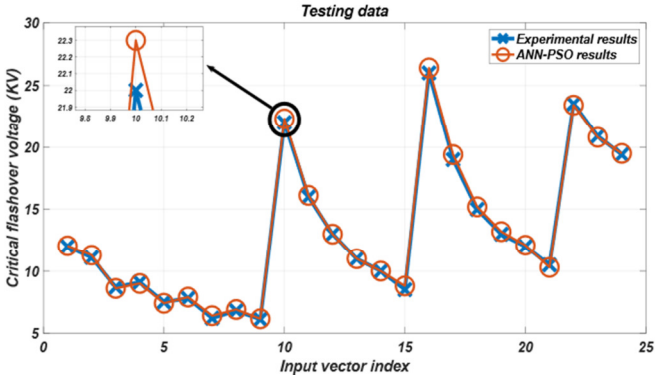

Fig. 5. Testing performance of the ANN-PSO model.

Figure 5 shows a comparison between the predicted data and test data, evaluating the predictive performance of the ANN-PSO model. This analysis demonstrates the model's ability to accurately estimate the critical flashover voltage for four distinct types of insulators. The results highlight the robustness and versatility of the ANN-PSO model in adapting to various insulator types and accurately generalizing its predictions to unseen test data. This capability underscores the model's potential for practical applications in predicting flashover voltage under varying conditions, contributing to the development of reliable and efficient solutions for polluted insulators. An effective approach to assess the accuracy of ANN-PSO is to analyze the correlation between the actual critical flashover voltage U_c and the values predicted by the model. A correlation value closer to 1 signifies a higher level of model performance and precision. Figures 6 and 7 display the correlation between the estimated versus actual values of U_c for the four insulator types for the training and testing data. The data points almost perfectly align with the line of best-fit, demonstrating the model's strong ability to accurately predict the duty ratio for the test dataset. Specifically, the correlation for the training data reached 0.9998, while for testing data it was 0.9997, showcasing the model's high accuracy in both scenarios.

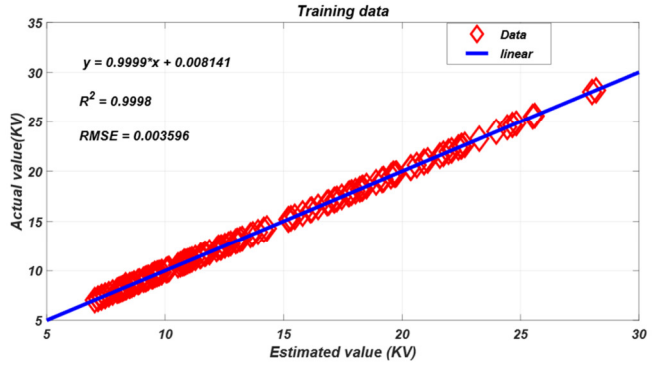

Fig. 6. Correlation between predicted and actual critical flashover voltage values for training data.

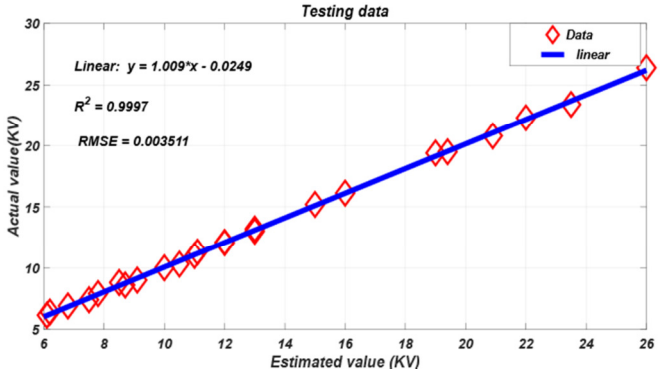

Fig. 7. Correlation between predicted and actual critical flashover voltage values for testing data.

Figure 8 provides a comparative analysis of the proposed model's results against experimental data from [17-18] and those of the LSSVM-PSO and LSSVM-ACO algorithms [20, 25]. The analysis highlights that the simulated results of the proposed model not only align closely with the experimental data but also demonstrate superior agreement compared to the other algorithms, underscoring its enhanced accuracy and reliability.

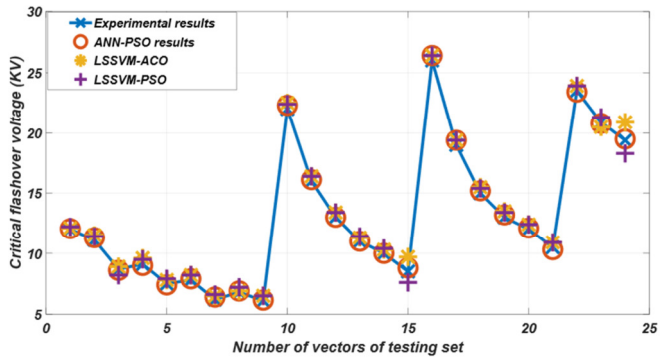

Fig. 8. Comparison of results with other models.

Given the difficulty in visually representing the accuracy and precision of the proposed model compared to previous approaches, statistical metrics such as RMSE, MAPE, and R^2 were used for evaluation. The validation results, summarized in Table V, clearly show that the proposed model achieved the highest regression accuracy. Additionally, in multiple cases, ANN-PSO outperformed other models by achieving the lowest RMSE and MAPE values and the highest R^2 values for the test dataset, highlighting its superior predictive performance.

TABLE V. COMPARISON OF ANN-PSO AGAINST OTHER MODELLING APPROACHES

Case	Models	Testing		
		RMSE	R^2	MAPE
1	ANN-PSO	0.0154	0.994	0.0114
	ANN [9]	-	-	-
	ANFIS [26]	-	-	-
	LSSVM -PSO [20]	-	-	-
	LSSVM-GWO [14]	0.0191	0.9921	0.0129
2	ANN-PSO	0.003511	0.997	0.842
	ANN [9]	0.4929	0.97	3.74
	ANFIS [26]	0.4766	0.9888	3.5185
	LSSVM -PSO [20]	0.0149	0.9934	1.0146
	LSSVM-GWO [14]	0.0132	0.9948	0.8891
3	ANN-PSO	0.0105	0.998	0.795
	ANN [9]	-	-	-
	ANFIS [26]	0.4766	0.9888	3.5185
	LSSVM -PSO [20]	0.0126	0.9972	0.8532
	LSSVM-GWO [14]	0.0125	0.9972	0.8298
4	ANN-PSO	0.0061	0.9994	0.3945
	ANN [9]	-	-	-
	ANFIS [26]	0.0472	0.9987	0.4610
	LSSVM -PSO [20]	0.0089	0.9984	0.5758
	LSSVM-GWO [14]	0.0064	0.9992	0.4196
5	ANN-PSO	0.0034	0.9999	0.2254
	ANN [9]	-	-	-
	ANFIS [26]	0.0417	0.9991	0.4072
	LSSVM -PSO [20]	0.0088	0.9990	0.5174
	LSSVM-GWO [14]	0.0037	0.9998	0.2458

The ANN-PSO model demonstrates superior predictive capability for flashover voltage compared to traditional ANN, ANFIS, LS-SVM-PSO, and LS-SVM-GWO models, largely due to its ability to efficiently optimize model parameters. Unlike conventional ANN, which relies on gradient-based optimization techniques that are prone to becoming trapped in local minima, the PSO algorithm utilizes a population-based search mechanism. This approach systematically explores the solution space using the collaborative behavior and information-sharing principles of particle swarms. As a result, PSO ensures the determination of near-optimal weights and biases for the ANN, significantly enhancing the model's accuracy, convergence speed, and generalization capabilities. In contrast, while LS-SVM has demonstrated commendable performance in certain predictive applications, its dependence on hyperparameter tuning methods, such as trial-and-error or grid search, poses limitations. These techniques are not only computationally intensive but also less adaptive, especially in complex dynamic scenarios. By integrating the adaptive and dynamic search mechanism of PSO, the ANN-PSO model circumvents these inefficiencies, ensuring a more robust optimization process that aligns with the multifactorial nature of flashover voltage prediction.

Furthermore, the ANN-PSO model excels in capturing complex nonlinear relationships and handling high-dimensional input data with greater precision. This is particularly crucial for

the prediction of flashover voltage, where multiple interrelated parameters, such as insulator spacing, leakage length, diameter, and environmental conditions, must be accurately modeled. The superior ability of ANN-PSO to adapt to these intricate interactions translates into improved predictive performance.

V. CONCLUSION

This study developed an ANN optimized using PSO to accurately predict the flashover voltage of polluted insulators. The ANN-PSO model effectively overcame the limitations of traditional methods by providing more reliable and precise predictions across different pollution levels and insulator types. Performance evaluations revealed that the ANN-PSO model surpassed conventional approaches, including standalone ANN, LS-SVM-PSO, and LS-SVM-GWO, in key accuracy metrics such as RMSE, MAPE, and R^2. A comparison of the model's predictions with experimental data and the results from other algorithms demonstrated its robustness and reliability, further validating its effectiveness. The results emphasize the significant potential of the ANN-PSO model for practical applications in power transmission systems, enabling better monitoring and prediction of insulator flashover under polluted conditions. This enhanced predictive capability provides a valuable tool for power utilities to improve system reliability, reduce maintenance costs, and mitigate the risk of flashover incidents. The proposed model is well-suited for integration into real-world deployment frameworks, such as SCADA-based monitoring systems, smart grid platforms, and predictive maintenance tools. Its ability to deliver real-time prediction can assist grid operators in making informed decisions to prevent insulation failure and ensure uninterrupted power delivery.

To construct a more comprehensive and adaptive predictive framework, future research could enhance this study by integrating additional environmental and climatic variables, such as temperature, humidity, wind speed, and varying levels of pollution. Incorporating real-time monitoring data from power systems would further enhance the precision and applicability of the model in dynamic operational settings. Moreover, exploring hybrid optimization techniques or ensemble learning approaches could enhance the model's performance, improving its predictive accuracy and robustness under complex scenarios.

REFERENCES

[1] D. Doufene, S. Benharat, S. Bouazabia, and S. A. Bessedik, "Hybrid Grey Wolf and Finite Element Method (GWO-FEM) Algorithm for Enhancing High Voltage Insulator String Performance in Wet Pollution Conditions," *Engineering, Technology & Applied Science Research*, vol. 12, no. 3, pp. 8765–8771, Jun. 2022, https://doi.org/10.48084/etasr.4978.

[2] M. Chen, H. Yang, Z. Song, F. Zhou, W. Shen, and L. Zhang, "Development path prediction of local arc over the wet contaminated insulator surface based on random walk theory," *Electric Power Systems Research*, vol. 241, Apr. 2025, Art. no. 111353, https://doi.org/10.1016/j.epsr.2024.111353.

[3] Y. Guo *et al.*, "Dynamic prediction of insulator pollution flashover and analysis of factors influencing pollution flashover-Part II," *IEEE Transactions on Dielectrics and Electrical Insulation*, pp. 1–1, 2025, https://doi.org/10.1109/TDEI.2025.3552014.

[4] R. Saleh, A. Musyafa, D. D. Risanti, I. Hasan, and Azhar, "Flashover Model of Polluted Medium Voltage AC Indoor Insulators Around Electric Arc Nickle Furnace Roof," in *2024 IEEE 4th International Conference in Power Engineering Applications (ICPEA)*, Pulau Pinang, Malaysia, Mar. 2024, pp. 187–192, https://doi.org/10.1109/ICPEA60617.2024.10498565.

[5] M. Faramarzi Palangar, M. Mirzaie, and A. Mahmoudi, "Improved flashover mathematical model of polluted insulators: A dynamic analysis of the electric arc parameters," *Electric Power Systems Research*, vol. 179, Feb. 2020, Art. no. 106083, https://doi.org/10.1016/j.epsr.2019.106083.

[6] H. Benguesmia, B. Bakri, Y. Mekkas, and N. M'ziou, "Experimental Study of the Flashover Process and the Leakage Current on the Surface of High Voltage Insulator Under AC Voltage," *Journal of Electrical Engineering & Technology*, vol. 20, no. 4, pp. 2553–2567, May 2025, https://doi.org/10.1007/s42835-024-02103-3.

[7] K. Belhouchet, I. Ghadbane, A. Zemmit, L. Ouchen, and A. Zorig, "Measurement and evaluation of the flashover voltage on polluted cap and pin insulator: An experimental and theoretical study," *Electric Power Systems Research*, vol. 236, Nov. 2024, Art. no. 110979, https://doi.org/10.1016/j.epsr.2024.110979.

[8] K. Kumar and R. V. Maheswari, "An experimental study of the flashover performances of outdoor high-voltage polymeric insulators under fan-shaped pollution and dry band location," *Electrical Engineering*, vol. 107, no. 5, pp. 5719–5735, May 2025, https://doi.org/10.1007/s00202-024-02825-7.

[9] L. Taibaoui, B. Zegnini, and A. Mahdjoubi, "An Approach To Predict Flashover Voltage on Polluted Outdoor Insulators Using ANN," in *2022 19th International Multi-Conference on Systems, Signals & Devices (SSD)*, Sétif, Algeria, May 2022, pp. 1842–1847, https://doi.org/10.1109/SSD54932.2022.9955667.

[10] H. Becha, "Estimation of the Flashover Voltage of Insulator Using Fuzzy Logic (FL)," *Prezglad Elektrotechniczny*, vol. 1, no. 7, pp. 103–109, Jul. 2024, https://doi.org/10.15199/48.2024.07.21.

[11] S. He *et al.*, "Intelligent prediction of 110kV insulator lightning flashover criteria based on random forest," *Electric Power Systems Research*, vol. 232, Jul. 2024, Art. no. 110423, https://doi.org/10.1016/j.epsr.2024.110423.

[12] M. Abdelhalim, Z. Boubakeur, and M. Belkheiri, "Prediction of critical flashover voltage of polluted insulators under sec and rain conditions using least squares support vector machines (LS-SVM)," *Diagnostyka*, vol. 20, no. 1, pp. 49–54, Nov. 2018, https://doi.org/10.29354/diag/99854.

[13] A. Belkebir, Y. Bourek, and H. Benguesmia, "Adaptive Neuro-Fuzzy Inference System Application of Flashover Voltage of High-Voltage Polluted Insulator," *Journal of Electrical Engineering & Technology*, vol. 19, no. 6, pp. 3839–3849, Aug. 2024, https://doi.org/10.1007/s42835-024-01862-3.

[14] S. A. Bessedik, R. Djekidel, and A. Ameur, "Performance of different kernel functions for LS-SVM-GWO to estimate flashover voltage of polluted insulators," *IET Science, Measurement & Technology*, vol. 12, no. 6, pp. 739–745, Sep. 2018, https://doi.org/10.1049/iet-smt.2017.0486.

[15] A. Mahdjoubi, B. Zegnini, M. Belkheiri, and T. Seghier, "Fixed least squares support vector machines for flashover modelling of outdoor insulators," *Electric Power Systems Research*, vol. 173, pp. 29–37, Aug. 2019, https://doi.org/10.1016/j.epsr.2019.03.010.

[16] F. V. Topalis, I. F. Gonos, and I. A. Stathopulos, "Dielectric behaviour of polluted porcelain insulators," *IEE Proceedings - Generation, Transmission and Distribution*, vol. 148, no. 4, 2001, Art. no. 269, https://doi.org/10.1049/ip-gtd:20010258.

[17] K. Ikonomou, G. Katsibokis, G. Panos, and I. A. Stathopoulos, "Cool fog tests on artificially polluted insulators," presented at the Fifth International Symposium on High Voltage Engineering, 1987, vol. 2.

[18] Z. Guan and R. Zhang, "Calculation of DC and AC flashover voltage of polluted insulators," *IEEE Transactions on Electrical Insulation*, vol. 25, no. 4, pp. 723–729, Aug. 1990, https://doi.org/10.1109/14.57096.

[19] R. Sundararajan and R. S. Gorur, "Dynamic arc modeling of pollution flashover of insulators under DC voltage," *IEEE Transactions on Electrical Insulation*, vol. 28, no. 2, pp. 209–218, Apr. 1993, https://doi.org/10.1109/14.212246.

[20] S. A. Bessedik and H. Hadi, "Prediction of flashover voltage of insulators using least squares support vector machine with particle swarm optimisation," *Electric Power Systems Research*, vol. 104, pp. 87–92, Nov. 2013, https://doi.org/10.1016/j.epsr.2013.06.013.

[21] T. F. Thien and W. S. Yeo, "A comparative study between PCR, PLSR, and LW-PLS on the predictive performance at different data splitting ratios," *Chemical Engineering Communications*, vol. 209, no. 11, pp. 1439–1456, Nov. 2022, https://doi.org/10.1080/00986445.2021.1957853.

[22] M. Belkin, D. Hsu, S. Ma, and S. Mandal, "Reconciling modern machine-learning practice and the classical bias–variance trade-off," *Proceedings of the National Academy of Sciences*, vol. 116, no. 32, pp. 15849–15854, Aug. 2019, https://doi.org/10.1073/pnas.1903070116.

[23] J. Guo, C. Chen, H. Wen, G. Cai, and Y. Liu, "Prediction model of goaf coal temperature based on PSO-GRU deep neural network," *Case Studies in Thermal Engineering*, vol. 53, Jan. 2024, Art. no. 103813, https://doi.org/10.1016/j.csite.2023.103813.

[24] B. Amel, B. Yacine, and B. Hani, "Particle swarm optimization of a neural network model for predicting the flashover voltage on polluted cap and pin insulator," *Diagnostyka*, vol. 23, no. 3, pp. 1–7, Sep. 2022, https://doi.org/10.29354/diag/154051.

[25] S. A. Bessedik, R. Djekidel, and A. Ameur, "Improved LS-SVM using ACO to estimate flashover voltage of polluted insulators," *Journal of Engineering Science and Technology*, vol. 12, no. 1, pp. 62–77, 2017.

[26] K. Erenturk, "Adaptive-Network-Based Fuzzy Inference System Application to Estimate the Flashover Voltage on Insulator," *Instrumentation Science & Technology*, vol. 37, no. 4, pp. 446–461, Jun. 2009, https://doi.org/10.1080/10739140903087873.