Electric Field Features and Its Application for Air Gap Breakdown Voltage Prediction

Zhibin Qiu and Jiangjun Ruan

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.72230

Abstract

Air gap breakdown voltage prediction by mathematical calculations instead of experiments has been a long sought goal in the area of high-voltage engineering. In this chapter, a prediction method is proposed based on the electric field features and support vector machine (SVM). Two sets of electric field features are defined on the shortest interelectrode path of sphere-sphere and rod (sphere)-plane gap to characterize their spatial structures, which can be extracted from the electric field calculation results by finite element method (FEM). A breakdown voltage prediction model is established by SVM, while the input parameters are the electric field features, and the output parameters are $C_0$ and $1$, respectively, characterizing withstanding and breakdown of an air gap under the applied voltage. The proposed method is used to predict the power frequency breakdown voltages of IEC standard sphere-sphere air gaps and the switching impulse discharge voltages of large sphere-plane air gaps. The prediction results coincide well with the experimental data, the mean absolute percentage error of the 260 sphere-sphere gaps is within 2% and that of the 16 sphere-plane gaps is 3.2%. The results preliminarily validate the validity and accuracy of the proposed method for air gap breakdown voltage prediction.

Keywords: electric field features, air gap discharge, breakdown voltage prediction, support vector machine (SVM), sphere-sphere gap, sphere-plane gap

1. Introduction

With the rapid development of numerical computation methods and calculation capability of computers, the multi-physics coupling analysis has been widely applied in structure design and condition assessment of electrical equipment [1–4]. The distributions of electric, magnetic, stress, and thermal field can be calculated by commercial software, and multi-physics problems in complex structures can be simulated using powerful numerical techniques, so as to
guide the optimal design of machines and devices. However, up to now, the insulation design of high-voltage equipment still depends on experimental verification.

The electric field distributions of complex geometries can be simulated very well, that is because electric field can be defined according to Gauss’s law and Maxwell’s relation, with clear constitutive relation and governing equations [5, 6]. Several numerical methods were developed for electric field calculation, such as finite element method (FEM), finite volume method (FVM), boundary element method (BEM), and those combining the advantages of different methods, like the control volume-based finite element method (CVFEM) [7–9]. However, even though the electric field distribution can be calculated accurately, the insulation strengths of dielectrics cannot be calculated or predicted, that is because there are no governing equations to describe the discharge process. Hence, the insulation calculation or prediction can be viewed as the short board for virtual design and manufacturing of high-voltage electrical equipment.

The dielectrics used in electric power system include air, sulfur hexafluoride, transformer oil, electroceramics, and silicone rubber, etc. while the most commonly used is the air. Even though air discharge phenomena have been extensively studied for more than 100 years, both experimentally and theoretically, the air discharge theory is still imperfect, and therefore the air gap breakdown voltage prediction is still a great challenge and one of the most important issues to be solved for external insulation design of power transmission and transformation projects. The Townsend discharge theory [10], streamer discharge theory [11, 12], and leader discharge theory [13–16] lay the foundation for the analysis and interpretation of air discharge phenomena. These classical theories were summarized from experimental measurements and observations, and they can be used to describe air discharge process. However, due to the complexity and randomness of air discharge process, it is difficult to realize air gap discharge voltage calculation with different gap configurations or under different voltage waveshapes and atmospheric environment.

Currently, discharge tests are still the main method to obtain air gap breakdown voltages. Various empirical [17, 18] and semiempirical [19, 20] formulas have been summarized for applications. However, the experimental studies cannot exhaustively reproduce all the gap configurations in actual engineering [21]. The discharge tests are usually conducted on air gaps with typical electrodes, like the rod-plane gaps, to obtain the basic characteristics of long air gap discharge, and the fitting formulas obtained by the experimental results are used to predict the breakdown voltages of engineering gaps, such as the conductor-tower window gaps, by taking the gap factor into consideration. However, the gap factors of different gap configurations should be determined by time-consuming and costly full-scale tests. When generalized to other gap structures, the fitting results are with certain deviations.

In order to find an alternative to replace the discharge tests, many scholars are devoted to the studies of air discharge mechanisms and try to achieve air gap breakdown voltage prediction by theoretical calculation. Based on air discharge theories and some simplifications and assumptions, numerous physical models have been proposed to simulate the entire duration of the discharge, including the first corona inception, streamer propagation, leader inception and propagation, and final jump [22–30]. However, long air gap discharge researches are faced
with some challenges, mainly including the inherent complexity of the discharge process and the scientific modeling, the limitations of observation and measurement of various physical quantities involved in the discharge process, and the uncertainty of the solutions obtained by these semiempirical models [31]. Hence, there are obvious errors between the computed results of these models and the experimental measurement results. At present, the research hotspot is still concentrated on further revealing the physical mechanisms of the complex discharge process, so as to establish more scientific simulation models, but the realization of discharge prediction still needs long-term studies.

The idea of predicting air gap breakdown voltage by some artificial intelligence algorithms has attracted interests of researchers for many years, but there are only a very few related studies and applications. The artificial neural network (ANN) [32, 33], fuzzy logic system [34, 35], and support vector machine (SVM) [36, 37] have been applied to predict the discharge voltages of air insulation gaps. In [36, 37], a method based on electric field features and SVM was proposed for discharge voltage prediction of air gaps, and it has been successfully applied to predict the breakdown voltages of air gaps with typical and atypical electrodes [38–40] and the corona onset voltages of rod-plane gaps, conductors, and valve hall fittings [41, 42]. Some features were extracted from the calculation results of the electric field distribution of an air gap to characterize its spatial structure, and the SVM was applied to establish the multidimensional nonlinear relationships between these features and the air gap discharge voltage. This method offers a possible way to achieve breakdown voltage calculation of air gaps, so as to guide the external insulation design of high-voltage electrical equipment.

Under a given applied voltage waveform and specific atmospheric environment, air gap breakdown voltage is determined by its structure, which can be characterized by the electrostatic field distribution. In previous studies, the electric field features were defined in the hypothetic discharge channel between two electrodes, on the surface of typical high-voltage electrode and on the shortest interelectrode discharge path, which contains different regions classified from the perspectives of volume, area, and line. However, the discharge channel and the electrode surface are difficult to be defined for air gaps with complicated configurations. Only the shortest interelectrode path can be defined for two-electrode air gaps with arbitrary structures. It is known that there is a one-to-one correspondence between the gap structure and the electric field distribution. If the distribution characteristics of a three-dimensional spatial electric field can be mapped to those along a one-dimensional path, it is possible to define some effective features on this path to characterize the gap structure and establish their relationships with the air gap breakdown voltage by SVM, so as to be applied for breakdown voltage prediction of air gaps with arbitrary structures.

In this chapter, the above ideas are achieved preliminarily. Two sets of electric field features are defined on the shortest path of sphere-sphere gaps and rod (sphere)-plane gaps, respectively, with the U-shaped curve and monotonously decreased curve of the electric field distribution along the shortest path. A prediction model is established by SVM, and the electric field features on the shortest path are taken as the input parameters, while the output parameters are \( -1 \) and \( 1 \), respectively, characterizing whether a gap will withstand or breakdown under a given applied voltage. This model is applied to predict the breakdown voltages of sphere-sphere air
gaps and large sphere-plane air gaps, and the predicted results are compared with the experimental data given in IEC 60052 and other references to verify its validity and accuracy.

2. Electric field features

2.1. Electric field distributions along the shortest path

The sphere-sphere air gaps given in IEC 60052 [43] (or IEEE Std 4 [44]) and the large sphere-plane air gaps given in [45, 46] are taken as the samples for breakdown voltage prediction studies in this chapter. The schematic diagrams of the sphere-sphere and the sphere-plane gaps are shown in Figure 1, where \( d \) is the gap distance. For sphere-sphere gaps, the sphere diameter \( D \) ranges from 5 to 200 cm, and the experimental breakdown voltages of these gaps are given in [43, 44], with different sphere diameters and gap distances. For the large sphere-plane air gap, it is composed of the sphere electrode with a mounting rod tube and the grounded plane electrode. The experimental 50% discharge voltages of these gaps, with the sphere diameter ranging from 25 to 95 cm and the gap distance ranging from 2 to 5 m, are given in [45], under the standard switching impulse voltage.

Since the sphere-sphere and the sphere-plane air gaps are with axisymmetric structures, the two-dimensional axisymmetric models are established by ANSYS, a finite element analysis software, to calculate their electric field distributions. The high-voltage sphere electrodes are applied unit voltage 1 V, and the grounded sphere electrode or the grounded plane electrode is applied zero potential. Taking a sphere-sphere gap with \( D = 10 \) cm and \( d = 3 \) cm and a sphere-plane gap with \( D = 45 \) cm and \( d = 3 \) m as examples, the cloud charts of their electric field distributions are shown in Figure 2. It can be seen that the maximum electric field strength appears at the bottom of the high-voltage sphere electrode, both for the short sphere-sphere gap and the long sphere-plane gap.

Taking the 3 cm short sphere-sphere gaps and the 3 m long sphere-plane gaps with different sphere diameters as examples, their electric field distributions along the shortest path are

![Figure 1. Schematic diagrams of sphere-sphere and large sphere-plane air gaps.](image-url)
shown in Figure 3. For sphere-sphere gaps, the electric field distribution along the shortest path is a U-shaped curve, while the field strengths near two sphere electrodes are higher than those in the middle of the path. For sphere-plane gaps, the electric field distribution along the shortest path is a monotonously decreased curve, while the field strength gradually decreases along the path from the sphere electrode to the plane electrode. The field strength reduces quickly within 1 m from the sphere electrode, while on the path from position of 1 m to the plane electrode, the field strength changes a little. In addition, the maximum field strength decreases with the increase of the sphere diameter, both for sphere-sphere and sphere-plane gaps.

Figure 3. Electric field distributions along the shortest path of the short sphere-sphere gaps and long sphere-plane gaps.
2.2. Definitions of electric field features

According to the electric field distribution characteristics of sphere-sphere gap and sphere-plane gaps, two sets of electric field features are, respectively, defined for these two different gap types. These features mainly include the electric field strength, electric field gradient, square of electric field strength, electric field integral (electric potential), path length, and some scaling parameters related to the above quantities used to characterize the electric field inhomogeneity.

2.2.1. Sphere-sphere air gap

Some equally spaced sampling points are selected along the shortest path, and the values of their electric field strength are extracted to calculate the features. The electric field features for sphere-sphere air gap are defined as follows:

(1) Electric field strength, including the maximum values \( E_{h\text{max}} \) and \( E_{l\text{max}} \), respectively, on the surface of the high-voltage and grounded sphere electrode; the minimum value \( E_{\text{min}} \) and the average value \( E_a \); and the variance and standard deviation of the electric field strength along the shortest path \( E_{\text{std}2} \) and \( E_{\text{std}} \):

\[
E_{h\text{max}} = \max E_i \quad (i = 1, 2, \ldots, m)
\]

\[
E_{l\text{max}} = \max E_i \quad (i = m + 1, m + 2, \ldots, n)
\]

\[
E_{\text{min}} = \min E_i \quad (i = 1, 2, \ldots, n)
\]

\[
E_a = \frac{1}{n} \sum_{i=1}^{n} E_i
\]

\[
E_{\text{std}2} = \frac{1}{n} \sum_{i=1}^{n} (E_i - E_a)^2 = \frac{1}{n} \sum_{i=1}^{n} E_i^2 - E_a^2
\]

\[
E_{\text{std}} = \sqrt{E_{\text{std}2}} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} E_i^2 - E_a^2}
\]

where \( n \) is the number of the sampling points along the shortest path, \( E_i \) is the electric field strength of the \( i \)th point, and \( m \) is the point number whose electric field strength is the minimum value.

(2) Electric field gradient, including the maximum gradient \( E_{gh\text{m}} \) on the path from the point of \( E_{h\text{max}} \) to that of \( E_{\text{min}} \), the maximum gradient \( E_{gl\text{m}} \) on the path from the point of \( E_{\text{min}} \) to that of \( E_{l\text{max}} \), and the average gradient \( E_{ga} \) on the whole shortest path:

\[
E_{gh\text{m}} = \max (|\text{grad} E_i|) \quad (i = 1, 2, \ldots, m)
\]

\[
E_{gl\text{m}} = \max (|\text{grad} E_i|) \quad (i = m, m + 1, \ldots, n)
\]
where $||$ is the absolute value sign.

(3) Square of electric field strength ($W$ and $W_a$), respectively, means the integral of the field strength square on the shortest path and its expected value:

$$W = \int_0^d E_i^2 \, dl = \sum_{i=1}^n E_i^2 \, d_i$$  \hspace{1cm} (10)

$$W_a = \frac{W}{d} = \frac{1}{n} \sum_{i=1}^n E_i^2$$  \hspace{1cm} (11)

where $d$ is the gap distance and $d_i$ is the length of each segment between two sampling points on the shortest path, $d_i = d/(n-1)$.

(4) Electric field integral, that is, electric potential, including the field integral $V_x$ on the path whose field strength exceeds $E_x = x\%E_{hmax}$, the potential $V_h$ between the high-voltage sphere electrode and the point of $E_{min}$, and the potential $V_1$ between the point of $E_{min}$ and the grounded sphere electrode:

$$V_x = \int_{E_i \geq E_x} E_i \, dl \approx \sum_{E_i \geq E_x} E_i \, d_i$$  \hspace{1cm} (12)

$$V_h = \int_{E_{min} \leq E_i \leq E_{hmax}} E_i \, dl \approx \sum_{E_{min} \leq E_i \leq E_{hmax}} E_i \, d_i$$  \hspace{1cm} (13)

$$V_1 = \int_{E_{min} \leq E_i \leq E_{hmax}} E_i \, dl \approx \sum_{E_{min} \leq E_i \leq E_{hmax}} E_i \, d_i$$  \hspace{1cm} (14)

(5) Path length $L_x$, including the length $L_{Ex}$ of the path whose electric field strength exceeds $E_x = x\%E_{hmax}$, the length $L_{gx}$ of the path whose electric field gradient exceeds $E_{gx} = x\%E_{ghm}$, the length $L_{Wx}$ of the path where the sum of the electric field strength square equals to $W_x = x\%W$, and the distance $L_{min}$ from the high-voltage sphere electrode to the point of $E_{min}$:

$$L_x = \sum_{i=1}^p d_i$$  \hspace{1cm} (15)

where $p$ is the number of the points on the shortest path which meet the related conditions about the electric field strength, gradient, and square and $d_i$ is the length of each segment.

(6) Electric field inhomogeneity, characterized by some scaling parameters related to the above five kinds of features. For electric field strength, these parameters include the field distortion factor $E_{d}$, the ratio of $E_{limax}$ to $E_{hmax}$, namely, $E_{rlh}$, and the ratio of $E_{min}$ to $E_{hmax}$, namely, $E_{rm}$:
\[ E_d = \frac{E_{h\text{max}}}{E_a} \quad (16) \]
\[ E_{rlh} = \frac{E_{l\text{max}}}{E_{h\text{max}}} \quad (17) \]
\[ E_{rm} = \frac{E_{min}}{E_{h\text{max}}} \quad (18) \]

For electric field square, the scaling parameter is a ratio \( W_{rx} \) of the sum of the electric field strength square on the path where \( E_i \geq E_x = x\% E_{h\text{max}} \) to \( W \):

\[ W_{rx} = \frac{\sum E_i \geq E_x E_i^2 d_i}{W}, \quad E_x = x\% E_{h\text{max}} \quad (19) \]

For electric field integral, the scaling parameters are the ratios of \( V_{x}, V_{h}, \) and \( V_1 \) to the applied voltage \( U \), namely, \( V_{tx}, V_{rh}, \) and \( V_{tl} \):

\[ V_{tx} = \frac{V_x}{U} \quad (20) \]
\[ V_{rh} = \frac{V_h}{U} \quad (21) \]
\[ V_{tl} = \frac{V_1}{U} \quad (22) \]

For path length, the scaling parameters are the ratios of \( L_{Ex}, L_{rgx}, L_{Wx}, \) and \( L_{\text{min}} \) to the gap distance \( d \), namely, \( L_{tx}, L_{rgx}, L_{Wx}, \) and \( L_{\text{min}} \). They all have the following expression:

\[ L_{rx} = \frac{L_x}{d} \quad (23) \]

The \( x\% \) is set as 90 and 75%; therefore, the electric field features are summarized in Table 1. There are altogether 38 features for sphere-sphere air gap.

<table>
<thead>
<tr>
<th>Category</th>
<th>Features</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electric field strength</td>
<td>( E_{h\text{max}}, E_{l\text{max}}, E_{min}, E_a, E_{std} )</td>
<td>6</td>
</tr>
<tr>
<td>Electric field gradient</td>
<td>( E_{gh\text{max}}, E_{glm}, E_{ga} )</td>
<td>3</td>
</tr>
<tr>
<td>Square of electric field strength</td>
<td>( W, W_a )</td>
<td>2</td>
</tr>
<tr>
<td>Electric field integral</td>
<td>( V_{90}, V_{75}, V_{h}, V_1 )</td>
<td>4</td>
</tr>
<tr>
<td>Path length</td>
<td>( L_{E90}, L_{E75}, L_{rg90}, L_{rg75}, L_{W90}, L_{W75}, L_{\text{min}} )</td>
<td>7</td>
</tr>
<tr>
<td>Electric field inhomogeneity</td>
<td>( E_d, E_{std} )</td>
<td>16</td>
</tr>
</tbody>
</table>

| \( L_{W90}, L_{W75}, L_{\text{min}} \) | 16 |

Table 1. Electric field features for sphere-sphere air gap.

### 2.2.2. Sphere-plane air gap

For sphere-plane air gap with a monotonously decreased curve of the electric field distribution along the shortest path, the electric field features are defined as follows:
1. The maximum, the minimum, and the average value of the electric field strength \( E_{\text{max}} \), \( E_{\text{min}} \), and \( E_a \) and the variance and standard deviation of the electric field distribution along the shortest path \( E_{\text{std2}} \) and \( E_{\text{std}} \). Their calculation formulas are similar to or the same with Eqs. (1)–(6).

2. The maximum and the average value of the electric field gradient \( E_{\text{gm}} \) and \( E_{\text{ga}} \), whose calculation formulas are similar to or the same with Eqs. (7)–(9).

3. The square of electric field strength \( W \) and \( W_a \), which can be calculated, respectively, by Eqs. (10) and (11).

4. The electric field integral of the path on which the field strength exceeds \( x\%E_{\text{max}} \) \( (V_x) \) and its ratio to the applied voltage \( U \) \( (V_{rx}) \). \( V_x \) and \( V_{rx} \) can be calculated, respectively, by Eqs. (12) and (20).

5. The length of the path on which the electric field strength exceeds \( x\%E_{\text{max}} \) \( (L_x) \) and the ratio of \( L_x \) to the gap length \( d \) \( (L_{rx}) \). \( L_x \) and \( L_{rx} \) can be calculated, respectively, by Eqs. (15) and (23).

For long sphere-plane air gaps, \( x\% \) includes 90, 75, 50, and 25%. Hence, there are altogether 25 features extracted from the shortest path used to characterize the electric field distribution of the sphere-plane air gap.

3. Breakdown voltage prediction method

3.1. Basic ideas

The proposed method for air gap breakdown voltage prediction is based on electric field features and SVM. A support vector classifier (SVC) is used to establish the prediction model. The withstanding and breakdown of an air gap under a given voltage are, respectively, denoted as \( -1 \) and \( 1 \). The input data of the prediction model are the abovementioned electric field features extracted from the electric field calculation results under different applied voltages, and the outputs are \( -1 \) and \( 1 \), respectively, means whether the gap will withstand or breakdown under the applied voltage, so as to transform the breakdown voltage prediction from a regression problem to a binary classification problem.

To be specific, if the critical breakdown voltage is \( U_b \), the interval \( [(1-a)U_b, U_b] \) is defined as withstand voltage interval, and \( [U_b, (1 + a)U_b] \) is defined as breakdown voltage interval. The value of \( a \) is determined by experience, which is set as 0.1 in this chapter. Set the step size as 0.01\( U_b \), and then the applied voltage values are \( 0.9U_b, 0.91U_b, ..., U_b, 1.01U_b, ..., 1.1U_b \). The SVM model should be trained by some training samples to make it have the learning ability for accurate classification. By this binary classification method, one training sample can be extended to 21 samples to train the SVM model. For test samples, if the model outputs \( -1 \) to 1 under the applied voltage from \( U_{b0-dU} \) to \( U_{b0} \), then \( U_{b0} \) is the predicted breakdown voltage.
3.2. Brief introduction of SVM

SVM is a machine learning algorithm developed on the basis of VC dimension in statistical learning theory and the principle of structural risk minimization [47]. The fundamentals of SVM were detailed and introduced in [47, 48]. Here, a brief introduction is provided.

Set a known training sample as \( T = \{ (x_i, y_i) \} \), in which \( x_i \in \mathbb{R}^k \), \( y_i \in \{-1, 1\} \), \( i = 1, 2, \ldots, n \). The implementation of SVM is to solve an optimization problem based on the maximum margin principle, which finds an optimal separating hyperplane to divide the sample data into two diverse classes. By application of the kernel trick, SVM transforms the original sample data into a high-dimensional Hilbert space \( H \) using nonlinear mapping. The training sample data are transformed to \( T_\Phi = \{ (\Phi(x_i), y_i) \} \), where \( \Phi(x_i) \in H \), \( y_i \in \{-1, 1\} \), \( i = 1, 2, \ldots, n \). In this feature space, the sample data can be linearly separated, and the decision function can be expressed as.

\[
 f(x) = \text{sgn} (w^T \cdot \phi(x_i) + b)
\]  

(24)

where \( w \) and \( b \) are, respectively, the weight vector and bias term of the separating hyperplane \((w \in \mathbb{R}^k)\) and \( b \) is a real number. The symbol “\( \text{sgn} \)” is the signum function. When \( w^T x_i + b > 0 \), the output is 1, while \( w^T x_i + b < 0 \), the output is \( -1 \). Meanwhile, \( w^T x_i + b = 0 \) is the classification hyperplane.

The optimization problem can be expressed as

\[
 \min_{w, b, \xi} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i \\
\text{s.t. } y_i (w \cdot \phi(x_i) + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i = 1, 2, \ldots, n
\]  

(25)

where \( \xi_i \) is the slack variable and \( C \) is the penalty factor, which determines the balance between the maximization of the margin and the minimization of the classification error [48].

In order to make the solution of the primal problem more simple and practicable, the Lagrange function is introduced:

\[
 L(w, b, \xi, \alpha, \beta) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \alpha_i \{ y_i ((w \cdot x_i) + b) - 1 + \xi_i \} - \sum_{i=1}^{n} \beta_i \xi_i
\]  

(26)

where \( \alpha = (\alpha_1, \ldots, \alpha_n)^T \) and \( \beta = (\beta_1, \ldots, \beta_n)^T \) are the Lagrange multiplier vectors. By solving the partial derivatives of Eq. (26) for \( w, b, \) and \( \xi \), according to the extremum conditions, the primal optimization problem (25) can be transformed to the following dual problems:

\[
 \max_{\alpha, \beta} -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) + \sum_{j=1}^{n} \alpha_j \\
\text{s.t. } \sum_{i=1}^{n} \alpha_i y_i = 0 \\
C - \alpha_i - \beta_j = 0, \quad i = 1, 2, \ldots, l \\
\alpha_i \geq 0, \quad \beta_j \geq 0, \quad i = 1, 2, \ldots, l
\]  

(27)
By solving Eq. (27), the decision function can be obtained. It can be seen that the function of transformation $\Phi$ is realized by inner product $(\Phi(x_i) \cdot \Phi(x_j))$. The kernel function can be expressed as

$$K(x_i, x_j) = (\phi(x_i) \cdot \phi(x_j))$$  \hspace{1cm} (28)

Hence, if the function $K$ is selected, it is not necessary to choose the transformation $\Phi$. $K(x_i, x_j)$ is used in training and classification instead of $\Phi(x)$. The generalization performance of SVM is determined by properly selecting kernel functions. In this chapter, the radial basis function (RBF) kernel is selected as the kernel function of SVM for its good generalization performance and high computational efficiency:

$$K(x_i, x_j) = \exp(-\gamma ||x_i - x_j||^2), \quad \gamma > 0$$ \hspace{1cm} (29)

where $\gamma$ is the kernel parameter.

The penalty factor $C$ and the kernel parameter $\gamma$ determine the classification performance of SVM. They can be optimized by grid search (GS) method or genetic algorithm (GA) based on $K$-fold cross validation or leave-one-out (LOO) cross validation, so as to obtain the optimal predictive model [36–39].

### 3.3. Implementation procedures of the prediction method

The flow chart of the prediction method is shown in Figure 4. The implementation procedures are depicted as follows.

Firstly, the training samples with known gap structures and experimental data of breakdown voltage are used to train the SVM model. The electric field features of each training sample are extracted from the FEM calculation results of the electric field distribution. These features are normalized to $[0, 1]$ by

$$\bar{x}_i = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}$$ \hspace{1cm} (30)

where $x_i$ is a feature, $\bar{x}_i$ is its normalized value, and $x_{\text{min}}$ and $x_{\text{max}}$ are, respectively, the minimum and maximum values of $x_i$. After normalization, the electric field features are taken as the input data to train the SVM model, while the outputs are $-1$ and $1$, respectively, correspond to the applied voltage in the withstand interval and the breakdown interval. Based on cross validation, the optimal penalty factor $C$ and kernel parameter $\gamma$ are searched by GS method or GA to obtain an optimal prediction model. Trained by the known experimental data, SVM establishes the multidimensional nonlinear relationships between the electric field features and the air gap breakdown voltage.

Then, the optimal SVM model is used to predict the breakdown voltages of test samples. For an air gap, an estimated breakdown voltage is set in the range $[U_{\text{min}}, U_{\text{max}}]$; the golden section search method is applied for the breakdown voltage prediction [49]. For each applied voltage,
the electric field features are extracted and input to the prediction model to judge whether the output is $-1$ or $1$. For example, the first applied voltage is $U_{t1} = \frac{U_{max}}{C_0} \times \frac{0.618}{C_2} (\frac{U_{max}}{C_0} - \frac{U_{min}}{C_0})$, and then the electric field features of this air gap are calculated under $U_{t1}$ and input to the SVM model. According to the output of the model, namely, $1$ or $-1$, the search interval will be narrowed to $[U_{min}, U_{t1}]$ to generate another applied voltage value, or otherwise, $U_{t2} = U_{min} + \frac{0.618}{C_2} (\frac{U_{max}}{C_0} - \frac{U_{min}}{C_0})$ is applied to calculate the electric field features which will be input to the SVM model to judge whether the output is $-1$ or $1$. So repeatedly, the search interval is narrowed constantly, and the iterative predictions are conducted until the convergence condition $\frac{U_{max}}{C_0} - \frac{U_{min}}{C_0} < \varepsilon$ is satisfied, where $\varepsilon$ is the convergence precision. The breakdown voltage prediction result is the average value of the last two applied voltages.

4. Breakdown voltage prediction of air gaps

The proposed method is applied to predict the power frequency breakdown voltages of sphere-sphere air gaps and the switching impulse discharge voltages of sphere-plane air gaps. The predicted results are compared with the experimental data given in references.
4.1. Power frequency breakdown voltage prediction of sphere-sphere air gaps

4.1.1. Training and test samples

The sample data of sphere-sphere air gaps are selected from IEC 60052 [43] and IEEE Std 4 [44]. The values of sphere diameter $D$ are 5, 6.25, 10, 12.5, 15, 25, 50, 75, 100, 150, and 200 cm. The shortest gap distance is 1 cm, for those with $D$ ranging from 5 to 25 cm, and the longest gap distance is 100 cm for that of $D = 200$ cm. There are altogether 271 sample data with different sphere diameters and different gap distances.

The proper selection of training samples is of vital importance for the generalization performance of the SVM model. Here, the training samples are selected according to the electric field nonuniform coefficient $f$, that is, the ratio of the maximum field strength $E_{\text{max}}$ to the average field strength $E_a = U/d$, where $U$ is the applied voltage. The maximum field strength of each sphere-sphere gap can be calculated by FEM, and therefore the electric field nonuniform coefficient $f$ can be obtained.

According to the calculation results, all of these samples are slightly nonuniform electric field, and the values of $f$ range from 1.00 to 1.55. Taking 0.05 as the step size, the values of $f$ can be divided into 11 intervals. The samples belong to each $f$ interval are collected together, and the sample sizes corresponding to each $f$ interval are summarized in Table 2. According to the electric field nonuniform coefficient $f$, the 271 samples are divided into 11 groups. The training sample set is constituted by random selection of one sample from each group. Hence, there are 11 training samples, and the other 260 samples are taken as the test samples to verify the validity of the prediction method. Since the training sample selection is conducted by computer program which has a certain randomness, three different selection results are successively taken as the training sample set to train the SVM model, and the three times of prediction results are compared to validate the accuracy. The three groups of training samples are shown in Table 3.

<table>
<thead>
<tr>
<th>$f$ interval</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00–1.05</td>
<td>58</td>
</tr>
<tr>
<td>1.05–1.10</td>
<td>57</td>
</tr>
<tr>
<td>1.10–1.15</td>
<td>41</td>
</tr>
<tr>
<td>1.15–1.20</td>
<td>23</td>
</tr>
<tr>
<td>1.20–1.25</td>
<td>19</td>
</tr>
<tr>
<td>1.25–1.30</td>
<td>18</td>
</tr>
<tr>
<td>1.30–1.35</td>
<td>15</td>
</tr>
<tr>
<td>1.35–1.40</td>
<td>13</td>
</tr>
<tr>
<td>1.40–1.45</td>
<td>9</td>
</tr>
<tr>
<td>1.45–1.50</td>
<td>7</td>
</tr>
<tr>
<td>1.50–1.55</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 2. Sample size of each $f$ interval.
4.1.2. Prediction results and analysis

Based on fivefold cross validation, the grid search method is applied to determine the optimal parameters of the SVM model. The search ranges of the penalty factor $C$ and the kernel parameter $\gamma$ are, respectively, set as $[2^3, 2^9]$ and $[2^{-8}, 2^{-2}]$, and the step sizes are both $2^{0.1}$. Taking the training sample set 1, for example, the parameter optimization results of $C$ and $\gamma$ by GS method are shown in Figure 5. It can be seen that the best $C = 90.5097$, $\gamma = 0.0167$, under which the SVM model, has the highest classification accuracy for the training samples, that is, 98.2684%. The GS-optimized SVM models under the three groups of training sample set shown in Table 3 are used for breakdown voltage prediction of the 260 test samples.

For each test sample, the breakdown voltage is predicted by the golden section search method, and the prediction results are compared with the experimental data given in [43, 44]. In order to evaluate the prediction accuracy of the proposed method and the SVM model, three error indices, including the root-mean-square error (RMSE), the mean absolute percentage error (MAPE), and the mean square percentage error (MSPE), are used to examine the errors of the prediction results, which can be calculated by

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} [U_i(i) - U_p(i)]^2}$$  \hspace{1cm} (31)

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \frac{|U_i(i) - U_p(i)|}{U_i(i)}$$  \hspace{1cm} (32)

$$\text{MSPE} = \frac{1}{n} \sqrt{\sum_{i=1}^{n} \left[\frac{U_i(i) - U_p(i)}{U_i(i)}\right]^2}$$  \hspace{1cm} (33)

<table>
<thead>
<tr>
<th>Training sample set 1</th>
<th>Training sample set 2</th>
<th>Training sample set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$ (cm)</td>
<td>$d$ (cm)</td>
<td>$D$ (cm)</td>
</tr>
<tr>
<td>5</td>
<td>1.4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>2.2</td>
<td>6.25</td>
</tr>
<tr>
<td>10</td>
<td>4.5</td>
<td>12.5</td>
</tr>
<tr>
<td>15</td>
<td>6.0</td>
<td>12.5</td>
</tr>
<tr>
<td>25</td>
<td>5.5</td>
<td>12.5</td>
</tr>
<tr>
<td>25</td>
<td>8.0</td>
<td>25</td>
</tr>
<tr>
<td>50</td>
<td>8.0</td>
<td>50</td>
</tr>
<tr>
<td>50</td>
<td>17</td>
<td>75</td>
</tr>
<tr>
<td>75</td>
<td>5.5</td>
<td>150</td>
</tr>
<tr>
<td>75</td>
<td>10</td>
<td>200</td>
</tr>
<tr>
<td>200</td>
<td>100</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 3. Three groups of training sample set.
where \( n \) is the number of the test samples and \( U_t(i) \) and \( U_p(i) \) are, respectively, the experimental and predicted breakdown voltages of the \( i \)th test sample.

The optimal parameters and the error indices of the three times of prediction results are summarized in Table 4. It can be seen that the prediction results are with high accuracy, while the MAPEs of the three times of prediction are, respectively, 1.88, 2, and 1.4%. Taking the prediction results by training sample set 2, for example, the comparisons between the predicted and experimental breakdown voltages of sphere-sphere air gaps with different diameters are shown in Figure 6, where \( U \) is the breakdown voltage, \( D \) is the sphere diameter, \( d \) is the gap distance, and T-value and P-value, respectively, mean the test value and the prediction value of the breakdown voltage. For better comparisons, the prediction results of training samples are also plotted in Figure 6.

It can be seen from Figure 6 that the predicted results coincide well with the experimental data, the trends of the breakdown voltages with the gap distance are the same, and the errors are within an acceptable range. The results shown in Table 4 and Figure 6 validate the feasibility and accuracy of the proposed method for sphere-sphere air gap breakdown voltage prediction.

4.2. Switching impulse breakdown voltage prediction of sphere-plane air gaps

4.2.1. Training and test samples

The sample data of sphere-plane air gaps are selected from [45, 46]. The positive switching impulse discharge tests of sphere-plane air gaps with the sphere diameter of 25, 45, 75, and 95 cm were conducted in [45]. The applied voltage waveform is the standard 250/2500 \( \mu \)s switching impulse voltage. The experimental data were corrected to standard atmospheric condition. In order to make the SVM model generalize to sphere-plane gaps with different sphere diameters and gap lengths, seven test data shown in Table 5 are selected as the training samples, where \( D \) is the sphere diameter ranging from 25 to 95 cm, \( d \) is the gap length ranging from 2 to 5 m, and \( U_{50} \) is the 50% discharge voltage.
The SVM model trained by the seven sample data is applied to predict the 50% discharge voltages of sphere-plane gaps with larger diameters, namely, 110, 150, and 200 cm. There are altogether 16 test samples. The predicted results will be compared with the experimental data cited from [46], as shown in Table 5. These experimental data also had been corrected to standard atmospheric condition.

### 4.2.2. Prediction results and analysis

Based on LOO cross validation, the penalty factor $C$ and the kernel parameter $\gamma$ are optimized by the GA method. The population quantity is set as 20, the maximum generation is 200, and the crossover probability is 0.9. The search scopes of $C$ and $\gamma$ are, respectively, set as $[10, 500]$ and $[0.005, 0.25]$. The fitness function is the classification accuracy of SVM for training samples. The parameter optimization results of $C$ and $\gamma$ by GA method are shown in Figure 7. It can be seen that the best $C = 85.3407$, $\gamma = 0.0926$, under which the SVM model has the highest classification accuracy for the training samples, that is, 97.2789%.

### Table 4. Optimal parameters and error indices of the sphere-sphere air gap breakdown voltage prediction results.

<table>
<thead>
<tr>
<th>Results</th>
<th>Training sample set 1</th>
<th>Training sample set 2</th>
<th>Training sample set 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C$</td>
<td>90.5097</td>
<td>78.7932</td>
<td>13.9288</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.0167</td>
<td>0.0313</td>
<td>0.0625</td>
</tr>
<tr>
<td>RMSE</td>
<td>16.668</td>
<td>8.048</td>
<td>9.143</td>
</tr>
<tr>
<td>MAPE</td>
<td>0.0188</td>
<td>0.0200</td>
<td>0.0140</td>
</tr>
<tr>
<td>MSPE</td>
<td>0.0015</td>
<td>0.0140</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

Figure 6. Comparisons between the predicted and experimental breakdown voltages of sphere-sphere air gaps (prediction results by training sample set 2).

The SVM model trained by the seven sample data is applied to predict the 50% discharge voltages of sphere-plane gaps with larger diameters, namely, 110, 150, and 200 cm. There are altogether 16 test samples. The predicted results will be compared with the experimental data cited from [46], as shown in Table 5. These experimental data also had been corrected to standard atmospheric condition.

![Figure 6](image-url)
The GA-optimized SVM model is used to predict the $U_{50}$ of large sphere-plane air gaps by the golden section search method. The initial applied voltage interval $[U_{\text{min}}, U_{\text{max}}]$ is set as 0–4000 kV, and the convergence precision $\epsilon$ is set as 1 kV. The prediction results are summarized in Table 6, where $U_{50}$ is the experimental data extracted from [46], $U_p$ is the predicted discharge voltage, and $\delta$ is the relative error.

It can be seen from Table 6 that the largest error of the prediction results is 9.7%, for the gap with $D = 110$ cm and $d = 3.9$ m. This is probably due to different experimental arrangements between [45, 46]. The $U_{50}$ of sphere-plane gaps with $D = 110$ cm and $d = 4$ m in [46] is even lower than that with $D = 95$ cm and $d = 4$ m in [45]. The SVM model is trained by the experimental data cited from [45], and therefore the prediction results of sphere-plane gaps with $D = 110$ cm may be larger than the experimental data obtained in [46]. Overall, the prediction errors are acceptable in the view of engineering applications, while the MAPE of the 16 test samples is only 3.2%.

The $U_{50}$ prediction results and the experimental data are summarized in the same graph for a better comparison, as shown in Figure 8. It can be seen that the predicted values of the discharge voltage agree well with the experimental data, with similar trends and acceptable errors. The results verify the validity and accuracy of the proposed model for discharge voltage prediction of sphere-plane air gaps, with large sphere diameter and long gap length. Within the range of certain precision, the prediction method can be used to replace the experiments, so as to reduce the testing expenses.

<table>
<thead>
<tr>
<th>$D$ (cm)</th>
<th>$d$ (m)</th>
<th>$U_{50}$ (kV)</th>
<th>$D$ (cm)</th>
<th>$d$ (m)</th>
<th>$U_{50}$ (kV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>2</td>
<td>864</td>
<td>25</td>
<td>3</td>
<td>1020</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1045</td>
<td>75</td>
<td></td>
<td>1222</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1186</td>
<td>95</td>
<td></td>
<td>1504</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1303</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Training and test samples of the sphere-plane air gaps.

<table>
<thead>
<tr>
<th>$D$ (cm)</th>
<th>$d$ (m)</th>
<th>$U_{50}$ (kV)</th>
<th>$D$ (cm)</th>
<th>$d$ (m)</th>
<th>$U_{50}$ (kV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>110</td>
<td>3.9</td>
<td>1461</td>
<td>150</td>
<td>8.5</td>
<td>2408</td>
</tr>
<tr>
<td></td>
<td>4.5</td>
<td>1527</td>
<td></td>
<td>10</td>
<td>2497</td>
</tr>
<tr>
<td></td>
<td>5.7</td>
<td>1675</td>
<td>200</td>
<td>3</td>
<td>2290</td>
</tr>
<tr>
<td></td>
<td>6.5</td>
<td>1796</td>
<td></td>
<td>4</td>
<td>2540</td>
</tr>
<tr>
<td></td>
<td>7.5</td>
<td>1973</td>
<td></td>
<td>5</td>
<td>2711</td>
</tr>
<tr>
<td>150</td>
<td>4</td>
<td>2101</td>
<td>6</td>
<td></td>
<td>2745</td>
</tr>
<tr>
<td></td>
<td>5.5</td>
<td>2242</td>
<td>7</td>
<td></td>
<td>2836</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>2316</td>
<td>8</td>
<td></td>
<td>2927</td>
</tr>
</tbody>
</table>
Figure 7. Parameter optimization results of C and γ by GA method.

<table>
<thead>
<tr>
<th>$D$ (cm)</th>
<th>$d$ (m)</th>
<th>$U_{50}$ (kV)</th>
<th>$U_p$ (kV)</th>
<th>$δ$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>110</td>
<td>3.9</td>
<td>1461</td>
<td>1602</td>
<td>9.7</td>
</tr>
<tr>
<td>4.5</td>
<td></td>
<td>1527</td>
<td>1647</td>
<td>7.9</td>
</tr>
<tr>
<td>5.7</td>
<td></td>
<td>1675</td>
<td>1745</td>
<td>4.2</td>
</tr>
<tr>
<td>6.5</td>
<td></td>
<td>1796</td>
<td>1807</td>
<td>0.6</td>
</tr>
<tr>
<td>7.5</td>
<td></td>
<td>1973</td>
<td>1910</td>
<td>–3.2</td>
</tr>
<tr>
<td>150</td>
<td>4</td>
<td>2101</td>
<td>2036</td>
<td>–3.1</td>
</tr>
<tr>
<td>5.5</td>
<td></td>
<td>2242</td>
<td>2132</td>
<td>–4.9</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>2316</td>
<td>2266</td>
<td>–2.2</td>
</tr>
<tr>
<td>8.5</td>
<td></td>
<td>2408</td>
<td>2413</td>
<td>0.2</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>2497</td>
<td>2560</td>
<td>2.5</td>
</tr>
<tr>
<td>200</td>
<td>3</td>
<td>2290</td>
<td>2348</td>
<td>2.5</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>2540</td>
<td>2513</td>
<td>–1.1</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>2711</td>
<td>2595</td>
<td>–4.3</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>2745</td>
<td>2700</td>
<td>–1.6</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>2836</td>
<td>2778</td>
<td>–2.0</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>2927</td>
<td>2893</td>
<td>–1.2</td>
</tr>
</tbody>
</table>

Table 6. 50% discharge voltage prediction results of the sphere-plane air gaps.

5. Conclusions

Two sets of electric field features defined on the shortest interelectrode path are, respectively, used to characterize the gap structure of the sphere-sphere air gap and the rod (sphere)-plane air gap. These features are taken as the input parameters of the SVM model, which is used to establish the breakdown voltage prediction model. The proposed method based on electric
field features and SVM is applied to predict the breakdown voltages of sphere-sphere and sphere-plane air gaps. Some conclusions can be drawn as follows:

1. The proposed electric field features extracted from the shortest interelectrode path are effective to characterize the spatial structure of sphere-sphere and sphere-plane air gaps, and the multidimensional nonlinear relationships between these features and the air gap breakdown voltage can be established by SVM, so as to achieve breakdown voltage prediction of air gaps without considering the complex and random discharge process.

2. Trained by only 11 sample data selected randomly according to the electric field nonuniform coefficient $f$, the SVM model is able to accurately predict the power frequency breakdown voltages of IEC standard sphere-sphere air gaps. The mean absolute percentage errors of the 260 test samples, with three times of prediction by different training sample sets, are within 2%. The results validate the validity and accuracy of the proposed method for breakdown voltage prediction of sphere-sphere air gaps.

3. The proposed method is able to predict the switching impulse discharge voltages of sphere-plane air gaps, with large sphere diameters and long gap distances. The prediction results agree well with the experimental data, with similar trends and acceptable errors. The mean absolute percentage error of the 16 test samples is 3.2%, which is acceptable for engineering applications. The results verify the feasibility of the proposed model for discharge voltage prediction of large sphere-plane air gaps, which may be useful to replace the time-consuming and costly discharge tests.

The authors are still engaged in improving this model. The following work will be carried out in the future. Firstly, the electric field features will be simplified by some feature selection approaches to make it easier for applications. Secondly, the applications of this method will be extended to breakdown voltage prediction of more complex air gaps such as the practical engineering gaps. It should be noted that new problems will inevitably appear in different applications, and therefore this topic is worthy to be studied in-depth. We hope that it is possible to achieve breakdown voltage prediction of arbitrary engineering gap configurations.
in the future, so as to guide the insulation design of high-voltage electrical equipment by mathematical calculations rather than costly experiments.

Acknowledgements

This work is supported by China Postdoctoral Science Foundation (2016 M602354). The authors would like to thank Dr. Shengwen Shu for his primary work on this topic during his study for PhD degree in Wuhan University. The experimental data of sphere-plane gaps are cited from [45, 46] and taken as the sample data in this paper. We would like to express thanks to those authors.

Author details

Zhibin Qiu* and Jiangjun Ruan

*Address all correspondence to: whuerliuming@whu.edu.cn

School of Electrical Engineering, Wuhan University, Wuhan, China

References


[38] Qiu ZB, Ruan JJ, Huang DC, PU ZH, Shu SW. A prediction method for breakdown voltage of typical air gaps based on electric field features and support vector machine. IEEE Transactions on Dielectrics and Electrical Insulation. 2015;22(4):2125-2135. DOI: 10.1109/TDEI.2015.004887

[39] Qiu ZB, Ruan JJ, Huang DC, Wei MT, Tang LZ, Huang CP, WJ X, Shu SW. Hybrid prediction of the power frequency breakdown voltage of short air gaps based on orthogonal design and support vector machine. IEEE Transactions on Dielectrics and Electrical Insulation. 2016;23(2):795-805. DOI: 10.1109/TDEI.2015.005398

[40] Qiu ZB, Ruan JJ, Huang CP, Xu WJ, Tang LZ, Huang DC, Liao YF. A method for breakdown voltage prediction of short air gaps with atypical electrodes. IEEE Transactions on Dielectrics and Electrical Insulation. 2016;23(5):2685-2694. DOI: 10.1109/TDEI.2016.7736827


[43] IEC 60052. Voltage measurement by means of standard air gaps; 2012


