A novel partial discharge localization method for GIL based on the 3D optical signal irradiance fingerprint and bagging-KELM

Yiming Zang¹ | Yong Qian¹ | Xiaoli Zhou² | Antian Xu² | Gehao Sheng¹ | Xiuchen Jiang¹

¹ Department of Electrical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Minhang, Shanghai 200240, China
² Department of Light Sources and Illuminating Engineering, Fudan University, 220 Handan Road, Yangpu, Shanghai 200433, China

Abstract
Partial discharge (PD) is one of the main reasons that endanger the safe operation of power equipment. The effective detection and localization of PD in a gas-insulated transmission line (GIL) is essential for timely finding the insulation defects to improve maintenance efficiency. Therefore, this paper proposes a method based on the three-dimensional optical signal irradiance fingerprint and bagging-kernel extreme learning machine (Bagging-KELM), which introduces optical simulation into the PD localization. Using a simulation model with the same size and optical sensor arrangement as the experimental GIL, optical simulation signals emitted from different PD sources are collected. Principal component analysis is used to extract the signal features to construct optical PD fingerprints that correspond to PD source locations. This paper builds all the simulated PD fingerprints into a PD fingerprint database. Afterwards, the Bagging-KELM is used to match PD optical fingerprints detected on site with the fingerprint database to identify the location of the PD sources. The experimental results show that the average localization error of this method is 0.93 cm, with 93.75% of the errors being less than 1.5 cm. The proposed method has better performance than the traditional KELM and BPNN, which can achieve precise PD localization in GIL.

1 | INTRODUCTION

With the growth of energy generation and industrialization, the need for the high voltage transmission of long distance and large capacity has increased sharply [1]. Gas-insulated transmission lines (GILs) have been used around the world in recent years because of their superior performance, such as low life cycle costs, high safety and large transmission capacity [2].

During the operation of GILs, a partial discharge (PD) will occur due to defects such as the abnormal roughness of the surface and free particles in the GIL. The PD will cause the deterioration of the insulating medium and even a breakdown, which seriously affects the operational safety of the GIL. Thus, the localization of the PD source is essential, which can effectively determine the location of insulation defects to help evaluate the insulation status and facilitate a timely maintenance strategy [3].

PD occurrences are accompanied by electrical, electromagnetic, optical, acoustic and optical phenomena [4]. Thus, apart from the ultra-high frequency (UHF) detection, current flow detection, radio frequency detection and acoustic detection of the PD, optical detection is a novel and effective method for PD monitoring and localization because of its advantages in terms of anti-electromagnetic and anti-acoustic interferences, high sensitivity and intrinsic characterization [5, 6]. However, only a few PD localization studies is aimed at the detection of optical signal irradiance in recent years. In paper [7], although it is possible to locate the PD source through the optical sensing array, it is difficult to install the array in a large area in the electric equipment. Therefore, this optical sensing array is only suitable for small-range PD localization, not long-distance GILs. In addition, some machine learning algorithms have been used for optical PD localization, such as the Gaussian mixture model, support vector machines (SVM) [8], rough set theory [9] and sparse representation classifier [10]. These machine learning-based localization methods need to construct a fingerprint database containing the spatial location information of the...
PD source in advance for identifying. However, the current construction method is to select some locations in the equipment to perform PD experiments practically to obtain PD fingerprints of the corresponding locations. This method is not only lot of workload, but also difficult to carry out in actual equipment outside the laboratory. Moreover, current fingerprint localization methods can only locate specific locations where PD experiments have been conducted, which significantly reduces the localization accuracy [11].

Therefore, to improve the detection range, accuracy and practicability of optical PD localization, this paper proposes a PD localization method based on a three-dimensional optical signal irradiance fingerprint (3D-OSIF). This method introduces optical simulation data into the PD localization for the first time and solves the problem that the fingerprint database is difficult to be obtained from actual GIL equipment conveniently. By building the same simulation model as the real GIL tank in TracePro software and using fitting algorithm, an optical simulation fingerprint database containing coordinate information of PD sources can be constructed. This method can realize the PD detection of any location in the GIL rather than locating a few specific PD sources where the PD test is conducted, which can greatly improve the localization accuracy.

Furthermore, the PD fingerprints matching is also an important part in the localization process. This paper combines the bagging ensemble learning algorithm with the kernel extreme learning machine (Bagging-KELM) to match the detected PD fingerprint with the optical simulation fingerprint in the fingerprint database. KELM is a machine learning algorithm that has stronger stability and higher computational efficiency than conventional algorithms such as neural network algorithms and the SVM algorithm. In addition, KELM solves the problem of hidden layer parameters and the random distribution of input layer parameters in ELM [12]. Moreover, the bagging framework is a kind of ensemble learning method that can improve the stability, generalization ability and accuracy of machine learning algorithms [13]. Thus, in the Bagging-KELM algorithm, KELM classifiers are embedded as base classifiers in the bagging framework to enhance the recognition accuracy.

In the matching processing, the dimensions of optical PD fingerprints increase with an increase in the number of optical sensors, which will make the dimension of the fingerprint too high, leading to a dimensional disaster and overfitting. Thus, principal component analysis (PCA) is applied to reduce the dimensions of the optical PD fingerprint and extract valid features. PCA is a mathematical tool that uses an orthogonal transformation to extract a set of linearly uncorrelated features from potentially related features, which have a good ability to reduce the data dimension [14].

In this paper, GIL PD experiments that use fluorescent fibre optic sensors installed on the GIL wall are carried out to obtain real optical PD fingerprints [15]. After the PCA process of the optical fingerprint, the Bagging-KELM algorithm is used for matching the actual optical PD fingerprint with the simulation fingerprint database based on 3D-OSIF to obtain the spatial coordinates of the PD source in GIL. Finally, based on a considerable number of experiments, the feasibility of this optical PD localization method is verified by comparison with other localization models.

2 | THE PRINCIPLE OF OPTICAL PD SIMULATIONS IN GIL

2.1 | Simulation settings of the PD sources and optical signals

In the simulation, the PD source is set as a spherical point light source, placed directly under the needle plate defect. It is assumed that the optical PD signal emitted from all directions in space is evenly distributed and perpendicular to the PD source surface. The total number of rays emitted by the PD source point is 250,000, and the total optical radiation flux is 100 watts. The optical refractive index of SF₆ in the GIL is set to 1.000783. The absorption spectrum of SF₆ is mainly concentrated in the mid-infrared band, which has minimal impact and can be ignored with respect to the propagation of the PD optical signal [16]. In addition, because the wavelength of the PD light radiation in SF₆ is mainly concentrated at approximately 500 nm, the radiation of the PD light source is set to green light with a wavelength of 546.1 nm [17].

To represent the intensity of the light signal received by the optical sensor, the concept of light irradiance $E_o$ (unit: W/m²) is introduced in this article:

$$E_o = \frac{dP_o}{dS}$$  \hspace{1cm} (1)

where $P_o$ (unit: watt) is the light radiation flux received by the sensor and $S$ (unit: m²) represents the receiving area of the sensor.

Because the PD source set in the simulation is not exactly the same as the optical signal generated by the actual PD, the above $E_o$ is a relative value instead of representing the actual optical signal irradiance. However, because the localization method in this paper is based on the distribution of optical signals between different sensors rather than using the actual value directly, the relative irradiance $E_o$ can fully represent the difference in the distribution of optical PD signals between each sensor.

2.2 | Simulation settings of the GIL tank material

The propagation and scattering of optical PD signals in the GIL tank have a great influence on the light irradiance detected by the sensor, so it is particularly important for the setting of the GIL tank simulation material.

The diffuse reflection model of the GIL surface material used in this simulation is a bidirectional reflection distribution function (BRDF) [18]. This model represents the reflected irradiance distribution produced by the incident light from different angles at each unit three-dimensional angle in the three-dimensional
where $\omega_i$, $\theta_i$, and $\phi_i$ represent the solid angle, elevation angle, and azimuth angle of the incident light, respectively; and $\omega_r$, $\theta_r$, and $\phi_r$ represent the solid angle, elevation angle and azimuth angle of the reflected light, respectively.

The definition of the BRDF model is

$$f(\theta_i, \phi_i, \theta_r, \phi_r) = \frac{dL(\theta_r, \phi_r)}{dE(\theta_i, \phi_i)}$$

where $dE$ is the incident light irradiance in a direction per unit area and $dL$ is the reflecting light irradiance in a direction per unit area.

In addition, the propagation of optical signals in the GIL must satisfy the following conditions:

$$\alpha + R + D = 1$$

where $\alpha$ represents the absorption coefficient, $R$ represents the specular reflection coefficient, and $D$ represents the diffuse reflectance coefficient. In this simulation, the internal material of the GIL tank is set to polished and oxidized medium smooth aluminium.

### 2.3 Simulation settings of the GIL tank structure

Since the length of one gas compartment in the GIL actually put into operation is approximately 100 m and the equipment condition is not suitable for the PD localization experiment [19], a GIL experimental tank that considers the basic structural characteristics of the GIL and the sensor arrangement is established in the laboratory to simulate the PD in the real GIL, as shown in Figure 2. In order to easily distinguish the sensors in different positions, the sensors in Figure 2 are marked with serial numbers. Meanwhile, a GIL simulation model that is identical to the experimental tank is set up in the TracePro software, as shown in Figure 3.

In Figures 2 and 3, the internal height of the GIL tank is 310 mm, the inner radius is 90 mm, the wall thickness is 10 mm, and the inner conductor diameter of the axis is 25 mm. A needle-plate defect model is connected to the axial conductor post that can rotate 360°, and the radial length of the defect model can be adjusted. The needle-plate spacing is always 6 mm, the length of the tip is 25 mm, the cross-sectional angle of the tip is 30°, and the diameter of the grounding disc below is 20 mm. By changing the height of the needle-plate defect model, the radial distance from the axis, and the angle of rotation during the simulation process, the optical PD sources at most locations can be simulated in the GIL simulation model.
To collect the light irradiance of the PD source and simulate the actual optical sensors, 9 simulation probes are set on the inner wall of the GIL simulation model. Nine simulation probes around the GIL model are divided into three columns at 120°, numbering 1 to 9, as shown in Figure 2. The distances from the three simulation probes on each column to the top surface of the GIL model are 70 mm, 160 mm, and 250 mm, as shown in Figure 3. The normal of each simulation probe is perpendicular to the inner wall of the GIL model and points to the cylindrical axis. The simulation probe has a radius of 10 mm and a thickness of 5 mm. This probe is a fully transmissive body model and does not have any refracting or absorbing effect on the light emitted by the PD source.

According to the settings of the GIL model, this research sets the plane of the x-y-axis on surface 2, as shown in Figure 4(a). The origin of the x-y-z coordinate system is located on the axis of the GIL model, and the y-axis passes through the circle centre of the No. 2 simulation probe. A random PD source (at 0 mm, 44 mm, 95 mm) is set for the simulation whose ray tracing graphs (Figure 4(a)) and the light irradiance map through three cross sections are shown in Figure 4(b), (c) and (d), for example.

3 | THE ESTABLISHMENT OF THE OPTICAL SIMULATION FINGERPRINT DATABASE

Through the above-mentioned optical PD simulation, this research solves the problem that the PD fingerprint database on-site are difficult to obtain and the location range is limited. Moreover, the simulation model can be built and changed according to different GIL structures, which has high applicability. Based on the optical simulation data, the PD simulation fingerprint database containing the spatial location information of the PD source is constructed, which lays the foundation for subsequent PD fingerprint matching.

Suppose an optical PD simulation experiment is performed at N locations in the model. Let \( I_j \) be the location of the PD source, where \( j = 1, 2, ..., N \). For each PD simulation experiment, the light irradiance is measured through \( M \) simulation probes. Let \( \varphi_i \) represent each simulation probe, where \( i = 1, 2, ..., M \). Therefore, when the simulated PD source is located at \( I_j \), the PD light irradiance detected by the simulation probe \( \varphi_i \) is expressed as \( \varphi_{i,j} \).

During construction of the fingerprint database, to avoid the influence of the PD signal intensity, the optical fingerprint of the PD source at the same location is normalized to \([-1, 1]\). The normalization rule is defined below:

\[
\varphi_{i,j} = \frac{2 \times (\varphi_{i,j} - \min(\varphi'_{1,j}, \varphi'_{2,j}, ..., \varphi'_{M,j}))}{(\max(\varphi'_{1,j}, \varphi'_{2,j}, ..., \varphi'_{M,j}) - \min(\varphi'_{1,j}, \varphi'_{2,j}, ..., \varphi'_{M,j})) - 1}
\]

(4)

where \( \varphi_{i,j} \) is the PD light irradiance value detected by the simulation probe after normalization.

Based on the processed data above, three kinds of fingerprint databases are proposed. The three kinds of fingerprint databases are denoted by K1, K2 and K3, as described below:

(K1) The normalized light irradiance detected by each probe is used directly as the fingerprint of a PD source. The structure
of the simulation fingerprint database \( \Psi_{k,1} \) is defined as

\[
\Psi_{k,1} = \begin{bmatrix}
\varphi_{1,1} & \varphi_{1,2} & \cdots & \varphi_{1,N} \\
\varphi_{2,1} & \varphi_{2,2} & \cdots & \varphi_{2,N} \\
\vdots & \vdots & \ddots & \vdots \\
\varphi_{M,1} & \varphi_{M,2} & \cdots & \varphi_{M,N}
\end{bmatrix}
\]  

(5)

where \( M \) is the total number of simulation probes and \( N \) is the total number of simulation PD sources. The column vector \( \Psi_{k,1} = [\varphi_{1,j}, \varphi_{2,j}, \ldots, \varphi_{M,j}]^T \) represents the optical simulation fingerprint at location \( L_p \).

(K2) The fingerprint is established by subtracting the light irradiance of each two simulation probes. The structure of the simulation fingerprint database \( \Psi_{k,2} \) is defined as

\[
H = \frac{M!}{(M-2)! \times 2!}
\]

(6)

\[
\delta_{k,j}^{(a)} = \varphi_{k,j}^{(a)} - \varphi_{k,j}^{(b)}, \text{s.t. } \begin{cases}
b = 1, 2, \ldots, H \\
a < k \\
a, b \in Z; \\
a, b \in [1, M].
\end{cases}
\]

(7)

\[
\Psi_{k,2} = \begin{bmatrix}
\delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,N} \\
\delta_{2,1} & \delta_{2,2} & \cdots & \delta_{2,N} \\
\vdots & \vdots & \ddots & \vdots \\
\delta_{H,1} & \delta_{H,2} & \cdots & \delta_{H,N}
\end{bmatrix}
\]

(8)

where \( H \) is the fingerprint dimension, \( \varphi_{k,j}^{(a)} \) and \( \varphi_{k,j}^{(b)} \) represent the light irradiance of the PD at \( N \) location detected by probes, \( \delta^{(a)} \) is the difference of the light irradiance between each two simulation probes, and \( \delta \) is the normalized \( \delta \) using the method of Equation (4).

(K3) K3 represents the first \( P \) principal components extracted from the K2 fingerprint database by the PCA algorithm.

PCA is an effective dimensionality reduction algorithm to reduce the fingerprint dimension from \( H \) to \( P \), which transforms the fingerprint of K2 to a new set of coordinates and orthogonal eigenvalues [20]. With the new coordinates obtained in this way, this research finds that most of the variances are contained in the first \( P \) coordinates and the variances in the subsequent coordinates are almost zero. Moreover, the cumulative contribution rate is calculated by the value of the variance, selecting the first \( P \) principal components corresponding to the cumulative contribution rate of approximately 99\%. Thus, the structure of the simulation fingerprint database \( \Psi_{k,3} \) is defined as

\[
\Psi_{k,3} = \begin{bmatrix}
\chi_{1,1} & \chi_{1,2} & \cdots & \chi_{1,N} \\
\chi_{2,1} & \chi_{2,2} & \cdots & \chi_{2,N} \\
\vdots & \vdots & \ddots & \vdots \\
\chi_{P,1} & \chi_{P,2} & \cdots & \chi_{P,N}
\end{bmatrix}
\]

(9)

where \( P \) is the fingerprint dimension (\( P < H \)) and \( \chi \) is the transformed fingerprint feature after PCA.

4 | BAGGING-KELM MATCHING MODEL

By the optical PD simulation above, a complete simulation database has been built. To match the actual PD signal detected by the optical sensor with the fingerprint in the simulation database, the Bagging-KELM model is used to identify the location of the PD source in this paper. Based on the bagging ensemble learning framework, \( C \) subsets are sampled from the simulation fingerprint database by the bootstrapped sample method. C KELM models are used as the base classifier to learn and test the \( C \) subsets. Finally, the location of the PD source is obtained by averaging the coordinates of the \( C \) base classifiers.

4.1 | KELM principle

The KELM is a single hidden layer feed-forward neural network (SLFN), which is an improvement in the ELM model. The KELM model is trained by minimizing the output weight norm and the training error instead of adjusting the network iteratively. The random hidden layer output matrix of the ELM is replaced by a kernel matrix, which shows better generalization performance than the ELM [21].

For SLFN with \( L \) hidden layer nodes, assuming the number of training sample datasets \( \{x_i, y_i; x_i \in \mathbb{R}^m, y_i \in \mathbb{R}^n, j = 1, 2, \ldots, N\} \) is \( l \), the expression of the model output is shown below:

\[
y(x) = \sum_{i=1}^{L} \hat{\beta}_i \phi_i(x) = \sum_{i=1}^{L} \hat{\beta}_i g_i(xw_i + b_i), \hat{\beta}_i \in \mathbb{R}^n
\]

(10)

where \( x_i = [x_{i1}, x_{i2}, \ldots, x_{im}]^T, y_i = [y_{i1}, y_{i2}, \ldots, y_{in}]^T \), each sample contains \( m \)-dimensional features, \( b_i(x) \) is the output function of the \( i \)-th hidden layer node, \( b_i \) is the bias of the \( i \)-th hidden layer, \( w_i \) is the input weight vector, and \( \hat{\beta}_i \) is the weight vector of the \( i \)-th hidden layer node and the output layer node.

In the KELM model, the hidden layer node is replaced by the radial basis function (RBF) node, and its activation function is as follows:

\[
g_i(w_i \cdot x + b_i) = \exp\left(-b_i\|x-w_i\|^2\right)
\]

(11)

When the excitation function can approach any \( l \) samples with zero error, that is,

\[
\sum_{j=1}^{N} \|y_j - t_j\| = 0
\]

(12)

The mathematical model of ELM can be derived as follows:

\[
\sum_{i=1}^{L} \hat{\beta}_i g_i(w_i \cdot x_j + b_i) = t_j
\]

(13)
Equation (13) can be expressed as follows:

\[
\mathbf{H} \mathbf{\beta} = \mathbf{T}
\]

(14)

where \( \mathbf{\beta} \) is the vector of output layer weights, \( \mathbf{T} \) represents the class label, and \( \mathbf{H} \) is the hidden layer output matrix:

\[
\mathbf{H} = \begin{bmatrix}
 g(\mathbf{w}_1 \cdot \mathbf{x}_1 + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_1 + b_L) \\
 \vdots & & \vdots \\
 g(\mathbf{w}_1 \cdot \mathbf{x}_N + b_1) & \cdots & g(\mathbf{w}_L \cdot \mathbf{x}_N + b_L)
\end{bmatrix}_{N \times L}
\]

(15)

Because Equation (14) is linear, \( \mathbf{\beta} \) can be obtained by the following equation:

\[
\mathbf{\beta} = \mathbf{H}^\dagger \mathbf{T}
\]

(16)

where \( \mathbf{H}^\dagger \) is the Moore-Penrose generalized inverse of the hidden layer output matrix [22]. To improve the generalization ability and accuracy of ELM, KELM is introduced to the kernel function to avoid the problem of stochastically generating bias values and the input weight on the ELM. The output layer weights of the KELM are determined as follows:

\[
\mathbf{\beta} = \mathbf{H}^\dagger \left( \frac{1}{C} + \mathbf{H} \mathbf{H}^\dagger \right)^{-1} \mathbf{T}
\]

(17)

where \( C \) is the penalty coefficient. The output function for the model is

\[
f(\mathbf{x}_j) = b(\mathbf{x}_j) \mathbf{\beta} = b(\mathbf{x}_j) \mathbf{H}^\dagger \left( \frac{1}{C} + \mathbf{H} \mathbf{H}^\dagger \right)^{-1} \mathbf{T}
\]

(18)

where \( b(\mathbf{x}_j) \) represents the output function of the hidden nodes and the feature mapping function that maps the data from the input space to the hidden layer feature space \( \mathbf{H} \). When \( b(\mathbf{x}_j) \) is unknown, the kernel function matrix is calculated as follows:

\[
\begin{cases}
\mathbf{\Omega}_{\text{ELM}} = \mathbf{H} \mathbf{H}^\dagger \\
\mathbf{\Omega}_{\text{ELM}}^{(i,j)} = b(\mathbf{x}_i) \cdot b(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)
\end{cases}
\]

(19)

where \( K(\mathbf{x}_i, \mathbf{x}_j) \) represents the RBF kernel function, which can be written as

\[
K(\mathbf{x}_i, \mathbf{x}_j) = \exp \left( -\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2} \right)
\]

(20)

where \( \sigma \) is the kernel function parameter factor. According to the equation above, the output function of KELM is

\[
f(\mathbf{x}) = \begin{bmatrix}
 K(\mathbf{x}, \mathbf{x}_1) \\
 K(\mathbf{x}, \mathbf{x}_2) \\
 \vdots \\
 K(\mathbf{x}, \mathbf{x}_N)
\end{bmatrix}^T \left( \frac{1}{C} + \mathbf{\Omega}_{\text{ELM}} \right)^{-1} \mathbf{T}
\]

(21)

### 4.2 Bagging framework

By bootstrap aggregating, the bagging framework implements the resampling of the optical simulation fingerprint database to form a sub-database \( \mathcal{C} \) that is similar in size to the simulation fingerprint database but is different. Bagging uses a randomly selected sub-database of the simulation fingerprint database to train each KELM model, which helps reduce variance and helps avoid overfitting [23]. The final PD source location is calculated by the average value of all KELM results. Based on the framework, the step of obtaining an actual PD source location detected by the optical sensor is shown as Algorithm 1.

#### Algorithm 1: Bagging-KELM

| Input: |
|-------|
| KELM: base classifier; |
| \( C \): base classifiers number; |
| \( \mathbf{X} \in \mathbb{R}^{N \times m} \): simulation fingerprint database of input training; |
| \( \mathbf{y} \): the test sample of the actually detected PD fingerprint; |
| \( \mathbf{v} \): the average of \( x, y, \) and \( z \)-axis coordinates in \( \mathbf{V}(i) \); |
| \( C = 1 \): |
| Get new sub-databases \( \mathbf{X}_c \in \mathbb{R}^{N \times m} \) from \( \mathbf{X} \) by using the bootstrap method; |
| Calculate the 3D-coordinate value \( \mathbf{V}(i) \) for the test sample \( \mathbf{y} \), using the KELM model trained by \( \mathbf{X}_c \); |
| Return vector \( \mathbf{V}(i) \) including \( C \) testing 3D-coordinate value of \( \mathbf{y} \); |
| Calculate \( \mathbf{V} \) : the average of \( x, y, \) and \( z \)-axis coordinates in \( \mathbf{V}(i) \); |
| Output: PD source location = \( \mathbf{V} \). |

### 5 EXPERIMENTAL VERIFICATION OF PD LOCALIZATION USING 3D-OSIF AND BAGGING-KELM

In this paper, a novel method of PD localization combining 3D-OSIF and Bagging-KELM is proposed, which mainly has two parts: 3D-OSIF database simulation and Bagging-KELM matching. The overall process of PD localization is shown in Figure 5.

#### 5.1 PD simulation experiment based on 3D-OSIF

In the simulation process, theoretically, PD simulations should be conducted at all locations in the GIS model. However, considering the possibility of operation in practice, this research selects as many PD sources as possible and then obtains the PD information of the remaining locations in the model by biharmonic spline interpolation fitting [24].

Based on the settings of the GIS simulation model in Section 2, 27 cross sections are selected inside the simulation model; these cross sections are separated by 10 mm between...
are fitted to obtain the optical signals collected by each probe when PD occurs at any position in the GIL model. Accordingly, a complete optical simulation database $\Psi$ is established. Taking one column of the simulation probes (No. 1 to No. 3 in Figure 2.) as an example, the relative light irradiance values collected by one simulation probe when a PD occurs at each location in the GIL model are recorded as the optical PD simulation fingerprint map of this probe, as shown in Figure 7. The optical PD simulation fingerprints of the remaining two columns of simulation probes can be obtained by rotating 120° and 240° around the axis.

However, when performing a calculation via a computer, the dimension of the fingerprint database cannot be infinite. Therefore, on the premise of ensuring the matching accuracy and efficiency, in this study, 21,033 PD fingerprints are sampled at a regular interval from the fitted fingerprint database to establish the final optical fingerprint database $\Psi_{\text{final}}$ whose dimension is $21,033 \times 9$. Each column vector $\Psi_{\text{final}}^j [\phi_{1,j}^\prime, \phi_{2,j}^\prime, \ldots, \phi_{9,j}^\prime]^T$ of $\Psi_{\text{final}}$ represents one optical PD fingerprint, where $\phi_{i,j}^\prime$ ($i = 1, 2, \ldots, 9; j = 1, 2, \ldots, 21,033$) is the feature of the PD fingerprint.

5.2 | Experiment of optical PD localization

To verify the effectiveness of the localization method in this paper, an experimental platform of PD localization is built, as shown in Figure 8.

In the platform, the voltage range of the corona-free AC voltage regulator is 0–150 kV. The GIL tank is an aluminium
experiment tank, as shown in Figure 2, with good airtight performance and no incident light. Nine openings on the GIL tank are installed with 9 identical optical sensors. Each optical sensor is composed of fluorescent fibres of the same type and length. The photon counter is HAMAMATSU H11890-210, whose spectral response range is 230–700 nm, and each counting threshold is 1000 ms. A digital PD detector (Haffley DDX 9121b) is used to detect whether PD occurs. The GIL tank is filled with SF6 at a pressure of 0.2 MPa.

In this paper, sixteen PD localization experiments are conducted with different locations of PD sources. During the experiment, the PD at different locations in the GIL tank is produced by adjusting the angle, height, and radial length of the needle-plate model. The space of the needle plate is always maintained at 6 mm, regarding the location of the needle tip as the actual location of the PD source. When a voltage is applied to the GIL tank until a stable PD phenomenon occurs, the optical signals generated by the PD are collected by optical sensors and the photon counter. Since the number of photons is proportional to the optical power, this article uses the number of photons as an index to measure the light irradiance of the PD [25]. To reduce the impact of the PD signal instability, a photon counter is used to collect the photon number of 60 thresholds whose average value \( \phi_{ij}^{\text{detect}} \) (\( i = 1, 2, \ldots, 9; j = 1, 2, \ldots, 16 \)) is recorded as the optical fingerprint feature of this PD source. Finally, sixteen optical detection fingerprints \( \Psi_{\text{detect}}^{j} \) are established by optical fingerprint features \( \phi_{ij}^{\text{detect}} \) to prepare for Bagging-KELM matching.

### 5.3 Optical PD fingerprint matching by bagging-KELM

Based on Sections 5.1 and 5.2, this research processes the optical fingerprint database \( \Psi_{\text{final}} \) and detection fingerprint \( \Psi_{\text{detect}}^{j} \) according to the construction methods of K1, K2 and K3. Three kinds of fingerprint databases are put into the Bagging-KELM model. In the bagging framework, the number \( C \) of base classifiers that is KELM is 10. For each testing fingerprint, the Bagging-KELM model will obtain 10 location coordinates, where the average of the \( x \)-axis, \( y \)-axis, and \( z \)-axis coordinates is recorded as the spatial location of this PD source.

To compare the performance of the matching model in this paper with other models, this research also uses the KELM
model and back propagation neural network (BPNN) model to match the PD fingerprint. The experimental results are shown in Section 6.

6 | RESULTS AND DISCUSSION

According to the experiment above, the test results of different classifiers for three kinds of fingerprints (K1, K2, K3) are shown in Table 1. The cumulative density function (CDF) of detection error is shown in Figure 9.

The average error of applying the Bagging-KELM for K3 is 0.93 cm, which is the smallest. Furthermore, the percent of detection error that is less than 1.5 cm accounts for 93.75% by applying Bagging-KELM for K3, which is much higher than that of the others. In addition, the root means square error (RMSE) and error variance for applying Bagging-KELM to K3 are 10.04 and 3.66, respectively, showing the best performance of all. Therefore, the results indicate that using the Bagging-KELM model to locate the fingerprint database in the form of K3 has the highest accuracy and stability, which can well meet the requirements of PD localization in the GIL.

For the three matching models, it can be seen that the average RMSE of the Bagging-KELM model to the three fingerprint databases is 11.77, which is smaller than the 14.35 and 57.03 of the KELM and BPNN models, respectively. The KELM is more stable than the BPNN but less accurate than the Bagging-KELM model. The BPNN has poor accuracy and stability. Thus, this shows that the Bagging-KELM model has strong adaptability while having a high localization level.

For the three kinds of fingerprint databases, the average RMSE of the K3 fingerprint database under the three matching models is 25.65, which is smaller than the 28.55 and 28.96 of the K1 and K2 fingerprint databases. This finding indicates that the K3 fingerprint database can better characterize the features of optical PD fingerprints for 3D-OSIF localization.

This research compares the localization results of this paper with some other current PD localization methods. For the existing optical PD localization method, the minimum detection error in References [8] and [9] is 8 cm, and only the position where the PD experiment is performed can be identified. The minimum detection error in this article can be adjusted according to the spatial resolution of the simulation fingerprint, and is not limited by the field PD experiment. For the current UHF PD localization method, the detection error reference [11] is 2.21 m, which is much larger than the 0.93 cm in this paper.

For PD acoustic localization method, the detection error in the transformer in reference [26] can reach 11.7 cm, which is also inferior to the localization accuracy in this article. Therefore, by comparing with the existing main PD localization methods, the method proposed in this paper has a better localization effect.

7 | CONCLUSION

In this paper, a novel method that can identify the location of PD sources is proposed. An optical PD simulation model of

FIGURE 9  CDF plots of the detection error for the three kinds of fingerprints using different matching methods. (a) CDF of K1, (b) CDF of K2, (c) CDF of K3.
the GIL has been set up to simulate the generation of PDs at different locations. The optical PD simulation signals captured by several simulation probes strategically arranged in particular positions are used to construct three kinds of fingerprint databases (K1, K2 and K3). At the same time, a GIL PD localization experimental platform identical to the simulation model is built. The optical PD fingerprints detected by real optical sensors mounted at the same position as the simulation probe are used to match with the optical simulation fingerprint database through the Bagging-KELM model. Finally, a comparative study was performed to prove the effectiveness of the locating method. These results lead to the following conclusions:

1. The characteristics of the PD fingerprints influence the localization results. The PD fingerprint database (K3) constructed by performing PCA processing on the differences in PD light irradiance detected by each sensor has richer PD location information, which can improve the accuracy of the localization.

2. Using the bagging framework to optimize the KELM model can improve the performance of the matching model. In particular, when the Bagging-KELM model is used to match the K3 fingerprint database, the average localization error can reach 0.93 cm, and 93.75% of the error is less than 1.5 cm. Meanwhile, the error variance of the Bagging-KELM method is only 3.66, which means that this method has high stability.

3. This article introduces the optical simulation fingerprint database based on 3D-OsIF into PD localization for the first time, which effectively solves the problem that a large number of PD localization data cannot be collected at the equipment site. Furthermore, the optical simulation model can simulate the optical signal emitted from PD sources and be constructed based on different real GIL equipment in TracePro software, which has a wide range of applications.

REFERENCES

1. Niu, H., et al.: Multi-physical coupling field study of 500-kV GIL: Simulation, characteristics, and analysis. IEEE Access 8, 131439–131448 (2020)
2. Tu, Y., et al.: ±100-kV HVDC SF6/N2 Gas-insulated transmission line. IEEE Trans. Power Delivery 35(2), 735–744 (2020)
3. Toigo, C., et al.: Partial discharge behavior of protrusion on high voltage conductor in GIS/GIL under high voltage direct current: Comparison of SF6 and SF6 alternative gases. In IEEE Trans. Dielectr. Electr. Insul. 27(1), 140–147 (2020)
4. Stone, G.C.: Partial discharge diagnostics and electrical equipment insulation condition assessment. IEEE Trans. Dielectr. Electr. Insul. 12(5), 891–903 (2005)
5. Ren, M., et al.: Optical partial discharge diagnosis in SF6 gas-insulated system with SiPM-based sensor array. IEEE Sens. J. 18(13), 5532–5540 (2018)
6. Han, X., et al.: Study on the combined characteristics of UHF and optical signals induced by partial discharge at spacer surface in GIS. IET Gener. Transm. Distrib. 14(16), 3332–3337 (2020)
7. Guzman, H., Kranz, H.-G.: Analysis of internal partial discharges in solid dielectrics using an electro-optical measurement system. In: Proceedings of the 1995 IEEE 5th International Conference on Conduction and Breakdown in Solid Dielectrics. Leicester, UK, pp. 239–243 (1995)
8. Biswas, S., et al.: A methodology for identification and localization of partial discharge sources using optical sensors. IEEE Trans. Dielectr. Electr. Insul. 19(1), 18–28 (2012)
9. Biswas, S., et al.: An approach based on rough set theory for identification of single and multiple partial discharge source. Int. J. Electr. Power Energy Syst. 46, 163–174 (2013)
10. Baug, A., et al.: Identification of single and multiple partial discharge sources by optical method using mathematical morphology aided sparse representation classifier. IEEE Trans. Dielectr. Electr. Insul. 24(6), 3703–3712 (2017)
11. Li, Z., et al.: UHF partial discharge localization algorithm based on compressed sensing. IEEE Trans. Dielectr. Electr. Insul. 25(1), 21–29 (2018)
12. Wang, M., et al.: Grey wolf optimization evolving kernel extreme learning machine: Application to bankruptcy prediction. Eng. Appl. Artif. Intel. 63, 54–68 (2017)
13. Yao, B., et al.: Bagging based metric learning for person re-identification. In: Proceedings of IEEE International Conference on Multimedia and Expo. Chengdu, China, 1–6 (2014)
14. Rizvi, Z.S., et al.: A kernel-based PCA approach to model reduction of linear parameter-varying systems. IEEE Trans. Control Syst. Technol. 24(5), 1883–1891 (2016)
15. Li, J., et al.: A novel GIS partial discharge detection sensor with integrated optical and UHF methods. IEEE Trans. Power Delivery 33(4), 2047–2049 (2016)
16. Xu, Y., et al.: Simulation analysis on the propagation of the optical partial discharge signal in l-shaped and L-shaped GILs. IEEE Trans. Dielectr. Electr. Insul. 25(4), 1421–1428 (2018)
17. Ren, M., et al.: Optical partial discharge diagnostic in SF6 gas insulated system via multi-spectral detection. ISA Trans. 75, 247–257 (2018)
18. Qi, C., et al.: Slant light suppression in BRDF measurement infrared optical system. Chin. Opt. Lett. 1(7), 398–400 (2003)
19. Miyazaki, A., et al.: Long-distance 275-kV GIL monitoring system using fiber-optic technology. IEEE Trans. Power Delivery 18(4), 1545–1553 (2003)
20. Moore, B.: Principal component analysis in linear systems: Controllability, observability, and model reduction. IEEE Trans. Autom. Control 26(1), 17–32 (1981)
21. Huang, G.-B., et al.: Extreme learning machine for regression and multi-class classification. IEEE Trans. Syst. Man Cybern. Part B Cybern. 42(2), 513–529 (2011)
22. Huang, G.-B., Zhu, Q.-Y., Siew, C.-K.: Extreme learning machine: Theory and applications. Neurocomputing 70(1–3), 489–501 (2006)
23. Breiman, L.: Bagging predictors. Mach. Learn. 24(2), 123–140 (1996)
24. Sandwell, D.T.: Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. Geophys. Res. Lett. 14(2), 139–142 (1987)
25. Cheung, J.Y., et al.: Low optical power reference detector implemented in the validation of two independent techniques for calibrating photon-counting detectors. Opt. Express 19(21), 20347–20363 (2011)
26. Wang, Y., et al.: Acoustic localization of partial discharge sources in power transformers using a particle-swarm-optimization-route-searching algorithm. IEEE Trans. Dielectr. Electr. Insul. 24(6), 3647–3656 (2017)

How to cite this article: Zang, Y., et al.: A novel partial discharge localization method for GIL based on the 3D optical signal irradiance fingerprint and bagging-KELM. IET Gener. Transm. Distrib. 15, 2240–2249 (2021). https://doi.org/10.1049/gtd2.12173