energies

MDPI

Article

A Virtual–Physical Sensing Method for Fault-Gas Source Localization in High-Voltage Switchgear

Yujian Zhou [1] and Hao Luo [2,*]

[1] CHN Energy Guangtou Beihai Power Generation Co., Ltd., Beihai 536000, China; 20024359@chnenergy.com.cn
[2] State Key Laboratory of Power Transmission Equipment Technology, School of Electrical Engineering, Chongqing University, Chongqing 400044, China
* Correspondence: haoluo@cqu.edu.cn

Abstract

Characteristic gas monitoring is an important approach for detecting incipient faults and supporting the intelligent operation and maintenance of high-voltage switchgear. However, existing methods primarily rely on concentration thresholds or steady-state features, making it difficult to simultaneously optimize sensor placement and accurately localize fault-gas sources. To address these limitations, this study proposes a virtual–physical sensing method for fault-gas source localization. A full-scale test platform and a 1:1 three-dimensional numerical model of 10 kV air-insulated switchgear (AIS) were constructed to obtain the time-series responses of CO, CO_2, NO, and O_3 for different fault-gas source locations. Candidate monitoring points were quantitatively evaluated using a combination of the Welch *t*-test, mutual information, and entropy difference. Multi-gas features from the selected monitoring points were then fused to identify nine fault-location classes. The results showed that the reported relative errors in final recorded concentrations did not exceed 9.09%; MP8, MP1, and MP2 were the three highest-ranked monitoring locations; under five-fold stratified cross-validation, the accuracy, macro-precision, macro-recall, and macro-F1 score of the random forest model all reached 100%. The proposed method provides a quantitative basis for reducing the number of gas sensors and localizing fault-gas sources in high-voltage switchgear.

Keywords: fault-gas source localization; fault-gas transport; high-voltage switchgear; random forest; sensor-layout optimization; virtual–physical sensing

Academic Editor: Marco Faifer

Received: 29 September 2026
Revised: 27 October 2026
Accepted: 3 November 2026
Published: 4 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

As a key component of power distribution systems, high-voltage switchgear plays an important role in ensuring the safety and reliability of power supply [1,2]. During long-term operation, contacts, busbar joints, cable terminations, and insulating supports are subjected to combined electrical, thermal, and mechanical stresses and may develop incipient faults such as partial discharge, contact overheating, and insulation aging. These processes can induce reactions in the enclosed air and organic insulating materials and may be accompanied by characteristic gases including ozone (O_3), nitric oxide (NO), carbon monoxide (CO), and carbon dioxide (CO_2). More generally, localization studies using spatially distributed measurements have shown that source-dependent propagation responses can provide useful spatial information [3,4].

 https://doi.org/10.3390/en19215296

A broad range of electrical, electromagnetic, and acoustic approaches—including transient earth voltage, ultra-high-frequency, ultrasonic, and pulse-current measurements—have been investigated for partial-discharge detection, diagnosis, and localization in enclosed high-voltage equipment [2,5–8]. These techniques directly reflect electromagnetic or acoustic phenomena associated with discharge activity, but practical measurements may be affected by electromagnetic interference, propagation paths, enclosure shielding, and sensor installation location. Gas-based monitoring provides a complementary route by measuring chemical responses inside the enclosure. Existing gas-based methods, however, commonly rely on single-point concentration thresholds or steady-state parameters and therefore contain limited spatial information. Machine-learning methods have also been applied to switchgear condition monitoring and fault-type recognition [5,6,8], but most studies establish a mapping between measured features and fault categories without explicitly exploiting the dynamic transport process from a source to multiple monitoring points. As a result, distinguishing the spatial origin of a gas source remains challenging.

Previous studies have investigated gas decomposition and transport in enclosed electrical equipment using computational fluid dynamics and multiphysics modeling [9,10]. Three-dimensional models show that internal structures and temperature-driven flow can redirect transport paths and produce localized accumulation regions. Temperature gradients and natural convection alter gas migration velocity and spatial distribution, while differences in molecular mass and transport properties produce gas-dependent accumulation behavior. Although these studies establish an important physical basis for gas-transport-based monitoring, full-scale validation with multi-point measurements and a direct connection among validated transport modeling, sensor-layout optimization, and spatial source localization remain limited. Consequently, a virtual–physical framework that explicitly links transport validation, monitoring-point selection, and multiclass source localization is still needed.

To address these issues, this study proposes a virtual–physical sensing method for fault-gas source localization in 10 kV AIS. Physical measurements are combined with virtual sensing data from a validated three-dimensional transport model to evaluate candidate monitoring locations and identify nine predefined source-location classes. The main contributions are as follows:

(1) A full-scale fault-gas transport platform and a corresponding 1:1 three-dimensional model of 10 kV AIS were established, retaining the principal internal structures and natural-convection conditions. The numerical model was validated against physical concentration responses before being used to generate virtual sensing data.

(2) Eight candidate monitoring locations were ranked using a multi-metric evaluation method combining the Welch *t*-test, mutual information, and entropy difference, with equal weights assigned to the three metrics. MP8, MP1, and MP2 ranked first, second, and third, with composite scores of 100, 72.38, and 54.65, respectively. These locations form a high-ranking three-site configuration within the candidate set. The minimum sufficient number of sensors, however, remains to be determined through direct comparison with one- and two-site arrangements.

(3) Multi-gas responses from the selected monitoring points were fused to formulate a nine-class source-localization problem. Five classifiers—Random Forest, Decision Tree, RBF-SVM, KNN, and GaussianNB—were compared to assess the discriminative capability and model dependence of the optimized sensor layout.

The remainder of this paper is organized as follows. Section 2 describes the overall framework, numerical model, sensor-layout evaluation method, and multiclass source-localization approach. Section 3 presents the full-scale experimental platform, numerical model, and virtual–physical validation results. Section 4 discusses the sensor-layout

optimization and nine-class localization results. Finally, Section 5 summarizes the principal findings, limitations, and directions for further validation.

2. Materials and Methods

2.1. Overall Framework and Technical Approach

The overall research framework of this paper is shown in Figure 1. First, a test platform for fault gas transport in 10 kV high-voltage switchgear was constructed, and a corresponding 1:1 three-dimensional simulation model was established based on the structural dimensions of the physical switchgear. The accuracy and reliability of the simulation model were verified by comparing the gas concentration response curves obtained from physical experiments and numerical simulations under different fault location conditions. Building on this, a simulation analysis of fault gas transport characteristics under complex structural and natural convection conditions was conducted. The simulations generated concentration time-series data for different fault-source locations, gas species, and monitoring points, revealing the effects of diffusion, convection, flow diversion, and local accumulation of characteristic gases within the switchgear. Finally, using statistical analysis methods such as *t*-tests, mutual information, and entropy differences, the ability of each candidate monitoring point to distinguish among fault-source locations was evaluated. Furthermore, fault-source localization performance was assessed using a random forest model, thereby enabling the optimized layout of fault gas sensors in high-voltage switchgear and the tracing of fault locations.

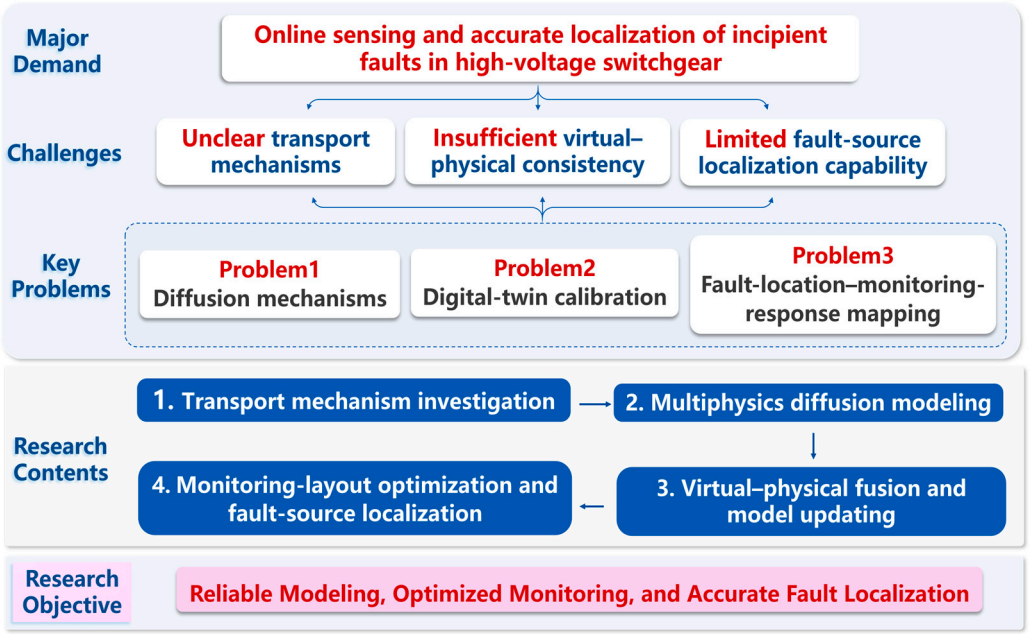


Figure 1. Overall framework of the proposed fault-gas source localization method.

2.2. Numerical Modeling of Fault-Gas Transport in High-Voltage Switchgear

A numerical model of fault-gas transport in high-voltage switchgear was established using ANSYS Fluent 2026 R1. Because the concentrations of the fault gases were low and their transport was dominated by molecular diffusion and natural convection in air, the species transport model was adopted to describe the air–gas mixture. The spatiotemporal distribution of gas concentrations is obtained by solving the continuity, momentum, energy, and species transport equations.

Air was treated as the background carrier medium, while CO, CO_2, NO, and O_3 were investigated as separate dilute tracer species. For each numerical scenario, only one target

gas was released at one predefined source location; the four gases were not simultaneously released as a four-component mixture. The same prescribed release rate and duration were used across source locations for comparison. All enclosure walls and internal solid surfaces were treated as no-slip, impermeable boundaries. The four gas-specific response features were combined only at the subsequent machine-learning feature-fusion stage.

The numerical procedure consisted of two stages. First, the steady-state temperature and buoyancy-driven airflow fields were calculated without tracer-gas release. Second, the converged fields were used to initialize the time-dependent species-transport simulations following activation of a prescribed gas source. Gravity was set to 9.81 m/s^2, the initial temperature to 25 °C, the reference pressure to 101,325 Pa, and the initial tracer-gas mass fraction to zero. The physical observation period does not itself determine whether the numerical solution is transient; here, time-dependent refers to the calculated temporal evolution of species concentrations. Model settings and gas properties are summarized in Tables 1 and 2.

Numerical verification and reporting. Numerical discretization uncertainty is distinguished from comparison with experimental measurements. Grid refinement was assessed by comparing at least three systematically refined grids under identical boundary and source conditions, with refinement ratios and grid-convergence behavior reported for representative quantities of interest. Here, appropriate quantities are the arrival time, peak concentration and final recorded concentration at selected monitoring points; the same quantities should be checked after reducing the time step. Iterative convergence should be judged jointly from scaled equation residuals, mass/species balance and the stability of monitored concentrations, rather than residual thresholds alone. The model description should also identify mesh type and cell counts, mesh-quality measures, time-step size, residual criteria, species-in-air diffusion coefficients, and the diffusion model used. The literature on CFD discretization uncertainty provides a procedure, but does not establish convergence of the present model without its original mesh and solver records.

As an illustrative grid-refinement example, three computational meshes containing 420,000, 960,000, and 2,050,000 cells are associated with MP8 peak concentrations of 25.80, 25.15, and 24.90 ppm, respectively. Relative to the finest-grid value, the coarse- and medium-grid differences are 3.61% and 1.00%. The example shows how monitoring-point peak concentration can be used to compare successive grid resolutions; these values are a methodological illustration rather than outputs from the present Fluent model.

Assessment of the assumed flow regime. Natural convection is induced by the approximately 10 ± 2 °C vertical temperature difference maintained in the physical enclosure; its intensity depends on geometry, buoyancy and thermal transport, not simply on the absence of a fan. For a velocity-based assessment, $Re = UL/\nu$, where U is a representative simulated circulation speed, L is a justified flow-path or hydraulic length scale and ν is the kinematic viscosity of air. For thermal buoyancy, $Ra = g\beta\Delta TL^3/(\nu\alpha)$, where β is the thermal expansion coefficient and α is thermal diffusivity. As an order-of-magnitude illustration only, using L = 0.8 m, ΔT = 10 K, β = 1/298 K^{-1}, $\nu = 1.5 \times 10^{-5}$ m^2/s and $\alpha = 2.1 \times 10^{-5}$ m^2/s gives $Ra \approx 5.4 \times 10^8$. This value depends strongly on the chosen L ($Ra \propto L^3$), does not represent an extracted CFD result, and cannot establish laminar flow on its own in this obstructed, nonuniformly heated enclosure. In buoyancy-driven enclosure flows, flow-regime transitions are influenced by enclosure and obstacle geometry, as well as by thermal boundary conditions. Accordingly, the laminar approximation is retained as a modeling assumption for the controlled case, subject to confirmation from the actual velocity field and, if required, a turbulence-model sensitivity comparison.

For a representative parameter comparison with L = 0.8 m, ϱ = 1.18 kg/m^3 and $\mu = 1.85 \times 10^{-5}$ Pa·s, characteristic speeds of 0.005, 0.010, 0.020 and 0.030 m/s correspond to

Reynolds numbers of approximately 255, 510, 1021 and 1531, respectively. This velocity range demonstrates how the flow-regime assessment depends on local circulation intensity; the appropriate model choice must also consider buoyancy and cavity geometry.

Table 1. Parameter settings for numerical simulation of fault-gas transport.

Project	Settings
Simulation Model	Species transport model, energy equation
Flow Conditions	Low-speed natural convection, laminar flow
Solver	Pressure-based solver
Calculation Type	Steady-state flow field calculation and transient species transport calculation
Pressure–Velocity Coupling	SIMPLE for steady-state, PISO for transient
Spatial Discretization Method	Finite Volume Method
Gradient Calculation Method	Element-based Least Squares Method
Wall Conditions	No-slip, impermeable walls
Initial Temperature	25 °C
Reference Pressure	101,325 Pa
Gravitational Acceleration	9.81 m/s^2

Table 2. Characteristic parameters of the simulated fault gases.

Gas Type	Relative Molecular Mass	Density (kg/m^3)	Specific Heat at Constant Pressure [J/(kg·K)]	Thermal Conductivity [W/(m·K)]	Dynamic Viscosity (Pa·s)
CO	28.01	1.145	1040	0.0248	1.75×10^{-5}
CO_2	44.01	1.842	846	0.0166	1.48×10^{-5}
NO	30.01	1.226	995	0.0258	1.80×10^{-5}
O_3	48.00	1.962	818	0.0170	1.90×10^{-5}

Because the relative molecular masses of CO and NO are relatively close to that of air, their transport processes are influenced by both molecular diffusion and natural convection; the relative molecular masses of CO_2 and O_3 are significantly greater than that of air, making them more prone to downward settling and bottom accumulation in areas of weak flow. Therefore, the simulation process simultaneously accounts for the effects of gravity, temperature gradients, structural barriers, and species diffusion to enhance the realism of fault-gas transport modeling [9,10]. The values in Table 2 are thermophysical properties used in the species-transport calculation; they are not randomly selected fault concentrations and do not represent diagnostic thresholds.

2.3. Data-Driven Methods for Sensor-Layout Optimization and Traceability

2.3.1. Sensor-Layout Optimization Based on Multi-Metric Evaluation

The type and installation location of sensors directly affect the completeness of status information, fault sensitivity, and diagnostic reliability; therefore, candidate locations should be evaluated according to their source-discrimination capability [4]. To evaluate the eight candidate monitoring points for nine source-location classes, a comprehensive sensor-layout evaluation method integrating the Welch *t*-test, mutual information, and entropy difference was developed. The Welch *t*-test characterizes differences in class-dependent response distributions [11]; information entropy characterizes distributional uncertainty [12]; and mutual information quantifies the statistical dependence between monitoring-point features and source-location labels [13]. All three metrics were normalized to a range of 0–100, and the composite score was calculated using equal weighting. The specific definitions of the metrics are as follows:

1. Welch's *t*-test

Used to assess the significance of differences in the distribution of gas concentrations between the gas-producing location and normal monitoring points. The larger the absolute T_t value, the more pronounced the difference in the distribution of data between the two locations, which aids in fault identification. The calculation formula is:

$$T_t = \frac{X_1 - X_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}} \tag{1}$$

In this equation, T_t is the statistical value of the T_t test; X_1, X_2, s_1^2, s_2^2, n_1, n_2 represent the mean, variance, and sample size of the gas concentrations at the gas production site and at the normal monitoring points, respectively.

2. Mutual Information

Mutual information measures the statistical dependence between the features extracted at a monitoring point and the corresponding fault-location labels. The calculation formula is as follows:

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{2}$$

In this equation, $I(X;Y)$ and $p(x,y)$ represent, respectively, the mutual information between the monitoring-point feature variable X and the fault-location label Y, and the joint probability distribution; $p(x)$ and $p(y)$ represent the marginal probability distributions of X and Y. Mutual information measures nonlinear statistical dependence between monitoring-point features and fault categories and has been widely used for supervised feature selection [13].

3. Entropy Difference

The entropy difference measures the uncertainty between variables, where the entropies of X and Y are calculated as follows: [12].

$$\begin{cases} H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \\ H(Y) = -\sum_{i=1}^{n} p(y_i) \log p(y_i) \end{cases} \tag{3}$$

The calculation of the entropy difference can be expressed as:

$$\Delta H = |H(X) - H(Y)| \tag{4}$$

For each source-location class, 25 samples were generated by applying controlled noise and small perturbations to the corresponding parent multi-gas time-series response [14]. This augmentation was introduced to represent local measurement variability and increase within-class diversity; the resulting 25 samples per class are therefore augmented realizations rather than 25 independent physical experiments. Because samples within a class share a common parent response, the subsequent cross-validation results should be interpreted as internal robustness and class-separability results for the augmented dataset, not as independent field-generalization estimates.

Finally, based on the three evaluation metrics described above, this paper constructs a comprehensive scoring model for the four fault gases:

$$S = \alpha \times |T_t| + \beta \times I(X;Y) + \gamma \times \Delta H \tag{5}$$

Here, S denotes the weighted sum of three dimensionless component scores, each normalized to 0–100 across the eight candidate monitoring positions. The baseline weights are $\alpha = \beta = \gamma = 1/3$, representing a transparent, nonpreferential reference choice rather than

a data-verified optimum. For min–max scaling, if that was the implemented procedure, Xj = 100(Xj − Xmin)/(Xmax − Xmin), with constant-valued components treated explicitly. A useful conservative sensitivity check is available without the original component-score matrix: for any two monitoring points i and j with component scores in [0, 100], a change in the weight vector from w0 to w changes their score difference by no more than 100 $||w - w0||_1$. Under the reported baseline scores, the margins MP8–MP1 = 27.62, MP1–MP2 = 17.73, and MP2–MP7 = 1.09. Thus the relative ordering of MP8 and MP1 is guaranteed for $||w - w0||_1 < 0.2762$, and that of MP1 and MP2 for $||w - w0||_1 < 0.1773$; by contrast the MP2–MP7 ordering can be guaranteed from these margins alone only for $||w - w0||_1 < 0.0109$. This bound is a mathematical worst-case sensitivity statement, not an empirical reranking of the eight sites. In particular, the stability of third place remains unverified without the original normalized metric scores. The orientation of the entropy-difference score and the Welch-test comparison groups must also match the actual implementation.

2.3.2. Nine-Class Fault-Source Localization and Performance Evaluation

Fault-gas source localization was formulated as a nine-class classification task using the predefined source positions as targets. The CO, CO_2, NO, and O_3 measurements were obtained in separate controlled gas-release scenarios, and their corresponding features were concatenated by source class to form a multi-gas feature representation. This feature fusion does not represent a physical measurement of the four gases released simultaneously.

Five classification algorithms with different inductive biases were compared: random forest, decision tree, radial-basis-function support vector machine (RBF-SVM), k-nearest neighbors (KNN), and Gaussian naive Bayes (GaussianNB) [15–18]. Their performance was evaluated consistently using the same nine-class augmented dataset. This comparison assesses classification separability under the present controlled conditions rather than external deployment accuracy. The schematic workflow of the random forest classifier is shown in Figure 2.

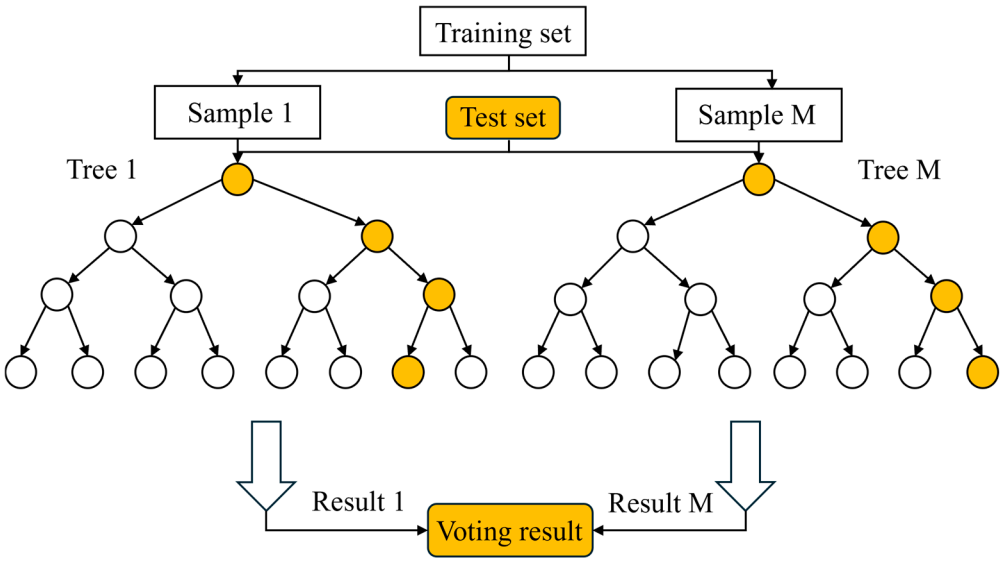


Figure 2. Schematic of the random forest classification process. (Yellow dots represent samples from the highlighted class, whereas white dots represent samples from the other classes. The final classification is determined by majority voting).

Five-fold stratified cross-validation was performed on the 225-sample augmented dataset. The samples were divided into five mutually exclusive folds while approximately preserving the proportions of the nine source-location classes; four folds were used for

training and one fold for validation in each round, and every fold served once as the validation set. The reported metrics were calculated from the resulting out-of-fold predictions. However, because the 25 samples within each class were generated by perturbing a common parent response, the folds are not equivalent to independent physical repetitions. The five-fold results are therefore used to compare classifiers and quantify separability within the controlled augmented dataset, rather than to claim unbiased external generalization. This limitation is explicitly considered when interpreting the 100% Random Forest result [19].

To comprehensively evaluate the performance of different sensor point combinations in identifying the nine fault locations, this paper selects accuracy (Accuracy), macro-precision (Macro-Precision), macro-recall (Macro-Recall), and macro-F1-score (Macro-F1) as evaluation metrics [20]. Accuracy reflects the overall correctness of the classifier's identification of all fault samples; Macro-Precision reflects the model's average ability to control misclassifications for each fault location; Macro-Recall reflects the average detection capability of the sensor layout for different fault locations; and Macro-F1-score comprehensively evaluates both misclassifications and missed detections across all fault types.

$$Accuracy = \frac{\sum_{i=1}^{K} M_{ii}}{\sum_{i=1}^{K}\sum_{j=1}^{K} M_{ij}} \tag{6}$$

For class i faults, the precision, recall, and F1-score are expressed as follows:

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \tag{7}$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \tag{8}$$

$$F1_i = \frac{2Precision_i Recall_i}{Precision_i + Recall_i} \tag{9}$$

In the formula, TP_i represents the number of samples from fault class i that are correctly identified; FP_i represents the number of samples from other fault classes that are incorrectly identified as belonging to class i; and FN_i represents the number of samples from fault class i that are incorrectly identified as belonging to other classes. A higher precision indicates that faults at other locations are less likely to be misclassified as the current fault location; a higher recall indicates that a higher proportion of faults occurring at the current fault location can be reliably detected by the sensor combination; and the F1-score is used to comprehensively reflect classification performance in terms of both misclassification and missed detection.

Since this paper involves nine types of fault locations, evaluating only the overall metrics across all samples may result in the evaluation results being influenced by categories that are easy to identify or those with a large number of samples. Therefore, this paper adopts a macro-average approach: first, the precision, recall, and F1-score for each of the nine fault types are calculated separately, and then the results for each category are averaged arithmetically:

$$Precision_{macro} = \frac{1}{K}\sum_{i=1}^{K} Precision_i \tag{10}$$

$$Recall_{macro} = \frac{1}{K}\sum_{i=1}^{K} Recall_i \tag{11}$$

$$F1_{macro} = \frac{1}{K}\sum_{i=1}^{K} F1_i \quad (12)$$

In this equation, K represents the number of fault categories; in this paper, $K = 9$. The macro-average method assigns equal evaluation weights to all nine fault categories, thereby preventing the overall evaluation results from being dominated by a single easily identifiable category or a category with a large number of samples. This approach is better suited for evaluating the sensor layout's ability to detect faults evenly across different spatial regions within the switchgear.

Therefore, during the sensor layout optimization process, it is necessary not only to focus on the overall accuracy of sensor point combinations but also to comprehensively analyze the macro-average precision, macro-average recall, and macro-average F1-score. If a particular sensor combination exhibits a high overall accuracy rate and macro-average F1-score, while the difference between the macro-average precision and macro-average recall is small, this indicates that the combination can both reduce misclassifications between different fault locations and minimize missed detections of faults at specific locations, demonstrating a relatively balanced recognition capability for all nine fault types. Based on this, The three highest-ranked monitoring points were therefore selected to construct the fused feature set for subsequent fault-source localization, thereby determining an optimized sensor layout that balances fault coverage, location discrimination capability, and recognition stability while controlling the number of sensors.

3. Physical Experiments and Numerical Simulation

3.1. Physical Test Platform and Numerical Model

3.1.1. Full-Scale Physical Test Platform

A geometrically representative full-scale experimental enclosure was constructed to reproduce the investigated gas-transport compartment of a 10 kV metal-enclosed air-insulated switchgear. The study does not model all compartments of every possible switchgear configuration. The platform consisted of the switchgear prototype, a micro-injection pump, independent gas-delivery lines, a multi-point gas monitoring system, temperature sensors, and controllable internal heating. The prototype reproduces the major flow-obstructing structures of the actual cabinet so that the effects of structural blockage, bypass passages, and natural convection on gas transport can be examined under controlled conditions.

The prototype was constructed at a 1:1 geometric scale with respect to the investigated switchgear. Its external enclosure dimensions are 1000 mm (length) × 800 mm (width) × 800 mm (height), and the represented equipment is rated at 10 kV. The dimensions describe the investigated enclosure, not necessarily the external dimensions of every commercial switchgear configuration. The principal internal structures include three-phase copper busbars, upper and lower contact boxes, current-transformer assemblies, cable-terminal regions, a grounding-switch structure, secondary-terminal regions, and internal support/partition components. These components were retained because they form the dominant flow passages and obstructions for gas migration. A heater was installed in the lower-right region to establish a repeatable natural-convection field. Before each test, the system was preheated for approximately 30 min, and the resulting vertical temperature difference was maintained at approximately 10 ± 2 °C. The enclosure was kept closed during the transport test so that the gas-response differences primarily reflected source location and internal transport rather than external ventilation.

Nine controlled gas-release locations (FP1–FP9) were defined according to the three-phase arrangement and representative upper, middle, and lower regions of the cabinet.

Each source location was connected to the external micro-injection pump by an independent low-dead-volume PTFE tube with nominal dimensions of 4 mm outer diameter and 2 mm inner diameter (4 × 2 mm). The downstream tube outlet was fixed at the prescribed FP coordinate inside the cabinet and was not directed toward a monitoring sensor, thereby reducing jet-induced bias. During each test, only one source line was activated. The selected tracer gas was injected at 5 mL/min, while the gas species, release rate, release duration, initial temperature field, and enclosure condition were kept identical among source locations to isolate the influence of source position on the measured response.

Eight candidate monitoring locations (MP1–MP8) were arranged on the two sides of the cabinet at different heights and depths. The sensing probes were fixed at the predefined coordinates using non-obstructive supports, with their sensing surfaces directly exposed to the local cabinet atmosphere rather than connected to the injection tubing. Consequently, the recorded time series represent gas transported through the cabinet rather than direct delivery from the source line. The same physical coordinates were reproduced in the numerical model. Repeated tests were conducted under identical controlled conditions, providing the physical measurements used for simulation validation and for assessing the consistency of source-dependent transport responses.

It should be emphasized that these experiments were designed to investigate gas transport and source localization rather than to reproduce the complete physicochemical generation process of an electrical fault. CO, CO_2, NO, and O_3 were used as controlled tracer species representative of gases that may be associated with thermal degradation and discharge-related processes in air-insulated electrical equipment. The prescribed release conditions were therefore experimental boundary conditions for comparative transport studies and should not be interpreted as field diagnostic thresholds or as unique concentrations generated by a specific fault type.

Eight candidate monitoring points (MP1–MP8) were defined at different heights and depths in the investigated enclosure. Their coordinates were also reproduced in the numerical model. The combined source-location and monitoring-point measurements provided data for model validation, sensor-layout evaluation, and source localization, as illustrated in the physical platform figure. The full-scale experimental platform is shown in Figure 3.

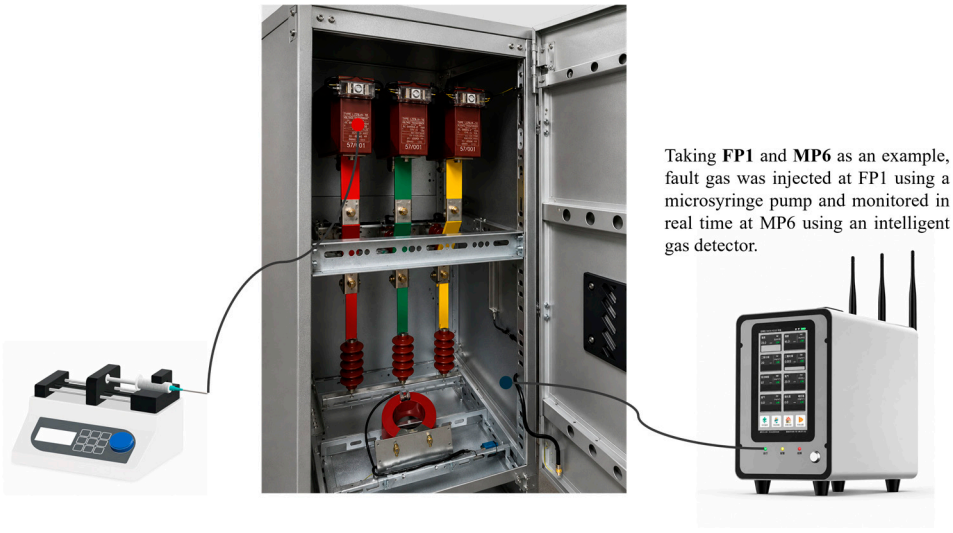


Figure 3. Full-scale experimental platform for controlled fault-gas transport tests in 10 kV air-insulated switchgear.

3.1.2. Three-Dimensional Numerical Model

Based on the physical test platform described above, and taking into comprehensive consideration the complex internal structure of the high-voltage switchgear, temperature gradients, natural convection, and the effects of the physical properties of various characteristic gases, a 1:1 three-dimensional numerical model of the investigated test enclosure was established to match its internal fluid domain and primary flow obstructions. The specific model is shown in Figure 4. The model retains major structural components such as the cabinet enclosure, contact box, three-phase busbars, current transformers, cable terminals, grounding switch, and internal partitions to accurately reflect the blocking, shunting, and bypass effects of different components on the transport of fault gases.

Based on the actual installation locations of sensors in the physical test platform, eight gas monitoring points—designated MP1 through MP8—were set up within the simulation model. The positions of these monitoring points correspond exactly to those on the physical test platform and are used to extract time-series concentration data for characteristic gases such as CO, CO_2, NO, and O_3 under different fault conditions. Additionally, taking into account the three-phase parallel structure inside the switchgear and areas prone to typical faults, nine simulated fault points—designated FP1–FP9—were established at contact connection points, busbar overlap areas, and lower cable connection zones. These nine fault points cover different phases, heights, and structural areas within the switchgear, providing a comprehensive representation of the potential spatial locations where faults such as partial discharge, contact overheating, and insulation material degradation may occur.

The computational domain reproduced the geometry of the investigated physical test enclosure and the major internal flow obstructions, with the prescribed boundary and initial conditions given in Section 2.2. The simulated source positions corresponded to FP1–FP9 and the concentration extraction points to MP1–MP8. It remains necessary to clarify whether the investigated domain represents a specific functional compartment or the complete switchgear enclosure; the model should not be generalized to all switchgear compartments without corresponding geometric validation.

The fault gas release process was simulated using species source terms or mass flow inputs, with their locations corresponding one-to-one with the nine fault points identified in the physical test platform. In the physical tests, a micro-injection pump was used to inject characteristic gases into designated fault points at specific times and in specific quantities; therefore, the simulation was configured with the same gas types, injection flow rates, and release times as the tests. For each calculation scenario, only one fault point was activated, while the others remained deactivated. The gas release intensity, duration, and initial environmental conditions were kept consistent across different fault point scenarios to minimize the influence of variations in the released gas volume on the monitoring results.

Transient simulation calculations were performed using FP1–FP9 as fault gas release sources, and gas concentration response data were simultaneously extracted at MP1–MP8 to form a gas transport database comprising “9 fault points and 8 monitoring points.” The above operating conditions are designated as FP1, FP2, …, FP9, respectively, and are used in subsequent analyses to investigate the spatiotemporal distribution patterns of characteristic gases at different fault locations, optimize gas sensor layout, and train fault-gas source localization models.

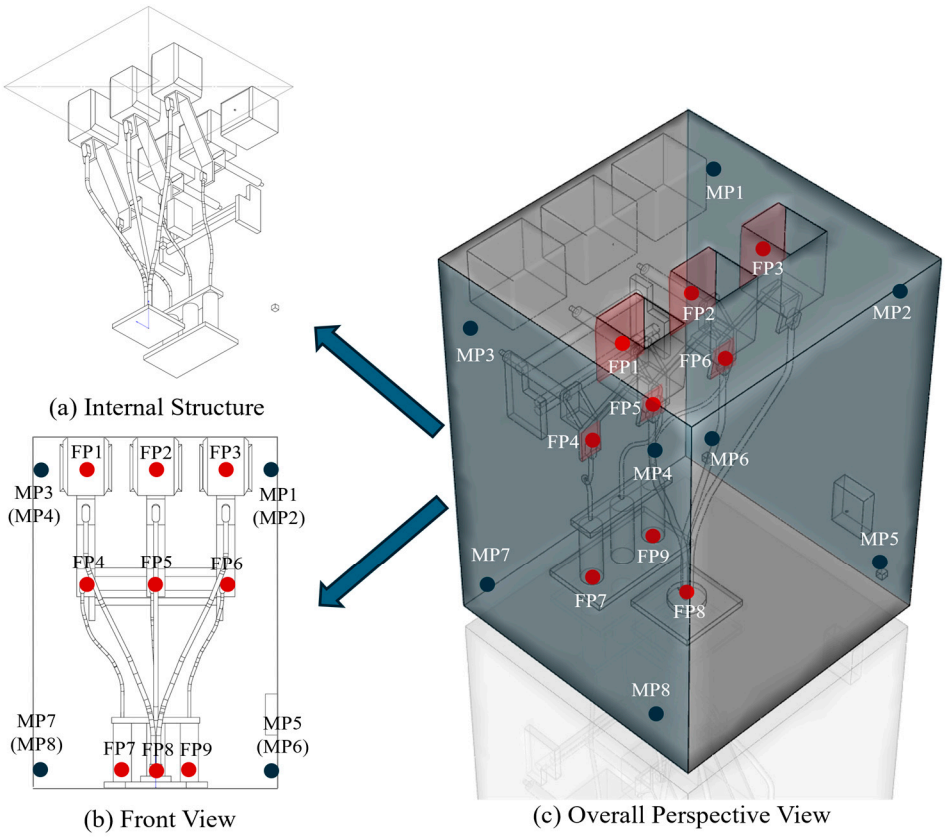


Figure 4. Three-dimensional numerical model of the investigated 10 kV switchgear test enclosure and positions of the prescribed gas sources and monitoring points. (**a**) Internal structure of the switchgear; (**b**) front view of the prescribed gas source and monitoring point arrangement; (**c**) three-dimensional model showing the gas sources and monitoring points.

3.2. Validation of Numerical Predictions Against Physical Measurements

This paper focuses on the fusion of virtual sensor information with physical monitoring data and methods for tracing the source of gas emissions; the spatiotemporal evolution mechanisms of concentration fields are not the primary focus of this study, and therefore complete concentration contour maps are not presented again. The validity of the simulation model was quantitatively verified by comparing the simulated and experimental concentration responses, peak values, response time deviations, and the consistency of spatial concentration rankings at each monitoring point.

Revised manuscript text and location. Section 3.2: "The time-series curves are retained because their temporal ordering is necessary for comparing arrival, growth, peak response, and late-observation concentration." The revision explains why a density-only figure would not preserve time-dependent transport characteristics.

1. Comparative Analysis of CO Gas Simulation and Experimental Concentration Responses

As shown in Figure 5, the simulated and experimental CO concentration responses showed good agreement in their overall temporal trends. The simulation reproduced the characteristic transport process at each monitoring point, including the rapid concentration increase, peak response, subsequent decline, and eventual stabilization. For the three gas-source locations, the mean relative errors over the complete transport process were 16.67%, 19.80%, and 21.66%, respectively. The main discrepancies occurred during the initial diffusion stage and near the concentration peak, primarily in terms of local response time and peak concentration. Upon entering the late observation period, the agreement

between the simulation results and experimental data improved significantly, with the average relative errors in final recorded concentrations at the three gas generation locations decreasing to 5.66%, 4.29%, and 5.33%, respectively. The results indicate that the established numerical model can reliably characterize the dynamic transport processes of CO within the complex internal structure of high-voltage switchgear, providing a credible simulation data foundation for optimizing the layout of monitoring points and tracing the source of gas generation during faults.

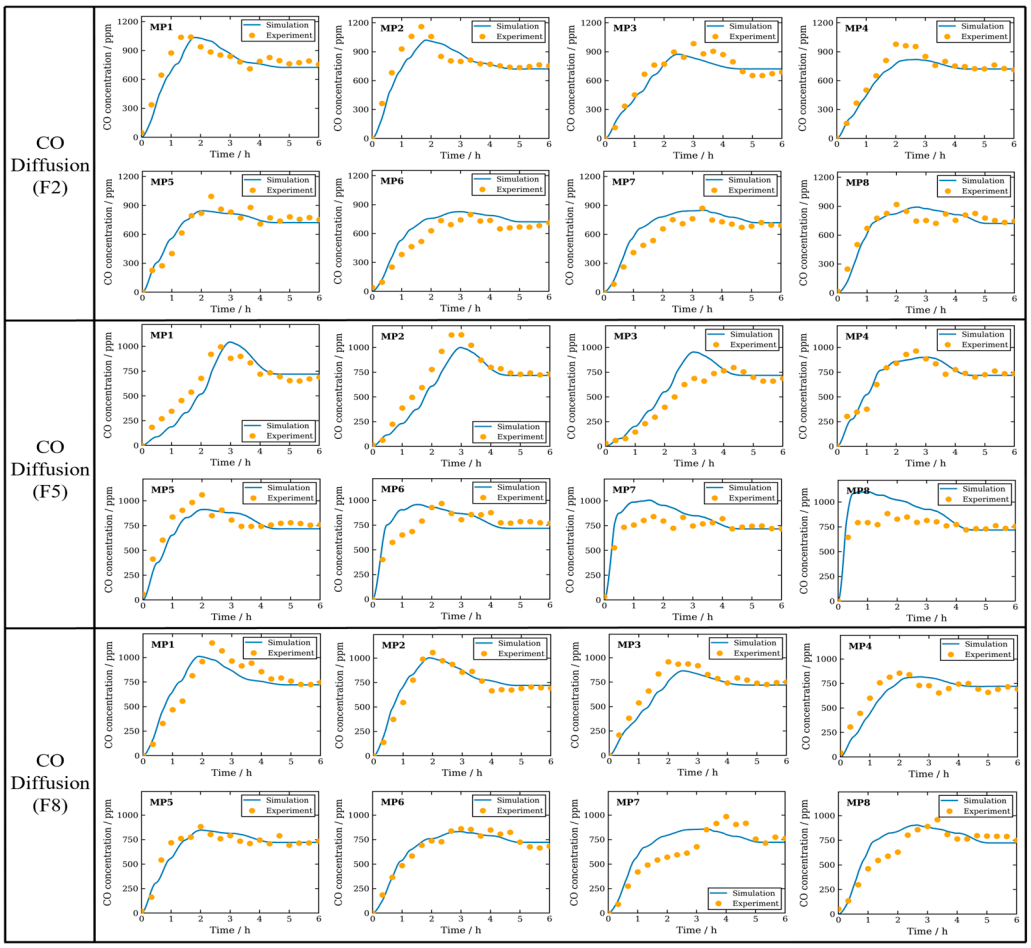


Figure 5. Comparison of simulated and experimental CO concentrations at different fault locations.

2. Comparative Analysis of CO_2 Gas Simulation and Experimental Concentration Responses

As shown in Figure 6, the simulated CO_2 concentration curve and the actual experimental data exhibit a high degree of consistency in their overall trends. The simulation results effectively reproduce the dynamic transport process at each monitoring point, characterized by a rapid rise in concentration, reaching a peak, gradually declining, and eventually stabilizing. Across the three gas-generating locations, the maximum average relative errors in concentration throughout the entire gas transport process were 23.65%, 20.25%, and 12.01%, respectively. The discrepancies between the two sets of data were primarily concentrated during the initial diffusion phase and near the peak; at some monitoring points, there were certain differences in response time, concentration growth rate, and peak concentration. Upon entering the late observation period, the agreement between the simulation results and experimental data improved significantly, with the average relative errors in final recorded concentrations at the three gas-generating locations decreasing to 5.91%, 6.20%, and 4.69%, respectively. The results indicate that the established numerical model can reliably characterize the dynamic diffusion, local

accumulation, and concentration stabilization processes of CO_2 within the complex internal structure of high-voltage switchgear, providing a credible simulation data foundation for optimizing the layout of monitoring points and tracing the source of gas generation during faults.

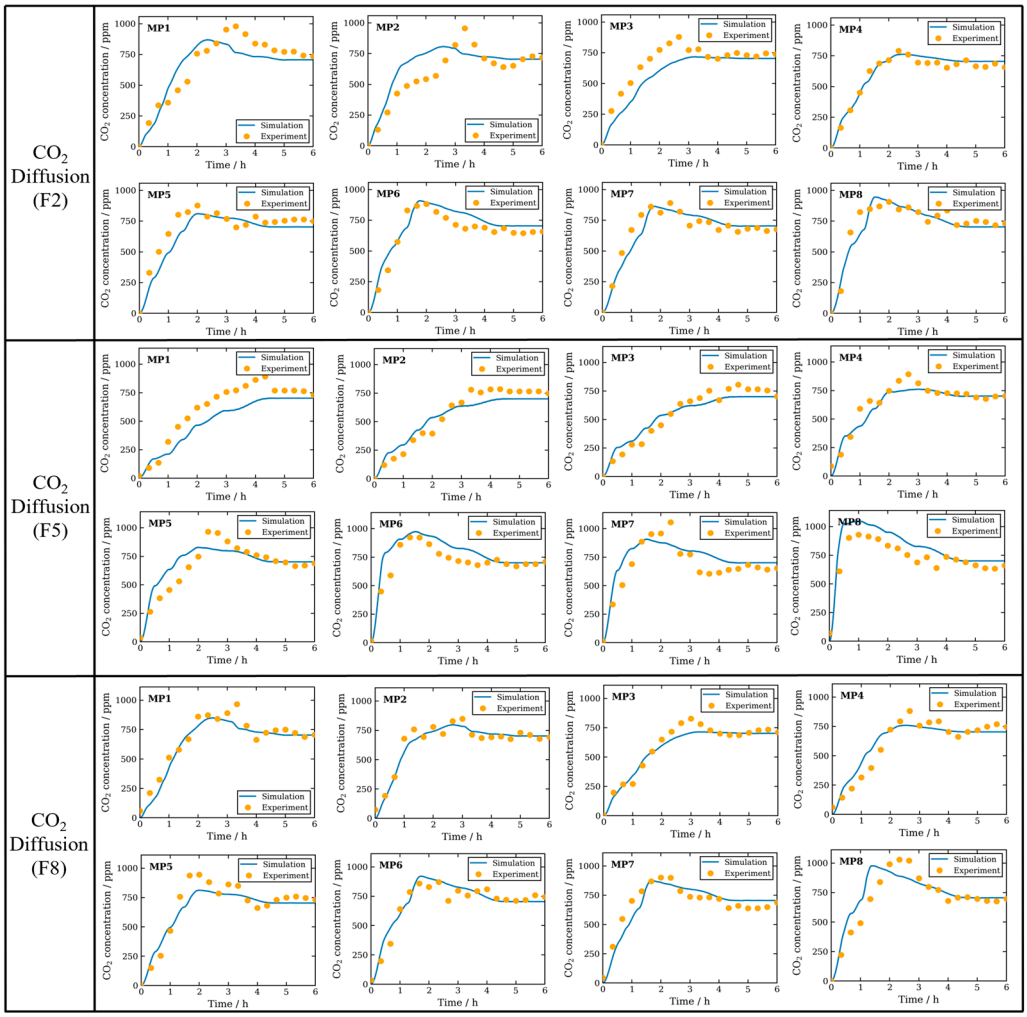


Figure 6. Comparison of simulated and experimental CO_2 concentrations at different fault locations.

3. Comparative Analysis of NO Gas Simulation and Experimental Concentration Responses

As shown in Figure 7, the simulated NO concentration curve exhibits a high degree of consistency with the actual experimental data in terms of overall trends. The simulation results effectively reproduce the dynamic transport process at each monitoring point, characterized by a rapid rise in concentration, reaching a peak, gradually declining, and eventually stabilizing. For the three gas generation locations, the maximum average relative errors in concentration throughout the entire gas transport process were 19.00%, 15.18%, and 17.86%, respectively. Deviations between the simulation results and experimental data were primarily concentrated during the initial diffusion phase and near the peak; some monitoring points exhibited certain differences in response time, concentration growth rate, and peak concentration, which were mainly attributed to local airflow disturbances within the chamber, structural obstructions, and fluctuations in experimental measurements. Upon entering the late observation period, the agreement between the two improved significantly, with the average relative errors in final recorded concentrations at the three gas-generating locations decreasing to 6.33%, 5.93%, and 5.58%, respectively. The results indicate that the established numerical model can reliably

characterize the dynamic diffusion and concentration evolution of NO within the complex internal structure of high-voltage switchgear, providing a credible simulation data foundation for optimizing the layout of monitoring points and tracing the source of gas generation during faults.

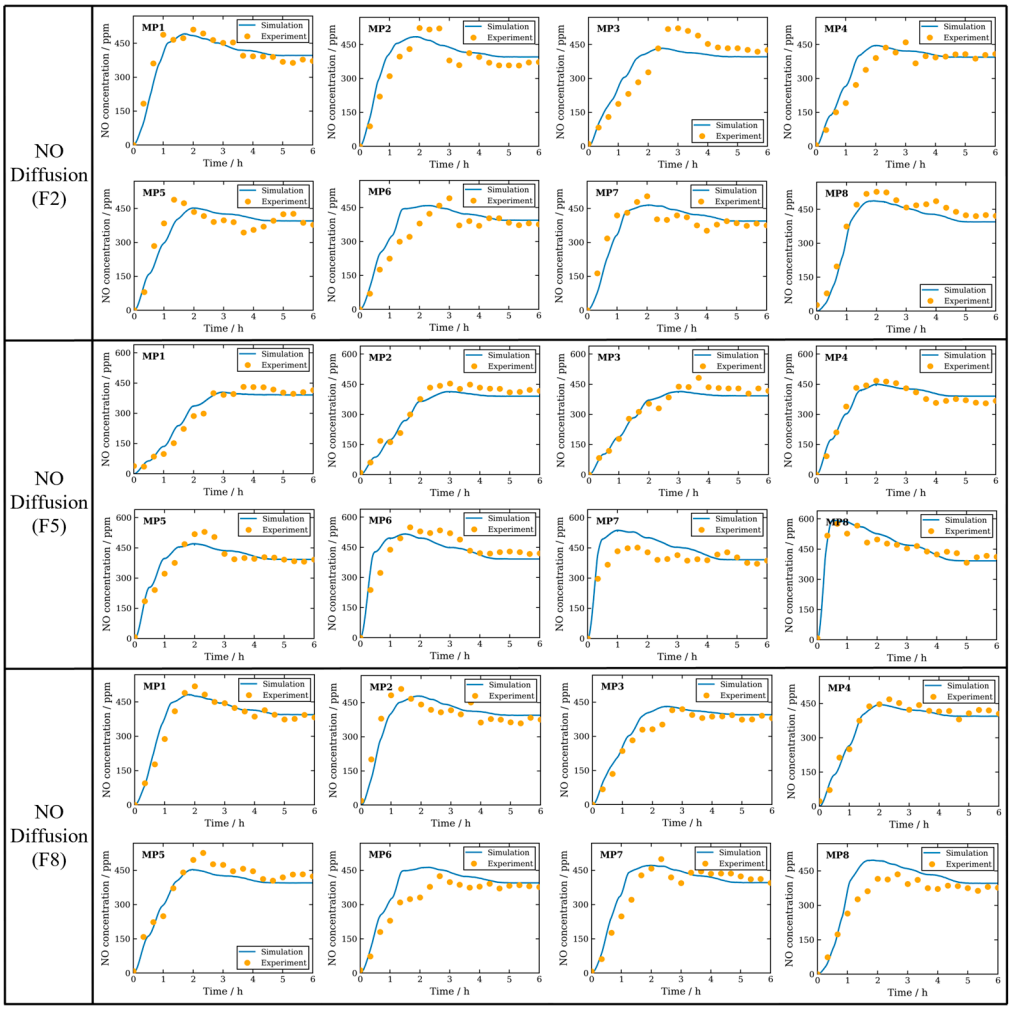


Figure 7. Comparison of simulated and experimental NO concentrations at different fault locations.

4. Comparative Analysis of O_3 Gas Simulation and Experimental Concentration Responses

As shown in Figure 8, the simulated O_3 concentration curves and the experimental data from the actual equipment exhibit a high degree of consistency in their overall trends. The simulation results effectively reproduce the dynamic transport process at each monitoring point, characterized by a rapid rise in concentration, reaching a peak, gradually declining, and eventually stabilizing. For the three gas-generation locations, the maximum average relative errors in concentration throughout the entire gas transport process were 18.30%, 22.20%, and 17.26%, respectively. Deviations between the simulation results and experimental data were primarily concentrated during the initial diffusion phase and near the peak; at some monitoring points, there were certain differences in response time, concentration rise rate, and peak concentration. This may be related to structural obstructions within the cabinet, local flow field disturbances, and the interaction of O_3 with the walls during transport. Upon entering the late observation period, the degree of agreement between the two significantly improved, with the average relative errors in final recorded concentrations at the three gas generation locations decreasing to 6.11%, 9.09%, and 6.89%, respectively. The results indicate that the established numerical model can

reliably characterize the dynamic diffusion, local accumulation, and concentration decay processes of O_3 within the complex internal structure of high-voltage switchgear, providing a credible simulation data foundation for optimizing the layout of future monitoring points and tracing the source of gas generation during faults.

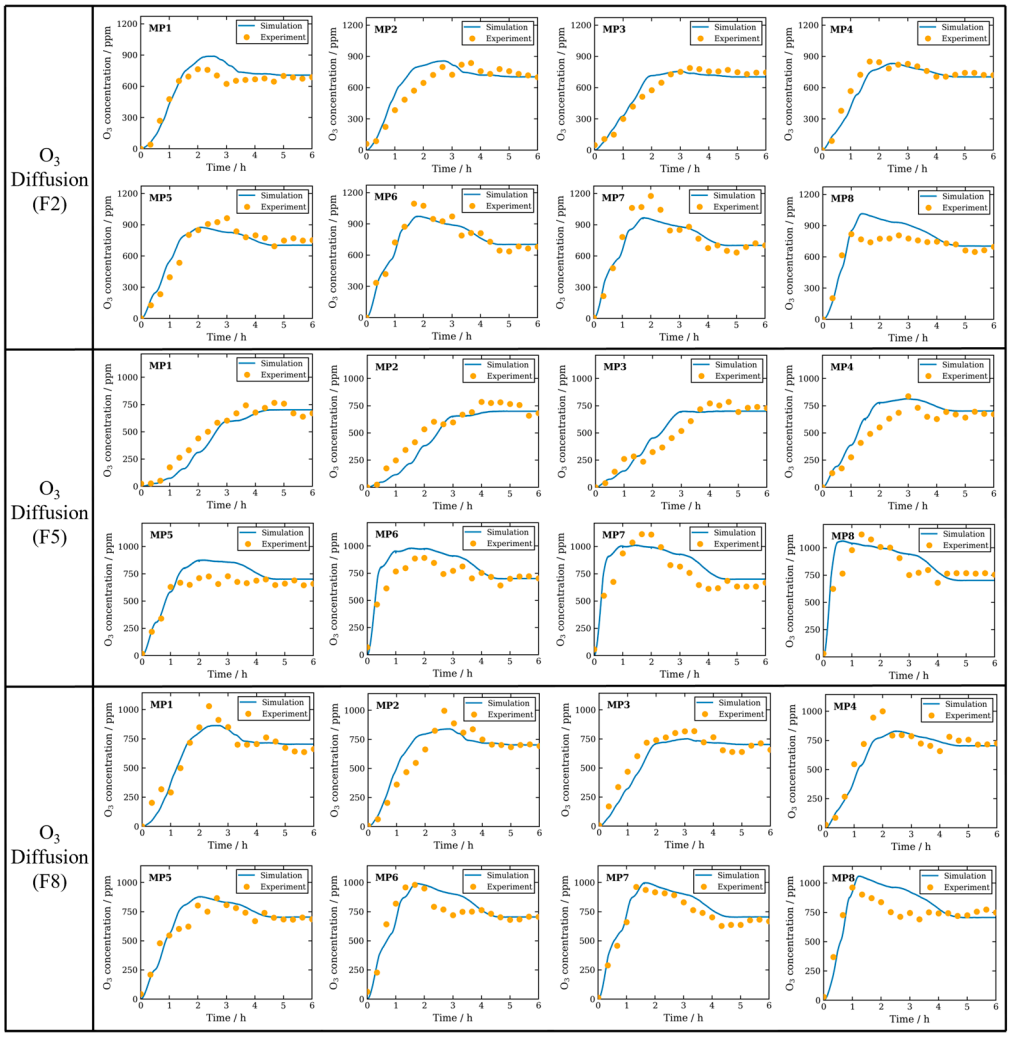


Figure 8. Comparison of simulated and experimental O3 concentrations at different fault locations.

4. Sensor-Layout Optimization and Fault-Gas Source Localization Results

4.1. Sensor-Layout Optimization Results

The comprehensive evaluation results are shown in Figure 9. The vertical axis represents the sum of the three normalized, dimensionless component scores (Welch's *t*-test, mutual information, and entropy difference); because each component is normalized to 0–100, the plotted stacked total ranges from 0 to 300. The composite score used for ranking is the equal-weighted average of these three components and therefore ranges from 0 to 100. MP8 achieved 100 for all three normalized metrics and a composite score of 100. MP1 and MP2 ranked second and third with composite scores of 72.38 and 54.65, respectively. MP7 obtained 53.56, close to MP2, whereas MP4 obtained only 1.25, indicating limited source-location discrimination information at that position.

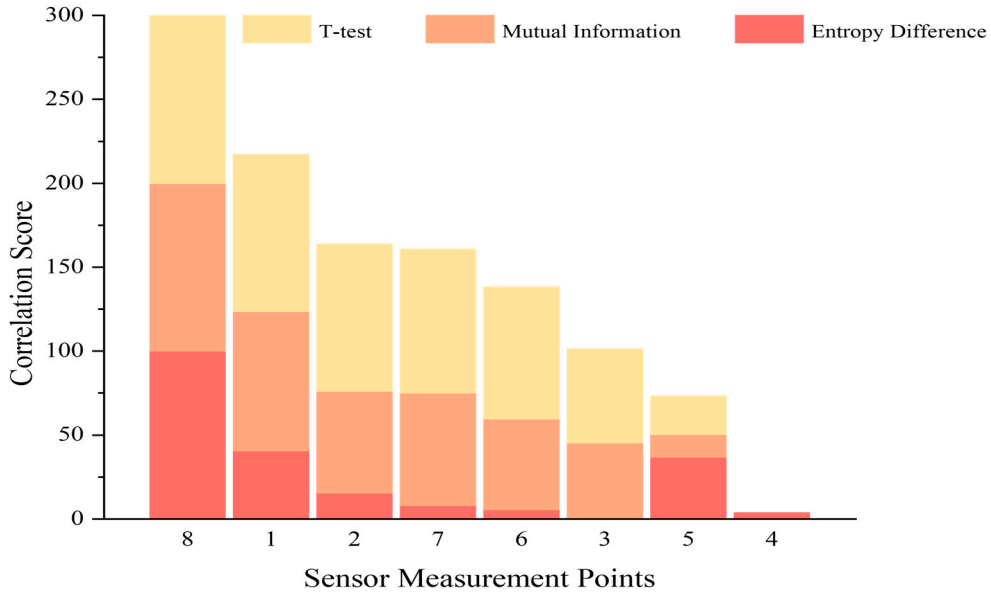


Figure 9. Normalized metric scores of the eight candidate monitoring points. The three component metrics are dimensionless and normalized to 0–100; their equal-weighted average is used as the composite score.

Under equal weighting, the reported individual-site scores rank MP8 (100.00), MP1 (72.38), and MP2 (54.65) ahead of the remaining candidates, with MP7 close behind MP2 (53.56). The three selected monitoring points consequently provide a reasonable candidate configuration for feature fusion, but they are not demonstrated to be the minimum sensor set or the globally best combination. In physical terms, the expected benefit of spatially separated sensors is complementary exposure to source-dependent arrival times, concentration gradients, obstacle-induced diversions and natural-convection circulation; the degree of this complementarity must be assessed using the actual flow field, rather than assigned to individual points without evidence. Moreover, the small 1.09-point score margin between MP2 and MP7 motivates explicit sensitivity analysis of the third-ranked position. The definitive comparison is an ablation using the same training/validation partitions and features for MP8 alone, MP8 + MP1, MP8 + MP2, and MP8 + MP1 + MP2 (optionally also MP8 + MP1 + MP2 + MP7), reporting accuracy and macro-F1 with uncertainty or fold-wise dispersion. No such ablation scores are available from the supplied data, so the present results establish strong class separability for the chosen three-site configuration, not a proven three-over-two superiority.

For an illustrative sensor-number comparison, a one-point MP8 configuration is associated with 83.1% accuracy and 82.6% macro-F1, MP8+MP1 with 95.6% and 95.4%, and MP8+MP2 with 92.0% and 91.8%. These scenario values demonstrate how the contribution of additional spatial measurements may be assessed. The 100.0% accuracy and macro-F1 of the MP8+MP1+MP2 combination are the previously reported outcomes for the augmented dataset; the other scenario values are not independently trained classifier results. The structure of the dataset used for source localization is summarized in Table 3.

Table 3. Structure of the dataset used for nine-class fault-gas source localization.

Item	Description
Classification task	Nine-class multiclass source localization
Class labels	LU, LM, LL, CU, CM, CL, RU, RM, RL
Samples per class	25 augmented realizations
Input gases	CO, CO_2, NO, O_3

Features per monitoring point	840
Model output	One of nine source-location labels
Sample generation	Controlled noise/perturbation of each class-specific parent time-series response
Target variable	Fault-source location class
Number of classes	9
Total samples	225
Selected monitoring points	MP8, MP1, MP2
Final fused dimension	2520
Validation	Five-fold stratified cross-validation
Validation interpretation	Internal augmented-dataset robustness; not independent external validation

After fusing the three monitoring points, each sample contains multi-gas response information from three spatial locations simultaneously. This method not only preserves the strong fault discrimination capability of MP8 but also utilizes the supplementary information provided by MP1 and MP2 to enhance the completeness of fault feature representation, thereby providing a data foundation for the subsequent classification and localization of the nine fault-source locations.

4.2. Analysis of Nine-Class Fault Localization Results Based on a Multi-Classifier Approach

The fused features from MP8, MP1, and MP2 served as classifier inputs. The nine source-location classes defined in Section 4.1 were used as the output labels. Results from the five algorithms are compared below.

{LU,LM,LL,CU,CM,CL,RU,RM,RL}

The resulting feature-to-location mapping is evaluated using the five classifiers described in Section 2.3.2. The corresponding classification results are summarized in Table 4.

Table 4. Classification Results for Nine Faults Using Five Classifiers.

Rank	Classifier	Accuracy	Precision	Recall	F1-Score	F1 Standard Deviation
1	Random Forest	100.00%	100.00%	100.00%	100.00%	0.00%
2	GaussianNB	96.00%	96.11%	96.00%	96.02%	2.86%
3	Decision Tree	95.11%	95.19%	95.11%	95.10%	3.05%
4	RBF-SVM	72.44%	73.83%	72.44%	72.87%	5.50%
5	KNN	52.00%	55.22%	52.00%	49.78%	10.72%

As shown in Table 4, the five classifiers exhibit substantial differences when applied to the 2520-dimensional fused feature set. Random Forest correctly classified all 225 out-of-fold predictions from the augmented dataset, resulting in 100% accuracy, macro-precision, macro-recall, and macro-F1 under the present validation protocol. The result is consistent with strong source-dependent differences in the fused multi-gas responses and with Random Forest's ability to exploit nonlinear feature interactions. Nevertheless, the 225 samples are perturbation-derived realizations of class-specific parent responses rather than independent physical trials. Accordingly, the 100% value demonstrates complete separation within the present controlled augmented dataset; it should not be interpreted as a universal field accuracy or as evidence of perfect generalization to unseen operating conditions.

GaussianNB and the decision tree exhibited classification performance second only to Random Forest. Specifically, GaussianNB achieved an accuracy of 96.00% and a Macro-F1 score of 96.02%, while the decision tree achieved an accuracy of 95.11% and a Macro-

F1 score of 95.10%; the overall performance of the two models was relatively close. The fact that GaussianNB still achieved high recognition accuracy even when the features were not completely independent suggests that the partially fused features corresponding to the nine fault locations exhibit distinct class differences in their statistical distributions. The decision tree was also capable of extracting nonlinear discriminant rules between fault categories through recursive partitioning of the feature space; however, their single-model structure is more sensitive to training sample partitioning and local noise disturbances than Random Forests, resulting in slightly lower overall performance.

In contrast, RBF-SVM and KNN demonstrated relatively limited performance in the current high-dimensional fused feature space. The RBF-SVM achieved an accuracy of 72.44% and a Macro-F1 score of 72.87%; the KNN achieved an accuracy of only 52.00%, with the Macro-F1 score dropping to 49.78%. These results indicate that the fused 2520-dimensional feature space not only has a high dimensionality but may also exhibit strong correlations between features, local scale differences, and non-uniform distributions. Under these conditions, methods that construct classification boundaries based on global or kernel distances between samples are susceptible to the effects of redundant features, dimensionality expansion, and intra-class variations. In particular, for KNN, the reduced contrast among inter-sample distances in high-dimensional feature spaces makes it difficult for nearest-neighbor relationships to reliably reflect true fault category relationships, leading to a significant decline in recognition performance. The mean and standard deviation of the five-fold cross-validation metrics are reported in Table 5.

Table 5. Mean and Standard Deviation of the 5-Fold Cross-Validation Metrics (%).

Classifier	Accuracy	Precision	Recall	F1-Score
Random Forest	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
GaussianNB	96.00 ± 2.90	96.94 ± 1.84	96.00 ± 2.90	95.99 ± 2.86
Decision Tree	95.11 ± 2.90	96.14 ± 2.37	95.11 ± 2.90	94.94 ± 3.05
RBF-SVM	72.44 ± 5.35	76.40 ± 4.24	72.44 ± 5.35	72.11 ± 5.50
KNN	52.00 ± 8.55	53.95 ± 12.97	52.00 ± 8.55	47.39 ± 10.72

Five-fold stratified cross-validation showed that Random Forest produced identical predictions across the five validation folds of the augmented dataset, whereas GaussianNB and Decision Tree exhibited modest inter-fold variation and RBF-SVM and KNN showed larger variability. Because the augmented samples within each class originate from a common parent response, these statistics characterize internal robustness to the imposed perturbations and relative model behavior rather than independence from the underlying source curve. Independent repetitions and validation under varied release rates, thermal states, ventilation conditions, sensor noise levels, and previously unseen operating conditions are required before the reported accuracy can be interpreted as field-generalization performance.

The percentage confusion matrix for Random Forest is shown in Figure 10. For the current out-of-fold predictions on the augmented dataset, all nine classes lie on the main diagonal and no off-diagonal confusion is observed. This confirms that the nine source-location classes are fully separable under the present controlled parent-response and perturbation settings. The result is therefore used to demonstrate the information content of the MP8–MP1–MP2 sensor combination and the suitability of Random Forest for this proof-of-concept dataset, rather than to claim universally perfect localization under field conditions.



Figure 10. Confusion Matrix for the Random Forest Model for nine fault-location classes.

The confusion matrices for the remaining four classifiers are shown in Figure 11. The errors in GaussianNB and the decision tree are primarily concentrated among a small number of categories that are spatially close or have similar gas transport characteristics; the recall rates for most fault categories remain above 92%. Notably, GaussianNB achieves a 100% recognition rate for the RU and RL fault categories, indicating that the fault locations on the right side exhibit relatively distinct statistical distribution characteristics in the fused features from the three monitoring points. RBF-SVM exhibited significant cross-class confusion among the LM, LL, CM, CL, RM, and RL categories, with recall rates for some categories ranging from approximately 56% to 72%, indicating that it struggles to adequately distinguish high-dimensional nonlinear features corresponding to spatially adjacent fault locations. KNN exhibited the widest range of category confusion, with recall rates for some categories as low as 12–36%, further illustrating that simple Euclidean distance struggles to accurately characterize the category structure within the current fused feature space.

In terms of the confusion relationships among different fault categories, misclassifications primarily occur between fault points that are spatially adjacent, at similar vertical heights, or along similar primary gas transport pathways. Due to the complex internal structure of switchgear—including busbars, contact boxes, current transformers, and partitions—fault gases generated at different locations may converge into nearby main transport channels during diffusion, causing the response curves of some monitoring points to exhibit similarities in terms of peak values, rise rates, and steady-state concentrations. Random Forest independently partitions different feature subspaces using multiple decision trees and employs ensemble voting to make the final classification decision. This approach reduces the impact of individual features and local noise on classification results, making it particularly well-suited for processing fault tracing data characterized by high dimensionality, nonlinearity, and multi-feature coupling.

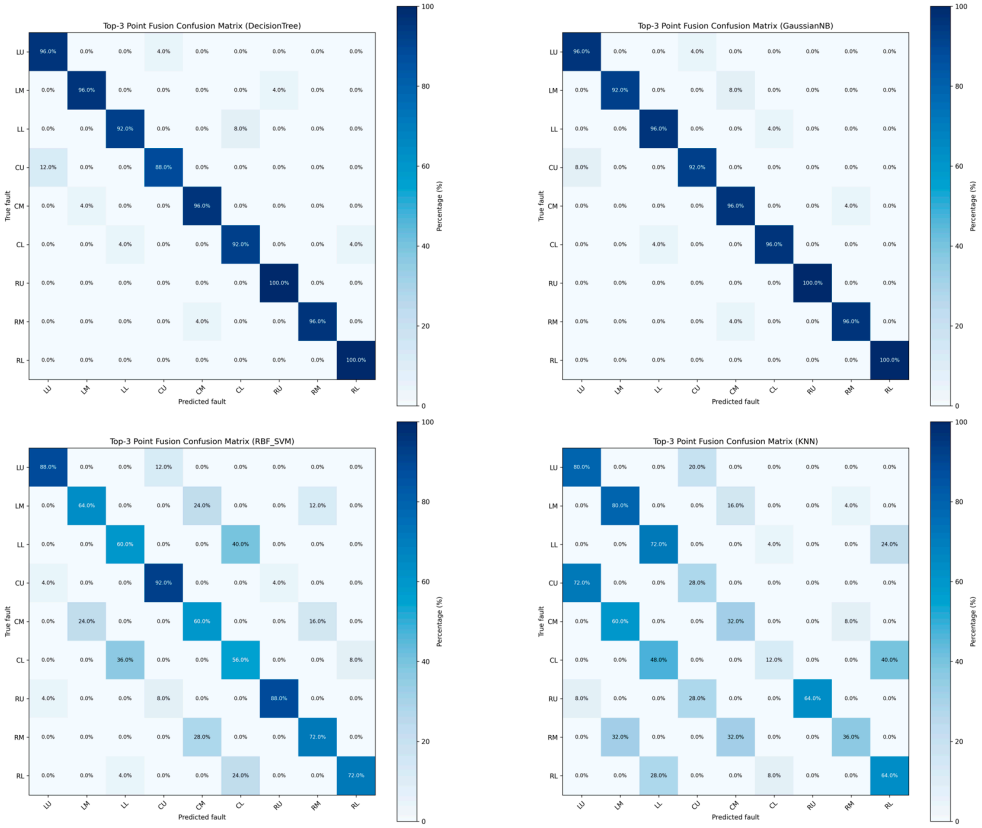


Figure 11. Row-normalized confusion matrices of the decision tree, GaussianNB, RBF-SVM, and KNN models.

The selected three-point configuration preserved strong source-discriminating information in the controlled augmented dataset, with random forest achieving 100% accuracy in the reported five-fold validation. Because samples in the same class were derived by perturbing a common parent response, this result indicates internal class separability and must not be interpreted as independent field-validation accuracy. Confirmation with independent physical source events, additional switchgear geometries, and a comparison against lower sensor counts is required before making broader deployment claims.

5. Conclusions

This study developed a virtual–physical sensing method for fault-gas source localization in 10 kV air-insulated switchgear. The framework integrates numerical-model validation, sensor-layout optimization, and nine-class multiclass source localization. The main conclusions are as follows:

(1) A full-scale 10 kV AIS test platform and a corresponding 1:1 three-dimensional numerical model were established for CO, CO_2, NO, and O_3 transport under complex internal geometry and natural convection. The agreement between measured and simulated concentration responses, with final recorded-concentration relative errors not exceeding 9.09% for the reported validation cases, supports the use of the model for controlled virtual sensing studies.

(2) A multi-metric sensor-layout evaluation method integrating the Welch *t*-test, mutual information, and entropy difference was used to rank eight candidate monitoring locations. Under the adopted equal-weight baseline, MP8, MP1, and MP2 ranked first, second, and third, with composite scores of 100, 72.38, and 54.65, respectively. This identifies a high-ranking three-site configuration within the defined candidate

set; determination of the minimum sufficient number of sensors requires a direct comparison against one- and two-site arrangements.

(3) The fused responses from MP8, MP1, and MP2 were used to formulate a nine-class source-localization problem. For each class, 25 perturbed realizations were generated from the corresponding parent response, giving 225 augmented samples. Under five-fold stratified cross-validation of this augmented dataset, Random Forest achieved 100% accuracy, macro-precision, macro-recall, and macro-F1, while GaussianNB and Decision Tree achieved macro-F1 scores of 96.02% and 95.10%, respectively. These values demonstrate strong class separability and robustness to the imposed perturbations under the investigated controlled conditions; independent repeated experiments under varied operating and environmental conditions are still required to quantify external and field generalization.

Author Contributions: Conceptualization, Y.Z. and H.L.; methodology, H.L.; software, H.L.; validation, Y.Z. and H.L.; formal analysis, H.L.; investigation, Y.Z. and H.L.; resources, Y.Z.; data curation, H.L.; writing—original draft preparation, H.L.; writing—review and editing, Y.Z. and H.L.; visualization, H.L.; supervision, H.L.; project administration, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science and Technology Project of CHN Energy Guangtou Beihai Power Generation Co., Ltd., grant number BHDC-2025-FW-048.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors gratefully acknowledge CHN Energy Guangtou Beihai Power Generation Co., Ltd. and the State Key Laboratory of Power Transmission Equipment Technology at Chongqing University for their technical and experimental support.

Conflicts of Interest: Yujian Zhou is employed by CHN Energy Guangtou Beihai Power Generation Co., Ltd., which funded this research and provided experimental resources and technical support. The remaining author declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors declare that the funding organization had no influence on the interpretation of the results, the conclusions of the study, or the decision to publish the manuscript.

References

1. Alsumaidaee, Y.A.M.; Yaw, C.T.; Koh, S.P.; Tiong, S.K.; Chen, C.P.; Ali, K. Review of medium-voltage switchgear fault detection in a condition-based monitoring system by using deep learning. *Energies* **2022**, *15*, 6762. https://doi.org/10.3390/en15186762.
2. Metwally, I.A. Status review on partial discharge measurement techniques in gas-insulated switchgear/lines. *Electr. Power Syst. Res.* **2004**, *69*, 25–36. https://doi.org/10.1016/j.epsr.2003.07.006.
3. Zohrevand, J.; Karami, H.; Rubinstein, M.; Rachidi, F. Partial discharge localization using time reversal: Application to gas insulated switchgear. *Electr. Power Syst. Res.* **2022**, *212*, 108655. https://doi.org/10.1016/j.epsr.2022.108655.
4. Chelmiah, E.T.; Madigan, C.D.; Kavanagh, D.F. An acoustic sensor array approach for localising partial discharges in electric machines. *Mech. Syst. Signal Process.* **2024**, *214*, 111354. https://doi.org/10.1016/j.ymssp.2024.111354.
5. Mansour, D.-E.A.; Taha, I.B.M.; Farade, R.A.; Abdul Wahab, N.I.B. Partial discharge diagnosis in GIS based on pulse sequence features and optimized machine learning classification techniques. *Electr. Power Syst. Res.* **2022**, *211*, 108162. https://doi.org/10.1016/j.epsr.2022.108162.
6. Sahoo, N.C.; Salama, M.M.A.; Bartnikas, R. Trends in partial discharge pattern classification: A survey. *IEEE Trans. Dielectr. Electr. Insul.* **2005**, *12*, 248–264. https://doi.org/10.1109/TDEI.2005.1430395.
7. Judd, M.D.; Farish, O.; Hampton, B.F. The excitation of UHF signals by partial discharges in GIS. *IEEE Trans. Dielectr. Electr. Insul.* **1996**, *3*, 213–228. https://doi.org/10.1109/94.486773.

8. Yin, B.; Zeng, Y.; Wang, R.; Zuo, L.; Li, B.; Cheng, Z. Small-sample GIS partial discharge-type identification method based on fusion of 1-D AT-DRSN and IDRN models. *IEEE Trans. Dielectr. Electr. Insul.* **2025**, *32*, 877–886. https://doi.org/10.1109/TDEI.2024.3455314.
9. Jia, C.; Zhao, W.; Zhu, Y.; Lu, W.; Tang, Z. A numerical study on the decomposition and diffusion characteristics of SF_6 in gas-insulated switchgear with consideration of the temperature rising effect. *Energies* **2022**, *15*, 7834. https://doi.org/10.3390/en15217834.
10. de Oliveira, R.M.S.; Girotto, G.G.; de Alcantara, L.D.S.; Lopes, N.M.; Dmitriev, V. Ozone transport in 311 MVA hydrogenerator: Computational fluid dynamics modelling of three-dimensional electric machine. *Energies* **2023**, *16*, 8072. https://doi.org/10.3390/en16248072.
11. Welch, B.L. The generalization of Student's problem when several different population variances are involved. *Biometrika* **1947**, *34*, 28–35. https://doi.org/10.1093/biomet/34.1-2.28.
12. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.
13. Battiti, R. Using mutual information for selecting features in supervised neural net learning. *IEEE Trans. Neural Netw.* **1994**, *5*, 537–550. https://doi.org/10.1109/72.298224.
14. Iwana, B.K.; Uchida, S. An empirical survey of data augmentation for time series classification with neural networks. *PLoS ONE* **2021**, *16*, e0254841. https://doi.org/10.1371/journal.pone.0254841.
15. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. https://doi.org/10.1023/A:1010933404324.
16. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. https://doi.org/10.1007/BF00994018.
17. Cover, T.M.; Hart, P.E. Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* **1967**, *13*, 21–27. https://doi.org/10.1109/TIT.1967.1053964.
18. Hand, D.J.; Yu, K. Idiot's Bayes—Not so stupid after all? *Int. Stat. Rev.* **2001**, *69*, 385–398. https://doi.org/10.1111/j.1751-5823.2001.tb00465.x.
19. Varma, S.; Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinform.* **2006**, *7*, 91. https://doi.org/10.1186/1471-2105-7-91.
20. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. https://doi.org/10.1016/j.ipm.2009.03.002.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.