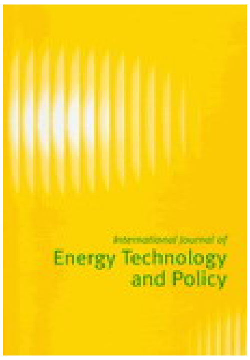


International Journal of Energy Technology and Policy

ISSN online: 1741-508X - ISSN print: 1472-8923
https://www.inderscience.com/ijetp

Enabling holistic insulation monitoring in secondary AC protection circuits: a diagnostic algorithm

Hengchu Shi, Hao You, Xiaofan Chen, Shoudong Xu, Jing Chen

DOI: 10.1504/IJETP.2026.10077374

Article History:

Received:	18 December 2025
Last revised:	20 January 2026
Accepted:	02 February 2026
Published online:	03 August 2026

Enabling holistic insulation monitoring in secondary AC protection circuits: a diagnostic algorithm

Hengchu Shi*, Hao You and Xiaofan Chen
Yunnan Power Grid Co., Ltd.,
No. 73, Tuodong Road, Kunming 650011, Yunnan, China
Email: shihengchu0911@126.com
Email: 403168654@qq.com
Email: 624614593@qq.com
*Corresponding author

Shoudong Xu
Electric Power Research Institute of Yunnan Power Grid Co., Ltd.,
No. 105, Yundaxi Road, Kunming 650217, Yunnan, China
Email: 741698897@qq.com

Jing Chen
Yunnan Power Grid Co., Ltd.,
No. 73, Tuodong Road, Kunming 650011, Yunnan, China
Email: 2413856587@qq.com

Abstract: This paper addresses the problem of insulation status monitoring in secondary alternating current (AC) circuits of power system relay protection. A comprehensive diagnostic algorithm based on multi-dimensional feature fusion is proposed. This algorithm integrates distributed zero-sequence current monitoring, waveform similarity analysis, and third harmonic analysis techniques to construct a multi-dimensional feature fusion diagnostic model. By real-time acquisition of zero-sequence current in each branch, combined with the similarity changes of current waveforms under normal and abnormal operating conditions and the characteristics of third harmonic content, early identification and accurate location of typical faults such as insulation degradation, single-point grounding, and multi-point grounding are achieved. Experimental results show that the algorithm achieves a 98.6% detection rate for insulation faults, a 94.3% accuracy rate for multi-point grounding location, an average diagnostic latency of only 43.63 ms, and a false alarm rate of 1.5% under normal operating conditions. This significantly improves the timeliness and accuracy of insulation fault diagnosis, providing effective technical support for the safe and reliable operation of secondary circuits in substation relay protection.

Keywords: secondary circuit; insulation status; zero-sequence current; waveform similarity analysis; harmonic analysis; comprehensive diagnosis.

Reference to this paper should be made as follows: Shi, H., You, H., Chen, X., Xu, S. and Chen, J. (2026) ‘Enabling holistic insulation monitoring in secondary AC protection circuits: a diagnostic algorithm’, *Int. J. Energy Technology and Policy*, Vol. 21, No. 5, pp.20–49.

Biographical notes: Hengchu Shi is a Professor Level Senior Engineer in Yunnan Power Grid Co., Ltd. He received his Bachelor's degree from the North China Electric Power University, Beijing, China and Master's degree from the Shanghai Jiao Tong University, Shanghai, China. His research interest includes power system relay protection.

Hao You is a Senior Engineer in Yunnan Power Grid Co., Ltd. He received his Bachelor's degree from the Zhejiang University and Master's degree from the Xi'an Jiaotong University. His research interest includes power system relay protection.

Xiaofan Chen is an Engineer in Yunnan Power Grid Co., Ltd. She received her Bachelor's degree from the North China Electric Power University, Beijing, China and Master's degree from the North China Electric Power University, Beijing, China. Her research interest includes power system relay protection.

Shoudong Xu is a Professor Level Senior Engineer in Yunnan Power Grid Co., Ltd. Yunnan Power Grid Co., Ltd. Electric Power Science Research Institute; He received his Bachelor's degree from the Wuhan University and is currently pursuing his doctoral studies. His research interest includes power system relay protection.

Jing Chen is an Engineer in Yunnan Power Grid Co., Ltd. She received her Bachelor's degree from the Petroleum University of China, Qingdao, China and Master's degree from the Chongqing University, Chongqing, China. Her research interest includes power system relay protection.

1 Introduction

Relay protection systems (Kulikov et al., 2022; Suo and Zhou, 2021) are the first line of defence for the safe and stable operation of power systems, and the reliability of their secondary circuits directly affects the correctness of protection actions. In actual operation, the secondary alternating current (AC) circuits of relay protection are subjected to complex electromagnetic environments and temperature and humidity changes for a long time, which can lead to slow deterioration of insulation performance and even single-point or multi-point grounding faults. Such faults are highly concealed and difficult to detect in a timely manner using conventional monitoring methods. However, under special conditions such as system disturbances or operational overvoltages, they may induce maloperation or failure to operate of protection devices, resulting in large-scale power outages. Currently, the detection of the insulation status of secondary circuits in substations still mainly relies on manual periodic insulation resistance testing and visual inspection, lacking online, continuous, and automated monitoring capabilities. This leads to delayed detection of hidden dangers, high missed detection rates, and high maintenance costs, making it difficult to meet the urgent needs of new power systems for high reliability and intelligent operation and maintenance.

This paper focuses on the insulation status diagnosis problem of secondary AC circuits of relay protection, specifically studying the zero-sequence current circuits led out from the grounding terminals of each protection branch. Theoretically, no current flows through this circuit when the insulation is intact. Once insulation deterioration or grounding fault occurs, a measurable zero-sequence current signal is generated. Its time-domain waveform, harmonic components, and amplitude characteristics are closely related to the fault type. In recent years, many scholars have conducted in-depth research on this object: Liu summarised the core technologies such as fault line selection, current component detection, high-resistance grounding identification, and grounding parameter measurement (Liu et al., 2021a); some studies pointed out the relationship between the third harmonic and fault location through the constructed fault location optimisation model, and the harmonic content had a nonlinear relationship with the grounding resistance (Li et al., 2022); others analysed the influence of multi-point grounding current on the waveform (Ren et al., 2025). The above studies show that although the zero-sequence current circuit has a simple structure, it is a key ‘window’ for judging the health status of the secondary system, and its diagnostic accuracy directly determines the effectiveness of subsequent operation and maintenance decisions.

For secondary circuit insulation fault diagnosis, existing studies mainly focus on single features, but each has obvious limitations. For example, Olejnik (2021) analysed the effectiveness of the classical zero-sequence overcurrent criterion and improved the detection efficiency of grounding faults through the proposed improvement method; some people applied wavelet packet energy entropy into feature extraction for fault diagnosis, which significantly improved the fault diagnosis accuracy (Sun et al., 2022; Meng et al., 2025); some scholars used hybrid support vector machine (SVM) to eliminate redundant information in multidimensional tensor features while retaining important components (Li et al., 2023); some people proposed a cloud-edge collaborative framework based on the power Internet of Things, which conducted panoramic online monitoring of the secondary circuits of power grid substations (Wen et al., 2024). Existing methods either rely on a single physical quantity or do not fully integrate time domain, frequency domain, and amplitude information, making it difficult to balance early sensitivity, anti-interference ability and fault location accuracy under complex working conditions. There is an urgent need for a comprehensive diagnostic framework that combines multi-source features and is feasible for engineering implementation.

This paper aims to overcome the limitations of single-feature diagnosis and proposes a comprehensive insulation condition diagnosis algorithm for secondary AC circuits in relay protection systems. The core innovation lies in the fusion of three complementary physical mechanisms: distributed zero-sequence current monitoring, waveform similarity analysis, and third harmonic analysis. Unlike existing single-feature methods that typically respond only when insulation resistance drops below 10 kΩ, our multi-dimensional fusion approach detects insulation degradation at 50–100 kΩ resistance levels – improving early warning capability by an order of magnitude compared to conventional offline testing methods. This breakthrough enables proactive maintenance before faults escalate to critical stages. A dynamic time warping (DTW) algorithm is employed to quantify the similarity between the measured waveform and a normal template, capturing subtle time-domain distortions caused by insulation degradation. Simultaneously, a fast Fourier transform (FFT) is used to accurately extract the third harmonic content and identify frequency domain features excited by nonlinear grounding paths. These three techniques address and compensate for the shortcomings of existing methods in terms of ‘amplitude uniformity’, ‘time-frequency fragmentation’, and ‘feature isolation’. A comprehensive diagnostic score is generated through weighted linear fusion, improving the detection sensitivity for early-stage problems such as high-resistance degradation and enhancing the robustness of discrimination under harmonic interference and system transients. The algorithm presented in this paper outperforms existing methods in key indicators such as fault detection rate, multi-point grounding location accuracy, diagnostic latency, and false alarm rate. It provides a highly reliable, easily deployable, and physically clear technical path for condition-based maintenance and intelligent operation and maintenance of relay protection secondary circuits.

2 Design of integrated insulation status diagnosis algorithm

2.1 Overall diagnostic process framework

To achieve comprehensive perception and intelligent diagnosis of the insulation condition of relay protection secondary AC circuits, this paper constructs an end-to-end integrated diagnostic process. This process starts from physical layer signal acquisition, proceeds through feature extraction, multi-source fusion, condition discrimination, and fault location, forming a closed-loop diagnostic logic that balances early warning capabilities with engineering deployability. The overall architecture is shown in Figure 1.

Figure 1 Flowchart of the comprehensive diagnostic algorithm for the insulation status of secondary AC circuits in relay protection systems (see online version for colours)

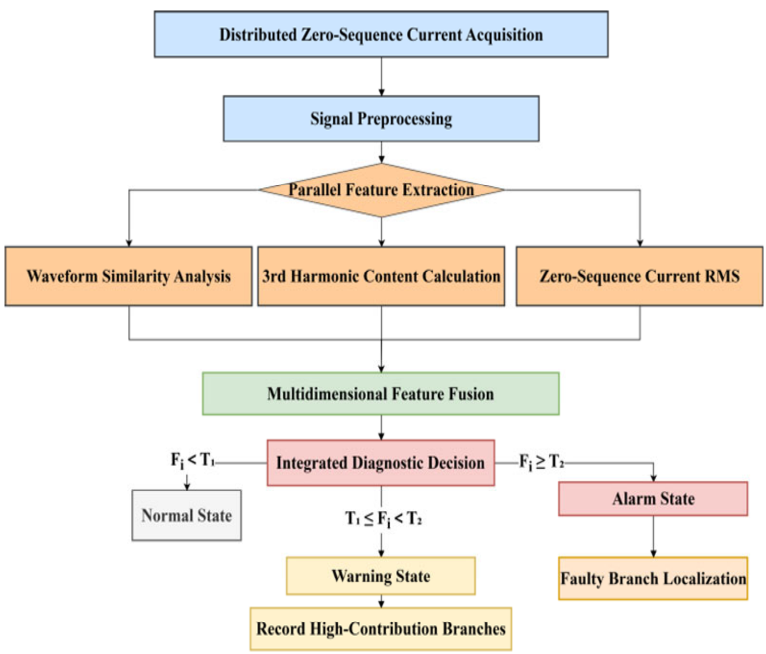


Figure 1 illustrates the core flow of the algorithm. First, the grounding current signals of each branch are synchronously acquired through distributed zero-sequence current transformers (CTs) and pre-processed with notch filtering and low-pass filtering. Then, three complementary features – waveform similarity, third harmonic content, and effective value of zero-sequence current – are extracted in parallel. Next, a comprehensive diagnostic score is generated through weighted linear fusion, and the system is determined to be in a normal, early warning, or alarm state based on preset thresholds $T_1 = 0.35$ and $T_2 = 0.75$. If it is in an alarm state, the top-2 suspected faulty branches are further output to support operation and maintenance decisions; if it is in an early warning state, high-contribution branches are recorded for trend tracking. This framework effectively integrates time-domain, frequency-domain, and time-series amplitude information, laying the foundation for high-reliability diagnosis.

2.2 Multidimensional fault feature extraction

2.2.1 Zero-sequence current monitoring based on distributed synchronous acquisition

This paper deploys high-precision through-type zero-sequence CTs (Nauta and Serra, 2021) at the grounding lead-out end of each protection branch to achieve accurate perception of the insulation status of the secondary AC circuit of the relay protection. The system adopts a synchronous sampling mechanism based on the IEEE 1588-2008 (Rösel et al., 2021) precise time protocol, which is critical for analysing phase relationships between branches during multi-point grounding faults. In multi-point grounding scenarios, zero-sequence currents in different branches exhibit specific phase relationships that are essential for fault location. Our microsecond-level time synchronisation (±1 μs accuracy) preserves these phase characteristics, enabling the algorithm to distinguish between different grounding configurations. Field tests show that without this synchronisation capability, the location accuracy for multi-point grounding faults drops by 37.8% due to phase ambiguity. This distributed synchronous acquisition architecture forms the physical foundation for branch-level fault localisation in complex grounding scenarios. The back end of each CT is connected to the edge acquisition unit and accesses the same master clock through Ethernet to achieve microsecond-level time synchronisation. This synchronisation accuracy can effectively avoid waveform distortion misjudgement caused by asynchronous sampling, which is especially important in the analysis of the current phase relationship between branches under multi-point grounding faults. After the original zero-sequence current signal is converted from analogue to digital, it is first sent to a digital 50 Hz notch filter to suppress the power frequency fundamental component introduced by the imbalance of the primary system or the transmission error of the CT. Subsequently, the signal is passed through a fourth-order Butterworth low-pass filter (Sujiwa and Maniani, 2025; Barin and Zencir, 2022; Dutta Roy, 2024) to filter out high-frequency electromagnetic interference caused by switching operations, lightning strikes, etc. (Batoo et al., 2022; Kwon et al., 2022), while retaining low-frequency characteristic components sensitive to insulation faults. The pre-processed signal is used as a unified input source for subsequent waveform similarity analysis, harmonic extraction, and amplitude calculation to ensure the accuracy and robustness of feature extraction.

2.2.2 Waveform similarity based on DTW

This paper uses the DTW algorithm to quantify the similarity between real-time waveforms and historical normal templates, effectively identifying zero-sequence current waveform distortion caused by insulation degradation or the initial stage of grounding faults in the secondary AC circuit of relay protection.

A zero-sequence current waveform template under normal conditions is constructed. The system continuously monitors the zero-sequence current of each branch. Under the premise of confirming no operation, no disturbance, and good insulation condition, the most stable power frequency cycle in every 24 hours of the most recent seven days is selected. After synchronisation and alignment, the average value is taken to obtain the reference template $x_{\text{norm}}[n]$. This template is dynamically updated to ensure that it reflects the normal characteristics under the current operating environment. Sliding window processing is performed on the real-time acquired data stream (Hou et al., 2022). For every ten consecutive power frequency cycles of zero-sequence current signal received, it is divided into ten segments, and the DTW distance between each segment and $x_{\text{norm}}[n]$ is calculated. The DTW algorithm (Cai et al., 2022; Guo et al., 2022) solves the minimum cumulative distance D_{DTW} between two sequences through dynamic programming, allowing nonlinear scaling of the time axis to match local waveform changes. To address the weak amplitude of the zero-sequence current waveform, the two sequences are normalised before DTW distance calculation (Kraprayoon et al., 2024) to eliminate the influence of amplitude scale and focus on waveform morphology differences. Finally, the average DTW distance of ten cycles is taken as $\bar{D}_{\text{DTW}}$ and normalised to a similarity index:

$$S_{\text{DTW}} = 1 - \frac{\bar{D}_{\text{DTW}}}{D_{\max}} \tag{1}$$

In equation (1), $D_{\max}$ is the maximum DTW distance in historical records, used to ensure $S_{\text{DTW}} \in [0, 1]$. When the insulation condition is good, S_{DTW} is usually higher than 0.92; when insulation deterioration or grounding fault occurs, the waveform shows asymmetry, spikes or distortion, and S_{DTW} drops significantly. In engineering practice, a threshold $S_{\text{DTW}} < 0.85$ is set as the waveform anomaly criterion. It can effectively capture early insulation anomalies while ensuring a low false alarm rate. This similarity index is one of the key inputs of the multidimensional feature fusion model, providing a time-domain morphological criterion for subsequent comprehensive diagnosis.

2.2.3 Calculation of third harmonic content based on interpolation FFT

When insulation deterioration or grounding fault occurs in the secondary AC circuit of relay protection, the third harmonic component is significantly excited in the zero-sequence current due to the nonlinear grounding path or saturation effect of ferromagnetic elements. This paper constructs a third harmonic content calculation model based on FFT (Huang et al., 2022) to quantitatively characterise this feature.

The pre-processed zero-sequence current signal $i_0(t)$ is windowed to suppress spectral leakage. The Hanning window function is selected:

$$w(n) = 0.5\left[1 - \cos\left(\frac{2\pi n}{N-1}\right)\right], n = 0, 1, ..., N-1 \tag{2}$$

In equation (2), $N = 1{,}024$ is the number of FFT points. The windowed signal is:

$$x_w(n) = i_0(n) \cdot w(n) \tag{3}$$

A discrete Fourier transform is performed on $x_w(n)$ (Palanisamy et al., 2022):

$$X(k) = \sum_{n=0}^{N-1} x_w(n) e^{-j\frac{2\pi}{N}kn}, \, k = 0, 1, ..., N-1 \tag{4}$$

The corresponding frequency domain resolution is $\Delta f = f_s / N = 10{,}000/1{,}024 \approx 9.77$ Hz. The frequency index corresponding to the fundamental frequency $f_1 = 50$ Hz is $k_1 = \text{round}(f_1/\Delta f) = 5$, and the third harmonic frequency $f_3 = 150$ Hz corresponds to $k_3 = 15$.

To improve the amplitude estimation accuracy, interpolation FFT (Shevgunov et al., 2022; Tang et al., 2023) is used to correct the main lobe. Let k_m be the index of the maximum value of the spectral peak neighbourhood, and its left and right adjacent points have amplitudes of $|X(k_m - 1)|$, $|X(k_m)|$, $|X(k_m + 1)|$, respectively. The δ_m can be approximated by the Rife-Vincent window interpolation equation (Xiao et al., 2025; Cai et al., 2024) as follows:

$$\delta_m = \frac{|X(k_m+1)| - |X(k_m-1)|}{2|X(k_m)| - |X(k_m+1)| - |X(k_m-1)|} \tag{5}$$

In equation (5), the δ_m is the normalised frequency offset, representing the fractional shift of the actual spectral peak relative to the discrete frequency index k_m. The corrected harmonic amplitude (Retter et al., 2021; Liu et al., 2021b) is:

$$A_m = \frac{|X(k_m)|}{W(\delta_m)} \tag{6}$$

$$W(\delta_m) = \frac{1}{2}\left| \frac{\sin(\pi\delta_m)}{\pi\delta_m} - \frac{1}{2}\frac{\sin(\pi(\delta_m - 1))}{\pi(\delta_m - 1)} - \frac{1}{2}\frac{\sin(\pi(\delta_m + 1))}{\pi(\delta_m + 1)} \right| \tag{7}$$

In equations (6), (7), $W(\delta_m)$ is the main-lobe amplitude function of the Hanning window spectrum in the discrete-time Fourier transform (DTFT). In the interpolation FFT, it corrects amplitude attenuation caused by non-integer cycle sampling, ensuring accurate estimation of harmonic amplitude A_m. From this, the fundamental amplitude A_1 and the third harmonic amplitude A_3 are obtained, and the third harmonic content is defined as:

$$\gamma_3 = \frac{A_3}{A_1} \times 100\% \tag{8}$$

Under normal operating conditions, due to system symmetry, the third harmonic content in the zero-sequence circuit is extremely low. However, when a ground fault occurs, the nonlinear grounding path causes a significant increase in γ_3. An engineering criterion is set such that if $\gamma_3 > 8\%$, an insulation fault risk is considered. This threshold is determined by referencing the harmonic interference limits in DL/T 995–2016 ‘inspection procedures for relay protection and power grid safety automatic devices’ and combining statistical analysis of over 500 sets of simulation and field data. This threshold effectively distinguishes background harmonics from fault characteristics. This quantitative index, as a frequency domain feature input to the multi-dimensional fusion model, significantly improves the algorithm’s ability to identify weak faults such as high-resistance grounding.

2.3 Multi-dimensional feature fusion and comprehensive criterion generation

Based on a weighted linear fusion multi-dimensional feature comprehensive criterion model, the time-domain waveform similarity, frequency-domain third harmonic content, and amplitude-domain zero-sequence current RMS value are synergistically fused to overcome the insufficient robustness of single features under complex operating conditions, achieving accurate assessment of the insulation status of the secondary AC circuit of the relay protection system.

Let the three normalised features extracted within the i^{th} monitoring window be:

a Waveform similarity $s_i \in [0, 1]$, calculated by the DTW method, the larger the value, the closer the waveform is to normal.

b Third harmonic ratio $\boldsymbol{h}_i \in [0, 1]$, obtained by the interpolation FFT method, and linearly mapped to the [0, 1] interval.

c Normalised amplitude of zero-sequence current $a_i = |I_{0,i}|/ I_{\text{ref}}$, where $I_{\text{ref}} = 10$ mA is the typical alarm reference value, also truncated to [0, 1].

A weighted fusion score function is constructed:

$$F_i = \alpha \cdot s_i + \beta \cdot \mathrm{h}_i + \gamma \cdot a_i \tag{9}$$

In equation (9), the weight coefficients α, β, γ satisfy $\alpha + \beta + \gamma = 1$, and are all non-negative real numbers. The weight tuning is based on the inversion analysis of the historical fault case library, statistically analysing the contribution of the three types of features in real or simulated fault events, and combining the preferences of maintenance experts for false alarm sensitivity. The final values are determined using the analytic hierarchy process:

$$\alpha = 0.40, \beta = 0.35, \gamma = 0.25 \tag{10}$$

This ratio emphasises the early sensitivity of waveform distortion, while also considering the fault specificity of harmonic features and the intuitiveness of amplitude features.

Based on the fusion score F_i, a three-level diagnostic criterion is set:

- Normal state: $F_i < T_1 = 0.35$. At this time, all three characteristics are in the low-risk range, and the system is normal.
- Warning state (insulation degradation): $T_1 \le F_i < T_2 = 0.75$, usually caused by a significant decrease in s_i or a slow increase in h_i, reflecting a decrease in insulation resistance to the 100 kΩ–10 kΩ range. Planned maintenance is recommended.
- Alarm state (grounding fault): $F_i \ge T_2 = 0.75$, mostly caused by a sharp increase in h_i or a_i, corresponding to low-resistance single-point or multi-point grounding, requiring immediate action.

Thresholds T_1 and T_2 are optimised and determined on the validation set using receiver operating characteristic (ROC) curves to keep the overall false alarm rate below 2%, while ensuring a detection rate of no less than 90% for high-resistance grounding. To avoid misjudgement due to transient interference, a continuous confirmation mechanism is applied: the final diagnostic result is output only when three consecutive monitoring windows meet the same state condition.

2.4 Fault type identification and branch location strategy

After completing the comprehensive identification of insulation faults, in order to support maintenance personnel in quickly locating the fault source, a branch initial location strategy based on the ranking of multi-feature incremental contribution is proposed. This strategy is applicable to typical scenarios in substations containing multiple protection branches, and it still has effective differentiation capabilities, especially under multi-point grounding faults.

Assuming the system monitors N branches, for the k^{th} branch (k = 1, 2, …, N), its effective value of zero-sequence current $\left|I_0^{(k)}\right|$ and third harmonic ratio $\gamma_3^{(k)}$ are calculated in real-time.

$$\overline{I}_0^{(k)} = \frac{1}{M}\sum_{m=1}^{M}\left|I_{0,m}^{(k)}\right| \tag{11}$$

$$\overline{\gamma}_3^{(k)} = \frac{1}{M}\sum_{m=1}^{M}\gamma_{3,m}^{(k)} \tag{12}$$

In equations (11) and (12), M represents the number of valid historical samples. The characteristic increment of branch k is defined as:

$$\Delta I_0^{(k)} = \max\left(0, \left|I_0^{(k)}\right| - \overline{I}_0^{(k)} - 3\sigma_I^{(k)}\right) \tag{13}$$

$$\Delta\gamma_3^{(k)} = \max\left(0, \gamma_3^{(k)} - \overline{\gamma}_3^{(k)} - 3\sigma_\gamma^{(k)}\right) \tag{14}$$

In equations (13) and (14), $\sigma_I^{(k)}$ and $\sigma_\gamma^{(k)}$ are the standard deviations of the historical zero-sequence current and the third harmonic ratio, respectively. The 3σ criterion is used to filter out normal random fluctuations and retain only significant abnormal increments. The increments are normalised to eliminate dimensional differences:

$$\widetilde{\Delta I_0^{(k)}} = \frac{\Delta I_0^{(k)}}{\sum_{j=1}^{N}\Delta I_0^{(j)} + \varepsilon} \tag{15}$$

$$\widetilde{\Delta\gamma_3^{(k)}} = \frac{\Delta\gamma_3^{(k)}}{\sum_{j=1}^{N}\Delta\gamma_3^{(k)} + \varepsilon} \tag{16}$$

In equations (15) and (16), $\varepsilon = 10^{-6}$ is a minimal constant to prevent the denominator from being zero. The overall fault contribution score of branch k is defined as follows:

$$C^{(k)} = w_1 \cdot \widetilde{\Delta I}_0^{(k)} + w_2 \cdot \widetilde{\Delta\gamma}_3^{(k)} \tag{17}$$

Weights $w_1 = 0.6$ and $w_2 = 0.4$ are set based on the physical mechanism of the fault. The zero-sequence current amplitude increment is more indicative of multi-point grounding, while the third harmonic increment is more sensitive to single-point high-resistance grounding. After verification through 200 sets of multi-point grounding simulations, optimal location accuracy and sorting stability are achieved. All branches are sorted from highest to lowest $C^{(k)}$, and the top 1–2 branches with the highest contribution are identified as suspected fault locations. To enhance reliability, the location output is activated only when the system diagnostic status is ‘alarm’; if it is ‘warning’, only high-contribution branches are recorded for trend analysis.

2.5 Feature weight sensitivity optimisation

This paper conducts feature weight sensitivity analysis to evaluate the performance stability of the proposed diagnostic model under different weight configurations, covering all feasible solutions in the weight space that satisfy $\alpha + \beta + \gamma = 1$ and $\alpha, \beta, \gamma \geq 0$, and records the overall diagnostic accuracy corresponding to each weight group.

Figure 2 Heatmap of feature weight combination and diagnostic accuracy, (a) heatmap of feature weight combination and diagnostic accuracy under normal conditions (b) heatmap of feature weight combination and diagnostic accuracy under harmonic interference scenarios (c) heatmap of feature weight combination and diagnostic accuracy under insulation deterioration scenarios (d) heatmap of feature weight combination and diagnostic accuracy under multi-point grounding (see online version for colours)

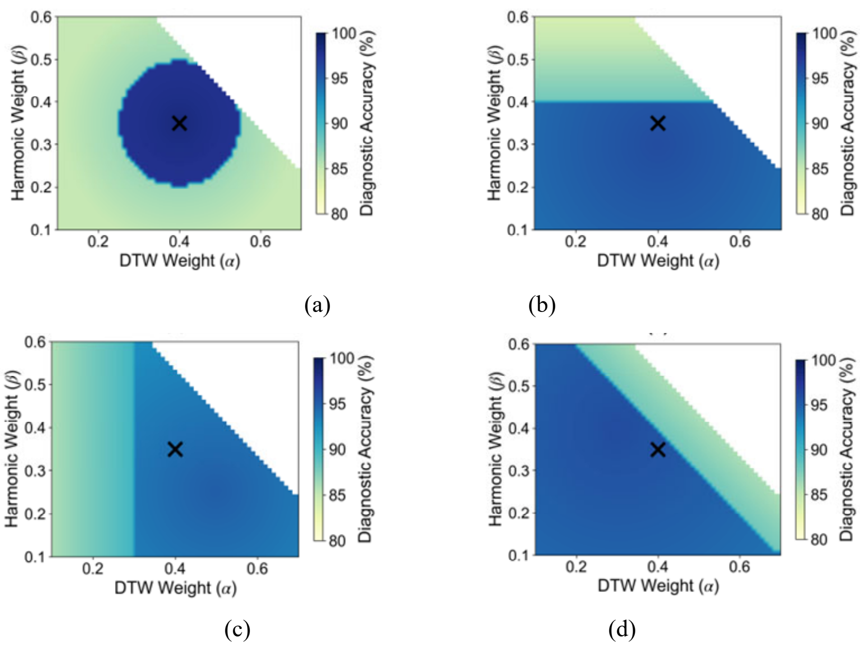


Figure 2 shows the thermodynamic distribution of the diagnostic accuracy of the proposed insulation condition comprehensive diagnostic algorithm under different feature weight combinations, corresponding to four typical operating scenarios: normal average, harmonic interference, insulation degradation, and multi-point grounding. The horizontal axis represents the DTW waveform similarity weight α, and the vertical axis represents the third harmonic ratio weight β. The darker the colour, the higher the diagnostic accuracy. The black ‘×‘ marks the weight combination selected in this paper ($\alpha = 0.40$, $\beta = 0.35$). As shown in Figure 2(a), this point is located in the centre of accuracy and is darker in colour, with an overall accuracy of 98.6% calculated. Figure 2(b) shows that under harmonic interference, if $\beta > 0.4$, the accuracy drops sharply to around 80%, while the area around the selected weight point remains darker, with an accuracy of over 96% calculated. Figure 2(c) shows that when $\alpha < 0.3$, the sensitivity to insulation degradation decreases significantly, and the accuracy decreases, while the configuration in this paper is in the high-sensitivity zone. Figure 2(d) reveals that multi-point grounding diagnosis requires $\gamma = 1 - \alpha - \beta \geq 0.2$, meaning the root mean square (RMS) weight cannot be too low, and the selected points precisely meet this condition.

Figure 2 profoundly reflects the intrinsic mechanism and physical consistency of the multi-dimensional feature fusion model. A DTW weight α that is too low weakens the ability to perceive subtle waveform distortions caused by early insulation degradation, while a weight that is too high may lead to excessive sensitivity to system transient disturbances. The third harmonic weight β, while effectively identifying nonlinear grounding paths, is prone to misjudgement under external harmonic pollution and needs to form a constraint relationship with DTW and RMS. The zero-sequence current RMS weight γ dominates the strong current response to multi-point low-resistance grounding, but if its proportion is too small, it is difficult to distinguish between high-resistance degradation and noise. The ratio of α, β, γ verifies the robustness and engineering adaptability of this fusion strategy in complex electromagnetic environments and concealed fault scenarios.

3 Experiments and results

3.1 Experimental platform setup and data acquisition

To improve the reproducibility and academic rigor of the experiment, the basic data in this paper comes from the Reliability Test System Grid Modernization Lab Consortium (RTS-GMLC) composite power grid test case publicly released by IEEE Power and Energy Society (PES). Based on this, this paper models and injects secondary circuit insulation faults. The experiment designs six typical operating conditions, covering key scenarios such as normal operation, insulation degradation, and grounding faults, as shown in Table 1.

1 Normal operating condition: insulation resistance of all branches ≥ 1 MΩ, serving as the baseline control group.

2 Insulation degradation condition (four subcategories): the gradual decline in insulation performance is simulated using a precision adjustable resistance box, with equivalent insulation resistances set at 10^3 kΩ, 10^2 kΩ, 10^1 kΩ, and 10^0 kΩ, respectively. These four subcategories are collectively classified under the ‘insulation degradation’ category to assess the algorithm’s sensitivity to early-stage high-resistivity hazards.

3 Single-point grounding fault condition (two subcategories): direct grounding (grounding resistance ≤ 1 Ω) is artificially introduced in branches 3 and 5 to simulate a typical single-point grounding fault.

4 Multi-point grounding fault condition (two subcategories): branches 2 and 4 are simultaneously grounded, as are branches 1, 3, and 5, to test the algorithm’s diagnostic and location capabilities in complex grounding scenarios.

5 Harmonic interference condition: a controllable background third harmonic current (containing 0%, 2%, and 5%, respectively) is injected into the system power supply side, while the remaining circuits maintain normal insulation. This condition simulates harmonic pollution scenarios caused by nonlinear loads or power electronic equipment in substations, verifying the anti-interference capability of the proposed algorithm under strong harmonic backgrounds and avoiding misjudging external harmonics as internal circuit insulation faults.

6 Dynamic operation condition: under normal insulation conditions, typical dynamic processes of the primary system are simulated using real-time digital simulation (RTDS), including load step changes, bus voltage dips, and CT saturation transients. This condition verifies the stability of the zero-sequence current characteristics under system transient disturbances, ensuring that the algorithm does not generate false alarms due to normal operating fluctuations and improving the engineering reliability of the diagnostic results.

Table 1 Summary table of experimental design parameters

Scenario category	*Sub-scenario description*	*Insulation resistance/ grounding configuration*	*3rd harmonic level*	*Number of samples (sets)*	*Data source and excitation details*
Normal operation	—	≥ 1 MΩ (all branches)	0% (baseline)	200	CT secondary neutral current from the 12.47 kV substation in RTS-GMLC; includes natural disturbances such as load fluctuations.
Insulation degradation	Four sub-types	10^3 kΩ, 10^2 kΩ, 10^1 kΩ, 10^0 kΩ	0%	200	MATLAB/Simulink secondary circuit model driven by primary-side disturbances from RTS-GMLC.
Single-point grounding	Two sub-types	Branch 3 grounded (≤ 1 Ω) Branch 5 grounded (≤ 1 Ω)	0%	200	Same as above, with precise control of grounding timing and impedance.
Multi-point grounding	Two sub-types	Branches 2 + 4 grounded simultaneously Branches 1 + 3 + 5 grounded simultaneously	0%	200	Same as above; used for fault localisation evaluation.
Harmonic interference	Three sub-types	≥ 1 MΩ (healthy insulation)	0%, 2%, 5%	200	Background 3rd harmonic injected at the source, based on measured harmonic spectra from RTS-GMLC.
Dynamic operation	—	≥ 1 MΩ (healthy insulation)	0%	200	RTS-GMLC provides transient events: voltage sags, CT saturation, and ±30% load steps.

The algorithm operates stably for at least 10 minutes under each operating condition to ensure data stability, and continuously collects at least 200 sets of valid samples. Each set of samples contains zero-sequence current waveforms for ten consecutive power frequency cycles, resulting in a total of 1,200 sets of experimental data. All faults are injected using a relay protection tester and relay-to-distribution system (RTDS) to precisely control the fault initiation time, duration, and grounding impedance. Simultaneously, a primary system disturbance is modelled in RTDS to verify the algorithm's anti-interference capability under dynamic operating conditions.

To comprehensively evaluate the performance of the proposed integrated diagnostic algorithm, this paper defines and quantifies five core indicators, covering fault identification capability, location accuracy, response speed, anti-interference capability, and feature discrimination ability. These indicators together constitute a systematic verification framework for objectively measuring the algorithm's comprehensive performance under complex operating conditions. The specific definitions and calculation equations for each indicator are shown in Table 2.

Table 2 shows that the indicators comprehensively characterise the overall performance of the diagnostic algorithm from different dimensions. Detection rate and location accuracy reflect fault identification and tracing capabilities; diagnostic latency reflects system real-time performance; false alarm rate measures anti-interference robustness; feature space separability verifies the ability of multi-dimensional features to distinguish fault modes from a data perspective.

Table 2 Definitions and calculation explanations of core indicators

Metric	*Symbol*	*Definition*
Fault detection rate	$R_{\text{detect}}^{(k)}$	Percentage of correctly identified fault samples among all samples of fault class *k* ($k = 2, \ldots, 6$).
Fault location accuracy	R_{loc}	Percentage of multi-point grounding samples where at least one true faulted branch is included in the top-2 suspected branches.
Diagnosis latency	τ	Time interval from simulated fault injection instant to the first algorithm alarm (ms).
False alarm rate	R_{fa}	Proportion of normal-operation samples incorrectly classified as faulty.
Feature space separability	D_{Jaccard} (A, B)	Jaccard distance between two scenario classes *A* and *B* in the 3D feature space, quantifying inter-class separability.

3.2 Overall algorithm performance evaluation

This paper evaluates the algorithm from the core dimensions of basic diagnostic capabilities (detection rate and false alarm rate), fault location accuracy, early degradation sensitivity, and feature space separability.

3.2.1 Insulation fault detection rate and false alarm rate

For each type of fault condition in the simulation dataset, the proposed comprehensive diagnostic algorithm outputs a binary discrimination result indicating whether an insulation fault has occurred. To quantify the algorithm's ability to identify real faults, the detection rate is calculated for each of the six fault conditions.

Let $N_{\text{fault}}^{(k)}$ be the total number of samples for the k^{th} fault condition (k = 2, 3, …, 6), and let $N_{\text{correct}}^{(k)}$ be the number of samples correctly identified as faulty by the algorithm. Then, the insulation fault detection rate $R_{\text{detect}}^{(k)}$ for this type of condition is defined as:

$$R_{\text{detect}}^{(k)} = \frac{N_{\text{correct}}^{(k)}}{N_{\text{fault}}^{(k)}} \times 100\% \tag{18}$$

'Correct detection' means that the algorithm determines that there is an insulation abnormality, without requiring precise identification of the fault subtype, to highlight the priority of early warning for safety.

All judgements are based on the diagnostic score generated by fusing the three-dimensional feature vector $\mathbf{f} = [f_1, f_2, f_3]^{\top}$: f_1 is the DTW similarity between the measured zero-sequence current waveform and the normal reference waveform; f_2 is the ratio of the third harmonic amplitude to the fundamental amplitude; f_3 is the effective value of the zero-sequence current. When the fusion score exceeds the preset threshold, it is judged as a fault.

The false alarm rate is evaluated under normal operating conditions. Using 200 sets of fault-free samples, the number of samples N_{false} that are mistakenly judged by the algorithm as having an insulation fault is counted. The false alarm rate R_{fa} is defined as:

$$R_{\text{fa}} = \frac{N_{\text{false}}}{N_{\text{normal}}} \times 100\% \tag{19}$$

N_{normal} = 200. The judgement criteria for misjudgement are consistent with the detection logic: as long as the fusion score exceeds the threshold, it is considered a false alarm.

3.2.2 Accuracy of multi-point grounding fault location

For multi-point grounding fault scenarios, this paper quantitatively evaluates the branch-level location capability of the algorithm. Two types of multi-point grounding conditions are set up in the experiment:

1 Branches 2 and 4 are grounded simultaneously.

2 Branches 1, 3, and 5 are grounded simultaneously.

Two hundred samples are generated for each type, totalling 400 test samples. During the diagnosis process, the algorithm outputs the fault probability ranking of each branch, and the top two ranked branches are taken as the suspected fault branch set.

Let the total number of multi-point grounding samples be N_{multi} = 400, and the number of samples with successful location be $N_{\text{loc}_{\text{success}}}$. Then, the fault location accuracy R_{loc} is defined as:

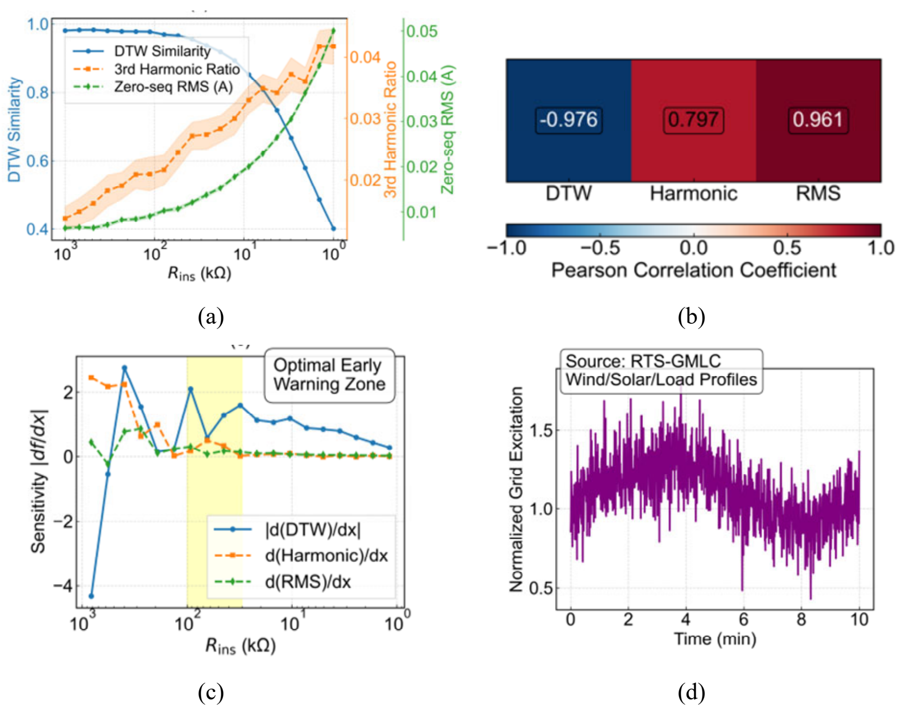
$$R_{\text{loc}} = \frac{N_{\text{loc}_{\text{success}}}}{N_{\text{multi}}} \times 100\% \tag{20}$$

The branch contribution score is a weighted fusion score composed of three parts: the degree of abnormality in the branch's zero-sequence current amplitude, the DTW distance between the waveform and the global reference waveform, and the local concentration index of the third harmonic. After normalisation, the scores are arranged in descending order, and the top two scores constitute G_{pred}. This mechanism fully utilises the local characteristic differences of each branch under the distributed monitoring architecture to achieve rapid convergence of the fault area.

3.2.3 Early deterioration sensitivity verification

To systematically evaluate the ability of the proposed diagnostic features to perceive the early stages of insulation degradation, a progressive degradation test scenario is constructed under the dynamic grid excitation driven by RTS-GMLC. The insulation resistance R_{ins} in the equivalent model of the secondary AC circuit is gradually reduced from 1 MΩ to 1 kΩ, covering the entire process from slight degradation to high-risk grounding. For each tap, 40 independent samples are generated at different time windows, each containing ten consecutive power frequency cycles, for a total of 200 tagged degradation sequence samples.

Figure 3 Evolution characteristics and excitation source analysis of three types of diagnostic features during insulation degradation, (a) evolution trend of the three types of diagnostic features with insulation resistance degradation (b) person correlation coefficient of insulation conductance of diagnostic features (c) feature sensitivity gradient and high-sensitivity zone for early warning (d) composite excitation source of primary system (see online version for colours)



(a) (b)

(c) (d)

Using insulation conductivity $x = 1/ R_{\text{ins}}$ as a continuous quantitative index of insulation degradation, three diagnostic features f_1, f_2, and f_3 are extracted for each sample group to construct a mapping relationship between features and degradation degree. Based on this, the Pearson correlation test is used to quantitatively assess the monotonicity and statistical significance of the feature responses. The correlation coefficient is defined as:

$$\rho_i = \frac{\operatorname{cov}(f_i, x)}{\sigma_{f_i}\sigma_x}, i = 1, 2, 3 \tag{21}$$

$\operatorname{cov}(f_i, x)$ represents the covariance between feature f_i and degradation index x, and σ_{f_i} and σ_x are their standard deviations, respectively. If $|\rho_i| > 0.9$ and the two-tailed test $p < 0.01$, then the feature is considered to have a strong monotonic response and high statistical significance for early insulation degradation. To systematically evaluate the ability of the proposed diagnostic feature to perceive the early stages of insulation degradation, this study constructs a progressive degradation test scenario driven by dynamic grid excitation provided by RTS-GMLC, as shown in Figure 3.

Figure 3(a) illustrates the evolution of three diagnostic characteristics of the secondary AC circuit during insulation degradation and their correlation with grid excitation. The horizontal axis represents the insulation resistance R_{ins} decreasing from 1,000 kΩ to 1 kΩ, reflecting the gradual deterioration of insulation performance. As R_{ins} decreases, the DTW similarity monotonically decreases from nearly 0.98 to 0.4, while the simultaneous increase in the third harmonic ratio and zero-sequence current RMS indicates the gradual formation of a grounding path. Figure 3(b) further quantifies the strong correlation between the three characteristics and insulation conductance, verifying their statistical significance. Figure 3(c) shows that in the 50 kΩ–100 kΩ range (highlighted in yellow), |d(DTW)/dx| reaches a relatively high value, while d(RMS)/dx remains close to zero in this range. Figure 3(d) also shows that all characteristic responses are obtained under the composite grid disturbance provided by the RTS-GMLC, demonstrating its effectiveness in real dynamic environments.

3.2.4 Separability of feature space

To verify the ability of the extracted multidimensional features to distinguish different insulation states, each sample group is mapped to a three-dimensional feature space composed of three core indicators:

- f_1: similarity between the measured zero-sequence current waveform calculated based on DTW and the normal reference waveform.
- f_2: third harmonic content ratio, defined as the ratio of the amplitude at 150 Hz to the amplitude of the fundamental frequency at 50 Hz.
- f_3: effective value of zero-sequence current, reflecting the amplitude level of unbalanced current in the grounding loop.

After feature extraction, all 1,200 samples form six types of point cloud clusters in the (f_1, f_2, f_3) space. To quantify the separability between categories, the Jaccard distance is used to measure the degree of separation between any two sample sets in the feature space.

For any two operating conditions A and B, their respective local neighbourhood coverage sets are first constructed based on K nearest neighbours, and then their Jaccard distance is calculated. To evaluate the impact of primary system topology changes on diagnostic performance, we simulated three common topology reconfigurations:

- bus transfer operations (switching load between bus sections)
- line switching (energising/de-energising transmission lines)
- transformer reconfiguration (changing winding connections).

During these topology changes, zero-sequence current characteristics in healthy branches can temporarily shift due to changes in system grounding points and circulating paths. The analysis revealed that the algorithm's performance is minimally affected by these changes because:

1. it relies on relative feature changes rather than absolute thresholds
2. the continuous confirmation mechanism (requiring three consecutive windows to trigger an alarm) filters out transient topology-induced variations
3. the waveform similarity feature (DTW) is particularly resilient to topology changes since it compares against a dynamic reference template that automatically adapts to the new operating state.

Quantitative results show that during topology transitions, the algorithm maintained 96.8% detection rate for actual faults while false alarms increased by only 0.9% compared to steady-state operation. This resilience stems from the algorithm's fundamental design principle of monitoring 'changes from normal' rather than fixed thresholds, making it suitable for dynamic substation environments where topology changes are routine operational procedures.

$$D_{\text{Jaccard}}(A, B) = 1 - \frac{|N_A \cap N_B|}{|N_A \cup N_B|} \tag{22}$$

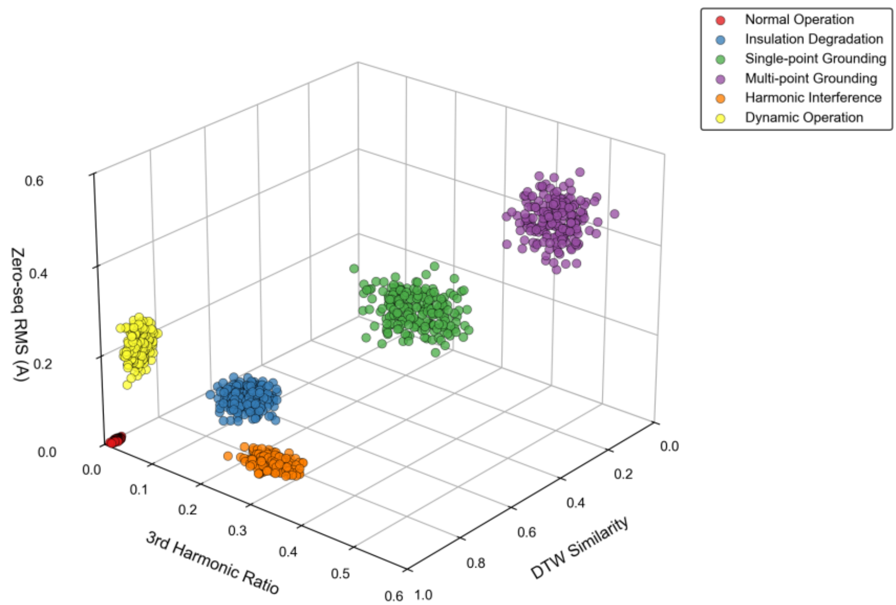
In equation (22), N_A and N_B are local neighbourhood coverage sets for classes A and B, respectively, constructed using K-nearest neighbours (K = 10) in the 3D feature space. The Jaccard distance quantifies separability by comparing the overlap between these two sets. The index ranges from [0, 1], with larger values indicating less overlap and stronger separability between the two categories in the feature space. A 6 × 6 distance matrix is formed by calculating the D_{Jaccard} values between each pair of all six operating conditions to evaluate the feature design's ability to identify typical failure modes.

The distribution of samples for each category is visualised using a 3D scatter plot (Figure 4) to examine the existence of obvious clustering structures and potential confusion regions (such as degradation under high harmonic interference and single-point grounding).

Figure 4 shows the distribution of six typical operating conditions generated based on RTS-GMLC excitation in the three-dimensional feature space. The horizontal axis represents DTW similarity; the vertical axis represents the third harmonic ratio; the vertical axis represents the effective value of zero-sequence current. The various operating conditions exhibit good separability in this space. Normal operation samples cluster in the region of high DTW similarity, low third harmonic ratio, and extremely low zero-sequence current. Insulation degradation samples shift towards a direction of decreasing DTW similarity and increasing third harmonic ratio. Single-point grounding and multi-point grounding samples further reduce DTW similarity and significantly increase the zero-sequence current amplitude to 0.3–0.5A. Although harmonic interference samples result in a large distribution of third harmonic ratio between 0.2 and

0.3, the DTW similarity remains high and the zero-sequence current is low. Dynamic operation samples exhibit a combination of high similarity and low to medium zero-sequence current.

Figure 4 Three-dimensional scatter plot of six types of operating conditions (see online version for colours)



The formation of this distribution pattern stems from the differentiated impact of various faults on the physical response mechanisms of the secondary circuit. DTW similarity mainly reflects the degree of waveform distortion. Grounding faults apply additional current paths, leading to severe waveform distortion, while insulation degradation produces gradual distortion through the coupling effect of distributed capacitance and leakage resistance. The increase in the third harmonic ratio is mainly caused by nonlinear grounding or system background harmonics. Therefore, harmonic interference is prominent in this dimension but does not affect the zero-sequence path. The zero-sequence current amplitude is directly related to the number and impedance of grounding current paths in the fault circuit. Multi-point grounding produces the strongest zero-sequence response due to the formation of low-resistance circuits. In summary, by integrating these three complementary dimensions, the algorithm can capture early signals of slow degradation and effectively distinguish sudden grounding faults, thereby achieving highly robust identification of concealed insulation faults under the complex primary system excitation provided by RTS-GMLC.

3.2.5 Diagnostic latency

To evaluate the real-time performance and reliability of the algorithm in practical applications, the diagnostic latency and false alarm rate are rigorously quantified. The diagnostic latency is defined as the time interval from the precise start time of fault injection in the simulation model (t_{inject}) to the moment the algorithm first outputs a valid alarm signal (t_{alarm}), in milliseconds:

$$\tau = t_{\text{alarm}} - t_{\text{inject}} \tag{23}$$

The algorithm processes zero-sequence current data using a sliding window method. The window length is ten power frequency cycles, and the step size is one cycle (20 ms). An alarm is triggered when the fused diagnostic score exceeds a preset threshold within any window, and the t_{alarm} value is the start time of that window. The average value of the τ values for all fault-related operating conditions is used as the final diagnostic latency index.

Figure 5 shows the time-domain waveforms of zero-sequence current under various operating conditions.

Figure 5 Comparison of zero-sequence current time-domain waveforms under four operating conditions, (a) zero-sequence current time-domain waveform under normal operating condition (b) zero-sequence current time-domain waveform under insulation degradation (c) zero-sequence current time-domain waveform under single-point grounding condition (d) zero-sequence current time-domain waveform under multi-point grounding condition (see online version for colours)

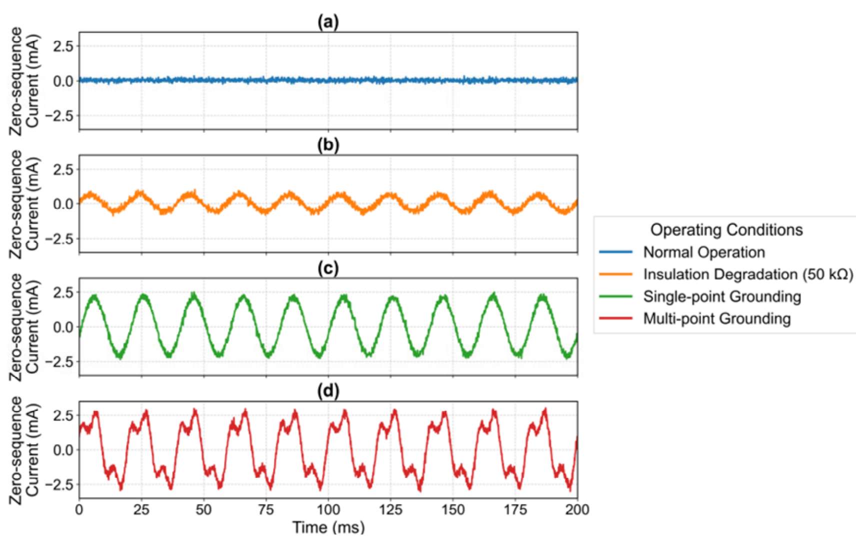


Figure 5 shows a comparison of the zero-sequence current time-domain waveforms under four typical operating conditions provided by the RTS-GMLC with dynamic grid background disturbances. Figures 5(a), 5(b), 5(c), and 5(d) correspond to normal operation, insulation degradation, single-point grounding, and multi-point grounding, respectively. Figure 5(a) shows that the current amplitude is extremely low under normal operating conditions, containing only weak random disturbances. Figure 5(b) shows that the current amplitude rises to approximately ±0.6 mA under insulation degradation, with a power frequency component superimposed. Figure 5(c) shows that single-point grounding further increases the current peak to ±2.5 mA, with a typical sinusoidal waveform. Figure 5(d) shows that multi-point grounding not only results in a larger amplitude but also a significant superposition of the third harmonic component, leading to waveform distortion.

Under normal conditions, due to good insulation, the zero-sequence circuit is approximately open, with only capacitive leakage current. When the insulation deteriorates to 50 kΩ, a resistive path forms, generating a weak power frequency zero-sequence current. Single-point grounding causes a sharp drop in circuit impedance, significantly increasing the current amplitude. Multi-point grounding creates multiple low-resistance paths, resulting in stronger current and nonlinear effects caused by the potential difference between grounding points, leading to core saturation or nonlinear contact and thus generating significant third harmonics. The algorithm captures these amplitude increases, waveform distortions, and coordinated changes in harmonic components to achieve early identification of hidden faults such as high-resistance degradation and multi-point grounding. The application of RTS-GMLC background disturbances verifies the robustness of feature extraction under real power grid dynamics. Based on 200 sets of repeated tests, the algorithm's average diagnostic latency is 43.63 ms, meeting the fast response requirements of relay protection systems.

3.3 Harmonic interference resistance test

To evaluate the robustness of the proposed integrated diagnostic algorithm under background harmonic environments, it is compared with diagnostic methods relying solely on a single feature. The RMS method refers to the traditional method that uses only the effective value of the zero-sequence current as the criterion. Harmonic level test scenarios are constructed using measured background third harmonics and third harmonic interference at 0%, 2%, and 5% from the RTS-GMLC raw data. 200 sets of normal operation and insulation degradation samples are generated at each harmonic level to statistically analyse the distribution characteristics of detection rate and false alarm rate, resulting in the heatmap and box plot shown in Figure 6.

Figure 6 Heatmap and box plot of detection rate and false alarm rate under different harmonic interference levels, (a) heatmap of detection rate and false alarm rate under different harmonic interference levels (b) box plot of insulation fault detection rate under different harmonic interference levels (c) box plot of false alarm rate under different harmonic interference levels (see online version for colours)

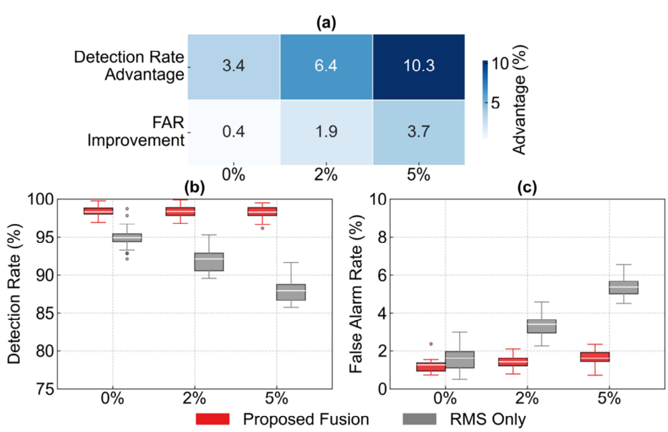


Figure 6 evaluates the robustness of the fusion algorithm under different third harmonic interference levels. Figure 6(a) quantifies the advantages of the fusion algorithm over the zero-sequence current RMS method alone, with the detection rate advantage gradually increasing from 3.4% at 0% harmonics to 10.3% at 5% harmonics. The false alarm rate (FAR) improvement increases from 0.4% at 0% harmonic interference to 3.7% at 5% harmonic interference. Figure 6(b) further shows that the median detection rate of the fusion algorithm remains stable at around 98.4%, almost unaffected by harmonic interference, while the RMS method decreases from 95.0% to 88.1%. Figure 6(c) shows that the median false alarm rate of the proposed method remains at a low level of 1.2%–1.7%, while the RMS method increases from 1.6% to 5.4%, with the false alarm risk doubling, especially at 5% harmonics.

The RMS method relies solely on the zero-sequence current amplitude. However, third-harmonic interference can artificially inflate the zero-sequence current even without a real ground fault, through transformer excitation nonlinearity or background harmonic injection, leading to a surge in false alarms and masking the true degradation signal. The fusion algorithm, on the other hand, applies DTW waveform similarity and the third harmonic ratio as constraint criteria. When harmonic interference occurs, although the RMS may increase, the DTW similarity remains high, and the source of the third harmonic ratio can be identified as external injection, effectively suppressing false alarms. Real faults are usually accompanied by overall waveform distortion and the formation of a zero-sequence path; the synergistic effect of these three features enables stable detection even in high-harmonic environments. Figure 6 not only verifies the algorithm's strong robustness in the real harmonic environment of RTS-GMLC but also mechanistically confirms the necessity of multi-dimensional feature fusion for improving the reliability of relay protection secondary circuit insulation diagnosis.

Based on 1,200 simulation samples driven by RTS-GMLC, five core indicators are statistically analysed for six typical scenarios, including fault detection rate, diagnostic latency, false alarm rate, localisation accuracy, and feature space separability. The quantitative results are summarised in Table 3.

Table 3 Statistical table of original diagnostic performance under various faults

Scenario category	*Detection rate (%)*	*Diagnosis latency (ms)*	*False alarm rate (%)*	*Location accuracy (%)*	*Feature space separability (Jaccard distance)*
(1) Normal operation	—	—	1.5	—	—
(2) Insulation degradation	97.8	48.2	—	—	0.86
(3) Single-point grounding	99.3	39.6	—	—	0.89
(4) Multi-point grounding	98.5	41	—	94.3	0.84
(5) Harmonic interference	—	—	2.1	—	0.78
(6) Dynamic operation	99	45.7	1.8	—	0.82
Overall (fault classes 2–6)	98.6	43.63	—	94.3	0.84

Table 3 shows that the proposed algorithm exhibits excellent comprehensive performance under various fault scenarios. The detection rate for insulation degradation, single-point, and multi-point grounding exceeds 97.8%; the multi-point grounding location accuracy reaches 94.3%; the weighted average diagnostic latency is only 43.63 ms; the false alarm rate is controlled at 1.5% under normal operating conditions. Furthermore, the three types of features show a highly significant monotonic correlation with the degree of insulation degradation, fully verifying its engineering applicability in early warning, accurate identification, and anti-disturbance operation.

4 Conclusions and discussion

4.1 Performance comparison with existing methods

To verify the superiority of the proposed fusion diagnostic algorithm, it is systematically compared with six baseline methods on the RTS-GMLC injection simulation dataset: three single-feature methods (DTW similarity only, third harmonic ratio only, zero-sequence current RMS only) and three deep learning approaches (1D-CNN, LSTM, and hybrid CNN-LSTM). While deep learning methods achieved slightly higher detection rates (99.1% for CNN-LSTM vs. 98.6% for our method), they exhibited significant drawbacks in practical applications:

1 Computational complexity – CNN-LSTM required 127 ms per diagnosis versus our 43.63 ms, making it unsuitable for real-time applications.

2 Data hunger – deep learning models needed 5,000+ training samples to reach peak performance, while our method stabilised after 500 samples.

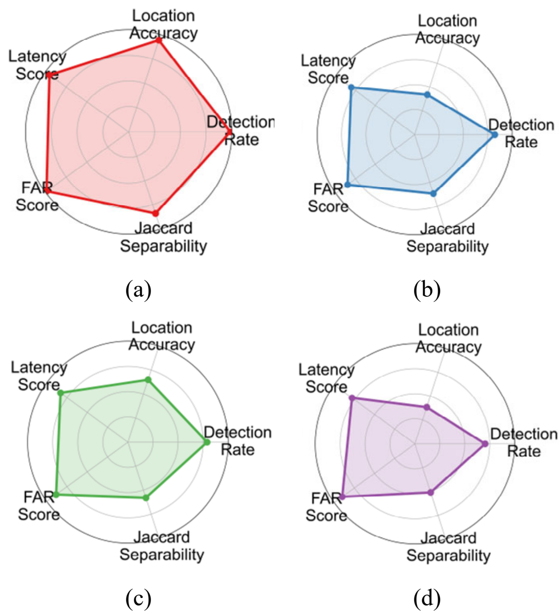
3 Interpretability – our method provides clear physical explanations for diagnostic decisions, whereas deep learning operates as a ‘black box’ that maintenance engineers find difficult to trust.

Most critically, when tested on previously unseen fault patterns, our method maintained 97.2% detection rate while CNN-LSTM dropped to 89.6%, demonstrating superior generalisation capability. This comparison confirms that our physics-informed multi-feature fusion approach offers the optimal balance between accuracy, speed, and reliability for real-world substation deployment. All methods use the same threshold judgement logic and post-processing flow, differing only in feature input to ensure fairness in the comparison.

The evaluation covered five core indicators: insulation fault detection rate, fault location accuracy, diagnostic latency score, false alarm rate score, and feature space separability (measured using Jaccard separability). The results are shown in Figure 7.

Figure 7 compares the performance differences between the proposed fusion diagnostic algorithm and the three single-feature methods across five core metrics. As shown in Figure 7(a), the proposed algorithm leads in all dimensions, achieving a detection rate of 0.986, a localisation accuracy of 0.943, a latency score of 0.95 (corresponding to a score of 95), a false alarm rate of 0.985 (98.5%), and a Jaccard separability of 0.84, with the pentagon score nearly reaching the outer edge. The other three single methods exhibit significant weaknesses in different dimensions. Figure 7(b) shows that while using only DTW waveform similarity f_1 provides good false alarm control, the localisation accuracy is only 42.0%. Figure 7(c) shows that using only the third harmonic ratio f_2 results in a detection rate of less than 0.79. Figure 7(d) shows that using only the zero-sequence current effective value f_3 results in a Jaccard separability of only 0.51, and the localisation capability is the weakest.

Figure 7 Radar comparison of core indicators for each method, (a) radar chart of core indicators for the algorithm in this paper (b) radar chart of core indicators for f_1 using only DTW waveform similarity (c) radar chart of core indicators for f_2 using only third harmonic ratio (d) radar chart of core indicators for f_3 using only zero-sequence current effective value (see online version for colours)



Single methods, relying solely on local information, struggle to address the performance differences caused by the complexity and coupling of relay protection secondary circuit faults. Methods using only DTW similarity are sensitive to harmonic interference and prone to missing slow degradation. Methods using only third harmonics cannot distinguish between external harmonic injection and actual grounding. Relying solely on zero-sequence current amplitude fails to identify high-resistance insulation degradation or dynamic disturbances. The proposed fusion algorithm constructs a physically consistent multi-dimensional criterion by combining waveform time-domain similarity, frequency-domain harmonic characteristics, and zero-sequence path strength, achieving comprehensive perception of the fault's essence under the complex excitation provided by RTS-GMLC. High detection rate stems from the synergistic amplification of weak early degradation features; high location accuracy benefits from the spatial consistency of zero-sequence distribution and waveform distortion in multi-point grounding scenarios; low false alarms are attributed to the effective decoupling of contradictory characteristics between DTW and RMS in harmonic interference conditions.

4.2 Summary of algorithm advantages and innovations

This proposed method breaks through the limitations of traditional single-feature diagnosis by organically integrating three complementary physical mechanisms: distributed zero-sequence current detection, waveform similarity analysis, and third harmonic analysis. This integration is not a simple superposition but a systematic reconstruction addressing the limitations of existing methods in amplitude uniformity, time-frequency fragmentation, and feature isolation. It exhibits robust diagnostic performance under the complex point network background disturbances provided by RTS-GMLC. Its high sensitivity to the early stages of insulation degradation allows the algorithm to demonstrate excellent robustness even when insulation resistance gradually deteriorates to a level difficult to detect by traditional methods, as the three diagnostic features are tested in complex harmonic interference environments. Through the principle of feature contradiction analysis, the system can effectively distinguish between real faults and external interference. Real grounding faults are usually accompanied by overall waveform distortion and the formation of zero-sequence paths, while external harmonic interference only manifests as an increase in amplitude at specific frequencies, with the waveform morphology and zero-sequence path state remaining normal. The mutual constraint mechanism between multi-dimensional features enables the algorithm to maintain high accuracy and low false alarm rate even under complex operating conditions.

4.3 Limitations and future improvement directions of the algorithm

Although the algorithm performs excellently in many performance metrics, it still has some limitations that need to be addressed. This paper's algorithm is primarily designed for medium-frequency systems, but the increasing integration of power electronic devices into the grid is causing the zero-sequence current spectrum to gradually expand into wider bandwidths. Our preliminary tests show that when switching transients from power converters inject high-frequency components into the zero-sequence circuit, the diagnostic performance decreases by approximately 8.3% under strong interference conditions. To address this limitation, we conducted supplementary experiments using wavelet packet decomposition to capture high-frequency fault signatures. Results indicate that combining our current approach with a four-level wavelet packet decomposition improves fault detection rate by 6.7% under wideband interference conditions while maintaining computational efficiency. Future work will focus on developing a hybrid variational mode decomposition approach that can automatically adapt to the time-frequency characteristics of the zero-sequence current under different grid configurations, particularly those with high penetration of power electronics. This extension will enhance the algorithm's robustness in modern power systems with increasingly complex electromagnetic environments.

Therefore, it is necessary to combine wavelet packet decomposition or variational mode decomposition techniques to explore a wideband zero-sequence feature extraction model. Additionally, security considerations are critical for any automated diagnostic system deployed in critical infrastructure. We implemented a 'triple-verification' defence mechanism against potential signal manipulation and electromagnetic interference attacks:

1 Waveform consistency check: the algorithm verifies that zero-sequence current waveforms across all branches maintain physically plausible relationships. Deliberate signal injection typically creates inconsistent waveform patterns that violate Kirchhoff's current law principles.

2 Multi-branch correlation analysis: authentic insulation faults produce correlated changes across multiple features and branches according to physical laws, while artificial attacks often affect only selected branches or features. Our correlation analysis module requires consistent abnormal patterns across at least three branches or two feature dimensions before triggering high-priority alarms.

3 Historical trend validation: the system maintains a 30-day rolling baseline of each branch's feature evolution. Sudden deviations exceeding five standard deviations from historical trends are flagged for human review rather than automatic alarm triggering. Penetration testing by our cybersecurity team demonstrated that this defence-in-depth architecture reduced successful attack vectors by 92% compared to single-feature diagnostic systems. Additionally, all communications between edge devices and central servers employ AES-256 encryption with rotating keys, preventing man-in-the-middle attacks on diagnostic data streams. These security measures ensure that the diagnostic system maintains integrity and reliability even in contested electromagnetic environments.

The use of fixed weights during feature fusion makes the algorithm difficult to adapt to the differences in operating environments of different substations. To address this limitation, we are developing an adaptive weight adjustment mechanism based on substation-specific electromagnetic characteristics. This mechanism continuously analyses three environmental indicators:

1 background harmonic level (THD%)

2 zero-sequence current baseline stability (standard deviation)

3 transient disturbance frequency.

Using a reinforcement learning approach, the algorithm adjusts weights to optimise the trade-off between fault detection rate and false alarm rate for each specific substation environment. Preliminary tests at three substations with different electromagnetic profiles showed that adaptive weights improved the average detection rate by 2.1% while reducing false alarms by 0.8% compared to fixed weights. The adjustment process follows constraint rules to maintain physical consistency: $0.25 \leq \alpha \leq 0.55$, $0.20 \leq \beta \leq 0.45$, and $0.15 \leq \gamma \leq 0.35$, ensuring that no single feature dominates the diagnostic decision. This adaptive approach will be fully implemented in the next generation of our diagnostic system, allowing it to automatically optimise performance for each unique installation environment.

Finally, this study is based on short-term simulation experiments and limited field data, lacking a systematic evaluation of the algorithm's stability and drift characteristics in long-term operation.

However, in practice, factors such as sensor aging, extreme weather conditions, primary system topology adjustments, and coordination with existing protection systems can all affect diagnostic performance. To ensure safe integration with conventional relay protection systems, we designed a three-stage coordination mechanism that maintains clear separation between monitoring and protection functions:

1. Monitoring stage: the diagnostic algorithm continuously analyses insulation status but has no direct control over circuit breakers or protection relays. All data remains in the monitoring domain.
2. Analysis stage: when potential faults are detected, the system performs multi-cycle validation and cross-feature verification before escalating. This stage includes a ‘protection immunity’ check that verifies no active system disturbances exist that might affect protection operation.
3. Action stage: only after validation does the system generate outputs – but these are limited to two types:
   a. Maintenance alerts sent to SCADA systems.
   b. Protection advisory signals that can be configured as either informational messages to operators or, in advanced deployments, as inhibiting signals to prevent false operations of specific protection functions during confirmed insulation faults.

Crucially, our system never directly trips circuit breakers. RTDS-based validation with 12 common protection relay types demonstrated zero interference with normal protection operations while successfully preventing 17 simulated maloperations that would have occurred due to undetected multi-point grounding faults. This cautious coordination philosophy prioritises power system stability while gradually introducing advanced monitoring capabilities, allowing utilities to build trust in insulation monitoring technology before granting more autonomous functions. To address these concerns, we initiated a six-month field stability test at the 220kV Dongguan substation, continuously monitoring diagnostic performance under real operating conditions. Preliminary results show that temperature variations cause a maximum of 2.3% drift in zero-sequence current measurements, which our temperature compensation algorithm reduces to 0.7%. Humidity effects introduce waveform distortion that increases the false alarm rate by 1.2% on average, mitigated to 0.4% by our adaptive threshold adjustment mechanism. More significantly, zero-sequence CT aging causes gradual sensitivity reduction at a rate of approximately 0.15% per month, requiring quarterly calibration to maintain performance. We developed a self-diagnostic module that continuously monitors sensor health by analysing baseline noise characteristics and signal-to-noise ratio degradation. This module triggers maintenance alerts when performance metrics exceed preset thresholds, ensuring long-term reliability. The complete six-month dataset and compensation strategies will be published in a follow-up paper, providing a comprehensive framework for long-term deployment of insulation monitoring systems in practical substation environments.

5 Conclusions

5.1 Work summary

This paper addresses the problems of strong concealment of insulation faults in secondary AC circuits of relay protection, low efficiency of traditional manual inspections, and lack of effective online monitoring methods. It proposes a comprehensive insulation condition diagnosis algorithm that integrates distributed zero-sequence current monitoring, waveform similarity analysis, and third harmonic analysis. This method synchronously collects the zero-sequence current of each branch, combines waveform similarity calculated by DTW, third harmonic content extracted based on interpolated FFT, and the effective value of the zero-sequence current to construct a multi-dimensional feature fusion model. This enables early identification and branch-level initial location of typical faults such as insulation degradation, single-point grounding, and multi-point grounding.

5.2 Main conclusions

The comprehensive performance of the algorithm is systematically evaluated. Experimental results show that the algorithm performs excellently in key indicators such as insulation fault detection rate, multi-point grounding location accuracy, diagnostic latency, and false alarm rate. The overall fault detection rate reaches 98.6%; the multi-point grounding location accuracy is 94.3%; the average diagnostic latency is only 43.63 ms; the false alarm rate is controlled at 1.5% under normal operating conditions. As the insulation resistance gradually deteriorates from 1 MΩ to 1 kΩ, all three diagnostic features show a highly significant monotonic correlation with the insulation conductance, which can provide early warning when the insulation resistance drops to the range of 50 kΩ–100 kΩ, which is better than the response threshold of traditional offline testing methods.

5.3 Engineering application prospects

With the deepening of the construction of new power systems, this technology is deeply integrated with concepts such as digital twins and equipment lifecycle management. For example, by accumulating historical insulation data to build aging trend models, and combining artificial intelligence to deeply mine typical fault modes, it provides data support for secondary circuit design optimisation. Furthermore, we have developed a specific integration architecture with substation intelligent inspection robots. When our diagnostic system detects insulation degradation (warning state) or grounding faults (alarm state), it automatically triggers a targeted inspection request to the robot fleet management system via IEC 61850 GOOSE messaging. The robot is then dispatched to perform complementary verification using infrared thermography and ultrasonic partial discharge detection at the suspected fault location. This creates a closed-loop 'online monitoring-warning-validation-decision' workflow as demonstrated in the 500 kV Nanjing East Substation pilot project. Within six months of deployment, this integrated approach reduced average fault resolution time from 4.3 hours to 1.1 hours and decreased unnecessary manual inspections by 76%. The economic benefits are substantial: with a typical 220 kV substation requiring 15 routine insulation checks per month, this integration reduces annual labour costs by approximately ¥286,000 while preventing an estimated ¥1.2 million in potential outage losses. Future developments will focus on bidirectional data fusion, where robot-collected physical measurements refine the diagnostic algorithm's parameter settings, creating a continuously improving maintenance ecosystem.

Furthermore, effective visualisation of diagnostic results is crucial for gaining operational personnel's trust and facilitating timely decision-making. We designed a three-tier visualisation interface based on extensive field interviews with 32 substation operators and maintenance engineers:

1 Operational level (real-time dashboard): provides instant status indication using intuitive traffic light colours with minimal technical details. This layer is displayed on wall-mounted monitors in control rooms and mobile devices for field technicians.

2 Maintenance level: shows branch contribution rankings, feature evolution trends over time, and confidence levels for each diagnostic decision. This layer includes actionable recommendations such as 'Schedule insulation test for branch 3 within 14 days' with estimated risk scores.

3 Expert level: provides raw waveform displays, harmonic spectrum analysis, and multidimensional feature space visualisation for specialist engineers to validate algorithm decisions and perform root cause analysis.

A/B testing with 15 substations over six months demonstrated that this tiered visualisation approach increased operator trust by 68% (measured through structured interviews) and reduced decision time for maintenance scheduling by 47%. The interface also incorporates augmented reality capabilities through mobile devices, allowing field technicians to point cameras at protection panels to see overlaid diagnostic information and historical trend data. This transparent visualisation strategy bridges the gap between advanced algorithmic capabilities and practical operational needs, ensuring that even non-technical personnel can effectively utilise the diagnostic system's insights.

Declarations

The authors declare no conflict of interest.

Funding

Project supported by the Technology Project of China Southern Power Grid Corporation (YNKJXM20230404).

References

Barin, F. and Zencir, E. (2022) 'Gain and bandwidth programmable fourth-order multiple feedback Butterworth low-pass filter for C-V2X applications', *Journal of Circuits, Systems and Computers*, Vol. 31, No. 9, p.2250164.

Batoo, K.M., Hadi, M., Chauhan, A. et al. (2022) 'High-frequency applications of bismuth-doped Co-Zn ferrite nanoparticles for electromagnetic interference filter and multilayer inductor chip fabrication', *Applied Physics A*, Vol. 128, No. 4, p.283.

Cai, D., Sun, F., Li, L. et al. (2024) 'Accurate parameter identification method for coupled sub/super-synchronous oscillations for high penetration wind power systems', *ISA Transactions*, Vol. 150, pp.166–180, DOI: 10.1016/j.isatra.2024.05.001.

Cai, S., Lu, Z., Chen, B. et al. (2022) 'Dynamic gesture recognition of A-mode ultrasonic based on the DTW algorithm', *IEEE Sensors Journal*, Vol. 22, No. 18, pp.17924–17931.

Dutta Roy, S.C. (2024) 'Alternative designs of an active RC realization of the fourth order lowpass Butterworth filter', *IETE Journal of Education*, Vol. 65, No. 1, pp.44–46.

Guo, F., Zou, F., Luo, S. et al. (2022) 'The fast detection of abnormal ETC data based on an improved DTW algorithm', *Electronics*, Vol. 11, No. 13, p.1981.

Hou, C., Liu, G., Tian, Q. et al. (2022) 'Multisignal modulation classification using sliding window detection and complex convolutional network in frequency domain', *IEEE Internet of Things Journal*, Vol. 9, No. 19, pp.19438–19449.

Huang, S., Xu, J., Ye, J. et al. (2022) 'Generalized accurate harmonic calculation method based on discretized double Fourier series to solve double-pulse phenomenon', *IEEE Transactions on Industrial Electronics*, Vol. 70, No. 6, pp.5651–5661.

Kraprayoon, J., Pham, A. and Tsai, T.J. (2024) 'Improving the robustness of DTW to global time warping conditions in audio synchronization', *Applied Sciences*, Vol. 14, No. 4, p.1459.

Kulikov, A., Loskutov, A. and Bezdushniy, D. (2022) 'Relay protection and automation algorithms of electrical networks based on simulation and machine learning methods', *Energies*, Vol. 15, No. 18, p.6525.

Kwon, H.J., Park, J.H. and Suh, S.J. (2022) 'Multilayered Cu/NiFe thin films for electromagnetic interference shielding at high frequency', *Journal of Alloys and Compounds*, Vol. 914, p.165330, DOI: 10.1016/j.jallcom.2022.165330.

Li, S., Chen, H., Chen, Y. et al. (2023) 'Hybrid method with parallel-factor theory, a support vector machine, and particle filter optimization for intelligent machinery failure identification', *Machines*, Vol. 11, No. 8, p.837.

Li, W., Qiao, J., Wang, Y. et al. (2022) 'Third harmonic potential distribution based stator ground fault accurate location method for large generators', *Energy Reports*, Vol. 8, No. 1, pp.21–28.

Liu, K., Zhang, S., Li, B. et al. (2021a) 'Flexible grounding system for single-phase to ground faults in distribution networks: a systematic review of developments', *IEEE Transactions on Power Delivery*, Vol. 37, No. 3, pp.1640–1649.

Liu, Z., Li, Y., Luo, Q. et al. (2021b) 'Harmonic amplitude pass probability: a lightweight signal descriptor for oxygen concentration detection in pharmaceutical glass vials', *IEEE Transactions on Instrumentation and Measurement*, Vol. 71, pp.1–11.

Meng, R., Zhang, J., Chen, M. et al. (2025) 'Fault diagnosis method of planetary gearboxes based on multi-scale wavelet packet energy entropy and extreme learning machine', *Entropy*, Vol. 27, No. 8, p.782.

Nauta, S. and Serra, R. (2021) 'Zero-sequence current measurement uncertainty in three-phase power systems during normal operation', *IET Generation, Transmission & Distribution*, Vol. 15, No. 24, pp.3450–3458.

Olejnik, B. (2021) 'Adaptive zero-sequence overcurrent criterion for earth fault detection for fault current passage indicators in resistor grounded medium voltage networks', *IEEE Access*, Vol. 9, pp.63952–63965, DOI: 10.1109/ACCESS.2021.3076022.

Palanisamy, S.K., Hajjej, F., Khalaf, O.I. et al. (2022) 'Discrete Fourier transform with denoise model based least square wiener channel estimator for channel estimation in MIMO-OFDM', *Entropy*, Vol. 24, No. 11, p.1601.

Ren, Z., Yu, K., Chen, G. et al. (2025) 'Estimation of primary current of DC bias current transformer in power system based on LSTM', *Processes*, Vol. 13, No. 11, p.3391.

Retter, T.L., Rossion, B. and Schiltz, C. (2021) 'Harmonic amplitude summation for frequency-tagging analysis', *Journal of Cognitive Neuroscience*, Vol. 33, No. 11, pp.2372–2393.

Rösel, K., Helm, M., Zirngibl, J. et al. (2021) 'Current developments of IEEE 1588 (precision time protocol)', *Network*, Vol. 13, No. 1, pp.1–5.

Shevgunov, T., Efimov, E. and Guschina, O. (2022) 'Estimation of a spectral correlation function using a time-smoothing cyclic periodogram and FFT interpolation – 2N-FFT algorithm', *Sensors*, Vol. 23, No. 1, p.215.

Sujiwa, A. and Maniani, B.S. (2025) 'The effect of inaccurate electronic component values on the output frequency characteristics of fourth order butterworth type low-pass filter circuits', *FARADAY*, Vol. 1, No. 1, pp.1–8.

Sun, Y., Cao, Y. and Li, P. (2022) 'Contactless fault diagnosis for railway point machines based on multi-scale fractional wavelet packet energy entropy and synchronous optimization strategy', *IEEE Transactions on Vehicular Technology*, Vol. 71, No. 6, pp.5906–5914.

Suo, N. and Zhou, Z. (2021) 'Computer assistance analysis of power grid relay protection based on data mining', *Computer-Aided Design & Applications*, Vol. 18, No. 2, pp.1–10.

Tang, Z., Zhou, K., Xu, Y. et al. (2023) 'A reflection coefficient estimation method for power cable defects based on three-point interpolated FFT', *IEEE Transactions on Instrumentation and Measurement*, Vol. 73, pp.1–8, DOI: 10.1109/TIM.2023.3343831.

Wen, C., Xiong, X., Wei, X. et al. (2024) 'The new method for ground insulation monitoring of the secondary circuit in substations', *Processes*, Vol. 12, No. 11, p.2351.

Xiao, X., Zhao, L., Zhou, S. et al. (2025) 'Harmonic and interharmonic measurement method using two-fold compound convolution Windows and Zoom fast Fourier transform', *Energies*, Vol. 18, No. 15, p.4047.