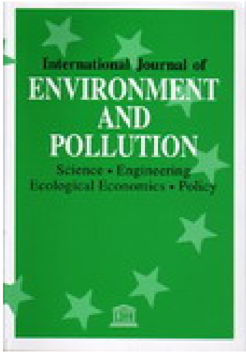


International Journal of Environment and Pollution

ISSN online: 1741-5101 - ISSN print: 0957-4352
https://www.inderscience.com/ijep

State diagnosis technology of metal enclosed gas insulation equipment based on Apriori algorithm in cloud computing environment

Jiayi Wang, Yuan Fang, Shaoqing Chen, Zongxi Zhang, Dianbo Zhou, Yuhang He, Jing Zhang

DOI: 10.1504/IJEP.2026.10075463

Article History:

Received:	22 July 2025
Last revised:	08 November 2025
Accepted:	03 December 2025
Published online:	24 March 2026

State diagnosis technology of metal enclosed gas insulation equipment based on Apriori algorithm in cloud computing environment

Jiayi Wang*

Electric Power Research Institute,
State Grid Sichuan Electric Power Company,
Chengdu, 610048, Sichuan, China

and

State Grid Southwest Electric Power Research Institute Co. Ltd.,
Chengdu 610072, China
Email: jiayi_wang2012@163.com
*Corresponding author

Yuan Fang

State Grid Sichuan Electric Power Company,
Chengdu, 610042, Sichuan, China
Email: 382888783@qq.com

Shaoqing Chen, Zongxi Zhang, Dianbo Zhou,
Yuhang He and Jing Zhang

Electric Power Research Institute,
State Grid Sichuan Electric Power Company,
Chengdu, 610048, Sichuan, China
Email: Chensq2126@sc.sgcc.com.cn
Email: Zhangzx3469@sc.sgcc.com.cn
Email: 13708225862@163.com
Email: 365373195@qq.com
Email: zhangjing2012@alu.cqu.edu.cn
*Corresponding author

Abstract: In recent years, with the increasing failure rate of gas-insulated switchgear (GIS), there has been a growing need for people to understand the most common insulator issues. Therefore, real-time monitoring of its operating status is crucial to ensure the safe and reliable operation of power lines. This study proposes the use of Apriori algorithm (CFSA-AA) for cloud based fault state analysis to predict discharge faults, mechanical faults, and abnormal mechanical vibrations in metal enclosed GIS. This data is sourced from the Kaggle repository used for VSB power line fault detection. This study summarises GIS abnormal heating faults, including circuit breakers, isolating switches, shell grounding, and disc insulator bolts. The experimental results show that compared with other existing models, the proposed CFSA-AA model improves the accuracy of fault diagnosis to 98.9%, pattern discovery rate to 97.4%, operating state detection rate to 95.6%, Mattew correlation coefficient ratio to 96.4%, and error rate to 7.4%.

Keywords: state diagnosis; metal enclosed gas insulation equipment; Apriori algorithm; cloud computing; fault detection; transmission lines.

Reference to this paper should be made as follows: Wang, J., Fang, Y., Chen, S., Zhang, Z., Zhou, D., He, Y. and Zhang, J. (2026) 'State diagnosis technology of metal enclosed gas insulation equipment based on Apriori algorithm in cloud computing environment', *Int. J. Environment and Pollution*, Vol. 76, No. 5, pp.77–99.

Biographical notes: Jiayi Wang received his Master's degree from Xi'an Jiaotong University, China. Now, he works as a Department Supervisor at the State Grid Sichuan Electric Power Research Institute. His research interests include state monitoring and fault diagnosis technology of electrical equipment.

Yuan Fang received his Master's degree from North China Electric Power University, China. Now, he works as a Vice Director in State Grid Sichuan Electric Power Company. His research interests include the technology of electrical equipment maintenance and operation.

Shaoqing Chen received his Doctorate from Xi'an Jiaotong University, China. Now, he works as Department Chief at the State Grid Sichuan Electric Power Research Institute. His research interests include state monitoring and fault diagnosis technology of electrical equipment.

Zongxi Zhang received his Doctorate from Wuhan University, China. Now, he works as Department Chief at the State Grid Sichuan Electric Power Research Institute. His research interests include state monitoring and fault diagnosis technology of electrical equipment.

Dianbo Zhou received his Master's degree from Chongqing University, China. Now, he works as a Department Supervisor at the State Grid Sichuan Electric Power Research Institute. His research interests include state monitoring and fault diagnosis technology of electrical equipment.

Yuhang He received his Master's degree from the University of Electronic Science and Technology of China, China. Now, he works as a Senior Engineer at the State Grid Sichuan Electric Power Research Institute. His research interests include state monitoring and fault diagnosis technology of electrical equipment.

Jing Zhang received his Doctorate from Chongqing University, China. Now, he works as a Senior Engineer at the State Grid Sichuan Electric Power Research Institute. His research interests include state monitoring and fault diagnosis technology of electrical equipment.

1 Introduction

For power utilities, ensuring the effective and dependable operation of power substations and their components is a difficult task. Widespread outages can result from any failure of substations, which are crucial for controlling the distribution of electricity (Peyghami et al., 2020). Gas-insulated switchgear (GIS), transformers, and circuit breakers all require routine maintenance and inspection. To control high-voltage levels and ensure a consistent electricity supply, these systems must operate flawlessly (Panmala et al., 2023). Errors must be quickly identified and corrected to avoid disruptions. For electricity substations to remain dependable and effective, modern technologies such as fault-detection algorithms, predictive analytics, and real-time monitoring are crucial.

Issues with the insulating systems of utility components often cause power outages. Power supply disruptions could occur directly or indirectly due to failures in these systems (Ghasemi et al., 2023). Preventing electrical malfunctions requires these systems. GIS, prized for its compact design and efficiency, is essential to modern ultra-high-voltage systems. GIS is crucial for monitoring and managing high-voltage electrical power in substations. Its compact design saves space and improves operating efficiency, but it is challenging to diagnose and maintain (Rozga et al., 2021). When a GIS unit fails, it can be challenging to access and repair internal components because of its enclosed design. This fundamental design makes troubleshooting more challenging and requires advanced diagnostic tools and methods to ensure the reliable operation of power systems.

Given the rising failure rate of GIS, it is more important than ever to comprehend common insulator issues. Insulators are essential for protecting GIS units from electrical failures, but they can deteriorate or fail over time (Purnomoadi et al., 2020). These failures have become more frequent recently, necessitating better knowledge and diagnostic techniques. A more reliable GIS operation can be ensured by identifying common issues, which can help predict and prevent malfunctions (Subramaniam et al., 2021). Knowledge gained makes it easier to create maintenance schedules, improve insulator designs, and advance monitoring systems that spot failures early.

Continuous monitoring of metal-enclosed GIS is necessary to guarantee the reliable and safe operation of the power line. GIS is an advanced electrical switchgear that uses a gas, typically sulphur hexafluoride (SF_6), to insulate its components (Han et al., 2022). GIS is small and effective due to its design, which also allows it to withstand high voltages. By enabling prompt detection of anomalies, such as unusual pressures or temperatures, continuous monitoring of GIS systems helps prevent minor issues from developing into more serious failures. Partial discharge evaluation, which entails identifying and examining small electrical discharges within the insulation, is necessary for monitoring (Hussain et al., 2021). Failure to address insulation issues, such as these leaks, could result in major interruptions. Because GIS units are sealed, insulation issues are common, making it difficult to identify faults and maintain them. Power utilities can mitigate GIS and electrical infrastructure failures they leveraging real-time monitoring to recognise insulation issues early (Faizol et al., 2023). By preventing unexpected outages and exorbitant maintenance costs, this improves the performance and safety of the electrical grid (Mahmoud et al., 2021). Current GIS fault-diagnosis technology faces a pair of bottlenecks: minimal data fusion across multiple sources and general inefficiencies in the process as a real-time task. Existing efforts tend to focus on single fault modes, which complicates the collaborative identification of multiple faults, such as discharge, mechanical, or vibration faults, under operational conditions. This research proposes a novel unified holistic space diagnostic of multi-fault types by pairing the elastic resource schedule capacities of cloud compute platforms to design a simple and integrated process that automates the integration of massive heterogeneous data collections with strong association rule mining utilising the Apriori algorithm to manage a whole process from data fusion to status monitoring conclusions.

Predicting abnormal mechanical vibrations, mechanical faults, and discharge faults in metal-enclosed GIS (Dai et al., 2023). The Apriori algorithm's capacity to identify patterns and correlations enables the detection of errors or malfunctions in operational data (Zixian et al., 2021). By analysing data for recurring patterns and relationships, the algorithm can identify anomalies that may indicate systemic problems. By detecting and addressing issues early, a predictive approach can improve the reliability of the power infrastructure (Dhinakaran and Prathap, 2022). Due to cloud storage's scalability, power utilities can effortlessly manage enormous amounts of historical and current data. Since it guarantees that even the most significant data streams can be handled and stored efficiently, adaptability is crucial (Mehmood and Anees, 2020). To continuously monitor and analyse operational data, power utilities may use cloud storage for long-term trend analysis and real-time diagnostics. This capacity is necessary to keep the electrical infrastructure reliable and efficient (Awaysheh et al., 2021). Because the cloud can scale up or down in response to data volume demands while maintaining cost-effectiveness and resource optimisation, it is well-suited for continuous power system monitoring and defect detection.

This paper presents a cloud-based fault-state analysis method (CFSA-AA) based on the Apriori algorithm, designed explicitly for fault diagnosis of GIS. This method integrates historical and real-time operational data via a cloud computing platform, leveraging pattern recognition technology to identify discharge, mechanical, and vibration-related faults efficiently. This significantly improves fault trend prediction capabilities and detection accuracy, effectively overcoming the shortcomings of traditional techniques.

The CFSA-AA system demonstrates superior performance during operation, with its fault-state recognition rate, pattern-mining efficiency, and diagnostic accuracy exceeding those of existing solutions, while significantly reducing the risk of false alarms. Leveraging the advanced data processing capabilities of the cloud platform, this method provides a more reliable and efficient approach to early fault warning and handling for GIS equipment.

An overview of the research follows. Section 2 describes a comprehensive literature and research methodology review. Section 3 describes the research plan, methodologies, and processing. Section 4 describes the results of the analysis. The conclusion and future work are in Section 5.

2 Literature survey

2.1 Design, operation, and challenges of gas-insulated switchgear (GIS)

Zebouchi and Haddad (2020) reported that Remote renewable energy sources, such as offshore wind farms and hydropower plants, are likely to increase globally. High-voltage direct current (HVDC) delivers large amounts of power over long distances with low loss. Current innovations include small HVDC gas-insulated transmission lines (GIL) and switchgear. Still, developing epoxy-cast resin spacers, or insulators, is difficult. This study discussed basic electric-field calculations, spacer design, insulation performance verification, constraints, and the fabrication of full-scale HVDC spacers. Real-size HVDC spacers have been studied little, restricting their practicality.

Xing et al. (2022) examined GIS insulator faults and failure types, classifying them as production-related, assembly-related, and operation-related errors caused by insulation deterioration. Engineers and scientists researching GIS solid-insulation faults and failure processes will find this paper helpful; it addresses the rising failure rate of GIS due to insulator failures. The study discussed the need to dynamically monitor the internal environment or identify faults to reduce the frequency of failures. Upon failure, the study cannot confirm the insulator because of a lack of direct proof.

Mahdi et al. (2022) examined more than 100 studies on partial discharge (PD) monitoring and diagnosis in GIS. It highlights four points: methods for detecting PD, sources of PD, evaluation of PD severity, and PD diagnosis. The paper addresses gas pressure, applied voltage, and contaminants as potential contributors to the degradation of insulating gas. According to the review, a GIS can detect PD and its causes when there is a positive association between SF6 breakdown byproducts and PD activity. According to the report, combining PD detection approaches in a GIS could increase performance and overall GIS monitoring systems.

2.2 Types, causes, and detection methods of insulation faults in GIS

He et al. (2021) investigated the diagnosis and detection of mechanical defects in GIS based on vibration signals (DDMD-GISVS). This study compares and contrasts the structural features of three-phase and single-phase insulated GIS by analysing their vibration processes. Research using a 110 kV GIS and an in-house vibration identification model showed that signal amplitude, waveform distortion rate, and frequency spectrum can be used for fault diagnosis. It is better to use a large current to identify faults. A negative contact fault was observed during field testing on a 220 kV GIS, demonstrating that the suggested solutions are feasible and practical. The limitations of the study will centre on the use of AI for mechanical problem diagnostics.

Zhuang et al. (2020) described GIS, which is essential for electrical safety but has many flaws. Machine learning-based SF_6 deconstructed components analysis (DCA) for discharge fault diagnosis is presented to increase GIS system safety. The Arrhenius chemical reaction model examined the empirical probability distribution of faults. Six machine learning methods were used to model discharge fault and main insulation defect severity. A probability model discovers numerous discharge fault types in GIS. The model identified new states and provided more sample data, information on insulation problems, and transparency into the SF_6 decomposition mechanism. The probability model for GIS fault evaluation would be more reliable.

2.3 Cloud computing applications in power system monitoring

Hashmi et al. (2021) presented a smart grid that controls power production, transmission, and distribution via digital communication. Traditional systems typically overlook energy customers, resulting in wasted energy and money. Energy savings in underdeveloped nations depend on demand-side management (DSM). EMS based on cloud computing and IoT is presented. Such technologies provide utility companies with consumer load profiles to regulate and distribute incentives. A project circuit board (PCB) with the EMS uploads data to Google Firebase to generate DSM load profiles.

AL-Jumaili et al. (2023) examined the recent literature on power management and battery charging control, focusing on CC-based power optimisation systems (CC-POS) for hybrid renewable power systems. The study monitors battery state of charge (SoC) and extends battery life. The study critically analyses 174 research papers and highlights significant findings and future directions. The study examined battery management, discharge/charge control, and the role of cloud computing in battery life and charging percentage. It promotes energy conservation and innovative use of renewable and other power sources. It is advised that sophisticated BMS algorithms, optimal cell-balancing models, and a digital twin method for real-time optimisation of battery storage systems be developed. The study stressed intelligent power systems and battery storage systems.

2.4 Apriori algorithm for fault detection and pattern recognition

Zhao and Huang (2024) demonstrated that applying the Apriori big data analysis and calculation approach (ABDAC) to power system operation data can enhance power systems by analysing transmission line failures. The candidate item set calculation time is reduced by 52% through optimisation of the traditional Apriori calculation method. The real-time analysis of the association between event index data and defects using the M-Apriori calculation approach enables automatic fault diagnosis. Using the Apriori algorithm, the study develops a framework for thorough data exploration that improves performance by 15%. The study's scope is limited to enhancing power data integration for completeness and uniformity.

Hassan et al. (2023) attempted to discover the key drug addiction rules using an open-source dataset of 474 occurrences and 212 people with an addiction. Eight essential rules were extracted using the Apriori technique. Personal concerns like smoking and thievery were more likely to induce addiction. The study seeks to analyse drug addiction and offer treatment recommendations to families and the government. However, no direct application approach predicts drug addiction rates. Limitations could employ data and intelligence systems to identify and address harmful policies.

2.5 State diagnosis technologies for metal-enclosed gas insulation equipment

Mier et al. (2024) presented a method for differentiating partial discharges from gas-insulated system interference (DPD-GISI). Four approaches are used to determine the characteristic impedance utilising high-frequency electric and magnetic sensors: peak value, PD charge magnitude, frequency spectrum, and peak-to-peak value. A PD calibrator in a coordinated GIS testbed and a full-scale GIS subjected to pulse overlap and noise were used to test the technique. Characteristic impedance may be determined using the four methods in a coordinated test bench, although the peak method yields the most accurate results. The technology enhances the monitoring and maintenance planning for HVDC GIS PD, helping reduce costs, particularly for offshore substations.

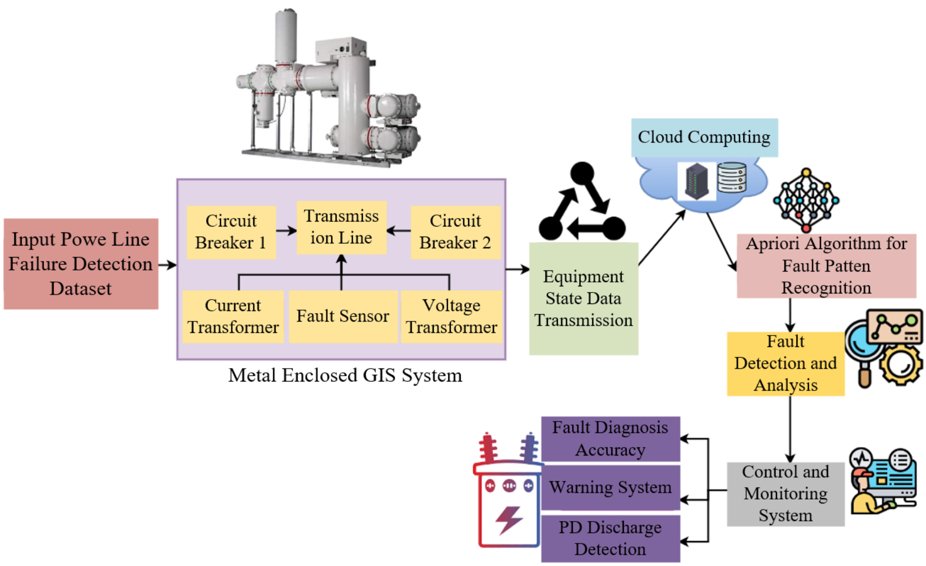
Sun et al. (2023) investigated partial discharge in GIS and GIL using three standard defect models. According to the study, air-gap faults and partial discharge of free-metal particles increase gradually with voltage. The partial discharge from the floating potential defect increased suddenly. While the other two faults were located on one side and produced a phase variance, the discharge quantity of free conductive particles, as determined by UHF identification, was evenly distributed on both sides of the smeared voltage peak. Despite ultrasonic particle-collision signals, there was no correlation between voltage and its phase. The integrated technique may better identify the three common issues using theoretical underpinnings and data from GIS and GIL voltage testing, online monitoring, and detection technologies.

3 Proposed methodology

Due to its small size, low maintenance requirements, and high reliability, GIS with metal enclosures has been widely used. Inevitable insulation flaws that develop during production, transit, assembly, or operation may lead to insulation deterioration or failure. Partial discharge (PD) is a condition that occurs before the insulating system fails. As a result, PD measurement is often used for defect detection and to prevent severe insulation failure. The ultra-high frequency (UHF) technique has seen extensive field application for GIS PD detection due to its strong anti-interference performance. However, it is still not easy to tell, particularly during field testing, whether UHF detection findings are due to real PD signals. Power delivery grids nowadays face numerous challenges. Firstly, there is a constant increase in urban and suburban areas, which means less space for power facilities. Secondly, there is a constant quest for better power quality and availability. Lastly, there are increasing concerns about the visual impact of high-voltage systems and their electromagnetic compatibility. However, due to its excellent insulation performance, GIS is expensive and still suitable for these applications. Conductors supported by solid insulators inside an enclosure filled with sulphur hexafluoride (SF_6) gas make up gas-insulated gearbox lines (GITL) and GIS. Hence, this study proposes the Cloud-enabled Fault State Analysis using the Apriori Algorithm (CFSA-AA) for predicting discharge faults, mechanical faults, and abnormal mechanical vibrations in metal-enclosed GIS.

Figure 1 shows the proposed CFSA-AA model. The data are from the Kaggle VSB power line failure dataset. This system's components are important, including fault sensors, voltage transformers, current transformers, and circuit breakers. At the same time, circuit testers (CTs) measure voltage levels, and circuit breakers (CBs) stop current flow in the event of a failure; they are an essential component in protecting the system. A fault sensor system is installed in the gearbox line. These sensors are positioned to detect any anomalies that may indicate a malfunction. This system collects raw data from CTs (current transformers), VTs (voltage transformers), and fault sensors. After local unit analysis, normalisation, and noise reduction preprocessing, the data is uploaded to a cloud database to ensure efficient and scalable storage. An Apriori algorithm module is deployed in the cloud that, based on frequent itemsets and association rules mined from GIS data, automatically identifies recurring patterns and their correlations to support anomaly detection and fault tracing. The system compares current data with standards to identify faults accurately. Finally, the visualisation module presents the analysis results in the form of dashboards and charts, helping maintenance personnel quickly grasp the fault status and impact. The overall architecture integrates cloud storage, intelligent analysis, and intuitive display to achieve intelligent monitoring and decision support for equipment status. In the event of a malfunction, this system is responsible for alerting the GIS system and instructing the circuit breakers to apply a fix. Additionally, it has an easy-to-use interface that system managers can use to monitor the system's state, view the outcomes of in-depth studies, and respond immediately when required. Through cutting-edge algorithms and cloud computing, this all-encompassing approach enhances the reliability and security of GIS transmission lines while simultaneously enabling efficient data management and discovery.

Figure 1 Proposed CFSA-AA model (see online version for colours)



3.1 *Numerical computation of electromagnetic and temperature field*

Joule heat loss due to electromagnetic fields acting on the GIL conductors and enclosures is used as the heat source for the temperature field calculation. The current flowing through the conductors is perpendicular to the current circulating through the enclosure. The solution domain medium features are connected to the electric fields E and the magnetic flux A in GIL. Consequently, expanding Maxwell's equations is required for describing the medium's properties.

$$\begin{cases} D = \varepsilon_0 \varepsilon_s E = \varepsilon E \\ A = \mu_0 \mu_r G = \mu G \\ \quad I = \rho E \end{cases} \tag{1}$$

As shown in equation (1), where ε denotes the dielectric constant, ρ symbolises the electrical conductivity, and μ represents the magnetic permeability. Because enclosure grounding acts as a shield, the GIL conductor's coefficient of proximity effect is one, and its impedance is modest, both of which contribute to a low Joule heat loss calculation.

As a result, it may disregard the influence of the imbalanced current and focus only on the skin effect.

An electric field is produced when alternating current (AC) flows through a conductor, which induces eddy currents in other conductors and creates a magnetic field similar to that of the AC magnetic field. Consequently, losses are amplified because currents are confined to the conductors' surfaces. Other neighbouring conductors impact the distribution of current and energy density. Here, the following form is taken by Maxwell's equations, which explain the relationships between the field quantities:

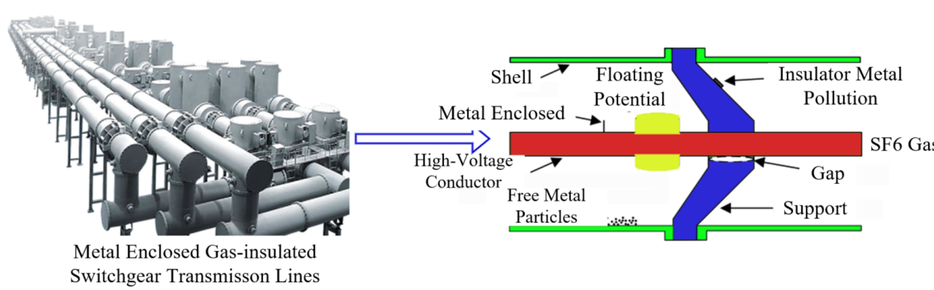
$$\nabla \times G = \rho E; \nabla \times E = -\mu \frac{\partial G}{\partial t} \tag{2}$$

$$\nabla \cdot E = 0; \nabla \cdot G = 0 \tag{3}$$

As found in equations (2) and (3), where G denotes the magnetic field strength, E indicates the electric field strength, ρ represents the material's electrical conductivity, μ denotes its magnetic permeability, $\nabla \times$ symbolises the curl, and ∇ denotes divergence operators. The solutions of these expressions yield the loss distributions.

Figure 2 illustrates the insulation fault type for GIS internal circuits. One of the most common reasons insulation fails is flaws in the GIS, namely, free-metal particles. Free metal particles can be released into the environment when metal components of a GIS rub against one another during installation, operation, or assembly. Because of their small size, these metallic particles will be propelled by the electric field and vibrate. Conductive channels or arc tunnels may be formed between the HV conductor and the shell if the particle movement range is adequate, which can severely damage the GIS. Several variables influence the path-forming process, including the applied voltage, particle size and shape, and particle location.

Figure 2 Insulation fault type illustration for the GIS internal circuits (see online version for colours)



This study treats insulating gas and air as incompressible, viscous fluids because their velocities are much slower than the speed of sound, and gas has difficulty absorbing and emitting radiation. The gas is thought to be transparent when subjected to heat radiation. In a one-phase GIL, the heat loss is uniformly distributed throughout, and the enclosure and conductor have isotropic compositions. Given that the gas's density changes with temperature, the GIL insulating gas is naturally convective. Since the gallery wall is the airfield's exterior boundary, it is unaffected by GIL heating. A conductor or enclosure with a length of 1 m may have its resistance represented as

$$R_j = L_f \frac{\sigma_m(T)}{W_j} \tag{4}$$

$$\sigma_m(T) = \sigma_{20}\left[1 + \beta_{20}(T - 293.15)\right] \tag{5}$$

As inferred from equation (8), where R_j denotes the resistance, L_f indicates the coefficients of skin effect, $\sigma_m(T)$ signifies the material resistivity, W_j represents the cross-section region, σ_{20} symbolises the material resistivity at 293.15 K, β_{20} specifies the material temperature coefficients of resistances at 293.15 K, and T denotes the thermodynamic temperatures.

Equations (6) and (7) may represent the Joule heat loss per conductor and per unit volume of the enclosure, respectively.

$$Q_{du} = \frac{J_d^2 R_d}{W_d} \tag{6}$$

$$Q_{lu} = \frac{J_l^2 R_l}{W_l} \tag{7}$$

As shown in equations (9) and (10), where J_l indicates the induced current of the enclosures, Q_{du} and Q_{lu} are the Joule thermal energy per unit volume of the conductors and enclosures, respectively, and J_d denotes the current of the conductors. Once the single-phase GIL length is less than 20 m, J_l = 0.95Id; else, $J_l = J_d$. R_d and R_l denote the resistance values of the conductors and enclosures determined via equation (4). W_d and W_l represent the cross-sectional areas of the conductors and enclosures, respectively.

Consideration of simply convective heat transfer and thermal radiation is required when the heat created and exchanged is the same. This is the moment when GIL's temperature is steady. The following equations so characterise a two-layer thermal equilibrium relationship:

$$Q_{du} + Q_{lu} = P_{lF} + P_{lD} \tag{8}$$

$$Q_{du} = P_{dF} + P_{dD} \tag{9}$$

As found in equations (8) and (9), where P_{lF} denotes the thermal radiation of the enclosures dissipating heat, P_{lD} indicates the heat dissipated by normal convection in the enclosures, P_{dF} represents the thermal radiation of the conductors dissipating heat, *and* P_{dD} indicates the heat dissipated by natural convection in the conductors.

The control-volume calculation of heat conduction provides the primary expression of heat transfer in enclosures and conductors. The heat transfer differential calculation must account for time when examining the temperature changes caused by changes in the GIL's ampacity, as the ampacity is a function of the period. This study may express the governing equation for GIL heat conduction as follows:

$$\sigma_T C_q \frac{\partial T}{\partial t} = \lambda_T \nabla^2 T + P \tag{10}$$

As inferred from equation (10), where σ_T represents the material densities of conductors or enclosures at T, C_q denotes the particular heat capacities at constant pressure, P indicates the thermal losses, λ_T signifies the thermal conductivity at T, and t is the run-period.

Assuming that mass, momentum, and energy are conserved, the following differential control equations describe the coupling connection between the temperature and insulating gas flow fields.

$$\frac{\partial \sigma_h}{\partial t} + \sigma_h \nabla \cdot u = 0 \tag{11}$$

$$\sigma_h \left(\frac{\partial u}{\partial t} + u \cdot \nabla u \right) = \eta_h \nabla^2 u + E \tag{12}$$

$$\sigma_h C_q \left(\frac{\partial T}{\partial t} + u \cdot \nabla T \right) = \lambda_T \nabla^2 T \tag{13}$$

As discussed in equations (11)–(13), where σ_h signifies the gas density, u indicates the gas flow rates, η_h denotes the aerodynamic viscosities, and E implies the generalised source terms.

The Joule or heat-loss term from the electromagnetic-field model is a heat source, f_g, in thermal or fluid-flow models. The principle of energy conservation must be satisfied by the second model. The internal energy of a gas confined to a volume can change due to gas-particle motion and heat conduction, since the kinetic and thermal energies of particles entering and exiting the volume can be adjusted. Shear stress affects the internal energy, further complicating the issue, and viscous friction, characterised by dynamic viscosity η, adds to the complexity. If the flow rate is low, it is safe to assume the flow is incompressible and omit this step. The energy equation is provided by

$$\nabla \frac{\partial}{\partial t}\left(c_q \sigma T\right) + \nabla \cdot \left(u c_q \sigma T\right) + \nabla \cdot \left(-\lambda \nabla T\right) = f_g \tag{14}$$

As shown in equation (14), where ρ denotes the density, c_q signifies the particular heat, λ indicates the heat conductivity, T symbolises the temperature, ∇ denotes the gradient operator, and u denotes the flow velocities. The governing equations for convective heat transmission must adhere to the concept of mass conservation. Based on the premise that chemical processes can neither create nor destroy matter, this research concludes that mass movement is the only possible source of local density fluctuation.

$$\frac{\partial \sigma}{\partial t} + \nabla \cdot \left(\rho u\right) = 0 \tag{15}$$

The flow of mass, as in momentum transfer, is necessary to satisfy conservation of momentum. In our studies, the momentum transfer is not free of sources; gas motion is induced by the buoyant forces acting on a gas particle of varying density. In the momentum equation, the acceleration due to gravity (g) and the density gradient (d) are the two variables that contribute to the initial momentum.

$$\frac{\partial(\sigma u)}{\partial t} + \nabla \cdot (\sigma u \otimes u) = -\nabla q + \eta \nabla^2 u + \sigma h \tag{16}$$

As inferred from equation (16), where $\otimes$ denotes the dyadic products, and q represents the pressure. Both the outside air and the inside SF_6 exhibited turbulent flow. The above equations must include additional variables in turbulent flow, which challenges the precision of the solutions. Various probabilistic approaches have been devised to streamline the computations. All of them result in the earlier versions of the calculations that describe laminar flow, yet with varying interpretations of the coefficient. They provide average heat-conductivity and viscosity values, along with a turbulent component. This research used the conventional k-ε turbulence model in its simulations. In our simulations, radiation manifested itself as distinct boundary conditions. This research examined a busbar setup that used high-voltage SF_6 gas insulation. The busbars were circular and located inside a cylindrical enclosure.

3.2 Apriori association rule mining for fault pattern recognition

Association rules are specified as $Y \Rightarrow X$ $(Y \subset j, X \subset j, y \cap z = \varphi)$. Y denotes the preceding rule, and x is the later rule item. The progression of associated rules refers to the expression $Y \Rightarrow X$ in the affairs libraries. The transaction database will comprise unlimited associated rules if efficient limits are not set. The support equation is provided in equation (1),

$$Support(Y, X) = \frac{|T(Y \cup X)|}{|T|} \tag{17}$$

As shown in equation (17), where the power line database comprises the number of transactions of $Y \cup X \mid T(Y) \mid, T(Y \cup X)$ is the overall number of transactions in the transaction libraries. The Confidence replicates the intensity of the associated rule $Y \Rightarrow X$. The calculation of Confidence in equation (2).

$$Confidence(Y \Rightarrow X) = \frac{|T(Y \cup X)|}{|T|} \tag{18}$$

As discussed in equation (18), where $|T(Y \cup X)|$ denotes the number of transactions containing $Y \cup X$ in the transaction databases, |T(Y)| denotes the number of transactions comprising Y in the database.

The lifting level specifies the association rule validity $Y \Rightarrow X$. When it is less than 1, association rules $Y \Rightarrow X$ has a disability; when it is higher than 1, association rules $Y \Rightarrow X$ is valid, and the higher the values, the tougher the validity; when it is equivalent to 1, it denotes that Y and X are independent of every other. The computation is in equation (3).

$$Lift = \frac{cofidence(Y \Rightarrow X)}{|T(Y)|} \tag{19}$$

As inferred from equation (19), where $|T(Z)|$ denotes the number of transactions that comprise Z in the databases.

The support of the itemset Y is at least the minimum support, and Y is termed a frequent itemset. If Y denotes frequent l-item sets, then Y is called a frequent l-item set. The minimum support is the lowest mathematical threshold for an itemset, and it must hold that the support of every item is at least as high as the minimum support. Minimal Confidence (minconf) is responsible for verifying the correctness of the association rules and determining the minimum Confidence. In the Apriori process, the minimum support is the lowest statistical threshold for item collections. Measuring the association is the minimum Confidence, and the frequent item in the association rule must have support of at least min-support.

Given that the accuracy of association rules depends on the reliability of the rules, the final computation results in practical applications are affected by the choices of minimal support and confidence. Currently, the Apriori method relies on artificially determined minimum support and confidence values to generate rules; determining which rules are applicable requires extensive experimentation across many configurations and diverse combinations. This will have various outcomes. This study employs weight coefficients to establish threshold functions for the minimal support level and the minimum confidence level for operating-state and fault detection in metal-enclosed GIS, thereby mitigating this challenge, reducing time, and enhancing the algorithm's effectiveness.

$$\begin{cases} f = \sum_{j=1}^{m} \lambda_j \cdot q_j \\ \sum_{j=1}^{2} \lambda_j = 1 \ \lambda_j > 0 \end{cases} \tag{20}$$

There are several significant benefits to using the Apriori algorithm to find errors in GIS systems: It is Important for identifying patterns of errors, and it excels at finding common itemsets and relationships in large datasets. It is well-suited for intricate GIS systems due to its ability to manage massive amounts of operational data. The produced rules are straightforward and helpful for troubleshooting. Modifying it to include domain-specific knowledge is a breeze. Because it iteratively narrows the search space, it uses fewer processing resources. It remains stable even in the face of minor data fluctuations, which are typical in GIS measurements.

3.3 Cloud computing for the fault warning system of GIS

First, in assessing the cloud models, let Y be a quantitative dataset, i.e., $Y = \{y\}$. This quantitative database is called the universe of discourses. C denotes a qualitative notion on the universe of discourse Y, and if $y \in Y$ is a random map of $\mu(y) \in [0,1]$, then the distribution of y on Y is termed a cloud, signified as $C(y)$. $\mu(y)$ is the cloud y of membership to C, and every $(y, \mu(y))$ is termed a cloud drop, signified as a drop $(y, \mu(y))$. The respective statistical association for fault diagnosis accuracy is as follows.

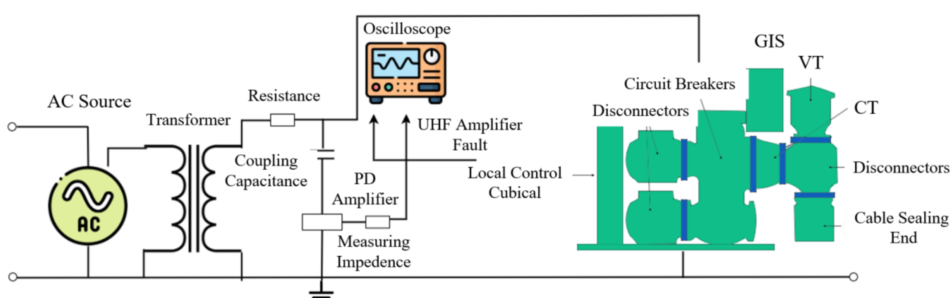
$$\mu : Y \to [0,1], \forall y \in Y, y \to \mu(y) \tag{21}$$

The evaluation cloud for each regulation layer index can be determined by entering the details of the indicators (such as ground conductors, foundation, towers, grounding fittings, insulator strings, auxiliary facilities, grounding devices, meteorological environment, channel environment) that primarily impact the line state of every regulation layer into the forward cloud model's digital cloud generator and setting the number of clouds to $n = 2300$. The following features of the cloud model become apparent throughout its build: First, cloud droplets are an ad hoc result of converting qualitative ideas into quantitative data; hence, creating cloud droplets is an example of the two types of ideas adapting to one another. The approach exemplifies a lack of clarity and predictability. To elaborate on the second point, the generated cloud becomes more precise at expressing the qualitative idea as the number of cloud droplets increases. A digital evaluation cloud was created that included the entire GIS operating status. Using the digital element and a setting of 2300 for the number of clouds, a thorough assessment cloud for the transmission line failure warning system was constructed.

3.4 Experimental arrangements

Figure 3 shows the partial discharge detection in a metal-enclosed GIS. Two signal output channels are available from the integrated sensor. There are two signals: the optical signal and the ultra-high-frequency signal. The optical signal is converted to a voltage signal employing a photoelectric transducer, and a 20 dB gain device amplifies the UHF signal. Then, a digital oscilloscope, the TEK4054, with a bandwidth of 500 MHz and a sampling rate of 2.5 Gsamples/s, is used to capture these two signals. The GIS high-voltage conductor is designed to generate an electric field concentration using a 26 mm long metallic protrusion as the flaw. The high-voltage conductor, which is not connected to the sensor, is the source of the flaw. At a voltage of 53 kV, the PD signals first show up in the optical channel. For the moment, UHF signals are unavailable. After reaching 69 kV, the PD signals are sent by optical and UHF channels. The optical approach yields a PD inception voltage of 53 kV, whereas the UHF method yields 69 kV. Compared with the UHF approach, the optical technique exhibits superior sensitivity. Different methods generate optical and UHF signals.

Figure 3 The partial discharge detection in metal-enclosed GIS (see online version for colours)



Fluorescence fibres pick up optical signals from photons emitted during electron attachment, while UHF signals are proportional to the rate and amplitude of discharge current changes. Both collision ionisation discharge and streamer discharge are methods for SF6 discharges. As the applied voltage increases, collision ionisation and steam discharge occur in that order. A dipole antenna, chosen for its wide bandwidth and exceptional HF response, is used to receive the PD signal in this investigation. It shows the attenuation properties recorded by the spectrum analysers (Agilent CSA) in the GTEM system, which stands for Gigahertz Transverse Electromagnetic. The GTEM system inputs the output of the spectrum analysers. A sequence of sine waves spanning 10 MHz to 2.5 GHz makes it up. The spectrum analyser takes the signal the antenna has picked up as input. Antenna attenuation characteristics and frequency response can be derived by comparing the input and output spectra of the spectrum analyser.

The relationship between attenuation and frequency response is well-known to be positive. The frequency response is good beyond 300 MHz, as shown. Thus, the dipole antenna is a suitable choice for the experiment. In this trial test, the efficient distance between the antennas and the PD test cells was maintained at 1 m, well below the typical value of 5 m or more. The experiment reveals that the antenna captures the electromagnetic wave in a plane and that the dipole antenna's orientation affects its effectiveness. It shows that the dipole antenna's orientation affects the measurement outcomes. Thus, according to antenna theory, a dipole antenna is best for detecting partial discharges when placed vertically, contrary to the direction of discharge propagation. The HF signal from the PD is collected using an Agilent DSO9254A digital oscilloscope with a 2.5 GHz bandwidth and a 10 GS/s sample rate. The sensor and oscilloscope are linked via the well-shielded wire. An HV power source rated up to 100 kV is connected to the PD test cell.

The cloud-based architecture of the CFSA-AA model enables it to process data in real time and detect faults. The Apriori algorithm is used to process data continuously collected from GIS sensors in the cloud. As a result, massive data streams can be analysed efficiently and scalably. The model processes incoming data in real-time and detects fault patterns by comparing it with learned association rules. It enables preventive maintenance by rapidly detecting anomalies and predicting probable failures. Real-time power system problem detection relies on the cloud's high availability and fast response times.

Based on the facts, the CFSA-AA methodology may apply to electrical systems other than GIS. Cloud computing for data storage/processing, the Apriori algorithm for pattern identification, and real-time sensor data analysis may apply to various power system components that create operational data. However, various equipment may require distinct defect kinds and detection procedures. More research is needed to confirm its efficacy on different electrical systems.

CFSA-AA, or Cloud-enabled Fault State Analysis using the Apriori Algorithm, is the proposed system for GIS. It connects circuit breakers, voltage and current transformers, and fault sensors by integrating with cloud-based infrastructure. The system collects data from these sources, preprocesses it locally, and then sends it to the cloud for analysis. It uses the Apriori method to identify potential errors by analysing data patterns. The system enhances the effectiveness and reliability of GIS by including an interface for real-time monitoring and visualisation, as well as simple defect reporting.

The proposed study enhances GIS failure detection by utilising the Apriori Algorithm in conjunction with Cloud-enabled failure State Analysis. This technique resolves issues with mechanical fault detection and partial discharge. The study enhances data processing and visualisation by leveraging cloud computing and sophisticated algorithms, potentially improving system reliability and defect identification. The methodology's comprehensive and flexible nature promises to improve GIS performance and maintenance, despite the high investment costs and challenging data processing.

In the field of fault diagnosis for GIS, this study proposes a cloud-enabled Apriori algorithm-based fault state analysis model, which is systematically compared with current mainstream technologies. Some existing methods focus on offline diagnosis or theoretical reviews of specific fault types, failing to construct a comprehensive framework that covers multiple fault types and supports real-time data stream processing. While some methods can perform specific functions, they have limitations in real-time performance, depth of data integration, or environmental adaptability. In contrast, this model, by integrating a cloud computing architecture with association rule mining algorithms, achieves simultaneous prediction and state monitoring of discharge faults, mechanical faults, and abnormal vibrations, demonstrating systematic advantages in diagnostic accuracy, pattern recognition efficiency, and real-time response capabilities.

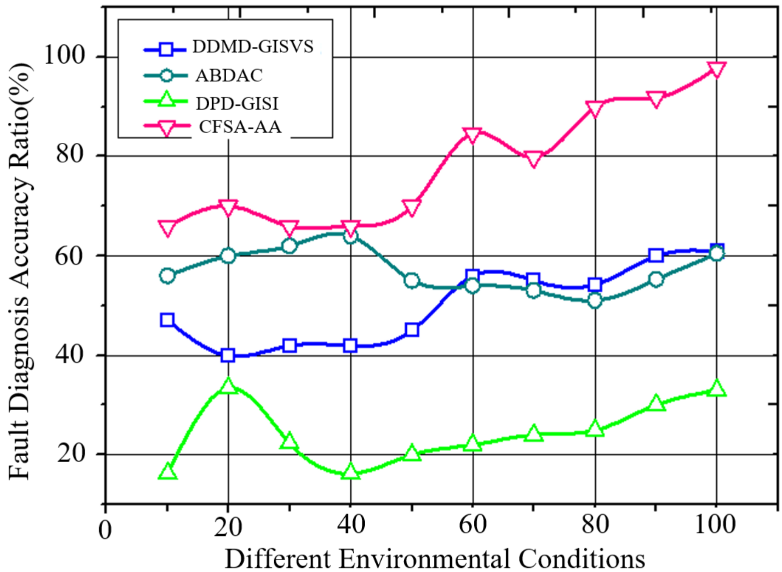
4 Results and discussion

To predict discharge faults, mechanical faults, and abnormal mechanical vibrations in metal-enclosed GIS, this study introduces the CFSA-AA. The data are from the VSB Power Line Failure Detection Kaggle Dataset. The terrible occurrence known as partial discharge may occur when electric transmission cables fail. Partially discharged equipment might become completely inoperable if not addressed promptly. To prevent partial discharges from causing permanent damage, they must be identified. Each signal is 800,000 voltage measurements taken over 20 ms from a power line. Since the underlying electric grid operates at 50 Hz, every signal encompasses one whole grid cycle. For the grid to function, all three phases of electricity must be measured simultaneously. 800,000 into eight measurements, as exported with yarrow.parquet version 0.11. The model's performance is enhanced by an extensive experimental setup that includes MATLAB R2024a, a powerful Intel Core i7 CPU with 32 GB of DDR4 RAM, and a high-performance NVIDIA GeForce GTX 1650 GPU. Fast data processing, precise problem diagnosis, and trustworthy pattern discovery are all guarantees of this cutting-edge software and hardware combo. The CFSA-AA model's overall performance in fault detection is greatly improved by the well-chosen setup, which increases the model's capacity to handle complicated calculations and big datasets.

4.1 Fault diagnosis accuracy

This study integrates the power grid's underground transmission line fault with an a priori algorithm to implement an intelligent monitoring system for underground transmission lines. It studies the power grid's underground transmission line problem by collecting real-time data and analysing it to detect faults. The objective is to stabilise the power system's functioning and enhance its intelligent monitoring function to better diagnose and analyse faults in real time. The study provides a thorough analysis of the clustering techniques used and explains the algorithms in detail. Partial discharge measurements facilitate online condition monitoring and problem detection, a crucial technique for assessing dielectric integrity. Decomposition component analysis (DCA) has shown that ratios of SF_6's key decomposition components can be used to detect power equipment failures. Based on equation (21), fault diagnosis accuracy has been identified. Figure 4 shows the fault diagnosis accuracy ratio.

Figure 4 Fault diagnosis accuracy ratio (see online version for colours)



4.2 Pattern discovery ratio

A pattern of discontinuous impulse currents with a maximum magnitude of 475 A was observed. The timing of the pulse's appearance is directly related to the polarity of the voltage. Learning how the phasor-measuring unit (PMU) recorder system produces such unusual waveforms is highly relevant, given the peculiar nature of this disaster. There must be a two-step analysis since the collected data showed a distinct pattern. The root of this problem cannot be determined without doing further research and computer simulations. Consequently, the discharge zone is discontinuous due to the negative and positive half-cycle partitioning throughout the cycle in the conventional phase-resolved PD fault pattern. This study shows how to generate PD scales in glacial coordinates and get the phase-resolved PD analysis pattern in polar coordinates. Improved defect type identification is possible using the apriori technique with the new discharge variables defined for each cluster derived from the novel coordinate pattern. A fault pattern has been identified using equations (17)–(19). Figure 5 shows the pattern discovery ratio.

4.3 Operation state detection ratio

There exists a connection between the operating condition of the power switchgear and the quality of the power supply provided by power utilities. The sensors and bolts were maintained in a satisfactory state of tightening to receive the longitudinal waves from the GIS. At the same time, coupling agents were used to enhance the matching qualities of vibration waves. This study aims to develop a cubicle-insulated switchgear diagnosis and monitoring model. In cubicle GIS, this system may remotely measure and analyse the gas region's degradation, as well as the operating state and integrity of a breaking unit, enabling diagnosis of the gear's current state. Based on equation (20), the operation state has been detected. Figure 6 shows the operational-state detection rate.

Figure 5 Pattern discovery ratio (see online version for colours)

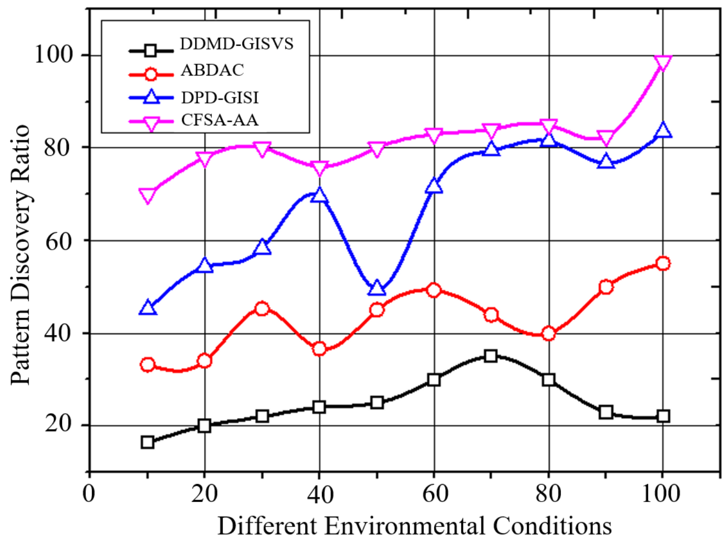
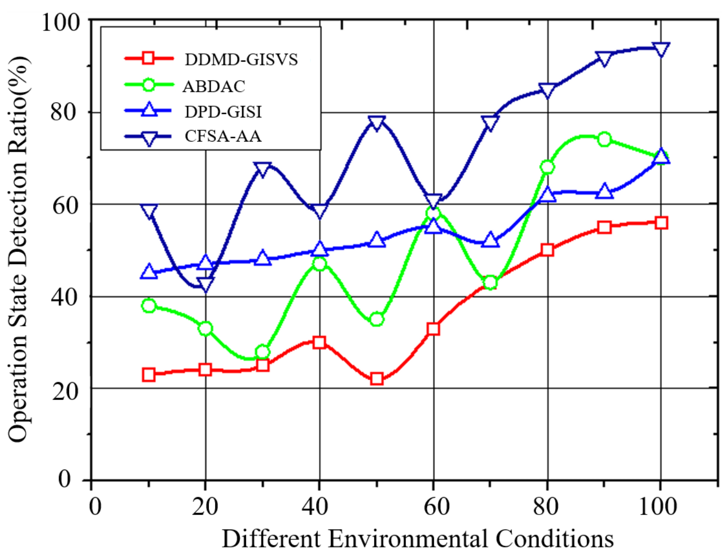


Figure 6 Operation state detection ratio (see online version for colours)



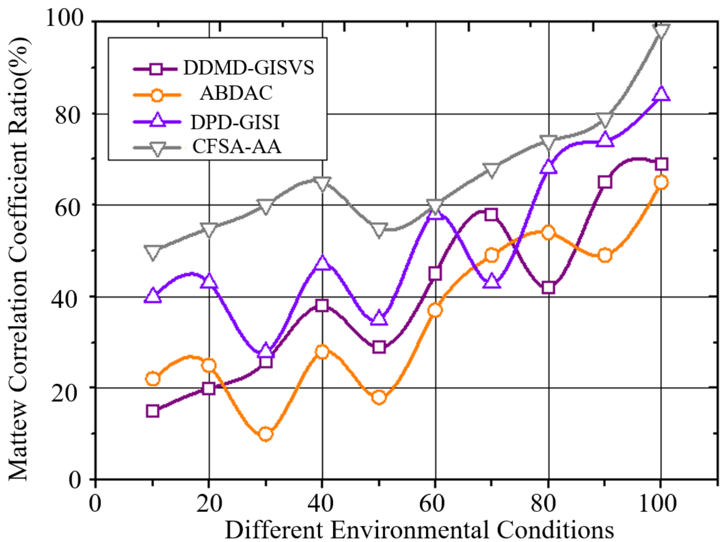
4.4 *Matthew correlation coefficient (MCC) ratio*

One metric for evaluating the efficacy of binary classification systems is the MCC. The MCC measures the degree of coherence between the actual and predicted fault classifications. Even if class sizes differ, MCC remains the balanced metric for true and false negatives and positives. MCC represents the model's correlation classification coefficient, which compares the observed and predicted models.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)\times(TP+FN)\times(TN+FP)\times TN+FN)}} \tag{22}$$

As shown in equation (22), where *TN* denotes the correctly labelled negative signal, *TP* indicates the correctly labelled positive signal, *FN* represents the incorrectly labelled negative signal, and *FP* signifies the incorrectly labelled positive signal. Figure 7 shows the Matthew correlation coefficient.

Figure 7 Matthew correlation coefficient ratio (see online version for colours)



The Matthew correlation coefficient (MCC) is an excellent indicator of the CFSA-AA model's fault-detection accuracy. It provides a comprehensive assessment of the model's performance, accounting for false positives, false negatives, true positives, and true negatives. The high MCC of 96.4% indicates a strong correlation between the actual and predicted fault classifications, even in skewed datasets. This robust metric ensures that unequal class distributions do not affect the model's accuracy when evaluating its fault-detection abilities across various GIS system contexts.

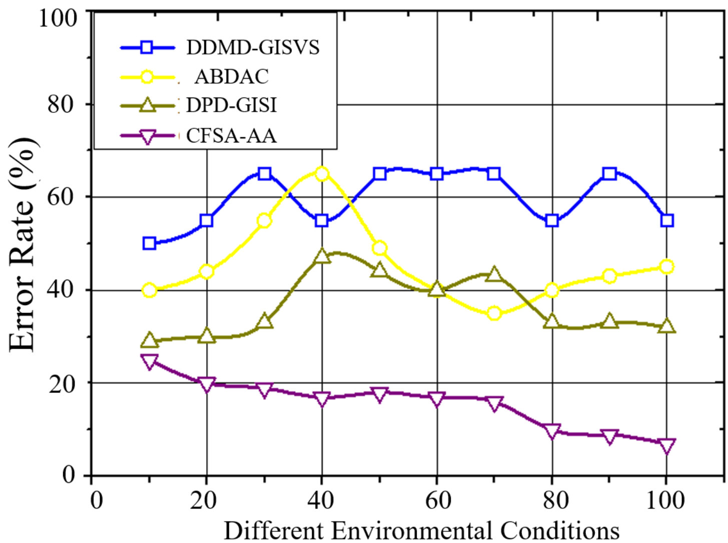
4.5 Error rate

Electrodes with protrusions or abnormalities, metallic particles in suspension, contamination of the spacer surface, and voids or gaps at the electrode/dielectric contact are the causes of PD. Mechanical abrasion of the wire during load cycling, manufacturing errors in GIS, vibration during the shipping and assembly of GIS equipment, undiscovered scratches on electrodes, and weak electrical connections are all potential causes of these faults. The error rate equation (23) is as follows.

$$Error\ Rate = \frac{FP + FN}{P + N} \tag{23}$$

The hot spot temperature estimation approach often yields non-negligible inaccuracies when the transformer load varies. Even for a basic procedure, the results of the thermal circuit model computation still need improvement. Numerous parameters are included in the numerical computation procedure. The intelligent a priori algorithm makes accurate oil temperature predictions easily. Over time, it will undoubtedly become a crucial resource for resolving the issue of transformer oil temperature prediction. Figure 8 shows the error rate.

Figure 8 Error rate (see online version for colours)



There is a strong correlation between the pattern finding ratio and the number of GIS system problems detected. It shows that the model can spot correlations and recurring error patterns in operational data. With a 97.4% accuracy, the CFSA-AA model clearly excels at identifying fault signs and their relationships. It enables a better knowledge of fault processes, predictive maintenance, and early fault identification. Improved dependability and reduced downtime in GIS operations result from the system's ability to detect these patterns and anticipate potential issues before they escalate.

5 Conclusion

This study presents the CFSA-AA for predicting discharge faults, mechanical faults, and abnormal mechanical vibrations in metal-enclosed GIS. This study suggested integrating online gas and electricity fault-monitoring devices into high-voltage switchgear to overcome the limitations of conventional PD monitoring methods, such as insufficient capacity for large-scale monitoring, stringent operator qualifications, and high inspection costs. Various influencing variables and SF_6 levels may be accounted for using the suggested method to determine the rate of temperature rise. Equipment insulation may be protected against damage and the effects of ageing using this method's improved technical design and practicality. As this study showed, the electrical equipment's internal temperature distribution is heavily influenced by the actual layout of the conductor and the enclosure, as well as the loss distribution. Outside the apparatus, radiation and airflow play a far greater role in transferring heat from the busbars to the enclosure than high-pressure SF_6 convection does within GIS, according to this research. To meet the functional needs of a cloud computing power plant, this study developed and implemented an intelligent fault-state diagnostic management system. The system's aim is realised by the interplay of its modules, which include technical supervision, electronic equipment account, data centre, expert database, and system authority management. This work uses the Apriori algorithm to build a rule-based knowledge base that includes failure occurrences, failure causes, and expert recommendations by correlating large-scale electric equipment vibration data with real-world fault phenomena. The experimental outcomes show that the suggested CFSA-AA model increases the fault diagnostic accuracy ratio to 98.9%, the pattern discovery ratio to 97.4%, the operation state detection ratio to 95.6%, the Mattew correlation coefficient ratio to 96.4%, and the error rate to 7.4% compared to other existing models. The model depends on the specific features and distribution of the dataset, and its generalisation ability has yet to be verified across different geographical regions or operating environments. The algorithm is sensitive to manually set minimum support and confidence thresholds, which may affect the stability of association rule mining. Data transmission latency and storage costs on cloud platforms pose potential constraints to practical deployment. The complex internal structure of GIS equipment and limitations in sensor layout may lead to incomplete collection of some fault signals, thus affecting the comprehensiveness of diagnosis. The coverage of historical data limits the model's ability to identify novel or rare fault modes. To expedite the development of power equipment condition monitoring and diagnosis, future work should accelerate the use of advanced technologies such as the Internet of Things (IoT), big data, and others.

Funding

This work is supported by the Science and Technology Project of State Grid Sichuan Electric Power Company (Project No.52199723001C).

Conflict of interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this paper.

References

Al-Jumaili, A.H.A., Muniyandi, R.C., Hasan, M.K., Singh, M.J., Paw, J.K.S. and Amir, M. (2023) 'Advancements in intelligent cloud computing for power optimization and battery management in hybrid renewable energy systems: a comprehensive review', *Energy Reports*, Vol. 10, pp.2206–2227, https://doi.org/10.1016/j.egyr.2023.08.069

Awaysheh, F.M., Aladwan, M.N., Alazab, M., Alawadi, S., Cabaleiro, J.C. and Pena, T.F. (2021) 'Security by design for big data frameworks over cloud computing', *IEEE Transactions on Engineering Management*, Vol. 69, No. 6, pp.3676–3693, https://doi.org/10.1109/TEM.2021.3082859

Dai, D., Liao, Q., Sang, Z., You, Y., Qiao, R. and Yuan, H. (2023) 'Audio general recognition of partial discharge and mechanical defects in switchgear using a smartphone', *Applied Sciences*, Vol. 13, No. 18, p.10153, https://doi.org/10.3390/app.131810153

Dhinakaran, D. and Prathap, P.J. (2022) 'Protection of data privacy from vulnerability using the two-fish technique with the Apriori algorithm in data mining', *The Journal of Supercomputing*, Vol. 78, No. 16, pp.17559–17593, https://doi.org/10.1007/s11227-022-04609-y

Faizol, Z., Zubir, F., Saman, N.M., Ahmad, M.H., Rahim, M.K.A., Ayop, O. and Yusoff, Z. (2023) 'Detection method of partial discharge on transformer and gas-insulated switchgear: a review', *Applied Sciences*, Vol. 13, No. 17, p.9605, https://doi.org/10.3390/app.13179605

Ghasemi, H., Farahani, E.S., Fotuhi-Firuzabad, M., Dehghanian, P., Ghasemi, A. and Wang, F. (2023) 'Equipment failure rate in electric power distribution networks: an overview of concepts, estimation, and modeling methods', *Engineering Failure Analysis*, Vol. 145, p.107034, https://doi.org/10.1016/j.engfailanal.2023.107034

Han, Y.J., Zhao, Q.Q. and Yun, W.Y. (2022) 'Optimal inspection and replacement strategy for 145 kV gas-insulated switchgear', *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, Vol. 236, No. 2, pp.339–347, https://doi.org/10.1177/1748006X211024562

Hashmi, S.A., Ali, C.F. and Zafar, S. (2021) 'Internet of things and cloud computing-based energy management system for demand-side management in smart grids', *International Journal of Energy Research*, Vol. 45, No. 1, pp.1007–1022, https://doi.org/10.1002/er.5758

Hassan, M.M., Zaman, S., Mollick, S., Hassan, M.M., Raihan, M., Kaushal, C. and Bhardwaj, R. (2023) 'An efficient apriori algorithm for frequent patterns in human intoxication data', *Innovations in Systems and Software Engineering*, Vol. 19, No. 1, pp.61–69, https://doi.org/10.1007/s11334-022-00461-y

He, L., Yang, J., Zhang, Z., Li, Z., Ding, D., Yuan, M. and Chen, M. (2021) 'Research on mechanical defect detection and diagnosis methods for GIS equipment based on vibration signals', *Energies*, Vol. 14, No. 17, p.5507, https://doi.org/10.3390/en14175507

Hussain, M.R., Refaat, S.S. and Abu-Rub, H. (2021) 'Overview and partial discharge analysis of power transformers: a literature review', *IEEE Access*, Vol. 9, pp.64587–64605, https://doi.org/10.1109/ACCESS.2021.3074824

Mahdi, A.S., Abdul-Malek, Z. and Arshad, R.N. (2022) 'SF_6 decomposed component analysis for partial discharge diagnosis in GIS: a review', *IEEE Access*, Vol. 10, pp.27270–27288, https://doi.org/10.1109/ACCESS.2022.3152985

Mahmoud, M.A., Md Nasir, N.R., Gurunathan, M., Raj, P. and Mostafa, S.A. (2021) 'The current state of the art in research on predictive maintenance in smart grid distribution network: fault types, causes, and prediction methods–a systematic review', *Energies*, Vol. 14, No. 16, p.5078, https://doi.org/10.3390/en14165078

Mehmood, E. and Anees, T. (2020) 'Challenges and solutions for processing real-time big data stream: a systematic literature review', *IEEE Access*, Vol. 8, pp.119123–119143, https://doi.org/10.1109/ACCESS.2020.3003852

Mier, C., Mor, A.R., Luo, T. and Vaessen, P. (2024) 'Partial discharge and interference discrimination in gas-insulated systems using electric and magnetic sensors', *International Journal of Electrical Power and Energy Systems*, Vol. 158, p.109911, https://doi.org/10.1016/j.ijepes.2024.109911

Panmala, N., Suwanasri, T. and Suwanasri, C. (2023) 'A fuzzy logic approach to health index determination for a gas-insulated switchgear', *Energies*, Vol. 16, No. 18, p.6605, https://doi.org/10.3390/en16186605

Peyghami, S., Palensky, P. and Blaabjerg, F. (2020) ‘An overview of the reliability of modern power electronic-based power systems’, *IEEE Open Journal of Power Electronics*, Vol. 1, pp.34–50.

Purnomoadi, A.P., Mor, A.R. and Smit, J.J. (2020) ‘Health index and risk assessment models for gas insulated switchgear (GIS) operating under tropical conditions’, *International Journal of Electrical Power and Energy Systems*, Vol. 117, p.105681, https://doi.org/10.1016/j.ijepes.2019.105681

Rozga, P., Kraslawski, A., Klarecki, A., Romanowski, A. and Krysiak, W. (2021) ‘A new approach for decision support of the selection of construction technology of high-voltage substations based on the AHP method’, *IEEE Access*, Vol. 9, pp.73413–73423, https://doi.org/10.1109/ACCESS.2021.3078852

Subramaniam, A., Sahoo, A., Manohar, S.S., Raman, S.J. and Panda, S.K. (2021) ‘Switchgear condition assessment and lifecycle management: standards, failure statistics, condition assessment, partial discharge analysis, maintenance approaches, and future trends’, *IEEE Electrical Insulation Magazine*, Vol. 37, No. 3, pp.27–41, https://doi.org/10.1109/MEI.2021.9388574

Sun, J., Fan, Y., Qiu, W., Song, S., Tan, X., Tian, Y. and Wu, N. (2023) ‘Applicability of the acoustic–electrical joint detection method to identify defects in a gas-insulated system’, *IET Science, Measurement and Technology*, Vol. 17, No. 4, pp.147–157, https://doi.org/10.1049/smt2.12145

Xing, Y., Wang, Z., Liu, L., Xu, Y., Yang, Y., Liu, S. and Li, C. (2022) ‘Defect types and failure modes of solid insulation in gas-insulated switchgear: in situ study and case analysis’, *High Voltage*, Vol. 7, No. 1, pp.158–164, https://doi.org/10.1049/hve2.2021.0037

Zebouchi, N. and Haddad, M.A. (2020) ‘A review on real-size epoxy cast resin insulators for compact high voltage direct current gas insulated switchgears (GIS) and gas insulated transmission lines (GIL)–current achievements and envisaged research and development’, *Energies*, Vol. 13, No. 23, p.6416, https://doi.org/10.3390/en13236416

Zhao, X. and Huang, C. (2024) ‘Intelligent processing of power operation data based on the improved Apriori algorithm’, *Strategic Planning for Energy and the Environment*, Vol. 43, No. 2, pp.477–498, https://doi.org/10.1080/23311916.2024.1878578

Zhuang, Y., Hu, X., Tang, B., Wang, S., Cui, A., Hou, K. and Chu, J. (2020) ‘Effects of SF_6 decomposition components and concentrations on the discharge faults and insulation defects in GIS equipment’, *Scientific Reports*, Vol. 10, No. 1, p.15039, https://doi.org/10.1038/s41598-020-71866-w

Zixian, Z., Xuning, L., Zhixiang, L. and Hongqiang, H. (2021) ‘Outburst prediction and influencing factors analysis based on Boruta-Apriori and BO-SVM algorithms’, *Journal of Intelligent and Fuzzy Systems*, Vol. 41, No. 2, pp.3201–3218, https://doi.org/10.3233/JIFS-210512

Website

https://www.kaggle.com/c/vsb-power-line-fault-detection/data