E-ISSN: 2997-9382
American Journal of Technology Advancement
https://semantjournals.org/index.php/AJTA

| Research Article

Check for updates

Faults Detection in Power System Grids using Isolation Forests

Hope Linus Edet[1], Nseobong Okpura[2], Akaninyene Obot[3], Mfonobong Umoren[4]
[1,2,3,4]Department of Electrical and Electronic Engineering, University of Uyo, Uyo, Akwa Ibom State, Nigeria

Abstract
The study examined faults detection in power system grids using Isolation Forests. In carrying out the study, a descriptive survey research design was adopted. The study was carried out in Akwa Ibom State. The target population of the study comprised all the licensed Electrical Engineers in Akwa Ibom State. A stratified sampling technique was used to select licensed Electrical Engineers from the three senatorial districts in Akwa Ibom State, represented by Uyo, Ikot Ekpene and Eket, giving a total sample size of 180 respondents for the study. The instrument used for data collection was a structured questionnaire entitled "Fault Detection in Power System Grids and Isolation Forests Questionnaire" (FDPSCIFQ). The face and content validity of the instrument were established by experts in Electrical Engineering to ensure the accuracy, appropriateness, clarity and completeness of the items in relation to the objectives of the study. The reliability of the instrument was determined and a reliability coefficient of 0.81 was obtained. This coefficient was considered sufficiently high enough to justify the use of the instrument for the study. The data generated from the study were analyzed using appropriate descriptive statistical techniques to answer the research questions. The results showed that Single Line-to-Ground Fault (L-G) was the major type of fault commonly encountered in power system grids, accounting for 23.89%, while Three-Phase-to-Ground Fault (L-L-L-G) recorded the least percentage of 16.67%. The results also showed that Extended Isolation Forest (EIF) was the major type of Isolation Forest identified as useful for detecting different forms of anomalies in datasets, with 24.44%, while Deep Isolation Forest (Deep IF) recorded the least percentage of 15.56%. Furthermore, Axis-Parallel Splitting Artifacts and Feature Space Constraints constituted the major challenge associated with the use of Isolation Forests for fault detection, with 28.33%, while High False-Positive Rates in High-Noise Sensor Environments recorded the least percentage of 21.67%. Regarding mitigating strategies, Extended Isolation Forest (EIF) with Random Hyperplanes recorded the highest percentage of 23.89%, while Physics-Informed Feature Engineering and SHAP Explanations recorded the least percentage of 15.00%. The study concluded that Isolation Forests offer a promising intelligent approach for improving the speed, flexibility, and effectiveness of fault detection in modern power system grids and can contribute to the development of more reliable, resilient, and intelligent electrical power networks. One of the recommendations made was that Power system operators should integrate Isolation Forests with existing protection systems to provide an additional intelligent layer for detecting abnormal electrical conditions and identifying potential faults before they develop into more serious grid disturbances.

Keywords: Faults Detection, Power System, Grids, Isolation Forests

This is an open-access article under the CC-BY 4.0 license

1. Introduction

Faults in power system grids are abnormal conditions that can significantly affect the reliability, stability, and safety of electricity supply. They may occur in transmission lines, distribution networks, transformers, generators, and other grid components as a result of insulation failure, lightning, equipment deterioration, conductor contact, environmental conditions, or human activities. When a fault occurs, electrical parameters such as voltage, current, frequency, and phase angles may deviate considerably from their normal operating conditions. Although conventional protection devices such as relays and circuit breakers are widely used to detect and isolate faults, the increasing complexity of modern power grids has created a need for intelligent techniques that can identify abnormal operating patterns quickly and accurately. Therefore, effective fault detection has become an important requirement for maintaining power-system reliability, preventing equipment damage, and reducing interruptions in electricity supply [1], [2].

The development of artificial intelligence and machine-learning techniques has provided new opportunities for improving fault detection in power systems. One technique that has attracted attention is Isolation Forest, an unsupervised machine-learning algorithm designed to identify unusual observations by isolating them from normal data. Isolation Forest is based on the principle that abnormal observations are generally fewer and easier to isolate than normal observations. It constructs a number of isolation trees in which data are repeatedly divided using randomly selected features and split points. Observations that become isolated after relatively few partitions receive higher anomaly scores and can consequently be identified as potential faults. Lesouple [3] noted that the isolation-based approach provides an efficient means of detecting anomalies without requiring extensive labelled datasets. This characteristic is particularly relevant to power-system fault detection because obtaining accurately labelled fault data for every possible operating condition can be difficult, expensive, and time-consuming.

The application of Isolation Forest to fault detection in power system grids therefore provides an opportunity to combine electrical measurements with intelligent anomaly detection. Data obtained from voltage, current, frequency, phase-angle measurements, and other grid-monitoring devices can be analysed to distinguish normal operating conditions from abnormal patterns associated with faults. The ability of Isolation Forest to process large and multidimensional datasets makes it potentially useful for modern grids characterized by distributed generation, renewable-energy integration, smart sensors, and continuously changing operating conditions. Feng [4] demonstrated the effectiveness of an improved Isolation Forest approach in detecting abnormal patterns in large-scale data, while Pelletier [5] highlighted its usefulness for identifying geometrically isolated observations. Consequently, faults detection in power system grids using Isolation Forest can provide a promising intelligent approach for improving the speed and effectiveness of fault identification, supporting timely protective actions, minimizing equipment damage, and contributing to the overall reliability and stability of modern electrical power systems.

Statement of the Problem

The increasing complexity of modern power system grids has made the detection of electrical faults a major concern for reliable power supply. Faults such as short circuits, abnormal current, voltage disturbances, and other unusual operating conditions can cause equipment damage, power interruptions, instability, and system failures if they are not detected early. Conventional fault detection methods often depend on predefined rules, system models, relay settings, and known fault characteristics, which may become less effective when the operating conditions or configuration of the power grid change. In addition, power-system fault datasets are often imbalanced because normal operating conditions occur much more frequently than actual fault

events. This makes it difficult for some supervised machine-learning techniques to obtain sufficient labelled examples of different fault types for effective training and detection [6].

The problem is further complicated by the possibility of detecting unusual or previously unseen fault conditions for which adequate labelled data may not be available. Isolation Forest provides an alternative anomaly-detection approach because it identifies observations that are different from normal patterns by isolating them through randomly generated decision trees. Isolation Forest is particularly designed for detecting anomalies and can operate efficiently on large and high-dimensional datasets Shakiba, Azizi, Zhou & Abusorrah, [7]. However, there is still a need to establish how effectively Isolation Forest can detect faults and abnormal conditions in power-system grids and whether it can provide reliable detection performance under varying grid conditions. Therefore, this study focuses on the application of Isolation Forest for fault detection in power system grids, with the aim of determining its effectiveness in identifying abnormal power-system behaviour and improving the reliability of fault detection.

Research Objectives

The study specifically seeks:

1. To find out the major types of faults commonly encountered in power system grids.
2. To examine the major types of Isolation Forests that are useful for detecting different forms of anomalies in datasets.
3. To determine the challenges associated with the use of Isolation Forests for fault detection in power systems.
4. Identify the mitigating strategies to the challenges associated with the use of Isolation Forests for fault detection in power systems**.**

Research Question

1. What are the major types of faults commonly encountered in power system grids?
2. What are the major types of Isolation Forests that are useful for detecting different forms of anomalies in datasets?
3. What are the challenges associated with the use of Isolation Forests for fault detection in power systems?
4. What are the mitigating strategies to the challenges associated with the use of Isolation Forests for fault detection in power systems?

Concept of Faults in Power System Grids

A fault in a power system grid refers to an abnormal electrical condition that occurs when the normal operating conditions of an electrical network are disturbed. Faults may result from unintended contact between conductors, contact between a conductor and the ground, insulation failure, equipment damage, or other disturbances. According to Grainger and Stevenson [8], faults are among the major disturbances considered in power-system analysis because they can produce excessive currents and abnormal voltages that threaten the safe operation of electrical equipment. Faults may occur in generating stations, transmission lines, distribution networks, transformers, and other components of the power grid.

Power-system faults are commonly classified into symmetrical and unsymmetrical faults. A symmetrical fault, particularly a three-phase fault, involves all three phases and produces approximately equal fault currents with a 120-degree phase displacement. Unsymmetrical faults affect the phases differently and include single line-to-ground, line-to-line, and double line-to-ground faults. Among these, single line-to-ground faults are generally the most frequently

occurring faults in transmission and distribution systems. Kundur, Balu, and Lauby [9] explain that understanding the nature and classification of faults is important for determining the magnitude of fault currents and designing appropriate protection systems.

Faults can arise from several causes, including lightning strikes, strong winds, falling trees, equipment deterioration, insulation breakdown, conductor contact, human errors, and animal interference. Environmental conditions such as heavy rainfall, flooding, pollution, and high temperatures may also contribute to insulation deterioration and equipment failure. When a fault occurs, the impedance between certain points in the network may become very low, resulting in a substantial increase in current. Excessive fault current can cause overheating, mechanical stress, equipment damage, voltage disturbances, and interruption of electricity supply if the fault is not cleared promptly.

The occurrence of faults therefore makes power-system protection an essential part of grid operation. Protective devices such as circuit breakers, relays, fuses, current transformers, and voltage transformers are used to detect abnormal conditions and isolate the faulty section from the healthy portion of the network. Modern protection systems increasingly employ digital and numerical relays capable of detecting faults rapidly and selectively. According to Blackburn and Domin, effective protection should operate quickly enough to minimize equipment damage while maintaining the continuity and stability of electricity supply to unaffected parts of the system.

Concept of Isolation Forests

Isolation Forest is an unsupervised machine-learning technique used primarily for detecting anomalies or unusual observations in datasets. The method is based on the principle that anomalous observations are usually few and different from normal observations and can therefore be isolated more easily. Unlike many conventional anomaly-detection methods that attempt to model normal observations, Isolation Forest directly isolates individual observations through random partitioning. According to Lesouple, Isolation Forest provides an efficient approach to anomaly detection by recursively partitioning the data and identifying observations that can be separated with relatively few partitions. This makes the technique particularly useful where labelled data are unavailable.

The basic operation of Isolation Forest involves constructing a collection of binary decision trees known as isolation trees. During tree construction, a feature is randomly selected and a random split point is chosen between the minimum and maximum values of that feature. This process continues recursively until observations are isolated. An observation that requires only a small number of partitions to become isolated is considered more likely to be anomalous, whereas an observation requiring a longer path through the tree is more likely to represent normal behaviour. Hannák & Horváth [10] explained that the computational efficiency of Isolation Forest results from its isolation-based approach, which does not depend on distance calculations and can be applied effectively to multidimensional numerical data.

Another important feature of Isolation Forest is its ability to handle large and complex datasets with relatively low computational requirements. The method can be useful in situations where data are continuously generated and where conventional supervised learning approaches are difficult to implement because labelled examples of anomalies are limited. Feng, for example, applied an improved Isolation Forest approach to network traffic anomaly detection and reported that the method was effective for detecting abnormal patterns in large-scale network traffic data. Their findings demonstrate that Isolation Forest can serve as an important tool for identifying unusual behaviour in environments characterized by large volumes of unlabelled data.

However, Isolation Forest is not without limitations. Although the method performs well when anomalies are sufficiently different from normal observations, it may experience difficulty when anomalous observations occur in dense groups or are not clearly separated from normal

observations. Pelletier examined the statistical properties of Isolation Forest and found that the method is particularly effective for detecting geometrically isolated observations, but dense clusters of anomalies may remain undetected because of what is described as a masking effect. This limitation has encouraged researchers to develop improved versions of the method that can address more complex data structures and different forms of anomalies.

Recent developments have continued to improve the application of Isolation Forest to modern data-analysis problems. He and Chen [11], for instance, developed an anomaly-detection approach based on Isolation Forest and adaptive feature selection to improve performance when dealing with large-scale and high-dimensional data. Their study showed that adjusting the contribution of different features can improve anomaly-detection performance. Similarly, research on variants of Isolation Forest has extended its application to streaming data, uncertain data, and other complex environments. Therefore, Isolation Forest can be conceptualized as a flexible, efficient and unsupervised anomaly-detection technique that identifies unusual observations based on how easily they can be isolated from the rest of a dataset.

Types of Faults in Power System Grids

The major types of faults commonly encountered in power system grids are as follows:

- **Single Line-to-Ground Fault (L-G)**

A single line-to-ground fault occurs when one phase conductor comes into contact with the ground or grounded equipment. It is the most frequently occurring type of fault in many power systems. It can be caused by lightning, insulation failure, fallen trees, broken conductors, or other environmental and equipment-related conditions. The severity of the fault depends on the system grounding arrangement and impedance of the fault path [12].

- **Line-to-Line Fault (L-L)**

A line-to-line fault occurs when two phase conductors come into contact with each other. This type of fault does not necessarily involve the ground and may occur because of conductor movement, insulation breakdown, mechanical damage, or adverse weather conditions. Line-to-line faults create unbalanced currents and can impose considerable thermal and mechanical stress on electrical equipment.

- **Double Line-to-Ground Fault (L-L-G)**

A double line-to-ground fault occurs when two phase conductors simultaneously make contact with the ground. It may result from damaged insulation, fallen conductors, equipment failure, or the development of a ground connection following a line-to-line fault. Such faults can produce substantial currents and voltage disturbances, making rapid detection and isolation important for maintaining grid reliability [13].

- **Three-Phase Fault (L-L-L)**

A three-phase fault occurs when all three phase conductors become electrically connected. It is described as a symmetrical fault because the currents in the three phases have approximately equal magnitudes and are separated by 120 degrees. Although three-phase faults occur less frequently than single line-to-ground faults, they can produce very high currents and therefore require adequate circuit-breaker and protection-system ratings [14].

- **Three-Phase-to-Ground Fault (L-L-L-G)**

A three-phase-to-ground fault occurs when the three phase conductors simultaneously come into contact with the ground. It is a serious disturbance that can produce high fault currents, severe voltage changes, and significant stress on generators, transformers, transmission lines, and other

grid components. Effective protective relaying and rapid circuit-breaker operation are required to isolate the fault and reduce its effect on the wider network [15].

Types of Isolation Forest

The major types of Isolation Forest are useful for detecting different forms of anomalies in datasets.

- **Extended Isolation Forest (EIF)**

Extended Isolation Forest is an improved version of the traditional Isolation Forest. It uses randomly oriented hyperplanes to divide data rather than relying only on splits along individual features. This allows it to identify anomalies more effectively in multidimensional and complex datasets. It is particularly useful where abnormal observations have complicated relationships among several variables [16].

- **Fuzzy Isolation Forest**

Fuzzy Isolation Forest is designed to handle data characterized by uncertainty and imprecise boundaries. It incorporates fuzzy concepts into the isolation process, allowing observations to have different degrees of membership rather than being treated as completely normal or abnormal. This makes the approach useful for datasets in which uncertainty is an important characteristic of the information being analysed [17].

- **Online Isolation Forest**

Online Isolation Forest is designed for continuously changing or streaming datasets. Instead of requiring the entire dataset to be available before analysis, the algorithm can process new observations as they arrive. This makes it suitable for real-time applications such as network monitoring, fraud detection, industrial systems, and cyber security, where new data are continuously generated [18].

- **Layered Isolation Forest (LIF)**

Layered Isolation Forest is an approach that improves anomaly detection by examining data through different levels or subspaces. It is particularly useful for identifying local anomalies that may appear normal when the entire dataset is considered but become abnormal within a particular group or region. The layered approach can therefore improve the detection of complex patterns that may be missed by a conventional Isolation Forest.

- **Deep Isolation Forest (Deep IF)**

Deep Isolation Forest is designed to improve anomaly detection in complex and high-dimensional datasets. It combines the isolation principle with deeper representations of the data, helping the algorithm identify anomalies that may not be easily separated using conventional methods. It is useful when the relationships between variables are nonlinear and difficult to capture using simpler isolation techniques.

Steps for Adoption of Isolation Forest in Fault Detection in Power System Grids

- **Collection of Power-System Data**

The first step in adopting Isolation Forest for fault detection in power system grids is the collection of relevant operational data. Data can be obtained from SCADA systems, smart meters, protective relays, phasor measurement units and other monitoring devices. Important parameters include voltage, current, frequency, active power, reactive power and phase angle. Both normal and abnormal operating data should be collected because Isolation Forest identifies observations that differ from normal operating patterns. Li [19] noted that power-dispatching data can be used with Isolation Forest to detect abnormal conditions in power-system operations.

- **Data Pre-processing and Cleaning**

The second step involves preparing the collected data for analysis. Power-system data may contain missing values, measurement errors, duplicated records and noise. These problems should be identified and corrected before the data are used to develop the Isolation Forest model. Proper data cleaning helps the model distinguish genuine faults from normal variations in power-system operation. Safari [20] demonstrated the usefulness of Isolation Forest for intelligent fault detection in power-electronic systems, particularly where labelled fault data are limited.

- **Feature Extraction and Selection**

The third step is the extraction and selection of important features from the prepared data. Features such as voltage variation, current magnitude, frequency deviation, phase-angle changes and fluctuations in active and reactive power can provide useful information about the condition of the grid. Selecting appropriate features helps the Isolation Forest model concentrate on characteristics that can distinguish normal conditions from abnormal ones. Huang and Gao [21] showed that combining signal-processing techniques with Isolation Forest can improve the identification of characteristics associated with electrical faults.

- **Development of the Isolation Forest Model**

The fourth step involves developing the Isolation Forest model using the selected power-system features. The algorithm creates several isolation trees by randomly selecting features and splitting observations. Abnormal observations are generally isolated more quickly than normal observations because their characteristics differ from the majority of the dataset. The model can therefore assign anomaly scores to observations and identify those that may represent faults. Cai and Lu [22] demonstrated the application of Isolation Forest to automatic fault location in a 10 kV distribution network.

- **Detection and Identification of Faults**

The fifth step is the detection and identification of faults using the developed model. New measurements from the power grid are supplied to the Isolation Forest model, which evaluates them and assigns anomaly scores. Observations with high anomaly scores can be identified as potential abnormal conditions and referred for further investigation. This approach can assist in identifying transmission-line defects, equipment abnormalities and other unusual operating conditions. Wang and Wang [23] found that an improved Isolation Forest approach could effectively identify transmission-line defects and hazards.

- **Real-Time Implementation and Monitoring**

The sixth step is the implementation of the Isolation Forest model for real-time monitoring of the power system. Modern power grids continuously generate large volumes of operational data, making rapid analysis important for early fault detection. The model can be integrated into a monitoring system to analyse incoming measurements and generate alerts when unusual conditions are detected. Xie [24] developed an adaptive Isolation Forest method for power-dispatching streaming data, emphasizing the need for anomaly-detection models that can adapt to changing operating conditions.

- **Performance Evaluation and Continuous Improvement**

The final step is evaluating and continuously improving the Isolation Forest fault-detection system. The performance of the model can be assessed using measures such as precision, recall, F1-score, false-alarm rate and detection time. The system should also be reviewed and updated as power-system conditions change. Improvements may involve updating the dataset, selecting better features, adjusting anomaly thresholds and modifying the model. Din [25] demonstrated an

enhanced Isolation Forest approach for smart-grid anomaly detection through normalization, feature selection and threshold analysis.

Challenges of Using Isolation Forests in Faults Detection in Power System

Isolation Forest (iForest) is an unsupervised anomaly detection algorithm widely used across various domains due to its linear time complexity and efficiency in isolating anomalies by randomly partitioning feature spaces. However, when applied to fault detection in modern electrical power systems—such as transmission lines, microgrids, photovoltaic (PV) arrays, and hydroelectric generating units—iForest encounters several fundamental limitations. The primary challenges of employing Isolation Forests for power system fault detection include:

- **Axis-Parallel Splitting Artifacts and Feature Space Constraints**

Traditional Isolation Forests use axis-aligned splits by selecting a single random attribute and a random cut point. In power systems, sensor data—such as synchronized Phasor Measurement Unit (PMU) measurements, current, voltage, and reactive power readings—are highly correlated across time and spatial nodes. Axis-parallel splitting creates artificial, rectangular decision boundaries, which can cause severe misclassifications in multivariate time-series data. This leads to high false-positive rates when subtle electrical transients occur, or false negatives when complex multidimensional phase shifts take place. Advanced variants, such as Extended Isolation Forest (EIF), attempt to resolve this using hyperplanes with random slopes, though at the expense of higher variance and computational complexity [26].

- **Sensitivity to Temporal Dependencies and Dynamic Operating Conditions**

Power systems operate under continuous, dynamic load fluctuations, switching operations, and intermittent renewable energy inputs (such as solar PV irradiance shifts). Standard Isolation Forests evaluate data points independently as static observations, completely ignoring temporal sequence context. Consequently, legitimate operational shifts—such as sudden high-demand load spikes or solar shading events—may be misidentified as severe equipment faults, while slow, incipient insulation degradations remain unisolated due to their gradual drift [27].

- **Hyperparameter Instability in Streaming and High-Dimensional Grid Environments**

The operational performance of Isolation Forest relies heavily on hyperparameter configurations, including sample size (S), tree depth (D), number of trees (T), and anomaly score thresholds (Th). In continuous power-dispatch streaming environments, fixed hyperparameter configurations fail to adapt to evolving grid topology and continuous sensor streams. This static nature frequently causes instability in anomaly scores, generating redundant alarms or missing low-level electrical faults during non-stationary grid states [28].

- **High False-Positive Rates in High-Noise Sensor Environments**

Smart grids rely on complex measurement infrastructures, which are susceptible to measurement noise, missing packets, communication latency, and calibration drifts. Because Isolation Forest operates on the principle that "few and different" data instances are anomalous, transient noise spikes in PMU data or communication errors are rapidly isolated at shallow tree levels, causing a high rate of false fault detections [29].

Mitigating Strategies to the Challenges of using Isolation Forests in Faults Detection in Power System Grids

To overcome the operational limitations of the standard Isolation Forest (iForest) algorithm in power system fault detection, tailored algorithmic modifications, structural integrations, and domain-specific enhancements must be applied. Below are the mitigating strategies paired directly with the challenges associated with deploying standard Isolation Forests in power system grids.

- **Solution to Axis-Parallel Splitting Artifacts: Extended Isolation Forest (EIF) with Random Hyperplanes**

To eliminate the axis-bias (rectangular decision boundaries) caused by axis-aligned cuts, standard iForest is replaced or enhanced using the Extended Isolation Forest (EIF) framework [30]. Instead of picking a single feature and a random value along its axis, EIF selects a random normal vector to construct cutting hyperplanes with random slopes across multi-variable feature spaces. Implementation in Power Grids: In electrical power networks where voltage and current phasors are bound by physical laws (e.g., Ohm's and Kirchhoff's laws), random slope hyperplanes allow the cuts to align obliquely across non-linear correlations. This significantly reduces false-positive anomaly scores along continuous operating trajectories.

- **Solution to Time-Series Ignorance: Hybrid Temporal Feature Extraction (FFT/Sliding-Window Integration)**

To capture sequential dynamics and frequency changes without discarding the computational speed of iForest, the raw static feature inputs are replaced with time-series transformed features. Implementation in Power Grids: High-frequency waveform signals obtained from Phasor Measurement Units (PMUs) or SCADA systems are first processed through sliding time windows, Fast Fourier Transforms (FFT), or wavelet transforms to extract temporal-domain features (such as rate-of-change of frequency [RoCoF], harmonic distortion indices, and moving variance) before feeding them into the isolation trees. Converting sequential dynamics into explicit time-frequency features enables iForest to isolate transient grid faults, voltage sags, and sub-synchronous resonance without losing temporal context.

- **Solution to High-Dimensional Noise and Sensor Distortions: Dimensionality Reduction & Subspace Isolation**

To prevent random sensor noise, communication delays, or measurement errors from dominating tree branch selection, iForest is coupled with unsupervised feature projection or subspace sampling techniques. Implementation in Power Grids: High-dimensional smart grid sensor networks deploy Principal Component Analysis (PCA) or Autoencoder embeddings prior to isolation tree construction. Reducing feature dimensions projects low-magnitude sensor noise into orthogonal spaces while preserving maximum variance. Furthermore, training ensemble trees on constrained random feature subsets prevents noisy or redundant telemetry channels from impairing the identification of low-magnitude faults (such as high-impedance ground faults).

- **Solution to Swamping and Masking Effects: Dynamic Sub-sampling and Density-Weighted Isolation**

To prevent large, consecutive clusters of abnormal data (e.g., during prolonged disturbances or sustained cyber-physical attacks) from collapsing isolation depth, models employ sub-sampling optimization and density-aware tree partitioning. Implementation in Power Grids: By training individual trees on very small, randomly sampled subsets of the power grid dataset (e.g., ψ=256 samples per tree), the probability that a dense cluster of anomalies appears together within a single sub-sample is drastically reduced. This forces clustered fault points to appear isolated in sub-sampled trees, successfully eliminating masking effects during continuous grid disturbances, islanding events, or False Data Injection Attacks (FDIAs).

- **Solution to Lack of Physics-Informed Interpretability: Physics-Informed Feature Engineering & SHAP Explanations**

To resolve the "black-box" scalar output issue, iForest is integrated with Physics-Informed Machine Learning (PIML) frameworks and post-hoc Explainable AI (XAI) metrics. Implementation in Power Grids: Power grid domain equations (such as power balance residuals, tracking errors, and physical operational limits) are encoded directly as explicit input features into

the isolation pipeline. When an anomaly is detected, model agnostic explainability tools—such as Shapley Additive exPlanations (SHAP)—calculate feature importance scores for individual tree paths. This maps the generic anomaly score directly to physical root causes (e.g., identifying phase-to-ground short circuits versus routine capacitor bank switching).

2. Methodology

In carrying out the study, a descriptive survey research design was adopted. The study was carried out in Akwa Ibom State. The target population of the study comprised all the licensed Electrical Engineers in Akwa Ibom State. A stratified sampling technique was used to select licensed Electrical Engineers from the three senatorial districts in Akwa Ibom State, represented by Uyo, Ikot Ekpene and Eket, giving a total sample size of 180 respondents for the study. The instrument used for data collection was a structured questionnaire entitled "Fault Detection in Power System Grids and Isolation Forests Questionnaire" (FDPSCIFQ). The face and content validity of the instrument were established by experts in Electrical Engineering to ensure the accuracy, appropriateness, clarity and completeness of the items in relation to the objectives of the study. The reliability of the instrument was determined and a reliability coefficient of 0.81 was obtained. This coefficient was considered sufficiently high enough to justify the use of the instrument for the study. The data generated from the study were analyzed using appropriate descriptive statistical techniques to answer the research questions.

3. Results and Discusions

Research Question 1:

The research question sought to examine the major types of faults commonly encountered in power system grids. To answer the research question, percentage analysis was performed on the data (see Table 1).

Table 1. Percentage Analysis of the Major Types of Faults Commonly Encountered in Power System Grids

Major Types of Faults Commonly Encountered in Power System Grids	**Frequency**	**%**
Single Line-to-Ground Fault (L-G)	43	23.89**
Line-to-Line Fault (L-L)	39	21.67
Double Line-to-Ground Fault (L-L-G)	36	20.00
Three-Phase Fault (L-L-L)	32	17.78
Three-Phase-to-Ground Fault (L-L-L-G)	30	16.67*
TOTAL	**180**	**100**

****The highest percentage frequency**
***The least percentage frequency**
SOURCE: Field survey

Table 1 presents the percentage analysis of the major types of faults commonly encountered in power system grids. The result of the data analysis showed that the highest percentage (23.89%) was recorded for "Single Line-to-Ground Fault (L-G)", while the least percentage (16.67%) was recorded for "Three-Phase-to-Ground Fault (L-L-L-G)". The finding is in support of Saadat, who stated that a single line-to-ground fault occurs when one phase conductor comes into contact with

the ground or grounded equipment. It is the most frequently occurring type of fault in many power systems. It can be caused by lightning, insulation failure, fallen trees, broken conductors, or other environmental and equipment-related conditions. Similarly, Kothari and Nagrath explained that three-phase faults occur less frequently than single line-to-ground faults, and they can produce very high currents and therefore require adequate circuit-breaker and protection-system ratings.

Research Question 2:

The research question sought to examine the major types of Isolation Forests that are useful for detecting different forms of anomalies in datasets. To answer the research question, percentage analysis was performed on the data (see Table 2).

Table 2. Percentage Analysis of the Major Types of Isolation Forests that are Useful for Detecting Different Forms of Anomalies in Datasets

Major Types of Isolation Forests that are Useful for Detecting Different Forms of Anomalies in Datasets	**Frequency**	**%**
Extended Isolation Forest (EIF)	44	24.44**
Fuzzy Isolation Forest	40	22.22
Online Isolation Forest	36	20.00
Layered Isolation Forest (LIF)	32	17.78
Deep Isolation Forest (Deep IF)	28	15.56*
TOTAL	**180**	**100**

****The highest percentage frequency**
***The least percentage frequency**
SOURCE: Field survey

Table 2 presents the percentage analysis of the major types of Isolation Forests that are useful for detecting different forms of anomalies in datasets. From the result of the data analysis, it was observed that the highest percentage (24.44%) was recorded for "**Extended Isolation Forest (EIF)",** while the least percentage (15.56%) was recorded for **"Deep Isolation Forest (Deep IF)".** The finding aligns with Wang and Wang, who noted that an improved Isolation Forest approach could effectively identify transmission-line defects and hazards. Cai and Lu observed that abnormal observations are generally isolated more quickly than normal observations because their characteristics differ from the majority of the dataset. The model can therefore assign anomaly scores to observations and identify those that may represent faults.

Research Question 3:

The research question sought to examine the challenges associated with the use of Isolation Forests for fault detection in power systems. To answer the research question, percentage analysis was performed on the data, as presented in Table 3.

Table 3. Percentage Analysis of the Challenges of Using Isolation Forests in Faults Detection in Power System

Challenges of using isolation forests in faults detection	Frequency	%
Axis-Parallel Splitting Artifacts and Feature Space Constraints	51	28.33**
Sensitivity to Temporal Dependencies and Dynamic Operating Conditions	47	26.11
Hyperparameter Instability in Streaming and High-Dimensional Grid Environments	43	23.89
High False-Positive Rates in High-Noise Sensor Environments	39	21.67*
TOTAL	**180**	**100**

****The highest percentage frequency**
***The least percentage frequency**
SOURCE: Field survey

Table 3 presents the percentage analysis of the challenges of using Isolation Forests for fault detection in power systems. From the result of the data analysis, it was observed that the highest percentage (28.33%) was recorded for "Axis-Parallel Splitting Artifacts and Feature Space Constraints," while the least percentage (21.67%) was recorded for "High False-Positive Rates in High-Noise Sensor Environments". The finding is in agreement with the opinion of De Santis and Costa, who stated that Axis-parallel splitting creates artificial, rectangular decision boundaries, which can cause severe misclassifications in multivariate time-series data. This leads to high false-positive rates when subtle electrical transients occur or false negatives when complex multidimensional phase shifts take place. Advanced variants, such as Extended Isolation Forest (EIF), attempt to resolve this using hyperplanes with random slopes, though at the expense of higher variance and computational complexity. Similarly, Wu, Ma, Alharbi, Wang, Xiong, Zhu, Qin, and Wang, affirmed that Power systems operate under continuous, dynamic load fluctuations, switching operations, and intermittent renewable energy inputs (such as solar PV irradiance shifts). Standard Isolation Forests evaluate data points independently as static observations, completely ignoring temporal sequence context. Consequently, legitimate operational shifts such as sudden high-demand load spikes or solar shading events may be misidentified as severe equipment faults, while slow, incipient insulation degradations remain unisolated due to their gradual drift.

Research Questions 4:

The research question sought to identify the mitigating strategies to the challenges of using isolation forests in faults detection in power system grids. To answer the research question, percentage analysis was performed on the data, as presented in Table 4.

Table 4. Percentage analysis of the mitigating strategies to the challenges of using isolation forests in faults detection in power system grids

Mitigating strategies to the challenges of using isolation forests in faults detection	Frequency	%
Solution to Axis-Parallel Splitting Artifacts: Extended Isolation Forest (EIF) with Random Hyperplanes	43	23.89**
--Solution to Time-Series Ignorance: Hybrid Temporal Feature Extraction (FFT/Sliding-Window Integration)	40	22.22
Solution to High-Dimensional Noise and Sensor Distortions: Dimensionality Reduction & Subspace Isolation	37	20.56
Solution to Swamping and Masking Effects: Dynamic Sub-sampling and Density-Weighted Isolation	33	18.33
Solution to Lack of Physics-Informed Interpretability: Physics-Informed Feature Engineering & SHAP Explanations	27	15.00*
TOTAL	**180**	**100**

****The highest percentage frequency**
***The least percentage frequency**
SOURCE: Field survey

The above table 4 presents the percentage analysis of the mitigating strategies to the challenges of using Isolation Forests in fault detection. The result of the data analysis showed that the highest percentage (23.89%) was recorded for “Solution to Axis-Parallel Splitting Artifacts: Extended Isolation Forest (EIF) with Random Hyperplanes”, while the least percentage (15.00%) was recorded for” Solution to Lack of Physics-Informed Interpretability: Physics-Informed Feature Engineering & SHAP Explanations. The finding agrees with the opinion of Hariri, who stated that to eliminate the axis-bias (rectangular decision boundaries) caused by axis-aligned cuts, standard iForest is replaced or enhanced using the Extended Isolation Forest (EIF) framework. Instead of picking a single feature and a random value along its axis, EIF se----------lects a random normal vector to construct cutting hyperplanes with random slopes across multi-variable feature spaces. Implementation in Power Grids: In electrical power networks where voltage and current phasors are bound by physical laws (e.g., Ohm's and Kirchhoff's laws), random slope hyperplanes allow the cuts to align obliquely across non-linear correlations. This significantly reduces false-positive anomaly scores along continuous operating trajectories.

4. Conclusion

In conclusion, faults in power system grids constitute major disturbances that can compromise the stability, reliability, safety, and continuous supply of electrical energy when they are not detected and cleared promptly. The different forms of faults, including single line-to-ground, line-to-line, double line-to-ground, three-phase, and three-phase-to-ground faults, produce abnormal changes in important electrical parameters such as current, voltage, frequency, and phase angle. Although conventional protective devices remain essential for fault isolation, the increasing complexity of

modern power grids, the integration of renewable energy sources, the growth of smart-grid technologies, and the continuous generation of large volumes of operational data require more intelligent approaches to fault detection. Isolation Forests provide a promising machine-learning approach because they can identify unusual patterns without requiring extensive labelled fault data and can analyse large and multidimensional datasets through the isolation of observations that differ significantly from normal operating conditions. Different variants, including Extended Isolation Forest, Fuzzy Isolation Forest, Online Isolation Forest, Layered Isolation Forest, and Deep Isolation Forest, further provide opportunities to improve anomaly detection under complex operating conditions. However, challenges such as sensor noise, temporal dependencies, changing grid conditions, high-dimensional data, hyperparameter selection, and false-positive detections may affect the effectiveness of the technique. Therefore, the successful application of Isolation Forests in power-system fault detection requires appropriate data preprocessing, careful model configuration, continuous monitoring, and integration with existing protection mechanisms. Power system operators should integrate Isolation Forests with existing protection systems to provide an additional intelligent layer for detecting abnormal electrical conditions and identifying potential faults before they develop into more serious grid disturbances. Adequate investment should be made in high-quality monitoring and data acquisition systems, including smart sensors, PMUs, and digital relays, while appropriate preprocessing techniques should be employed to reduce noise, missing data, and measurement errors that may increase false-positive fault detections. Researchers and power-system engineers should develop adaptive and improved Isolation Forest models capable of handling real-time data, changing load conditions, renewable-energy fluctuations, temporal dependencies, and high-dimensional grid measurements to improve the accuracy and reliability of automated fault detection.

REFERENCES

[1] J. L. Blackburn and T. J. Domin, *Protective Relaying: Principles and Applications*, 5th ed. CRC Press, 2020.

[2] J. D. Glover, M. S. Sarma, and T. J. Overbye, *Power System Analysis and Design*, 7th ed. Cengage Learning, 2022.

[3] J. Lesouple, C. Baudoin, M. Spigai, and J.-Y. Tourneret, "Generalized isolation forest for anomaly detection," *Pattern Recognition Letters*, vol. 149, pp. 109–119, 2021.

[4] Y. Feng, W. Cai, H. Yue, J. Xu, Y. Lin, J. Chen, and Z. Hu, "An improved X-means and isolation forest based methodology for network traffic anomaly detection," *PLoS ONE*, vol. 17, no. 1, e0263423, 2022.

[5] B. Pelletier, "On the statistical properties of the isolation forest anomaly detection method," *Electronic Journal of Statistics*, vol. 18, no. 2, pp. 4322–4381, 2024.

[6] M. Haque and A. Kashtiban, "Machine learning applications in power system fault diagnosis: Research advancements and perspectives," *Engineering Applications of Artificial Intelligence*, vol. 106, Art. no. 104504, 2021.

[7] F. Shakiba, M. Azizi, M. Zhou, and A. Abusorrah, "Application of machine learning methods in fault detection and classification of power transmission lines: A survey," *Artificial Intelligence Review*, vol. 56, pp. 5799–5836, 2023.

[8] W. D. Stevenson and J. J. Grainger, *Power System Analysis*. McGraw-Hill Education, 2020.

[9] P. Kundur, N. J. Balu, and M. G. Lauby, *Power System Stability and Control*. McGraw-Hill Education, 2021.

[10] G. Hannák and G. Horváth, "Bilateral-weighted online adaptive isolation forest for anomaly detection in streaming data," *Statistical Analysis and Data Mining: The ASA Data Science Journal*, 2023.

[11] M. He and X. Chen, "Anomaly detection algorithm for big data based on isolation forest algorithm," *Journal of Computational Methods in Sciences and Engineering*, vol. 25, no. 5, pp. 4143–4154, 2025.
[12] H. Saadat, *Power System Analysis*. PSA Publishing, 2022.
[13] J. Arrillaga and N. R. Watson, *Computer Modelling of Electrical Power Systems*, 3rd ed. Wiley, 2021.
[14] D. P. Kothari and I. J. Nagrath, *Modern Power System Analysis*. McGraw-Hill Education, 2021.
[15] O. I. Elgerd, *Electric Energy Systems Theory: An Introduction*. McGraw-Hill, 2020.
[16] S. Hariri, M. C. Kind, and R. J. Brunner, "Extended Isolation Forest," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 4, pp. 1479–1489, 2021, doi: 10.1109/TKDE.2019.2947676.
[17] M. Abdallah, M. Elhoseny, *et al.*, "Fuzzy Isolation Forest for anomaly detection," *Procedia Computer Science*, vol. 207, pp. 916–925, 2022.
[18] F. Leveni, G. W. Cassales, B. Pfahringer, A. Bifet, and G. Boracchi, "Online Isolation Forest," in *Proc. 41st Int. Conf. Mach. Learn.*, 2024, pp. 27288–27298.
[19] X. Li, X. Gao, B. Yan, C. Chen, B. Chen, J. Li, and J. Xu, "An approach of data anomaly detection in power dispatching streaming data based on isolation forest algorithm," *Power System Technology*, vol. 43, no. 4, 2023.
[20] A. Safari, M. Sabahi, and A. Oshnoei, "ResFaultyMan: An intelligent fault detection predictive model in power electronics systems using unsupervised learning isolation forest," *Heliyon*, vol. 10, e35243, 2024, doi: 10.1016/j.heliyon.2024.e35243.
[21] C. Huang and W. Gao, "Series arc fault diagnosis method for photovoltaic system based on ultrasonic sensor and isolation forest," *Electrical Technology*, vol. 26, no. 5, pp. 10–16, 2025.
[22] L. Cai and X. Lu, "Research on automatic fault location of 10 kV distribution network based on Isolation Forest algorithm," *Electric Engineering*, vol. 46, no. S1, pp. 158–159, 2025, doi: 10.19768/j.cnki.dgjs.2025.25.054.
[23] W. Wang and G. Wang, "A study of improved isolation forest algorithm for data management of transmission line defects and hazards," *Energy Informatics*, vol. 8, Art. no. 130, 2025, doi: 10.1186/s42162-025-00585-7.
[24] T. Xie, B. Li, Y. Zhen, H. Wei, K. Xu, and J. Huang, "Anomaly detection method for power dispatch streaming data based on adaptive isolation forest and self-supervised learning," *Frontiers in Physics*, vol. 13, Art. no. 1704495, 2026, doi: 10.3389/fphy.2025.1704495.
[25] J. Din, G. Wang, S. Hongsheng, and S. Ali, "Robust unsupervised anomaly detection in smart grids using median-MAD normalized isolation forest," *Electric Power Systems Research*, vol. 261, Art. no. 113473, 2026, doi: 10.1016/j.epsr.2026.113473.
[26] R. B. de Santis and M. A. Costa, "Extended Isolation Forests for fault detection in small hydroelectric plants," *Sustainability*, vol. 12, no. 16, Art. no. 6421, 2020, doi: 10.3390/su12166421.
[27] S. Wu, H. Ma, A. M. Alharbi, B. Wang, L. Xiong, S. Zhu, L. Qin, and G. Wang, "Integrated energy system based on isolation forest and dynamic orbit multivariate load forecasting," *Sustainability*, vol. 15, no. 20, Art. no. 15029, 2023, doi: 10.3390/su152015029.
[28] D. S. Pastor, J. V. B. Martínez, L. A. Muñoz, and C. C. Asensi, "Multivariate automatic tuning of isolation forest for anomaly detection in critical infrastructures: A solution for intelligent information systems," in *Proc. 17th Int. Joint Conf. Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2025)*, vol. 2, 2025, pp. 393–400.
[29] L. V. Utkin, A. Y. Ageev, and A. V. Konstantinov, "Improved anomaly detection by using the attention-based isolation forest," *arXiv*, 2022, doi: 10.48550/arXiv.2210.02558.
[30] S. Hariri, M. C. Kind, and R. J. Brunner, "Extended Isolation Forest," *IEEE Transactions on Knowledge and Data Engineering*, vol. 33, no. 4, pp. 1479–1489, 2021.