

REVIEW

Open Access



Application of machine learning in power grid fault detection and maintenance

David Olojede^{1*} , Stephen King¹ and Ian Jennions¹

*Correspondence:

David Olojede
olojededavid@outlook.com
¹Faculty of Engineering and
Applied Science, Cranfield
University, Bedford, United
Kingdom

Abstract

The power grid infrastructure serves as the backbone of modern society, providing essential electricity supply to meet the demands of various sectors. Ensuring a reliable and efficient power grid amidst increasing demand remains paramount. This paper provides a literature assessment of the United Kingdom's (UK) power grid, with a focus on fault occurrences, maintenance techniques, and the use of new technology for monitoring and maintenance. According to the research, insulation degradation is the most common source of power grid problems. The power grid's maintenance cycle is then investigated, including preventive, predictive, and corrective maintenance techniques. The study emphasises the significance of regular inspections, condition-based monitoring, and asset management strategies in improving grid dependability and longevity. The paper then addresses the concept of Integrated Vehicle Health Management (IVHM) and how it relates to power grid infrastructure. It studies the role of IVHM systems in real-time monitoring, diagnostics, and prognostics for grid assets, allowing for predictive maintenance and informed decision-making. Furthermore, the article studies the use of machine learning approaches to power grid health monitoring and maintenance. This article discusses machine learning methodologies such as supervised and unsupervised learning, as well as reinforcement learning, and how they are used in defect detection, classification, and predictive maintenance. Overall, this paper provides an overview of the UK power grid, its fault management strategies, maintenance cycles, and the integration of machine learning techniques for health monitoring and maintenance, offering insights into enhancing grid reliability and performance in the face of evolving challenges.

Keywords Electrical power grid, Machine learning, Condition based maintenance, Integrated vehicle health management (IVHM)

Introduction

Electrical power grids are essential in modern life, powering homes, industries, and important services like healthcare and transportation. Their dependability and longevity are critical to economic development and social stability. As global energy use increases, the grid becomes more interconnected and complex [1]. This complexity is further complicated by the integration of distributed generation, renewable energy, and digital

technologies, which makes it more susceptible to failures. A significant example of such vulnerability happened in August 2003, when a Major blackout affected over 500,000 people in London and Kent, which was caused by a transformer malfunction and faulty relay functioning, resulting in a widespread power outage. Although power was restored within two hours, the economic cost was significant. Using Ofgem's Value of Lost Load (VoLL) methodology, the cost of unserved energy during the thirty-seven-minute outage was estimated at between £6–9 million, with wider economic disruption likely raising the total to over £10 million [2]. Power cables play a crucial role in the power system network, functioning as the backbone of the power grid by transmitting electrical energy from generation stations to consumers. Insulation failure in power cables can lead to widespread power outages, impacting numerous customers. Thus, ensuring the reliability of power systems requires the early detection of incipient faults in power cables [3]. Partial discharge (PD) activity measurements are widely employed to assess the condition of insulation materials in high-voltage equipment [4]. According to IEC60270, PD is defined as a localized electrical discharge that partially bridges the insulation between two conductors. In power cables, PD refers to localized discharges within the insulation or occasionally along the cable surface, often caused by flaws, voids, or defects in the insulation material. These discharges are critical as they can degrade the insulation over time, potentially resulting in cable failure or breakdown [5]. Consequently, reliable PD detection system is essential for power companies to enhance safety and minimize the risk of cable breakdowns.

Hussain et al.[6] reported that insulation degradation and PD account for over 70 percent of high-voltage failures on the grid. These failures can result in costly unplanned outages and asset damage. Given the ageing infrastructure in many power networks and the increased strain on existing assets, there is a clear need for advanced, proactive strategies to monitor insulation health and detect faults at an early stage. This has led to growing interest in predictive maintenance and intelligent diagnostic techniques capable of operating in real time. To address these challenges, research in power system maintenance has shifted from reactive or schedule-based approaches to more intelligent, condition-based strategies.

Problem statement

Current fault detection methods predominantly rely on rigid, schedule-driven inspections that are not responsive to real-time operational conditions. Consequently, they struggle to detect subtle indicators of insulation deterioration, especially in aging infrastructure or under high electrical stress. For example, in the UK, National Grid primarily conducts snapshot-based assessments, such as helicopter-based visual and thermal imaging surveys [7]. While effective at identifying obvious defects like corrosion or overheating, these methods are not sensitive to early insulation stress or pre-PD deterioration. Machine learning (ML) techniques are increasingly applied to PD diagnostics, however, their use is typically reactive, with models deployed only after PD activity is detected. Thirdly, Prior studies have largely utilized supervised learning & unsupervised learning algorithms to classify PD type or location. These models and their use cases would be discussed further in later sections. Despite achieving high classification accuracies, these models are not designed to detect the onset of degradation but rather to analyse already active PD phenomena. In contrast, Integrated Vehicle Health

Management (IVHM)—a proven framework in aerospace and advanced manufacturing for continuous sensing, diagnostics, and prognostics—has seen limited application in the context of overhead power transmission systems. This study aims to bridge that gap by developing a novel machine learning-based IVHM approach tailored to high-voltage transmission infrastructure. The goal is to enable early-stage detection of insulation degradation and accurate classification of PD, ultimately improving the reliability and resilience of the UK's power grid.

Gap analysis

Research on partial discharge (PD) in high-voltage transmission systems has significantly expanded over the past decade. However, the dominant focus remains reactive, with most studies classifying insulation health in binary terms either faulty or healthy without modelling the continuous degradation process. This binary framing limits operators' ability to monitor the progression of PD and implement timely preventive maintenance. The absence of a graded health index or degradation trajectory hampers effective condition-based strategies and risk-informed decision-making. Machine learning (ML) methods dominate the PD diagnostics landscape, particularly supervised models such as support vector machines (SVMs), artificial neural networks (ANNs), and decision trees, which are primarily employed for post-fault classification. In contrast, reinforcement learning (RL)—which is well-suited for adaptive Maintenance scheduling and sequential decision-making under uncertainty remains largely unexplored. A structured Scopus search conducted in July 2025 revealed 8,610 publications relating to PD and ML, 9,171 addressing PD and degradation, but only 377 that mentioned RL, and just five that referred to Integrated Vehicle Health Management (IVHM). Despite its success in aerospace and advanced manufacturing, IVHM has seen little application in power transmission networks. Current PD monitoring practices in this domain remain predominantly snapshot-based—for example, periodic helicopter surveys using thermal and visual imaging—which often fail to capture early signs of insulation stress. This study aims to bridge these critical gaps by introducing a unified IVHM-inspired architecture for transmission systems. The proposed framework incorporates supervised ML for degradation classification, integrates RL for adaptive scheduling, and shifts the operational paradigm from reactive fault response to anticipatory health management.

Chapter summary and paper structure

This chapter introduced the context and motivation for this research, outlining the increasing vulnerability of high-voltage transmission systems to insulation-related faults and the limitations of current maintenance practices. It reviewed recent developments in CBM and machine learning applications, identified critical gaps in the literature, and formulated the problem that this study seeks to address. The chapter also presented the rationale for integrating IVHM and machine learning in transmission line monitoring. The remainder of this paper is structured as follows: "[Overview of the UK power grid](#)" provides an overview of the UK power grid, with a focus on its Transmission infrastructure. "[Faults on the electrical power grid](#)" examines the types and impacts of Transmission Line faults, particularly insulation degradation and partial discharge. "[Maintenance cycle of the power grid](#)" reviews Maintenance strategies, including preventive, predictive, and corrective approaches. "[IVHM & Health monitoring](#)" explores the principles

and components of IVHM and its potential application in power systems. "[Machine learning methodologies for health monitoring and maintenance](#)" evaluates Machine learning methods for fault detection, classification, and Maintenance decision-making, with special emphasis on reinforcement learning. "[Conclusion](#)" concludes the study, summarising the key findings and suggesting directions for future research.

Overview of the UK power grid

The power grid is an interconnected network of electric generators and loads, a transmission and distribution network that delivers electricity from the generation source to its end-users [1, 8]. The electrical power grid can be divided into three sections based on their voltage levels: The generation network, transmission network and the distribution network. These different sections of the grid are discussed in subsequent sections.

The generation network

Figure 1 above illustrates the hierarchical flow of electricity in the UK power grid based on voltage levels from the generation source to the end consumer. As shown in Fig. 1 the power system starts at the generation stage marked on the left side. The electricity generation process involves creating electrical energy through primary sources of energy [8, 9]. Currently, over 2,000 generation stations across a variety of power generation sources are in use in the UK. Many of these stations use renewable source of generation, and they tend to be typically smaller than conventional power plants [10]. Electricity generation sources are grouped into low carbon energy sources and high carbon energy sources. Low carbon energy sources produce energy with lower levels of carbon emissions while high carbon energy sources generate energy with high levels of carbon emissions. Examples of high carbon energy sources are coal, oil, natural gas and examples of low carbon energy sources are nuclear, solar, hydro, wind, and biomass [11]. Currently, most of the electricity generated in the UK comes from high-carbon sources with gas, coal and nuclear power stations providing the base load, they can generate continuously throughout the day [10]. However, there has been an increase in the use of low carbon energy sources in recent years. Over the last ten years, low carbon energy generation has experienced a progression from about 10% of the total energy generated to approximately 21% of the energy generated in the UK [12]. This is mostly due to the United Kingdom's low carbon transition plan. The UK's low carbon Transition plan aims to reduce emissions by at least 80% by 2050 through investment in energy efficiency and clean energy technologies [13]

Transmission network

The generated electricity is transmitted from the generation station to the distribution substation via a transmission network [8]. The transmission network, depicted in the middle section of Fig. 1, consists of interconnected high-voltage conductors that

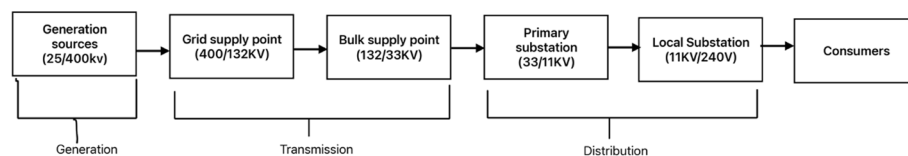


Fig. 1 Electrical power grid

facilitate the long-distance flow of electrical energy [8]. To prevent energy loss while Transmitting across long distances, electricity is Transmitted at high voltages, around 400KV or 132KV [1, 14]. The transmission network primarily transmits three-phase, high-voltage alternating current (HVAC). However, in certain cases, the transmission line transmits single-phase AC and high voltage direct current (HVDC) Examples of these are: electrification of the electric Trainlines in the UK, 2000MW England–France interconnector (connects the French and the English transmission systems), and the England–Netherlands 1000MW interconnector [15]. Some major industries take their electricity straight off the transmission line, but a majority of commercial and other end users receive electricity at much lower voltage levels. Transmission powerlines are divided into two categories: overhead powerlines and underground powerlines. Overhead power lines are structures used in power transmission and distribution to transmit electrical energy over long distances. They are made up of one or more electrical cables that are supported by towers and poles. Overhead lines are cheaper to implement compared to underground power lines. However, they are more prone to powerline defects than the underground powerline [16]. The UK Transmission network has over 24,000 km of overhead cables. Although underground cables are less prone to powerline problems compared to overhead powerline cables, they can still be susceptible to mechanical faults, such as thermal runaway due to the compacted soil acting as an insulator. From the name, underground cables are cables buried underneath the ground to transmit electricity. Underground cables are primarily used in urban areas where overhead power lines may be aesthetically unacceptable and pose safety risks, such as electrocution or falling cables. Additionally, they help minimize electromagnetic interference (EMI), which can disrupt communication networks. The UK Transmission network has over 1,000 km of underground cables [10].

Distribution network

Electricity is transmitted from the distribution substation to end consumers via the distribution network [1, 9]. The distribution network is depicted on the right-hand side of Fig. 1.

The distribution grid is grouped into the primary distribution network and the secondary distribution network [1]. The primary distribution is the high and medium voltage network. On the UK power grid, voltage is Transmitted at 33KV and 11KV. The secondary distribution is the low voltage network. The secondary distribution steps down electricity from the primary distribution network with step-down Transformers before being Transmitted to the end-users. They carry voltages ranging between 100V and 240V. There are three main types of distribution architecture: loop, radial or interconnected. The Loop Distribution system circles back to the point of origin after looping around the service area. Loss reduction, voltage management, and avoidance of cable congestion are all advantages of the loop network. In the radial distribution, different feeders emanate from a substation, feeding the distributors at one end of the system [17]. The type of architecture is commonly adopted in areas with low population density. As a result, a radial distribution system's most distinctive feature is that electricity flows in only one direction – it only ever flows in one direction, from high potential to low. They are also the most cost-effective to deploy. However, it is the least dependable since any failure on the power line can cause the feeder to shut down. Finally, an interconnected

distribution system comprises a ring main feeder powered by two or more substations or generating units. This type of architecture is implemented in major cities and metropolitan areas. This distribution architecture guarantee's reliability in the event of a fault at a feeder.

Synchronization on the power grid

As discussed in earlier sections, the power grid utilises a variety of generation sources to generate electricity. If the connection to the utility network is not properly synchronized and controlled, it can cause grid stability problems [18, 19]. As a result, before connecting the different electricity generation sources to the utility network, factors such as power generation, safe operation, and grid synchronisation must be considered. The goal of synchronisation is to monitor and automatically match the frequency of a generator with other sources on the network to ensure differences in voltage, frequency and phase angle are minimised [18, 19]. The UK grid operates as a wide area synchronous grid with a nominal frequency of 50 hertz.

Faults on the electrical power grid

Faults on the grid are abnormal conditions or disturbances that occur in an electrical power system. More than 70% of electrical power grid failures occur on the transmission and distribution side (Wang, 2017). These failures can cause an interruption in the delivery of energy, as well as potential equipment damage or safety issues. Faults on an electrical power grid can be categorized by their level of impedance or by their time interval of occurrence. These two categories of faults are discussed in subsequent sections.

Faults grouped by impedance level

Low impedance/Bolted fault

A bolted fault, also known as a short-circuit fault or a low impedance fault, is an electrical system fault state in which two or more conductors make direct contact with each other or with a grounded surface, resulting in a low impedance fault channel. This results in an almost zero resistance current flow route, resulting in a high fault current [20, 21]. These faults can be caused by, but are not limited to insulation failure, corrosion, loose connection, vandalism and external damage. Examples of these faults are described in Table 1 below.

High impedance fault

High impedance faults are faults that have enough impedance to prevent a protective device from operating [21, 23]. In some circumstances, high impedance failures may not result in outages but can cause significant power quality issues and could lead to serious equipment damage. High Impedance faults can either be intermittent or permanent faults. High impedance faults can be caused by, but not limited to, corrosion, insulation degradation, arcing, lose connections, environmental factors, and manufacturing defects. Examples of these faults are described in Table 2 below.

Table 1 Fault category, fault, and grid behaviour (Bolted fault)

No	Fault name	Grid behaviour	Key insights for this study	References
1	<i>Single line to ground (SLG):</i> A single line to ground fault occurs when one of the wires in a three-phase electrical system comes into contact with the ground or another grounded item.	Voltage at affected phase dips Voltage on non-faulted phases increases above the typical level. Phase current upstream of the fault on the affected line surges till protection trips	Could be an early indicator of insulation failure on one phase. Highlights how PD evolution on a single phase can be detected using ML to prevent further insulation breakdown.	[22]
2	<i>Double line to ground (DLG):</i> Like a single-phase fault, this occurs when two conductors in a three-phase power system come into contact with the ground or a grounded item at the same time.	The voltage across the two faulted phases decreases significantly due to the fault current flow. Meanwhile, the voltage of the healthy phase remains relatively stable.	Indicates more severe degradation. Highlights the need for robust ML-based multi-phase fault detection. Although DLG faults are less common but more severe, involving two-phase insulation failure. Highlights how progressive PD on multiple phases can evolve into catastrophic multi-phase faults if not detected early.	[22]
3	3-phase fault: A three-phase fault, also known as a symmetrical or balanced fault, is a fault that occurs simultaneously and equally on all three phases of a three-phase electrical power system.	High fault currents due to the simultaneous short circuiting of all three phases. A significant drop on voltage causes by the high currents that can lead to a blackout. Transient instability of the grid can be caused by the imbalance in currents and voltage	Though rare, these faults are severe. They indicate system-wide insulation failure, usually catastrophic. Demonstrates the cost and risk of not implementing predictive maintenance like PD-based monitoring.	[4, 22]
4	<i>Phase – Phase:</i> A phase-to-phase fault, also known as a line-to-line fault, is an electrical problem that occurs when two separate phase wires in a three-phase electrical power system make direct contact.	Current in the two phases involved in the fault increase significantly, causing an imbalance in the three-phase system. The voltage imbalance can lead to abnormal voltages across the system, affecting the operation of connected equipment. A phase-to-phase fault may lead to arcing at the fault location, resulting in transient voltages and currents	Occurs from weak spots between phases, often preceded by PD in stressed regions. Validates why insulation condition monitoring should cover inter-phase regions, not just phase-to-ground.	[5, 22]

Faults grouped by their time interval occurrence

Transient fault

Transient faults, also known as temporary faults, occur in an electrical power grid for a relatively short time. They often resolve themselves automatically without any manual intervention to correct the fault. These types of faults are monitored by protection systems installed within the power grid. Although transient faults are often self-correcting, they can still cause momentary power outages and voltage fluctuation. Hence power system operators employ the use of protective relays and circuit breakers to monitor and minimise the impact of transient faults to ensure a reliable operation of the grid.

Intermittent fault

An intermittent fault, is a malfunction that occurs periodically or infrequently. Unlike a transient fault, which appears for a brief period and then disappears, an intermittent fault appears at unpredictable intervals, making them difficult to diagnose and resolve.

Table 2 Fault category, fault, and grid behaviour (high impedance)

No	Fault name	Grid behaviour	Key insights	References
1	<i>Single line to ground with arcing</i> : Defined as a type of fault that happens when one of the phase conductors in a power system comes into contact with a high impedance material surface, causing an arc between the conductor and the high impedance material.	The fault current will flow from the faulted phase conductor to the ground. However, due to the presence of an arc, the fault current can be intermittent and may include high-frequency transients caused by the arc extinguishing and re-igniting. These transients can cause additional stress on the system and equipment. The presence of arcing at the fault location can cause rapid changes in current flow, leading to voltage fluctuations in the system. The voltage drop at the fault point may vary depending on the arc's characteristics and the impedance of the faulted section. The arc can be self-sustaining due to the continuous supply of fault current, even when the fault impedance is relatively high. Arc flash events can result in high temperatures, intense light, and the release of hot gases.	Highlights the importance of wideband PD sensors and ML models that capture time-domain and frequency-domain features. Arc-induced faults produce high-frequency transients similar to later PD characteristics.	Taft [22]
2	<i>Phase - phase, arcing</i> : A phase-to-phase fault with arcing is a type of fault in which two separate phase conductors in a power system come into direct contact with one other, resulting in the production of an arc between the faulted phases.	A phase-to-phase fault with arcing can result in a high fault current flowing between the faulted phases. The arc can cause the fault current to be intermittent and may exhibit high-frequency transients due to the arc's extinguishing and re-igniting. The presence of arcing at the fault location can cause rapid fluctuations in current flow, leading to voltage fluctuations in the system. The arc's characteristics and the fault impedance will influence the magnitude of the voltage fluctuations. Arcing generates intense heat and pressure that can lead to thermal and mechanical stress on equipment and conductors in the vicinity of the fault.	Arcing between phases produces signatures that resemble advanced PD. ML models must distinguish between sustained PD and full arcing faults to support timely intervention. Intermittent arc discharges accelerate insulation aging, similar to repeated PD events. Provides a strong link between PD signatures and evolving arcing faults, supporting your study's focus on PD pattern recognition for predictive maintenance.	Taft [22]

Permanent fault

A permanent fault also known as a prolonged fault or continuous fault, is a type of fault that remains in an electrical system until manually handled and corrected. Permanent faults consistently interrupt system operation until the damaged component or condition is fixed or replaced.

Evolving fault

Evolving faults, also known as multistage faults, are defects that begin as one kind or involve one phase or pair of phases but subsequently transform to another type or involve more phases over time. An example of this evolving fault is partial discharge (PD). PD is considered an evolving fault because it gradually degrades the insulation causing localized damage that evolves over time, increasing the likelihood of more frequent and severe discharges. Among the several types of faults that can affect the reliability and safety of the power grid, partial discharge (PD) faults stand out due to their subtle yet destructive nature. This study focuses on PD faults because they are often

precursors to insulation breakdown, leading to severe system failures. By understanding and simulating PD behaviour, particularly in high-voltage transmission lines, we can develop more effective detection and mitigation strategies to enhance grid resilience.

Insulation degradation and partial discharge (PD)

Insulation degradation describes the decrease of a material's insulating characteristics over time. Insulation degradation and partial discharge is responsible for 70 percent of high voltage problems on the power grid [6]. As a result, partial discharge can be utilised as a useful indicator for insulation degradation for assessing and monitoring insulation systems [6, 24]. PD happens when the insulation material fails due to significant electrical stress in the surrounding medium. PD can occur in a variety of high voltage equipment, including transformers, cables, and switchgear. PD can be induced by a variety of factors, including degradation of insulating materials, mechanical stress on the insulation, and contamination of the insulation material. PD can be defined as the breakdown of the insulating material under high voltage stress and occurs only in certain localised areas within the insulation. PD can be caused by the following but not limited to manufacturing defects in the insulation such as voids, cracks, air pockets and aging of the insulation material due to thermal, mechanical, or electrical stress. Figure 2 below shows a simple cross-section of a transmission cable with a void.

When the electric field in a localized defect (such as an air void in solid insulation) exceeds the local dielectric strength of the insulation material, the air inside ionizes, creating a discharge. Although PD does not immediately cause total insulation failure, it gradually weakens the insulation material, eventually leading to complete breakdown over time. It emits high-frequency electromagnetic waves, ultrasonic signals, and sometimes gas and light emissions, all of which can be detected using specialized sensors. Common PD detection methods include ultrasonic analysis, radiofrequency measurements, and electrical pulse monitoring to identify early signs of insulation degradation. A detailed discussion of partial discharge and its measurement techniques is discussed in Section "Insulation degradation curve and measurement technique" below. The magnitude of the PD does not always correlate with the amount of damage on a conductor [6]. Some PD of a small magnitude may cause more damage to higher voltage cables than lower voltage cables. PD pattern representation entails methods for studying and

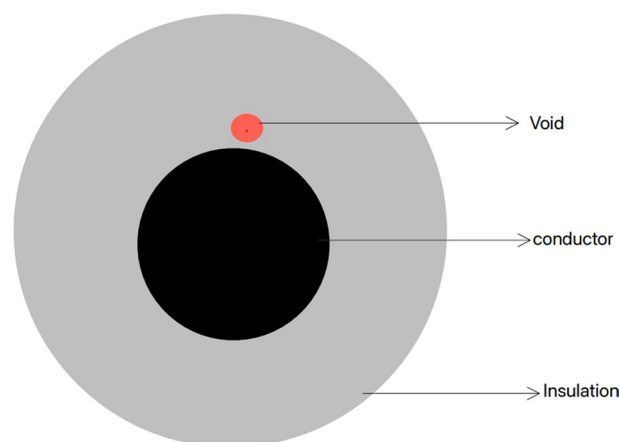


Fig. 2 Simple cross section of a transmission cable

understanding the electrical signals produced by PD in electrical equipment. There are two main approaches to identify and characterise PD activities on the power grid: Time resolved partial discharge (TRPD), and phase resolved partial discharge (PRPD). TRPD representation involves analysing and interpreting PD signals based on their temporal characteristics. This strategy focuses on recording and understanding the dynamics of PD pulses through the time of PD occurrence. Time-resolved analysis alone cannot accurately pinpoint the location of PD events within the insulation system. It may identify the presence of PD but not the origin, because TRPD does not consider the phase angle of the PD signal in relation to the power system voltage waveform [25]. This means it may not provide information about the exact timing or phase location of PD events within the electrical system. Time-resolved analysis also only focuses on the temporal characteristics of PD signals, such as pulse duration and repetition rate [25]. It may also have limitations in distinguishing between types of PD sources such as corona discharge, surface discharge or internal discharge which may exhibit similar temporal patterns. In PRPD the amplitude of PD pulses is plotted against their respective phase angles inside the AC voltage cycle creating a graphical depiction. The graphical depiction of the plotted pulses aids in the analysis and interpretation of PD signal characteristics, identifying discharge sources, and gaining insight into the status of electrical equipment as illustrated in Fig. 3 below. Phase resolved pattern representation is more widely accepted for PD classification because of its ability to represent the physical process at the PD location. Also, the phase resolved pattern recognition waveform can better differentiate between different PD sources and types using the phase angle.

Insulation degradation curve and measurement technique

The insulation potential-functional (P-F) failure curve is a graphical representation used in reliability engineering and preventive maintenance that shows the relationship between the failure rate of a system and the effectiveness of preventive measures over time. The aim is to assess how maintenance can reduce the risk of system failure. Figure 4 below represents the Condition-Based Monitoring (CBM) approach used for detecting equipment failures. The graph shows how equipment condition deteriorates over time, from initial detection of a potential issue to actual failure, with various CBM technologies employed at different stages to monitor and detect failures. It is used to assess the condition of the insulation and predict remaining useful life.

Figure 4 above illustrates the complete degradation cycle indicating the characteristics at each stage and the different measuring sensors that could be used to detect PD. At

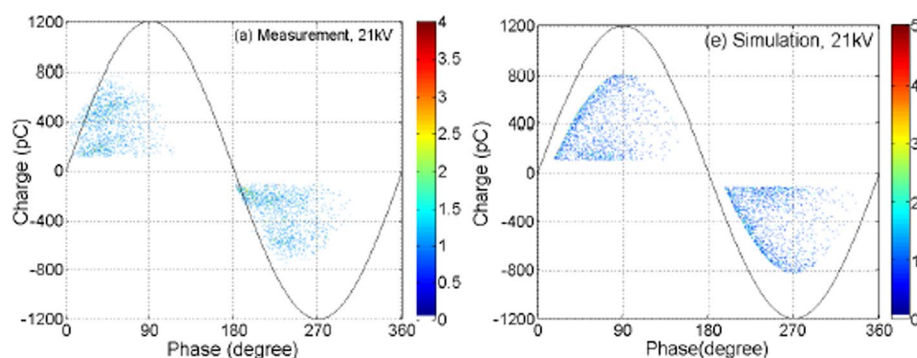


Fig. 3 PRPD representation adapted from [3]

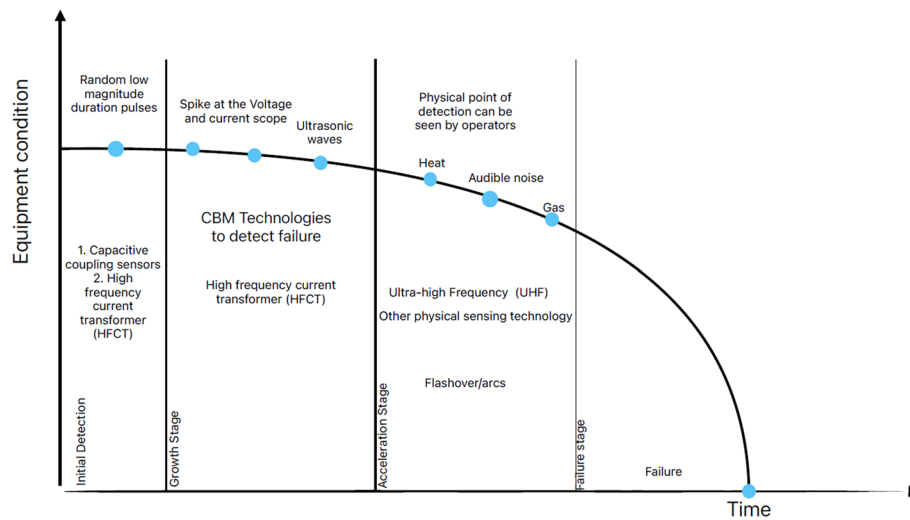


Fig. 4 Insulation degradation curve

the onset stage of partial discharge, the degradation begins with minor changes in the insulation quality. When a PD pulse occurs, electrons move rapidly from one side of the void to the other. Because electrons travel at near-light speed over a short distance, the pulse is often only a few nanoseconds long [6]. These electrons carry a charge, and each discharge produces a current pulse. In addition to the electron current flow, positive ions will flow in the opposite direction (which occurs when electrons are ionised from gas molecules). The PD current in the void causes a disturbance, resulting in pulse current and voltage flowing away from the PD point. These changes may not be immediately noticeable and are often difficult to detect. At this stage of the PD, electrical detection methods are commonly used. The common electrical detection approaches used for condition monitoring of PD is the capacitor coupling method [6]. A PD coupling device is sensitive to high frequencies and can be used to measure PD signals in the form of voltage impulses that propagate through the insulation. The measuring test is performed using the coupling capacitor to the high voltage components for continuous monitoring [6].

At the growth stage, the PDs become more frequent and intense indicating the growth of the insulation degradation. At this stage, depending on the magnitude of the PD, the insulation condition is often repairable by splicing, re-terminating, or applying insulating materials to restore the cable's integrity. There would be an increase in the number of discharge pulses which indicates further deterioration of the insulation. PD analysis using the electromagnetic (EM) method shows that it is possible to get a linear connection between the PD charge and potential signal sources at a fixed PD position and consistent EM signal propagation [26]. Ultra-High Frequency (UHF) detection is the process of detecting electromagnetic signals that are produced by PD occurrences in the UHF region (usually 300 MHz to 3 GHz) [27]. Localised electrical discharges caused by PD within insulating materials release electromagnetic waves, including UHF signals. UHF detection process captures electromagnetic signals in the UHF range which are generated by the rapid charge movement during PD events. The UHF sensors are installed through specialized couplers placed near the high voltage equipment. UHF detection is also affected by radio interference and noise from the detection device. The noise can be

eliminated by applying the appropriate denoising technique [28]. Various current transformer types have been coupled with UHF sensors, including Rogowski coils, high-frequency current transformers, and radiofrequency current transformers, have also been explored as sensors for the PD detection for power grid maintenance [29, 30].

At the acceleration stage, the PD pulses start to occur at a rapid rate and shorter intervals, increased PD frequency is experienced, elevated PD magnitudes, more severe discharges are also experienced. Other physical signs like arcing, audible noise and smell of gas would also be noticed at this stage. The insulation's ability to withstand the voltage stress becomes increasingly compromised and the risk of failure increases. At this stage, the ultrasonic signals Manifest as pressure waves with distinct characteristics. These signals have frequencies ranging from 40kHz to 400kHz. These ultrasonic signals can be localized and detected using Acoustic emissions (AE) techniques [31, 32]. The drawbacks of this method are the complex behaviour of acoustic emission, and low intensity detected signals. AE detection devices explored include microphones, piezoelectric transducers, accelerometers, and fibre optic sensors [32], (Kim et al. 2017.). By denoising and optimization techniques, noise due to internal transformer design can be eliminated, and multi-PD sources can be detected. Dissolve gas analysis makes use of the idea of detecting gases emitted because of the insulating material's deterioration. One of the most produced gases during PD events is hydrogen. It results from the dissociation of hydrogen-containing molecules within the insulating substance. Monitoring hydrogen is important because it indicates the risk of arcing and other high-energy discharges, which accelerate insulation degradation. Although it is a valuable diagnostic tool in enclosed systems, it is less applicable to overhead lines, where other techniques are better suited for fault detection and insulation assessment. Another gas that may be created during PD episodes is methane. It occurs when the insulation's carbon-containing components disintegrate. Like hydrogen, the presence of methane can indicate increased PD activity and arcing potential. Other gases that could be identified at the PD emission include ethylene, acetylene, carbon monoxide, carbon dioxide, and water vapour. However, PD detection at this stage might be extremely risky and ineffective due to the proximity to the failure stage.

The failure stage is the final stage of insulation degradation, where the insulation dielectric strength has decreased to the point where it can no longer withstand the applied voltage stress. At this stage, the insulation experiences a complete breakdown, leading to electrical failure. Given the degradation process described above, and relevant sensing options, the P-F curve can serve as a useful conceptual guide for implementing monitoring solutions to assess the health status of high-voltage electrical equipment. Operators can decide whether to maintain, repair, or replace the insulation to avoid catastrophic failures and guarantee the steady operation of the electrical system by routinely assessing the insulation and comparing the results with the P-F curve.

Maintenance cycle of the power grid

Maintenance is the act of preserving a condition or state of a component [33]. In some circumstances, the cost of maintenance and the repercussions of failure can exceed the equipment's original cost. As a result, maintenance optimisation becomes vital, with the goal of attaining the required level of reliability while maintaining the system at a low overall cost and meeting customer and regulatory needs [34, 35].

Types of maintenance on the power grid

Corrective/reactive maintenance

Corrective maintenance is performed after a fault has been found with the intention of returning the component to a state where it can perform a needed function [36]. The component is used until it breaks down. Corrective maintenance may be viewed as a last resort however, in some cases, it may be used as an immediate reaction, or strategically. The maintenance policy implemented will be determined by an understanding of the likelihood/risk of failure, the consequence of failure and the cost of impact of the failure. There are several circumstances when this maintenance policy is the best option, such as circuit breakers and fuses. Given that resources are focused on other, more critical assets, corrective maintenance might be the appropriate method for a component group [36] [37]. A downside of this maintenance policy is a fault can shut down the entire system until it is repaired [38].

Preventive maintenance

Preventive maintenance is the routine and regular repair of equipment and assets to keep them functioning and prevent costly unplanned downtime due to unexpected equipment breakdown [36]. There are two types of preventive maintenance: scheduled maintenance and condition-based maintenance [33].

Scheduled maintenance refers to any planned repair or routine inspection carried out on a system at predetermined intervals to ensure optimal performance, prevent breakdowns and extend the life of the asset. The predetermined intervals are either given by the manufacturer or from network operator experience [33–35, 39]. Scheduled maintenance combines acceptable availability with comparable high maintenance costs based on interval length. Performing scheduled maintenance at these generalized intervals can lead to unnecessary component maintenance and increasing costs [40]. Nonetheless, legislation sometimes mandates routine scheduled maintenance for high-value assets.

Condition-based maintenance

Is conducted depending on the system's performance and parameters [33]. Performance and parameter monitoring may be scheduled on request, or monitoring may be continuous. Machine data can be collected through the onsite monitoring equipment, indicating the maintenance condition of the equipment to assess asset health. Maintenance, in this case, is initiated when the condition variables pass a pre-set threshold [36]. In most cases, these thresholds are based on some measure of degradation and are set by industry experts. Condition-based maintenance is used on extremely high voltage, high voltage, and medium voltage maintenance of the power grid [39].

Proactive maintenance

Proactive maintenance is a maintenance management strategy that focuses on preventing equipment failures and minimising downtime by resolving issues before they cause breakdowns [41]. It entails planned, methodical operations to assure equipment reliability, availability, and optimal performance. Proactive maintenance seeks to discover and correct potential problems as early as possible, thereby improving asset efficiency and durability.

Reliability centred maintenance

Reliability-Centred Maintenance (RCM) is a system for completing maintenance chores in an efficient, cost-effective, dependable, and safe manner by establishing the best maintenance plan for various assets [41, 42]. RCM investigates the performance standards of the system, the cause of each failure, the impact of the failure on the system, how to prevent and contain the failure [39]. It was initially developed by the aerospace industry and later adopted by other sectors. RCM was introduced to a nuclear power station in 1978. In 1990 it was introduced to hydroelectric power generation [42]. RCM can be applied to individual pieces of equipment as well as the entire network [39, 43]. Studies show at the component level, RCM has been applied on voltage regulators, underground and overhead transmission networks, circuit breakers and power transformers ([34, 35], Kim et al. 2021, [33]. One of the main drawbacks of Reliability-Centred Maintenance (RCM) is the need for an in-depth analysis of each asset, its potential failure modes, and their impact on overall system reliability. This process can be complex and requires significant resources and expertise to evaluate every component and system thoroughly. As a result, it can become expensive, especially in terms of time, personnel, and the tools needed to perform the detailed analysis, making it a resource intensive.

Digital twins and AI for maintenance planning and reduced MTTR

Digital Twin (DT) technology is increasingly seen as a critical instrument for altering maintenance techniques in modern power systems. A digital twin is best defined as a high-resolution, real-time digital copy of a physical object or process. Unlike static models, DTs are dynamically linked to their physical counterparts through continuous data interchange, allowing for real-time monitoring, simulation, and predictive analysis. This bidirectional synchronisation enables DTs to represent not only the current state, but also the projected future behaviour of grid components under changing operational conditions [44].

In the context of electricity transmission and distribution networks, Digital Twins provide an integrated platform that combines sensor data, AI-driven analytics, and physics models. This fusion enables the move from scheduled and reactive maintenance to data-driven, condition-based treatments. DTs allow operators to detect early warning signs of failure, estimate component Remaining Useful Life (RUL), and schedule maintenance with greater accuracy and efficiency by replicating asset performance and degradation trajectories in real time [45].

Several practical applications highlight the expanding use of Digital Twins in power systems. For example, DTs have been used to monitor transformer insulation health by combining thermal stress data with simulation models, resulting in better failure prediction and maintenance timing [46].

In fault diagnostics, DTs can be combined with machine learning models trained on phasor measurement unit (PMU) data to identify and localise Transmission Line faults significantly faster than conventional inspection techniques, resulting in up to a 70% reduction in fault localisation time. Furthermore, DTs help with load forecasting, contingency simulation, and frequency management by modelling system behaviour under changing demand and fault conditions.

Beyond operational diagnostics, Digital Twins offer asset risk prioritising. By generating condition-based health indices in real time, DT platforms enable grid operators

to evaluate assets based on failure likelihood and criticality, resulting in more strategic resource allocation and increased grid dependability. Furthermore, DTs are increasingly being used to simulate cyber-physical attack scenarios, which improve grid resilience by testing and evaluating cybersecurity solutions in a virtual environment.

Although DTs have great promise, there are still significant hurdles. These include high implementation costs, the complexities of integrating several data sources, and concerns about data security and interoperability across platforms [44]. Nonetheless, the combination of DTs and AI provides a strong foundation for current grid maintenance, consistent with the principles of Integrated Vehicle Health Management (IVHM). This framework enables real-time diagnostics, prognostics, and intelligent decision-making, hence improving system resilience, reliability, and sustainability.

IVHM & Health monitoring

Integrated Vehicle Health Management (IVHM)

The term "IVHM" refers to a system's unified capability to assess a member system's current or future state of health and integrate that health within a framework of available resources and operational demand [47]. Due to the high initial cost of design and installation, IVHM is better suited for high-value, complex products. In this context, complex products can be electrical, mechanical, electromechanical, or a combination of the three. IVHM is used in a wide range of complex systems, including aircraft, ships, automobiles, trains, medical scanners, power generation equipment, and power transmission networks.

IVHM Cycle

The Integrated Vehicle Health Management (IVHM) cycle is a systematic framework for monitoring, diagnosing, and predicting the health of vehicle systems, encompassing five key stages: sense, acquire, transfer, analyse, and act as seen Fig. 5 below. The sense stage involves the collection of data through various sensors, while acquire focuses on gathering and preparing this data for processing [47]. The transfer stage transmits the collected data to centralized systems, where the analyse phase utilizes advanced algorithms and machine learning techniques to extract actionable insights, such as fault diagnosis and predictive maintenance. Finally, the act phase implements corrective, or optimization actions based on the analysis. Over time, the dominance within the IVHM cycle has shifted significantly. Two decades ago, the sense phase was the primary focus due to the need for reliable and robust sensor technologies, as early systems were constrained by limited sensor availability and integration capabilities. Today, advancements in data science, machine learning, and computational power have elevated the analyse phase to prominence, as it enables the extraction of actionable insights from vast sensor datasets in real time, enhancing vehicle reliability, safety, and operational efficiency. This evolution highlights how technological advancements have reshaped the priorities within the IVHM cycle, emphasizing the growing importance of transforming raw data into meaningful, proactive decisions.



Fig. 5 IVHM Cycle

Synergy between IVHM and the power grid

IVHM and power grids share an underlying synergy due to their reliance on continuous monitoring, complex interdependent systems, and the need for predictive maintenance. While IVHM has been applied in aircraft health management, its core principles leveraging sensor data for real-time health assessment and informed decision-making—are equally valuable in the context of power grids. Both domains benefit from advancements in sensor technology, enabling more intelligent and proactive management of their assets.

Power grids, like aircraft systems, are critical assets that demand high reliability and efficiency. They consist of intricate networks of transformers, substations, power lines, and other components, all of which must work seamlessly to ensure uninterrupted energy distribution. Embedded sensors within these systems provide a wealth of performance and health data, like the sensor networks used in aerospace for IVHM. These sensor-driven insights are instrumental in monitoring the condition of critical infrastructure, detecting anomalies, and predicting potential disruptions. The shared focus on preventing failures before they occur highlights the synergy between IVHM and power grids. In aerospace, IVHM ensures the reliability and efficiency of flights by predicting component failures and enabling timely maintenance. Similarly, in power grids, IVHM principles can be applied to identify early signs of degradation, wear, or faults, thereby enhancing grid reliability, and minimizing unplanned downtime. Moreover, the transition toward renewable energy sources and the increasing complexity of modern power grids makes the adoption of IVHM principles even more critical. By leveraging IVHM techniques, power grid operators can manage the variability and distributed nature of renewable energy sources more effectively, ensuring stable and reliable energy delivery. This proactive approach aligns with the overarching goals of IVHM, fostering resilience and sustainability in essential systems. Advancements in machine learning further amplify the potential of IVHM in power grid applications. Machine learning enables the analysis of vast amounts of sensor data to uncover patterns, predict failures, and enhance operational efficiency. By applying machine learning techniques, power grid operators can detect subtle signs of degradation in transformers, identify line faults, and predict substation equipment failures. These algorithms not only provide predictive insights but also optimize maintenance strategies, reducing downtime and improving the overall reliability of the grid. As the power grid becomes increasingly reliant on smart infrastructure, machine learning's role in IVHM frameworks is crucial to ensuring robust and efficient energy distribution.

Business value of IVHM in power grid

Integrated Vehicle Health Management (IVHM) delivers substantial business value to power grid operations by enhancing cost efficiency, reducing operational disruption, and improving safety outcomes. Through the integration of diagnostics and prognostics, IVHM enables the early detection of emerging faults, thereby facilitating timely maintenance interventions before failures escalate, significantly reducing unplanned maintenance costs. Moreover, IVHM supports a shift from reactive to proactive maintenance strategies. By continuously monitoring system health, operators can schedule maintenance activities in advance, thereby minimising service interruptions. As demonstrated by [33], condition monitoring increases system availability and helps justify the initial

investment in IVHM by reducing both downtime and operational disruptions. From a safety perspective, IVHM acts as a preventative safeguard, detecting incipient faults early and ensuring that critical repairs are prioritised without the inefficiencies of over-maintenance. In this way, IVHM not only optimises asset management but also contributes to a more resilient and secure grid infrastructure.

Decentralised IVHM and scalability on large complex systems

The increasing digitalisation of smart grids, characterised by high data volumes from distributed energy resources (DERs) and Internet of Things (IoT) sensors, has prompted a paradigm shift from centralised analytics to decentralised, edge-based diagnostic frameworks. Rather than transmitting all sensor data to a remote control centre, lightweight artificial intelligence (AI) models are now being deployed at substations, DER units, and distribution-level assets. This strategy facilitates near real-time fault diagnostics while reducing network communication loads and dependency on central infrastructure. Edge Intelligence (EI)—the integration of edge computing and AI—supports both on-site inference and, in some cases, federated model updates. By relocating computation closer to the data source, EI enhances the responsiveness of power systems and mitigates constraints related to network bandwidth, latency, and privacy. [48] highlight that edge-deployed AI systems enable real-time condition monitoring and anomaly detection, effectively addressing the limitations inherent in cloud-centric architectures. Their review further identifies decentralised, low-latency applications as key opportunities for improving smart grid resilience and scalability. In critical fault scenarios such as high impedance faults (HIFs), which typically produce low-magnitude and intermittent signatures, real-time local processing is essential. [49] demonstrate the deployment of deep learning models on embedded edge devices for HIF detection, achieving a latency of 115.2 ms and a detection accuracy of 98.67%. The system performs feature extraction and fault classification entirely at the edge, thereby reducing the reliance on high-throughput communication links and enabling prompt response.

Singh et al. [50] present the FaultEdge framework, which employs AI-driven predictive analytics and anomaly detection mechanisms directly at the grid edge. Their model introduces adaptive reconfiguration techniques that anticipate and mitigate potential failures, enhancing fault tolerance and operational reliability even under intermittent connectivity conditions. This proactive edge-based framework addresses key requirements for decentralised power system monitoring in dynamic environments. To meet the computational and energy constraints of embedded grid nodes, [51] advocate the use of Tiny Machine Learning (TinyML) approaches. Their study explores the integration of compact deep learning models into ultra-low-power microcontrollers, capable of inference at milliwatt-level power budgets. These TinyML deployments are particularly advantageous for monitoring tasks in remote substations and resource-limited distribution infrastructure.

Additionally, the orchestration of computational tasks across edge, fog, and cloud layers remains a central consideration. [52] analyse offloading strategies that balance workload distribution, energy efficiency, and communication latency. They propose reinforcement learning and metaheuristic algorithms to optimise dynamic task allocation in heterogeneous smart grid environments.

Collectively, these approaches represent a transition toward scalable, decentralised health monitoring frameworks for smart grids. By leveraging lightweight AI at the grid edge, utilities can reduce communication overheads, accelerate fault response, and improve the overall robustness of power system diagnostics.

Cybersecurity in IVHM-Enabled power systems

The integration of Integrated Vehicle Health Management (IVHM) into power grid infrastructures offers substantial benefits through enhanced connectivity, made possible by IoT-enabled sensors, remote diagnostics, and cloud-based analytics. This interconnected architecture facilitates real-time monitoring and intelligent decision-making for predictive maintenance. However, it also introduces new cybersecurity vulnerabilities. Malicious actions such as false data injection, sensor spoofing, and denial-of-service (DoS) attacks pose a significant risk to data integrity and operational reliability by either concealing actual faults or triggering unnecessary maintenance operations [53, 54]. To safeguard these systems, cybersecurity must be embedded into IVHM frameworks at the architectural level. Recommended countermeasures include end-to-end encryption protocols (e.g., AES, SSL/TLS), secure communication infrastructures, and AI-enhanced intrusion detection systems for real-time anomaly identification [53, 54]. Additionally, blockchain-based authentication techniques have been proposed to enhance sensor data provenance and trustworthiness in distributed smart grid environments. Integrating such multilayered cybersecurity mechanisms ensures that IVHM-enabled maintenance decisions are grounded in accurate and uncompromised system health data, thereby supporting grid resilience, safety, and operational continuity.

Machine learning methodologies for health monitoring and maintenance

Numerous studies have concentrated on using machine learning and optimisation techniques to monitor the condition of power system components as well as to improve and enhance the accuracy of diagnosing faults. A concise and accepted definition of machine learning can vary depending on its application and the problem its being used to solve. Regardless of the area of application and its function, machine learning is described as a methodology that focuses on combining statistical techniques and data-driven methods to enable computers to learn from data and make intelligent judgements using algorithms and models. Machine learning algorithms are designed to automatically learn from vast amounts of data without the need for explicit programming for each individual circumstance [55]. These algorithms can spot patterns, correlations, and trends in data, allowing the system to make predictions and classifications, on unseen data. Machine learning methodology has several learning regimes as discussed below.

Machine learning regimes

Machine learning regimes are paradigms or frameworks for developing and applying machine learning Algorithms. Examples of these regimes include supervised and unsupervised learning, transfer learning, reinforcement learning, and semi-supervised learning. Although transfer learning and semi-supervised learning have demonstrated significant success and popularity in applications such as image and voice recognition, their use cases in fault detection and maintenance, particularly in the context of partial discharge detection, are limited. Due to this lack of relevant references and applications

in the literature, these learning regimes have been excluded from consideration in this paper.

As discussed earlier in section "Insulation degradation and Partial discharge (PD).", insulation degradation accounts for 60% of high-voltage issues on the power grid. Consequently, it serves as a valuable diagnostic tool for assessing and monitoring insulation systems [6, 24]. Therefore, this paper focuses on machine learning applications for partial discharge fault classification. These machine learning algorithms have been grouped into their respective learning regimes as seen in Fig. 6 above. These machine learning regimes and applications are further discussed in the following subsections below.

Supervised learning

Supervised learning is a statistical machine learning regime in which an algorithm learns to make classifications based on labelled training data. In supervised learning, the algorithm is given a dataset comprising input features and their corresponding labels. The algorithm's purpose is to learn a mapping from inputs to outputs so that it can accurately predict outputs for new inputs [55, 56]. The learning problem can be expressed as:

$$\theta = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^N \mathcal{L}(f_{\theta}(x_i) y_i)$$

where (x_i) is the input y_i is the label $\mathcal{L}$ is the mean square error or cross entropy [57].

This formulation provides a unifying mathematical foundation for a wide range of supervised learning algorithms, including linear regression, logistic regression and support vector machines

Most machine learning approaches used for identifying, classifying, and localising partial discharges (PDs) in various areas of the power grid have relied on supervised learning techniques. Below, some prominent supervised learning techniques applied to PD detection are discussed. One of these supervised learning techniques is the support vector machine (SVM). At its foundation, SVM aims to identify the optimum hyperplane for separating data points into multiple classes. The goal is to find the hyperplane that maximises the margin, which is the distance between the hyperplane and the nearest data points in each class. These closest data points are known as support vectors. Although SVMs were originally designed for linearly separable data, they can be extended to

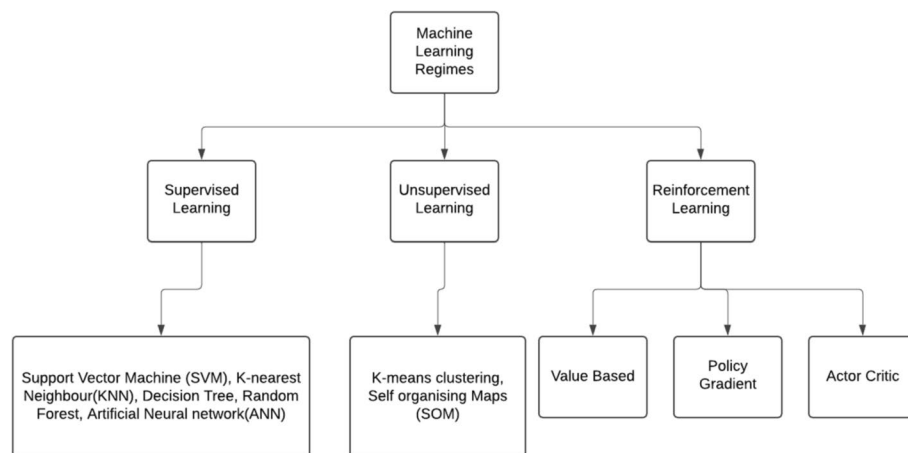


Fig. 6 Machine learning regimes

manage nonlinear data by using kernels functions. Kernels transform the original feature space into a higher-dimensional space where data points can become linearly separable. Common kernel functions include the linear, polynomial, radial basis function (RBF), and sigmoid kernels. [24] proposed the use of a wide-bandwidth PD online measuring system composed of a radio frequency current transducer (RFCT) sensor and a digital storage oscilloscope. The wide bandwidth PD sensor's measurement data was pre-processed using wavelet analysis. Correlation analysis was then used to analyse the data and categorise the discharges into distinct groups. SVM was then employed to distinguish between the various PD sources.

Tang et al. [58] presented a feature selection strategy based on minimum redundancy and Maximum relevance for PD severity evaluation using an SVM classifier on gas insulated switchgear. The suggested approach uses statistical PRPD features, including the skewness of the positive and negative half cycles, the whole cycle, amplitude, the discharge time ratio between positive and negative half cycles, and the cross-correlation coefficient of positive and negative half cycles, all extracted from the UHF signal as input. The SVM classifier achieved an accuracy of 88%- 93% across the severity levels. For high-voltage DC lines, [59] created an SVM-based PD detection algorithm. Pulse signal statistical data and discrete wavelet transform (DWT) based features were fed into the SVM algorithm, the findings revealed that DWT-based features were more reliable than others. [60] developed an ML based classification method to detect different types of PD in power transformers using acoustic emission signals. Energy features extracted by DWT are used as the input and SVM. The study showed that SVM achieved the best performance compared to k-nearest neighbour (kNN), Decision Tree (DT), and ensemble bagged DT. The main disadvantage is that SVMs are sensitive to the optimal choice of kernel. Additionally, they are computationally inefficient with a large dataset [61]. Consequently, more feature engineering is required for SVMs. Another supervised learning regime is the k nearest neighbour (kNN). The kNN uses a similarity metric to calculate the distance between each data point. Data points with similar sample features are referred to as near neighbours. Outliers are those that deviate significantly from other data points. The distance between an outlier and its nearest neighbour, known as the anomaly index, is used by kNN to detect anomalies [62]. kNN have been applied for various PD applications on the power grid. PD localisation [63, 64], PD classification [65–67], and PD detection [68]. Although kNN is simple to implement it has fewer application in detecting PD on the power grid because the computational cost increases significantly when the sample size is large, and the power grid is a complex system with a large data size. Another example of supervised learning is the decision tree.

Decision trees (DT) are a graphical representation of a decision-making process that resembles a flowchart with nodes and branches. Each internal node represents a decision or test based on a specific feature property. Each branch corresponds to the outcome of that test, while each leaf or terminal node represents a class label or, in the case of regression tasks, a numerical value [69]. A DT- based classifier can be applied to PD classification [70]. This research by [70] described the use of machine learning approaches for reliable partial discharge categorisation in oil-paper insulation systems. Classifiers include Random Forest, Gradient Boosting, SVM, Decision Tree (DT), and Linear Discriminant Analysis (LDA). Random Forest combined with Gradient Boosting algorithms demonstrated superior classification accuracy for partial discharge (PD)

detection compared to other algorithms. These algorithms demonstrated improved classification accuracy in scenarios where the test and training data are obtained under varied conditions, such as differences in location or oil temperature.

Random Forest is a supervised machine learning algorithm that constructs multiple decision trees during training and combines their outputs to improve accuracy. The Random Forest algorithm works by constructing a multitude of decision trees during the training phase. Each tree is built using a random subset of the data and features. This introduces diversity among the trees, reducing the risk of overfitting. When making predictions, the forest combines the outputs of all trees through majority voting (for classification) or averaging (for regression), leading to a more robust and accurate result [63, 64] proposed a PD localisation method for substations that employs wavelet packet decomposition to extract location-dependent characteristics from PD signals recorded by three Radio frequency sensors and RF regression to localise PD. [71] created an FFT-based feature and use various ML classifiers to achieve acoustic PD classification in a high-voltage bushing system. Even when the signal is distorted with noise, the RF classifier outperforms the other ML classifiers.

Artificial neural networks (ANN) are a collection of interconnected nodes called neurons. Each neuron receives an input, performs a weighted sum, applies an activation function, and produces an output. Typically, neurons are aggregated into layers. Different layers may perform different transformations on their inputs. When used for anomaly detection, neural networks can function as both one-class and multi-class classifiers. Neural networks are trained using both normal and faulty instances. However, in certain scenarios, the neural network could be trained on normal instances and then evaluated against anomaly examples. Following that, the network should be able to identify anomalous instances in a test data set. Soltani et al [72] developed a denoising algorithm using ANN, aiming to suppress white noise in PD diagnostics applications using R-F signals. Other BPNN use case include [73] where the authors used the wavelet-entropy vector extracted from transient earth voltage time-domain signal and BPNN to detect and classify PD [74, 75] applied DWT to extract distinguishable features from ultrasonic signals of PD and then employed BPNN based on wavelet neural network to achieve PD classification in an oil-filled submarine cable terminal. Neural networks have low interpretability because of their black-box nature. Table 3 presents a summary of selected supervised machine learning algorithms that have been applied to partial discharge (PD) detection, classification, and analysis in high-voltage insulation systems. The table highlights the algorithm type, key features or input representations, and the main application context as reported in the literature.

Unsupervised learning

Unsupervised learning is a statistical ML regime that involves an algorithm learning patterns and structures from unlabelled data. Unsupervised learning, as opposed to supervised learning, in which the algorithm is given labelled examples to learn from, entails the detection of similarities within data without specified labels. Unsupervised learning's primary purpose is to discover hidden patterns, groupings, or representations in data. A widely studied example of unsupervised learning is clustering. Among clustering techniques, the most common is *k-means clustering*, which seeks to minimise the within-cluster variance through the following optimisation problem

Table 3 Summary of supervised machine learning models

Study	ML method	Sensors/Data	Feature extraction	Task	Performance/key insight
Hao and Lewin [24]	SVM	RFCT	Cross-correlation, FFT	PD/noise classification	Effective noise filtering and PD signal detection
Tang et al. [58]	SVM	UHF	PRPD statistical features	Multi-level classification	Accuracy up to 93%; using selected PRPD features
Kunicki and Wotzka [60]	SVM, k-NN, Decision Tree	Acoustic Emission	DWT	PD type recognition	SVM outperformed DT and kNN on robustness
(Xiaoqi et al. [76])	CNN	UHF	PRPD parameters	Source classification	83% accuracy
Morette et al. [59]	SVM	HVDC PD signals	DWT	Pattern recognition	Robust under high-voltage noise
Ephraim et al. [63, 64]	Random Forest Regressor	RF sensors	Wavelet packet decomposition	PD source localisation	Low MAE; effective for localisation
Singh et al. [50]	Convolutional Neural Network	UHF PD signals	Raw waveform input	Fault type classification	Auto-feature extraction reduced preprocessing

$$\min \sum_{i=1}^N ||x_i - \mu_{C_i}||^2$$

Where x_i = data point μ_{C_i} = centroid of cluster C_i [77].

This formulation embodies the principle of variance minimisation within clusters, yielding compact and well-separated groups.

Clustering analysis is a technique for determining the distribution of patterns in data sets. In general, the purpose of clustering is to divide a data set into groups (clusters) so that data items within a cluster are more similar to one another than data elements from different clusters. Clustering, in the instance of separating PD signals, seeks to automatically assign each input signal to a cluster so that signals in the same cluster are more similar in shape to those in other clusters [78]. K- means clustering was used in the following literature for PD clustering and classification [79–81] Another unsupervised machine learning algorithm is the self-organising map. A self-organizing map (SOM) is a form of artificial neural network that is trained using unsupervised learning to generate a map that is a low-dimensional discretised representation of the training samples' input space. Nodes are typically distributed in a hexagonal or rectangular grid with consistent spacing. SOMs are automatically associated with the nodes of a regular, usually two-dimensional, grid in an orderly method, such that more comparable input data points are automatically connected with nodes that are adjacent in the grid, while less similar input data are situated farther apart in the grid. This organisation, a kind of model similarity diagram, allows for insight into the topographic relationships of data, particularly high-dimensional data items. If the data items are assigned to one of several predefined classes, the model can be trained to classify new input items based on these categories. When a new, unknown input is encountered, it is classified by determining which node in the model has the most similar vector representation, based on a similarity metric (e.g., Euclidean distance). The model assigns the input to the class corresponding to the closest node, effectively categorizing the input based on its similarity to the predefined class vectors. [82] applied SOM for anomaly detection of PD characteristics on wind

turbines using multidimensional data collected on SCADA. The results demonstrated that the method was highly effective.

Reinforcement learning & deep reinforcement learning (DRL)

Reinforcement learning (RL) involves training an agent to make sequential decisions in an environment, aiming to maximise a cumulative reward. The agent interacts with the environment, taking actions that lead to either rewards or penalties, which serve as feedback. Over time, the agent learns to optimise its behaviour by adjusting its actions to increase the total reward it accumulates. Reinforcement learning use exploration and exploitation strategies to balance the exploration of new actions to learn more about the environment and exploration of knowledge to maximize rewards.

$$\pi = \arg \max_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \right]$$

were,

s_t = state at time

a_t = action,

R= reward function,

$\gamma \in [0,1]$ = discount factor [83].

A popular algorithm that uses this form of training is the Q-learning reinforcement algorithm. The Q learning algorithm makes use of the optimal action value function, which is the maximum expected cumulative reward achievable from taking action a in state s under the optimal policy. These state-action pairs $Q(s, a)$ are then stored in a tabular form known as a Q-table. Conventional reinforcement learning algorithm usually have limitations when handling complex data with high-dimensional inputs due to its state action pair being stored in tabular forms. DRL, on the other hand, addresses the restrictions of conventional reinforcement learning by employing a neural network for function approximation, which obtains features from models, value functions, or rules and then seeks to generalise from them to construct an approximation of the full function. Another advantage of the DRL is that it can learn useful features from raw data automatically, eliminating the need for manual feature engineering. Traditional RL, on the other hand, frequently relies on handcrafted features, which can be time-consuming because they require manual selection, and tuning of relevant input features before training the RL agent. Deep neural networks can approximate complex, nonlinear functions, allowing DRL models to capture complex interactions between states, actions, and rewards. As a result, they can be suited to tasks in which the optimal policy is exceedingly complex, and they can operate in continuous action spaces. DRL in this paper is grouped into value-based reinforcement learning, policy gradient reinforcement learning, and actor-critic reinforcement learning. These classifications are discussed in further detail in subsequent sections. Reinforcement learning and DRL have been explored for various power system applications.

Value-based reinforcement learning

Value-based reinforcement learning (VBRL) is a reinforcement learning approach that focuses on evaluating the worth of various states or state-action combinations in an

environment. The core notion of value-based RL is to learn a value function that assists an agent in making decisions that maximise its cumulative benefit over time. This value function is also known as the state-value function (V) or the action-value function (Q). Popular examples of this method are the Q-learning and deep Q-learning. Q-learning is a model-free, off-policy reinforcement learning algorithm designed to solve issues in which an agent must make a series of decisions to maximise a cumulative reward. The agent interacts with its surroundings by performing actions and changing states. At each time step, it must determine whether to investigate a new action (exploration) or to exploit the action with the largest cumulative payoff depending on the current Q-table. This trade-off is often managed by an exploration-exploitation method such as epsilon-greedy, in which the agent explores with a specific probability (epsilon) and exploits the rest of the time. After taking an action and observing the resulting state and reward, the Q-value (expected cumulative reward) for the current state-action pair is updated in the Q-table. The agent continues to interact with the environment, and over time, the Q-values converge to their optimal values. These converged Q-values represent the best policy for the agent to follow to maximize cumulative rewards. Q-learning has been applied on multiple parts of the power systems ranging from energy demand response to market pricing. The purpose of Q-learning is to develop a Q-table that depicts the predicted cumulative rewards for actions taken in different states. Instead of a Q-table, Deep Q-network approximates the Q-function using a deep neural network (DNN). The neural network is fed the current state and outputs Q-values for each feasible action. Each action in the network has one output node, and the Q-values are the values associated with those actions. Backpropagation is used to train the DNN to minimise the loss between predicted and target Q-values.

Policy gradient reinforcement learning

Policy gradient is a reinforcement learning strategy in which the agent directly learns a parameterised policy. Policy gradient approaches, as opposed to Q-learning, seek to optimise policy parameters to maximise the expected cumulative reward. The key idea is to compute gradients of expected reward with respect to policy parameters and then adjust the policy to enhance the predicted reward. Examples of this learning strategy include Monte-Carlo reinforcement learning, trust region policy optimisation and proximal policy optimisation. The Monte Carlo reinforcement learning method calculates the value of states or state-action pairings by averaging the returns gained from sampling several episodes of interaction with the environment. Monte Carlo methods learn from direct experience rather than a model of the environment. The policy gradient methods Trust Region Policy Optimisation (TRPO) and Proximal Policy Optimisation (PPO) are two innovative reinforcement learning algorithms. Both approaches are intended to address the difficulties associated with developing stable and sample-efficient policies in reinforcement learning tasks. Although they have some similarities, they also have unique qualities and trade-offs. TRPO is an algorithm that aims to optimise policy in a stable and theoretically sound manner. It addresses some of the problems that previous policy gradient approaches had, such as large variance and susceptibility to policy changes. TRPO employs a trust region technique to keep policy modifications as small as possible. It guarantees that the new policy does not depart too far from the old policy, which aids in training stability [84]. More details of this algorithm can be seen in

reference [85]. Proximal policy optimisation (PPO) is another policy gradient method that tries to increase the stability and sample efficiency of reinforcement learning systems. It is well-known for its simplicity, ease of implementation, and strong empirical performance. PPO adds a clipped surrogate objective function that limits policy modifications. It promotes the new policy to be similar to the existing policy but does not necessitate solving a constrained optimisation problem like TRPO does. The step size of policy update is controlled by an adaptive learning rate mechanism in PPO which aids in the achievement of consistent training. More details about this algorithm can also be found in reference [85]. In summary, TRPO and PPO are both policy gradient methods meant to overcome the difficulties associated with training reinforcement learning agents.

Actor-critic methods

Policy-based (actor-only) approaches explicitly learn a policy that implicitly maximises the discounted cumulative reward; yet they are hampered by substantial variance in gradient estimations. Value-based (critic-only) approaches leverage the Bellman optimality relationship to generate policy from the learnt value function, resulting in decreased variance in predicted return estimations. The Bellman optimality equation describes the relationship between a state or state-action pair's value and the values of its successor states or successor state action pairs. The actor-critic methods combine the advantages of the actor-only and critic-only methods, which are discussed in reference [86, 87]. The actor-critic algorithms parameterise both policy and value functions and update them simultaneously during training, consequently, they fall under both value-based and policy-based categories. They frequently yield superior empirical results than value-based and policy-based strategies. A commonly explored actor critic algorithm is the Deterministic policy gradient (DPG). DPG is a reinforcement learning framework that focuses on learning deterministic policies in continuous action spaces, such as robotic control, autonomous vehicles, and reinforcement learning continuous control tasks. In contrast to stochastic policies, which generate probability distributions over actions, deterministic policies directly translate states to specified actions. DPG is particularly well-suited to situations requiring precise and continuous actions. The policy of DPG is deterministic, which means that it directly maps states to actions without the use of chance. A neural network with learnable parameters is commonly used to represent it. DPG frequently employs a critic to estimate the action-value function (Q-function). The Q-function calculates the expected return on a certain activity in a given state while adhering to the policy. The critic contributes to the evaluation of the quality of the policy's actions. To update the policy, DPG applies the deterministic policy gradient theorem. The expected return gradient with respect to policy parameters is computed and used to alter the policy in a direction that enhances the expected return. Gradients are commonly computed in DPG using action-value gradients (Q-gradients), which are the gradients of the Q-function with respect to the action [88]. Deep Deterministic Policy Gradient (DDPG) is a reinforcement learning system that extends Deterministic Policy Gradient (DPG) ideas to handle high-dimensional state spaces and continuous action spaces. DDPG was created primarily for situations requiring continuous and precise actions, such as robotic control and autonomous navigation. DDPG employs a deterministic policy that maps states to continuous actions without the introduction of randomness. A neural network

with learnable parameters represents the policy. DDPG uses an actor-critic architecture, which is made up of two neural networks: the actor and the critic. The deterministic policy is represented by the actor network. It accepts states as input and returns continuous action values as output. The critic network, on the other hand, estimates the action-value function (Q-function). It accepts as input pairs of states and actions and predicts the expected cumulative reward of performing an action in a given state while adhering to the actor's policy [88]. DDPG stores and samples previous experiences (state-action-reward-next state transitions) in an experience replay buffer. This replay buffer aids in breaking down correlations in the data and enables a stable learning process. To stabilise training, DDPG employs target networks for both the actor and the critic. These target networks are gradually updated using a soft target update technique, which helps keep the training process stable. DDPG uses gradient descent to update the actor and critic networks. The actor aims to maximise the expected return, while the critic minimises the Mean Squared Error (MSE) between its predicted Q-values and the actual values. MSE is a loss function commonly used in regression tasks to measure the average squared difference between predicted and true values, calculated by averaging the squared differences over all data points. DDPG has been used successfully in robotics and simulation environments for a range of continuous control tasks. It is well-known for its capacity to deal with high-dimensional state spaces and continuous action spaces, making it ideal for tasks requiring precision control. While DDPG performs well in many domains, it is important to note that it may require careful tweaking of hyperparameters for optimal results and can be sensitive to task-specific factors. Other actor-critic algorithms are soft actor critic (SAC), Advantage actor-critic(A2C), and Asynchronous Advantage actor-critic(A3C). SAC is an actor-critic approach, which means that it makes use of both a policy network (the actor) and a value function network (the critic). SAC implements a soft policy update, which means it attempts to maximise a trade-off between expected return and entropy, so encouraging the agent to explore effectively while learning. To manage the level of exploration, SAC employs a soft Bellman backup operator. The soft Bellman backup operator is a mathematical operation that updates a policy's value function. It includes a softness parameter that determines the balance of exploration and exploitation. This operator is widely utilised in algorithms such as SAC and DDPG. SAC works effectively in situations with high-dimensional action spaces and is well-suited for continuous action areas. The Advantage Actor-Critic (A2C) algorithm is a reinforcement learning system that combines policy-based and value-based features. It is an actor-critic algorithm, which means it learns a policy (actor) and a value function (critic) at the same time. A2C is an on-policy algorithm, which means that it collects data while adhering to the present policy and updates it based on the data. The critic (value network) is used to estimate the advantage of each action in each state along the trajectory. The advantage reflects how much better or worse an action is compared to the average action in that stage. It aids in discriminating between good and bad behaviour. The actor (policy network) is updated to maximise the expected return, which requires modifying the policy to favour activities that result in higher advantages. This is often accomplished using policy gradient approaches. The critic (value network) is updated to minimise the Mean Squared Error (MSE) between its anticipated and observed values during the journey [89]. This improves the accuracy of value assessments. When compared to pure policy gradient approaches, A2C has

lower variation, making it more stable and efficient. A2C strikes a balance between exploration through the actor and evaluation by using the critic, resulting in faster learning. Because it leverages advantages to direct policy changes, A2C are sample-efficient. A3C makes use of several concurrent agents, each interacting with its own copy of the environment and updating a shared global policy and value function. When compared to A2C, the parallelism in A3C can result in significantly faster training, particularly on multi-core CPUs or in distributed computing environments. This acceleration is especially beneficial when dealing with complex scenarios or utilising diverse computational resources. A3C's asynchronous nature fosters exploration, as multiple agents independently navigate the environment. This increases the likelihood of discovering innovative and potentially superior policies, which is particularly advantageous in environments with sparse rewards or intricate state spaces. Additionally, due to asynchronous updates from multiple agents, A3C tends to be more stable during training compared to A2C. This stability contributes to faster algorithm convergence. A3C's parallelism contributes to this reduction by averaging the experiences of numerous agents, resulting in more reliable updates [88]. A3C can scale to a greater number of parallel agents and processing resources than A2C. This makes A3C an excellent candidate for distributed training on clusters, where several CPUs or GPUs can be used. In reference (Oh et al. 2015) A3C was compared with other RL methodologies like DQN and Q-learning on Atari games and it performed better in terms of accuracy and faster convergence.

Application of reinforcement learning in the power grid

Reinforcement Learning (RL) has been used for a variety of power grid applications, including but not limited to energy management, demand response, control, fault detection & maintenance, market pricing, and voltage regulation. Some of these applications will be covered in the following sections below.

Energy management

Lu and Hong [90] investigated the benefits of using DRL on a smart grid to optimise schedules for energy management systems. The learning strategy was examined using two methods: Deep Q-learning and Deep Policy Gradient. Both algorithms recommended solutions effectively to reduce energy costs and improve net energy efficiency. To reduce energy costs, a variable power price signal is examined to encourage users to shift their consumption to low-price, off-peak periods. (Mbuwir et al. 2017.) addressed the difficulty of providing flexibility by operating storage devices and supply systems to fulfil household thermal and electrical demand in a combined heat and power system with photovoltaic (PV) installation. The problem was addressed as a sequential decision-making process under uncertainty, caused by unpredictable future electricity usage and weather-dependent PV generation. To solve this, the researchers employed fitted Q-iteration, a batch Reinforcement Learning (RL) technique. The study, which was conducted using data from the Belgian grid, demonstrated that RL can handle varied energy carriers in the system optimally even without a comprehensive Multi Carrier Energy System(MCES) model. The findings reveal that batteries efficiently provide flexibility by storing excess PV energy and discharging during low PV output periods, resulting in a significant reduction in power injection into the grid. Al-Saffar et al. [86] introduced a novel method for distributed Battery Energy Storage System (BESS) control to address

voltage rise issues in power distribution networks due to high residential photovoltaic (PV) penetration. This approach, known as Coordinated State of Charge Control (CSOCC), combines Monte Carlo Tree Search-based Reinforcement Learning (MCTS-RL) and Model Predictive Control (MPC). The technique optimises network-wide BESS capacity to avoid voltage violations and reduce power transfer losses. The study shows two scenarios: one without CSOCC, which requires a considerable increase in BESS penetration from 9% to 18% per impacted bus, and one with CSOCC, which mitigates voltage rise difficulties by using assisting buses with only 20% BESS participation. This coordination of scattered storage units produces an optimised aiding zone, lowering voltage concerns with minimal system losses. The model was tested on the distribution system of Lloydminster, Alberta, Canada. The results indicate that the distributed control approach effectively mitigates overvoltage issues while ensuring optimal system performance.

Demand response

Lu and Hong [90] presented a dynamic pricing demand response (DR) algorithm for smart grids using reinforcement learning (RL). This algorithm adjusts electricity prices in real-time to balance service provider (SP) profitability and customer (CU) costs, improving grid reliability and reducing energy expenses. The system model includes both SP and CU, where SP aims to maximise profit by setting optimal prices, and CUs respond by adjusting their energy consumption. The dynamic pricing problem is formulated as a Markov decision process (MDP) and solved using Q-learning, a model-free RL algorithm. This approach allows the SP to learn optimal pricing strategies through iterative updates based on rewards received from different pricing actions. Simulation results show that the algorithm effectively reduces overall energy consumption and enhances the balance between energy supply and demand. The optimal prices determined by the algorithm lead to significant energy reductions, benefiting both SP and CUs. The study concludes that the RL-based dynamic pricing algorithm is a viable solution for real-time demand management in smart grids. Babar et al. [87] formulated a methodology for agile demand response (DR) using mathematical micromodels. This approach includes complex demand bidding, which integrates price relationships with varying load demand and marginal costs and considers temporal and profit constraints. The paper proposes a decentralised Q-learning technique to plan flexible demand, offering a straightforward yet effective method for active distribution networks. Simulations involving several households equipped with the complex bidding mechanism and decentralised learning demonstrate the methodology's application. Agile demand response aims to enhance fair distribution of DR benefits among various stakeholders, including customers, retailers, and system operators. The proposed methodology involves multiple types of agents and learning objectives, exploring trade-offs in a day-ahead market context. It emphasises the importance of planning demand flexibility and the fair allocation of DR benefits to ensure optimal outcomes. In summary, Babar et al. presented a novel approach to demand response that combines complex bidding rules with reinforcement learning to improve market efficiency supporting the transition to more responsive and efficient electricity markets.

Control

Yin et. al, [91] proposed an innovative approach to enhance Automatic Generation Control (AGC) in large-scale power systems. The authors introduced Deep Forest Reinforcement Learning (DFRL) as a solution. DFRL integrates deep forest and subsidiary RL to predict systemic states (emergency or normal) and adapt control strategies accordingly. This method allows for real-time updates and mitigates dimensionality issues by splitting RL parameters for different states, reducing memory requirements. DFRL was tested on various power systems, including a two-area load frequency control system, a three-area system, and the China Southern Power Grid. It outperformed ten conventional AGC algorithms, achieving superior control performance and reducing emergency situations. In conclusion, DFRL offers an effective preventive strategy for AGC, enhancing system stability and performance by predicting and adapting to different operational states while efficiently handling large disturbances. [74] proposed a DRL algorithm to address the Volt-Var Control (VVC) problem in power distribution systems, which typically requires detailed and accurate network topology and parameter information. The VVC problem is an important optimisation task in electric power distribution systems. the goal is to manage voltage levels and reactive power to increase system efficiency, dependability, and stability. By formulating the VVC problem as a Constrained Markov Decision Process (CMDP), the authors developed an off-policy DRL algorithm called CSAC (Constrained Soft Actor-Critic). This algorithm integrates a device-decoupled policy network with an ordinal encoding scheme, which enhances its scalability and efficiency. The proposed CSAC algorithm is validated through numerical studies on IEEE 4-bus¹, 34-bus², and 123-bus³ distribution test feeders. The IEEE buses are described in the appendix at the end of the paper. The results demonstrate that CSAC outperforms existing state-of-the-art reinforcement learning algorithms and traditional optimisation-based methods in terms of sample efficiency, scalability, and constraint satisfaction. Specifically, CSAC achieves higher returns without violating voltage constraints, indicating its effectiveness in practical applications. The research highlights the potential of DRL algorithms in solving complex control problems in power systems without relying on precise network models. Gao et al. [92] proposed a novel approach to optimize the reconfiguration of power distribution networks using a Batch-Constrained Soft Actor-Critic (BCSAC) algorithm. The authors leverage a Conditional Variational Autoencoder (CVAE) to model the behaviour policy from historical operational data, addressing challenges posed by high-dimensional state-action spaces and mixed variable types. The BCSAC algorithm integrates this generative model to guide the learning process, ensuring policy updates remain within the feasible region defined by the historical data. The proposed method diverges from traditional actor-critic frameworks by utilising a batch-constrained soft policy iteration and finite action space policy gradient for updating the actor network. This approach aims to improve the efficiency and stability of learning policies for dynamic distribution network reconfiguration (DNR). The algorithm showed significant improvements in operational performance compared to existing methods. The study concludes that BCSAC, combined with CVAE, offers a robust

¹ A simple, small power system model with 4 buses (nodes) and typically used for basic power flow and stability analysis.

² A medium-sized test case with 34 buses, designed to represent a more realistic power distribution network.

³ A larger, more complex test system consisting of 132 buses, often used for detailed studies of transmission networks.

solution for enhancing the reliability and efficiency of power distribution systems through dynamic reconfiguration. Ruisheng Diao [93] proposed the Grid Mind framework, a novel approach leveraging DRL for autonomous voltage control in modern power grids. Traditional power grids, with their ageing infrastructure, face significant challenges due to the increasing penetration of renewable energy sources and the dynamic nature of electricity demand. Conventional control methods, which rely on offline simulations and predefined strategies, often fail to adapt to real-time changes, leading to inefficiencies in voltage regulation, increased power losses, and decreased system reliability. Grid Mind addresses these issues by training an AI agent through extensive offline simulations and historical data. This agent learns optimal control policies to Manage voltage levels effectively in real time, even under varying load conditions and grid topologies. Some of the distribution topologies have been discussed in section 2. The framework utilises a Deep Q Network (DQN) to handle the continuous state space of power grids, ensuring robust performance through continuous learning and adaptation.. Case studies on the IEEE 14-bus system demonstrate the efficacy of the Grid Mind framework. Key findings include, The DRL agent successfully Managed voltage control under random load changes, showing significant improvement in reward values over 10,000 episodes, indicating enhanced performance and fewer control actions required as learning progressed in cases without contingencies, the agent efficiently adapted maintaining robust performance under emergency conditions, modifying the reward structure to prioritize single-action solutions led to high reward values, confirming the agent's superior performance in resolving voltage issues efficiently. These results validate the effectiveness of the DRL-based Grid Mind framework in real-time voltage control, demonstrating its potential for improving grid reliability and efficiency. Huang et al. [94] presented a novel approach to enhance grid stability and reliability through DRL. The study focuses on two emergency control schemes: generator dynamic braking and under-voltage load shedding (UVLS). For generator dynamic braking, the DRL agent is trained and tested on a two-area, four-machine system. The agent, using a deep Q-network (DQN), effectively achieved robust performance in generator dynamic braking simulations. The results demonstrate that the DRL-based control outperforms Traditional methods by operating the braking more efficiently, leading to better system stability and higher rewards. In the UVLS scheme, the DRL agent is tested on a modified IEEE 39-bus system to handle Fault-Induced Delayed Voltage Recovery (FIDVR) issues. The agent successfully learns a load-shedding policy that meets voltage recovery requirements while minimizing the load shedding amount. The Training process, which involved 1.2 million interaction steps, resulted in a robust and adaptive control policy tested across 960 scenarios with varying conditions. The DRL-based UVLS scheme showed superior performance compared to traditional relay-based UVLS and Model Predictive Control (MPC) methods in terms of execution time and system rewards. Extensive case studies in both the two-area, four-machine system and the IEEE 39-bus system have demonstrated the excellent performance and robustness of the proposed schemes. The reward difference between the DRL and the MPC shows that DRL-based control has slightly better performance than the MPC in 57.22% of test scenarios). Overall, the DRL approach demonstrated effectiveness, robustness, and adaptability in emergency power system control, indicating its potential for real-world application in enhancing grid resilience. Lee et al [95] proposed the use of a Graph Convolutional Proximal Policy Optimization

(GC-PPO) method to enhance frequency control within the UK power grid. This algorithm was designed to optimize grid frequency stability by leveraging spatial and temporal dependencies inherent in the grid structure. Additionally, the study extended the application of RL techniques to improve renewable energy investment planning and dynamically balance electricity networks under varying demand and supply conditions. The proposed methods demonstrated significant advancements, including improved frequency stabilization, cost-effective carbon reduction through strategic energy investments, and enhanced demand-supply equilibrium compared to conventional approaches. These findings highlight the transformative potential of reinforcement learning in modern power grid management, particularly as the grid adapts to increasing renewable energy integration.

Fault detection

Rocchetta et al. [96] investigated the use of DRL for emergency control in power systems to enhance resilience and stability. The study addresses the challenges of traditional power system control methods, which often struggle with real-time adaptation to dynamic conditions. The authors proposed a DRL-based framework that leverages neural networks to learn optimal control policies from simulated grid environments. The developed DRL algorithm, demonstrates its effectiveness in managing grid emergencies by balancing between preventive and corrective actions. The results show that the DRL approach can significantly improve the system's ability to withstand and recover from disturbances, compared to conventional methods. Specifically, the algorithm adapts to varying load demands and generator degradation states, optimising the expected rewards through efficient decision-making processes. Comparative analysis with the Bellman's optimality and other synthetic policies reveals that the DRL policy closely approximates the optimal policy while offering practical advantages in terms of computational feasibility and adaptability. The DRL policy also outperforms expert-based solutions by reducing the frequency of costly corrective maintenance actions and improving overall grid reliability. The convergence of the DRL algorithm to near-optimal solutions highlight its potential as a robust tool for adaptive power system control.

Market pricing

Chen and Su [97] addressed the modelling of energy trading behaviour within a local energy market using DRL. The research proposes a holistic market model that enables domestic consumers to trade energy directly, rather than exclusively through utility companies, thereby eliminating the price gap associated with traditional trading methods. The study models the decision-making process of consumers as a Markov Decision Process (MDP) with continuous variables, leveraging deep Q-learning (DQL) algorithms to facilitate efficient energy trading without requiring comprehensive market knowledge or analytical calculations. The DRL technique, particularly with experience replay, is highlighted for its ability to adapt to new data and improve trading strategies over time. The results, evaluated through simulations, demonstrate that the proposed algorithm significantly enhances the efficiency of local energy trading. Consumers using DQL achieved higher economic benefits compared to those using rule-based or random strategies. The DRL approach effectively managed market prices, battery storage, and the stochastic nature of renewable generation and demand, resulting in optimal trading decisions and

improved economic returns. Rahimiyan and Rajabi Mashhadi [98] explored the application of Q-learning (QL) to determine optimal bidding strategies for suppliers in electricity markets. The study addresses the complexity of bidding in pay-as-bid (PAB) auctions. PAB is a pricing mechanism in which participants pay the exact amount they bid for goods or services, rather than a single clearing price determined by the highest approved offer (as in uniform pricing systems). The research solved this by modelling the problem as a self-play scenario where suppliers learn to adapt their strategies based on the observed actions of their competitors and market conditions. The research presents a detailed QL-based approach for solving the bidding problem, demonstrating that the algorithm converges to the optimal solution provided by the QL-based approach. The study uses simulations to replicate a realistic power market environment, showing that suppliers can effectively adjust their bidding strategies in response to rival behaviour and other influential factors. The study concludes that QL provides a robust method for suppliers to optimise their bidding strategies in electricity PAB auctions, demonstrating significant potential for application in real-world power markets. The results demonstrate that the QL algorithm can successfully learn and adapt optimal bidding strategies, achieving performance comparable to traditional model-based methods. Notably, suppliers using QL initially make suboptimal decisions but quickly improve as they accumulate experience. The study finds that the QL-based method effectively adjusts to market dynamics and competitor actions, offering a robust solution for determining optimal bid prices under uncertainty.

RL and Bayesian networks for asset prioritization

Reinforcement Learning (RL) and Bayesian Networks (BNs) offer powerful, complementary tools for risk-informed asset prioritisation in power grid maintenance. RL enables agents to continually adapt maintenance strategies based on evolving operational data, such as asset health metrics, environmental conditions, and cost considerations. This approach has been demonstrated in frameworks developed by [96], as well as in the multi-objective deep RL model (MO-DCMAC) proposed by [99] aimed at optimizing infrastructure maintenance. In parallel, BNs provide a structured probabilistic framework to model dependencies between key indicators (e.g., temperature, load, vibration, and partial discharge activity). As new data becomes available, BNs can update failure probabilities dynamically, contributing to informed fault diagnostics—an approach showcased in the Dempster–Shafer–enhanced model by [100].

Conclusion

According to the research, more than 70% of power grid problems occur on the Transmission and distribution side of the electrical power grid. As a result, future research will focus on defects and fault detection in the Transmission and distribution networks. The Literature also identifies that insulation degradation accounts for more than 60% of high-voltage problems on the power grid. Due to this it can be helpful to develop diagnostics for high voltage problems using insulation degradation. [6, 24]. The study will focus on partial discharge (PD) data, specifically high-frequency signals captured using sensors such as High-Frequency Current Transformers (HFCTs), as primary indicators of insulation degradation and fault initiation in high-voltage components. From the literature, all the algorithms used to explore PD either predicted the type of PD or have been able to

detect between faulty and non-faulty signals. Supervised learning methods, reviewed in Section "[Supervised learning](#)", are highly effective in classifying known fault types but rely heavily on labelled data and could potentially struggle with novel fault scenarios. Unsupervised methods, as discussed in Section "[Unsupervised learning](#)", offer advantages in detecting anomalies without labelled data, which is valuable in complex grid environments, though they lack specificity in pinpointing precise fault types. Reinforcement learning (RL) approaches, and particularly Asynchronous Advantage Actor-Critic (A3C), discussed in Section "[Application of reinforcement learning in the power grid](#)", have demonstrated superior adaptability and performance improvements in complex environments in comparison to other supervised and unsupervised algorithms, many of the studies indicated a 10-20% performance improvement with the usage of RL [101]. Unlike supervised and unsupervised approaches, RL, including A3C, adapts through iterative interaction with the environment, providing robust solutions for real-time fault detection and management. In most investigations, other deep learning approaches and actor-critic methods have not been explored as much as Q-learning, deep Q-learning, and their variations. Also, majority of the studies tackle operational issues and energy management. Future evaluations will benchmark the proposed A3C-based models against traditional classifiers such as Support Vector Machines (SVM), k-Nearest Neighbours (KNN), and Deep Q-Learning (DQL) frameworks. Performance will be assessed using standard metrics including classification accuracy, precision, and statistical significance via McNemar's test. Future research will investigate multi-agent implementations of the Asynchronous Advantage Actor-Critic (A3C) framework to support distributed coordination among substations and Distributed Energy Resource (DER) nodes. This decentralised approach is expected to enhance responsiveness, scalability, and fault resilience across large-scale transmission networks. The study will maintain a focus on partial discharge (PD) as a primary diagnostic indicator for insulation degradation and fault initiation in high-voltage lines, leveraging PD's early detectability for predictive maintenance. Building on A3C's capability for hierarchical representation learning, subsequent models will be designed to detect and classify anomalies across varying levels of degradation severity. Such stratification is critical for transitioning from binary fault detection to nuanced health state estimation. The integration of reinforcement learning (RL) into power grid diagnostics presents a transformative opportunity to optimise maintenance planning—enabling adaptive decision-making that reduces downtime, enhances reliability, and reduce operational cost.

Acknowledgement

The authors would like to express their heartfelt appreciation to Cranfield University's Integrated Vehicle Health Management Centre.

Author contributions

D.O. conceptualized the study and wrote the main manuscript text. S.K. and I.J. supervised the research and reviewed the manuscript. All authors approved the final version of the manuscript.

Funding

The research conducted for this study did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability

No datasets were generated or analysed during the current study.

Declarations**Ethics approval and consent to participate**

Not applicable. This study did not involve human participants, personal data, or interventions requiring ethical approval.

Consent for publication

No restrictions apply on publication in this research.

Competing interests

The authors declare no competing interests.

Received: 12 June 2025 / Accepted: 30 August 2025

Published online: 26 September 2025

References

1. Pagani GA, Aiello M (2013) The power grid as a complex network: a survey. *Physica A Stat Mech Appl* 392(11):2688–2700. <https://doi.org/10.1016/j.physa.2013.01.023>
2. London Economics (2013) The Value of Lost Load (VoLL) for Electricity in Great Britain Final Report for OFGEM and DECC Acknowledgements. www.londecon.co.uk
3. Illias HA, Tunio MA, Bakar AHA, Mokhlis H, Chen G (2016) Partial discharge phenomena within an artificial void in cable insulation geometry: experimental validation and simulation. *IEEE Trans Dielectr Electr Insul* 23(1):451–459. <https://doi.org/10.1109/TDEI.2015.005155>
4. Callender G, Lewin PL (2020) Modeling_partial_discharge_phenomena.
5. Arief YZ, Izzati WA, Adzis Z (2012) Modeling of partial discharge mechanisms in solid dielectric material. *Int J Eng Innov Technol (IJET)* 1(4):315–320
6. Hussain MR, Refaat SS, Abu-Rub H (2021) Overview and partial discharge analysis of power transformers: a literature review. *IEEE Access* 9:64587–64605
7. National Grid (2024) What Do National Grid's Helicopters Do?
8. Blaabjerg F, Yang Y, Yang D, Wang X (2017) Distributed power-generation systems and protection. *Proc IEEE* 105(7):1311–1331
9. Fang Xi, Misra S, Xue G, Yang D (2012) Smart grid - the new and improved power grid: a survey. *IEEE Commun Surv Tutor* 14(4):944–980. <https://doi.org/10.1109/SURV.2011.101911.00087>
10. Ofgem (2006) 13537-Elecgenfactsfs
11. Ellabban O, Abu-Rub H, Blaabjerg F (2014) Renewable energy resources: current status, future prospects and their enabling technology. *Renew Sustain Energy Rev* 39:748–764. <https://doi.org/10.1016/j.rser.2014.07.113>
12. Beis (2021) UK energy in brief (2021) www.gov.uk/government/statistics/uk-energy-in-brief-2021.
13. HM Government (2009) The UK Low Carbon Transition Plan.
14. Barker PP, De Mello RW (2000) Determining the Impact of Distributed Generation on Power Systems: Part 1 - Radial Distribution Systems. In *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*, Vol. 3 Institute of Electrical and Electronics Engineers Inc, Pp. 1645–56
15. National grid (2013) 13784-High Voltage Direct Current Electricity – Technical Information
16. Qiao K, Zhu A, Wang B, Di C, Junwei Yu, Zhu Bo (2020) Characteristics of heat resistant aluminum alloy composite core conductor used in overhead power transmission lines. *Materials*. <https://doi.org/10.3390/MA13071592>
17. Saradarzadeh M, Farhangi S, Schanen JL, Jeannin PO, Frey D (2010) The Benefits of Looping a Radial Distribution System with a Power Flow Controller. *PECon2010 - 2010 IEEE International Conference on Power and Energy*, 723–28 <https://doi.org/10.1109/PECON.2010.5697674>.
18. Behera RR, Thakur AN (2016) An Overview of Various Grid Synchronization Techniques for Single-Phase Grid Integration of Renewable Distributed Power Generation Systems. In *International Conference on Electrical, Electronics, and Optimization Techniques, ICEEOT 2016*. Institute of Electrical and Electronics Engineers Inc, Pp. 2876–80
19. Timbus A, Liserre M, Teodorescu R, Blaabjerg F (2005) Synchronization Methods for Three Phase Distributed Power Generation Systems. An Overview and Evaluation. In *PESC Record - IEEE Annual Power Electronics Specialists Conference*. Vol. 2005. Pp. 2474–81
20. Asman SH, Ab Aziz NF, Ungku Amirulddin UA, Ab Kadir MZA (2021) Transient fault detection and location in power distribution network: a review of current practices and challenges in Malaysia. *Energies*. <https://doi.org/10.3390/en14112988>
21. Mar A, Pereira P, Martins JF (2019) A survey on power grid faults and their origins: a contribution to improving power grid resilience. *Energies*. <https://doi.org/10.3390/en12244667>
22. Taft JD (2017) Fault intelligence: distribution grid fault detection and classification. *Pac Northwest Natl Lab* 1:1–38
23. Santos WC, Lopes FV, Brito NSD, Souza BA (2017) High-impedance fault identification on distribution networks. *IEEE Trans Power Delivery* 32(1):23–32. <https://doi.org/10.1109/TPWRD.2016.2548942>
24. Hao L, Lewin PL (2010) Partial discharge source discrimination using a support vector machine. *IEEE Trans Dielectr Electr Insul*. <https://doi.org/10.1109/TDEI.2010.5412017>
25. Ma H, Chan JC, Saha TK, Ekanayake C (2013) Pattern Recognition Techniques and Their Applications for Automatic Classification Pattern Recognition Techniques and Their Applications for Automatic Classification of Artificial Partial Discharge Sources. *IEEE Trans Dielectr Electr Insulation* 20(2):468–478
26. Troeger A, Riechert U, Burrow S, Tenbohlen S *Conference Proceedings of CMD2010*.
27. Li S, Li J (2017) Condition monitoring and diagnosis of power equipment: review and prospective. *High Voltage* 2(2):82–91
28. Coenen S, Tenbohlen S, Markalous SM, Strehl T (2008) Sensitivity of UHF PD measurements in power transformers. *IEEE Trans Dielectr Electr Insul*. <https://doi.org/10.1109/TDEI.2008.4712657>
29. Rahman Abd MS, Lewin PL, Rapisarda P (2016) Autonomous localization of partial discharge sources within large transformer windings. *IEEE Trans Dielectr Electr Insul* 23(2):1088–1098
30. Seo J, Ma H, Saha TK (2017) A joint vibration and arcing measurement system for online condition monitoring of on-load tap changer of the power transformer. *IEEE Trans Power Delivery* 32(2):1031–1038. <https://doi.org/10.1109/TPWRD.2016.2531186>
31. Kweon DJ, Chin SB, Kwak HR, Kim JC, Song KB (2005) The analysis of ultrasonic signals by partial discharge and noise from the transformer. *IEEE Trans Power Deliv* 20(3):1976–1983. <https://doi.org/10.1109/TPWRD.2004.833923>

32. Qian S, Chen H, Xu Y, Su L (2018) High sensitivity detection of partial discharge acoustic emission within power transformer by Sagnac fiber optic sensor. *IEEE Trans Dielectr Electr Insul* 25(6):2313–2320. <https://doi.org/10.1109/TDEI.2018.007131>
33. Nilsson J, Bertling L (2007) Maintenance management of wind power systems using condition monitoring systems. *IEEE Pow Eng Soc Gen Meet* 22(1):223–229. <https://doi.org/10.1109/PES.2007.385616>
34. Endrenyi J, Anders GJ, Leite da Silva AM (1998) Probabilistic Evaluation of the Effect of Maintenance on Reliability - an Application. *IEEE Trans Power Syst* 13(2):576–583
35. Endrenyi J, Anders GJ, Leite Da Silva AM (1998) Probabilistic Evaluation of the Effect of Maintenance on Reliability - an Application. *IEEE Trans Power Syst* 13(2):576–583. <https://doi.org/10.1109/59.667385>
36. Sunday AA, Omolayo MI, Okwilagwe O, Moses ME, Bernard AA (2021) Evaluation of effective maintenance and reliability operation management-a review. <https://doi.org/10.1051/e3sconf/202130>
37. Mobley RK (2002) An Introduction to Predictive Maintenance. Butterworth-Heinemann
38. Gertsbakh I, Gertsbakh IB (2000) Reliability Theory: With Applications to Preventive Maintenance. Springer Science
39. Schneider J, Gaul AJ, Neumann C, Höggräfer J, Wellßow W, Schwan M, Schnettler A (2006) Asset management techniques. *Int J Electr Power Energy Syst* 28(9):643–654. <https://doi.org/10.1016/j.jepes.2006.03.007>
40. Hilber P, Lindquist T (2003) The Maintenance Process. Sweden, Swedish
41. Carnero MC, Gómez A (2017) Maintenance strategy selection in electric power distribution systems. *Energy* 129:255–272. <https://doi.org/10.1016/j.energy.2017.04.100>
42. Yssaad B, Khiaf M, Chaker A (2014) Reliability centered maintenance optimization for power distribution systems. *Int J Electr Power Energy Syst* 55:108–115. <https://doi.org/10.1016/j.jepes.2013.08.025>
43. Bertling L, Allan R, Eriksson R (2005) A reliability-centred asset maintenance method for assessing the impact of maintenance in power distribution systems. *IEEE Pow Eng Soc Gen Meet*. <https://doi.org/10.1109/pes.2005.1489068>
44. Batty M (2018) Digital twins. *Environ Plann B: Urb Anal City Sci* 45(5):817–820
45. Harries T, Hartnoll M, Hafezianrazavi M, Meek H, Nassehi A (2023) Digital twins for predictive maintenance. *Procedia CIRP*. 118:306–311
46. Mansour DEA, Numair M, Zalhaf AS, Ramadan R, Darwish MMF, Huang Qi, Hussien MG, Abdel-Rahim O (2023) Applications of IoT and digital twin in electrical power systems: a comprehensive survey. *IET Gener Transm Distrib* 17(20):4457–4479
47. Jennions IK (2011) Integrated Vehicle Health Management: Perspectives on an Emerging Field. SAE International, Warrendale, PA
48. Gooi HB, Wang T, Tang Y (2023) Edge intelligence for smart grid: a survey on application potentials. *CSEE J Pow Energy Syst* 9(5):1623–1640. <https://doi.org/10.17775/CSEEJPES.2022.02210>
49. Sirojan T, Shibo Lu, Phung BT, Zhang D, Ambikairajah E (2022) Sustainable deep learning at grid edge for real-time high impedance fault detection. *IEEE Trans Sustain Comput* 7(2):346–357. <https://doi.org/10.1109/TSUSC.2018.2879960>
50. Singh T, Kumar S, Singh SK, Chilkoti V, Bansal S, Chui KT, Gupta BB (2024) Intelligent FaultEdge: AI-Driven Fault-Tolerant Edge Framework for Smart Grid Monitoring in IoT. In *International IEEE Conference proceedings, IS. Institute of Electrical and Electronics Engineers Inc.*
51. Alajlan NN, Ibrahim DM (2022) Tinyml: enabling of inference deep learning models on ultra-low-power IoT edge devices for AI applications. *Micromachines*. <https://doi.org/10.3390/mi13060851>
52. Arcas GI, Cioara T, Anghel I, Lazea D, Hangan A (2024) Edge offloading in smart grid. *Smart Cities* 7(1):680–711
53. Hossain MM, Fotouhi M, Hasan R (2015) Towards an Analysis of Security Issues, Challenges, and Open Problems in the Internet of Things. In *Proceedings - 2015 IEEE World Congress on Services, SERVICES 2015, Institute of Electrical and Electronics Engineers Inc*, pp. 21–28
54. Yan Ye, Yi Qian, Sharif H, Tipper D (2012) A survey on cyber security for smart grid communications. *IEEE Commun Surv Tutor* 14(4):998–1010
55. Mahesh B (2018) Machine learning algorithms-a review. *Int J Sci Res*. <https://doi.org/10.21275/ART20203995>
56. Nasteski V (2017) An overview of the supervised machine learning methods. *Horizons. B* 4:51–62. <https://doi.org/10.2054/horizons.b.04.1.17.p05>
57. Shalev-Shwartz S, Ben-David S (2012) Understanding Machine Learning: From Theory to Algorithms. Cambridge University Press
58. Tang Ju, Jin M, Zeng F, Zhou S, Zhang X, Yang Yi, Ma Y (2017) Feature selection for partial discharge severity assessment in Gas-Insulated Switchgear based on minimum redundancy and maximum relevance. *Energies*. <https://doi.org/10.3390/en10101516>
59. Morette N, Ditchi T, Oussar Y (2020) Feature extraction and ageing state recognition using partial discharges in cables under HVDC. *Electr Power Syst Res*. <https://doi.org/10.1016/j.epsr.2019.106053>
60. Kunicki M, Wotzka D (2019) A classification method for select defects in power transformers based on the acoustic signals. *Sensors (Switzerland)*. <https://doi.org/10.3390/s19235212>
61. Patle A, Chouhan DS (2013) SVM Kernel Functions for Classification.
62. Jiang H, Zhang JJ, Gao W, Wu Z (2014) Fault detection, identification, and location in smart grid based on data-driven computational methods. *IEEE Trans Smart Grid* 5(6):2947–2956. <https://doi.org/10.1109/TSG.2014.2330624>
63. Iorkyase ET, Tachtatzis C, Lazaridis P, Glover IA, Atkinson RC (2019) Low-complexity wireless sensor system for partial discharge localisation. *IET Wirel Sensor Syst* 9(3):158–165. <https://doi.org/10.1049/iet-wss.2018.5075>
64. Iorkyase ET, Tachtatzis C, Glover IA, Lazaridis P, Upton D, Saeed B, Atkinson RC (2019) Improving RF-based partial discharge localization via machine learning ensemble method. *IEEE Trans Power Deliv* 34(4):1478–1489. <https://doi.org/10.1109/TPWRD.2019.2907154>
65. Janani H, Kordi B (2018) Towards automated statistical partial discharge source classification using pattern recognition techniques. *High Voltage* 3(3):162–169. <https://doi.org/10.1049/hve.2018.5048>
66. Zang Y, Qian Y, Liu W, Yongpeng Xu, Sheng G, Jiang X (2019) A novel partial discharge detection method based on the photoelectric fusion pattern in GIL. *Energies*. <https://doi.org/10.3390/en12214120>
67. Zhang S, Li C, Wang Ke, Li J, Liao R, Zhou T, Zhang Y (2016) Improving recognition accuracy of partial discharge patterns by image-oriented feature extraction and selection technique. *IEEE Trans Dielectr Electr Insul* 23(2):1076–1087. <https://doi.org/10.1109/TDEI.2015.005226>

68. Tra V, Duong BP, Kim JM (2019) Improving diagnostic performance of a power transformer using an adaptive over-sampling method for imbalanced data. *IEEE Trans Dielectr Electr Insul* 26(4):1325–1333. <https://doi.org/10.1109/TDEI.2019.008034>
69. Mitchell TM (1997) *Machine Learning*. McGraw-Hill
70. Woon WL, El-Hag A, Harbaji M (2016) Machine learning techniques for robust classification of partial discharges in oil-paper insulation systems. *IET Sci Meas Technol* 10(3):221–227. <https://doi.org/10.1049/iet-smt.2015.0076>
71. Hussein R, Shaban KB, El-Hag AH (2017) Robust feature extraction and classification of acoustic partial discharge signals corrupted with noise. *IEEE Trans Instrum Meas* 66(3):405–413. <https://doi.org/10.1109/TIM.2016.2639678>
72. Soltani AA, El-Hag A (2019) Denoising of radio frequency partial discharge signals using artificial neural network. *Energies*. <https://doi.org/10.3390/en12183485>
73. Luo G, Zhang D, Tseng KJ, He J (2016) Impulsive noise reduction for transient Earth voltage-based partial discharge using wavelet-entropy. *IET Sci Meas Technol* 10(1):69–76. <https://doi.org/10.1049/iet-smt.2014.0203>
74. Wang W, Nanpeng Yu, Gao Y, Shi J (2020) Safe off-policy deep reinforcement learning algorithm for Volt-Var control in power distribution systems. *IEEE Trans Smart Grid* 11(4):3008–3018. <https://doi.org/10.1109/TSG.2019.2962625>
75. Wang Y, Zhang X, Li Y, Li L, Gao J, Guo N (2020) Multi-scale analysis and pattern recognition of ultrasonic signals of PD in a liquid/solid composite of an oil-filled terminal. *Energies*. <https://doi.org/10.3390/en13020366>
76. Xiaoqi W, Hui S, Lingling L, Zhe L, Gegho S, Xiuchen J (2018) 2018 Condition Monitoring and Diagnosis (CMD). IEEE.
77. Murphy K (2012) *Machine Learning: A Probabilistic Perspective*. MIT press
78. IEEE Staff, and IEEE Staff. n.d (2012) IEEE Power and Energy Society General Meeting.
79. Chatpattananan V, Pattanadech N, Yutthagowith P (2006) Partial Discharge Classification on High Voltage Equipment with K-Means.
80. Lin YH (2011) Using K-means clustering and parameter weighting for partial-discharge noise suppression. *IEEE Trans Power Deliv* 26(4):2380–2390. <https://doi.org/10.1109/TPWRD.2011.2162858>
81. Peng X, Zhou C, Hepburn DM, Judd MD, Siew WH (2013) Application of K-means method to pattern recognition in on-line cable partial discharge monitoring. *IEEE Trans Dielectr Electr Insul* 20(3):754–761
82. Du M, Ma S, He Q. A (2016) A SCADA Data Based Anomaly Detection Method for Wind Turbines. In *China International Conference on Electricity Distribution, CIGED*. Vols. 2016, IEEE Computer Society.
83. Schulman J, Levine S, Abbeel P, Jordan M, Moritz P (2017) Trust Region Policy Optimization. <http://arxiv.org/abs/1502.05477>
84. Shani L, Efroni Y, Mannor S Adaptive Trust Region Policy Optimization: Global Convergence and Faster Rates for Regularized MDPs. www.aai.org.
85. Li Y (2017) Deep Reinforcement Learning: An Overview. <http://arxiv.org/abs/1701.07274>.
86. Al-Saffar M, Al-Saffar M, Musilek P (2020) Reinforcement learning-based distributed BESS management for mitigating overvoltage issues in systems with high PV penetration. *IEEE Trans Smart Grid* 11(4):2980–2994. <https://doi.org/10.1109/TSG.2020.2972208>
87. Babar M, Nguyen PH, Cuk V, Kamphuis IG, Bongaerts M, Hanzelka Z (2018) The evaluation of agile demand response: an applied methodology. *IEEE Trans Smart Grid* 9(6):6118–6127. <https://doi.org/10.1109/TSG.2017.2703643>
88. Zhang Z, Zhang D, Qiu RC (2020) Deep reinforcement learning for power system applications: an overview. *CSEE J Power Energy Syst* 6(1):213–225
89. Yang T, Zhao L, Li W, Zomaya AY (2020) Reinforcement learning in sustainable energy and electric systems: a survey. *Annu Rev Control* 49:145–163
90. Lu R, Hong SH (2019) Incentive-based demand response for smart grid with reinforcement learning and deep neural network. *Appl Energy* 236:937–949. <https://doi.org/10.1016/j.apenergy.2018.12.061>
91. Yin L, Zhao L, Tao Yu, Zhang X (2018) Deep Forest Reinforcement Learning for Preventive Strategy Considering Automatic Generation Control in Large-Scale Interconnected Power Systems. *Appl Sci (Switzerland)*. <https://doi.org/10.3390/app8112185>
92. Gao Y, Wang W, Shi J, Yu N (2020) Batch-constrained reinforcement learning for dynamic distribution network reconfiguration. *IEEE Trans Smart Grid* 11(6):5357–5369. <https://doi.org/10.1109/TSG.2020.3005270>
93. Diao R, Wang Z, Shi D, Chang Q, Duan J, Zhang X (2019) 2019 IEEE Power & Energy Society General Meeting (PESGM), IEEE.
94. Huang Q, Huang R, Hao W, Tan J, Fan R, Huang Z (2020) Adaptive power system emergency control using deep reinforcement learning. *IEEE Trans Smart Grid* 11(2):1171–1182. <https://doi.org/10.1109/TSG.2019.2933191>
95. Lee Y, Choi H, Pagnier L, Kim CH, Lee J, Jhun B, Kim H, Kurths J, Kahng B (2024) Reinforcement learning optimizes power dispatch in decentralized power grid. *Chaos Solitons Fractals*. <https://doi.org/10.1016/j.chaos.2024.115293>
96. Rocchetta R, Bellani L, Compare M, Zio E, Patelli E (2019) A reinforcement learning framework for optimal operation and maintenance of power grids. *Appl Energy* 241:291–301. <https://doi.org/10.1016/j.apenergy.2019.03.027>
97. Chen T, Su W (2018) Local energy trading behavior modeling with deep reinforcement learning. *IEEE Access* 6:62806–62814. <https://doi.org/10.1109/ACCESS.2018.2876652>
98. Rahimiyan M, Mashhadi HR (2008) Supplier's optimal bidding strategy in electricity pay-as-bid auction: comparison of the Q-learning and a model-based approach. *Electr Power Syst Res* 78(1):165–175. <https://doi.org/10.1016/j.epsr.2007.01.009>
99. van Remmerden J, Kenter M, Roijers DM, Andriotis C, Zhang Y, Bukhsh Z (2025) Deep Multi-Objective Reinforcement Learning for Utility-Based Infrastructural Maintenance Optimization. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-024-10954-0>
100. Wu X, Zhao H, Wentao Xu, Pan W, Ji Q, Hua X (2024) Fault diagnosis of the distribution network based on the D-S evidence theory Bayesian network. *Front Energy Res*. <https://doi.org/10.3389/fenrg.2024.1422639>
101. Perera ATD, Kamalaruban P (2021) Applications of reinforcement learning in energy systems. *Renew Sustain Energy Rev*. <https://doi.org/10.1016/j.rser.2020.110618>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



DSSPACE

<https://dspace.org/>

Application of machine learning in power grid fault detection and maintenance

Olojede, David; King, Stephen; Jennions, Ian

2025-09-26

Springer

<https://doi.org/10.1186/s42162-025-00574-w>

Olojede D, King S, Jennions I. (2025) Application of machine learning in power grid fault detection and maintenance. Energy Informatics, Volume 8, September 2025, Article number 119