Data-Driven Approaches to Reliability Modeling in Critical Energy Infrastructure: Analytical Perspectives on Early Failure Detection

Derrick Ohene Adusei

Northeastern University, Boston, MA

Abstract:
The growing complexity and interconnection of critical energy infrastructure, which stems out of integration of renewables, decentralization of generation, and cyber-physical integration, has increased the pressure on the sophisticated reliability modeling and early warning of failure systems. Although traditional probabilistic methods are grounded, they have disadvantages in terms of being state-based and less adaptive to high-dimensional operation data. This paper is a PRISMA systematic literature review of data-driven reliability models in critical energy infrastructure based on publications in ScienceDirect, having selected on the use of structured Boolean search operators. Eighty-eight studies were retained out of an original 296 records after a process of removal of duplicates and systematic screening. The review constructs a taxonomy of statistical, machine learning and deep learning, and hybrid physics-informed models, and assesses their applicability in early failures estimation of transformers, wind turbines, photovoltaic systems, nuclear systems as well as smart grids. A comparative analysis shows increasing popularity of ensemble and deep learning architectures, specifically, to anomaly detection and remaining useful life (RUL) prediction. Nevertheless, there are chronic problems such as rare-event imbalance, benchmarking fragmentation, limited cross-domain transferability, interpretability limitations as well as a lack of validation on real-world operational data that limit cumulative improvements. The paper lists such strategic research priorities as the integration of physics-informed AI, federated and edge learning models, frameworks of transfer learning, unlabelled data self-supervised learning, and standardized benchmarking data. The results give an all-purpose analytical starting point in the Standard of predictive, scalable, and robust early failure detection systems in the future energy infrastructure.

Keywords: Data-driven reliability modelling, Early failure detection, Predictive maintenance, Critical energy infrastructure, Physics-informed artificial intelligence**.**

1. INTRODUCTION

The critical energy infrastructure has become a highly interdependent cyber-physical system which contributes to national economic stability, safety of people, and resilience. The high rate of adoption of renewable energy, distributed generation, storage technologies and smart protection systems have fundamentally redesigned the architecture of traditional grids, ensuring that complexity in operations and interdependence of the system has risen. The current power systems have gone digital and physical, with SCADA platforms, edge devices, and cloud analytics, and the high-resolution data streams constantly change and exchange with each other. Although such improvements increase the efficiency and flexibility, they also increase the exposure to cascading failures, cyber disruptions and operational instability. The use of AI-based automation and predictive diagnostics is also becoming a significant enabler of the safe and resilient infrastructure management (Opoku & Adeoye, 2025; Bonsu & Adeoye, 2025; Adeoye et al., 2025). General surveys of AI application in the smart grid and renewable systems also focus on the strategic application of smart monitoring to protect energy resilience (Banad et al., 2025; Wang et al., 2025). As a result, the concept of reliability modeling in critical energy infrastructure should not be presented as a technical role but as a national resilience need.

Classical reliability modeling has been based on deterministic and probabilistic models like the Weibull distributions, Markov chain, and fault tree analysis to estimate failure rate and hazard functions. These methods are still fundamental to reliability engineering (Nor et al., 2021), but they are usually limited by the assumptions of stationarity, lack of scalability, and inability to adapt quickly to the changing operational circumstances. Statistical distribution of failures in digitally intensive systems with renewable intermittency and cyber-physical interaction, nonlinear degradation processes and new fault mode structures cannot be readily addressed with a class of statistical distributions. The reviews of the programmable logic controllers that have become increasingly unreliable in the advanced nuclear automation and AI-driven diagnostics in the nuclear control systems indicate the increasing inapplicability of the purely model-based or rule-based reliability frameworks (Adeoye, 2025; Bonsu & Adeoye, 2025). In the same fashion, the joint analyses of digital twins and predictive maintenance in energy systems indicate that adaptive and data-responsive models that can learn on the basis of real-time sensor streams are necessary (Zhong et al., 2023; Das et al., 2024; Amin et al., 2025). The spread of the high-resolution monitoring technologies such as the phasor measurement units, vibration sensors and IoT-enabled devices has thus expedited a paradigm shift in the data-driven modeling of reliability.

The advent of machine learning, deep learning, and hybrid physics-informed computations has changed the view of reliability assessment as a forecast of failure probability at a given operating point to a pattern search and discovery in prognostics. The real-time diagnostics driven by AI, self-correcting control schemes, and robotics-aided inspection systems have been shown to have a higher sensitivity of anomaly detection and responsiveness when it comes to nuclear and high-risk industrial environments (Opoku & Adeoye, 2025; Adeoye et al., 2025; Arthur et al., 2025). Data-driven reliability assessment frameworks, ensemble learning models and systems of anomaly detection are also increasingly applied to distribution networks, photovoltaic systems, wind turbines and transformers in renewable applications, and grid applications (Obatola & Junjie, 2024; Ukwuoma et al., 2025; Safari et al., 2024). The situation with early failure detection is especially important as it minimizes disastrous failures, allows depreciating in advance, and increases the accuracy of remaining useful life (RUL) prediction (Liu et al., 2025). Nevertheless, literature is still in disjointed form in infrastructure domains, mostly methodological in form and seldom structured in a system-level manner. The predictive analytics and process optimization reviews also indicate inconsistencies with benchmarking, validation datasets, and lead-time reporting of detectives (Ndlovu & Oware, 2025; Shopeju & Oware, 2025). Additionally, the transferability of analytics in relation to cross infrastructure is a critical area of need, considering the advanced geospatial and computational studies in risk prediction are still not well developed (Yator & Aliu, 2026). It is against this background that the given study will do a PRISMA-based systematic review to come up with a structured taxonomy of data-based reliability models, assess early detection paradigms, benchmarking gaps, and positive research roadmap of resilient and infrastructure-integrated energy systems.

2. REVIEW METHODOLOGY

The paper takes a PRISMA-based systematic review approach that will enhance methodological rigor and replicability. On 20/02/2026, the search was carried out in the ScienceDirect library under the Boolean operator: ("data-driven reliability" OR "predictive maintenance" OR "PHM") AND ("power systems" OR "energy infrastructure") AND ("early failure detection" OR "fault prediction"). The search resulted in 296 records, among which 209 articles were published in 2021-2025, which is indicative of the active pace of the development of AI-based reliability studies in the last several years. In the dataset, there were 80 review articles and 106 research articles (186 total journal articles), and 96 publications were open access or open archive. After identifying the duplicates, relevancy sorting, and the initial screening, 88 studies were left to undergo full-text evaluation, thereby suggesting that all removals (duplicates and unrelated records) were done before this ultimate screening process. Only English-language peer-reviewed journal articles were considered in the review to determine the quality and consistency of the research and eliminate conference papers, editorials, and non-scholarly publications.

Table 1. PRISMA Study Selection

Stage	Description	Number (n)
Identification	Records identified through ScienceDirect search (20/02/2026)	296
Time Filter	Records from 2021–2025	209
Article Type Filter	Review Articles (80) + Research Articles (106)	186
Accessibility	Open Access / Open Archive	96
Screening	Records after duplicates removal and sorting	88

The selection of the studies was done in accordance with PRISMA stages: identification (296 records), exposure to duplicates and sorting, title/abstract screening, and full-text eligibility evaluation. The studies were used when they concentrated on critical energy infrastructure, employed explicit methods of reliability modelling, and contained an early warning element, including anomaly detection or remaining useful life (RUL) forecasting. Questionable studies that this investigation was restricted to maintenance planning, non-energy systems, or even the hypothetical although not empirically validated discussions were excluded. An identical data extraction protocol was then implemented in the final eligible studies and the data included infrastructure type, data size, sensor modality, model type (statistical, machine learning, deep learning, or hybrid), lead time of detection, and evaluation metrics (e.g., Accuracy, F1-score, ROC-AUC, RMSE). Such a methodology provides transparency, cross-study comparability and a strong analytical base on which taxonomy can be developed, and synthesis benchmarking can be conducted.

Table 2. Study Characteristics

Author	Infrastructure	Model	Data Type	Detection Focus	Metric
Obatola & Junjie (2024)	Grid-Connected PV	Deep Learning + Hybrid Optimization	Electrical + Operational Data	Reliability Assessment	Accuracy, RMSE
Ukwuoma et al. (2025)	Transmission Systems	Comparative ML Models	Electrical Fault Signals	Fault Detection & Classification	F1, ROC-AUC
Safari et al. (2024)	Power Electronics	Isolation Forest (Unsupervised)	SCADA Time-Series	Anomaly Detection	Precision, Recall
Cancemi et al. (2025)	Nuclear Power Components	Hybrid Neural + Statistical	Operational Time-Series	Predictive Monitoring & RUL	RMSE
Lyu et al. (2025)	Nuclear Power Applications	Deep Learning Dual Monitoring	SCADA + Forecast Data	Fault Detection	Accuracy
Aghahadi et al. (2025)	Distribution Grids	Comparative ML Models	Grid Operational Data	Fault Prediction	ROC-AUC
Rengasamy & Rajesh (2025)	Wind Turbines	RF-XGBoost Hybrid	Vibration + SCADA	Fault Detection	Accuracy, F1
Rafati & Shaker (2024)	District Heating Networks	Data-Driven Predictive Models	Sensor + Thermal Data	Reliability & Asset Monitoring	RMSE

3. ANALYTICAL EVOLUTION OF RELIABILITY MODELING

Classical probabilistic models based on failure rate modeling and hazard function estimation are the analytical basis of reliability modeling of critical energy infrastructure. Probabilistic risk assessment methods used in the past to support nuclear plants, transmission systems, and industrial energy assets have included Weibull distributions, exponential failure laws, Markov chains, and fault tree analysis (Nor et al., 2021). These models measure mean time to failure and reliability of the system on predefined conditions during operation. Their predictive ability, however, is limited by assumptions of both static and stationarity, constant transition probabilities and scalability in high dimensional settings. The contemporary energy system and renewable intermittency lead to a dynamic degradation pattern, cyber-physical integration, and decentralized control architecture, instead of being distributed over a fixed distribution, patterns. Surveys of the architecture of programmable logic controllers and AI-based diagnostics in large-scale nuclear systems highlight the increasing lack of viability of rule-based and static reliability models in digitally intensive systems (Adeoye, 2025; Bonsu & Adeoye, 2025). As a result, classical reliability modeling is inherently less accurate to capture real time dynamics of degradation and early signs of faults.

These constraints of the static probabilistic modelling led to the emergence of condition-based monitoring (CBM) that added real time operational awareness to the reliability management (Nor et al., 2021; Rafati & Shaker, 2024). CBM changed the maintenance approaches depicted by time scheduling to that of performance intervention through SCADA platforms, IoT-enabled sensors, and continuous vibration, thermal, and electrical monitoring (Tan et al., 2023). The main part of this strategy is the development of health indices, the aggregated measures, which depict the degradation conditions of equipment (Rafati & Shaker, 2024). The use of predictive monitoring systems on nuclear components, wind turbines, and other energy infrastructure illustrates how real-time sensing can be used to promote the detection of anomalies in stages and operational sensitivity (Cancemi et al., 2025; Rengasamy & Rajesh, 2025). However, initial CBM applications tended to be threshold based or shallow statistical based, prohibiting their capability to identify nonlinear and multivariate trends in degradation (Nor et al., 2021). Therefore, CBM marked an intermediate step-the introduction of dynamic data streams in the reliability assessment, which, however, had not yet been fully used in the form of predictors (Zhong et al., 2023; Das et al., 2024).

The modern data-driven paradigm signifies a radical conceptual shift between the estimation of the probability of failing to prediction of patterns and prognostic intelligence. Random Forest, Support Vector machines and Gradient Boosting are supervised learning models that learn how to map engineered features to fault states, which enhances the classification and early fault detection accuracy of photovoltaic infrastructures and transmission systems (Ukwuoma et al., 2025). The further application of deep learning architecture, such as spatial signal interpretation with CNN and time dependence modeling with LSTM, facilitates the automated learning of representation of high-dimensional SCADA and sensor streams. A good example of this trend toward self-correcting and adaptive reliability frameworks is hybrid neural-statistical models and AI-based real-time diagnostic in nuclear systems (Adeoye et al., 2025; Opoku & Adeoye, 2025).

4. TAXONOMY OF DATA-DRIVEN RELIABILITY MODELS

A methodological development of interpretable probabilistic inference to high-dimensional representation learning is indicated by the taxonomy of data-driven reliability models in critical energy infrastructure. The approaches to reliability modeling in the nuclear systems, in renewable energy installations, smart grids, and hybrid cyber-physical architectures are divided into three broad categories: (i) statistical and probabilistic learning models, (ii) machine learning models, and (iii) deep learning and hybrid architectures. This classification is consistent with the conceptual clarification that hazard-rate estimation is no longer relevant, but the focus of success ought to be on patterns and earlier detection of failure and anticipatory therapy (Nor et al., 2021; Rafati & Shaker, 2024; Liu et al., 2025).

4.1 Statistical and Probabilistic Learning Models

The advantage of statistical learning models in the context of reliability modeling is that they are interpretable, and regulators allow them to be used in systems of high safety. Logistic regression is also used extensively in binary fault classification in photovoltaic infrastructure and transmission systems where the failure probability is modeled as a parametric(function) of operational variables (Ukwuoma et al., 2025; Obatola & Junjie, 2024). The structure of its coefficients gives it transparency, and it can directly be interpreted regarding the effects of temperature gradients, load stress or harmonic distortions on failure likelihood. The Bayesian inference is an expansion of the classical risk assessment based on probabilistic risk analysis, which includes priori knowledge and a posteriori failure probability update as sensor evidence is acquired- a solution that becomes more and more applicable in nuclear automation and control systems (Bonsu & Adeoye, 2025; Adeoye, 2025). In the next generation nuclear systems, Bayesian reasoning is applied to uncertainty-aware AI diagnostics to improve reliability decision-making in automated structures of monitoring (Adeoye et al., 2025).

The survival analysis and especially the Cox proportional hazards modeling can be useful as it bridges the classical hazard-rate theory with operational data in time, so that it is possible to model a time-to-failure in the district heating systems, grid assets and energy conversion technologies (Rafati & Shaker, 2024; Ndlovu & Oware, 2025). These models prove to be beneficial where the datasets available are small or when their interpretation is required by the regulatory control. Nonlinear degradation dynamics that are typical of cyber-physical energy systems are, however, predicted by them only to the extent that they are linear, but with proportional hazard constraints. The more complex the infrastructure is, the less sensitive to weak early-failure signals hidden in high-dimensional sensor streams purely statistical methods are (Banad et al., 2025; Shadi et al., 2025). Therefore, statistical models offer interpretable baselines but their performance on its own in the context of the early detection of failure in the contemporary energy systems is moderate.

4.2 Machine Learning Models

Machine learning (ML) algorithms handle nonlinear interaction of degradation and multifaceted association of features that traditional statistical tools are unable to follow. The algorithms based on the use of random forests (RF) take advantage of the use of an ensemble of decision trees to enhance strength and eliminate overfitting, showing excellent results in the classification of transformer faults, photovoltaic reliability evaluations, and diagnostics of transmission systems (Ukwuoma et al., 2025; Obatola & Junjie, 2024; Azmi et al., 2025). They offer semi-interpretability, which is the importance metrics embedded within them, which allows identifying the major drivers of failures. High-dimensional spaces, where the nonlinear kernel transformations increase the separation of classes, are the settings where Support Vector Machines (SVMs) work effectively and, therefore, it can be used in vibration based and electrical signal based fault detection (Safari et al., 2024; Rengasamy & Rajesh, 2025). Gradient Boosting models are optimistic to predictive quality and have been shown to have better rare event forecasting performance in the research of smart grids and faults in distribution (Aghahadi et al., 2025; Senapati et al., 2025).

Nevertheless, the quality of feature engineering is still critical to ML models. The domain indicators that are frequency-domain descriptors, statistical moments, and domain-specific indicators rely directly on the transformation of the raw sensor signals. There is also a strong class imbalance in reliability datasets with failure events being sparse when compared to normal operations. In the absence of countermeasures, e.g. synthetic minority over-sampling, cost-sensitive learning or ensemble balancing, ML models will be prone to being biased concerning the majority classes and suffer low early detection sensitivity. This imbalance issue is highly documented in research of faults in the transmission systems and diagnostics of renewable assets (Ukwuoma et al., 2025; Aghahadi et al., 2025). Also model training can be made infrastructure-specific, which can reduce cross domain transferability requiring retraining of heterogeneous assets. In spite of these issues, ML models always excel in the ability to detect faults early, especially when labelled data sets are sufficiently large and powerful enough to preprocess pipelines.

The other important benefit of ML approaches is that it is compatible with predictive maintenance architectures. The example of AI-driven hazard detection in the mining industry and hazardous industrial setting shows how ensemble-based learning contributes to increasing the sensitivity of anomalies in the initial stages but with stability in the operations. These results are in line with nuclear predictive automation experiments that show better diagnostic responsiveness in the case of nonlinear learning architecture incorporated in control architectures (Opoku & Adeoye, 2025; Adeoye et al., 2025). Therefore, ML solutions are a good compromise between predictive capability and feasibility.

4.3 Deep Learning and Hybrid Models

The most sophisticated level of data-driven reliability modeling is the deep learning (DL) architecture, which allows extracting hierarchical features in high-dimensional sensor data in an automated manner. Convolutional Neural Networks (CNNs) are specifically useful at analyzing the spatial degradation, such as vibration spectrograms and thermal images in rotating machines and transformer systems (Khan & Byun, 2024; Zhou et al., 2025), whereas Long Short-Term Memory (LSTM) networks are effective at analyzing the temporal processes of the SCADA time-series to predict the remaining useful life (RUL) in both nuclear and grid systems (Cancemi et al., 2025; Liu et al., 2025). Autoencoders particularly within unsupervised conditions learn the normal operation behavior and identify variations that would signify the onset of faults (Safari et al., 2024; Shadi et al., 2025). These architectures minimize the use of manual feature engineering and can much better be better able to detect faults very early because they can directly infer latent degradation behavior through the use of data.

Hybrid frameworks also enhance the predictive reliability, which combines the physics-based constraints and neural learning. Physics-informed neural networks (PINNs) incorporate, in the training process, equations that control operating processes, which provides a better generalization and physically consistent prediction. The next-generation nuclear systems will involve AI-based diagnostics and robotics-controlled predictive maintenance (Opoku & Adeoye, 2025; Adeoye et al., 2025; Bonsu & Adeoye, 2025) as one example of such hybridization, and geospatial computational analytics will apply the same principles to mining and subsurface risk prediction (Yator & Aliu, 2026). Although deep and hybrid models have a better nonlinear modeling ability, they are data-intensive and computationally expensive, as well as are not interpretable in safety-critical energy contexts (Adeoye, 2025; Shadi et al., 2025). However, these methods have the greatest potential to identify a failure early in a complicated energy system when backed by efficient data infrastructures.

Table 3. Model Taxonomy Comparison

Model Class	Strength	Weakness	Data Volume	Early Detection Suitability
Statistical & Probabilistic	High interpretability; uncertainty quantification; low data need	Linear assumptions; limited nonlinear capture	Low	Moderate
Machine Learning	Nonlinear modeling; robust classification; moderate interpretability	Feature dependence; class imbalance sensitivity	Medium	High
Deep Learning & Hybrid	Automated representation learning; superior nonlinear detection; scalable	Data-hungry; interpretability challenges; computational cost	High	Very High

5. EARLY FAILURE DETECTION MECHANISMS

The early failure detection systems in critical energy infrastructure may be divided into four main paradigms: classification-based, anomaly detection, remaining useful life (RUL) prediction, as well as probabilistic prediction. Classification-based detection is a technique based on supervised learning to classify operational states into a predetermined set of categories such as "normal" or "faulty" typically with the help of an ensemble or deep learning classifier that is trained on labeled failure data (Ukwuoma et al., 2025; Rengasamy & Rajesh, 2025). Although it is very effective in detecting the known types of faults, it also has a limited detection horizon due to historical labelling and it might have difficulties in detecting latent modes of failure. Conversely, anomaly detection, which is often trained with the help of autoencoders or isolation-based models, learns normal working behavior and identifies deviations without fault examples on them (Safari et al., 2024; Shadi et al., 2025). This paradigm can be used especially in incipient or rare defects, which increases the detection horizon at the cost of often making false-positive results caused by sensitivity to benign operational variability.

Remaining Useful Life (RUL) prediction models are not binary fault prediction models, but degradation-trajectory models, predicting time-to-failure based on temporal models (i.e. LSTM or hybrid neural-statistical models) (Cancemi et al., 2025; Liu et al., 2025). RUL approaches offer greater detection horizons and can be used in predictive intervention planning but need large longitudinal data sets. Probabilistic forecasting Relying on Bayesian updating or uncertainty-aware models, probabilistic forecasting predicts the risk of failure in different operational settings and is interpretable and can be regulated in safety-critical settings (Bonsu & Adeoye, 2025; Adeoye, 2025). In any paradigm, there exists a trade-off whereas a researcher increases sensitivity of early detection, false negatives are decreased but false positives increase and cause unneeded interventions. So, to achieve the best deployment, a balance among detection lead time, model confidence and operational risk tolerance is needed especially in nuclear and grid settings where false alarms and missed failures both are of great consequence (Opoku & Adeoye, 2025; Adeoye et al., 2025).

Table 4: Early Detection Comparison

Technique	Lead Time	Data Need	Interpretability	Deployment Readiness
Classification-Based	Moderate (known faults)	Medium-High (labeled data)	Moderate	High
Anomaly Detection	High (incipient faults)	Low-Medium (normal data sufficient)	Low-Moderate	Medium
RUL Prediction	Very High (degradation trajectory)	High (longitudinal data)	Moderate	Medium
Probabilistic Forecasting	Moderate-High	Low-Medium	High	High (regulated systems)

The models of classification are feasible and implementation ready when the classes of faults are clear and well-defined, and the reliance on the data on classified failures restricts the flexibility to new faults (Ukwuoma et al., 2025; Azmi et al., 2025; Zampini et al., 2025). Anomaly detection is suitable in the detection horizon and also useful in the case of rare occurrences, however it can result in increased false-positive rates when the variability of operations is mistaken to be degradation (Safari et al., 2024; Shadi et al., 2025; Khan & Byun, 2024). The most actionable, long-term perspective of planning is included in RUL prediction, which is also time-consuming and requires considerable time-series data and computing power (Cancemi et al., 2025; Liu et al., 2025; Al-Selwi et al., 2024). Probabilistic forecasting is extremely interpretable and quantifies uncertainty, which is why it is very applicable to systems where safety is paramount (nuclear and grid), but it does not necessarily have the nonlinear sensitivity of deep learning forecasting (Banad et al., 2025; Shadi et al., 2025). These paradigms do not exclude each other, but they complement each other. Combined designs of anomaly detection to identify early signals, classification to confirm faults, RUL prediction to predict when to intervene, and probabilistic forecasting to determine the level of risk are the best place forward towards resilient early detection of failure in critical energy infrastructure (Opoku & Adeoye, 2025).

6. INFRASTRUCTURE-LEVEL APPLICATIONS

There has been a wide variation in the reliability modeling strategies across the various areas of infrastructure because of variations in the physics of degradation and data structure. Structured tabular datasets are more suitable in transformer health monitoring dissolved gas analysis and electrical load data than structured tabular data with ensemble machine learning and probabilistic classifiers (Azmi et al., 2025; Ukwuoma et al., 2025). The high-frequency vibration signals are fundamental in wind turbine gears breakdown, and consequently, CNN-based spectral analysis and LSTM time-series models are specifically useful to detect incipient fault (Khan & Byun, 2024; Rengasamy & Rajesh, 2025). The SCADA electrical parameters and environmental variability in the solar inverter degradation usually necessitates the use of the hybrid nonlinear models to describe the performance drift (Obatola & Junjie, 2024). Meanwhile, cyber-physical failures of smart grids require heterogeneous multidimensional electrical and communication data streams, which requires anomaly detection and probabilistic prediction frameworks that can manage multidimensional uncertainty (Banad et al., 2025; Aghahadi et al., 2025).

There are three patterns of structure that are found in domains. First, the type of data determines the choice of models: vibration-dense systems are better served with deep learning, and structured datasets of electrical data are better with ensemble and probabilistic models. Second, there is a sensor density related to the detection horizon, high-resolution monitoring can be detected sooner, but it requires more computational resources. Third, model transferability is not as high, because mechanisms of infrastructure degradation-specific to the infrastructure are limiting cross-asset generalization without domain adaptation (Opoku & Adeoye, 2025;

Adeoye et al., 2025). Infrastructure-level reliability performance is therefore not established entirely on the sophistication of the algorithm but the compatibility between the sensing architecture, data quality and the physical system attributes.

7. PERFORMANCE EVALUATION AND BENCHMARKING

One of the most fragmented stream in the literature is the performance assessment in the reliability modeling that is based on data. Accuracy, F1-score and ROC-AUC are the most common metrics reported in the classification-based detection literature. Accuracy gives an approximate idea of the number of correct predictions but is inaccurate when there is a very imbalanced reliability dataset with normal operating states being mostly more than operations in failure cases (Ukwuoma et al., 2025; Aghahadi et al., 2025). In turn, F1-score, a more balanced score in terms of precision and recall, provides more informative data when the main goal is to find rare faults. ROC-AUC also assesses classifier separability with the variation of threshold, and it gives an idea of discrimination capacities that are independent of class distribution. Root Mean Square Error (RMSE) is the most dominant in regression-based metrics in the remaining useful life (RUL) prediction studies, which measure the error between the predicted and actual failure time (Cancemi et al., 2025; Liu et al., 2025). Nonetheless, RMSE by itself is not an indicator of operational implications of early and late prediction, which restrict its applicability in safety-critical settings.

Heterogeneous datasets, different sensor densities, inconsistent preprocessing pipelines, and infrastructure-related degradation properties further complicate the process of cross-study comparison. The datasets of transformer DGA do not resemble wind turbines vibration spectra or SCADA photovoltaic logs, and their reporting of performance scores do not provide universal comparative performance scores of any kind (Obatola & Junjie, 2024; Rengasamy and Rajesh, 2025). Also, the issue of persistent dataset imbalance gives models an artificial advantage of majority of health classes, artificially inflating accuracy, and obscuring poor sensitivity to early faults (Ukwuoma et al., 2025; Safari et al., 2024). Even though resampling or cost-sensitive learning can be used in some studies, standardized imbalance-handling protocols are hardly described. Another constraint is the unavailability of open benchmarking repositories of energy infrastructure reliability especially in nuclear and grid systems where security restrictions place limits on the sharing of data (Adeoye, 2025; Bonsu & Adeoye, 2025). Losing reproducibility and making objective comparison of modeling paradigms impossible is due to the lack of shared datasets and uniform assessment programs.

Table 5. Benchmarking Summary

Study	Model	Dataset	Metric	Reported Score
Ukwuoma et al. (2025)	Ensemble ML	Transmission Fault Signals	F1, ROC-AUC	High AUC (>0.90 reported)
Obatola & Junjie (2024)	Deep Learning + Hybrid Optimization	Grid-Connected PV SCADA	Accuracy, RMSE	High Accuracy (>90%)
Cancemi et al. (2025)	Hybrid Neural-Statistical	Nuclear Component Time-Series	RMSE	Low RMSE (improved RUL precision)
Rengasamy & Rajesh (2025)	RF–XGBoost Hybrid	Wind Turbine Vibration	Accuracy, F1	Improved F1 vs baseline
Aghahadi et al. (2025)	Comparative ML Models	Distribution Grid Fault Data	ROC-AUC	Strong separability (>0.85 AUC)

8. CRITICAL CHALLENGES AND METHODOLOGICAL GAPS

Although the methodological progress has been fast, the state of the art in early failure detection of critical energy infrastructure is limited to structural data and modeling. The most important of these is the problem of rare events imbalance: failure events have a limited proportion by definition when compared to normal operational states, which results in biased datasets causing bias in the classifiers due to the prevalence of majority classes (Ukwuoma et al., 2025; Safari et al., 2024). Partial mitigation of imbalance is achieved by resampling methods and cost-sensitive learning, but synthetic bias is frequently created, or physical degradation distributions are distorted. False negatives (missed failures) in safety critical areas such as nuclear systems are catastrophic, whereas the tuning of the sensitivity to aggressively minimize false positives and disruption of operation leads to false positives (Bonsu & Adeoye, 2025; Opoku & Adeoye, 2025). This unresolved tradeoff between sensitivity of detection and operational reliability is one of the underlying methodological issues especially in cases where there are weak early-failure signals that are hidden in sensor streams with high noise.

The second significant gap is in the overfitting of models to infrastructure-specific data and cross-domain low transferability. Numerous studies perform well in controlled experimental environments, e.g. wind turbine vibration data or photovoltaic SCADA monitoring, but fail to justify models on heterogeneous resources or operating conditions (Rengasamy & Rajesh, 2025; Obatola & Junjie, 2024). There is no standardized benchmark on multi-infrastructure, and models that are trained on one site tend to perform poorly when they are put into other environmental or load settings. People are trying to incorporate domain knowledge into learning structures to make them more robust (Adeoye et al., 2025; Adeoye, 2025) with examples of AI-based nuclear automation and hybrid diagnostic systems, but no systematic cross-domain validation has been done. In the same vein, the use of geospatial predictive systems in the mining case also emphasizes the significance of contextual physics in the model stability, which explains why strictly data-driven models with no domain adaptation tend to be brittle (Yator & Aliu, 2026). Therefore, domain generalization and transfer learning are undeveloped in the reliability research.

There are emerging interpretability crisis and dilemma of cybersecurity vulnerability in the field. Deep learning and hybrid neural networks provide high-quality nonlinear detection, but their decisions are opaque, which will not be regulatory in the grid and nuclear setting (Shadi et al., 2025; Adeoye, 2025). The early failure signal would not be trusted to operate without a reason to have an impact, which restricts the deployment capability. Meanwhile, the growing dependence on IoT-based surveillance and cyber-physical interconnection broadens the attacker base to attack adversarial control and data poisoning especially in the context of smart grids (Banad et al., 2025). Besides, there are many suggested models that are tested on simulated or partially anonymised datasets (as opposed to actual operational data of the longitudinal nature) which reduces their empirical credibility (Cancemi et al., 2025; Liu et al., 2025). The imbalance bias, limited transferability, interpretability limitations, cybersecurity risk, and lack of field validation prove that the methodological sophistication is not enough. To enhance the state of the early failure detection, integrated solutions that will unite resilient validation protocols, model-driven based on domain knowledge, safe data architecture, and explainable AI frameworks specific to safety-critical energy infrastructure are needed.

9. RESEARCH ROADMAP

To spearhead the development of early warning systems in the critical energy infrastructure, there should be a transition of physics-inspired and hybrid AI systems, which combine domain information with adaptability to data. Purely black-box models are powerful, but not robust when there are changes in the regimes of operation as well as in cases where there are rare failure data. Generalization can be enhanced by embedding governing equations, thermodynamic constraints, and degradation physics within neural architectures via physics-informed neural networks and hybrid control-learning systems, which can also guarantee physically consistent predictions. Uncertainty-conscious diagnostics and probabilistic reasoning should be included in AI systems in nuclear and grid settings to maintain the safety, explainability, and regulatory compliance. The roadmap thus gives more weight to the models that do not just provide accurate models, but also models that are physically based and operationally credible.

The second strategic path is distributed intelligence and cross asset knowledge transfer. Since wind farms, distributed renewable installations, and the substations produce huge, localized data streams, edge AI can be used to perform real-time anomaly detection on a device or substation-level to minimize latency and cybersecurity liability. Federated learning, complementary to it, enables joint model training via geographically separated resources without the central distribution of data, which is a key feature of security-sensitive systems like nuclear systems and smart grids. In order to resolve the infrastructure-specific overfitting issue, transfer learning approaches should be created so that models, trained on one type of asset (e.g., wind turbines), can be transferred to other assets (e.g., hydro generators or transformers) with lower retraining costs, and with higher scalability. It is consistent with the demands of more generalizable predictive maintenance schemes across energy systems.

The discipline needs to resolve the shortage of data and fragmentation of benchmarking by using self-supervised learning and standardized evaluation models. Self-supervised methods are able to utilize the rich unlabeled signals of operation to be able to learn representations of degradation without necessarily being fault-labeled, thus reducing the issue of imbalance in rare-event occurrences. At the same time, by creating openly and standardized benchmarking sets, including transformers, wind turbines, photovoltaic systems, and cyber-physical grid elements, it would be possible to compare the detection paradigms objectively as well as expedite the convergence of methods. Unless common standards are established and aligned reviewing guidelines, reported performance gains are localized and cannot be transferred. Reliability modeling should not just be algorithmically innovative in the future, but ecosystem-level integrated: physics-informed intelligence, distributed, and secure learning systems, transferable models and transferable benchmarking standards.

10. CONCLUSION

In this review, it is evident that the field of reliability modeling has moved away from classical probabilistic reliability models to machine learning, deep learning, and hybrid physics-informed models. Structured fault classification problems in transformers and grid systems are dominated by ensemble models and CNNs, LSTM and autoencoders are becoming the norm of early failures detection problems based on automated feature learning and time-related degradation modeling A new category of robust solutions are also coming to play in safety-critical domains, which involves the synthesis of predictive intelligence and regulatory-compliant interpretability, which is referred to as hybrid and uncertainty-aware systems. Anomaly detection and RUL prediction paradigms have the longest detection horizons, as compared to other detection paradigms, whereas probabilistic forecasting is important in making risk-sensitive decisions.

Nevertheless, there are still outstanding challenges. The lack of cross-domain applicability, an imbalanced dataset, and the absence of hyperdomain fragmentation, inadequate cross-world validation, and reproducibility limits reproducibility and confidence of deployment. Well-known metrics of performance like Accuracy, F1-score, ROC-AUC and RMSE are not always reported consistently, restricting the cross-study comparability. To deal with these gaps, physics-informed AI integration, benchmark datasets of multi-infrastructures that are standardized, federated and edge learning deployment frameworks, and imbalance-sensitive evaluation protocols are needed. The future of critical energy infrastructure is transferring away reactionary fault response to predictive, adaptive, and resilient intelligence structures that are able to foresee failure prior to being disrupted.

REFERENCES:

1. Adeoye, M. B., Annankra, J. A., & Yakin, Z. (2025). A review of programmable logic controllers in advanced nuclear plant automation: Challenges and future prospects. *International Journal of Frontline Research in Science and Technology, 5*(1), 1–8
2. Adeoye, M. B., Annankra, J. A., & Yakin, Z. (2025). AI-driven real-time diagnostics and self-correcting control schemes for next-generation nuclear energy systems. *World Journal of Advanced Research and Reviews, 27*(3), 1550–1557
3. Aghahadi, M., Bosisio, A., Pegoiani, A., Forciniti, S., Merlo, M., & Berizzi, A. (2025). Predicting faults in power distribution grids during heatwaves: A comparative study of machine learning models applied to Milan distribution network. *Sustainable Energy, Grids and Networks*, *43*, 101741.
4. Amin, U., Kim, D., Ahmed, F. N., Ahmad, G., & Hossain, M. J. (2025). Digital twins for smart asset management in the energy industry: State-of-the-art. *Expert Systems with Applications*, *289*, 128358.
5. Arthur, A. A., Annankra, J. A., & Yakin, Z. (2025). Examining the role of AI and machine learning in improving hazard detection and predictive analytics for accident prevention in mining operations. *World Journal of Advanced Engineering Technology and Sciences, 15(3), 640-646.*

6. Azmi, P. A. R., Yusoff, M., & Sallehud-din, M. T. M. (2025). Improving transformer failure classification on imbalanced DGA data using data-level techniques and machine learning. *Energy Reports*, *13*, 264-277.
7. Banad, Y. M., Sharif, S. S., & Rezaei, Z. (2025). Artificial intelligence and machine learning for smart grids: From foundational paradigms to emerging technologies with digital twin and large language model-driven intelligence. *Energy Conversion and Management: X*, *28*, 101329.
8. Bonsu, J. O., & Adeoye, M. B. (2025). AI-Driven Fault Detection and Diagnostics in Nuclear Power Plant Control Systems: A Review. *Journal Of Engineering And Computer Sciences, 4(9), 552-559.*
9. Cancemi, S. A., Angelucci, M., Chierici, A., Paci, S., & Frano, R. L. (2025). Hybrid neural network and statistical forecasting methodology for predictive monitoring and residual useful life estimation in nuclear power plant components. *Nuclear Engineering and Design*, *433*, 113900.
10. Das, O., Zafar, M. H., Sanfilippo, F., Rudra, S., & Kolhe, M. L. (2024). Advancements in digital twin technology and machine learning for energy systems: A comprehensive review of applications in smart grids, renewable energy, and electric vehicle optimisation. *Energy Conversion and Management: X*, *24*, 100715.
11. Khan, P. W., & Byun, Y. C. (2024). Fault prognosis of wind turbines using multimodal machine learning. *Energy Reports*, *12*, 4938-4952.
12. Liu, Y., Wen, J., & Wang, G. (2025). A comprehensive overview of remaining useful life prediction: From traditional literature review to scientometric analysis. *Machine Learning with Applications*, 100704.
13. Lyu, M., Gong, H., Chen, Z., Wang, J., Zhong, M., Wang, Z., ... & Pan, Z. (2025). Deep learning-based dual monitoring system for power forecasting and fault detection in nuclear power applications. *Energy and AI*, *20*, 100515.
14. Ndlovu, N. T., & Oware, D. (2025). A review of predictive analytics in food waste-to-energy technologies: Opportunities for enhancing sustainability in the United States. *EPRA International Journal of Climate and Resource Economic Review, 13(5), 15*
15. Nor, A. K. M., Pedapati, S. R., & Muhammad, M. (2021). Reliability engineering applications in electronic, software, nuclear and aerospace industries: A 20 year review (2000–2020). *Ain Shams Engineering Journal*, *12*(3), 3009-3019.
16. Obatola, S. O., & Junjie, T. (2024). A data-driven approach to grid-connected PV system reliability assessment: Combining deep learning and hybrid optimization. Energy Reports, 12, 5582-5593.
17. Opoku, J. A., & Adeoye, M. B. (2025). Harnessing AI and Robotics-Assisted Automation for Predictive Maintenance of Next-Generation Nuclear Energy Infrastructure. *Journal Of Innovative Science, 1(2), 1-8*
18. Rafati, A., & Shaker, H. R. (2024). Predictive maintenance of district heating networks: A comprehensive review of methods and challenges. *Thermal Science and Engineering Progress*, *53*, 102722.
19. Rengasamy, P., & Rajesh, R. (2025). Explainable Artificial Intelligence Framework for Wind Turbine Fault Detection using Random Forest-Extreme Gradient Boosting Hybrid Model. *Results in Engineering*, 107446.
20. Safari, A., Sabahi, M., & Oshnoei, A. (2024). ResFaultyMan: An intelligent fault detection predictive model in power electronics systems using unsupervised learning isolation forest. *Heliyon*, *10*(15).
21. Senapati, M., Panigrahi, P. K., Mohanty, A., Rajamony, R. K., Ray, P. K., & Allasi, H. L. (2025). Intelligent Protection Systems for Grid-Connected Renewables: A Review of AI Techniques and Applications. *Results in Engineering*, 107863.
22. Shadi, M. R., Mirshekali, H., & Shaker, H. R. (2025). Explainable artificial intelligence for energy systems maintenance: A review on concepts, current techniques, challenges, and prospects. *Renewable and Sustainable Energy Reviews*, *216*, 115668.
23. Shopeju, O., & Oware, D. A. (2025). Enhancing the efficiency of waste-to-energy plants through advanced process optimization and control strategies. *EPRA International Journal of Multidisciplinary Research (IJMR), 11(8), 65*

24. Tan, Y., Zhou, L., Xue, X., & Duan, B. (2023). Exploration of key technologies for equipment operation and maintenance based on new power systems. *International Journal of Thermofluids*, *20*, 100482.
25. Ukwuoma, C. D., Cai, D., Ukwuoma, C. C., Otuka, C. I., & Huang, Q. (2025). Comparative analysis of data-driven models on detection and classification of electrical faults in transmission systems: Explainability, applicability and industrial implications. *Alexandria Engineering Journal*, *127*, 75-91.
26. Wang, Q., Li, Y., & Li, R. (2025). Integrating artificial intelligence in energy transition: A comprehensive review. *Energy Strategy Reviews*, *57*, 101600.
27. Yator, K. Y., & Aliu, A. (2026). Advanced Geospatial and Computational Analytics for Predicting Subsidence and Slope Failure in the USA Mining Regions.
28. Zhong, D., Xia, Z., Zhu, Y., & Duan, J. (2023). Overview of predictive maintenance based on digital twin technology. *Heliyon*, *9*(4).
29. Zhou, Q., Chai, B., Guo, Y., Li, T., Zhou, S., Wang, K., & Ye, Y. (2025). Multimodal mechanical wear fault diagnosis: Fusion of signal characterization and image information. *Results in Engineering*, 107204.