sustainability

MDPI

Review

A Digital-Twin-Enabled Resilience Framework (DTERF) for Machine-Learning-Based Anomaly Detection in High-PV Cyber–Physical Smart Grids

Franco Fernando Yanine [1], Mauricio Hidalgo [2,*], Jonathan Frez [3], Challa Krishna Rao [4] and Sarat Kumar Sahoo [5]

[1] Facultad de Ingeniería, Universidad Finis Terrae, Santiago 7501015, Chile; fyanine@uft.cl
[2] Facultad de Ingeniería, Universidad San Sebastián, Santiago 8420524, Chile
[3] Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Santiago 8370191, Chile; jonathan.frez@udp.cl
[4] Department of Electrical and Electronics Engineering (EEE), Aditya Institute of Technology and Management, Srikakulam 532203, India; krishnarao.challa@gmail.com
[5] Department of Electrical Engineering, Parala Maharaja Engineering College, Biju Patnaik University of Technology, Berhampur 761003, India; sksahoo.ee@pmec.ac.in
* Correspondence: mauricio.hidalgob@uss.cl

Abstract

The rapid integration of solar photovoltaic (PV) generation, distributed energy resources, and advanced communication infrastructures is transforming conventional power systems into highly interconnected cyber–physical smart grids. Although this transition improves sustainability and operational flexibility, it also increases grid-management complexity and introduces cyber–physical vulnerabilities, including false data injection attacks, communication failures, equipment degradation, and renewable-induced operational instabilities. This paper presents the Digital-Twin-Enabled Resilience Framework (DTERF), a conceptual reference architecture for anomaly detection in high-PV cyber–physical smart grids. DTERF integrates heterogeneous cyber–physical data acquisition, Digital Twin-based contextual representation, machine-learning analytics, explainable decision support, adaptive operational response, continuous learning, and self-healing capabilities within a unified resilience cycle. The framework is grounded in a structured review and comparative assessment of contemporary machine-learning approaches and recent integrated smart-grid research. Its architecture is conceptually evaluated through requirements-to-architecture traceability, examining functional coverage and internal consistency across the complete operational cycle. The analysis shows that DTERF provides explicit architectural mechanisms addressing the principal requirements identified in the literature, including contextual anomaly analysis, interpretability, cybersecurity robustness, resilience support, and operational integration. Rather than proposing a new anomaly detection algorithm or claiming empirical performance superiority, DTERF provides a technology-agnostic architectural foundation for coordinating complementary capabilities required for resilient anomaly management. Future work should empirically validate the framework using Digital Twin simulation environments, representative high-PV distribution systems, cyber–physical anomaly scenarios, and real or utility-derived operational data.

Keywords: digital twin; cyber–physical smart grids; machine learning; anomaly detection; explainable artificial intelligence; self-healing smart grids; grid resilience

Academic Editor: Babatunde Olubayo

Received: 21 July 2026
Revised: 12 August 2026
Accepted: 21 August 2026
Published: 26 August 2026

1. Introduction

The global transition toward low-carbon energy systems has accelerated the deployment of smart grid technologies as a key enabler of sustainable electricity networks. Smart grids combine advanced sensing, communication, automation, and control technologies to support real-time monitoring, bidirectional power flows, distributed energy resources (DERs), energy storage systems (ESS), demand response programs, and advanced metering infrastructure (AMI) [1,2]. These capabilities are essential for enabling large-scale renewable energy integration, improving energy efficiency, and increasing operational flexibility across modern power systems [3].

Among renewable technologies, solar photovoltaic (PV) generation has experienced particularly rapid growth, driven by declining installation costs and increasingly ambitious decarbonization policies. However, high penetration levels of inverter-based renewable resources also introduce important operational challenges related to voltage regulation, frequency stability, power quality, reverse power flows, protection coordination, and system resilience. At the same time, the extensive digitalization of grid infrastructure has transformed electricity networks into complex cyber–physical systems (CPS), in which physical processes are tightly interconnected with communication networks, data analytics platforms, and automated control mechanisms [4,5].

This convergence between renewable integration and digitalization improves observability, flexibility, and operational efficiency, but it also increases system complexity and expands the range of possible disturbances. Cyber-attacks, communication failures, equipment malfunctions, data corruption, renewable generation variability, and abnormal operating conditions can compromise grid stability, reliability, and security [4,6]. Because of the strong interdependence between physical and cyber components, localized disturbances may propagate across multiple layers of the system and potentially trigger cascading failures or large-scale service disruptions.

Recent blackout events have further highlighted the importance of improving situational awareness and resilience in modern power systems. Examples include the widespread outage that affected Chile in February 2025 and the major disturbance experienced across the Iberian Peninsula in April 2025 [7,8]. Although the technical causes and operating conditions associated with these events differ, they illustrate the growing complexity of electricity systems characterized by renewable integration, reduced system inertia, distributed control, and cyber–physical interdependencies. These developments reinforce the need for monitoring and anomaly detection mechanisms capable of identifying abnormal conditions before they evolve into severe operational events.

Traditional anomaly detection methods in power systems are generally based on threshold rules, residual analysis, statistical monitoring, and model-driven techniques. Although these approaches remain useful under well-defined and relatively stable operating conditions, they often struggle to identify subtle, nonlinear, and evolving anomalies in dynamic smart grid environments characterized by distributed generation, variable demand patterns, heterogeneous information sources, and high-dimensional data streams [9]. Their dependence on predefined assumptions may also limit their ability to adapt to previously unseen disturbances, coordinated cyber-attacks, or renewable-induced operational deviations.

Machine learning (ML) has emerged as a promising alternative because it enables complex system behavior to be learned directly from operational data. ML techniques can identify multidimensional relationships, characterize normal operating patterns, and detect deviations without relying exclusively on fixed rules or simplified physical models [10]. Supervised, unsupervised, semi-supervised, and deep-learning approaches have been applied to cyber intrusion detection, false data injection attacks, equipment degradation, abnor-

mal load profiles, and renewable-induced instabilities [11]. In particular, autoencoders, Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Graph Neural Networks (GNNs) provide complementary capabilities for representing the temporal, spatial, and topological dependencies inherent to modern power systems.

Despite these advances, several limitations continue to restrict the operational adoption of ML-based anomaly detection. Most existing studies concentrate on improving individual algorithms or evaluating predictive performance for specific anomaly classes, datasets, or operating scenarios. As a result, machine-learning analytics, cybersecurity mechanisms, renewable integration, explainability, operational response, and resilience are frequently addressed as separate research problems. This fragmented perspective limits the development of end-to-end architectures capable of connecting anomaly detection with contextual interpretation, operational decision-making, mitigation, and system recovery in high-PV cyber–physical smart grids. Practical limitations related to data quality, model interpretability, scalability, cybersecurity robustness, and real-time deployment also continue to hinder large-scale implementation [12].

A further limitation is that conventional ML pipelines generally process measurements as isolated analytical inputs, without maintaining a continuously synchronized representation of the physical system and its operational context. Consequently, detected anomalies may be difficult to distinguish from legitimate variations caused by renewable intermittency, changing topology, asset conditions, or communication disturbances. Digital-twin technologies provide an opportunity to address this limitation by maintaining an updated virtual representation of the grid that connects heterogeneous observations with system topology, operational constraints, asset status, and historical behavior. When integrated with machine learning and explainable decision support, this representation can support a more contextualized and resilience-oriented approach to anomaly detection.

To address these challenges, this paper introduces the Digital-Twin-Enabled Resilience Framework (DTERF), a conceptual architecture for machine-learning-based anomaly detection in high-PV cyber–physical smart grids. Rather than proposing a new detection algorithm, DTERF integrates heterogeneous data acquisition, digital-twin representation, feature engineering, machine-learning analytics, explainable decision support, adaptive operational response, continuous learning, and self-healing capabilities within a unified resilience cycle. In this architecture, anomaly detection is not treated as an isolated classification task, but as an enabling component of a broader operational process that supports situational awareness, mitigation, recovery, and adaptation.

The framework is grounded in a structured analysis of current machine-learning approaches and their suitability for resilient smart grid operation. This analysis is used to identify complementary analytical capabilities, implementation limitations, and architectural requirements that inform the design of DTERF. The resulting architecture is conceptually evaluated through requirements-to-architecture traceability, examining its functional coverage and internal consistency across the complete resilience cycle.

The main contributions of this study are summarized as follows:

- The introduction of the Digital-Twin-Enabled Resilience Framework (DTERF), a conceptual architecture that integrates digital-twin representation, machine-learning analytics, explainable artificial intelligence, adaptive operational response, continuous learning, and self-healing capabilities for high-PV cyber–physical smart grids.
- A structured analysis of machine-learning approaches for smart grid anomaly detection, classifying representative methods according to their learning paradigm, data requirements, operational applicability, interpretability, scalability, and contribution to resilient grid operation.

- The definition of an end-to-end resilience cycle that connects heterogeneous data acquisition, contextual system representation, anomaly detection, explainable decision support, mitigation actions, recovery verification, and operational feedback.
- The identification of implementation challenges and emerging research opportunities involving trustworthy artificial intelligence, federated learning, edge intelligence, autonomous grid operation, and resilient digital-twin technologies.
- A conceptual evaluation of DTERF based on requirements-to-architecture traceability, assessing functional coverage and internal consistency between the requirements identified in the literature and the architectural mechanisms proposed by the framework.

The remainder of this paper is organized as follows. Section 2 reviews the evolution of smart grids, the principal sources of cyber–physical anomalies, and machine-learning approaches for anomaly detection. Section 3 presents the research methodology used to develop and evaluate the proposed framework. Section 4 introduces the architecture and functional layers of DTERF. Section 5 discusses representative application scenarios. Section 6 examines implementation challenges and research implications. Finally, Section 7 presents the conclusions and future research directions.

2. Literature Review

The rapid expansion of renewable energy and the continuous digitalization of electricity networks have fundamentally reshaped the architecture and operation of modern power systems. Contemporary smart grids extend beyond conventional electricity delivery by integrating advanced sensing technologies, communication infrastructures, distributed energy resources (DERs), energy storage systems (ESS), intelligent control mechanisms, and data-driven decision support into a unified cyber–physical environment [13]. While this transformation enhances operational flexibility, efficiency, and sustainability, it also increases system complexity and introduces new operational vulnerabilities that challenge traditional monitoring and protection strategies.

As smart grids evolve toward highly interconnected cyber–physical systems (CPS), anomaly detection has become a fundamental capability for ensuring secure, reliable, and resilient grid operation [14,15]. The growing penetration of renewable generation, particularly solar photovoltaic (PV) systems, further amplifies these challenges by introducing greater operational variability and reducing the effectiveness of conventional model-based monitoring approaches. Consequently, machine learning has emerged as one of the most promising technologies for identifying abnormal operating conditions under increasingly dynamic and data-intensive environments.

Rather than providing a descriptive survey of existing techniques, this section critically reviews the current literature to identify the technological and architectural challenges that motivate the proposed Digital-Twin-Enabled Resilience Framework (DTERF). The review examines the impact of high PV penetration on smart grid operation, analyzes current machine learning approaches for anomaly detection, and identifies the research gaps that continue to hinder the deployment of resilient, explainable, and operationally integrated anomaly detection solutions.

2.1. Smart Grids with High Solar PV Penetration

The global transition toward low-carbon electricity systems is driving a profound transformation of conventional power networks. Traditional grids based on centralized synchronous generation are progressively evolving into decentralized smart grids characterized by extensive integration of distributed energy resources (DERs), advanced communication infrastructures, intelligent control systems, and energy storage technologies [1,3].

This evolution has significantly improved operational flexibility while enabling greater participation of renewable energy sources in electricity generation.

Among renewable technologies, solar photovoltaic (PV) generation has experienced particularly rapid deployment due to declining installation costs, favorable regulatory policies, and ambitious decarbonization targets. As a result, high-PV penetration has become a defining characteristic of modern distribution networks, contributing substantially to the reduction of greenhouse gas emissions and the transition toward more sustainable energy systems.

Despite these benefits, large-scale PV integration also introduces important operational challenges that directly affect grid reliability and resilience. High penetration levels of inverter-based resources (IBRs) may cause voltage fluctuations, harmonic distortion, reverse power flows, protection coordination issues, and reduced system inertia, particularly in distribution systems where local renewable generation may exceed consumer demand [16–18]. These conditions increase the variability of system operating states and complicate the distinction between normal renewable-induced fluctuations and genuine anomalous events.

The simultaneous integration of advanced digital technologies—including Internet of Things (IoT) devices, cloud computing, artificial intelligence (AI), edge computing, and multi-agent control systems—has further increased the complexity of smart grid operation [19]. Although these technologies significantly improve system observability, forecasting accuracy, automation, and operational efficiency, they also create additional cyber–physical dependencies and expand the range of potential failure mechanisms. Consequently, disturbances may no longer originate exclusively from physical infrastructure but can also emerge from communication failures, software faults, cybersecurity incidents, or interactions among distributed intelligent components.

From a sustainability perspective, maintaining secure and reliable operation under high renewable penetration is essential for achieving long-term decarbonization objectives. However, accomplishing this objective requires monitoring strategies capable of adapting to highly dynamic operating conditions while simultaneously considering both physical and cyber dimensions of the grid. Therefore, anomaly detection should no longer be viewed as an isolated monitoring function, but rather as a core capability supporting resilient, adaptive, and self-healing smart grid operation. This evolving operational context provides the foundation for the increasing adoption of machine learning techniques and motivates the development of integrated resilience-oriented architectures such as the Digital-Twin-Enabled Resilience Framework proposed in this study.

2.2. Sources of Anomalies in Cyber–Physical Smart Grids

Modern smart grids operate as complex cyber–physical systems (CPS), where physical electrical infrastructure is continuously interconnected with communication networks, sensing technologies, data analytics platforms, and automated control systems [5,20]. This close interaction enables advanced monitoring and control capabilities but also creates multiple pathways through which abnormal operating conditions can emerge and propagate across the grid. Consequently, anomaly detection must simultaneously consider disturbances originating from both physical processes and cyber infrastructure.

Physical anomalies encompass a wide range of events, including equipment failures, transformer degradation, transmission and distribution line faults, voltage instability, frequency deviations, and renewable generation intermittency. In smart grids with high solar PV penetration, rapid fluctuations in solar irradiance, changing weather conditions, and dynamic load variations can produce operating states that differ significantly from traditional power system behavior [3,21]. These conditions complicate the distinction between normal

renewable-induced variability and genuinely anomalous events, increasing the need for intelligent monitoring techniques capable of interpreting complex operational patterns.

Cyber-related anomalies originate from failures or malicious activities affecting the digital infrastructure that supports grid operation. Representative threats include false data injection (FDI) attacks, denial-of-service (DoS) attacks, malware infiltration, communication failures, software vulnerabilities, and unauthorized access to critical infrastructure [4,22]. Owing to the strong interdependence between cyber and physical components, disturbances initiated within communication or information systems can rapidly propagate to the physical layer, potentially triggering cascading failures, incorrect control actions, equipment damage, or widespread service interruptions.

Beyond isolated physical or cyber events, modern smart grids are increasingly exposed to hybrid cyber–physical anomalies in which failures occurring across multiple layers interact simultaneously. For example, a communication disruption may compromise measurement integrity during periods of high renewable generation, while a cyber-attack targeting supervisory control systems may amplify the operational impact of naturally occurring voltage or frequency disturbances. These coupled events substantially increase operational uncertainty and challenge conventional monitoring approaches that analyze physical and cyber domains independently.

The growing operational complexity of renewable-rich smart grids therefore requires monitoring systems capable of detecting, contextualizing, and prioritizing anomalies across heterogeneous data sources in real time. Rather than simply identifying abnormal measurements, future monitoring architectures must support situational awareness by relating detected anomalies to the current operational state of the grid and their potential impact on system resilience.

The limitations of conventional rule-based and model-driven monitoring techniques have consequently accelerated the adoption of data-driven approaches capable of identifying complex, nonlinear, and evolving patterns within large-scale power system datasets. Among these approaches, machine learning has emerged as one of the most promising technologies due to its ability to learn representative operational behavior directly from historical and real-time data while adapting to continuously changing grid conditions. The following subsections review the principal machine learning paradigms currently applied to anomaly detection and discuss their respective strengths, limitations, and applicability to resilient cyber–physical smart grid operation.

2.3. Machine Learning Approaches for Anomaly Detection

Conventional anomaly detection in power systems has traditionally relied on threshold-based alarms, statistical process control, state-estimation residual analysis, and physics-based monitoring techniques. Although these methods remain effective for identifying well-defined faults under relatively stable operating conditions, they often struggle to detect subtle, nonlinear, and evolving anomalies in highly dynamic smart grid environments characterized by distributed generation, renewable intermittency, and heterogeneous data sources [9].

The increasing availability of high-resolution measurements from phasor measurement units (PMUs), supervisory control and data acquisition (SCADA) systems, smart meters, and distributed sensing infrastructures has accelerated the adoption of machine learning (ML) for anomaly detection. Unlike conventional rule-based approaches, ML algorithms learn representative operational patterns directly from historical and real-time data, enabling them to adapt to changing system conditions and identify anomalies that would otherwise remain undetected [23]. Depending on data availability and operational requirements, current research has explored supervised, unsupervised, semi-supervised, and deep

learning paradigms, each offering distinct advantages and limitations for cyber–physical smart grid applications.

2.3.1. Supervised Learning

Supervised learning algorithms are trained using labeled datasets containing examples of both normal and anomalous operating conditions. Representative techniques include Support Vector Machines (SVM), Random Forests (RF), Artificial Neural Networks (ANNs), and Extreme Gradient Boosting (XGBoost) [24]. These methods have demonstrated strong performance in detecting known cyber-attacks, equipment faults, and abnormal load behaviors, particularly when comprehensive and representative training datasets are available.

The primary strength of supervised learning lies in its ability to achieve high classification accuracy for previously observed anomaly classes. However, its performance depends heavily on the availability of labeled datasets, which are often scarce in operational power systems due to the rarity of anomalous events and the high cost of expert annotation. Consequently, supervised models generally exhibit limited capability to recognize previously unseen anomalies or novel attack strategies [25].

2.3.2. Unsupervised Learning

Unsupervised learning techniques detect anomalies without requiring labeled training data. Instead, these methods characterize the normal operating behavior of the system and identify observations that deviate significantly from learned patterns [26]. This characteristic makes them particularly attractive for operational environments where anomaly labels are incomplete, unavailable, or continuously evolving.

Representative unsupervised techniques include K-Means clustering, Isolation Forests, and autoencoders. Isolation Forests identify anomalous observations by recursively partitioning the feature space, whereas autoencoders learn compact representations of normal operating conditions and detect anomalies through elevated reconstruction errors. Because these approaches do not rely on predefined anomaly categories, they are well suited for identifying emerging faults, unknown attack patterns, and unexpected operational disturbances in renewable-rich smart grids.

2.3.3. Semi-Supervised Learning

Semi-supervised learning combines a relatively small amount of labeled data with substantially larger collections of unlabeled observations. In smart grid applications, these models are commonly trained using normal operational data and subsequently identify anomalous conditions as deviations from the learned representation of normal system behavior [25].

Since abnormal events occur infrequently during routine grid operation, semi-supervised learning represents an effective compromise between supervised and unsupervised approaches. Recent studies have reported promising results in equipment fault diagnosis, intrusion detection, and cyber–physical anomaly detection, particularly in scenarios where obtaining comprehensive labeled datasets is impractical.

2.3.4. Deep Learning Architectures

The temporal, spatial, and topological characteristics of modern smart grid data have driven the increasing adoption of deep learning architectures capable of extracting hierarchical feature representations from large-scale datasets [11]. These models are particularly effective in capturing complex nonlinear relationships that cannot easily be represented using conventional machine learning algorithms.

Long Short-Term Memory (LSTM) networks are extensively used for modeling sequential measurements such as voltage, frequency, and power flow time series. Convo-

lutional Neural Networks (CNNs) efficiently capture spatial correlations within multidimensional monitoring data, while Graph Neural Networks (GNNs) explicitly exploit grid topology to identify localized disturbances and cascading failures across interconnected electrical networks.

Although deep learning models frequently outperform traditional machine learning techniques in complex anomaly detection scenarios—including coordinated cyber-attacks, renewable-induced instabilities, and large-scale cyber–physical disturbances—their practical deployment remains constrained by substantial computational requirements, increased training data demands, and limited model interpretability [27]. These limitations have motivated growing interest in hybrid approaches that balance predictive performance with operational transparency.

Table 1 summarizes representative machine learning approaches currently applied to anomaly detection in smart grids. The comparison illustrates that no single technique simultaneously maximizes predictive performance, interpretability, scalability, computational efficiency, and robustness. While deep learning architectures such as LSTM networks and GNNs generally achieve superior performance in complex cyber–physical environments, traditional machine learning algorithms remain attractive because of their lower computational requirements, faster deployment, and greater interpretability. More recently, research has increasingly shifted toward hybrid, physics-informed, and explainable machine learning approaches that seek to combine the predictive capabilities of advanced learning algorithms with domain knowledge and operational transparency.

Table 1. Comparison of Representative ML Approaches for Smart Grid Anomaly Detection.

Approach	Application	Advantages	Limitations
Random Forest	Fault detection	High accuracy and good interpretability	Requires labeled datasets and limited generalization to unseen events
SVM	Intrusion detection	Effective in high-dimensional feature spaces	Sensitive to parameter tuning and computationally expensive at scale
Isolation Forest	Unknown anomaly detection	Unsupervised and scalable	Limited temporal awareness
Autoencoders	FDI detection	Detects unseen anomalies through reconstruction error	Limited explainability
LSTM	Time-series monitoring	Captures temporal dependencies effectively	High computational cost
CNN	Pattern recognition	Automatic feature extraction	Requires large datasets
GNN	Topology-aware detection	Incorporates grid connectivity information	Complex implementation
Physics-Informed Models	Cyber–physical anomaly detection	Improved robustness and physical consistency	Requires domain expertise

Overall, the literature demonstrates that machine learning has significantly advanced anomaly detection capabilities in smart grids. Nevertheless, existing approaches predominantly optimize the analytical performance of individual models rather than addressing how these models should be integrated into a comprehensive operational architecture capable of supporting real-time decision-making, resilience enhancement, adaptive response, and self-healing grid operation. This observation motivates the need for broader resilience-oriented frameworks that combine complementary machine learning techniques with contextual system representation and operational decision support.

2.4. Challenges and Research Gaps

Despite the substantial progress achieved in machine-learning-based anomaly detection, several technical and operational challenges continue to hinder its large-scale adoption in real-world smart grids. These challenges extend beyond predictive accuracy and include data quality and availability, scalability, real-time processing, interpretability, cybersecurity, model robustness, and integration with operational decision-making processes [9,12,23].

One of the most persistent limitations concerns the availability and representativeness of data. Many anomaly detection models are trained and evaluated using simulated systems, laboratory testbeds, synthetic attack scenarios, or controlled benchmark datasets. Although these resources have supported algorithm development and comparative experimentation, they do not fully reproduce the uncertainty, heterogeneity, evolving operating conditions, and complex cyber–physical interactions observed in operational power systems [9,24]. Moreover, anomalous events are inherently infrequent compared with normal operating conditions, resulting in highly imbalanced datasets. This imbalance complicates model training, reduces generalization capability, and may increase false-positive and false-negative rates during deployment.

Scalability and real-time implementation represent additional barriers. Contemporary smart grids continuously generate large volumes of heterogeneous information from phasor measurement units, SCADA systems, smart meters, intelligent electronic devices, distributed energy resources, photovoltaic installations, and communication infrastructures. Processing these high-dimensional and multimodal data streams with sufficiently low latency requires architectures capable of combining distributed computation, cloud services, edge intelligence, and adaptive data management [1,12,13]. Although recent studies highlight the role of IoT, cloud, and edge computing in strengthening anomaly detection, their coordinated implementation in operational smart grids remains technically demanding.

Model interpretability also constitutes a critical obstacle to practical adoption. Deep learning architectures can identify complex temporal, spatial, and multimodal patterns, but their internal reasoning is frequently difficult for operators to understand [11,27]. In safety-critical infrastructures, an anomaly alert is operationally useful only when decision-makers can understand its probable origin, affected assets, level of confidence, and potential consequences. Consequently, anomaly detection systems must move beyond opaque classification outputs and provide interpretable evidence that supports human supervision and accountable operational action. However, explainability remains insufficiently integrated into end-to-end smart grid monitoring and response architectures.

Cybersecurity introduces a further layer of complexity. Smart grid infrastructures are exposed to false data injection, data manipulation, communication disruption, malware, coordinated cyberattacks, and cascading failures capable of propagating between cyber and physical components [4,6,20,22]. Machine learning models may themselves become attack targets through adversarial inputs, data poisoning, model manipulation, or corrupted training data. Therefore, future anomaly detection systems must protect not only the underlying electrical and communication infrastructures but also the integrity, robustness, and trustworthiness of the analytical models used to support operational decisions.

These challenges are intensified in grids with high solar photovoltaic penetration. Large-scale PV integration introduces voltage variations, reverse power flows, frequency instability, reduced effective inertia, forecasting uncertainty, and increased sensitivity to rapidly changing environmental conditions [16–18]. Although intelligent technologies and machine learning can improve renewable forecasting, grid management, and predictive decision-making, most studies address these functions separately rather than as components of an integrated resilience process [2,3,10,19]. Consequently, abnormal operating

conditions caused by renewable variability may be difficult to distinguish from equipment failures, communication anomalies, or deliberate cyberattacks.

Beyond these technical limitations, the literature reveals a broader architectural gap. Existing research frequently concentrates on individual algorithms, isolated anomaly categories, specific datasets, or narrowly defined operational functions. Machine-learning-based detection, cyber–physical security, renewable integration, grid digitalization, resilience assessment, and autonomous response are commonly developed as parallel research streams, even though they are strongly interdependent in practical smart grid operation [5,14,21]. This fragmentation provides limited guidance regarding how heterogeneous data acquisition, contextual system representation, intelligent analytics, operator interpretation, mitigation, recovery, and continuous learning should be coordinated within a unified operational architecture.

Digital and virtual representations of grid assets provide a promising mechanism for reducing this fragmentation. Digitized grid environments and virtual microgrid representations can improve system visibility, contextual analysis, and service-quality diagnosis [21,28]. Nevertheless, their integration with machine-learning-based anomaly detection, explainable decision support, resilience-oriented control, and closed-loop adaptation remains insufficiently developed. In many studies, anomaly detection continues to be treated as an isolated analytical function rather than as one stage of a broader cyber–physical resilience cycle involving detection, interpretation, response, recovery, and system adaptation.

Recent studies have begun to reduce this fragmentation by combining virtual system representations, cyber–physical monitoring, machine-learning analytics, and resilience mechanisms [5,14,21,28]. These developments represent an important transition toward more integrated smart grid architectures. However, the reviewed literature also shows that these capabilities are generally addressed with different scopes and levels of integration, particularly regarding their connection with anomaly detection, explainable decision support, adaptive operational response, recovery, and continuous learning [14,21,28].

Accordingly, the contribution of DTERF does not lie in introducing Digital Twins, machine learning, or explainable artificial intelligence individually, nor in claiming their first combined use for anomaly detection. Its contribution is the organization of these complementary capabilities within a resilience-oriented reference architecture specifically conceived for high-PV cyber–physical smart grids. DTERF therefore emphasizes the end-to-end coordination of heterogeneous data acquisition, Digital Twin-based contextualization, anomaly analytics, explainable decision support, adaptive response, recovery, and operational feedback.

Table 2 synthesizes the principal research directions identified in the reviewed literature and compares their dominant contributions with the integrated capabilities required for resilient anomaly management in high-PV cyber–physical smart grids. The comparison does not imply that the cited studies pursue identical objectives. Instead, it illustrates how the relevant technological capabilities remain distributed across separate bodies of research.

As shown in Table 2, the reviewed literature provides substantial contributions within each individual research direction. Machine-learning studies offer increasingly sophisticated detection mechanisms; cybersecurity research characterizes threats and attack propagation; renewable-energy studies examine the operational consequences of high PV penetration; digitalization initiatives improve system representation; and resilience research emphasizes the capacity to withstand, recover from, and adapt to disruptive events. However, these capabilities remain unevenly represented across the reviewed research directions, particularly when considering their coordination within an end-to-end resilience cycle for high-PV cyber–physical smart grids.

Table 2. Comparison of representative research directions and their coverage of capabilities required for resilient anomaly management in high-PV cyber–physical smart grids.

Research Direction	ML Analytics	Cyber–Physical Context	High-PV Focus	Virtual Representation	Resilience Focus	Closed-Loop Response	References
Machine-learning anomaly detection	✓	Partial	Partial	–	Partial	–	[9,23,25,26]
Deep and spatiotemporal learning	✓	Partial	Partial	–	Partial	–	[11,27]
Cyberattack detection and smart grid security	✓	✓	–	–	Partial	Partial	[4,6,20,22,24]
Renewable integration and high-PV operation	Partial	Partial	✓	–	Partial	Partial	[2,3,16–19]
Grid digitalization and virtual representation	Partial	✓	Partial	✓	Partial	Partial	[1,13,21]
Cyber–physical resilience	Partial	✓	Partial	Partial	✓	Partial	[5,14]
Proposed DTERF	✓	✓	✓	✓	✓	✓	This work

A further limitation concerns the lack of standardized evaluation methodologies, representative benchmark datasets, and common architectural reference models. This situation restricts reproducibility, complicates objective comparison across studies, and slows the transfer of research results into operational smart grid environments [9,23,24]. Future research must therefore move beyond isolated improvements in detection accuracy and address the coordination of data acquisition, cyber–physical context, intelligent analytics, explainable decision support, adaptive intervention, recovery, and continuous system feedback.

The challenges and gaps identified in this review motivate the Digital-Twin-Enabled Resilience Framework (DTERF) proposed in this study. DTERF does not introduce another anomaly detection algorithm. Instead, it addresses the architectural fragmentation observed in the literature by integrating heterogeneous cyber–physical data, Digital Twin-based contextual representation, machine-learning analytics, explainable artificial intelligence, resilience-oriented decision support, adaptive response, self-healing, and continuous learning within a unified operational cycle. This integrated perspective is intended to bridge the gap between advances in anomaly detection research and the practical requirements of resilient, sustainable, and operationally deployable high-PV smart grids.

3. Research Methodology

This study adopts a design-oriented conceptual research methodology to develop the Digital-Twin-Enabled Resilience Framework (DTERF) for machine-learning-based anomaly detection in high-PV cyber–physical smart grids. Rather than proposing, implementing, or benchmarking a specific machine learning algorithm, the objective is to synthesize current knowledge in anomaly detection and resilience engineering into a unified conceptual architecture that supports intelligent, explainable, and adaptive smart grid operation.

The methodological approach combines a structured literature review with comparative analysis and conceptual framework design. Existing studies were systematically analyzed to identify current capabilities, implementation challenges, and architectural limitations associated with machine-learning-based anomaly detection in renewable-rich cyber–physical power systems. These findings were subsequently integrated into the design of DTERF, a resilience-oriented framework intended to bridge the gap between individual analytical techniques and their operational deployment in modern smart grids.

The research process consists of four sequential stages: (i) literature identification and knowledge synthesis, (ii) comparative assessment of machine learning approaches,

(iii) DTERF development, and (iv) qualitative conceptual evaluation. Figure 1 summarizes the methodological workflow adopted in this study.

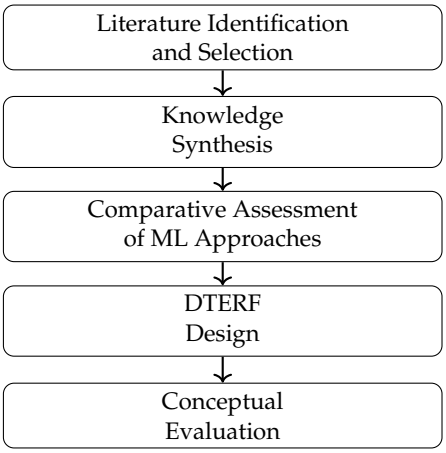

Figure 1. Research methodology adopted for the development of the Digital-Twin-Enabled Resilience Framework (DTERF).

3.1. Literature Identification and Knowledge Synthesis

The first stage consisted of a structured review of the scientific literature addressing anomaly detection in cyber–physical smart grids. The objective was not to produce a bibliometric analysis, but rather to identify the principal technological developments, architectural trends, implementation challenges, and research gaps that characterize the current state of the field.

The review focused on peer-reviewed journal articles and conference proceedings covering one or more of the following research topics:

- Machine learning and deep learning for anomaly detection;
- Cyber–physical security and resilient smart grid operation;
- High solar photovoltaic (PV) penetration and inverter-based resources;
- Digital Twin technologies for power systems;
- Explainable Artificial Intelligence (XAI), physics-informed learning, and resilience engineering.

The selected studies were examined to identify common architectural components, machine learning paradigms, operational capabilities, reported benefits, implementation limitations, and emerging research directions. Rather than evaluating individual algorithms in isolation, the review emphasized how these complementary research areas collectively contribute to the development of integrated anomaly detection architectures. The resulting knowledge synthesis established the conceptual foundation for the design of DTERF.

The literature search focused primarily on recent peer-reviewed studies published between 2023 and 2025 (ref. Table 2), reflecting the rapid evolution of machine learning, Digital Twin technologies, explainable artificial intelligence, and cyber–physical smart grids. Earlier publications were also retained when they provided foundational concepts, widely adopted methodologies, or essential background on smart grid operation, cyber–physical systems, and anomaly detection. Rather than aiming for exhaustive bibliographic coverage, the review prioritized studies that contributed to the analytical, architectural, resilience, and operational dimensions required for the development of DTERF.

3.2. Comparative Assessment of Machine Learning Approaches

Following the literature review, representative machine learning approaches were comparatively analyzed according to characteristics that directly influence their applica-

bility to operational smart grids. The comparison considered each technique's learning paradigm, data requirements, scalability, interpretability, computational complexity, operational maturity, and suitability for cyber–physical anomaly detection.

The assessment included four broad categories:

- Supervised learning methods;
- Unsupervised learning methods;
- Deep learning architectures;
- Hybrid and physics-informed approaches.

Rather than identifying a single superior technique, this comparative analysis sought to determine how the complementary strengths of different learning paradigms could be integrated within a unified resilience-oriented architecture. The results therefore served as architectural design inputs rather than as a conventional algorithmic performance comparison.

3.3. Development of the Digital-Twin-Enabled Resilience Framework (DTERF)

The third stage consisted of developing the Digital-Twin-Enabled Resilience Framework (DTERF). The framework was derived directly from the knowledge synthesized during the literature review and the comparative assessment of existing machine learning approaches.

Instead of defining a fixed analytical pipeline, DTERF was conceived as a conceptual reference architecture that coordinates heterogeneous data acquisition, digital-twin representation, feature engineering, machine learning analytics, explainable decision support, adaptive operational response, continuous monitoring, and self-healing mechanisms within a single resilience-oriented operational cycle.

The framework was specifically designed to address the principal limitations identified throughout the literature, including fragmented analytical workflows, heterogeneous data sources, limited model interpretability, cybersecurity concerns, scalability requirements, and the increasing operational complexity introduced by high PV penetration. By integrating these complementary dimensions, DTERF provides a conceptual foundation for resilient anomaly detection in renewable-rich cyber–physical smart grids.

3.4. Conceptual Evaluation

Because DTERF is proposed as a conceptual framework rather than an implemented software platform, its evaluation is qualitative rather than experimental. The objective of this stage is therefore to assess whether the proposed architecture addresses the principal technical and operational requirements consistently identified throughout the literature.

The conceptual evaluation was conducted through a requirements-to-architecture traceability assessment. First, the technical and operational requirements identified through the literature review in Section 2 were consolidated into seven evaluation dimensions: anomaly detection capability, scalability, interpretability, cybersecurity robustness, resilience support, operational applicability, and architectural integration. Second, each dimension was mapped to the functional components of DTERF described in Section 4, examining whether the architecture provides an explicit mechanism for addressing the corresponding requirement. Third, the consistency of these mechanisms was examined across the complete operational sequence, from heterogeneous data acquisition and Digital Twin contextualization to anomaly detection, explainable decision support, adaptive response, and operational feedback.

The assessment therefore considers three aspects of architectural soundness: functional coverage, traceability, and internal consistency. Functional coverage examines whether the principal requirements identified in the literature are represented within DTERF. Traceability examines whether these requirements can be associated with specific architectural

components and responsibilities. Internal consistency examines whether the interactions among these components form a coherent closed-loop resilience process.

The scenarios presented in Section 5 are used as conceptual walkthroughs to illustrate how these architectural responsibilities interact under representative cyber, physical, and cyber–physical anomaly conditions. These scenarios are not considered experimental validation and therefore do not provide evidence of quantitative performance or superiority over existing architectures.

Accordingly, the conceptual evaluation supports conclusions regarding the functional coverage, traceability, and internal coherence of DTERF. It does not establish detection accuracy, computational performance, real-time feasibility, or operational superiority. These properties require future implementation and empirical validation using representative grid datasets, test systems, and Digital Twin simulation environments.

4. Digital-Twin-Enabled Resilience Framework (DTERF)

The literature reviewed in the previous sections demonstrates that machine learning has substantially improved anomaly detection capabilities in cyber–physical smart grids. However, most existing studies focus on individual analytical components—such as anomaly detection algorithms, cybersecurity mechanisms, digital twins, explainable artificial intelligence, or resilience strategies—rather than on how these technologies can be integrated into a coherent operational architecture. As a result, current solutions often provide high analytical performance while offering limited guidance for their deployment within the increasingly complex operational environment of renewable-rich power systems.

This limitation becomes particularly evident in smart grids with high solar photovoltaic (PV) penetration, where renewable intermittency, reduced system inertia, heterogeneous sensing infrastructures, and cyber–physical interactions continuously reshape system behavior. Under these conditions, effective anomaly detection requires considerably more than accurate prediction models. It demands an architecture capable of continuously integrating heterogeneous information, maintaining situational awareness, supporting explainable operational decisions, and coordinating adaptive response mechanisms that enhance overall grid resilience.

To address these challenges, this study proposes the Digital-Twin-Enabled Resilience Framework (DTERF), a conceptual reference architecture that integrates digital-twin technology, machine learning, explainable artificial intelligence (XAI), resilience engineering, and self-healing concepts into a unified operational framework. Rather than introducing a new anomaly detection algorithm, DTERF defines how complementary technologies can be orchestrated to support intelligent monitoring, contextual anomaly interpretation, adaptive decision-making, and continuous operational improvement within renewable-rich cyber–physical smart grids.

The central premise of DTERF is that anomaly detection should be viewed as a continuous resilience process rather than as an isolated analytical task. Within this process, a Digital Twin maintains a synchronized virtual representation of the physical grid by continuously integrating operational measurements, environmental conditions, asset information, and cybersecurity events. Machine learning models analyze this contextualized representation to identify anomalous operating conditions, while explainable decision-support mechanisms provide interpretable recommendations that assist grid operators in selecting appropriate mitigation strategies. The outcomes of these operational decisions subsequently feed back into the Digital Twin, allowing the framework to continuously refine its knowledge and support system resilience through an adaptive learning cycle.

Figure 2 illustrates the overall conceptual architecture of the proposed Digital-Twin-Enabled Resilience Framework.

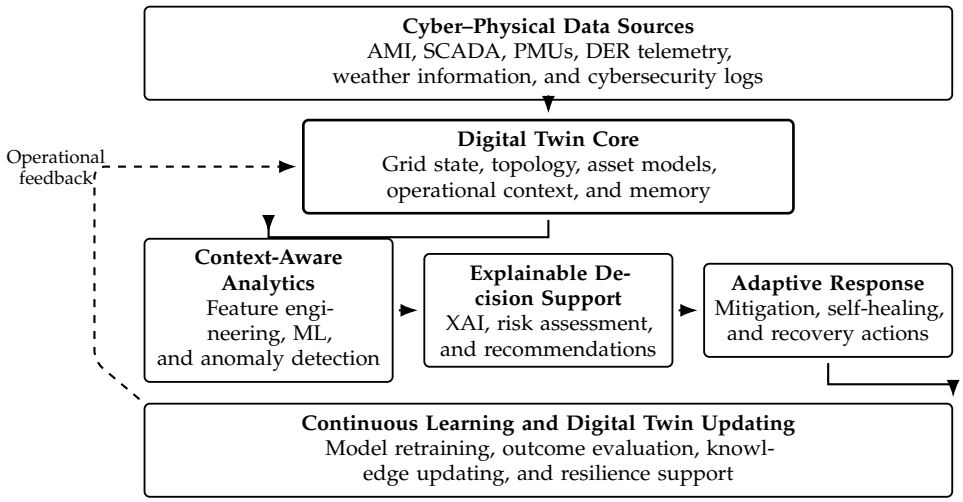

Figure 2. Closed-loop conceptual architecture of the proposed Digital-Twin-Enabled Resilience Framework (DTERF).

As illustrated in Figure 2, DTERF is organized as a closed-loop operational architecture that integrates data acquisition, contextual system representation, intelligent analytics, explainable decision support, adaptive response, and continuous learning. Unlike conventional anomaly detection pipelines that terminate once an anomaly has been identified, DTERF establishes a continuous operational feedback cycle in which every detected event contributes to updating both the analytical models and the Digital Twin. This feedback mechanism enables the framework to progressively update its representation of the physical system in response to changing operating conditions.

For clarity, Figure 2 groups the analytical workflow into a single context-aware analytics layer, whereas Table 3 describes its internal functional components—including processing and feature engineering, machine-learning analytics, and anomaly detection—in greater detail.

The information and decision flow represented in Figure 2 follows a closed-loop sequence. The Digital Twin provides the analytical layers with a contextualized system state comprising synchronized measurements, grid topology, asset status, environmental conditions, operational constraints, and relevant cybersecurity information. The machine-learning analytics layer processes this representation and produces model-specific outputs such as anomaly scores, predicted classes, confidence indicators, and the identification of potentially affected variables or assets. These outputs, together with the corresponding Digital Twin context, are passed to the explainable decision-support layer, which translates them into operator-oriented information concerning the probable origin, severity, affected components, confidence, and potential operational consequences of the detected anomaly.

Based on this contextualized interpretation, the adaptive-response layer determines the appropriate response according to the event type, severity, confidence, and current grid state. The resulting mitigation or recovery action and its observed operational outcome are subsequently returned to the Digital Twin. This feedback updates the system state and operational memory and provides information for outcome evaluation, model retraining, threshold adjustment, or decision-rule refinement when appropriate. Consequently, information flows forward from system observation to detection, interpretation, and response, while operational outcomes flow backward through the feedback mechanism to support subsequent monitoring and adaptation.

The framework combines operational, electrical, environmental, asset, and cybersecurity information within a unified digital representation of the smart grid. This integrated view allows machine learning algorithms to evaluate anomalies within their operational

context instead of relying solely on isolated measurements. Furthermore, by incorporating explainable artificial intelligence and resilience-oriented operational strategies, DTERF facilitates both automated analytical processing and human-centered decision support, providing operators with contextual information for anomaly interpretation and mitigation decisions.

Another distinguishing characteristic of DTERF is its modular architecture. Each functional layer can incorporate different machine learning paradigms—including supervised, unsupervised, deep learning, hybrid, or physics-informed models—according to the operational characteristics, computational resources, and data availability of a particular deployment scenario. Consequently, the framework is intended as a technology-agnostic architectural reference rather than as a platform dependent on any specific learning algorithm.

The principal functional layers that compose DTERF are summarized in Table 3.

Table 3. Functional Layers of the Digital-Twin-Enabled Resilience Framework (DTERF).

Layer	Primary Function	Representative Components
Cyber–Physical Data Sources	Acquisition and synchronization of heterogeneous operational information from the physical and cyber domains.	AMI, SCADA, PMUs, DER telemetry, weather services, cybersecurity logs.
Digital Twin	Dynamic virtual representation of the smart grid that maintains contextual awareness of assets and operating conditions.	Grid state, asset models, topology, operational context.
Processing and Feature Engineering	Preparation, synchronization, and transformation of heterogeneous data into meaningful analytical features.	Data cleaning, normalization, synchronization, feature extraction.
Machine-Learning Analytics	Execution of predictive models for anomaly detection and operational intelligence.	Random Forest, SVM, Isolation Forest, Autoencoders, LSTM, CNN, GNN.
Anomaly Detection	Identification, characterization, and prioritization of abnormal operational conditions.	Alerts, anomaly scores, reconstruction errors, confidence indicators.
Explainable Decision Support	Interpretation of detected anomalies and recommendation of mitigation strategies.	XAI, dashboards, root-cause analysis, risk assessment.
Adaptive Response	Selection and execution of operational actions to mitigate detected anomalies and enhance resilience.	DER coordination, network reconfiguration, protection actions, predictive maintenance.
Continuous Learning	Incorporation of operational feedback to continuously improve analytical models and the Digital Twin.	Model retraining, knowledge updating, resilience optimization.

The following subsections describe each functional layer in detail and explain how their interaction enables DTERF to provide an integrated, explainable, and resilience-oriented architecture for anomaly detection in high-PV cyber–physical smart grids.

4.1. Cyber–Physical Data Layer

The Cyber–Physical Data Layer constitutes the foundation of the Digital-Twin-Enabled Resilience Framework (DTERF), providing the heterogeneous information required to maintain an accurate and continuously updated representation of the smart grid. Unlike conventional monitoring architectures that rely primarily on electrical measurements, DTERF adopts a holistic perspective by integrating operational, environmental, asset, and cybersecurity information originating from multiple interconnected sources. This comprehensive data acquisition strategy enables the framework to capture not only the electrical behavior of the grid but also the contextual factors that influence system resilience and anomaly propagation.

Because high-PV smart grids operate as tightly coupled cyber–physical systems, abnormal conditions frequently emerge from the interaction of multiple domains rather than from a single isolated event. Consequently, DTERF considers heterogeneous data acquisition as a prerequisite for contextual anomaly detection, allowing machine learning models to evaluate system behavior within its complete operational environment instead of analyzing individual measurements independently.

The principal data sources incorporated into the framework include:

- Advanced Metering Infrastructure (AMI): Provides high-resolution measurements of electricity consumption, voltage, current, power quality, and customer demand profiles. These measurements contribute to the detection of abnormal consumption patterns, localized disturbances, and demand-side anomalies.
- Supervisory Control and Data Acquisition (SCADA): Supplies supervisory monitoring information from transmission and distribution networks, including equipment status, operational commands, switching events, alarms, and control signals. SCADA data provide the operational context necessary for understanding system-wide behavior.
- Phasor Measurement Units (PMUs): Deliver synchronized phasor measurements with high temporal resolution, enabling real-time monitoring of voltage angles, frequency variations, oscillatory behavior, and dynamic grid stability. PMUs significantly enhance situational awareness during fast-changing operating conditions.
- Distributed Energy Resources (DERs): Provide operational information from photovoltaic systems, battery energy storage systems, inverter-based resources, electric vehicles, and microgrids. These data are particularly important in high-PV environments, where renewable intermittency introduces substantial operational variability.
- Cybersecurity Monitoring Platforms: Collect network traffic statistics, authentication logs, intrusion detection alerts, communication performance indicators, and other cybersecurity events. Integrating cyber information enables DTERF to identify anomalies originating from malicious activities as well as failures affecting communication infrastructures.
- Environmental and Weather Information: Includes solar irradiance, ambient temperature, wind speed, cloud coverage, and other meteorological variables that directly influence renewable generation and grid operating conditions. Incorporating environmental information provides contextual information for distinguishing expected renewable variability from abnormal system behavior.

The heterogeneous information acquired from these sources is continuously synchronized before being incorporated into the Digital Twin. Rather than treating each data stream independently, DTERF establishes a unified cyber–physical representation in which electrical measurements, operational events, environmental conditions, and cybersecurity observations are analyzed together. This integrated perspective enables richer contextual awareness, facilitating the identification of complex anomalies that may emerge from interactions across multiple system layers.

By providing a comprehensive and continuously updated representation of the operational environment, the Cyber–Physical Data Layer establishes the information foundation upon which the remaining components of DTERF operate. The synchronized data are subsequently transferred to the Digital Twin, where they are contextualized, processed, and transformed into the analytical inputs required for machine learning, explainable decision support, adaptive response, and continuous resilience enhancement.

4.2. Digital Twin and Data Processing Layer

The Digital Twin constitutes the core integration component of DTERF, serving as the virtual representation through which heterogeneous information from the Cyber–Physical

Data Layer is transformed into contextual operational knowledge. Unlike conventional monitoring architectures that treat data repositories as passive storage mechanisms, the Digital Twin continuously synchronizes physical assets, operational conditions, environmental variables, and cybersecurity events to maintain an up-to-date representation of the smart grid throughout its operating lifecycle.

Within DTERF, the Digital Twin fulfills three complementary roles. First, it integrates heterogeneous information originating from multiple sensing and monitoring platforms into a coherent representation of the current system state. Second, it preserves contextual knowledge regarding network topology, asset conditions, renewable generation levels, operational constraints, historical events, and cyber–physical dependencies. Third, it provides a common analytical environment that enables machine learning models to evaluate anomalies within their operational context, thereby maintaining consistency between physical processes and the digital infrastructure that supports grid operation.

By maintaining this synchronized virtual representation, the Digital Twin extends anomaly detection beyond the identification of abnormal measurements. Instead, detected events can be interpreted according to their location within the network, their relationship with surrounding operating conditions, and their potential impact on system resilience. Consequently, the Digital Twin provides the contextual awareness required to support explainable and operationally meaningful anomaly detection.

Before analytical processing, the heterogeneous information acquired from the Cyber–Physical Data Layer undergoes a series of preprocessing operations designed to improve data quality, consistency, and interoperability. Typical preprocessing activities include:

- Data cleaning and validation;
- Missing-value detection and imputation;
- Noise filtering and outlier removal;
- Time synchronization across heterogeneous data sources;
- Data normalization and feature scaling.

Within DTERF, these operations define a common preprocessing sequence for heterogeneous data streams. Incoming measurements are first validated and cleaned, after which missing or corrupted observations are identified and treated. Noise filtering is then applied when required, followed by temporal alignment of measurements originating from sources with different sampling rates. Finally, normalization and feature scaling prepare the synchronized observations for feature engineering and machine-learning analysis. The specific techniques and parameter values used at each stage remain deployment-dependent, since DTERF defines the required processing functions rather than prescribing a particular software implementation.

These preprocessing operations ensure that measurements originating from different sensing technologies can be consistently integrated despite differences in sampling frequency, communication latency, data quality, and measurement precision. Such harmonization is particularly important in renewable-rich smart grids, where heterogeneous sensing infrastructures continuously generate high-volume, multi-rate data streams.

Synchronization within DTERF is therefore based on temporal alignment and contextual association rather than on the assumption that all data sources operate at identical sampling rates. Each observation is associated with its acquisition time, source, and corresponding grid or cyber asset before being incorporated into the Digital Twin. The synchronized observations can then be fused according to their common temporal and operational context, allowing electrical measurements, DER telemetry, environmental variables, asset information, and cybersecurity events to contribute to a consistent representation of the current grid state. The required synchronization interval is not fixed by the framework and

should be selected according to the dynamics, communication constraints, and monitoring requirements of each deployment.

DTERF also supports a distributed edge–cloud processing model. Time-critical and computationally lightweight functions can be executed at the edge, close to the corresponding sensing or control infrastructure. These functions may include initial data validation, filtering, normalization, feature extraction, and lightweight anomaly screening. The cloud or centralized computational environment maintains the global Digital Twin representation and supports functions requiring broader system context or greater computational capacity, including multi-source data fusion, global-state analysis, computationally intensive machine-learning models, historical analysis, model training, and Digital Twin-based simulation. Only the information required by the global analytical process needs to be propagated from the edge, reducing unnecessary communication while preserving system-level situational awareness.

The allocation of functions between edge and cloud is not fixed by DTERF. Instead, it depends on latency requirements, computational resources, communication availability, data sensitivity, and the spatial scope of the corresponding analytical task. This separation allows latency-sensitive monitoring to remain close to physical assets while the global Digital Twin retains the system-wide context required for coordinated anomaly interpretation, decision support, and resilience analysis.

Following preprocessing, feature engineering transforms the synchronized data into representative indicators that capture both electrical system dynamics and cyber–physical operating conditions. Rather than relying solely on raw measurements, DTERF derives higher-level features that provide richer information for machine learning models while preserving their operational meaning.

Representative electrical features include voltage deviations, frequency fluctuations, harmonic distortion, active and reactive power variations, load ramps, voltage stability indicators, and statistical descriptors extracted from time-series measurements. In parallel, cyber-related features include communication latency, abnormal packet rates, authentication failures, network traffic statistics, intrusion detection alerts, communication reliability metrics, and network integrity indicators. Environmental variables, including solar irradiance, temperature, wind conditions, and weather forecasts, are also incorporated to contextualize renewable generation variability and distinguish expected operational fluctuations from anomalous system behavior.

The integration of domain knowledge during feature engineering constitutes an important characteristic of DTERF. Rather than allowing machine learning algorithms to operate exclusively on generic statistical features, the framework incorporates indicators that reflect the physical behavior of electrical networks, renewable generation characteristics, and cyber–physical interactions. This approach is intended to support operationally meaningful analysis by ensuring that extracted features maintain a clear relationship with real-world grid behavior.

By combining a continuously synchronized Digital Twin with context-aware preprocessing and domain-informed feature engineering, DTERF provides machine learning models with substantially richer analytical inputs than conventional anomaly detection pipelines. The resulting contextual representation provides an information basis for distinguishing normal renewable variability from genuine anomalies, interpreting detected events within their operational context, and supporting explainable decision support, adaptive response, and resilience-oriented smart grid operation.

4.3. Machine-Learning Analytics and Anomaly Detection

The Machine-Learning Analytics and Anomaly Detection layer transforms the contextualized information provided by the Digital Twin into actionable indicators of abnormal system behavior. Rather than prescribing a single algorithm, DTERF supports different learning paradigms according to data availability, anomaly characteristics, computational constraints, and operational requirements. Supervised models can support the recognition of previously characterized faults or attacks, whereas unsupervised and semi-supervised approaches can identify deviations from normal operating patterns. Deep learning and physics-informed models can further capture temporal, spatial, topological, and physical dependencies in complex grid environments.

The outputs of this layer include anomaly scores, predicted classes, confidence indicators, affected variables or assets, and temporal or spatial patterns associated with the detected event. Multiple analytical models may operate in parallel when complementary perspectives are required. In this configuration, each model preserves its model-specific output, which is associated with the same contextual state maintained by the Digital Twin. The collaborative logic then consists of comparing these outputs according to their anomaly indication, confidence, affected assets, and temporal or spatial consistency. Agreement among complementary models can strengthen the evidence associated with a detected event, whereas conflicting outputs are retained as uncertainty and forwarded for contextual interpretation rather than being automatically resolved.

DTERF does not prescribe a fixed fusion algorithm because the appropriate aggregation mechanism depends on the models, output types, and operational requirements of each deployment. When aggregation is appropriate, implementations may use decision-level mechanisms such as confidence-weighted combination, voting, or rule-based consensus. Alternatively, heterogeneous model outputs can remain separate and be jointly contextualized by the Explainable Decision Support layer against the current Digital Twin state. This approach allows supervised, unsupervised, deep-learning, and physics-informed models to contribute complementary analytical evidence without assuming that their outputs are directly interchangeable, while preserving the technology-agnostic character of the framework.

Because the machine-learning pipeline may itself become a target of cyberattacks, DTERF incorporates model-security responsibilities across the analytical lifecycle. At the data-input stage, provenance verification, integrity checks, access control, and consistency validation can be applied before observations are incorporated into model training or inference. These mechanisms are intended to reduce the risk of corrupted measurements, unauthorized data modification, and data-poisoning attacks affecting the analytical process.

At the model level, DTERF considers controlled model versioning, integrity verification, authenticated model deployment, and monitoring for unexpected changes in model behavior. During inference, anomalous or low-confidence inputs can be cross-checked against the physical and operational context maintained by the Digital Twin before their outputs are used for operational decision support. Similarly, model retraining within the Continuous Learning mechanism should only use validated observations and confirmed operational outcomes rather than automatically incorporating every newly observed event.

These controls do not constitute a specific cybersecurity implementation; instead, they define security responsibilities that should accompany the machine-learning lifecycle within DTERF. Their objective is to preserve data provenance, model integrity, and analytical trustworthiness while reducing exposure to adversarial inputs, data poisoning, unauthorized model manipulation, and contamination of the continuous-learning process.

4.4. Explainable Decision Support

The Explainable Decision Support layer connects anomaly detection with operational decision-making. Its primary responsibility is to transform model outputs into information that can be interpreted and acted upon by grid operators. Detected anomalies are evaluated together with the contextual information maintained by the Digital Twin, including grid topology, asset status, operating conditions, renewable generation levels, environmental variables, and relevant cybersecurity events.

Depending on the underlying analytical model, the Explainable Decision Support layer can incorporate complementary XAI techniques. Feature-attribution methods such as SHAP can quantify the contribution of individual electrical, environmental, or cybersecurity variables to a detected anomaly, while local explanation methods such as LIME can provide case-specific interpretations of individual predictions [29]. Model-specific mechanisms can also be used when appropriate, including feature-importance measures for tree-based models, reconstruction-error analysis for autoencoders, and attention or saliency-based interpretations for deep learning architectures. Rule-based explanations can further translate analytical outputs into operational conditions that are directly understandable by grid operators [30].

Rather than exposing operators to raw model outputs, these explanations are organized into dispatcher-oriented interpretation dimensions: (i) probable anomaly type and origin; (ii) affected assets, measurements, or network areas; (iii) variables contributing most strongly to the detection; (iv) detection confidence and anomaly severity; (v) relevant temporal, spatial, or cyber–physical context; (vi) potential operational consequences, including possible propagation or cascading effects; and (vii) urgency and recommended mitigation or verification actions. The Digital Twin provides the contextual information required to relate these explanations to the current grid state. This layer therefore does not replace operator judgment; rather, it provides contextualized evidence and risk-oriented recommendations that support transparent and accountable decision-making before an operational response is selected.

4.5. Adaptive Response and Continuous Learning

The Adaptive Response layer translates the interpreted anomaly into mitigation, recovery, or resilience-enhancing actions. Depending on the type, severity, and operational context of the detected event, these actions may include operator alerts, measurement validation, DER coordination, network reconfiguration, protection actions, communication isolation, or predictive maintenance. DTERF does not prescribe a universal autonomous control policy; the degree of automation depends on the operational criticality, confidence of the analytical output, and control policies of the specific deployment environment.

Following the response, operational outcomes are returned to the Digital Twin through the Continuous Learning mechanism. This feedback includes the observed system response, effectiveness of the selected action, subsequent system state, and eventual confirmation or rejection of the original anomaly assessment. Such information updates the contextual memory of the Digital Twin and, once its integrity and operational validity have been verified, can support model retraining, threshold adjustment, or refinement of decision rules when appropriate. This validation step prevents the Continuous Learning mechanism from automatically incorporating unverified observations or potentially manipulated feedback into subsequent model updates.

5. Illustrative Application Scenarios

The Digital-Twin-Enabled Resilience Framework (DTERF) is intended as a conceptual operational architecture rather than as a framework designed for a single anomaly detection

task. Consequently, its applicability extends across multiple cyber–physical scenarios that characterize modern renewable-rich smart grids. To illustrate its potential operational use, this section presents representative application scenarios frequently discussed in the literature. These examples are not intended as experimental validation of DTERF; instead, they illustrate how the proposed architecture integrates heterogeneous data, machine learning analytics, Digital Twin technology, explainable decision support, and resilience-oriented response mechanisms to address different classes of anomalies.

Collectively, the scenarios illustrate the versatility of the framework in supporting cybersecurity, operational reliability, asset management, and cyber–physical resilience under conditions of high solar PV penetration.

5.1. Scenario 1: False Data Injection Attack Detection

False Data Injection (FDI) attacks represent one of the most critical cybersecurity threats affecting modern smart grids. By manipulating measurements collected from sensors, smart meters, or supervisory control systems, attackers attempt to corrupt state estimation processes, mislead operational decisions, and potentially trigger inappropriate control actions without immediately revealing the attack.

Within DTERF, measurements obtained from AMI, SCADA systems, PMUs, and cybersecurity monitoring platforms are continuously synchronized within the Digital Twin, enabling machine learning models to analyze measurement consistency in the context of the current operating state of the grid. Algorithms such as Autoencoders, Isolation Forests, Support Vector Machines, and Long Short-Term Memory (LSTM) networks can identify abnormal temporal and statistical patterns that deviate from expected operational behavior.

Rather than relying solely on anomaly scores, the Digital Twin provides contextual information regarding affected assets, network topology, renewable generation conditions, and communication events. The Explainable Decision Support layer subsequently evaluates the severity and probable origin of the anomaly, allowing operators to distinguish between measurement errors, equipment malfunctions, and malicious cyberattacks. Based on this assessment, mitigation strategies—including measurement validation, communication isolation, state estimation verification, or operator alerts—can be initiated before the integrity of grid operation is compromised.

This scenario illustrates how DTERF can support cyber resilience by connecting anomaly detection with contextual interpretation and coordinated operational response.

5.2. Scenario 2: Renewable-Induced Operational Instability

The rapid integration of solar photovoltaic (PV) generation introduces significant operational variability due to fluctuating irradiance, cloud transients, inverter dynamics, and reduced system inertia. These factors can produce voltage deviations, frequency fluctuations, reverse power flows, and reactive power imbalances that are difficult to distinguish from abnormal operating conditions using conventional monitoring approaches.

Within DTERF, the Digital Twin continuously integrates electrical measurements with environmental information such as solar irradiance, temperature, weather forecasts, and distributed energy resource operating conditions. Machine learning models—including Convolutional Neural Networks (CNNs), LSTM networks, and Graph Neural Networks (GNNs)—analyze temporal, spatial, and topological relationships to differentiate expected renewable variability from genuinely anomalous system behavior.

When abnormal operating conditions are detected, the Explainable Decision Support layer evaluates their probable causes and potential operational impact. Depending on the severity of the event, the Adaptive Response layer may recommend corrective actions

such as reactive power compensation, inverter parameter adjustment, demand response activation, battery energy storage coordination, or network reconfiguration.

Within this conceptual scenario, incorporating environmental context into anomaly detection is intended to improve the distinction between expected renewable variability and genuinely anomalous behavior, thereby providing an architectural basis for reducing false alarms. However, the extent of such improvement requires quantitative validation under representative stochastic PV fluctuation scenarios.

5.3. Scenario 3: Predictive Equipment Monitoring

Asset degradation remains one of the principal contributors to operational failures in modern power systems. Critical infrastructure—including transformers, circuit breakers, protection devices, power electronic converters, and inverter-based resources—typically exhibits measurable changes in operational behavior long before complete failure occurs.

Within DTERF, historical operating records and real-time sensor measurements are continuously incorporated into the Digital Twin, allowing machine learning models to identify gradual degradation trends that may not be observable through traditional threshold-based monitoring. Algorithms such as Isolation Forests, Random Forests, Autoencoders, and other anomaly detection models evaluate multidimensional asset behavior to identify deviations associated with incipient failures.

Unlike conventional condition monitoring systems, DTERF relates detected anomalies to the operational context of each asset, considering loading conditions, renewable generation levels, environmental variables, and historical maintenance information. The Explainable Decision Support layer subsequently estimates the potential operational impact of the detected degradation and recommends appropriate maintenance actions.

Within this conceptual scenario, such predictive maintenance capabilities could support utilities in optimizing maintenance scheduling, reducing unexpected equipment failures, extending asset lifetime, and improving overall grid reliability while minimizing unnecessary maintenance interventions.

The previous scenarios primarily illustrate independent cyber, operational, and asset-related anomalies. However, future renewable-rich smart grids are expected to experience increasingly complex disturbances involving simultaneous interactions between physical processes and cyber infrastructures. Such events motivate the need for integrated cyber–physical anomaly detection strategies.

5.4. Scenario 4: Coordinated Cyber–Physical Anomaly Detection in High-PV Distribution Networks

Future smart grids will increasingly face coordinated cyber–physical disturbances in which cyberattacks occur simultaneously with legitimate operational variability introduced by renewable generation. Under these conditions, distinguishing malicious behavior from normal fluctuations becomes considerably more challenging than detecting either phenomenon independently.

Within DTERF, heterogeneous information originating from PMUs, SCADA systems, AMI, distributed energy resources, weather services, and cybersecurity monitoring platforms is continuously synchronized within the Digital Twin. Machine learning models capable of capturing temporal, spatial, and topological dependencies—including Graph Neural Networks (GNNs), LSTM networks, and physics-informed learning approaches—evaluate the consistency between observed measurements and the expected operating state of the grid.

Consider, for example, a coordinated False Data Injection attack targeting photovoltaic monitoring devices during a period of rapidly changing solar irradiance. Conventional monitoring systems may interpret manipulated measurements as normal renewable vari-

ability, delaying the identification of the attack. In contrast, DTERF correlates electrical measurements, communication network behavior, environmental conditions, and historical operating patterns through the Digital Twin. This contextual analysis provides a basis for identifying inconsistencies that may be difficult to recognize when individual data sources are analyzed independently.

Once the anomaly has been identified, the Explainable Decision Support layer estimates its severity, identifies the affected assets, evaluates possible cascading effects, and recommends appropriate mitigation strategies. Depending on the operational context, these actions may include validating suspicious measurements using redundant sensors, isolating compromised communication channels, modifying inverter control strategies, reconfiguring distribution feeders, or activating additional protection mechanisms to preserve system stability.

This scenario illustrates the principal contribution of DTERF: the conceptual integration of machine learning, Digital Twin technology, explainable artificial intelligence, and resilience-oriented operational intelligence within a unified architecture designed to support the management of complex cyber–physical disturbances in renewable-rich smart grids.

Overall, these illustrative scenarios show how DTERF is conceptually applicable across a broad spectrum of operational situations involving cybersecurity, renewable integration, predictive maintenance, and coordinated cyber–physical anomaly detection. More importantly, they illustrate how the framework extends beyond anomaly identification by integrating contextual system representation, explainable decision support, adaptive operational response, and continuous learning within a single resilience-oriented architecture. These scenarios therefore serve as conceptual walkthroughs of the proposed architecture and should not be interpreted as evidence of empirical performance or comparative superiority.

6. Discussion and Implementation Challenges

The literature reviewed throughout this study demonstrates that machine learning has become a key enabling technology for anomaly detection in modern cyber–physical smart grids. As electrical networks evolve toward highly distributed, renewable-rich infrastructures, conventional monitoring techniques based on fixed thresholds, deterministic rules, or simplified physical models are increasingly challenged by the growing operational complexity of these systems. Machine learning addresses many of these limitations by enabling adaptive pattern recognition, continuous learning, and the analysis of heterogeneous data generated across the cyber and physical layers of the grid.

However, the analysis also indicates that advances in anomaly detection algorithms alone are insufficient to address the operational requirements of future smart grids. Most existing research concentrates on improving predictive performance through increasingly sophisticated analytical models, while comparatively less attention has been devoted to the architectural integration of machine learning within operational environments. Consequently, the practical value of anomaly detection depends not only on the accuracy of individual algorithms but also on their ability to interact with monitoring systems, decision-support platforms, Digital Twins, and resilience-oriented control mechanisms.

The comparative assessment performed in this review highlights that each machine learning paradigm offers complementary advantages. Supervised learning methods generally achieve high classification accuracy when representative labeled datasets are available, making them well suited for detecting known faults and previously observed cyberattacks. Nevertheless, their dependence on labeled data significantly limits their applicability in operational environments where anomalous events are rare and continuously evolving.

Unsupervised learning approaches, including Isolation Forests and Autoencoders, partially overcome this limitation by identifying deviations from learned normal operating behavior without requiring anomaly labels. These methods are particularly attractive for renewable-rich smart grids, where novel operational conditions and previously unseen disturbances emerge frequently. Nevertheless, because detected anomalies are often represented through abstract anomaly scores or reconstruction errors, their operational interpretation may remain challenging without complementary contextual information.

Deep learning architectures—including Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Graph Neural Networks (GNNs)—have demonstrated remarkable capability for capturing the temporal, spatial, and topological characteristics of modern power systems. Their ability to learn hierarchical feature representations enables the identification of subtle cyber–physical anomalies that may not be detectable using conventional machine learning techniques. However, these improvements are typically accompanied by increased computational complexity, larger training data requirements, longer deployment times, and reduced model transparency.

The growing interest in hybrid and physics-informed machine learning reflects an important evolution in the field. Rather than relying exclusively on statistical learning, these approaches combine data-driven intelligence with domain-specific knowledge describing electrical system behavior. Such integration can improve model robustness, help reduce false alarms, preserve physical consistency, and facilitate operational interpretation. As renewable penetration continues to increase, these characteristics are expected to become increasingly important for supporting reliable decision-making under uncertain operating conditions.

One of the principal observations emerging from this review is that no single machine learning technique consistently outperforms all others across every anomaly detection scenario. The suitability of a particular approach depends on multiple factors, including data availability, anomaly characteristics, grid topology, computational resources, latency requirements, and operational objectives. Consequently, future anomaly detection systems will likely evolve toward hybrid analytical architectures capable of combining complementary learning paradigms within integrated operational frameworks such as the Digital-Twin-Enabled Resilience Framework proposed in this study.

Beyond analytical performance, the literature also indicates an important conceptual transition in smart grid operation. Future power systems are expected not only to detect anomalies but also to understand their operational context, estimate their potential consequences, recommend appropriate mitigation strategies, and continuously improve their analytical capabilities through operational feedback. In this context, machine learning becomes one component of a broader resilience-oriented decision-making process rather than an isolated prediction engine. DTERF reflects this emerging perspective by integrating Digital Twin technology, machine learning, explainable artificial intelligence, adaptive response, and continuous learning within a unified operational architecture.

Despite these advances, several implementation challenges continue to limit the large-scale adoption of machine-learning-enabled anomaly detection systems.

A fundamental limitation concerns the availability of representative datasets. Many publicly available benchmark datasets are generated through simulation or laboratory environments and therefore cannot fully capture the diversity, uncertainty, and operational complexity of real electrical networks. Moreover, high-impact anomalies occur infrequently during normal grid operation, resulting in highly imbalanced datasets that complicate both model training and objective performance evaluation. Improving data availability through standardized datasets, secure data sharing, and realistic Digital Twin environments represents an important direction for future research.

Computational scalability constitutes another important challenge. Modern smart grids generate massive volumes of heterogeneous information from PMUs, SCADA systems, Advanced Metering Infrastructure (AMI), distributed energy resources, intelligent electronic devices, and cybersecurity monitoring platforms. Processing these data streams in real time while maintaining low latency and high detection accuracy requires scalable computational infrastructures capable of supporting distributed analytics, edge computing, and cloud-based processing. Although these technologies continue to mature, their deployment remains associated with significant infrastructure investment and operational complexity.

Model interpretability also remains a major concern. While advanced deep learning architectures frequently provide superior predictive performance, their black-box nature limits operator confidence and complicates decision-making in safety-critical environments. Explainable Artificial Intelligence (XAI) has emerged as a promising mechanism for improving transparency by identifying the factors influencing model predictions and facilitating root-cause analysis. Nevertheless, integrating explainability into operational smart grid environments remains an active research area, particularly when explanations must be generated in real time without compromising analytical performance.

Cybersecurity introduces an additional dimension of complexity because machine learning models themselves may become targets of adversarial attacks, data poisoning, model inversion, or manipulation attempts intended to degrade anomaly detection performance. Consequently, future resilience-oriented architectures must protect not only the physical power system but also the analytical models responsible for supporting operational decisions. This observation reinforces the importance of integrating cybersecurity considerations throughout the complete analytical pipeline rather than treating them as independent system components.

Finally, interoperability with existing utility infrastructures continues to represent a practical deployment challenge. Many transmission and distribution operators continue to rely on legacy SCADA platforms, heterogeneous communication protocols, and equipment that was not originally designed to support advanced machine learning or Digital Twin technologies. Consequently, successful implementation will require modular architectures capable of incremental deployment while maintaining compatibility with existing operational systems and industry standards.

Overall, the findings of this review suggest that the future of anomaly detection in cyber–physical smart grids will depend less on the development of individual machine learning algorithms and more on the integration of complementary technologies into coherent operational architectures. Advances in Digital Twin technology, explainable artificial intelligence, physics-informed learning, distributed analytics, cybersecurity, and autonomous grid control are expected to play a central role in this evolution. Within this context, the proposed Digital-Twin-Enabled Resilience Framework provides a conceptual foundation that integrates these complementary capabilities to support intelligent, explainable, adaptive, and resilience-oriented operation of renewable-rich smart grids.

More specifically, the implementation challenges identified in this review can be directly related to architectural responses within DTERF. Data heterogeneity and quality are addressed through the Cyber–Physical Data Layer and the preprocessing and synchronization functions of the Digital Twin and Data Processing Layer. Scalability and real-time processing are supported through the distributed edge–cloud model, which separates latency-sensitive local processing from computationally intensive system-level analytics. Limited model interpretability is addressed through the Explainable Decision Support layer, which transforms analytical outputs into contextualized, operator-oriented explanations. Cybersecurity and model robustness are addressed through data provenance and integrity

controls, model-security responsibilities, contextual cross-checking against the Digital Twin, and validation of information used for continuous learning. The challenge of connecting anomaly detection with operational action is addressed through the Adaptive Response and Self-Healing mechanisms, which link detection and interpretation with mitigation, recovery verification, and feedback. Finally, interoperability with heterogeneous and legacy infrastructures is supported by the modular and technology-agnostic design of DTERF, which defines functional interfaces and responsibilities without prescribing a specific implementation platform. These architectural responses do not eliminate the corresponding implementation challenges; rather, they specify how DTERF organizes the capabilities required to address them within a coherent resilience-oriented architecture.

Current Limitations and Future Validation

DTERF is currently proposed and evaluated as a conceptual reference architecture, and several limitations should therefore be acknowledged. First, the framework has not yet been implemented or evaluated using real grid datasets, standardized distribution test systems, or Digital Twin simulation environments. Consequently, the present study does not provide empirical evidence regarding detection accuracy, false-alarm rates, computational latency, scalability under different grid sizes, or comparative performance against existing monitoring architectures. The conceptual evaluation presented in Section 3.4 is limited to functional coverage, requirements traceability, and internal architectural consistency.

Second, DTERF intentionally defines functional responsibilities and information flows without prescribing specific machine-learning algorithms, Digital Twin platforms, communication protocols, or deployment configurations. This technology-agnostic design increases the adaptability of the framework but leaves implementation-dependent aspects—including model selection, synchronization frequency, computational resource allocation, communication latency, and response thresholds—to be determined for each operational environment.

Third, the adaptive response, continuous learning, and self-healing mechanisms represent architectural capabilities rather than experimentally validated autonomous control functions. Their deployment in operational power systems would require explicit safety constraints, operator supervision policies, cybersecurity safeguards, and validation of response actions before automated interventions are permitted.

Future validation should therefore proceed incrementally from architecture to implementation. An initial stage should instantiate DTERF in a Digital Twin simulation environment using representative high-PV distribution test systems and anomaly scenarios covering cyber, physical, and cyber–physical disturbances. Subsequent evaluation should quantify detection performance, false-alarm behavior under renewable variability, computational and communication latency, scalability, and recovery effectiveness. Finally, validation using real or utility-derived operational data would be required to assess the practical applicability of the framework under realistic grid conditions. These stages would provide the empirical evidence necessary to complement the conceptual evaluation conducted in this study.

7. Conclusions and Future Research Directions

This study reviewed recent advances in machine-learning-based anomaly detection for cyber–physical smart grids and proposed the Digital-Twin-Enabled Resilience Framework (DTERF) as a conceptual architecture for supporting resilient, explainable, and sustainable smart grid operation under high solar photovoltaic (PV) penetration. Rather than focusing on a specific anomaly detection algorithm, the study adopted an architectural perspective that integrates Digital Twin technology, machine learning, explainable artifi-

cial intelligence, adaptive operational response, and continuous learning within a unified resilience-oriented framework.

The literature review confirms that machine learning has become a fundamental enabling technology for anomaly detection in modern power systems. Supervised, unsupervised, semi-supervised, and deep learning approaches each contribute complementary capabilities, with their suitability depending on data availability, operational objectives, computational constraints, and the characteristics of the anomalies being monitored. In particular, deep learning architectures such as Long Short-Term Memory (LSTM) networks, Convolutional Neural Networks (CNNs), and Graph Neural Networks (GNNs) have demonstrated considerable potential for modeling the temporal, spatial, and topological complexity of renewable-rich cyber–physical power systems. At the same time, hybrid and physics-informed learning approaches have emerged as promising alternatives for improving robustness, physical consistency, and model interpretability.

A central finding of this review is that future advances in smart grid anomaly detection will depend not only on improving the predictive performance of individual machine learning algorithms but also on integrating complementary technologies into coherent operational architectures. Existing research has largely evolved through independent developments in anomaly detection, Digital Twin technology, explainable artificial intelligence, cybersecurity, resilience engineering, and renewable energy integration. Although each of these research areas has achieved significant progress, relatively few studies have explored how they can be systematically combined to support intelligent and adaptive grid operation.

The Digital-Twin-Enabled Resilience Framework proposed in this study addresses this gap by providing a conceptual reference architecture that integrates heterogeneous cyber–physical data acquisition, contextual system representation, domain-informed feature engineering, machine learning analytics, explainable decision support, adaptive operational response, and continuous learning within a closed-loop resilience cycle. Unlike conventional anomaly detection pipelines that terminate after identifying abnormal conditions, DTERF positions anomaly detection as one stage of a broader operational process in which detected events continuously contribute to improving system awareness, operational decision-making, and long-term resilience.

The conceptual evaluation conducted through requirements-to-architecture traceability indicates that DTERF provides explicit architectural mechanisms for the principal technical and operational requirements identified in the literature. Specifically, the assessment supports the functional coverage, traceability, and internal consistency of the proposed architecture across anomaly detection, scalability, interpretability, cybersecurity robustness, resilience support, operational applicability, and architectural integration. These findings concern the architectural soundness of DTERF and should not be interpreted as evidence of detection accuracy, computational performance, real-time feasibility, or operational superiority, which require empirical validation.

The application scenarios provide conceptual walkthroughs of how these architectural mechanisms can interact across representative operational situations, including cyberattacks, renewable-induced operational instability, predictive asset monitoring, and coordinated cyber–physical disturbances. These scenarios illustrate how contextual system representation and machine-learning analytics can be coordinated to distinguish potentially anomalous conditions from expected operational variability while supporting explainable and operationally meaningful decisions. They are intended to demonstrate architectural applicability rather than quantitative performance.

Despite these opportunities, several important challenges remain before machine-learning-enabled anomaly detection can be widely deployed in operational smart grids. Data scarcity, highly imbalanced datasets, computational scalability, interoperability with

legacy infrastructures, cybersecurity protection, and model interpretability continue to represent significant barriers to large-scale implementation. Addressing these challenges will require closer collaboration between machine learning researchers, power system engineers, cybersecurity specialists, and utility operators to ensure that analytical advances translate into practical operational solutions.

Future research should therefore move beyond algorithm-centric investigations toward integrated cyber–physical intelligence architectures. Several research directions appear particularly promising. First, Explainable Artificial Intelligence (XAI) should continue to evolve toward real-time, operator-oriented explanation mechanisms capable of increasing trust in automated decisions. Second, federated learning offers opportunities for collaborative model development across utilities while preserving data privacy and regulatory compliance. Third, Digital Twin technologies should be expanded beyond visualization platforms to support predictive simulation, scenario analysis, and online operational validation of anomaly detection strategies.

Additional research is also required in edge artificial intelligence to enable low-latency distributed analytics, physics-informed machine learning to improve physical consistency and generalization, and robust machine learning techniques capable of resisting adversarial attacks, data poisoning, and model manipulation. Finally, future work should include the implementation and experimental validation of DTERF using realistic smart grid testbeds or Digital Twin simulation environments. Such validation would enable quantitative assessment of the framework with respect to detection performance, scalability, explainability, computational efficiency, and resilience under representative cyber–physical operating conditions.

In summary, this study contributes a comprehensive synthesis of current research together with a resilience-oriented conceptual architecture that extends the role of machine learning beyond anomaly detection alone. By integrating Digital Twin technology, explainable artificial intelligence, adaptive response mechanisms, and continuous learning within a unified operational framework, DTERF provides a foundation for the development of next-generation intelligent smart grids. As renewable penetration continues to increase and cyber–physical interactions become more complex, architectures capable of combining advanced analytics with contextual awareness and resilience-oriented decision support can provide an important foundation for improving the reliability, cybersecurity, and sustainability of future electricity networks.

Author Contributions: Conceptualization, F.F.Y. and M.H.; methodology, F.F.Y.; formal analysis, J.F. and M.H.; investigation, S.K.S. and C.K.R.; writing—original draft preparation, M.H.; writing—review and editing, F.F.Y., C.K.R. and S.K.S.; supervision, S.K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: During the preparation of this manuscript, the author(s) used ChatGPT (GPT-5, OpenAI, https://chatgpt.com/) for the purposes of language editing and manuscript refinement. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
AMI	Advanced Metering Infrastructure
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
CPS	Cyber–Physical System
DER	Distributed Energy Resource
DTERF	Digital-Twin-Enabled Resilience Framework
DoS	Denial of Service
DT	Digital Twin
ESS	Energy Storage System
FDI	False Data Injection
GNN	Graph Neural Network
IBR	Inverter-Based Resource
IoT	Internet of Things
LSTM	Long Short-Term Memory
ML	Machine Learning
PMU	Phasor Measurement Unit
PV	Photovoltaic
RF	Random Forest
SCADA	Supervisory Control and Data Acquisition
SVM	Support Vector Machine
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

References

1. Al-Shetwi, A.Q.; Hannan, M.A.; Al-Masri, H.M.; Sujod, M.Z. Latest advancements in smart grid technologies and their transformative role in shaping the power systems of tomorrow: An overview. *Prog. Energy* **2025**, *7*, 012004. [CrossRef]
2. Kiasari, M.; Ghaffari, M.; Aly, H.H. A comprehensive review of the current status of smart grid technologies for renewable energies integration and future trends: The role of machine learning and energy storage systems. *Energies* **2024**, *17*, 4128. [CrossRef]
3. Cavus, M. Advancing power systems with renewable energy and intelligent technologies: A comprehensive review on grid transformation and integration. *Electronics* **2025**, *14*, 1159. [CrossRef]
4. Paul, B.; Sarker, A.; Abhi, S.H.; Das, S.K.; Ali, M.F.; Islam, M.M.; Saqib, N. Potential smart grid vulnerabilities to cyber-attacks: Current threats and existing mitigation strategies. *Heliyon* **2024**, *10*, e37980. [CrossRef]
5. Yu, X.; Xue, Y. Smart grids: A cyber–physical systems perspective. *Proc. IEEE* **2016**, *104*, 1058–1070. [CrossRef]
6. Bouslimani, M.; Benbouzid-Si Tayeb, F.; Amirat, Y.; Benbouzid, M. Cyber-Physical Security in Smart Grids: A Comprehensive Guide to Key Research Areas, Threats, and Countermeasures. *Appl. Sci.* **2025**, *15*, 12367. [CrossRef]
7. Reuters. Chile Power Outage Plunges Capital into Darkness, Hits Major Copper Mines. 25 February 2025. Available online: https://www.reuters.com/world/americas/power-outage-hits-vast-swaths-chile-largest-copper-mine-santiago-streets-2025-02-25/ (accessed on 16 April 2026).
8. Rouco, L. The blackout of the Iberian Peninsula on April 28, 2025. In Proceedings of the Champéry Power Conference, Champéry, Switzerland, 8–13 February 2026.
9. Madabhushi, S.; Dewri, R. A survey of anomaly detection methods for power grids. *Int. J. Inf. Secur.* **2023**, *22*, 1799–1832. [CrossRef]
10. Olawumi, M.A.; Oladapo, B.I. Enhancing grid stability with machine learning: A smart predictive approach to residential energy management. *Energy Build.* **2025**, *338*, 115729. [CrossRef]
11. Saffari, M.; Khodayar, M. Spatiotemporal deep learning for power system applications: A survey. *IEEE Access* **2024**, *12*, 93623–93657. [CrossRef]
12. Baimukhanov, S.; Ali, H.; Yazici, A. Enhancing ML-based anomaly detection in data management for security through integration of IoT, cloud, and edge computing. *Expert Syst. Appl.* **2025**, *293*, 128700. [CrossRef]
13. Wei, T.; Li, H.; Miao, J. Integration and Development Path of Smart Grid Technology: Technology-Driven, Policy Framework and Application Challenges. *Processes* **2025**, *13*, 2428. [CrossRef]

14. Longo, A.; Aghazadeh, A.A.; Lazari, A.; Ficarella, A. Cyber–Physical Resilience: Evolution of Concept, Indicators, and Legal Frameworks. *Electronics* **2025**, *14*, 1684. [CrossRef]
15. Córdova, F.; Yanine, F.; de Saint Pierre, T.; Rother, H.; Aly, M. Intelligent Grid Support with VPPs: A Scalable Solution for PMGD Integration Challenges. In Proceedings of the 2025 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), Valparaíso, Chile, 28–30 October 2025; pp. 1–6. [CrossRef]
16. Haque, M.M.; Wolfs, P. A review of high PV penetrations in LV distribution networks: Present status, impacts and mitigation measures. *Renew. Sustain. Energy Rev.* **2016**, *62*, 1195–1208. [CrossRef]
17. Karimi, M.; Mokhlis, H.; Naidu, K.; Uddin, S.; Bakar, A.A. Photovoltaic penetration issues and impacts in distribution network–A review. *Renew. Sustain. Energy Rev.* **2016**, *53*, 594–605. [CrossRef]
18. Zhang, Q.; Hu, Q.; Sun, S.; Mei, D.; Liu, S.; Liu, X. Voltage and frequency instability in large PV systems connected to weak power grid. *Front. Energy Res.* **2023**, *11*, 1210514. [CrossRef]
19. Rao, C.K.; Sahoo, S.K.; Yanine, F.F. A review of IoT-enabled intelligent smart energy management for photovoltaic power forecasting and generation. *Unconv. Resour.* **2025**, *9*, 100279. [CrossRef]
20. Liberati, F.; Garone, E.; Di Giorgio, A. Review of cyber-physical attacks in smart grids: A system-theoretic perspective. *Electronics* **2021**, *10*, 1153. [CrossRef]
21. Córdova, F.; Yanine, F.; Carrasco, S.; Rother, H.; Duarte, V. Diagnosing grid's service quality issues: The virtual microgrid and the digitization and innovations in the power distribution grid. *Procedia Comput. Sci.* **2024**, *242*, 1307–1314. [CrossRef]
22. Rajkumar, V.S.; Ştefanov, A.; Presekal, A.; Palensky, P.; Torres, J.L.R. Cyber attacks on power grids: Causes and propagation of cascading failures. *IEEE Access* **2023**, *11*, 103154–103176. [CrossRef]
23. Rahaman, M.O.; Mahin, M.R. Machine Learning Techniques for Anomaly Detection in Smart Grids. *J. Humanit. Soc. Sci. Stud.* **2024**, *6*, 159–165. [CrossRef]
24. Mohammed, S.H.; Al-Jumaily, A.; Singh, M.S.J.; Jiménez, V.P.G.; Jaber, A.S.; Hussein, Y.S. A Review on the Evaluation of Feature Selection Using Machine Learning for Cyber-Attack Detection in Smart Grid. *IEEE Access* **2024**, *12*, 44023–44042. [CrossRef]
25. Qi, R.; Rasband, C.; Zheng, J.; Longoria, R. Detecting Cyber Attacks in Smart Grids Using Semi-Supervised Anomaly Detection and Deep Representation Learning. *Information* **2021**, *12*, 328. [CrossRef]
26. Schummer, P.; del Rio, A.; Serrano, J.; Jimenez, D.; Sánchez, G.; Llorente, Á. Machine Learning-Based Network Anomaly Detection: Design, Implementation, and Evaluation. *AI* **2024**, *5*, 2967–2983. [CrossRef]
27. Zhou, F.; Wen, G.; Ma, Y.; Geng, H.; Huang, R.; Pei, L. A Comprehensive Survey for Deep-Learning-Based Abnormality Detection in Smart Grids with Multimodal Image Data. *Appl. Sci.* **2022**, *12*, 5336. [CrossRef]
28. Yanine, F.; Masrur, H.; Montoya, F.; Hidalgo, M. Enhancing grid stability and sustainability through virtual power plants: A case study of Chile. *Ain Shams Eng. J.* **2026**, *17*, 104333. [CrossRef]
29. Hassan, S.U.; Abdulkadir, S.J.; Zahid, M.S.M.; Al-Selwi, S.M. Local interpretable model-agnostic explanation approach for medical imaging analysis: A systematic literature review. *Comput. Biol. Med.* **2025**, *185*, 109569. [CrossRef]
30. Pelosi, D.; Cacciagrano, D.; Piangerelli, M. Explainability and Interpretability in Concept and Data Drift: A Systematic Literature Review. *Algorithms* **2025**, *18*, 443. [CrossRef]