International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(5s) 2026 pp. 996-1010
https://doi.org/10.70917/ijcisim-2026-2847
Article

Artificial Intelligence and Internet of Things-Based Smart Grid Technologies for Intelligent Energy Management Systems: A Comprehensive Systematic Review and Future Perspectives

Akshay Akhade[1], Sudhir Lande[2], Dineshkumar Adokar[3]

[1]*Department of Electrical Engineering, Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Baramati, Maharashtra, India

[1]Eamil Id: akshay.akhade@vpkbiet.org [1]ORCID ID: 0009-0006-9208-7351

[2]Department of Electronics and Telecommunication Engineering, Vidya Pratishthan's Kamalnayan Bajaj Institute of Engineering and Technology, Baramati, Maharashtra, India

[2]Eamil Id: sudhir.lande@vpkbiet.org [2]ORCID ID: 0000-0001-5204-8944

[3]Department of Electrical Engineering, Adsul's Technical Campus, Chas, Ahilyanagar, Maharashtra, India

[3]Eamil Id: dadokar@gmail.com [3]ORCID ID: 0009-0009-4693-9420

***Corresponding Author Email Id**: akshay.akhade@vpkbiet.org

Abstract: The traditional power grid model is no longer viable; with the increasing use of variable renewable sources, electric vehicles, and heating systems combined with an aging infrastructure and increasing peak demand the power grid is being stretched to the limit. To address these challenges, a new cyber-physical system called the "smart grid" is emerging. Smart grids will monitor the flow of electricity through the grid using sensor networks and smart meters. In addition to monitoring the flow of electricity, smart grids will also enable two-way communications between utility companies and customers. These two-way communications will allow utility companies to automatically control customer side loads based on changing grid conditions. As a result of these changes, there will be increased opportunities for utility companies to better manage their electrical distribution assets. Therefore, utility companies need to understand what smart grid technologies exist today, how they work together, and how they can be used to improve the performance of the utility company's distribution assets. One of the key drivers of smart grids is the widespread adoption of advanced wireless sensor networks. Advanced wireless sensor networks are a type of "Internet of Things" (IoT). IoT enables the connection of physical objects to a network so that the object may send or receive information. Sensor networks may consist of thousands of nodes that are dispersed throughout a geographic area. Each node may include multiple sensors that measure different environmental factors. For example, a node might include temperature, humidity and light sensors. Advanced wireless sensor networks are a technology enabler for smart grids because they provide the means by which utility companies can collect large amounts of data about the current operating state of their distribution assets.
Keywords: Power Grid, Renewable Sources, Electric Vehicles, Sensor Networks, IoT.

1. Introduction

1.1 Background and Motivation

A shift to new models will be required because electricity drives most aspects of both the modern economy and society. Moreover, growing demand globally for electricity results from multiple trends advancing simultaneously (industrialization, urbanization, digitization, and the widespread adoption of electric vehicles). The original electric utility grid model was based upon a single direction flow from large centralized generating plants with no consideration given to either the reverse flow of power or consumer participation. This model worked effectively for the 20th Century but has no ability to adapt to a rapidly changing environment in which the number of controllable end use devices is increasing and where millions of consumers are equipped with rooftop solar panels, community batteries, and/or electric vehicles. Fang et al. characterized the transformation as a movement away from the "legacy" grid toward a "new and improved" power grid whose function depends equally on digital communication & control as it does on copper and steel [1]. Similarly, Gharavi and Ghafurian framed the smart grid as the electric energy system of the future, with two-way communication between the network and its customers [2].

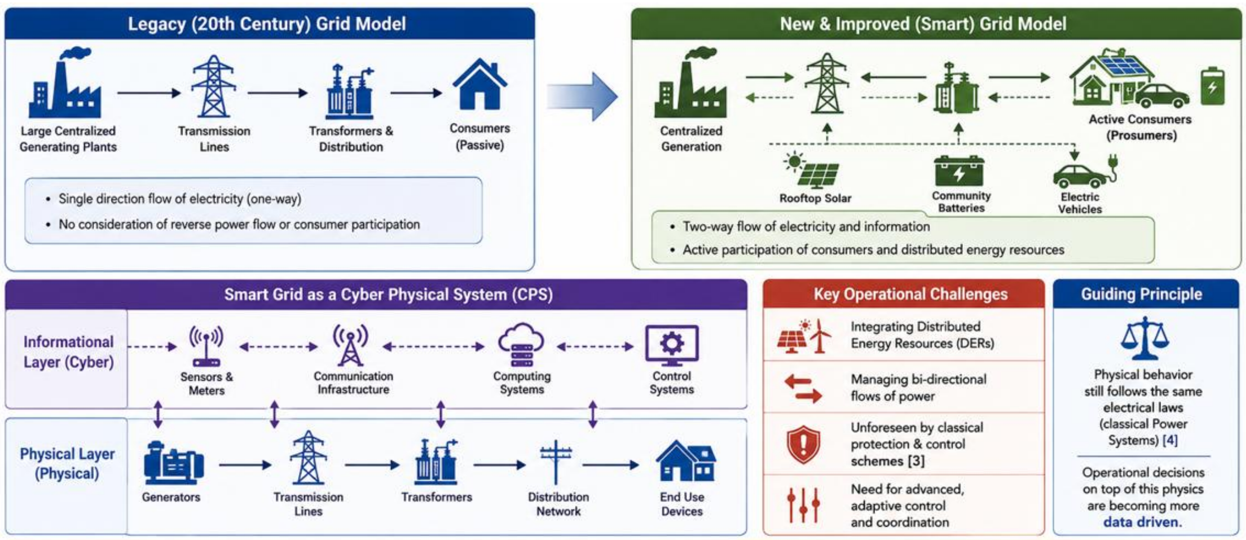

Figure 1. From Legacy Grid to Smart Grid.

The Smart Grid is best thought of as a Cyber Physical System (CPS) in which the physical layers of generators, transmission lines, transformers, etc., are mirrored by an informational layer consisting of sensors, communications infrastructure, computing systems and control systems. Many of the operational challenges associated with managing this "Grid of the Future," as envisioned by Ipakchi and Albuyeh, include accommodating Distributed Energy Resources (DERs), and controlling bi-directional flows of power. These were unforeseen by classical protection and control schemes [3] that are now being used in today's Smart Grid. The physical behavior of these CPSs—i.e. voltage stability, frequency regulation, power quality—is still subject to the same electrical laws that describe classical Power Systems [4]; however, the operational decisions made on top of this physics are becoming more data driven. (Figure 1.)

1.2 Limitations of the Conventional Grid

There are many physical/structural reasons why there should be a switch from traditional to smart energy management. The first reason is that the majority of our existing infrastructure is old (i.e., aging) and operating at near maximum levels with very little room for renewable resources which will cause large swings in power due to weather. Secondly, peak demand growth is higher than average demand growth. Therefore, the utility companies need to purchase and maintain costly reserve generating units that run nearly all of the time. Thirdly, as mentioned previously, the sun doesn't always shine or the wind blow; therefore, the suppliers of electricity (solar/wind) generate electricity when it isn't needed. This is an example of what is called a stochastic process in supply terms. The traditional method of scheduling supply/demand does not work well with such processes [5]. Fourth, faults are generally identified "after the fact" i.e., by the time a customer loses service, because we do not monitor the medium and low voltage distribution systems where faults typically occur. Finally, until recently, consumers did not possess the information required to move their load to reduce stress on the overall grid system nor was it economically beneficial to them to do so. Together, these factors support a continuous sensing of the status of the grid, forecasting of potential problems/disruptions and autonomous action by the grid (all features enabled by Internet of Things (IoT), Artificial Intelligence (AI)) [6].

1.3 The Convergence of AI and IoT

IoT and AI can actually work together instead of against each other. IoT provides the body (the "nervous system") for the network — through smart meters, phasor measurement units, distribution sensors, inverters and customer side equipment that provide electrical information and report on the condition of the electrical system via both wired and wireless connections. AI provides the brain ("cognition"): through statistical learning, neural networks and reinforcement-learning agents that process the electrical information provided by IoT and turn that into forecasts and control strategies. In fact, Saleem et al., have demonstrated the integration of these two technologies; they implemented an IoT based smart energy-management system using cloud computing to provide an efficient way to manage peak demands, and created a hierarchical model (perception-layer, communication-layer, cloud-analytics-layer) where IoT based smart energy-management devices collect data, the communication layer sends it to the cloud and the cloud analytics closes the control-loop [9]. Similarly, Hussain et al., reviewed multi-level energy-management systems and agreed that until intelligent analytics are embedded throughout the appropriate layers of the hierarchy that the value of increasing amounts of "data exhaust" produced by the grid cannot be realized [10].

1.4 Contributions of this Review

The body of work concerning AI and/or IoT in Power Systems has grown significantly in size and speed of growth; however, a significant proportion of this body of work is fragmented (i.e., focused upon a singular method/algorithm/applications/enabling technology), with a substantial amount being narrative instead of systematic. This document contributes to the area as follows:

• By applying an openly described, replicable evaluation process, which was conducted under the PRISMA 2020 Statement; whereby the identification, selection, suitability and inclusion of studies were documented in such a manner to allow auditing and updating of the evidence base.

• By documenting a unifying layering of IoT sensor technologies, communications standards, the edge-fog-cloud computing continuum and AI data analysis tools into one unified conceptual framework/model.

• By providing a well-structured taxonomy of the various forms of AI analysis applied to concrete energy management applications/tasks that have been substantiated by recent empirical results/evidence as opposed to claims/assertions.

• By analyzing four new paradigms- federated learning, blockchain enabled trading platforms, digital twins, and explainable AI are transitioning from prototype development towards implementation/deployment.

• By providing a critical review/discussion of the current status of unresolved challenges/open issues regarding the use of AI/IoT in power system applications; along with proposing a high-priority research agenda.

1.5 Organisation of the Paper

The remaining sections of this paper are structured as follows. In Section 2, a detailed explanation of the methodology that was employed to conduct the PRISMA-based review will be presented. This section will include an explanation of the search strategy for the literature review; a description of the selection criteria used to select studies that were included in the review; and finally, it will provide an overview of how studies were selected from those identified by the search process.

In Section 3, the authors will describe their proposed layered AI–IoT smart-grid architecture, which includes descriptions of each of the layers of technology within the proposed architecture and also includes explanations of the communication technologies that enable data exchange between these different layers. Section 4 will explore the role of artificial intelligence (AI) in facilitating intelligent energy management through energy systems. And similarly, in Section 5, the authors will examine the potential role of Internet-of-Things (IoT) technologies in supporting such systems. Section 6 provides an integration of examples of representative applications that have been developed with the support of both AI and IoT technologies. In addition, Section 7 will present an evaluation of current enabling and emerging technologies that may further facilitate development of AI/IoT enabled smart-grids. And finally, Section 8 will identify key challenges associated with successful implementation of smart-grids that employ AI and IoT. Section 9 will outline future research direction opportunities that may emerge if or when smart-grids that incorporate AI and IoT become widely accepted. And finally, Section 10 summarises and draws conclusions regarding findings related to AI/IoT based smart-grids.

2. Review Methodology

The aim of a systematic literature review is to provide a high degree of transparency and reproducibility, therefore it was performed according to the four phases (identification, screening, eligibility, and inclusion) as described in the PRISMA 2020 guidelines. Reporting structure and terminology are presented in accordance with the above mentioned guideline [7] and supplemented by further detailed recommendations concerning systematic literature reviews in engineering and computer science [8].

2.1 Review Questions and Protocol

A preliminary agreement on a protocol for research that would be conducted prior to a search in order to provide focus for the future, as well as guide the study with four key questions were established to help define the scope of this study.

RQ1: How are architecture and communications technologies used in developing AI- and IoT-based Smart Grids?

RQ2: What type of AI applications are being developed to support the three main areas of energy management such as: forecasting, fault detection & management and demand response? And what is their effectiveness.

RQ3: In what way can an IoT device/layer enable real time intelligence for energy management?

RQ4: What are some of the emerging paradigms and challenges which will develop into the next generation of intelligent energy management systems.

2.2 Information Sources and Search Strategy

Six databases were used as part of this literature study. They included: IEEE Xplore; Elsevier's Science Direct (which contains the journals, books and reference works of the company Elsevier); SpringerLink which is the portal for all online journals and books published by Springer; MDPI (Multidisciplinary Digital Publishing Institute) which publishes over 300 Open Access Journals across various disciplines; Web of Science (WoS) and Scopus – both are commercial database services designed to help researchers identify relevant materials they need to carry out their research. Citation tracking was also done via forward and backward citation tracking ("snowballing") from the major reviews. These searches employed Boolean operators to combine three concept groups. Group (i): An enabling technology group (“Artificial Intelligence”, “Machine Learning”, “Deep Learning”, “Internet of Things”, “IoT”, “Federated Learning”, “Digital Twin”, “Blockchain” etc.). Group (ii): A Domain Group (“Smart Grid”, “Energy Management”, “Power System”, “Demand Response”, “Load Forecasting”, etc.). Group (iii): Outcome Groups (“Forecasting”, “Fault Detection”, “Optimisation”, “Security”, “Privacy”, etc.). Only Journal Articles, Conference Papers and Review Articles were searched. The time frame for these searches was from 1st January 2019 to mid-2025.

2.3 Eligibility Criteria

Inclusion and exclusion criteria were defined explicitly to keep screening consistent between reviewers. They are summarised in Table 1.

Table 1. Inclusion and exclusion criteria used during screening.

Criterion	Inclusion	Exclusion
Topic	AI and/or IoT applied to smart-grid or intelligent energy-management problems	Studies unrelated to electrical energy systems or with only incidental mention of the grid
Period	Published 2019–2025	Published before 2019 unless a seminal, frequently cited foundation

Type	Peer-reviewed journal, conference or review article	Editorials, patents, non-archival white papers, unrefereed preprints without substantive results
Language	English	Non-English without an English full text
Substance	Reports a method, architecture, empirical result or structured review	Purely promotional or duplicate content; abstract-only records

2.4 Study Selection

Title and Abstract Screening

After title and abstract screening based on the inclusion/exclusion criteria, 706 records were removed as they did not fit the topic, fell outside of the scope of interest or were clearly promotional. This left 262 records to be assessed via full-text.

Full-Text Assessment Of the 262 full-text records that were accessed for assessment, 174 full-texts were reviewed thoroughly. A total of 88 full-text records were excluded from the final synthesis with justifications: - Insufficient Methodological Detail (n=31) - No Smart Grid Application (n=27) - Redundancy with Stronger Study (n=19) - Full Text Could Not Be Retrieved (n=11)

The remaining 86 full-text records that met the inclusion criteria formed the qualitative synthesis pool from which the most representative and methodologically robust studies will be referenced directly within this review. (Figure 2.)

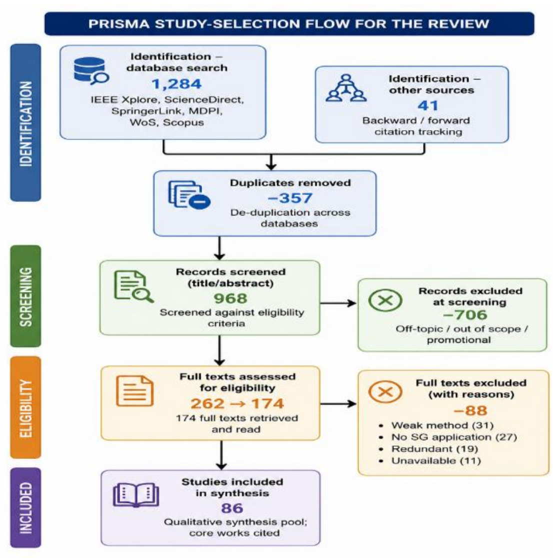

Figure 2. PRISMA study-selection flow for the review.

2.5 Data Extraction and Synthesis

In terms of the data for every single study that was included in this analysis, an overall extraction form was developed to record all relevant bibliographic information along with the two main types of technologies (i.e., AI/IoT) covered, the particular application area being studied within the context of energy management, the actual data sets or testing environments employed, the results of the studies' performances and the limitations recognized by the authors. Due to the wide variety of objectives and data sets presented by the individual studies along with the different types of performance metrics employed by those same studies, it was neither possible nor practical to undertake a systematic quantitative meta-analysis across all of them; instead, all of the studies were synthesized on a qualitative basis using a thematic-narrative synthesis format structured to address the four key research questions outlined above. The advantages of this method are twofold: first, the variability inherent in the results from these studies is preserved as well as secondly, commonalities across the individual studies can be noted — e.g. the use of hybrid deep-learning architectures to forecast future performance, and/or the growing need to develop methods that preserve users' privacy when learning — and can be discussed/identified and documented.

3. AI–IoT Smart-Grid Architecture

Although some variation may exist across different individual deployments, the overall structure and layering of an AI- and IoT-enabled smart grid is generally consistent within the body of literature. Understanding this general structure will help identify both the appropriate place for sensing, communication, computation, and control functions to reside, and where many of the major engineering trade-offs are located.

3.1 A Layered Reference Model

In addition, the structure and layering of the architecture can be defined by identifying four cooperating layers. The first layer, or Perception Layer (also known as Device Layer), contains the physical components of the grid. Examples include smart meters; current and/or voltage sensors; PMUs (phasor measurement units); transformer and substation monitors; renewable-based inverters; energy storage systems; and load controllers. Each device continually collects data regarding electrical characteristics, including voltage, current, frequency, active and reactive power; and, more recently, environmental factors (such as temperature and irradiance) which directly influence renewable-based energy production [9], [11].

The second layer, or Communication Layer (Network Layer), provides for transmission of sensor data from each monitoring location up to higher levels of management within the grid system; and transmission of control commands down to lower levels of management within the grid system. In order to satisfy all of the competing requirements for the Communication Layer (wide geographic area coverage; minimal energy consumption for battery powered field devices; sufficient bandwidth to support high resolution data; and minimal delay times for fault detection and control), no single technology exists which meets all of the criteria listed above. Real-world deployments therefore utilize multiple technologies (e.g., short range and long range standards). More detailed information regarding the use of these various standards is provided in Section 3.3.

The third layer of the architecture is referred to as Edge-Fog Computing Layer. This layer is responsible for processing collected sensor data locally at the point-of-data collection. Local processing enables the removal of extraneous data prior to back haul to remote locations; minimizes delay times for time critical processes such as fault detection and isolation; and eliminates the need to send large amounts of sensitive customer data over public networks. The last layer of the architecture is referred to as Cloud Layer. The Cloud Layer includes compute-intensive analytics (i.e., large scale weather forecasts; network wide optimization; wholesale market operations; etc.), as well as, energy management software that utility companies interact with [10].

3.2 Core Components

The integrated smart grid model is made up of many layers with the functional areas being as follows:

• The smart meter and AMI (advanced metering infrastructure) will collect a large amount of detail in both directions to report real-time consumption and production.

• Wide area visibility of the current state of the electrical distribution system through intelligent sensors and phasor measurement units.

• Distributed Energy Resources; rooftop solar photovoltaic panels, small wind turbines and battery storage systems for prosumer.

• Renewable generators along with their associated power electronic interfaces that produce an unpredictable output that can be anticipated by the predictive analytics layer.

• Communication networks that exist within the three different domain levels; Field Area Network, Neighbourhood Area Network and Wide Area Network.

• Energy Management Systems that have forecasting, optimization and controls functionality.

• Edge and Cloud Computing Platforms are required to support the high performance computational needs for edge-based AI

3.3 Communication Technologies

The choice of communication standard shapes what analytics are possible. Short-range technologies connect devices within a home or substation, while low-power wide-area and cellular technologies connect dispersed field assets to control centres. Table 2 compares the technologies most frequently reported in the reviewed literature.

Table 2. Comparison of communication technologies for AI–IoT smart grids (indicative values).

Comparison of Communication Technologies for AI–IoT Smart Grids
(Indicative Values)

Technology	Range	Data rate	Power	Typical smart-grid use
Wi-Fi	~50–100 m	High	High	Home energy management, in-building metering
ZigBee	~10–100 m	Low	Very low	Home-area networks, appliance-level control
LoRaWAN	2–15 km	Very low	Very low	Rural metering, distributed sensor telemetry
NB-IoT	1–10 km	Low	Low	Massive smart-meter roll-outs, deep-indoor sensors
5G / 5G-NR	Cellular	Very high	Moderate	Low-latency protection, real-time wide-area control
PLC (power-line)	Grid-bound	Low–med	N/A	Metering over existing conductors

In practice, hierarchical deployments are common: ZigBee or Wi-Fi within premises, NB-IoT or LoRaWAN for neighbourhood metering, and 5G or fibre for latency-critical protection and wide-area monitoring. The forthcoming move toward 6G is expected to push ultra-low-latency, high-density connectivity further, enabling tighter closed-loop control at the grid edge [34].

4. The Role of Artificial Intelligence in Intelligent Energy Management

AI adds value wherever the volume, velocity or complexity of grid data exceeds what conventional analytic and rule-based tools can handle. This section taxonomises the principal techniques and then maps them to the core energy-management tasks, drawing on recent empirical studies.

4.1 A Taxonomy of AI Techniques

A number of classes exist within the broader groupings of the approaches in the papers reviewed. The classical machine learning approach (e.g., support vector machines, random forest classifiers, gradient-boosted tree) is still widely used for tabular based tasks like fault classification and intrusion/ cyber-attacks since this method is both efficient with respect to data requirements, and has good interpretability. Artificial neural networks have become dominant for sequence or image-type problems. In particular, Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTMs) and Gated RNNs are often applied to modeling the temporal relationships present in time series data related to loads and generations. Convolutional Neural Networks can be applied for extracting localized and spatial features. Reinforcement Learning can frame sequential decision-making problems (i.e., when should you change your batteries, what prices would you need to set for demand response programs?) as the optimization of an expected future value. More recently, hybrid models which integrate various different techniques (signal decomposition, convolutional feature extraction, Temporal Modelling using Recurrent Neural Networks, and Attention Mechanisms) into one processing flow have produced superior results than individual approaches [13], [14] as shown in Table 3 below.

Table 3. Taxonomy of AI techniques and their principal smart-grid applications.

AI family	Representative methods	Primary grid applications
Classical ML	SVM, Random Forest, XGBoost	Fault classification, cyber-attack detection, theft detection
Neural / deep networks	ANN, CNN, LSTM, GRU, Transformer	Load & renewable forecasting, state estimation, anomaly detection
Reinforcement learning	Q-learning, DQN, DDPG, PPO	Demand response, storage & EV scheduling, voltage control
Hybrid / ensemble	EMD–LSTM, CNN–LSTM–attention, metaheuristic–DL	High-accuracy forecasting, multi-objective optimisation
Distributed learning	Federated learning	Privacy-preserving forecasting and theft detection

4.2 Electricity Load Forecasting

Load forecasting is a key element in both economic and secure grid operation. It provides load data for unit commitment, reserve procurement, network switching and market bidding. Deep learning has become more popular than statistical models in this field of study. Mounir et al., used an empirical mode decomposition with a bidirectional layered lstm to forecast short term loads on a smart-grid energy management system. They found that decomposing the signal before training significantly improved accuracy when dealing with non-stationary demand [13]. Mughees et al. Also showed that deep sequence to sequence bidirectional lstm architectures outperform shallower baselines when forecasting day-ahead peak loads. The value of these predictions are high due to the fact that capacity costs are primarily driven by peak loads [14]. Therefore, the broader literature confirm that there is a general trend where recurrent and convolutional architectures; especially those hybrids that use attention have lower forecast error rates than classical time series model and single layer neural networks. However, these architectures require large amounts of data and computational power [13], [14]. Additionally, reviews of the area note that the value of forecasting has grown as smart meters raised the temporal resolution of available data which made real-time and probabilistic forecast feasible.

4.3 Renewable Generation Forecasting

The main challenge in today's grid from the variable nature of both solar and wind is the forecasting of this variability and how much we can rely on these sources. To increase renewable penetration, machine learning models that take into account both numerical weather forecasts and historical data are the most used method for forecasting solar (photovoltaic) and wind energy. Hybrid deep learning architectures that integrate feed forward networks with recurrent neural networks and optimization algorithms provide good results for the joint forecast of PV and demand [35] . This work has been followed by the development of probabilistic and scenario-based approaches that evaluate the uncertainty associated to each source [36], because this is what operators really need when they schedule reserve.

4.4 Fault Diagnosis and Predictive Maintenance

Because IoT instrumentation allows for continuous monitoring of asset conditions whereas intermittent monitoring was used in the past; AI will be able to transition Maintenance schedules (reactive/fixed) into predictive/condition based. AI classifiers trained using signatures collected from sensors can identify impending failures as opposed to transient anomalies, and locate fault areas; additionally, remain useful life estimates may be made regarding transformers, cables and switchgear. This has a direct impact on both outage prevention prior to failure cascades and concentration of maintenance efforts where there is a true need, thus reducing costs. One of the main themes throughout the literature is the lack of labelled fault data that prompts both transfer learning and synthetic creation of fault scenarios with increasing use of digital twin environments (Section 7.3).

4.5 Demand Response and Reinforcement Learning

Demand Response (DR) — using price or reliability information to shift/curtail end-use consumption at times of high load or when there are excesses of renewable generation — reduces peak demand and can help stabilize supply. The sequential, uncertainty-based nature of DR makes it a good fit for reinforcement learning. Vázquez-Canteli & Nagy have provided an extensive review which maps various forms of reinforcement learning (RL) algorithms/models with various forms of demand-response modeling [16]. Bahrami et al. used Deep RL for demand response in Distribution Networks; as such they were able to learn the optimal control policy through experience without having to use an explicit system model [17]. In even more recent work authors include pricing/incentive design into the objective function of the RL algorithm so that the agent optimizes tariffs that will direct aggregate behavior towards desired System Goals while honoring consumer preference [18]. In addition to its other contributions, a recent review of RL for the evolving grid has also highlighted both the many applications of RL including Storage Dispatch, EV Charging, Voltage Regulation, Building Control — and the need for significant improvement in areas related to Safety, Sample Efficiency, Transferability etc., before we see widespread adoption of RL in these applications [19].

4.6 Energy Scheduling and Price Forecasting

AI provides support to the economic aspect of the management of a power system. Similarly, AI has improved in its ability to forecast electricity prices. This will allow both market participants and utilities to use price responsive resources to manage the supply and demand of electricity, which has the potential to benefit from AI's improvements in load forecasting. AI has the ability to optimize the amount of electric generation, the amount of storage available and how much flexibly a utility can use different types of loads. The optimization techniques currently being employed include; metaheuristics and learning based scheduling algorithms that provide utilities with an advantage over traditional (deterministic) scheduling techniques when it comes to managing the variability associated with renewable generation.

5. The Internet of Things in Smart Grids

If AI is the cognition of the smart grid, IoT is its sensory and motor system. This section examines the IoT device ecosystem, the metering infrastructure that anchors it, the edge–cloud continuum that processes its data, and the operational benefits it delivers.

5.1 The IoT Device Ecosystem

IoT in the grid has a very large variety of device types ranging from low cost microcontrollers used in simple sensor devices to highly advanced devices monitoring substations. Smart meter devices, transformer devices, feeder devices, distributed generator devices and customer premise equipment are all sending information back and forth with their respective control centers and as such provide the distinction between today's smart grid and yesterday's traditional grid. Abir and co-authors identified both the application opportunities enabled by IoT connections and the challenges (scale, heterogeneity, security) that are introduced when they do so; however, given that there will be literally hundreds of millions of end points in a typical country-wide deployment, it is because of the massive scale and speed of these data inputs that Artificial Intelligence is required to make sense of them.

5.2 Smart Metering and Advanced Metering Infrastructure

Smart metering is currently the most mature and deployed application in the grid that falls under IoT. Advanced meter infrastructure measures energy flows accurately and at high frequency bidirectional and allows automatic reading of meters and billing based on time of use. Crucially, granular consumption data that demand response and intrusive load monitoring rely on also comes from rich data. Saleem and colleagues showed how a management system anchored to meters and connected through middleware to cloud analytics analytics can deliver real time monitoring and measurable savings for demand side management [9]. Rich data though is very personal and can reveal occupancy and behavior which is why preserving privacy has become a central theme of research (Section 7.1).

5.3 The Edge–Fog–Cloud Continuum

Not every type of grid intelligence will be able to reside in the Cloud. Since latency sensitive functions like protection, fast voltage control and local fault isolation have decision times that are much less than 1 ms, they cannot afford to send a signal (in either direction) back and forth from a remote data center. As such edge and fog computing is used to move inference closer to where it is needed and reserve the Cloud for heavy, non-time critical analysis and inter network coordination. There is also a privacy aspect; by processing the raw data at the device and sending only aggregate data or updates to models, the amount of personal identifiable information exposed is limited. Thus there is an architectural challenge to determine how to divide up intelligence among the three layers of the continuum (edge, fog and Cloud), which is currently being explored through active research efforts and is directly related to the development of Federated Learning

5.4 Operational Benefits

The existing body of literature is in agreement that there is a consistent set of advantages provided by IoT instrumentation, which can be summarised as follows:

• The ability to continuously monitor and have real-time visibility into the network's state, particularly those portions of the distribution system that were previously un-monitored.

• The capability to remotely monitor and control equipment thereby minimizing the requirement for hands-on involvement.

• A faster and more precise method of detecting faults, along with automated methods for rapidly healing from failures (i.e. reconfiguring).

• Automatic meter readings; additionally, the flexibility to provide time-based rates/tariffs to consumers.

• A reduction in both technical and commercial loss through the use of data-based methods to detect fraud/theft.

• Increased levels of participation among consumers due to increased transparency of their usage, and greater control over their usage.

• A higher level of operational efficiency due to the elimination of many manual processes through the use of automation.

6. Integration and Applications of AI and IoT

The greatest value arises when AI and IoT operate as a single loop: IoT devices sense the grid, edge and cloud analytics interpret the data, and control actions flow back to actuators. Table 4 draws together representative recent studies that exemplify this integration across the main application domains, illustrating both the diversity of methods and the maturity of the field.

Table 4. Representative recent studies integrating AI and IoT for intelligent energy management.

Study / focus	Approach	Application	Key contribution
Saleem et al., 2023 [9]	IoT + cloud EMS	Demand-side management	Layered IoT–cloud system delivering real-time control and savings
Mounir et al., 2023 [13]	EMD–Bi-LSTM	Short-term load forecasting	Signal decomposition improves accuracy on non-stationary demand
Bahrami et al., 2021 [17]	Deep RL	Demand response	Model-free control policies for distribution networks
Fekri et al., 2022 [15]	Federated RNN	Distributed load forecasting	Accurate forecasting without centralising smart-meter data
Wen et al., 2022 [21]	Federated learning	Energy-theft detection	Privacy-preserving detection across many meters
Lin et al., 2023 [24]	Deep RL	Cyber-security (FDIA)	Adaptive detection of false-data-injection attacks
Jafari et al., 2023 [31]	Digital twin review	Grid monitoring & planning	Maps DT challenges and opportunities for the grid

Two conclusions result from synthesizing the literature. First, the state-of-the-art of the field is moving away from centralized, single-purpose systems toward decentralized, privacy-friendly, and secure (i.e., hardening) systems. As a consequence, several works on federated forecasting, federated theft detection, and learning-based intrusion detection were produced since 2022. Second, it is exactly through the synergy created by integrating the aforementioned components of the system for creating added-value for users that we increase the attack surface as well as the user's privacy risk. This is why, although they could potentially be useful for deploying such systems responsibly, the key-enabling-technologies and challenges described below represent critical issues.

7. Enabling and Emerging Technologies

Four paradigms are reshaping how intelligence is distributed, secured and trusted in the grid. Each addresses a specific limitation of the centralised AI–IoT model.

7.1 Federated Learning

Federated Learning is an approach that allows a single model to be developed across multiple locations (i.e., devices) utilizing the data that each location has available. However, this is done in a manner that does not require the individual device data to leave their respective location. Instead, what leaves the site is the update(s) to the model that were generated from the individual's training dataset. Thus, it is the first practical solution to address one of the primary privacy concerns associated with Smart Meter data collection.

In addition to the theoretical basis for its use in addressing privacy concerns, there is also empirical evidence demonstrating the feasibility of Federated Learning for Distributed Load Forecasting. The authors in reference [15] successfully demonstrated how Federated Learning could be used to develop Recurrent Neural Networks (RNNs) trained on Smart Meter Data collected at different geographic locations. They demonstrated that the RNNs developed via Federated Learning achieved performance levels equivalent to those models that had been developed using all of the collective Smart Meter Data (i.e., Centralized Training). Furthermore, they accomplished this without ever collecting all of the data into a centralized repository where it would become accessible to unauthorized individuals.

7.2 Blockchain and Peer-to-Peer Energy Trading

Blockchain technology allows decentralized trading through a transparent, tamper-proof record which settles all peer-to-peer energy transactions directly among participants (without an intermediary) by using smart-contracts [12]. Smart Contracts represent the rules or terms in each trade. Waseem et al. examined the use of blockchain in many different types of smart grid applications and the architecture issues and barriers that arise from such use [27]. Uddin et al. presented a conceptual framework and commercial practices for developing next generation blockchain-based electric power systems [28]. Hua et al. also identified the combination of blockchain and artificial intelligence as potentially very valuable for supporting prosumers because it has both the ability to predict energy demand/production and optimize production/demand via AI [29]. However, this potential is dependent on overcoming significant challenges related to deploying blockchain technology at scale including, but not limited to, scaling up the number of users that can transact within a reasonable amount of time (transaction volume), increasing the rate at which a block chain can process transactions (transaction throughput), reducing energy consumption associated with running blockchain (energy usage), and eliminating uncertainty regarding how governments will regulate blockchain based markets.

7.3 Digital Twins

Digital Twins enable real time synchronization to a virtual duplicate of a physical asset or the grid as a whole through IoT data. Digital Twins can be used to monitor, optimize, predictively maintain and safely test (i.e. with no risk of disrupting the actual system) extreme/antagonistic scenarios. Jafari et al. surveyed digital-twin technology across Smart Grids, Transport, Smart Cities etc. and identified the key technologies that enabled it and key challenges to implementing it [31]. The authors of reference [32] also provided an overview of digital twins for the future power system with a focus on how they will help bridge the gap between planning, operations and cyber-defense using a common, updated data model. Digital Twins provide another type of artificial intelligence training data in addition to simulated fault cases, attack scenarios and other less frequently occurring states that may occur when an asset is operated under normal operating conditions; all types of data needed to train supervised learning models.

7.4 Explainable AI

As A.I. transitions from providing advisory functions to making operational control decisions, it will be necessary for all of those who work in this environment or are subject to regulatory oversight to be able to understand and trust the decision-making processes of these models. The use of Explainable A.I. (X.A.I.) aspires to allow users to understand how and why a model behaves without having to sacrifice too much of the models' performance. Other areas of energy demonstrate that there is value in using the explainable methods. For example, researchers Tsoka et al., used explainable methods to develop their own method of classifying the energy performance of buildings which demonstrated that it was possible to have both predictive capability and interpretability when developing models [33]. It is also expected to provide additional benefits in terms of accountability; i.e. in addition to being accountable for ensuring safe and reliable supplies of electricity, utilities will need to ensure that each time they make a decision regarding pricing or reliability, that decision could be audited. This represents an emerging area of study; integrating explainability into a system of such complexity as a large scale utility system's distributed architecture and the high levels of security required to protect the integrity of that system presents numerous challenges that require significant amounts of research.

7.5 Edge AI and Next-Generation Communications

The last pattern observed in the future of smart grids is a trend for "intelligence" to migrate from central locations (the Cloud) toward the edges of the power system through use of smaller and lower-power processors and ultra-low latency communication systems. Edge-AI will allow for real-time inferencing on critical issues such as fault-isolation, rapid-voltage adjustment, and localized load management within seconds or less. This reduced latency will also minimize potential data exposures that may result from remote cloud processing. Anticipated capabilities associated with 6G, such as an order of magnitude greater numbers of devices, and deterministic low latency, should further accelerate this transition allowing for the implementation of tighter closed-loop controls on the grid-edge that complement but do not replace large scale cloud-based analytics.

8. Challenges and Open Issues

Despite clear progress, several interlocking challenges stand between the current state of the art and reliable, secure, large-scale deployment.

8.1 Cyber-security

The addition of millions of new smart devices on the grid increases the size of its target area. In this regard, false data injection attacks (FDIAs), where an attacker generates measurement values which meet all of the grid's validation checks, pose a particularly significant threat. Since these types of attacks do not contain 'bad' data in the classical sense, they will evade standard "bad-data" detection techniques. A comprehensive taxonomy of FDIA models, targets and impacts was developed by Reda et al. in [25] who demonstrated that FDIA attacks could be highly varied and adaptable. Learning based detection is now being pursued, with Lin et al. using deep reinforcement learning to develop a method to detect FDIAs which are too diverse to be detected using fixed detectors [24]; while Han et al. have utilized a modified Temporal Multi-Graph Convolutional Network to take advantage of the spatial-temporal structure of the network for effective detection [26]. Nonetheless, cyber security must be considered at the systems level and as a defence-in-depth problem, encompassing both device authentication and secure communications as well as resilient controls.

8.2 Data Privacy

The high sensitivity of metering data makes it possible to determine if an occupant in a dwelling is home/occupied, awake, or absent; centralizing this information for analytic purposes creates a single privacy risk and single point of failure. The primary driver of Federated Learning (FL), Differential Privacy (DP) and Edge Processing (EP) [15], [20], [22], is the tension created by the need for accurate analytical results and strong privacy guarantees. Reconciling these two competing requirements continues to be challenging and the appropriate balance will likely be determined as much through regulations and ethics as through technology.

8.3 Interoperability and Standardisation

Grid ecosystems are inherently diverse in terms of device type (from multiple different vendors) as well as protocol used to communicate. Interoperability among these various devices has been difficult due to lack of standards for data modelling and interface implementation. Data integration costs have also been high as analytics applications require clean, consistent data to perform optimally; however, IoT streams often provide messy and/or inconsistent data. For AI-IoT solutions to scale beyond pilot implementations, there needs to be some level of standardization around data format(s), communication protocol(s) and security baseline(s).

8.4 Scalability, Cost and Latency

A large-scale deployment of sensors, communications equipment and computers to a whole network represents significant investment for most organizations; therefore, the cost of deploying these systems must be compared with the long- term and largely unquantifiable benefits provided by such deployments. At scale, communication latency and bandwidth limitations reduce both the amount of data which may be transferred and the speed at which decisions can be generated. These factors have furthered the argument presented above as to why edge-cloud partitioning is necessary. Algorithms developed in a lab environment are typically very accurate, but must still be capable of being implemented on hundreds-of-thousands or even millions of endpoints.

8.5 Data Quality and Model Robustness

The quality of an AI model is only as good as the quality of the training data it was trained on. When there is missing information (missing values), or when sensor information changes over time (sensor drift) or if fault labels have been incorrectly applied to faulty equipment, the overall performance of the AI model can deteriorate in dangerous ways in control systems. In addition, in order to be reliable for use in control applications, the ability of a grid AI system to be robust to these types of issues, and also to gracefully degrade its performance when the model's confidence level has dropped below some threshold, are both essential but currently poorly addressed.

8.6 Explainability and Trust

Finally, the fact that the best models have opaque results has made it difficult to use them. The people who operate the systems, the regulators of those systems, and the end-users of those systems all want some level of transparency into how the system works before they will allow the system to make consequential decisions on their behalf. For this reason, we believe that advancing the explanation of artificial intelligence at the same time that performance improves (rather than waiting until after) will be key in allowing people to develop enough trust in these systems so that people will actually deploy them [33].

9. Future Research Directions

Drawing together these threads of review, there are several directions that stand out as both promising and necessary.

1. Federated learning with Privacy preservation and robustness: Advancing federated methods which simultaneously are accurate, communication efficient and resistant to poisoning and inversion attacks so collaborative intelligence does not come at the cost of either Privacy or security [15], [20], [22], [23].

2. Trustworthy and explainable AI: Developing interpretable models and post-hoc explanation techniques specifically for grid control to allow operators to safely understand, audit and act on ai recommendations during normal and contingency operation [33].

3. Digital twins as learning environments: Using high-faithful Digital twins to generate scarce fault and attack data to validate control policies safely before deployment and close the loop between simulation and live operation [31], [32].

4. Edge AI and 6G enabled control: Designing lightweight architectures and models that exploit next generation low latency communications to bring closed loop intelligence to the grid edge [34].

5. Blockchain enabled decentralised markets: Overcoming the obstacles of scalability, throughput and regulatory restrictions currently limiting peer to peer energy trading ideally in combination with forecasting and settlement driven by artificial intelligence [27]-[30].

6. Reinforcement learning: It is safe and sample efficient; making reliable enough to be used in critical infrastructure via constraints on safety, transfer learning and rigorous validation [17]-[19].

7. Hybrid (physics-based model / machine learning) A.I: It is sustainable optimisation; coordinating physics-based models, learning and metaheuristics to manage the extreme uncertainty introduced by very high renewable penetration and emerging vectors such as green hydrogen [6].

10. Conclusion

IoT (Internet of Things) and artificial intelligence (AI), working together, have transformed the conventional electric power distribution network from one based on physical processes into a smart, data driven, energy management system. The IoT provides the ubiquitous real-time sensor and actuators necessary for making the power distribution system visible and controllable; whereas AI provides the forecasting, optimization and autonomous control needed to convert that visibility into decision-making that results in improved performance. A PRISMA-based method was used for this review to synthesize recent studies and it demonstrated that machine learning approaches (in particular hybrid deep-learning and reinforcement-learning methods) consistently outperform traditional methods of load forecasting, renewable forecasting, fault diagnosis, predictive maintenance, and demand response. To be competitive with traditional methods, sufficient data must be available, as well as computational resources. Beyond centralized, single-task analytics, the research frontiers in AI-IoT applications now include distributed, privacy aware and security hardened architectures in which federated learning, blockchain-based trading, digital twin and explainable AI provide the key architecture components. However, the connectedness that generates value for the electric power distribution network presents significant challenges including cyber-security risks like false-data injection attacks, increased exposure to private information, interoperability issues between systems, increased costs, delay in processing data due to high volume of transactions generated per second and lack of transparency within complex black-box models. All of these can be addressed; however, they cannot be solved without the creation of new technologies that are inherently designed to preserve user-privacy, be transparent to users and stakeholders, provide strong assurance against intentional or unintentional failure (i.e., be robust), and be compatible with existing standards. Additionally, success requires continued cooperation and collaboration among researchers, utility companies, technology suppliers and policy makers.

References

1. X. Fang, S. Misra, G. Xue, and D. Yang, "Smart grid—The new and improved power grid: A survey," IEEE Commun. Surveys Tuts., vol. 14, no. 4, pp. 944–980, 2012.
2. H. Gharavi and R. Ghafurian, "Smart grid: The electric energy system of the future," Proc. IEEE, vol. 99, no. 6, pp. 917–921, 2011.
3. A. Ipakchi and F. Albuyeh, "Grid of the future," IEEE Power Energy Mag., vol. 7, no. 2, pp. 52–62, 2009.
4. G. Andersson, Modelling and Analysis of Electric Power Systems. Zurich, Switzerland: ETH Zürich, 2008.
5. R. Khan, A. Mahmood, A. Safdar, Z. A. Khan, and N. A. Khan, "Load forecasting, dynamic pricing and DSM in smart grid: A review," Renew. Sustain. Energy Rev., vol. 54, pp. 1311–1322, 2016.
6. M. SaberiKamarposhti, H. Kamyab, S. Krishnan, M. Yusuf, S. Rezania, S. Chelliapan, and M. Khorami, "A comprehensive review of AI-enhanced smart grid integration for hydrogen energy: Advances, challenges, and future prospects," Int. J. Hydrogen Energy, vol. 67, pp. 1009–1025, 2024.
7. M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," BMJ, vol. 372, art. n71, 2021.
8. B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," Keele Univ. and Univ. Durham, EBSE Tech. Rep. EBSE-2007-01, 2007.
9. M. U. Saleem, M. Shakir, M. R. Usman, M. H. T. Bajwa, N. Shabbir, P. Shams Ghahfarokhi, and K. Daniel, "Integrating smart energy management system with Internet of Things and cloud computing for efficient demand side management in smart grids," Energies, vol. 16, no. 12, art. 4835, 2023.
10. S. Hussain, C. Z. El-Bayeh, C. Lai, and U. Eicker, "Multi-level energy management systems toward a smarter grid: A review," IEEE Access, vol. 9, 2021.
11. S. A. A. Abir, A. Anwar, J. Choi, and A. S. M. Kayes, "IoT-enabled smart energy grid: Applications and challenges," IEEE Access, vol. 9, pp. 50961–50981, 2021.
12. S. Aggarwal, N. Kumar, S. Tanwar, and M. Alazab, "A survey on energy trading in the smart grid: Taxonomy, research challenges and solutions," IEEE Access, vol. 9, pp. 116231–116253, 2021.
13. N. Mounir, H. Ouadi, and I. Jrhilifa, "Short-term electric load forecasting using an EMD-BI-LSTM approach for smart grid energy management system," Energy Build., vol. 288, art. 113022, 2023.
14. N. Mughees, S. A. Mohsin, A. Mughees, and A. Mughees, "Deep sequence to sequence Bi-LSTM neural networks for day-ahead peak load forecasting," Expert Syst. Appl., vol. 175, art. 114844, 2021.
15. M. N. Fekri, K. Grolinger, and S. Mir, "Distributed load forecasting using smart meter data: Federated learning with recurrent neural networks," Int. J. Electr. Power Energy Syst., vol. 137, art. 107669, 2022.
16. J. R. Vázquez-Canteli and Z. Nagy, "Reinforcement learning for demand response: A review of algorithms and modeling techniques," Appl. Energy, vol. 235, pp. 1072–1089, 2019.
17. S. Bahrami, Y. C. Chen, and V. W. S. Wong, "Deep reinforcement learning for demand response in distribution networks," IEEE Trans. Smart Grid, vol. 12, no. 2, pp. 1496–1506, 2021.
18. E. J. Salazar, M. Jurado, and M. E. Samper, "Reinforcement learning-based pricing and incentive strategy for demand response in smart grids," Energies, vol. 16, no. 3, art. 1466, 2023.
19. N. Xu, Z. Tang, C. Si, J. Bian, and C. Mu, "A review of smart grid evolution and reinforcement learning: Applications, challenges and future directions," Energies, vol. 18, no. 7, 2025.
20. M. A. Husnoo, A. Anwar, N. Hosseinzadeh, S. N. Islam, A. N. Mahmood, and R. Doss, "A secure federated learning framework for residential short-term load forecasting," IEEE Trans. Smart Grid, early access, 2023.
21. M. Wen, R. Xie, K. Lu, L. Wang, and K. Zhang, "FedDetect: A novel privacy-preserving federated learning framework for energy theft detection in smart grid," IEEE Internet Things J., vol. 9, no. 8, pp. 6069–6080, 2022.
22. M. M. Badr, M. M. E. A. Mahmoud, Y. Fang, M. Abdulaal, A. J. Aljohani, W. Alasmary, and M. I. Ibrahem, "Privacy-preserving and communication-efficient energy prediction scheme based on federated learning for smart grids," IEEE Internet Things J., vol. 10, no. 9, pp. 7719–7736, 2023.
23. Z. Zhang, S. Rath, J. Xu, and T. Xiao, "Federated learning for smart grid: A survey on applications and potential vulnerabilities," arXiv preprint arXiv:2409.10764, 2024.

24. X. Lin, D. An, F. Cui, and F. Zhang, "False data injection attack in smart grid: Attack model and reinforcement learning-based detection method," Front. Energy Res., vol. 10, art. 1104989, 2023.
25. H. T. Reda, A. Anwar, and A. Mahmood, "Comprehensive survey and taxonomies of false data injection attacks in smart grids: Attack models, targets, and impacts," Renew. Sustain. Energy Rev., vol. 163, art. 112423, 2022.
26. Y. Han, H. Feng, K. Li, and Q. Zhao, "False data injection attacks detection with modified temporal multi-graph convolutional network in smart grids," Comput. Secur., vol. 124, art. 103016, 2023.
27. M. Waseem, Z. Adnan Khan, A. Khan, and I. A. Sajjad, "Incorporation of blockchain technology for different smart grid applications: Architecture, prospects, and challenges," Energies, vol. 16, no. 2, art. 820, 2023.
28. S. S. Uddin, M. S. H. Lipu, and S. Ansari, "Next-generation blockchain enabled smart grid: Conceptual framework, key technologies and industry practices review," Energy AI, vol. 12, art. 100228, 2023.
29. W. Hua, J. Chen, M. Qadrdan, J. Jiang, H. Sun, and J. Wu, "Applications of blockchain and artificial intelligence technologies for enabling prosumers in smart grids: A review," Renew. Sustain. Energy Rev., vol. 161, art. 112308, 2022.
30. T. Wang, H. Hua, Z. Wei, and J. Cao, "Challenges of blockchain in new generation energy systems and future outlooks," Int. J. Electr. Power Energy Syst., vol. 135, art. 107499, 2022.
31. M. Jafari, A. Kavousi-Fard, T. Chen, and M. Karimi, "A review on digital twin technology in smart grid, transportation system and smart city: Challenges and future," IEEE Access, vol. 11, pp. 17471–17484, 2023.
32. Z. Song, C. M. Hackl, A. Anand, A. Thommessen, J. Petzschmann, O. Kamel, R. Braunbehrens, A. Kaifel, C. Roos, and S. Hauptmann, "Digital twins for the future power system: An overview and a future perspective," Sustainability, vol. 15, no. 6, art. 5259, 2023.
33. T. Tsoka, X. Ye, Y. Chen, D. Gong, and X. Xia, "Explainable artificial intelligence for building energy performance certificate labelling classification," J. Cleaner Prod., vol. 355, art. 131626, 2022.
34. P. Varga, Á. I. Jászberényi, D. Pásztor, B. Nagy, M. Nasar, and D. Raisz, "How beyond-5G and 6G makes IIoT and the smart grid green—A survey," Sensors, vol. 25, no. 13, art. 4222, 2025.
35. C. F. Mbey, V. J. Foba Kakeu, A. T. Boum, and F. G. Yem Souhe, "Solar photovoltaic generation and electrical demand forecasting using multi-objective deep learning model for smart grid systems," Cogent Eng., vol. 11, no. 1, art. 2340302, 2024.
36. F. Zhang, Z. Leng, L. Chen, and Y. Zhang, "Joint probabilistic forecasting of wind and solar power exploiting spatiotemporal complementarity," Sustainability, vol. 17, no. 8, art. 3584, 2025.