energies

MDPI

Review

Machine Learning and Artificial Intelligence for Data-Driven Photovoltaic Power Systems: A Review

Yuxin Wu and Xueqian Fu *

College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China; 2023308130612@cau.edu.cn
* Correspondence: fuxueqian@cau.edu.cn

Abstract

At present, photovoltaic (PV) systems are becoming the core of low-carbon power systems, but their large-scale integration is still limited by weather-driven intermittency, heterogeneous data, equipment failures, operational uncertainty, and life-cycle sustainability requirements. Unlike specific task reviews that only focus on photovoltaic forecasting, fault diagnosis, or general artificial intelligence applications in renewable energy, this review develops an integrated data-driven perspective for machine learning and artificial intelligence in photovoltaic power generation systems. It links data governance, feature engineering, prediction, and uncertainty quantification, fault diagnosis and predictive maintenance, energy management, market participation, and carbon-aware optimization within a framework for photovoltaic systems. This review indicates that traditional machine learning, deep learning, graph learning, reinforcement learning, generative artificial intelligence, and physics-based artificial intelligence are suitable for different photovoltaic tasks based on data structure, time range, operational constraints, and deployment maturity. The main contribution is cross-task integration, which links the output of artificial intelligence models, including scheduling, storage scheduling, maintenance planning, virtual power plant operation, and low-carbon management, with actual decision-making. The review further identified the most critical deployment barriers, such as incomplete benchmarks, weak cross-site generalization, insufficient uncertainty calibration, limited interpretability, network security risks, and computational costs. The resulting methodological approach emphasizes data management, uncertainty awareness, physical constraints, decision orientation, and sustainability-driven photovoltaic intelligence.

Keywords: artificial intelligence; photovoltaic systems; PV forecasting; fault diagnosis; intelligent operation; energy management; low-carbon optimization

Check for updates

Academic Editor: Silvano Vergura

Received: 18 May 2026
Revised: 22 June 2026
Accepted: 29 June 2026
Published: 2 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Photovoltaics in Energy Transition

A general shift to low-carbon, renewable, and electrified configurations is taking place across the global energy sector, with solar photovoltaic generation standing out as one of the most rapidly developing renewable power technologies in this process [1]. The International Energy Agency indicates that renewable electricity capacity will continue to expand rapidly toward 2030, with solar PV serving as a major contributor to new renewable power additions [1]. The World Energy Outlook 2024 further shows that clean electricity is becoming increasingly central to future energy-system development as electrification and decarbonization accelerate [2]. From a climate-mitigation perspective, large-scale

renewable deployment is also a key pathway for reducing energy-related greenhouse gas emissions [1,2]. Therefore, PV generation is evolving from a supplementary clean-energy technology into a core resource for future low-carbon power systems.

The rapid expansion of PV is closely related to both policy support and economic competitiveness. IRENA reported that renewable power generation costs have continued to decline, and utility-scale solar PV has become one of the most cost-competitive electricity-generation options in many regions [3]. SolarPower Europe also documented strong growth in the global solar market, with both centralized and distributed PV systems contributing to capacity expansion [4]. Meanwhile, REN21 emphasized that renewable energy development increasingly depends on system flexibility, infrastructure coordination, and grid integration rather than capacity growth alone [5]. These trends indicate that the future value of PV will depend not only on installed capacity but also on the ability to integrate PV reliably into complex power systems.

However, high PV penetration also introduces operational uncertainty. The output power of photovoltaic systems is governed by solar radiation intensity, temperature, cloud cover, and other weather variables, thus exhibiting strong intermittency and instability [6]. Solar forecasting studies have shown that forecasting errors can influence dispatch schedules, reserve allocation, and grid-balancing decisions [6]. Reviews of PV power forecasting further emphasize that accurate prediction is essential for improving the grid friendliness and operational reliability of PV generation [7]. In addition, large-scale PV plants must satisfy increasingly advanced grid requirements, including voltage regulation, frequency support, active/reactive power control, and fault ride-through capability [8]. Thus, modern PV systems require intelligent methods that can support forecasting, uncertainty management, operational coordination, and secure grid integration.

Broadly speaking, solar photovoltaic resources have shifted their functional orientation in the energy transition, moving from simple clean electricity supply to in-depth participation in system planning, power dispatch, market operation, and low-carbon management. As PV systems become more distributed, data-rich, and operationally coupled with storage, loads, and grids, conventional monitoring and scheduling approaches are no longer sufficient. This motivates the introduction of AI as a key enabling technology for improving the perception, prediction, diagnosis, and optimization of PV systems.

1.2. Why Artificial Intelligence Is Needed for PV Systems

Traditional PV modeling methods have supported PV analysis for many years, but their limitations become increasingly evident under complex operating conditions. Physical models can describe PV output using irradiance, temperature, module parameters, and electrical characteristics, but their accuracy depends strongly on parameter availability and operating-condition assumptions. Empirical and statistical models are computationally efficient, yet many of them rely on simplified assumptions such as linearity, stationarity, or fixed functional relationships [7]. In practical PV plants, however, output power is influenced by weather fluctuations, partial shading, dust accumulation, aging, inverter behavior, and grid constraints [8]. Therefore, conventional methods are useful as baselines but are insufficient for fully representing the nonlinear, multi-source, and dynamic characteristics of modern PV systems.

AI provides a data-driven modeling paradigm that can complement physical and statistical methods. Relevant bibliometric research indicates a growing application trend of AI and machine learning in energy-system prediction, optimization, planning, and regulation [9]. Reviews of renewable energy applications also indicate that AI methods are effective in handling nonlinear relationships, heterogeneous data, and operational uncertainty [10]. Valuable feature patterns can be effectively extracted via machine learning

and deep learning from weather datasets, power generation records, electrical monitoring signals, image samples, and system operation records [11]. Therefore, AI is well-suited to PV systems because these systems generate diverse data streams and exhibit complex relationships between weather conditions, equipment status, and power output.

PV power forecasting is currently one of the most mature AI applications in PV systems. Deep learning-based renewable energy forecasting studies show that neural models can capture nonlinear temporal dependencies in renewable generation [12]. A large number of previous review studies in the solar PV domain indicate that convolutional and recurrent neural networks, along with LSTM, GRU, attention mechanisms, and Transformer variant models, have been broadly used in PV power prediction applications [13]. Solar PV forecasting reviews also show that data-driven methods can integrate historical power, meteorological variables, and numerical weather prediction information to improve forecasting performance [14]. Thus, AI-based forecasting provides a technical foundation for dispatch planning, storage control, market bidding, and uncertainty-aware operation [15].

AI is also becoming important for fault diagnosis and intelligent maintenance. Systematic reviews of PV fault identification show that AI methods can use electrical measurements, I–V curves, infrared images, and unmanned-aerial-vehicle inspection images to classify faults and identify abnormal states [16]. Fault-diagnosis reviews further demonstrate that AI can support anomaly detection, defect localization, degradation assessment, and condition monitoring in PV systems [17]. For utility-scale PV farms, these functions hold great practical significance since manual checking is expensive and unable to monitor all components continuously. So AI enables PV maintenance to move from periodic manual inspection toward predictive and condition-based maintenance.

Beyond forecasting and maintenance, AI can support PV control, energy management, and low-carbon operation. Existing literature reviews the implementation of AI in renewable energy systems, covering key areas such as renewable power output forecasting, resource coordination management, operational optimization, and systemic decision making [18,19]. In PV systems, these functions include maximum power point tracking, PV-storage coordination, microgrid energy management, virtual power plant operation, and carbon-aware dispatch. However, AI deployment also faces challenges such as weak interpretability, data-quality dependence, physical inconsistency, cybersecurity risk, and limited generalization across different regions or climates [20]. Therefore, AI should not be regarded merely as a tool for improving model accuracy, but as a broader framework for developing reliable, interpretable, and deployable intelligent PV systems.

To give readers a clearer entry point and to move beyond a descriptive list of methods, Figure 1 presents the roadmap and analytical logic of this review. The figure links PV data foundations, AI technique selection, application tasks, system-level decisions, and deployment challenges. It highlights how forecasting, uncertainty quantification, diagnosis, maintenance, energy management, and low-carbon optimization are connected within a data-driven PV intelligence framework.

1.3. Review Scope, Literature Methodology, and Contributions

Existing reviews have provided important foundations for AI-related PV research, but most of them remain task-specific. PV forecasting reviews mainly summarize forecasting horizons, input variables, evaluation metrics, and model categories [9,16], while deep learning reviews focus on neural architectures for PV prediction [13]. Fault-diagnosis reviews mainly discuss fault identification, classification, and maintenance support [18,19], and recent uncertainty reviews focus on probabilistic modeling, deep generative models, and scenario generation [21]. Compared with these studies, this review emphasizes an integrated PV power-system perspective that links data foundation, forecasting, uncer-

tainty representation, diagnosis, maintenance, operation management, and low-carbon optimization within one analytical framework.

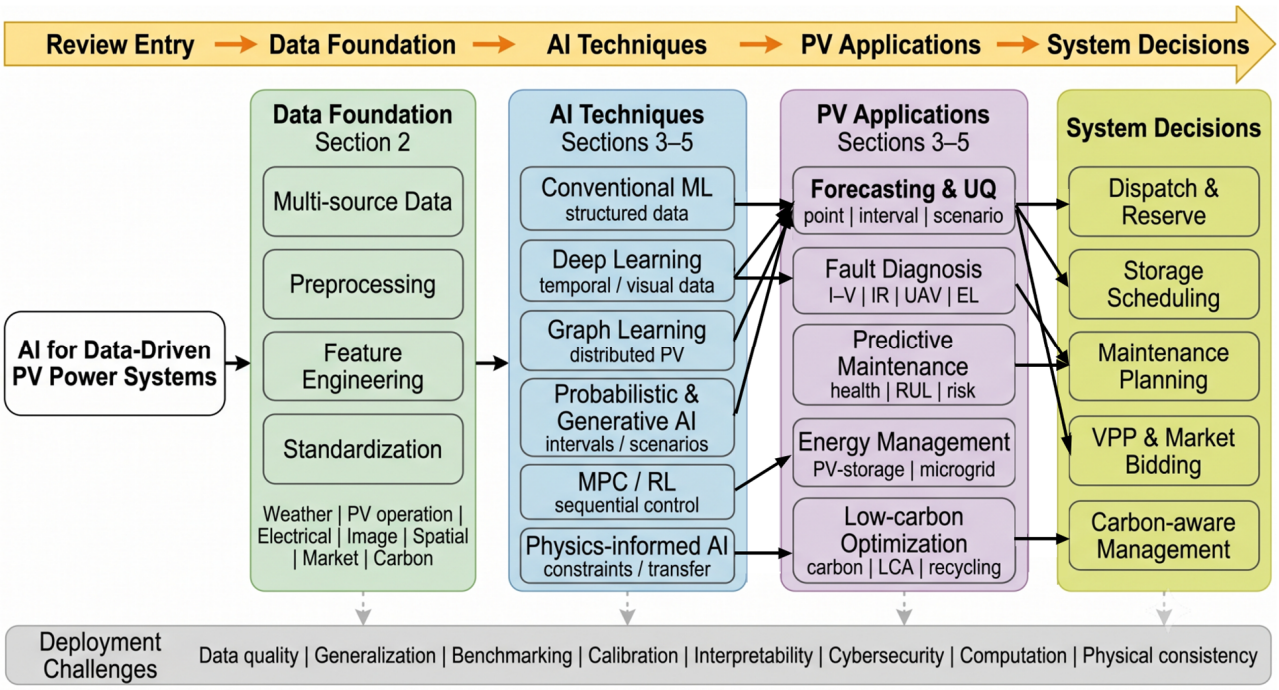


Figure 1. Roadmap and decision-oriented taxonomy of AI-enabled photovoltaic power systems.

This review focuses on machine learning and artificial intelligence for data-driven photovoltaic power systems and organizes the literature around the logic of "data foundation—AI methods—PV tasks—system decisions—deployment challenges." The data foundation includes meteorological data, PV operation data, electrical signals, image data, spatial information, market data, and carbon-related data. The core tasks include PV forecasting, uncertainty quantification, fault diagnosis, health assessment, intelligent maintenance, control, energy management, and carbon-aware optimization. The application scenarios include distributed PV systems, PV-storage systems, microgrids, virtual power plants, electricity markets, and life-cycle-oriented sustainable PV management. Thus, the review contribution is not another list of AI algorithms, but a system-level synthesis of how AI outputs are converted into forecasting, maintenance, operation, and low-carbon decisions.

The literature search was designed to improve transparency and reproducibility. Following PRISMA 2020 and PRISMA-S principles, this review reports the databases, search strings, retrieval period, screening rules, and selection numbers [20,21]. Literature was collected from Web of Science, Scopus, IEEE Xplore, ScienceDirect, SpringerLink, and Google Scholar for the period 2020–2026, while several classic and authoritative sources published before 2020 were retained when they provided fundamental methods, standards, or background evidence [22,23]. The main query string combined PV-related and AI-related terms, such as "photovoltaic" or "solar PV" or "PV power system" and "artificial intelligence" or "machine learning" or "deep learning" or "reinforcement learning" or "graph neural network" or "generative AI" and "forecasting" or "fault diagnosis" or "maintenance" or "energy management" or "microgrid" or "virtual power plant" or "carbon-aware optimization".

The search initially identified 742 records. After removing 130 duplicates, 612 records were screened by title and abstract. A total of 301 records were excluded because they were not related to PV systems, did not involve AI/ML methods, or lacked relevance to forecasting, diagnosis, operation, or low-carbon optimization. The remaining 311 full-text papers

were assessed for eligibility, and 113 were excluded because they lacked methodological detail, did not provide PV-specific evidence, or were outside the defined review scope. Finally, 198 studies were included and coded by data type, AI method, PV task, application scenario, evaluation evidence, deployment maturity, and major limitations.

The main contributions of this review are threefold. Firstly, it constructs a hierarchical classification system that links photovoltaic data resources, artificial intelligence method families, application tasks, and system-level decisions. Secondly, it introduces a thematic coding framework to compare artificial intelligence methods based on task types, data requirements, typical metrics, operating conditions, maturity, and deployment barriers, rather than just listing their advantages and disadvantages. Finally, it reached an engineering-oriented conclusion on when different artificial intelligence methods are more suitable: traditional machine learning for structured and medium-sized data, deep learning for temporal and visual representation, graph learning for distributed PV, reinforcement learning for sequential control, generative artificial intelligence for uncertainty and scene modeling, and physically informed artificial intelligence for physical consistency and transferable deployment. Figure 2 summarizes the literature screening and analysis workflow used in this review.

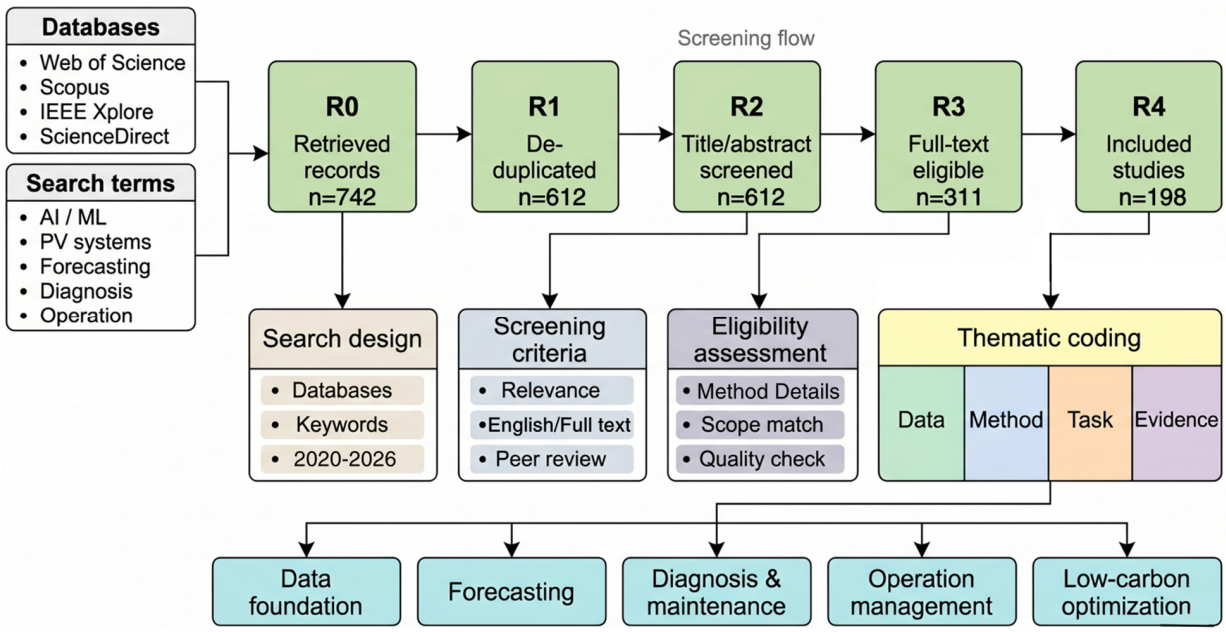


Figure 2. Literature screening and analytical workflow of this review.

The following three clear criteria will be used in the subsequent section to evaluate artificial intelligence methods, rather than just listing previous research.

(1) Each method is evaluated based on available PV data systems such as tabular meteorological and operational records, long time series, image or multimodal inspection data, spatially relevant multi-site data, and physically constrained simulation data.
(2) Compare these methods based on the output forms required for downstream decision-making, such as point prediction, calibration prediction interval, scenario trajectory, failure probability, health index, control policy, and carbon awareness plan.
(3) Deployment maturity is evaluated through computational burden, label cost, cross-site generalization, interpretability, network security risks, and compatibility with grid code and operational workflows.

2. Data Foundation and Feature Engineering for AI-Based PV Systems

2.1. Multi-Source Data in PV Systems

AI-based PV modeling depends on heterogeneous data that describe solar resources, meteorological conditions, system operation, equipment status, and external operating environments. As a pivotal input for PV-related analysis, solar irradiance can be acquired from large solar radiation repositories. The National Solar Radiation Data Base, in particular, provides long-period irradiance and meteorological records for solar resource assessment and photovoltaic model development [24]. Regional solar radiation databases have also been developed to estimate PV performance across different geographical conditions, indicating that spatially resolved solar resource data are essential for PV performance evaluation [25]. Reanalysis and satellite-derived datasets further allow researchers to construct long-term PV output series and analyze temporal variability over multi-year periods [26]. Therefore, solar resource and meteorological data form the first data layer for AI-based PV prediction and system assessment.

In addition to environmental data, PV operation and component-level data provide direct information about system behavior. Public PV monitoring platforms such as PVDAQ collect system metadata and time-series performance data from experimental and commercial PV systems, supporting data-driven studies of PV performance, degradation, and reliability [27,28]. Open-source modeling tools such as pvlib integrate solar-position calculation, irradiance transposition, temperature modeling, and PV system simulation, making it possible to transform environmental and system parameters into analyzable PV output series [29,30]. Classical PV performance models also show that irradiance, cell temperature, module parameters, and electrical characteristics are required to estimate PV array output accurately [31]. Similarly, the Sandia PV array performance model demonstrates that PV simulation requires the joint use of solar resources, weather, module, and inverter-related parameters [32]. Thus, AI-based PV models should not rely only on historical power values but should incorporate operation and physical-parameter data when available.

Visual and spatial data have become increasingly important as PV systems move toward short-term forecasting and intelligent monitoring. Studies on solar irradiance and PV forecasting show that historical power, meteorological variables, and multi-time-scale solar data are commonly used to describe temporal patterns in PV generation [33]. Recent reviews of computer-vision-based solar forecasting indicate that sky images, satellite images, and ground-based observations can capture cloud movement and irradiance variability that are difficult to represent using scalar meteorological variables alone [34]. Multimodal forecasting frameworks further show that sky images can be combined with meteorological data and Transformer-based models to improve ultra-short-term solar irradiance forecasting [33].

2.2. Data Preprocessing and Feature Construction

Raw PV data are rarely ready for direct AI modeling because they often contain missing values, outliers, sensor noise, inconsistent sampling intervals, and misaligned timestamps. Missing-value studies in day-ahead PV forecasting show that imputation methods can significantly affect prediction performance when meteorological or power generation data are incomplete [35]. PV production-data reconstruction studies further indicate that missing measurements can disrupt time-series continuity and reduce the reliability of data-driven models under constrained operating environments [36,37]. Researchers have also explored generative imputation algorithms, including WGAN-GP, to fill in missing PV power values and boost the availability of incomplete time-series datasets [38]. Therefore, data cleaning and missing-value imputation are necessary steps before training AI models for PV prediction, diagnosis, or operation.

Feature construction determines how raw data are transformed into meaningful model inputs. Feature selection studies based on kernel conditional density estimation show that identifying relevant input variables can reduce redundancy and improve PV forecasting efficiency [39]. Case studies combining feature selection with artificial neural networks further demonstrate that selecting highly correlated meteorological and operational variables can enhance PV power prediction performance [40]. Hybrid methods that combine random forest feature selection with signal decomposition also show that feature engineering can help extract short-term fluctuation patterns in PV generation [41]. It can be concluded from the above results that feature construction goes beyond simple data preprocessing and acts as a core modeling link with significant effects on accuracy and generalization ability.

For temporal and meteorological PV modeling, feature construction should reflect both physical drivers and time-dependent patterns. Reviews of machine learning-based solar radiation forecasting show that input-variable selection, time-lag construction, and meteorological preprocessing are important for improving model performance [42]. Hourly PV generation characteristics can be derived from numerical weather prediction data and further utilized in neural networks for solar energy decision-making [43]. Multimodal applications further demand image enhancement, cloud-motion description, and spatial feature extraction to derive effective AI inputs from sky and satellite remote sensing images [34,35].

2.3. Data Quality, Public Datasets, and Standardization

Although AI methods have advanced rapidly, their performance and reproducibility remain strongly constrained by data quality and dataset availability. Open-data studies for PV reliability show that environmental measurements, operation-monitoring data, image-inspection data, and material-level datasets are all valuable but remain unevenly available across PV research tasks [44]. Energy forecasting competitions such as GEFCom2014 demonstrate that public datasets and unified evaluation rules can promote fair comparison among forecasting models [45]. Broader reviews of energy forecasting also emphasize that open data, reproducible benchmarks, and standardized metrics are critical for evaluating forecasting methods objectively [46]. Therefore, the development of AI-based PV research depends not only on algorithms but also on accessible, high-quality, and well-documented data resources [47].

Several public platforms provide useful foundations for PV modeling, but their purposes and data coverage differ. The DKA Solar Centre provides downloadable PV system and meteorological data that can support PV performance analysis and forecasting studies [48]. The Photovoltaic Geographical Information System offers solar resource and PV-potential information for geospatial PV assessment in Europe and other regions [49]. The Global Solar Atlas provides global solar resource and PV-potential data for regional assessment, planning, and site-selection analysis [50]. These systems support data-driven PV analysis at diverse spatial and temporal scales, while variations in variable settings, time resolution, metadata quality, and data access terms remain obstacles to direct comparison among existing studies.

For this reason, data standardization lays a necessary foundation for subsequent AI-enabled PV studies. As defined by the FAIR principles, scientific data should follow findable, accessible, interoperable, and reusable criteria, which form a basic framework to enhance PV data governance and sharing mechanisms [47]. For PV applications, this implies that datasets should include clear metadata, standardized units, transparent preprocessing descriptions, and consistent evaluation protocols. Without such standardization, model performance may reflect dataset-specific artifacts rather than generalizable AI capability. Consequently, before discussing forecasting, diagnosis, and operation models

in Section 3, it is necessary to recognize that data quality, feature design, and dataset standardization jointly determine the reliability and transferability of AI-based PV systems.

Based on the above data discussion, Table 1 has been organized as a comparison of task-oriented photovoltaic data requirements. It not only describes the data source but also emphasizes which artificial intelligence methods are more desirable under each data condition, and which obstacles may limit on-site deployment.

Table 1. Task-oriented PV data requirements, suitable AI methods, and deployment barriers.

Refs.	Data Condition	Preferable AI Use	Engineering Judgment and Deployment Barriers
[26–28,33,49,50]	Irradiance, temperature, wind, cloud cover, and NWP/reanalysis data; imputation and time alignment.	Forecasting and UQ; ML for tabular data, LSTM/Transformer for sequential or multivariate inputs.	High value for dispatch and reserves; limited by NWP bias, missing data, and spatial mismatch.
[29–32,44,48]	PV power, voltage, current, inverter logs, and I-V/P-V curves; metadata and sensor checks.	Anomaly/fault/health assessment; SVM/RF for engineered features, autoencoders, or PINN for sparse labels/physics.	Deployable in monitored plants; limited by sensor calibration, metadata quality, and fault-label scarcity.
[33,34,44]	IR, UAV, EL, sky, and satellite images; quality checking, annotation, and registration.	Defect localization and cloud-motion tasks; CNN/YOLO/U-Net/ViT and transfer or multimodal learning.	Strong localization; limited by labeling cost, image-quality variation, and domain shift.
[25,26,49,50]	Site location, terrain, regional irradiance, resource potential, and topology/correlation information.	Regional and distributed PV forecasting; GNN/GAT when site relations or grid topology are stable.	Useful for multi-site PV; limited by coarse spatial resolution, regional bias, and grid-data gaps.
[29,44,46,48]	Fault logs, maintenance records, failure time, inspection reports, and replacement history.	Predictive maintenance and RUL; survival models, RF/boosting, LSTM, and Bayesian learning.	Direct maintenance value; limited by sparse faults, class imbalance, and inconsistent O&M records.
[45–47]	Public benchmarks, metadata, standardized units, and repeatable train-test protocols.	Reproducible comparison, benchmark learning, transfer learning, and federated learning.	Improves transparency; limited by missing PV-specific variables, privacy issues, and uneven metadata.
[30–32,38]	pvlib/digital twins, reconstructed data, synthetic scenarios, and rare-fault samples.	Imputation, data augmentation, and scenario generation; GAN/VAE/diffusion when real data are scarce.	Useful for stress testing; must verify physical consistency, capacity limits, and simulation-reality gap.

Building on the data categories summarized in Table 1, Figure 3 presents the general pipeline through which multi-source PV data are preprocessed, engineered, and fused into model-ready inputs for different AI tasks.

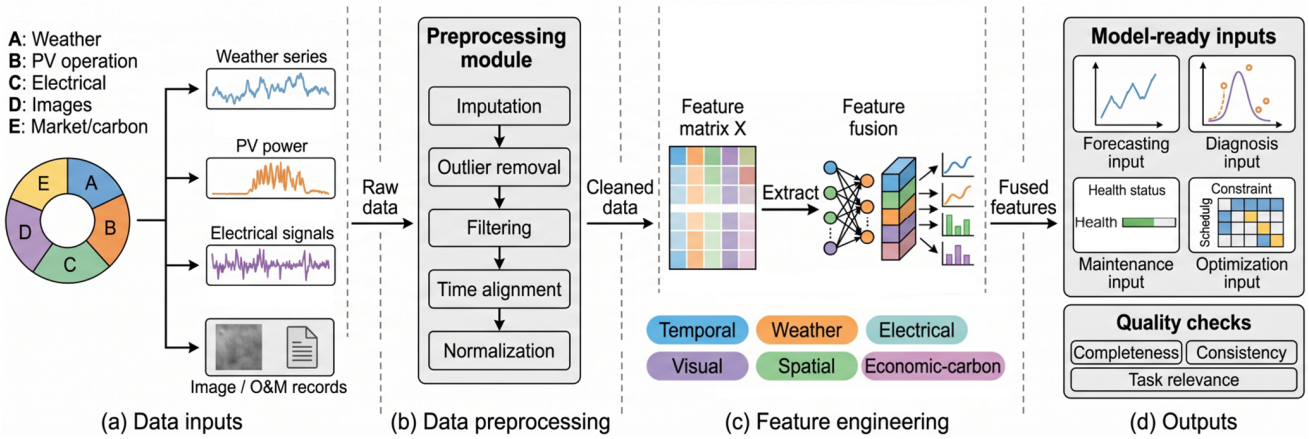


Figure 3. Multi-source PV data processing and feature engineering.

3. AI-Based PV Forecasting and Uncertainty Quantification

3.1. Forecasting Tasks and Time Horizons

PV power forecasting is a core task in AI-enabled PV systems because it transforms weather-dependent and intermittent generation into actionable information for scheduling, storage control, and grid operation. Solar forecasting studies commonly classify forecasting tasks according to time horizon, data source, and application objective, because different horizons correspond to different operational decisions [6]. Reviews of PV power forecasting further show that forecasting accuracy directly affects unit commitment, reserve allocation, grid balancing, and PV integration performance [7]. Therefore, PV forecasting should be understood not as a single prediction problem, but as a group of time-scale-dependent decision-support tasks.

PV power forecasting can be classified by time horizon into four categories: ultra-short-term, short-term, medium-term, and long-term forecasting. Covering a timescale from seconds to hours, ultra-short-term prediction mainly serves real-time power dispatch, ramp-rate management, energy storage scheduling, and local power grid regulation [6]. Short-term forecasting covers the timescale from hours up to two days ahead, and it is widely applied in day-ahead scheduling, electricity market transactions, and energy allocation optimization [14]. Medium- and long-term forecasting are more relevant to maintenance planning, resource assessment, investment analysis, and capacity expansion [51]. This time-scale classification is important because it determines whether the model should rely more on real-time measurements, sky images, numerical weather prediction, historical generation, or long-term climate data [33,52].

PV forecasting can also be classified according to output form. Point forecasting provides a single expected value and is widely used because of its simplicity and compatibility with conventional dispatch models [7]. However, point forecasts cannot fully represent the uncertainty caused by cloud movement, weather forecast errors, sensor noise, and model uncertainty [46]. Probabilistic forecasting provides prediction intervals, quantiles, or probability distributions, making it more suitable for risk-aware dispatch and renewable energy integration [46]. Scenario-based forecasting further generates possible future trajectories of PV output, which can be used in stochastic optimization, storage scheduling, and market-risk analysis [53]. Thus, the evolution of PV forecasting is moving from deterministic power prediction toward uncertainty-aware decision support.

The selection of forecasting input variables is closely related to both the time horizon and the forecasting target. Solar radiation forecasting studies show that meteorological variables, time features, and historical solar data are fundamental inputs for machine learning-based irradiance prediction [42]. Relevant literature regarding PV forecasting

confirms that researchers commonly employ historical PV production, irradiance data, temperature records, cloud conditions, and numerical weather forecasts to depict time-dependent changes in photovoltaic output [16,17]. Apart from that, the literature reviews adopting text mining techniques for solar irradiance and photovoltaic forecasting indicate that scholarly interest is gradually moving beyond basic statistical models toward data-driven hybrid strategies and AI-enabled methodologies [33]. Therefore, the forecasting task provides the methodological bridge from the data foundation discussed in Section 2 to the AI models discussed in the following subsections.

3.2. Machine Learning and Deep Learning Techniques for Time Series Forecasting

Conventional machine learning methods remain indispensable for PV temporal forecasting owing to their efficiency and interpretable baseline performance. Common machine learning approaches such as support vector regression, random forest, gradient boosting, and artificial neural networks are prevalently employed in PV output and solar irradiance forecasting research [17,42]. Recent PV forecasting reviews also show that machine learning models are effective when historical power and meteorological variables are available, but the dataset size is limited [51]. Studies using actual grid-connected PV operating data further confirm that machine learning algorithms can capture useful relationships between weather conditions and PV production. Therefore, traditional machine learning remains suitable for lightweight PV forecasting tasks, especially when interpretability and computational efficiency are important.

Proper feature selection and signal preprocessing are key to optimizing the overall performance of ML-based PV prediction systems. A KCDE-based feature selection method demonstrated that selecting relevant input variables can reduce redundancy and improve PV forecasting efficiency [39]. A case study combining feature selection with artificial neural networks showed that carefully selected meteorological and operational variables can enhance PV power prediction accuracy [40]. Random forest feature selection combined with CEEMD was also used to extract meaningful short-term fluctuation patterns from PV generation time series [41]. Numerical-weather-data-driven neural forecasting further indicates that transforming NWP variables into model-ready features can support hourly PV production prediction [43]. Related studies demonstrate that machine learning prediction accuracy depends on model configuration as well as the quality of feature construction.

Deep learning methods have become increasingly important because PV generation exhibits nonlinear and time-dependent patterns. Relevant overview studies on deep learning in PV forecasting reveal that models built on CNN, RNN, LSTM, GRU, attention mechanisms, and Transformer frameworks are widely researched for photovoltaic power prediction [13]. LSTM can effectively learn temporal correlations from sequential inputs, making it a commonly used approach for solar PV output forecasting [54]. Hybrid LSTM-convolutional networks further combine local pattern extraction and temporal memory, improving the representation of PV power fluctuations [55]. CNN, LSTM, and GRU-based models are therefore particularly useful when PV output depends on both short-term weather changes and previous generation states.

Recent studies increasingly use hybrid and attention-based architectures to improve forecasting robustness. TCN-ECANet-GRU integrates temporal convolution operations, attention mechanisms, and GRU networks for the task of short-term PV power prediction [56]. To boost short-term photovoltaic prediction performance across diverse operational modes, researchers have integrated K-means clustering, ensemble learning, and feature expansion methods [57]. Researchers have further introduced Transformer-based models to photovoltaic prediction tasks, benefiting from the ability of attention mechanisms to mine long-distance correlations in multivariate temporal sequences [58]. Transformer-enabled

ultra-short-term PV forecasting highlights the great potential of self-attention in modeling fast-changing PV power dynamics [59]. Therefore, deep learning is gradually shifting PV forecasting from shallow feature mapping toward long-sequence and multi-scale temporal representation.

Despite these advances, deep learning is not always superior in every forecasting setting. LSTM-based forecasting on limited datasets indicates that data availability can strongly affect model performance [60]. Comparative studies of forecasting techniques and optimization methods also show that model performance depends on data quality, forecasting horizon, weather variability, and hyperparameter tuning [61,62]. Meta-learning has been proposed to dynamically combine multiple forecasting models and improve short-term PV forecasting generalization across changing conditions [63]. Multi-step Transformer forecasting has also been explored for day-ahead PV power prediction, further indicating that attention-based models can be extended from ultra-short-term prediction to longer-horizon forecasting tasks [64]. Therefore, temporal PV forecasting should not be reduced to selecting the most complex model; rather, the model should be matched with data scale, forecasting horizon, computational constraints, and application requirements.

From a practical application perspective, the selection of photovoltaic forecasting methods cannot rely solely on the error metrics obtained from a single dataset. Instead, a comprehensive evaluation integrating data volume, forecast time horizon, types of input variables, computational resources, and dispatching requirements is necessary. For datasets with high structural integrity and a moderate number of samples, traditional machine learning approaches, including support vector regression, random forest, gradient boosting, and shallow neural networks, are generally easier to deploy and can deliver stable baseline performance. When short-term temporal dependencies dominate the forecasting challenge, LSTM, GRU, and temporal convolutional networks excel at capturing the continuous variation characteristics of photovoltaic power output. If historical power records, numerical weather predictions, calendar information, and multi-dimensional meteorological variables need to be processed simultaneously, Transformer-based models demonstrate superior performance in multivariate long-sequence modeling. Graph learning methods are more applicable to distributed photovoltaic forecasting or multi-station prediction scenarios, provided that clear geographic, meteorological, or electrical correlations exist among power stations. In the absence of credible spatial relationships, arbitrarily constructing graph structures may only increase model complexity without significantly improving forecasting performance for dispatching decision-making.

3.3. Spatiotemporal and Graph-Based Forecasting for Distributed PV

With the rapid growth of distributed PV, forecasting is increasingly shifting from single-site prediction to regional and multi-site forecasting. Long-term PV output studies based on reanalysis and satellite data show that regional PV generation exhibits clear spatial and temporal patterns [26]. Relevant literature reviews reveal an increasing research emphasis on spatial features, satellite observations, and multi-location datasets in the field of solar irradiance and PV power prediction [33]. Computer-vision-based solar forecasting further demonstrates that sky images and satellite images can capture cloud movement and spatial irradiance variability that cannot be fully described by single-site scalar variables [34]. Therefore, distributed PV forecasting requires models that can jointly represent temporal dynamics and spatial correlations.

For multi-site PV forecasting tasks, graph-structured learning serves as an intuitive paradigm: each photovoltaic station is abstracted as a node, while the inter-station relevance is defined as graph edges. To jointly mine temporal evolution and spatial connections among PV plants, scholars have proposed spatiotemporal graph neural networks for multi-

location power forecasting [65]. Interpretable temporal-spatial graph attention networks further show that attention mechanisms can identify important temporal and spatial contributions in multi-site PV prediction [66]. Explainable spatiotemporal GNNs have also been used for multivariate solar power forecasting, indicating that graph-based models can improve both predictive performance and interpretability [67]. Thus, graph learning is particularly suitable for distributed PV systems where geographic distance, weather propagation, and site correlation are important.

Recent studies have further extended graph learning toward more complex spatial and multimodal forecasting settings. DEST-GNN was proposed to explore spatiotemporal dependencies more deeply in multi-site PV power forecasting [68]. Literature examining distributed solar prediction based on remote sensing and deep learning indicates that regional photovoltaic forecasting now leans heavily on fusing remote sensing technology, deep learning approaches, and spatial modeling tools [69]. Combining satellite-ground observational data with hybrid neural network methods aids regional PV forecasting and illustrates the advantages of multi-source spatial observation integration [70]. Multimodal-learning frameworks using sky images further show that visual data can improve ultra-short-term solar irradiance prediction by representing cloud dynamics [33]. Therefore, spatiotemporal forecasting forms a key extension from conventional temporal prediction to networked and distributed PV intelligence.

However, spatiotemporal forecasting also introduces additional modeling challenges. Graph structures may be constructed according to geographic distance, meteorological similarity, power output correlation, or electrical topology, and different definitions can lead to different forecasting results [65,66]. Visual and remote sensing data can enrich spatial representation, but they also introduce image-quality variation, cloud-motion uncertainty, and multimodal alignment problems [34,69]. Consequently, the success of graph-based PV forecasting depends on both model architecture and the quality of spatial data construction. This reinforces the importance of the data and feature engineering pipeline discussed in Section 2.

3.4. Probabilistic Forecasting and Scenario Generation

Although deterministic PV forecasting is useful for estimating expected generation, it cannot fully support risk-aware operation under high renewable penetration. Relevant literature reviews concerning probabilistic PV prediction indicate that it is necessary to adopt prediction intervals and probability distributions to quantify the uncertainties inherent in PV output and electricity usage [71]. As probabilistic solar prediction is increasingly integrated into power grids, it has been proven that forecasts considering uncertainty are able to support operational scheduling, reserve allocation, storage regulation, and risk management practices [72]. Weather-scenario-generation methods provide another way to propagate meteorological uncertainty into solar power forecasts and represent possible future weather-driven PV trajectories [73]. Therefore, uncertainty quantification is a necessary extension of AI-based PV forecasting.

Probabilistic PV forecasting can be implemented through several methodological families. Quantile regression is commonly used to estimate conditional quantiles and construct prediction intervals for renewable generation and electrical load forecasting tasks [74]. Scenario quality assessment studies further emphasize that generated renewable energy scenarios should preserve statistical properties, temporal dependence, and practical decision value [75]. More general studies of uncertainty and probabilistic forecasting in power systems also indicate that uncertainty representation is essential for decision-making under variable renewable generation [76]. These studies suggest that PV forecasting outputs

should move from single deterministic values toward intervals, probability distributions, and decision-oriented scenarios.

Scenario generation provides a further step by producing multiple possible future PV trajectories. Literature reviews regarding solar power scenario generation indicate a clear evolution from statistical methods toward deep learning-based solutions [77]. Comprehensive literature reviews covering stochastic scenario generation of wind and PV power reveal that generated scenarios are essential for stochastic scheduling, storage layout planning, and the integration of renewable resources [78]. Deep generative models have also been applied to renewable energy scenario generation because they can learn complex distributions and generate diverse future trajectories [79]. Therefore, scenario generation should be evaluated not only by visual similarity but also by whether the generated trajectories can support reliable operational decisions.

Recent probabilistic and generative approaches further expand the modeling capacity of uncertainty-aware PV forecasting. Neural network-based lower–upper bound estimation provides a direct way to construct prediction intervals without assuming a fixed output distribution [80]. Researchers have begun applying conformal prediction to ultra-short-term PV forecasting, delivering reliable calibrated uncertainty assessments within data-driven frameworks [81]. Diffusion probabilistic models have also been introduced for probabilistic energy forecasting, providing a new approach for modeling complex uncertainty distributions [82]. Comparative studies of deep generative models show that GANs, VAEs, normalizing flows, and diffusion models each have different strengths in probabilistic energy forecasting [83]. Thus, generative AI expands PV forecasting from error estimation to data-driven uncertainty representation and scenario construction.

Nevertheless, probabilistic forecasting and scenario generation must satisfy both statistical and physical requirements. Generated PV scenarios should respect daily solar patterns, capacity limits, ramping behavior, and weather-driven temporal dependence [75,78]. Probabilistic forecasts should be evaluated using reliability, sharpness, calibration, and decision-oriented usefulness rather than point-error metrics alone [71,76]. As shown in Figure 4, AI-based PV forecasting should be understood as a pipeline that transforms multi-source inputs into point forecasts, probabilistic forecasts, and scenario trajectories for downstream operation. Table 2 further compares representative AI methods for PV forecasting in terms of input data, model characteristics, forecasting outputs, advantages, and limitations. The following section extends the discussion from PV generation prediction to fault diagnosis and intelligent maintenance.

Probabilistic photovoltaic forecasting results that incorporate uncertainty are not merely used to evaluate forecasting errors. Instead, they can directly support subsequent operational decision-making. Prediction intervals enable operators to determine the margin of reserve capacity and the safe buffer range for the state of charge of energy storage systems. For instance, lower quantiles serve as conservative power output estimates for dispatching scenarios with stringent reliability requirements, while higher quantiles help identify potential surplus power generation and photovoltaic curtailment risks. Complete probability distributions further facilitate chance-constrained dispatch, reliability assessment, and risk-aware market bidding, allowing operational strategies to balance economic gains against risks.

In contrast, time-series scenario curves are better suited for stochastic optimization and model predictive control. This is because energy storage scheduling, virtual power plant bidding, and system dispatch schemes can all be validated under multiple plausible photovoltaic power output trajectories. Accordingly, the value of probabilistic forecasting should be assessed not only by interval width or distribution accuracy but also by its effec-

tiveness in reducing operational costs, risks of insufficient reserve capacity, photovoltaic curtailment risks, and decision-making uncertainty.

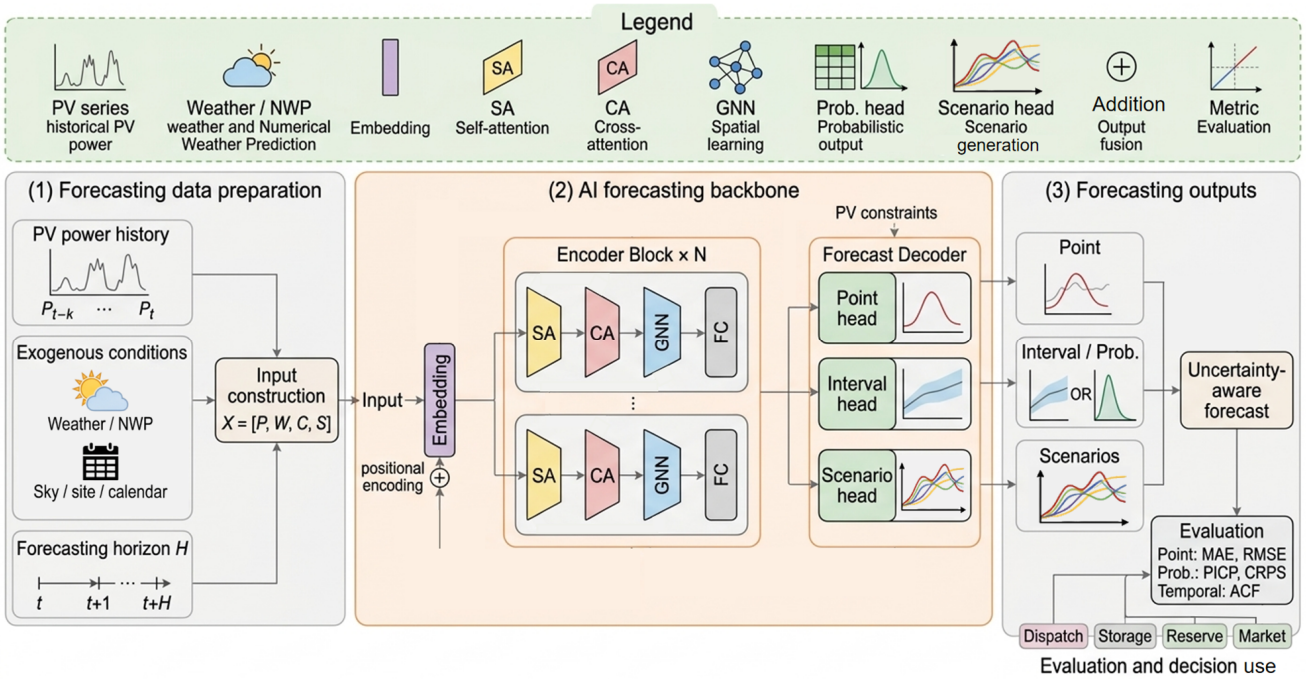


Figure 4. AI-based PV forecasting and uncertainty-aware output forms.

Table 2. Benchmarking AI Models for Solar PV Forecasting and Uncertainty Assessment.

Refs.	Method Category	Representative Models	Forecasting Target	Strengths	Limitations	Suitable Scenarios
[6,7,14,17,33,51–53]	General PV forecasting frameworks	Statistical models, ML, DL, hybrid models	Point forecasting, horizon-based forecasting	Provides task classification and baseline methodology	Often broad; limited task-specific depth	Review framework, task definition, method classification
[42,51,53,61,62]	Traditional machine learning	SVR, RF, XGBoost, KNN, ANN	Short-term PV power forecasting	Lightweight, efficient, relatively interpretable	Relies on feature engineering; limited sequence modeling	Small- to medium-scale datasets; engineering baseline models
[39–41,43]	Feature selection and decomposition-based methods	KCDE, RF feature selection, CEEMD, NWP-based neural models	Short-term and hourly PV forecasting	Reduces redundancy and improves input quality	Performance depends on feature design and preprocessing	Feature-driven forecasting; NWP-assisted prediction
[13,54,55,60]	Recurrent and hybrid deep learning	LSTM, CNN-LSTM, limited-data LSTM	Temporal PV power forecasting	Captures nonlinear temporal dependencies	Data-demanding; may overfit under limited samples	Short-term and ultra-short-term PV forecasting
[56,57,63]	Ensemble and hybrid deep models	TCN-ECANet-GRU, clustering ensemble, meta-learning	Short-term PV forecasting	Improves robustness under changing patterns	More complex training and tuning	Variable weather conditions; multi-model fusion tasks

Table 2. *Cont.*

Refs.	Method Category	Representative Models	Forecasting Target	Strengths	Limitations	Suitable Scenarios
[58,59,64]	Transformer-based forecasting	Transformer, TransPVP, multi-step Transformer	Ultra-short-term and multi-step PV forecasting	Captures long-range temporal dependencies	High computational cost; needs sufficient data	Long-sequence, multivariate, and multi-step prediction
[26,33,34,65–70]	Spatiotemporal and graph-based forecasting	GNN, GAT, ST-GNN, Graph Transformer, satellite-ground fusion	Distributed and regional PV forecasting	Captures spatial correlation and weather propagation	Graph construction and multimodal alignment are difficult	Multi-site PV plants; distributed PV; regional forecasting
[71–74,76,80,81]	Probabilistic forecasting	Quantile regression, BNN, prediction intervals, conformal prediction	Interval and probabilistic forecasting	Represents forecast uncertainty and risk	Requires calibration and probabilistic evaluation	Risk-aware dispatch, reserve allocation, storage control
[21,73,75,77–79]	Scenario generation	Weather scenarios, stochastic scenarios, deep generative models	Future PV trajectories	Supports stochastic optimization and scenario-based decisions	Scenario quality and physical consistency are hard to ensure	Storage scheduling, market bidding, renewable integration
[82,83]	Advanced generative forecasting	GAN, VAE, normalizing flow, diffusion models	Probabilistic energy forecasting and scenario generation	Learns complex distributions and diverse trajectories	High model complexity; validation remains challenging	Advanced uncertainty modeling; data-driven scenario generation
[46,71,75,76]	Evaluation and uncertainty assessment	MAE, RMSE, MAPE, CRPS, PICP, PINAW, reliability, sharpness	Forecast quality evaluation	Enables fair comparison of deterministic and probabilistic models	Metrics may not fully reflect decision value	Benchmarking, model selection, uncertainty-quality assessment

4. AI for PV Fault Diagnosis and Intelligent Maintenance

4.1. Fault Types and Monitoring Data

After PV forecasting estimates future generation uncertainty, high-precision fault diagnosis becomes essential for sustaining the efficiency, operational reliability, and safety of PV systems. Existing reviews show that PV faults may occur at the cell, module, string, inverter, and system levels, and that different faults require different monitoring evidence and maintenance responses [16–19].

Different PV faults require different monitoring data. Electrical faults are commonly reflected in voltage, current, power, and I–V curve variations, while module-level defects and thermal anomalies are often detected through infrared, electroluminescence, visible-light, or UAV inspection images [84,85]. Long-term degradation studies confirm that PV modules experience a gradual performance decline under field operating conditions, which makes degradation monitoring necessary for reliability assessment [86]. IEA PVPS reports on PV module failures also show that field failures include cell cracks, delamination, discoloration, glass breakage, junction-box failures, and interconnection defects [87,88]. Thus, a comprehensive diagnosis system should integrate electrical, visual, environmental, and maintenance data rather than depend on a single data source.

Monitoring data also determines the granularity and accuracy of AI-based diagnosis. Open PV reliability datasets show that useful diagnostic evidence may come from

environmental measurements, operation-monitoring records, image-inspection data, and material-level information [44]. Infrared thermography and digital image processing have been used to detect degradation effects in field-aged crystalline-silicon solar cells, showing that thermal and visual information can reveal faults that are difficult to identify from power signals alone [89]. High-temperature exposure studies further indicate that thermal stress can accelerate reliability problems in PV modules [90]. Therefore, the diagnostic value of data lies not only in detecting faults but also in linking fault symptoms with their physical and operational causes.

From a system perspective, fault diagnosis should be understood as a process that connects monitoring, abnormal state recognition, fault localization, and maintenance decision-making. Early degradation studies show that performance loss may appear before obvious system failure, which requires continuous monitoring rather than post-failure inspection [91]. AI-based diagnostic reviews also emphasize that the field is moving from manual rule-based detection toward automated classification, localization, and condition assessment [18,19]. Consequently, this section first discusses sensor- and signal-based diagnosis, then image-based detection, and finally predictive maintenance and health assessment.

4.2. Sensor- and Signal-Based Diagnosis

By utilizing measured operating parameters like voltage, current, power, temperature, and I–V characteristics, sensor and signal-based approaches can recognize abnormal operating states of PV modules. Voltage and current measurement-based methods serve for PV fault diagnosis, since electrical discrepancies effectively reveal mismatch, open-circuit, short-circuit, and partial shading anomalies [92]. Reviews of PV fault diagnosis also show that electrical signals are widely used because they can be obtained from inverters, string monitors, SCADA systems, or module-level devices [84]. Therefore, signal-based diagnosis is practical for online monitoring and can be integrated into existing PV operation platforms.

I–V characteristics provide richer information than single-point voltage or power measurements. Full I–V curves have been used with machine learning techniques to diagnose PV panel faults, because changes in the curve shape can reflect shading, degradation, mismatch, and other abnormal conditions [93]. Recent work has further used I–V characteristics with random forest-based fault detection and diagnosis, confirming that curve-based features can improve fault discrimination [94]. Since I–V curves embed both electrical behavior and operating-state information, they are particularly suitable for fault classification and parameter-related diagnosis.

Machine learning methods are widely adopted for signal-based diagnosis because they can classify complex nonlinear fault patterns. For electrical fault diagnosis and identification of PV arrays, SVM and XGBoost models have been widely employed, verifying the capability of supervised learning to classify different fault types [95]. Modulated photocurrent combined with machine learning has also been used for PV array fault diagnosis and localization, indicating that enriched electrical measurements can improve spatial fault identification [96]. Stacking ensemble learning has been proposed for PV array fault diagnosis, showing that ensemble models can improve classification robustness by combining multiple learners [97]. These studies suggest that conventional machine learning remains effective when electrical features are carefully constructed and labeled data are available.

Deep learning and semi-supervised methods are increasingly used to reduce dependence on handcrafted features and complete labels. Leveraging variational autoencoders, semi-supervised deep learning can realize automatic fault detection for grid-connected PV systems and shows satisfactory capacity for anomaly recognition when labeled data are limited [98]. Literature surveys concerning the application of artificial neural networks to

PV fault diagnosis demonstrate that MLP, PNN, RBF, CNN, stacked autoencoders, and other analogous neural architectures have drawn extensive research attention for feature mining and fault categorization [99]. Stacked sparse autoencoder integrated hybrid deep learning techniques can autonomously mine diagnostic features out of PV-related data [100]. Adaptive deep belief networks and LightGBM-based frameworks further show that combining deep feature learning with efficient classifiers can enhance PV fault diagnosis [101]. Therefore, deep models expand signal-based diagnosis from manual feature classification toward automated representation learning.

However, sensor- and signal-based diagnosis still faces practical limitations. Electrical measurements are sensitive to sensor noise, irradiance fluctuation, operating-point variation, and incomplete fault labels [16]. Reviews of PV fault detection techniques also indicate that similar electrical symptoms may correspond to different physical faults, which can reduce classification reliability under real operating conditions [85]. Therefore, signal-based diagnosis is efficient and deployable, but it often needs to be complemented by image-based inspection and maintenance records for more accurate fault localization and severity assessment.

4.3. Image-Based Fault Detection and Localization

Image-based diagnosis has become an important complement to electrical signal diagnosis because it can directly reveal spatially distributed defects in PV cells and modules. Reviews of electroluminescence images show that EL inspection is widely used to detect cracks, inactive regions, broken fingers, and other cell-level defects that are difficult to observe during normal operation [102]. Reviews of deep learning for PV module defect detection further indicate that image-based AI methods can extract defect features from EL, infrared, and visible-light images and support automatic visual inspection [103]. Therefore, image-based diagnosis provides a direct visual basis for fine-grained PV fault detection and localization.

EL image-based diagnosis is especially suitable for cell-level and module-level defect recognition. Deep learning-driven automatic inspection methods serve for PV cell defect recognition in EL images, and convolutional models perform well in capturing defect characteristics from image databases [104]. Automatic defect-identification methods using EL images also demonstrate that deep learning can reduce manual inspection dependence and improve the consistency of PV module defect recognition [105]. Thus, EL imaging is particularly valuable for detecting internal defects and structural abnormalities that cannot be easily captured by conventional electrical measurements alone.

Infrared thermography provides another important visual modality for PV fault detection. Transfer learning-based CNNs have been used to classify PV module faults from infrared images, indicating that pretrained networks can improve fault recognition when task-specific thermal image samples are limited [106]. Deep learning models based on infrared thermographic images have also been used to detect hotspots in PV modules, showing their ability to identify abnormal thermal patterns [107]. Deeplab-YOLO further combines segmentation and detection for hotspot defects in infrared PV images, suggesting that segmentation can help focus detection on relevant panel regions [108]. These studies indicate that infrared image diagnosis can link thermal symptoms with spatial fault localization.

Recent object detection, segmentation, and multimodal vision models have further expanded the capability of image-based PV diagnosis. ST-YOLO based on YOLOv8s has been proposed for PV module defect detection, indicating that lightweight object detection networks can improve inspection efficiency [109]. Deep learning-based EL defect detection and quantification studies show that visual models can not only classify defects but also

support quantitative assessment of defect regions [110]. Other EL image defect-detection studies further demonstrate that model reliability is affected by image resolution, defect diversity, annotation quality, and domain variation [111]. Automatic defect detection using infrared, EL, and RGB images shows that multimodal visual inspection can identify different types of PV module defects by combining complementary image sources [112]. Studies comparing YOLO and Transformer vision methods in thermal image-based PV fault diagnosis prove that these architectures feature unique merits in precision, inference velocity, and small fault perception ability [113]. Vision Transformer-based PV panel detection also shows that attention mechanisms can improve visual representation and detection performance [114]. CNN-based EL defect detection studies provide additional evidence that convolutional architectures remain effective for automatic PV cell defect recognition [115]. Aerial imagery combined with deep learning has also been used to improve hotspot and dirt-accumulation detection in large PV plants [116]. Therefore, image-based diagnosis is moving from single-modality defect classification toward multimodal, architecture-aware, and field-deployable fault localization.

Despite these advances, image-based diagnosis still faces practical deployment barriers. EL and infrared inspection often require standardized imaging conditions, reliable annotation, and sufficient defect samples to train robust models [102,103]. Small sample size, class imbalance, domain shift, imaging resolution, and environmental variation can reduce the generalization ability of deep learning-based visual diagnosis models [104,111]. Therefore, image-based PV fault diagnosis should be combined with data augmentation, transfer learning, multimodal fusion, and standardized inspection protocols to improve field reliability.

Practical diagnostic model selection should therefore be matched to the available evidence. Electrical feature-based SVM, RF, and boosting models are appropriate for small or medium supervised datasets with reliable voltage/current/I-V features and clearly defined fault categories. Autoencoders and semi-supervised methods are preferable when fault labels are sparse, but abundant normal operation data are available. CNN, YOLO, U-Net, and ViT models are more suitable for infrared, EL, UAV, or satellite images when annotation quality is sufficient. Transfer learning is especially useful when the target site has few labeled images but shares visual patterns with source datasets. The main deployment barriers are not the availability of algorithms alone, but the cost of fault labeling, class imbalance, weather-dependent image quality, sensor calibration, and the lack of uniform inspection protocols.

4.4. Predictive Maintenance and Health Assessment

Fault diagnosis identifies current abnormal states, whereas predictive maintenance aims to estimate future risk and optimize maintenance actions. Reviews of PV system predictive maintenance show that the field is moving from corrective and preventive maintenance toward condition-based and prediction-driven maintenance strategies [117]. Staged review systems for AI-enabled PV predictive maintenance and fault monitoring stress that fault pattern recognition, severity grading, and critical fault prediction constitute the core of maintenance decision-making [118]. Therefore, predictive maintenance extends PV fault diagnosis from "what has happened" to "what is likely to happen and when intervention is needed."

Health assessment and degradation modeling are central to predictive maintenance. Deep learning-based prognostics and health management reviews indicate that PV health management can integrate signal data, image data, and operational history to support fault prognosis and condition assessment [119]. These diagnostic strategies enable the correlation of detection outcomes with degradation evolution, health status judgment,

and remaining useful life estimation. For this reason, system health assessment can offer reliable analytical support to link fault detection outcomes with asset scheduling and maintenance arrangements.

AI methods are increasingly used to support predictive maintenance and fault prognosis in PV systems. AI-based predictive maintenance reviews report that ensemble learning, CNNs, LSTM autoencoders, and other data-driven models can identify maintenance needs and improve PV system reliability [120]. Reviews of recent advances in solar plant predictive maintenance further show that IoT monitoring, ML/DL models, anomaly detection, and cybersecurity protection are becoming integrated into intelligent O&M frameworks [121]. Therefore, AI enables predictive maintenance to combine monitoring, diagnosis, risk estimation, cybersecurity awareness, and maintenance scheduling within a more integrated PV operation framework.

Predictive maintenance also has system-level implications for PV operation and energy management. Comprehensive reviews of PV predictive maintenance indicate that early fault warning and proactive intervention can reduce downtime, improve energy yield, and extend equipment lifetime [122]. Intelligent maintenance approaches further show that corrective, preventive, and predictive strategies should be coordinated to improve PV performance and reliability [123]. The adoption of AI-enabled predictive maintenance techniques in renewable power systems shows an increasingly tight linkage between fault identification, maintenance arrangement, and the comprehensive operation, as well as reliability control of renewable energy projects [124]. Consequently, intelligent maintenance provides a bridge from equipment-level fault diagnosis to system-level operation optimization.

In summary, AI-based PV fault diagnosis and intelligent maintenance form a closed chain from monitoring data to maintenance decisions. Sensor-based methods are suitable for online abnormal state detection, while image-based methods provide stronger spatial localization and defect recognition. Predictive maintenance further transforms diagnostic information into health assessment, remaining-risk estimation, and maintenance scheduling. This theoretical basis of diagnosis and maintenance is vital for further research. PV systems are no longer limited to power production assets and can function as controllable resources to engage in energy allocation, storage linkage, market operation, and low-carbon optimization regulation.

To connect fault recognition with maintenance decision-making, Figure 5 illustrates the workflow through which AI converts PV monitoring data into diagnosis, localization, health assessment, and predictive O&M actions.

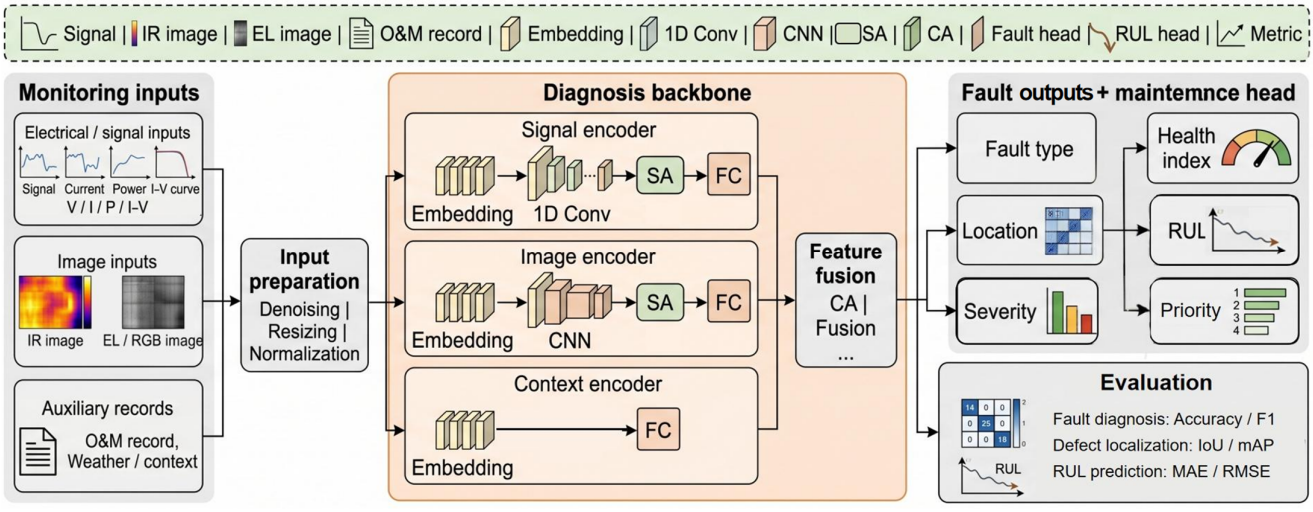


Figure 5. AI-enabled PV fault diagnosis and intelligent maintenance workflow.

In order to conclude the discussed data origins, diagnostic objectives, AI models, and maintenance outputs, Table 3 presents a comparison of classic AI methods in PV fault diagnosis and intelligent maintenance fields.

Table 3. AI-based PV fault diagnosis and intelligent maintenance techniques.

Refs.	Fault or Task	Data Source	AI Method	Output	Advantages	Limitations
[16,17,84,85]	General PV fault diagnosis	Voltage, current, power, temperature, I–V curves, images	ML, DL, hybrid AI	Fault type, abnormal state	Provides systematic diagnosis framework	Broad classification; limited task-specific detail
[87,88]	Module failure classification	Field failure records, module inspection data	Failure analysis, reliability assessment	Failure modes, field risk	Clarifies typical PV module failures	Limited real-time diagnostic capability
[86,90,91]	Degradation and aging assessment	Long-term performance data, temperature exposure, field-aged modules	Statistical analysis, degradation modeling	Degradation rate, performance loss	Supports lifetime and reliability evaluation	Long-term data required
[92]	Electrical fault detection	Voltage and current observations	Rule-based evaluation, signal analysis	Fault alarm, abnormal operating state	Simple and suitable for online monitoring	Sensitive to irradiance and operating-point changes
[93,94]	I–V curve-based diagnosis	Full I–V characteristics	ML classifiers, RF	Fault type, fault discrimination	Rich electrical health information	I–V measurement may be costly or discontinuous
[95–97]	Electrical signal-based classification and localization	Voltage, current, photocurrent, PV array signals	SVM, XGBoost, stacking ensemble, ML	Fault class, fault location	Effective for supervised multi-fault diagnosis	Requires labeled fault samples
[98,100,101]	Deep and semi-supervised signal diagnosis	PV operation data, electrical features	VAE, SSAE, DBN, LightGBM, SVM	Anomaly detection, fault classification	Reduces manual feature dependence	Training complexity and data-quality sensitivity
[99]	Neural network-based fault diagnosis	Electrical measurements, PV system signals	MLP, PNN, RBF, CNN, SAE	Fault identification	Summarizes ANN-based diagnostic routes	Model comparison may vary by dataset
[89,106–108]	Infrared hotspot detection	Infrared thermographic images	CNN, transfer learning, VGG-16, DeepLab-YOLO	Hotspot detection, thermal fault localization	Direct spatial fault visualization	Affected by weather, imaging angle, and thermal noise
[102,104,105,110,111,115]	EL image defect detection	Electroluminescence images	CNN, deep learning detection networks	Cracks, inactive areas, cell defects	Fine-grained cell/module defect recognition	Requires high-quality EL images and annotations
[109,113,116]	Object detection for PV defects	Infrared, thermal, aerial images	YOLO, YOLOv8, Transformer-based detectors	Defect boxes, hotspot/dirt localization	Fast inspection for large PV plants	Small defects and domain shifts remain difficult

Table 3. *Cont.*

Refs.	Fault or Task	Data Source	AI Method	Output	Advantages	Limitations
[112]	Multimodal image-based defect detection	Infrared, RGB, EL images	Deep learning, multimodal detection	Multiple defect types	Combines surface, thermal, and internal features	Multimodal alignment and data acquisition are complex
[114]	PV panel visual detection	Remote sensing or visual images	Vision Transformer	PV panel detection, localization	Captures global visual context	High computation and data requirements
[117,120, 122,123]	Predictive maintenance	O&M data, monitoring records, fault history	ML, DL, ensemble learning, LSTM autoencoder	Maintenance needs, warning signal	Reduces downtime and supports proactive O&M	Requires continuous monitoring and reliable labels
[118,119]	PHM and health assessment	Signals, images, operation history	AI-driven PHM, deep learning	Health state, severity, risk level	Links diagnosis with prognosis	Engineering deployment remains challenging
[44,121, 124]	Reliability, cybersecurity, and intelligent O&M	Reliability datasets, IoT data, solar plant records	AI, anomaly detection, cybersecurity-aware methods	Reliability support, risk warning	Extends diagnosis to system-level O&M	Data privacy, cyber risk, and standardization issues

5. AI-Driven PV Operation, Energy Management, and Low-Carbon Optimization

5.1. AI-Based Control and Maximum Power Point Tracking

Even though future power production trends and device operating states can be obtained via PV prediction technology and fault detection means, it is still essential for photovoltaic systems to adopt efficient control tactics to realize the conversion of solar resources into stable usable electric energy. As a core control function of photovoltaic systems, maximum power point tracking (MPPT) plays a vital role in judging the operation state of photovoltaic arrays, enabling them to work close to the maximum power point amid changing illumination and temperature environments [125]. Simple structure and low deployment cost make classic MPPT strategies widely researched and used, mainly including perturbation observation, incremental conductance, and hill climbing methods [126]. However, these methods may suffer from oscillation, slow tracking, or convergence to local maximum power points under rapidly changing weather or partial shading conditions [127]. Therefore, AI-based MPPT is introduced to improve tracking accuracy, dynamic response, and global optimization capability under complex operating conditions.

AI methods can enhance MPPT by learning nonlinear relationships between environmental conditions, operating states, and optimal control actions. Early reviews of AI techniques in PV applications show that neural networks, fuzzy logic, evolutionary algorithms, and hybrid intelligent methods have been used for PV modeling, control, prediction, and optimization [128]. Relevant comprehensive studies show AI aids control and decision-making of nonlinear and uncertain renewable power systems [10]. By integrating fuzzy logic regulation and tunable Gaussian process regression in current studies, it is proven that intelligent MPPT schemes are feasible to be constructed to cope with intricate external environments and partial shading effects [129]. Thus, AI-based MPPT extends traditional control from rule-driven adjustment toward adaptive and data-informed power extraction.

Reinforcement learning and neural control methods further expand AI-based PV control from static MPPT to dynamic system coordination. Reinforcement neural network-

based control has been applied to grid-integrated PV systems and battery management, indicating that learning-based controllers can jointly consider PV generation, storage operation, and grid interaction [130]. From existing summaries on AI utilization within renewable energy systems, it can be found that control methods driven by artificial intelligence can achieve effective fusion with energy administration, storage resource planning, and intelligent power grid running work [20]. AI-based PV energy management control studies further show that intelligent controllers can coordinate PV output and system demand through adaptive control strategies [131]. Consequently, MPPT should not be regarded as an isolated device-level function, but as the first control layer of a broader AI-driven PV operation framework.

Nevertheless, AI-based control still needs to satisfy engineering constraints. Intelligent MPPT methods may improve tracking performance, but their reliability depends on training data quality, sensor accuracy, computational resources, and robustness under unseen weather conditions [132]. In grid-connected PV systems, control actions must also respect converter limits, battery constraints, grid code requirements, and safety boundaries [130]. Therefore, the practical value of AI-based control lies not only in higher power extraction but also in its ability to operate reliably within the physical and grid-connected constraints of PV systems.

5.2. PV-Storage Coordination and Microgrid Energy Management

PV-storage coordination is a key step in transforming PV from an intermittent generation source into a dispatchable energy resource. Because PV output is weather-dependent and uncertain, probabilistic solar forecasting is often needed to support storage scheduling, reserve planning, and risk-aware operation [72]. Energy forecasting reviews also emphasize that forecasting should be connected with downstream operational decisions rather than evaluated only by prediction accuracy [46]. Battery energy storage systems can mitigate PV fluctuations, reduce curtailment, support peak shaving, and improve microgrid reliability, but their operation must consider state of charge, degradation, load demand, and electricity price signals. Therefore, AI-driven PV-storage coordination requires both predictive information and optimization-oriented control.

AI-based PV-storage management provides a flexible way to coordinate renewable generation, storage devices, load demand, and grid interaction. Numerous summaries concerning AI practical usage in renewable energy fields have illustrated that under working scenarios with nonlinear characteristics and uncertain factors, intelligent technologies can effectively assist optimization design, system control, storage planning, as well as smart grid operational management [132]. MPC is widely applied in PV-battery microgrids to embed prediction data into receding horizon decision making. A PV-battery microgrid study showed that MPC can manage variable PV generation and load conditions by updating control actions according to predicted system states [133]. Deep reinforcement learning has also been applied to microgrids with flexible demand, showing its ability to learn energy management policies under dynamic operating conditions [134]. Real-time economic energy management using deep reinforcement learning further demonstrates that DRL can address uncertainties in PV generation, load demand, and price signals [135]. Therefore, both predictive optimization and learning-based decision-making provide important technical routes for PV-storage energy management.

AI-based EMS methods are also being extended to different microgrid scales and operating structures. Deep learning optimized energy management is researched for PV-battery household microgrids, enabling AI to optimize local energy scheduling and raise PV self-consumption level [136]. Cooperative optimization of grid-connected microgrids using MPC further demonstrates that predictive control can coordinate multiple micro-

grids while considering operational constraints [137]. MPC has also been simulated in AC/DC microgrids for energy management and load control tasks, indicating its applicability to hybrid microgrid structures [138]. These studies suggest that AI-driven EMS is moving from single-site PV-storage scheduling toward coordinated and constraint-aware microgrid operation.

Recent studies further show that PV-storage management is expanding toward integrated solar-storage systems, networked microgrids, and hierarchical control. Reviews of AI for solar systems with integrated storage indicate that AI methods can improve storage sizing, dispatch, and operational optimization, while still facing challenges related to data quality, uncertainty, and practical deployment [139]. Reviews of intelligent control strategies for networked microgrids further show that AI, MPC, and hybrid methods are increasingly used to coordinate distributed energy resources at larger system scales [140]. Reinforcement learning is adopted for battery management in multi-level microgrids, revealing its superiority in hierarchical energy control [141]. Thus, PV-storage energy management forms the operational bridge between device-level PV control and system-level distributed energy coordination.

5.3. Virtual Power Plants and Electricity Market Participation

As distributed PV systems grow, their operational value increasingly depends on aggregation and market participation. Virtual power plants (VPPs) aggregate distributed PV, storage, flexible loads, and other distributed energy resources so that small-scale resources can participate in electricity markets and system services [142]. Studies on VPPs and virtual aggregators prove that VPPs can join energy markets, auxiliary services, flexible markets, and demand response projects [143]. Since PV generation, load response, and electricity prices are all uncertain, VPP operation requires forecasting, optimization, bidding, and risk-management functions. Therefore, AI provides a technical foundation for transforming distributed PV from passive behind-the-meter generation into actively managed market resources.

Market participation requires VPPs to make coordinated decisions across energy, reserve, and ancillary services products. Studies on VPP bidding strategies show that VPPs can participate simultaneously in energy and reserve markets through optimized bidding decisions [144]. Optimal bidding in combined electricity and ancillary services markets further demonstrates that demand response prices and consumption satisfaction can affect VPP revenue and operational strategy [145]. When carbon markets and demand response are considered, VPP bidding strategies can also incorporate carbon-related constraints and cooperative benefit allocation [146]. Thus, VPP operation is becoming a multi-objective decision problem involving profit, uncertainty, flexibility, carbon cost, and user response.

AI and reinforcement learning are increasingly used to support VPP scheduling and decision-making. Reinforcement learning has been applied to VPP decision-making in urban environments, showing its potential for adaptive operation under changing conditions [147]. Reviews of AI-driven VPPs indicate that AI can support resource aggregation, forecasting, scheduling, bidding, and ancillary services provision [148]. Data-driven intelligent scheduling studies also show that deep reinforcement learning and data-based methods can improve VPP scheduling decisions [149]. Therefore, AI enables VPPs to move from static aggregation platforms toward adaptive and intelligent market participants.

However, VPP deployment also depends on market mechanisms and response capability evaluation. Reviews of VPP response capability emphasize that flexibility evaluation, market mechanisms, and response reliability are essential for commercial operation [150,151]. Energy and probabilistic forecasting studies further show that uncertainty representation is necessary when renewable resources participate in markets and system

operation [46,76]. Probabilistic solar forecasting is especially relevant for PV-dominated VPPs because it helps quantify generation risk before market bidding and dispatch decisions are made [72]. Intelligent VPP market participation demands precise prediction and scheduling tools, alongside standardized uncertainty processing and market-adaptive decision systems.

5.4. Carbon-Aware and Life-Cycle-Oriented PV Optimization

PV generation is central to the low-carbon energy transition, but its sustainability should be evaluated from both operation and life-cycle perspectives. Global energy-transition studies identify renewable energy deployment as a major pathway for deep decarbonization [3]. Renewable power cost reports also show that the declining cost of PV strengthens its economic role in low-carbon power systems [3]. It cannot be ignored that a series of ecological effects will emerge in the full lifespan of PV facilities, ranging from raw material acquisition and product fabrication to distribution delivery, system erection, routine upkeep, scrapping disposal, and subsequent cyclic reuse [152]. Carbon-oriented PV optimization needs to surpass real-time control and integrate life-cycle-based evaluation.

To comprehensively analyze the environmental influences of PV systems excluding operational processes, researchers commonly adopt the life-cycle assessment framework. Existing overview studies on LCA application in solar panel evaluation have generalized key environmental metrics, research boundary settings, inventory data presuppositions, as well as various methodological difficulties in relevant technical assessments [152]. IEA PVPS life-cycle inventory reports provide standardized information on energy inputs, material flows, and emissions across PV system life cycles [153]. A great many LCA investigations targeting the present recycling modes of solar cell modules have proven that different recycling techniques will directly affect the ecological benefits of photovoltaic systems during their decommissioning phase [154]. A variety of new recovery techniques applied to crystal silicon solar modules have undergone comprehensive LCA verification. The results demonstrate that well-planned recycling schemes can effectively cut down adverse environmental impacts [155]. Therefore, LCA provides the basis for extending PV optimization from short-term operational performance to whole-life environmental assessment.

Recent LCA studies further indicate that both contemporary system conditions and earlier life-cycle evidence should be considered when evaluating PV sustainability. Updated LCA studies of utility-scale PV systems provide evidence for assessing carbon footprint, energy payback time, and environmental performance under current deployment conditions [156]. Earlier life-cycle analyses of photovoltaics also show that environmental performance depends on module technology, manufacturing energy, system lifetime, and local energy mix [157]. Growing PV installed capacity raises the significance of end-of-life management. IRENA and IEA-PVPS studies stress that effective collection, recycling, and resource recovery measures are needed for massive retired PV panels in the future [158]. Reviews of PV waste management further highlight the need to address environmental risks, recycling technologies, policy support, and sustainable waste governance [159]. Thus, circular economy thinking should be integrated into PV planning, operation, and retirement-stage management.

AI may support PV recycling and circular economy decisions by improving process recognition, material sorting, and recovery-route selection. Reviews of end-of-life PV module recycling identify opportunities for AI in automated disassembly, material recognition, process optimization, and recycling decision support [160]. These capabilities can help connect recycling information with LCA databases, operational records, and long-term sustainability assessment. Therefore, AI-assisted recycling can extend PV intelligence from power generation and operation to end-of-life resource recovery.

AI can also contribute to carbon-aware PV operation by integrating forecasts, storage states, demand flexibility, and carbon signals into operational decisions. Carbon-aware dispatch studies show that energy systems can incorporate carbon pricing, demand response, and stepped carbon cost mechanisms into scheduling models [161]. Cost- and carbon-aware sizing of local PV generation and energy storage further demonstrates that PV-storage systems can be optimized under both economic and emissions-related objectives [162]. Carbon-aware mobile energy storage scheduling based on carbon-emission flows also provides a reference for coordinating storage operation with low-carbon objectives [163]. These studies suggest that future PV energy management should optimize cost, reliability, flexibility, and carbon intensity simultaneously.

When combined with LCA data, carbon factors, operation forecasts, storage states, and recycling information, AI can support a more complete low-carbon decision chain from generation to retirement. Scenario generation reviews for renewable energy uncertainty also indicate that stochastic and AI-generated scenarios can support planning and operational decisions under uncertain PV output [78]. Thus, carbon-aware and life-cycle-oriented optimization extends AI-enabled PV operation from short-term dispatch to long-term sustainability management.

In summary, AI-driven PV operation connects the informational outputs of forecasting and diagnosis with practical control, scheduling, market participation, and low-carbon decision-making. MPPT improves device-level energy extraction, PV-storage coordination enhances system flexibility, VPPs enable distributed PV market participation, and life-cycle-oriented optimization embeds sustainability into both operation and planning. The next section therefore compares the main AI methods used across forecasting, diagnosis, control, and optimization, and discusses their common challenges and future research directions.

For operation and control, AI techniques should be distinguished by decision role and safety requirements. Classical optimization and model predictive control are preferable when system constraints, forecasts, and device models can be explicitly formulated. Reinforcement learning is more suitable for sequential PV-storage or VPP decisions only when a safe simulator, digital twin, or high-quality offline dataset is available for training and validation. Generative scenarios should be coupled with stochastic optimization rather than used as stand-alone forecasts. In field deployment, the limiting factors are often safety validation, communication latency, grid code compliance, privacy and cybersecurity requirements, and operator trust.

To synthesize the control, scheduling, market, and low-carbon optimization methods discussed in this section, Table 4 summarizes representative AI applications in PV operation, energy management, and life-cycle-oriented optimization.

Table 4. AI applications in PV operation, energy management, and low-carbon optimization.

Refs.	Application Area	Main Data/Inputs	AI or Control Methods	Optimization Target	Advantages	Limitations
[125–127]	Conventional MPPT baseline	Irradiance, temperature, PV voltage/current	P&O, incremental conductance, hill climbing	Maximum power extraction	Simple, low-cost, easy to implement	Oscillation, slow response, local optimum under partial shading
[128,129, 131]	AI-based MPPT and local PV control	PV operating states, weather variables, voltage/current	Fuzzy logic, neural networks, Gaussian process regression, neuro-fuzzy control	Fast and accurate maximum power tracking	Better nonlinear adaptation under complex conditions	Requires training data and controller tuning

Table 4. *Cont.*

Refs.	Application Area	Main Data/Inputs	AI or Control Methods	Optimization Target	Advantages	Limitations
[10,20,130, 132]	Intelligent PV control and battery management	PV output, grid status, battery SOC, control signals	Reinforcement learning, neural control, AI-based optimization	Coordinated PV-grid-battery control	Supports adaptive and dynamic control	Robustness and safety constraints remain critical
[133,137, 138,140]	MPC-based microgrid EMS	PV forecast, load forecast, SOC, grid constraints	Model predictive control, cooperative MPC	Rolling-horizon energy scheduling	Handles forecasts and operational constraints explicitly	Depends on model accuracy and prediction quality
[134,135, 141]	Reinforcement-learning-based EMS	PV generation, load, price, SOC, flexible demand	RL, DRL, hierarchical RL	Real-time economic dispatch	Learns adaptive policies under uncertainty	Training stability and safety during deployment are challenging
[136,139]	PV-storage optimization	PV output, battery data, load demand, storage capacity	Deep learning, AI optimization, hybrid EMS	Storage dispatch, curtailment reduction, self-consumption improvement	Improves PV flexibility and energy utilization	Battery degradation and cost modeling may be incomplete
[46,53,72, 76]	Forecasting-assisted operation	Point/probabilistic forecasts, uncertainty information, load, and price forecasts	Forecasting–optimization coupling, stochastic decision support	Risk-aware scheduling and reserve planning	Links prediction with operational decisions	Forecast uncertainty may propagate into dispatch errors
[142,143, 151]	VPP aggregation and response capability	Distributed PV, storage, flexible loads, demand response resources	Aggregation models, flexibility evaluation, market mechanism analysis	DER aggregation and flexibility service provision	Enables small PV resources to participate in system services	Requires communication, coordination, and market-rule support
[144–146,150]	VPP bidding and market participation	PV forecast, price forecast, reserve demand, carbon price, demand response	Stochastic optimization, game theory, bidding optimization	Energy/reserve market revenue and risk control	Supports multi-market participation	Sensitive to market rules, price uncertainty, and user response
[147–149]	AI-driven VPP scheduling	DER states, market signals, load response, historical operation data	RL, DRL, data-driven scheduling	Adaptive VPP operation and dispatch	Improves flexibility under changing conditions	Explainability and operational security remain concerns
[152,153, 156,157]	PV life-cycle assessment	Material flows, energy inputs, emissions, installation, and operation data	LCA, life-cycle inventory analysis	Carbon footprint and environmental impact assessment	Extends evaluation beyond operation phase	Results depend on system boundary and inventory assumptions
[154,155, 158–160]	PV recycling and circular economy	End-of-life modules, recycling process data, material recovery data	LCA, recycling-route evaluation, AI-assisted recycling	Waste reduction and material recovery	Supports circular PV development	Recycling economics and large-scale process standardization remain difficult

Table 4. *Cont.*

Refs.	Application Area	Main Data/Inputs	AI or Control Methods	Optimization Target	Advantages	Limitations
[161–163]	Carbon-aware dispatch and low-carbon operation	Carbon factor, electricity price, demand response, storage state, PV output	Carbon-aware optimization, stepped carbon pricing, carbon-flow-based scheduling	Cost-carbon coordinated dispatch	Integrates emissions into operational decisions	Carbon signals and policy mechanisms vary by region
[3,78,152–163]	Life-cycle-oriented low-carbon PV optimization	Renewable deployment targets, PV cost, LCA data, scenario uncertainty, recycling data	Scenario-based optimization, LCA-assisted AI, carbon-aware EMS	Long-term sustainable PV planning	Connects operation, planning, and end-of-life management	Requires integrated data across operation, carbon, and recycling stages

6. Comparative Analysis, Challenges, and Future Research Directions

6.1. Cross-Task Comparison of AI Methods in PV Systems

The preceding sections show that AI-enabled PV systems are no longer limited to isolated forecasting models but have gradually expanded into fault diagnosis, intelligent maintenance, energy management, market participation, and low-carbon optimization. A large number of bibliometric statistics and literature summaries focusing on intelligent energy systems have confirmed that both shallow and deep learning algorithms are commonly adopted to fulfill prediction, layout design, operational control, and performance optimization work [11,13]. A great deal of summary research focusing on intelligent technologies within renewable energy sectors has further proven that various artificial intelligence approaches are gradually penetrating into prediction analysis, resource scheduling, strategic decision formulation, and operational performance improvement [12,20]. Therefore, a cross-task comparison is necessary because the suitability of an AI method depends strongly on the PV task, data structure, time scale, and deployment requirement.

Traditional machine learning methods are still valuable when the input features are well defined and the dataset size is limited. Due to the advantages of fast calculation speed and simple deployment process, multiple intelligent algorithms, including support vector regression, random forest, gradient boosting, and artificial neural networks, have become mainstream tools in the field of photovoltaic forecasting [17,53]. Similar machine learning classifiers are also used in PV fault diagnosis because electrical signals and I–V curve features can be converted into structured inputs for fault classification [84]. However, these models usually depend heavily on feature engineering and may have limited ability to capture long-term temporal dependence or multimodal information [11]. Thus, traditional machine learning is more suitable for engineering baseline models than for highly heterogeneous and large-scale PV intelligence.

Deep learning performs better in PV research with nonlinear temporal features, visual data, and high-dimensional inputs. Related reviews confirm CNN, RNN, LSTM, GRU, attention, and Transformer models are widely studied for PV power forecasting [13]. In fault diagnosis and predictive maintenance, deep learning can also extract features from electrical signals, infrared images, EL images, and operation histories [118]. Nevertheless, deep learning models often require large datasets, careful tuning, and sufficient computing resources [15,118]. Therefore, deep learning improves representation capability but also raises requirements for data scale, annotation quality, and deployment resources.

Graph learning, reinforcement learning, and generative AI represent more task-specific extensions of PV intelligence. Graph-based models are particularly suitable for distributed PV forecasting because they can represent spatial correlations among multiple PV sites and networked resources [65,66]. Reinforcement learning is more suitable for energy management, storage control, and VPP scheduling because these tasks involve sequential decisions under uncertainty [139,148]. Generative and probabilistic models are useful for uncertainty quantification, scenario generation, and risk-aware operation because they describe possible future PV trajectories rather than only single expected values [21,83]. Therefore, advanced AI methods should be selected according to the task structure rather than treated as universally superior alternatives.

Generally, the decisive issue is not which model achieves the lowest error in isolated experiments, but which model provides decisions that are useful, physically reasonable, and operationally affordable outputs. When the output of a predictive model reduces equilibrium costs, reserve uncertainty, or market risk, the predictive model is valuable. When diagnostic models trigger timely maintenance with an acceptable false positive rate, they are valuable. When the operational model respects physical constraints, market rules, and security requirements, they are valuable. Low-carbon artificial intelligence is only valuable when it is reflected in carbon signals, life-cycle data, and economic goals. This cross-task perspective was used in Tables 5 and 6 to compare data requirements, deployment maturity, and remaining obstacles.

6.2. Data Quality, Generalization, and Benchmarking

Data quality remains one of the most fundamental constraints on AI-enabled PV systems. Missing-value studies show that incomplete meteorological and PV generation data can significantly affect day-ahead PV forecasting performance [35]. Feature selection studies also indicate that redundant or weakly relevant variables may reduce forecasting efficiency and model robustness [39]. Open PV reliability dataset reviews further show that environmental data, operation-monitoring data, image-inspection data, and material-level data remain unevenly available across different PV research tasks [44]. Therefore, the performance of PV-AI models is often limited not only by algorithm design but also by the completeness, consistency, and representativeness of the underlying data.

The problem becomes more serious when PV models are transferred across regions, seasons, devices, and climate conditions. Transfer learning theory shows that models trained in one domain may need adaptation before being applied to another domain with different data distributions [164]. Domain-adaptation studies further demonstrate that model performance may degrade when source and target distributions differ, and target labels are unavailable [165]. In PV applications, such distribution shifts may be caused by regional irradiance differences, seasonal weather patterns, module aging, inverter replacement, sensor drift, and changes in operation strategy. Thus, cross-site and cross-season generalization should be treated as a core evaluation requirement rather than an optional extension.

Concept drift and dataset shift are also important for long-term PV deployment. Concept-drift studies show that data streams may change over time, causing models trained on historical data to become less reliable [166]. Dataset-shift detection research further emphasizes that deployment systems should be able to identify when new data no longer match the training distribution [167]. Dataset bias studies also indicate that benchmark performance may be misleading if the test data do not reflect real operating diversity [168]. Therefore, PV-AI systems need continuous data monitoring, drift detection, model updating, and validation under changing environmental and equipment conditions.

Table 5. Comparison of various AI techniques for photovoltaic systems.

Refs.	AI Method Category	Main PV Tasks	Typical Input Data	Strengths	Limitations	Future Directions
[11,17,42,53,61,62,84]	Traditional machine learning	PV forecasting; fault classification; performance evaluation	Weather data, PV power, voltage, current, I–V features	Efficient, interpretable, easy to deploy	Relies on feature engineering; limited sequence modeling	Hybrid ML with physical constraints and automated feature selection
[13,54–60,99,104,115,118]	Deep learning	Temporal forecasting; image defect detection; predictive maintenance	Time series, EL images, infrared images, operation logs	Strong nonlinear representation; suitable for high-dimensional data	Data-hungry; weak interpretability; risk of overfitting	Explainable DL, lightweight DL, cross-site transfer learning
[55–57,63,100,101]	Hybrid and ensemble learning	Short-term forecasting; fault diagnosis; anomaly detection	Multi-source features, decomposed signals, electrical measurements	Improves robustness and prediction accuracy	Complex structure; difficult tuning	Adaptive ensemble selection and task-specific hybrid design
[58,59,64,113,114]	Transformer and attention models	Multi-step forecasting; ultra-short-term forecasting; visual fault detection	Multivariate time series, thermal images, visual images	Captures long-range dependencies and global context	High computation cost; requires sufficient data	Efficient attention, edge-friendly Transformer, interpretable attention
[26,33,34,65–70]	Spatiotemporal and graph learning	Distributed PV forecasting; regional PV prediction; multi-site coordination	Multi-site PV data, satellite data, sky images, topology data	Captures spatial correlation and weather propagation	Graph construction and multimodal alignment are difficult	Physics-guided graphs and dynamic graph learning
[21,71–83]	Probabilistic and generative AI	Uncertainty quantification; prediction intervals; scenario generation	PV power, weather variables, NWP data, historical scenarios	Supports risk-aware dispatch and stochastic optimization	Calibration and physical feasibility are challenging	Physics-constrained generative models and decision-oriented evaluation
[125–132]	AI-based control and MPPT	MPPT; inverter control; PV-grid-battery control	Irradiance, temperature, voltage, current, SOC, grid status	Improves adaptive control under complex conditions	Safety and stability constraints must be guaranteed	Safe RL, robust control, physically constrained AI controllers
[133–141]	MPC, RL, and DRL for EMS	PV-storage scheduling; microgrid energy management; flexible load control	PV forecast, load forecast, SOC, electricity price, constraints	Handles sequential decisions and uncertainty	Training stability and real-time feasibility remain issues	Safe DRL, hybrid MPC-RL, reliability-aware EMS
[142–151]	AI for VPP and market participation	VPP aggregation; bidding; ancillary services; demand response	DER states, PV forecasts, price signals, market rules, carbon price	Enables distributed PV to participate in markets	Sensitive to market uncertainty and user response	Multi-agent RL, explainable bidding, carbon-market integration

Table 5. *Cont.*

Refs.	AI Method Category	Main PV Tasks	Typical Input Data	Strengths	Limitations	Future Directions
[152–163]	Carbon-aware and LCA-assisted AI	Low-carbon dispatch; PV-storage sizing; recycling and circular economy	Carbon factors, LCA data, recycling data, cost data, PV output	Extends PV optimization to sustainability goals	LCA boundary and carbon-factor uncertainty	Life-cycle-aware AI and carbon cost-coordinated optimization
[169–175]	Explainable and trustworthy AI	Forecast interpretation; diagnosis explanation; risk-aware decision support	Model outputs, feature attribution, uncertainty estimates	Improves transparency and user trust	Explanation quality is hard to evaluate	Standardized XAI metrics and human-in-the-loop PV decision support
[20,121, 176,177]	Cybersecurity-aware AI	Secure monitoring; attack detection; intelligent O&M protection	IoT data, inverter data, communication logs, grid measurements	Protects data integrity and operational reliability	Cyber-physical attack data are limited	Resilient PV-AI architecture and attack-aware model validation
[178–182]	Physics-informed AI	PV forecasting; parameter estimation; physically consistent modeling	Physical equations, irradiance, temperature, power, sky images	Improves physical consistency and generalization	Requires accurate physical priors and model coupling	Physics-informed forecasting, diagnosis, and control
[183–186]	Digital twin AI	Monitoring; simulation; predictive maintenance; operation optimization	Real-time data, system models, control actions, maintenance records	Enables simulation, validation, and closed-loop optimization	Model calibration and real-time updating are challenging	PV digital twins for forecasting–diagnosis–operation integration
[33,34,70, 112,187]	Multimodal AI	Solar forecasting; image diagnosis; regional PV analysis	Weather data, sky images, satellite images, EL/IR/RGB images, topology	Integrates complementary information	Data alignment and modality imbalance are difficult	Unified multimodal representations and cross-modal reasoning
[188–191]	Federated learning and privacy-preserving AI	Cross-site forecasting; multi-plant diagnosis; privacy-aware collaboration	Decentralized PV data, local model parameters, encrypted updates	Enables data collaboration without raw-data sharing	Communication cost, system heterogeneity, privacy leakage risk	Federated PV benchmarks and secure cross-plant learning
[190–198]	Edge and sustainable AI	Real-time inference; edge diagnosis; low-power deployment; green AI	Edge sensor data, compressed models, local inference data	Reduces latency and communication cost	Accuracy–efficiency trade-off; AI carbon footprint	Lightweight, low-carbon, and energy-aware PV-AI deployment

Benchmarking and reproducibility are necessary to compare AI methods fairly. Energy forecasting competitions such as GEFCom2014 demonstrate that public datasets and unified evaluation rules can promote objective comparison among forecasting models [45]. Reviews of energy forecasting also emphasize the importance of open data, standard metrics, transparent baselines, and reproducible experimental design [46]. As specified by the FAIR guiding principles, all scientific datasets need to be easy to locate, openly available, mutually compatible, and capable of repeated utilization. This standard is highly applica-

ble to photovoltaic data construction and relevant model validation work [47]. Therefore, future PV-AI research should move from isolated case studies toward shared datasets, standardized preprocessing, common evaluation protocols, and transparent reporting.

Table 6. Quantitative comparison of dimensions and deployment maturity of major PV-AI application areas.

PV-AI Area	Typical Metrics	Data Requirement	Representative Methods	Deployment Maturity	Main Gap
PV forecasting	MAE, RMSE, MAPE, CRPS, PICP	Historical PV power, weather/NWP, sky/satellite data	ML, LSTM/GRU, Transformer, GNN, probabilistic models	Mature for point forecasting; emerging for probabilistic/scenario forecasting	Open benchmarks and calibrated uncertainty for decision use
Fault diagnosis	Accuracy, Precision, Recall, F1, IoU, mAP	I–V curves, voltage/current, IR/EL/RGB/UAV images	SVM, RF, CNN, YOLO, ViT, multimodal DL	Mature in laboratory/inspection settings; partly deployed in large plants	Class imbalance, annotation cost, cross-site generalization
Predictive maintenance	Health index error, RUL MAE/RMSE, alarm precision	Fault logs, maintenance records, degradation histories, sensor streams	Autoencoder, LSTM, PHM models, ensemble learning	Emerging	Sparse failure labels and limited long-term reliability datasets
PV-storage EMS	Cost reduction, self-consumption, SOC violation, curtailment	PV/load forecasts, SOC, prices, constraints	MPC, DRL, hybrid EMS	Partly mature for MPC; emerging for safe DRL	Safety constraints, battery degradation, real-time reliability
VPP/market operation	Revenue, risk, reserve error, flexibility index	Distributed PV, storage, flexible loads, market prices	Stochastic optimization, DRL, bidding models	Emerging	Market-rule dependence and uncertainty propagation
Low-carbon optimization	CO_2 reduction, carbon cost, life-cycle indicators	Carbon factors, LCA inventory, recycling data	Carbon-aware dispatch, LCA-assisted AI	Emerging	Integration of operation, LCA, recycling, and policy data

A more rigorous review and benchmarking culture is also needed. PRISMA 2020 provides a structured framework for transparent literature identification, screening, eligibility assessment, and inclusion [20]. PRISMA-S further emphasizes the need to report databases, search strategies, search limits, and retrieval procedures in systematic reviews [21]. General literature review methodology also suggests that reviews should synthesize evidence through clear classification logic instead of merely listing publications [24,25]. For PV-AI research, this implies that future review studies should report not only algorithm categories but also dataset sources, sample sizes, geographic coverage, data resolution, evaluation metrics, and deployment conditions. In this way, benchmarking can become a foundation for credible comparison and practical model selection.

6.3. Interpretability, Trustworthiness, and Cybersecurity

As AI models become more deeply embedded in PV operation, interpretability and trustworthiness become as important as accuracy. Explainable AI research shows that

users need understandable reasons for model predictions, especially when model outputs support operational decisions [169,170]. SHAP-based interpretation provides a unified way to estimate how different input variables contribute to model outputs [171]. More general studies on interpretable machine learning also emphasize that interpretability should be evaluated according to the decision context and user needs [172]. Therefore, PV-AI models should explain why a forecast, diagnosis, or control decision is produced, rather than only report numerical performance.

Interpretability is especially important for PV forecasting and fault diagnosis. Explainable PV output mapping based on weather data shows that XAI can reveal the contribution of meteorological variables to PV output estimates [173]. Trustworthy PV forecasting frameworks further combine explainability with uncertainty estimation to distinguish model anomalies from data anomalies [174]. In image-based PV fault detection, comparative studies of YOLO- and Transformer-based models show that model choice affects not only accuracy but also inference speed, small-target sensitivity, and practical usability [113]. Thus, interpretability helps identify whether a model is relying on physically meaningful features or dataset-specific artifacts.

Uncertainty quantification is another key component of trustworthy PV intelligence. Reviews of probabilistic PV forecasting show that prediction intervals and probability distributions can express the uncertainty of PV production and electricity consumption [71]. Studies on uncertainty and probabilistic forecasting in power systems also emphasize that uncertainty information is necessary for risk-aware decision-making under renewable variability [76]. Researchers can adopt neural network-driven bilateral boundary estimation strategies to realize the establishment of prediction interval ranges [80]. Scholars introduce conformal prediction methods into the field of ultra-short-term PV power prediction, aiming to optimize the rationality and accuracy of probabilistic fitting [81]. Therefore, trustworthy PV-AI should provide both predictions and explicit uncertainty boundaries.

Cybersecurity is a further requirement for deployable PV-AI systems. Reviews of AI in renewable energy systems indicate that data integrity, cyber risk, model robustness, and secure deployment should be considered when AI is applied to energy infrastructures [20]. IoT security and forensics studies show that connected devices face risks such as unauthorized access, data tampering, and post-incident traceability problems [175,176]. Smart-grid security research also shows that false data injection attacks can mislead monitoring and control systems if detection mechanisms are insufficient [177]. Since PV systems increasingly rely on inverters, sensors, communication networks, cloud platforms, and market interfaces, cybersecurity must be integrated into AI-enabled PV operation from the design stage.

For practical implementation, a trustworthy PV-AI architecture should include four operational layers: data integrity checking at the sensing layer, anomaly and attack detection at the communication layer, uncertainty- and explainability-based model validation at the decision layer, and fallback control or human review at the actuation layer. In engineering deployment, AI outputs should be accepted only when prediction uncertainty, data-quality indicators, and safety constraints remain within predefined thresholds; otherwise, the system should switch to conservative rule-based control or request operator verification. Recent anomaly detection studies for solar-grid batteries also show that Markov-constrained isolation forests can support early warning of storage anomalies in PV-grid applications.

Trustworthy AI also requires governance and risk management. In order to standardize artificial intelligence risk control, the NIST framework clarifies multiple key evaluation indicators, including effectiveness, operational stability, safety performance, information security, responsibility supervision, interpretability, and privacy protection [175]. For PV

systems, these principles imply that AI models should be validated under diverse operating conditions, protected against data manipulation, monitored after deployment, and audited when they influence control or market decisions. Therefore, the future development of AI-enabled PV systems should combine explainable modeling, calibrated uncertainty, cybersecurity protection, and governance mechanisms. These requirements provide the basis for the next discussion on physics-informed, multimodal, generative, digital twin, edge-deployed, and sustainable AI approaches.

To compare the applicability, advantages, limitations, and future directions of different AI methods across PV forecasting, fault diagnosis, operation, and low-carbon optimization, Table 5 summarizes the major AI categories discussed in this review.

To strengthen the cross-task comparison, Table 6 further distinguishes mature and emerging PV-AI areas according to typical evaluation metrics, data requirements, deployment readiness, and remaining engineering barriers.

As a synthesis of the operation-oriented and forward-looking discussions, Figure 6 integrates AI-driven PV operation, low-carbon optimization, and future research directions into a unified roadmap.

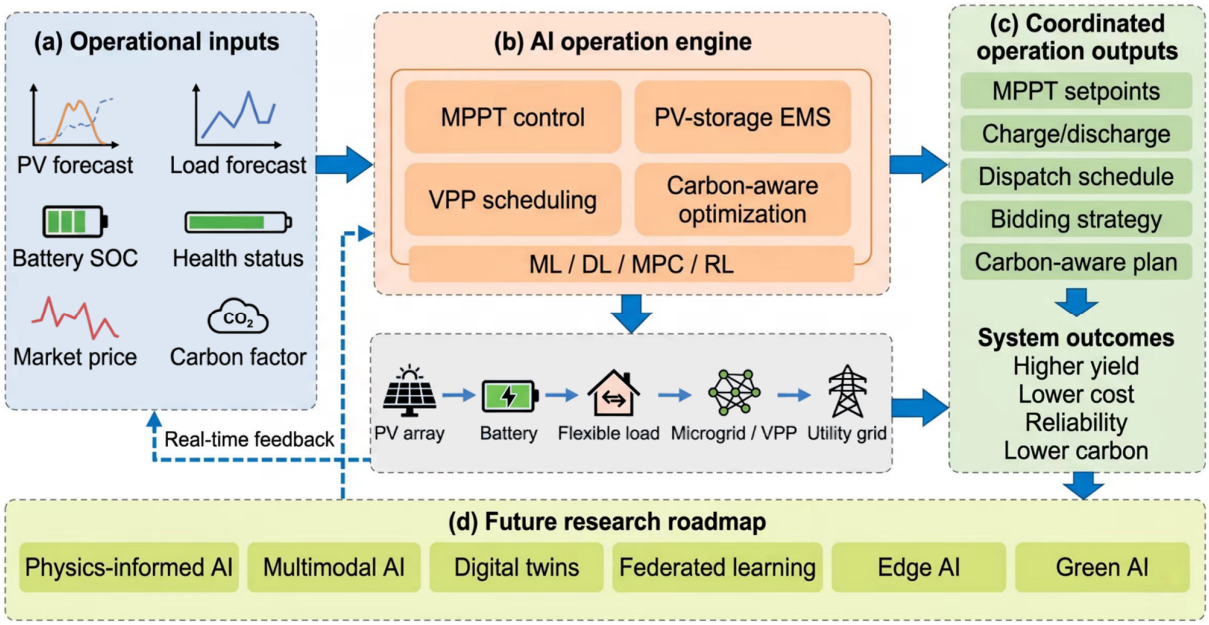


Figure 6. AI-driven PV operation, low-carbon optimization, and future research roadmap.

7. Conclusions

This review shows that machine learning and artificial intelligence are most valuable for PV power systems when they are connected to practical decisions rather than used only to improve isolated model accuracy. Across forecasting, diagnosis, maintenance, operation, and low-carbon optimization, the reviewed literature indicates that data governance, uncertainty representation, physical constraints, decision linkage, and sustainability objectives are the five key conditions for deployable PV intelligence.

Three actionable recommendations follow from this synthesis. First, PV-AI studies should report standardized datasets, preprocessing steps, evaluation metrics, and operating conditions so that forecasting, diagnosis, and energy management results can be compared objectively. Second, method selection should be task-specific: conventional machine learning is suitable for structured and moderate-scale data; deep learning is effective for temporal and visual representation; graph learning is suitable for distributed PV; reinforcement learning is promising for sequential control; generative AI is useful

for uncertainty and scenario modeling; and physics-informed AI can improve physical consistency. Third, future systems should link AI outputs with dispatch, storage scheduling, maintenance planning, VPP bidding, and carbon-aware operation through uncertainty thresholds, safety constraints, and fallback control rules.

The review also identifies several limitations in the existing literature. Many studies are still evaluated on single-site or laboratory datasets, with limited open benchmarks, insufficient cross-site validation, weak uncertainty calibration, incomplete reporting of computational cost, and limited evidence of industrial deployment. These gaps mean that reported gains in accuracy or classification performance do not always translate into robust field performance.

In conclusion, future PV-AI research should move toward data-governed, uncertainty-aware, physically constrained, decision-oriented, and sustainability-driven intelligence. Under this pathway, AI can help PV systems evolve from intermittent renewable generators into reliable, flexible, and low-carbon energy assets in future power systems.

Author Contributions: Conceptualization, Y.W. and X.F.; methodology, Y.W.;validation, Y.W. and X.F.; formal analysis, Y.W.; investigation, Y.W.; resources, X.F.; data curation, Y.W.; writing—original draft preparation, Y.W.; writing—review and editing, Y.W. and X.F.; visualization, Y.W.; supervision, X.F.; funding acquisition, X.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

Acknowledgments: The authors used AI-assisted language editing tools to improve grammar, readability, and formatting consistency during manuscript preparation. All scientific content, the literature interpretation, references, and conclusions were checked and approved by the authors.

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

References

1. International Energy Agency. *Renewables 2025: Analysis and Forecast to 2030*; IEA: Paris, France, 2025. Available online: https://www.iea.org/reports/renewables-2025 (accessed on 28 April 2026).
2. International Energy Agency. *World Energy Outlook 2025*; IEA: Paris, France, 2025. Available online: https://www.iea.org/reports/world-energy-outlook-2025 (accessed on 28 April 2026).
3. International Renewable Energy Agency. *Renewable Power Generation Costs in 2024*; IRENA: Abu Dhabi, United Arab Emirates, 2025. Available online: https://www.irena.org/Publications (accessed on 28 April 2026).
4. Europe, SolarPower. *Global Market Outlook for Solar Power 2025–2029*; SolarPower Europe: Brussels, Belgium, 2025. Available online: https://www.solarpowereurope.org/insights/outlooks/global-market-outlook-for-solar-power-2025-2029/detail (accessed on 28 April 2026).
5. REN21. *Renewables 2025 Global Status Report*; REN21 Secretariat: Paris, France, 2025. Available online: https://www.ren21.net/gsr-2025/ (accessed on 28 April 2026).
6. Inman, R.H.; Pedro, H.T.C.; Coimbra, C.F.M. Solar forecasting methods for renewable energy integration. *Prog. Energy Combust. Sci.* **2013**, *39*, 535–576. [CrossRef]
7. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de-Pison, F.J.; Antonanzas-Torres, F. Review of photovoltaic power forecasting. *Sol. Energy* **2016**, *136*, 78–111. [CrossRef]
8. Cabrera-Tobar, A.; Bullich-Massagué, E.; Aragüés-Peñalba, M.; Gomis-Bellmunt, O. Review of advanced grid requirements for the integration of large scale photovoltaic power plants in the transmission system. *Renew. Sustain. Energy Rev.* **2016**, *62*, 971–987. [CrossRef]
9. Entezari, A.; Aslani, A.; Zahedi, R.; Noorollahi, Y. Artificial intelligence and machine learning in energy systems: A bibliographic perspective. *Energy Strategy Rev.* **2023**, *45*, 101017.
10. Shoaei, M.; Ebrahimi, M.; Ghaderi, S.F.; Noorollahi, Y. A review of the applications of artificial intelligence in renewable energy systems. *Energy Convers. Manag.* **2024**, *304*, 118207.

11. Forootan, M.M.; Larki, I.; Zahedi, R.; Ahmadi, A. Machine learning and deep learning in energy systems: A review. *Sustainability* **2022**, *14*, 4832. [CrossRef]
12. Ying, C.; Li, M.; Li, X.; Yang, Y.; Wang, W. Deep learning for renewable energy forecasting: A taxonomy and systematic literature review. *J. Clean. Prod.* **2023**, *384*, 135414.
13. Yu, J.; Li, Y.; Wang, X.; Zhang, Y. Deep learning models for PV power forecasting: Review. *Energies* **2024**, *17*, 3973. [CrossRef]
14. Iheanetu, K.J. Solar photovoltaic power forecasting: A review. *Sustainability* **2022**, *14*, 17005. [CrossRef]
15. Mellit, A.; Massi Pavan, A.; Ogliari, E.; Leva, S.; Lughi, V. Advanced methods for photovoltaic output power forecasting: A review. *Appl. Sci.* **2020**, *10*, 487. [CrossRef]
16. Islam, M.; Mekhilef, S.; Huda, A.S.N.; Shah, N.M.; Soon, T.K. Artificial intelligence in photovoltaic fault identification and diagnosis: A systematic review. *Energies* **2023**, *16*, 7417. [CrossRef]
17. Sepúlveda-Oviedo, E.H.; Trujillo, C.L.; Franco, E.; Cortés, C.A. Fault diagnosis of photovoltaic systems using artificial intelligence: A review. *Heliyon* **2023**, *9*, e22855.
18. Xiang, H.; Wang, X.; Li, Y.; Zhang, L. Artificial intelligence in renewable energy systems: Applications and security challenges. *Energies* **2025**, *18*, 1931. [CrossRef]
19. Yang, L.; Fu, X. A review of photovoltaic uncertainty modeling based on statistical relational AI. *Energies* **2026**, *19*, 1509. [CrossRef]
20. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef] [PubMed]
21. Rethlefsen, M.L.; Kirtley, S.; Waffenschmidt, S.; Ayala, A.P.; Moher, D.; Page, M.J.; Koffel, J.B.; PRISMA-S Group. PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst. Rev.* **2021**, *10*, 39. [CrossRef] [PubMed]
22. Snyder, H. Literature review as a research methodology: An overview and guidelines. *J. Bus. Res.* **2019**, *104*, 333–339. [CrossRef]
23. Xiao, Y.; Watson, M. Guidance on conducting a systematic literature review. *J. Plan. Educ. Res.* **2019**, *39*, 93–112.
24. Sengupta, M.; Xie, Y.; Lopez, A.; Habte, A.; Maclaurin, G.; Shelby, J. The National Solar Radiation Data Base (NSRDB). *Renew. Sustain. Energy Rev.* **2018**, *89*, 51–60. [CrossRef]
25. Huld, T.; Müller, R.; Gambardella, A. A new solar radiation database for estimating PV performance in Europe and Africa. *Sol. Energy* **2012**, *86*, 1803–1815. [CrossRef]
26. Pfenninger, S.; Staffell, I. Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data. *Energy* **2016**, *114*, 1251–1265. [CrossRef]
27. Deline, C.; Perry, K.; Deceglie, M.; Muller, M.; Sekulic, W.; Jordan, D. *Photovoltaic Data Acquisition (PVDAQ) Public Datasets*; Open Energy Data Initiative: Golden, CO, USA, 2021.
28. De Soto, W.; Klein, S.A.; Beckman, W.A. Improvement and validation of a model for photovoltaic array performance. *Sol. Energy* **2006**, *80*, 78–88. [CrossRef]
29. Holmgren, W.F.; Hansen, C.W.; Mikofski, M.A. pvlib python: A Python package for modeling solar energy systems. *J. Open Source Softw.* **2018**, *3*, 884. [CrossRef]
30. Kratochvil, J.A.; Boyson, W.E.; King, D.L. *Photovoltaic Array Performance Model*; Sandia National Laboratories: Albuquerque, NM, USA, 2004.
31. Yang, D.; Kleissl, J.; Gueymard, C.A.; Pedro, H.T.C.; Coimbra, C.F.M. History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Sol. Energy* **2018**, *168*, 60–101. [CrossRef]
32. Paletta, Q.; Terrén-Serrano, G.; Nie, Y.; Li, B.; Bieker, J.; Zhang, W.; Dubus, L.; Dev, S.; Feng, C. Advances in solar forecasting: Computer vision with deep learning. *Adv. Appl. Energy* **2023**, *11*, 100150. [CrossRef]
33. Liu, J.; Zang, H.; Cheng, L.; Ding, T.; Wei, Z.; Sun, G. A Transformer-based multimodal-learning framework using sky images for ultra-short-term solar irradiance forecasting. *Appl. Energy* **2023**, *342*, 121160.
34. Kim, T.; Ko, W.; Kim, J. Analysis and impact evaluation of missing data imputation in day-ahead PV generation forecasting. *Appl. Sci.* **2019**, *9*, 204. [CrossRef]
35. Liu, Z.; Xuan, L.; Gong, D.; Xie, X.; Liang, Z.; Zhou, D. A WGAN-GP approach for data imputation in photovoltaic power prediction. *Energies* **2025**, *18*, 1042. [CrossRef]
36. De-Paz-Centeno, I.; García-Ordás, M.T.; García-Olalla, Ó.; Alaiz-Moretón, H. Imputation of missing measurements in PV production data within constrained environments. *Expert Syst. Appl.* **2023**, *217*, 119510. [CrossRef]
37. Macaire, J.; Zermani, S.; Linguet, L. New feature selection approach for photovoltaic power forecasting using KCDE. *Energies* **2023**, *16*, 6842. [CrossRef]
38. Ali, M.; Rabehi, A.; Souahlia, A.; Guermoui, M.; Teta, A.; Tibermacine, I.E.; Rabehi, A.; Benghanem, M.; Agajie, T.F. Enhancing PV power forecasting through feature selection and artificial neural networks: A case study. *Sci. Rep.* **2025**, *15*, 22574. [CrossRef] [PubMed]

39. Niu, D.; Wang, K.; Sun, L.; Wu, J.; Xu, X. Short-term photovoltaic power generation forecasting based on random forest feature selection and CEEMD: A case study. *Appl. Soft Comput.* **2020**, *93*, 106389. [CrossRef]
40. Voyant, C.; Notton, G.; Kalogirou, S.; Nivet, M.L.; Paoli, C.; Motte, F.; Fouilloy, A. Machine learning methods for solar radiation forecasting: A review. *Renew. Energy* **2017**, *105*, 569–582. [CrossRef]
41. Nicoletti, F.; Bevilacqua, P. Hourly photovoltaic production prediction using numerical weather data and neural networks for solar energy decision support. *Energies* **2024**, *17*, 466. [CrossRef]
42. Chen, X.; Li, B.; Braid, J.L.; Byford, B.; Colvin, D.J.; Glaws, A.; Jost, N.; Pierce, B.; Rabade, S.; Springer, M.; et al. Open data sets for assessing photovoltaic system reliability. *Appl. Energy* **2025**, *395*, 126132. [CrossRef]
43. Hong, T.; Pinson, P.; Fan, S.; Zareipour, H.; Troccoli, A.; Hyndman, R.J. Probabilistic energy forecasting: Global Energy Forecasting Competition 2014 and beyond. *Int. J. Forecast.* **2016**, *32*, 896–913. [CrossRef]
44. Hong, T.; Pinson, P.; Wang, Y.; Weron, R.; Yang, D.; Zareipour, H. Energy forecasting: A review and outlook. *IEEE Open Access J. Power Energy* **2020**, *7*, 376–388. [CrossRef]
45. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 160018. [CrossRef] [PubMed]
46. Desert Knowledge Australia Solar Centre. *DKA Solar Centre Data Download*; Desert Knowledge Australia Solar Centre: Alice Springs, Australia, 2008.
47. European Commission Joint Research Centre. *Photovoltaic Geographical Information System (PVGIS)*; European Commission: Brussels, Belgium, 2025.
48. World Bank Group; Energy Sector Management Assistance Program; Solargis. *Global Solar Atlas*; World Bank Group: Washington, DC, USA, 2017.
49. Raza, M.Q.; Nadarajah, M.; Ekanayake, C. On recent advances in PV output power forecast. *Sol. Energy* **2016**, *136*, 125–144. [CrossRef]
50. Diagne, M.; David, M.; Lauret, P.; Boland, J.; Schmutz, N. Review of solar irradiance forecasting methods and a proposition for small-scale insular grids. *Renew. Sustain. Energy Rev.* **2013**, *27*, 65–76. [CrossRef]
51. Ahmed, R.; Sreeram, V.; Mishra, Y.; Arif, M.D. A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization. *Renew. Sustain. Energy Rev.* **2020**, *124*, 109792.
52. Konstantinou, M.; Peratikou, S.; Charalambides, A.G. Solar photovoltaic forecasting of power output using LSTM networks. *Atmosphere* **2021**, *12*, 124. [CrossRef]
53. Wang, K.; Qi, X.; Liu, H. Photovoltaic power forecasting based LSTM-Convolutional Network. *Energy* **2019**, *189*, 116225. [CrossRef]
54. Xiang, X.; Ma, M.; Ma, X.; Chen, L.; Cai, W.; Yan, X.; Liu, J. A short-term forecasting method for photovoltaic power generation based on the TCN-ECANet-GRU hybrid model. *Sci. Rep.* **2024**, *14*, 6233.
55. Wang, H.; Zhang, R.; Liu, Y.; Wang, L. Short-term photovoltaic power forecasting based on K-means clustering, ensemble learning and feature rise-dimensional approach. *Sustainability* **2023**, *15*, 15594.
56. Piantadosi, G.; Dutto, S.; Galli, A.; De Vito, S.; Sansone, C.; Di Francia, G. Photovoltaic power forecasting: A Transformer based framework. *Energy AI* **2024**, *18*, 100444. [CrossRef]
57. Wang, J.; Wang, J.; Zhang, Y.; Li, Y.; Chen, X. TransPVP: A Transformer-based method for ultra-short-term photovoltaic power forecasting. *Energies* **2024**, *17*, 4426.
58. De, V.; Teo, T.T.; Woo, W.L.; Logenthiran, T. Photovoltaic power forecasting using LSTM on limited dataset. In Proceedings of the 2018 IEEE Innovative Smart Grid Technologies—Asia (ISGT Asia), Singapore, 22–25 May 2018.
59. Asiedu, S.T.; Adebanji, B.; Ojo, O.; Munda, J.L. Machine learning forecasting of solar PV production using operating data from a grid-connected PV system. *Sci. Rep.* **2024**, *14*, 8298.
60. Lari, A.J.; Alharthi, M.; Khan, B.; Ullah, Z.; Khan, I. Using machine learning algorithms to forecast solar photovoltaic power output. *Electronics* **2025**, *14*, 866.
61. Sarmas, E.; Spiliotis, E.; Stamatopoulos, E.; Marinakis, V.; Doukas, H. Short-term photovoltaic power forecasting using meta-learning and numerical weather prediction independent Long Short-Term Memory models. *Renew. Energy* **2023**, *216*, 118997.
62. Kim, J.; Lee, J.; Kim, H.; Kang, Y. Multi-step photovoltaic power forecasting using Transformer networks. *Renew. Sustain. Energy Rev.* **2024**, *197*, 114388.
63. Simeunović, J.; Schubnel, B.; Alet, P.-J.; Carrillo, R.E. Spatio-temporal graph neural networks for multi-site PV power forecasting. *IEEE Trans. Sustain. Energy* **2022**, *13*, 1210–1220.
64. Simeunović, J.; Schubnel, B.; Alet, P.-J.; Carrillo, R.E. Interpretable temporal-spatial graph attention network for multi-site PV power forecasting. *Appl. Energy* **2023**, *327*, 120127.
65. Verdone, A.; De Felice, M.; Bosisio, A.; Merlo, M. Explainable spatio-temporal graph neural networks for multivariate solar power forecasting. *Appl. Energy* **2024**, *353*, 122049.

66. Yang, Y.; Liu, Y.; Wang, S.; Chen, Y.; Li, Z. DEST-GNN: A double-explored spatio-temporal graph neural network for multi-site photovoltaic power forecasting. *Appl. Energy* **2025**, *377*, 124575.
67. Chu, Y.; Wang, Y.; Yang, D.; Chen, S. A review of distributed solar forecasting with remote sensing and deep learning. *Renew. Sustain. Energy Rev.* **2024**, *198*, 114403. [CrossRef]
68. Xie, Y.; Chen, Y.; Zhang, H.; Wang, J.; Sun, H. An efficient approach for regional photovoltaic power forecasting based on satellite-ground observation fusion and hybrid neural networks. *Appl. Energy* **2025**, *385*, 125235.
69. Van der Meer, D.W.; Widén, J.; Munkhammar, J. Review on probabilistic forecasting of photovoltaic power production and electricity consumption. *Renew. Sustain. Energy Rev.* **2018**, *81*, 1484–1512. [CrossRef]
70. Li, B.; Zhang, J. A review on the integration of probabilistic solar forecasting in power systems. *Sol. Energy* **2020**, *207*, 777–795.
71. Sun, M.; Feng, C.; Zhang, J. Probabilistic solar power forecasting based on weather scenario generation. *Appl. Energy* **2020**, *266*, 114823. [CrossRef]
72. Xu, C.; Sun, Y. Quantile regression based probabilistic forecasting of renewable energy generation and building electrical load: A state of the art review. *J. Build. Eng.* **2023**, *79*, 107772. [CrossRef]
73. Pinson, P.; Tastu, J.; Browell, J.; Guo, Z. Quality assessment of renewable energy scenarios. *Appl. Energy* **2019**, *250*, 769–784.
74. Pinson, P.; Girard, R.; Kariniotakis, G.N. Uncertainty and probabilistic forecasting in power systems. *Int. J. Forecast.* **2019**, *35*, 1312–1330.
75. Kousounadis-Knousen, M.A. A review of solar power scenario generation methods: From statistical to deep learning approaches. *Energies* **2023**, *16*, 5600.
76. Zheng, K.; Sun, Z.; Song, Y.; Zhang, C.; Zhang, C.; Chang, F.; Yang, D.; Fu, X. Stochastic scenario generation methods for uncertainty in wind and photovoltaic power outputs: A comprehensive review. *Energies* **2025**, *18*, 503. [CrossRef]
77. Chen, Y.; Zhang, B.; Wang, J.; Sun, H. Scenario generation for renewable energy integration using deep generative models. *Energy* **2021**, *226*, 120379.
78. Khosravi, A.; Nahavandi, S.; Creighton, D.; Atiya, A.F. Lower upper bound estimation method for construction of neural network-based prediction intervals. *IEEE Trans. Neural Netw.* **2011**, *22*, 337–346. [PubMed]
79. Song, Z.; Zhang, H.; Pinson, P.; Yang, D. Probabilistic ultra-short-term solar photovoltaic power forecasting with conformal prediction. *Energy Rep.* **2025**, *12*, 418–431.
80. Hernandez Capel, E.; Dumas, J. Denoising diffusion probabilistic models for probabilistic energy forecasting. *arXiv* **2022**, arXiv:2212.02977.
81. Dumas, J.; Dvorkin, Y.; Arroyo, J.M. A comparison of deep generative models for probabilistic energy forecasting. *IEEE Trans. Power Syst.* **2024**, *39*, 3892–3905.
82. Mellit, A.; Tina, G.M.; Kalogirou, S.A. Fault detection and diagnosis methods for photovoltaic systems: A review. *Renew. Sustain. Energy Rev.* **2018**, *91*, 1–17. [CrossRef]
83. El-Banby, G.M.; Abdelgawad, A.F.; Abdelaziz, A.Y.; El-Sehiemy, R.A. Photovoltaic system fault detection techniques: A review. *Neural Comput. Appl.* **2024**, *36*, 589–613. [CrossRef]
84. Jordan, D.C.; Kurtz, S.R. Photovoltaic degradation rates—An analytical review. *Prog. Photovolt. Res. Appl.* **2013**, *21*, 12–29.
85. Köntges, M.; Kurtz, S.; Packard, C.; Jahn, U.; Berger, K.A.; Kato, K.; Friesen, T.; Liu, H.; Van Iseghem, M. *Review of Failures of Photovoltaic Modules*; IEA PVPS Task 13, Report IEA-PVPS T13-01:2014; International Energy Agency Photovoltaic Power Systems Programme: Paris, France, 2014.
86. Köntges, M.; Kurtz, S.; Packard, C.; Jahn, U.; Berger, K.A.; Kato, K.; Friesen, T.; Liu, H.; Van Iseghem, M. *Assessment of Photovoltaic Module Failures in the Field*; IEA PVPS Task 13, Report IEA-PVPS T13-09:2017; International Energy Agency Photovoltaic Power Systems Programme: Paris, France, 2017.
87. Kaplani, E. Detection of degradation effects in field-aged c-Si solar cells through IR thermography and digital image processing. *Int. J. Photoenergy* **2012**, *2012*, 396792.
88. Kurtz, S.; Whitfield, K.; TamizhMani, G.; Koehl, M.; Miller, D.; Joyce, J.; Wohlgemuth, J.; Bosco, N.; Kempe, M.; Zgonena, T. Evaluation of high-temperature exposure of photovoltaic modules. *Prog. Photovolt. Res. Appl.* **2011**, *19*, 954–965.
89. Munoz, M.A.; Alonso-García, M.C.; Vela, N.; Chenlo, F. Early degradation of silicon PV modules and guaranty conditions. *Sol. Energy* **2011**, *85*, 2264–2274. [CrossRef]
90. Pei, T.; Hao, X. A fault detection method for photovoltaic systems based on voltage and current observation and evaluation. *Energies* **2019**, *12*, 1712. [CrossRef]
91. Li, B.; Delpha, C.; Migan-Dubois, A.; Diallo, D. Fault diagnosis of photovoltaic panels using full I–V characteristics and machine learning techniques. *Energy Convers. Manag.* **2021**, *248*, 114785. [CrossRef]
92. Amiri, A.F.; Benhmida, M.; Bacha, S.; Migan-Dubois, A.; Diallo, D. Faults detection and diagnosis of PV systems based on machine learning approach and I–V characteristics. *Energy Convers. Manag.* **2024**, *306*, 118191.
93. Suliman, F.; Omer, S.; Riffat, S.; Al-Sulaiman, F. Electrical faults analysis and detection in photovoltaic arrays using support vector machine and extreme gradient boosting. *Sustainability* **2024**, *16*, 1102.

94. Tao, Y.; Li, Z.; Wang, H.; Zhang, Y.; Chen, X. Photovoltaic array fault diagnosis and localization based on modulated photocurrent and machine learning. *Sensors* **2025**, *25*, 136. [PubMed]
95. Adhya, D.; Bera, P.; Chatterjee, A. Stacking ensemble-based fault diagnosis approach for photovoltaic arrays. Energy Sources Part A Recovery Util. *Environ. Eff.* **2022**, *44*, 8964–8981.
96. Harrou, F.; Sun, Y.; Hering, A.S.; Madakyaru, M.; Dairi, A. Automatic fault detection in grid-connected photovoltaic systems using semi-supervised deep learning. *Energy Convers. Manag.* **2024**, *313*, 118681.
97. Yuan, Z.; Wang, Y.; Zhang, C.; Li, X.; Liu, W. Artificial neural network for fault diagnosis of solar photovoltaic systems: A survey. *Energies* **2022**, *15*, 8693. [CrossRef]
98. Bougoffa, M.; Talha, A.; Chaoui, H.; Benaouda, O. Hybrid deep learning for fault diagnosis in photovoltaic systems. *Machines* **2025**, *13*, 378. [CrossRef]
99. Kull, K.; Ahmed, S.; Niazi, K.A.K.; Yang, Y. Adaptive deep belief networks and LightGBM-based fault diagnosis for photovoltaic systems. *Electronics* **2025**, *14*, 3649.
100. Hussain, T.; Haider, S.A.; Muhammad, K.; Ullah, A.; Ullah, I.; Lee, M.Y. A review on defect detection of electroluminescence images of solar panels. *Energies* **2023**, *16*, 4012.
101. Masita, K.; Rahman, M.M.; Hannan, M.A.; Ker, P.J.; Begum, R.A.; Mahlia, T.M.I. Deep learning in defects detection of PV modules: A review. *Sol. Energy Adv.* **2025**, *11*, 100868.
102. Wang, J.; Bi, L.; Sun, P.; Jiao, X.; Ma, X.; Lei, X.; Luo, Y. Deep-learning-based automatic detection of photovoltaic cell defects in electroluminescence images. *Sensors* **2023**, *23*, 297. [PubMed]
103. Tang, W.; Yang, Q.; Hu, X.; Yan, W. Deep learning based automatic defect identification of photovoltaic module using electroluminescence images. *Sol. Energy* **2020**, *201*, 453–460. [CrossRef]
104. Tang, R.; Li, B.; Delpha, C.; Migan-Dubois, A.; Diallo, D. Fault classification of photovoltaic module infrared images using transfer learning-based convolutional neural networks. *Sol. Energy* **2024**, *273*, 112545.
105. Ukiwe, E.K.; Obiora, S.; Ajao, L.A.; Adewuyi, O.B. Deep learning model for detection of hotspots using infrared thermographic images of solar photovoltaic modules. *Smart Grids Sustain. Energy* **2024**, *9*, 19.
106. Lei, Y.; Wang, X.; An, A.; Guan, H. Deeplab-YOLO: A method for detecting hot-spot defects in infrared image PV panels by combining segmentation and detection. *J. Real-Time Image Process.* **2024**, *21*, 54.
107. Xie, H.; Zhang, Y.; Li, X.; Wang, S.; Chen, L. ST-YOLO: A defect detection method for photovoltaic modules based on YOLOv8s. *PLoS ONE* **2024**, *19*, e0310742. [PubMed]
108. Pratt, L.; Mattheus, J.; Klein, R. Defect detection and quantification in electroluminescence images of photovoltaic modules using deep learning. *Renew. Energy* **2021**, *178*, 1211–1222.
109. Liu, Y.; Zhang, H.; Li, X.; Wang, J.; Chen, Y. Deep learning-based method for defect detection in electroluminescence images of photovoltaic cells. *Opt. Express* **2024**, *32*, 17295–17312. [PubMed]
110. Lai, Y.S.; Hsieh, C.C.; Liao, T.W.; Huang, C.Y.; Kuo, C.F.J. Deep learning-based automatic defect detection of photovoltaic modules in infrared, electroluminescence, and red–green–blue images. *Energy Convers. Manag.* **2025**, *332*, 119783. [CrossRef]
111. Shamisavi, M.; Segovia Ramírez, I.; Gómez Muñoz, C.Q. Comparative evaluation of YOLO- and Transformer-based models for photovoltaic fault detection using thermal imagery. *Energies* **2026**, *19*, 845. [CrossRef]
112. Guo, Z.; Lu, J.; Chen, Q.; Liu, Z.; Song, C.; Tan, H.; Zhang, Y. TransPV: Refining photovoltaic panel detection accuracy through a vision transformer-based deep learning model. *Appl. Energy* **2024**, *355*, 122282.
113. Akram, M.W.; Li, G.; Jin, Y.; Chen, X.; Zhu, C.; Ahmad, A. CNN based automatic detection of photovoltaic cell defects in electroluminescence images. *Energy* **2019**, *189*, 116319. [CrossRef]
114. Vourkos, E.G.; Ioannou, A.; Fintikakis, N.; Tsekouras, G.J. Improving photovoltaics hotspot and dirt accumulation detection in aerial imagery using deep learning. *Sustain. Energy Technol. Assess.* **2025**, *74*, 104221.
115. Bosman, L.B.; Leon-Salas, W.D.; Hutzel, W.; Soto, E.A. PV system predictive maintenance: Challenges, current approaches, and opportunities. *Energies* **2020**, *13*, 1398. [CrossRef]
116. Hamza, A.; El-Aroussi, M.; Saadane, R.; El Maghraoui, A. A multi-stage review framework for AI-driven predictive maintenance and fault diagnosis in photovoltaic systems. *Appl. Energy* **2025**, *393*, 125838.
117. Chang, Z.; Li, X.; Wang, Y.; Zhang, H.; Zhou, D. Prognostics and health management of photovoltaic systems: A deep learning perspective. *Renew. Sustain. Energy Rev.* **2024**, *202*, 114613.
118. Vichare, R.V.; Nalawade, S.M.; Shinde, A.; Kulkarni, P. AI-based predictive maintenance of solar photovoltaics systems: A comprehensive review. *Energy Inform.* **2025**, *8*, 129. [CrossRef]
119. Ledmaoui, Y.; El Maghraoui, A.; El Aroussi, M.; Saadane, R. Review of recent advances in predictive maintenance and cybersecurity for solar plants. *Sensors* **2025**, *25*, 206. [CrossRef] [PubMed]
120. Ahmed, A.M.; Ali, M.; Mahmoud, K.; Lehtonen, M. Predictive maintenance of solar photovoltaic systems: A comprehensive review. *IET Renew. Power Gener.* **2025**, *19*, 70152. [CrossRef]

121. Marangis, D.; Makrides, G.; Georghiou, G.E. Intelligent maintenance approaches for improving photovoltaic system performance and reliability. *Sol. RRL* **2025**, *9*, 2500289. [CrossRef]
122. Apata, O.; Nwulu, N.; Dike, D.O.; Okoye, C.O. Artificial intelligence for predictive maintenance and fault detection in renewable energy systems. *Energies* **2026**, *19*, 536.
123. Hohm, D.P.; Ropp, M.E. Comparative study of maximum power point tracking algorithms. *Prog. Photovolt. Res. Appl.* **2003**, *11*, 47–62.
124. Esram, T.; Chapman, P.L. Comparison of photovoltaic array maximum power point tracking techniques. *IEEE Trans. Energy Convers.* **2007**, *22*, 439–449. [CrossRef]
125. Ishaque, K.; Salam, Z. A review of maximum power point tracking techniques of PV system for uniform insolation and partial shading condition. *Renew. Sustain. Energy Rev.* **2013**, *19*, 475–488. [CrossRef]
126. Mellit, A.; Kalogirou, S.A. Artificial intelligence techniques for photovoltaic applications: A review. *Prog. Energy Combust. Sci.* **2008**, *34*, 574–632. [CrossRef]
127. Yılmaz, M.; Çorapsız, M.R.; Çorapsız, M.F. A novel maximum power point tracking approach based on fuzzy logic control and optimizable Gaussian Process Regression for solar systems under complex environment conditions. *Eng. Appl. Artif. Intell.* **2025**, *139*, 109780.
128. Thajeel, S.M.; Atilla, D.Ç. Reinforcement Neural Network-Based Grid-Integrated PV Control and Battery Management System. *Energies* **2025**, *18*, 637.
129. Al Smadi, T.; Al-Omari, A.; Alshoubaki, Y.; Alshammari, N. Artificial intelligent control of energy management PV system. *Energy Rep.* **2024**, *10*, 450–461.
130. Ukoba, K.; Olatunji, K.O.; Adeoye, E.; Jen, T.-C. Optimizing renewable energy systems through artificial intelligence: Review and prospects. *Energy Environ.* **2024**, *35*, 3833–3879. [CrossRef]
131. Hu, J.; Xu, Y.; Cheng, K.W.; Guerrero, J.M. A model predictive control strategy of PV-Battery microgrid under variable power generations and load conditions. *Appl. Energy* **2018**, *221*, 195–203. [CrossRef]
132. Nakabi, T.A.; Toivanen, P. Deep reinforcement learning for energy management in a microgrid with flexible demand. *Sustain. Energy Grids Netw.* **2021**, *25*, 100413. [CrossRef]
133. Liu, D.; Zang, C.; Zeng, P.; Li, W.; Wang, X.; Liu, Y.; Xu, S. Deep reinforcement learning for real-time economic energy management of microgrid system considering uncertainties. *Front. Energy Res.* **2023**, *11*, 1163053. [CrossRef]
134. Alam, M.M.; Rahman, M.H.; Ahmed, M.F.; Chowdhury, M.Z.; Jang, Y.M. Deep learning based optimal energy management for photovoltaic and battery energy storage integrated home micro-grid system. *Sci. Rep.* **2022**, *12*, 15133. [CrossRef] [PubMed]
135. Lim, S.; Lee, J.; Lee, S. Model Predictive Control-Based Energy Management System for Cooperative Optimization of Grid-Connected Microgrids. *Energies* **2025**, *18*, 1696.
136. Nassereddine, K.; Turzynski, M.; Bielokha, H.; Strzelecki, R. Simulation of energy management system using model predictive control in AC/DC microgrid. *Sci. Rep.* **2025**, *15*, 5388. [CrossRef] [PubMed]
137. Areola, R.I.; Adebiyi, A.A.; Moloi, K. Artificial Intelligence for Optimizing Solar Power Systems with Integrated Storage: A Critical Review of Techniques, Challenges, and Emerging Trends. *Electricity* **2025**, *6*, 60. [CrossRef]
138. Nkambule, M.S.; Hasan, M.; Musasa, K. A review of intelligent control strategies for energy management in networked microgrids. *Energy Rep.* **2025**, *13*, 318–335.
139. Hosseini, E.; Shotorbani, A.M.; Khazaei, J.; Guerrero, J.M. Reinforcement learning-based energy management system for battery energy storage systems in multilevel microgrids. *Sustain. Energy Technol. Assess.* **2025**, *72*, 104011.
140. Naval, N.; Yusta, J.M. Virtual power plant models and electricity markets—A review. *Renew. Sustain. Energy Rev.* **2021**, *149*, 111393.
141. Kaiss, M.; Mekhilef, S.; Chand, A.A.; Seyedmahmoudian, M.; Horan, B. Review on Virtual Power Plants/Virtual Aggregators: Concepts, market participation, and future perspectives. *Renew. Sustain. Energy Rev.* **2025**, *207*, 114914.
142. Nguyen-Duc, H.; Nguyen-Hong, N. A study on the bidding strategy of the Virtual Power Plant in energy and reserve market. *Energy Rep.* **2020**, *6*, 622–626. [CrossRef]
143. Mei, S.; Tan, Q.; Liu, Y.; Trivedi, A.; Srinivasan, D. Optimal bidding strategy for virtual power plant participating in combined electricity and ancillary services market considering dynamic demand response price and integrated consumption satisfaction. *Energy* **2023**, *284*, 128592. [CrossRef]
144. Liu, R.; Zhang, Y.; Wang, J.; Li, Y.; Chen, X. Bidding strategy for the virtual power plant based on cooperative game theory considering carbon market and demand response. *Int. J. Electr. Power Energy Syst.* **2024**, *160*, 110120.
145. Liu, C.; Li, Z.; Wang, Y.; Sun, H. Supporting virtual power plants decision-making in urban environments using reinforcement learning. *Sustain. Cities Soc.* **2023**, *96*, 104673.
146. Li, J.; Zhang, H.; Wang, Y.; Chen, X. AI-Driven Virtual Power Plants: A Comprehensive Review. *Energies* **2026**, *19*, 1084. [CrossRef]
147. Zhou, J.; Chen, Y.; Li, M.; Wang, X. Data-Driven Intelligent Scheduling of Virtual Power Plants. *Energies* **2025**, *18*, 6341.

148. Jin, Y.; Li, S.; Zhang, X.; Wang, C. Market Applications and Uncertainty Handling for Virtual Power Plants. *Energies* **2025**, *18*, 3743. [CrossRef]
149. Huang, J.; Zhang, L.; Xu, H.; Liu, Y. Review of Virtual Power Plant Response Capability: Concepts, Evaluation, and Market Mechanisms. *Technologies* **2025**, *13*, 216.
150. Muteri, V.; Cellura, M.; Curto, D.; Franzitta, V.; Longo, S.; Mistretta, M.; Parisi, M.L. Review on Life Cycle Assessment of Solar Photovoltaic Panels. *Energies* **2020**, *13*, 252. [CrossRef]
151. Frischknecht, R.; Stolz, P.; Heath, G.; Raugei, M.; Sinha, P.; de Wild-Scholten, M.; Fthenakis, V.; Kim, H.C.; Alsema, E.; Held, M. *Life Cycle Inventories and Life Cycle Assessments of Photovoltaic Systems*; IEA PVPS Task 12, Report T12-19:2020; International Energy Agency Photovoltaic Power Systems Programme: Paris, France, 2020.
152. Stolz, P.; Frischknecht, R.; Wambach, K.; Sinha, P.; Heath, G. *Life Cycle Assessment of Current Photovoltaic Module Recycling*; IEA PVPS Task 12, Report T12-13:2018; International Energy Agency Photovoltaic Power Systems Programme: Paris, France, 2018.
153. Latunussa, C.E.L.; Ardente, F.; Blengini, G.A.; Mancini, L. Life Cycle Assessment of an innovative recycling process for crystalline silicon photovoltaic panels. *Sol. Energy Mater. Sol. Cells* **2016**, *156*, 101–111. [CrossRef]
154. Smith, B.L.; Margolis, R.; Feldman, D.; Woodhouse, M.; Heath, G. *An Updated Life Cycle Assessment of Utility-Scale Solar Photovoltaic Systems Installed in the United States*; NREL/TP-7A40-87372; National Renewable Energy Laboratory: Golden, CO, USA, 2024.
155. Fthenakis, V.; Kim, H.C. Photovoltaics: Life-cycle analyses. *Sol. Energy* **2011**, *85*, 1609–1628. [CrossRef]
156. IRENA; IEA-PVPS. *End-of-Life Management: Solar Photovoltaic Panels*; International Renewable Energy Agency and International Energy Agency Photovoltaic Power Systems Programme: Abu Dhabi, United Arab Emirates, 2016.
157. Babaei, A.; Shafiee, M.; Dinmohammadi, F.; Flynn, D. A Review of Photovoltaic Waste Management from a Sustainable Perspective. *Clean Technol.* **2024**, *6*, 1405–1430.
158. Lu, J.; Zhang, Y.; Wang, H.; Chen, X.; Li, Y. Recycling of end-of-life photovoltaic modules: A review of technologies, challenges, and artificial intelligence opportunities. *Chem. Eng. J.* **2025**, *507*, 160112.
159. Yan, C.; Zhang, Y.; Wang, J.; Li, X. Carbon-Aware Dispatch of Industrial Park Energy Systems with Demand Response and Stepped Carbon Pricing. *Sustainability* **2025**, *17*, 9472.
160. Figini, E.; Paolone, M. Achieving dispatchability in data centers: Carbon and cost-aware sizing of energy storage and local photovoltaic generation. *Sustain. Energy Grids Netw.* **2025**, *42*, 101920.
161. Prabawa, P.; Liu, Y.; Chen, X.; Wang, J. Carbon-aware mobile energy storage system scheduling in coupled power distribution and transportation systems. *Sustain. Energy Grids Netw.* **2025**, *43*, 101771.
162. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359.
163. Gama, J.; Žliobaitė, I.; Bifet, A.; Pechenizkiy, M.; Bouchachia, A. A survey on concept drift adaptation. *ACM Comput. Surv.* **2014**, *46*, 44. [CrossRef]
164. Kouw, W.M.; Loog, M. A review of domain adaptation without target labels. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 766–785. [PubMed]
165. Rabanser, S.; Günnemann, S.; Lipton, Z.C. Failing loudly: An empirical study of methods for detecting dataset shift. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 1396–1408.
166. Torralba, A.; Efros, A.A. Unbiased look at dataset bias. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011*; IEEE: New York, NY, USA, 2011; pp. 1521–1528.
167. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017*; Curran Associates: Red Hook, NY, USA, 2017; pp. 4768–4777.
168. Ribeiro, M.T.; Singh, S.; Guestrin, C. "Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 1135–1144.
169. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]
170. Gomes, E.; Sousa, J.M.; Dourado, A. Leveraging explainable artificial intelligence in solar photovoltaic power output mapping based on weather data. *Energies* **2025**, *18*, 1282.
171. Doshi-Velez, F.; Kim, B. Towards a rigorous science of interpretable machine learning. *arXiv* **2017**, arXiv:1702.08608.
172. Saad, B.; El Hannani, A.; Aqqal, A.; Errattahi, R. Toward trustworthy AI: An explainable framework for deterministic PV power forecasting using uncertainty estimation to separate model and data anomalies. *Results Eng.* **2025**, *27*, 107645.
173. National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*; NIST: Gaithersburg, MD, USA, 2023.
174. Conti, M.; Dehghantanha, A.; Franke, K.; Watson, S. Internet of Things security and forensics: Challenges and opportunities. *Future Gener. Comput. Syst.* **2018**, *78*, 544–546. [CrossRef]

175. Musleh, A.S.; Chen, G.; Dong, Z.Y. A survey on the detection algorithms for false data injection attacks in smart grids. *IEEE Trans. Smart Grid* **2020**, *11*, 2218–2234.
176. Raissi, M.; Perdikaris, P.; Karniadakis, G.E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **2019**, *378*, 686–707. [CrossRef]
177. Karniadakis, G.E.; Kevrekidis, I.G.; Lu, L.; Perdikaris, P.; Wang, S.; Yang, L. Physics-informed machine learning. *Nat. Rev. Phys.* **2021**, *3*, 422–440. [CrossRef]
178. Willard, J.; Jia, X.; Xu, S.; Steinbach, M.; Kumar, V. Integrating scientific knowledge with machine learning for engineering and environmental systems. *ACM Comput. Surv.* **2022**, *55*, 66. [CrossRef]
179. Parsa, S.M. Physics-informed machine learning meets renewable energy systems: A comprehensive review. *Appl. Energy* **2025**, *402*, 126093.
180. Li, H.; Zhang, Y.; Wang, J.; Chen, X. RTI-Net: Physics-informed deep learning for photovoltaic power forecasting with sky images. *Renew. Energy* **2025**, *238*, 121915.
181. Angelova, D.D.; Mavromatidis, G.; Dimeas, A.; Hatziargyriou, N. A review on digital twins and its application in the photovoltaic sector. *Energies* **2024**, *17*, 1227.
182. Elnosh, A.; Calais, M.; Parlevliet, D. A systematic literature review of digital twin research for photovoltaic systems: Trends, challenges, and opportunities. *Renew. Sustain. Energy Rev.* **2026**, *218*, 116326.
183. Tao, F.; Zhang, H.; Liu, A.; Nee, A.Y.C. Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inform.* **2019**, *15*, 2405–2415.
184. Qi, Q.; Tao, F. Digital twin and big data towards smart manufacturing and Industry 4.0: 360 degree comparison. *IEEE Access* **2018**, *6*, 3585–3593. [CrossRef]
185. Zhang, J.; Wang, Y.; Chen, X.; Li, H. Multimodal learning for renewable energy forecasting: A review. *Renew. Sustain. Energy Rev.* **2025**, *210*, 115207.
186. McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; Arcas, B.A.Y. Communication-efficient learning of deep networks from decentralized data. In Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, FL, USA, 20–22 April 2017; PMLR: Brookline, MA, USA, 2017; pp. 1273–1282.
187. Li, T.; Sahu, A.K.; Talwalkar, A.; Smith, V. Federated learning: Challenges, methods, and future directions. *IEEE Signal Process. Mag.* **2020**, *37*, 50–60. [CrossRef]
188. Kairouz, P.; McMahan, H.B.; Avent, B.; Bellet, A.; Bennis, M.; Bhagoji, A.N.; Bonawitz, K.; Charles, Z.; Cormode, G.; Cummings, R.; et al. Advances and open problems in federated learning. *Found. Trends Mach. Learn.* **2021**, *14*, 1–210. [CrossRef]
189. Zhang, C.; Li, S.; Xia, J.; Wang, W.; Yan, F.; Liu, Y. BatchCrypt: Efficient homomorphic encryption for cross-silo federated learning. In *Proceedings of the 2020 USENIX Annual Technical Conference, Virtual Event, 15–17 July 2020*; USENIX Association: Berkeley, CA, USA, 2020; pp. 493–506.
190. Han, S.; Mao, H.; Dally, W.J. Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding. In Proceedings of the International Conference on Learning Representations, San Juan, Puerto Rico, 2–4 May 2016.
191. Hinton, G.; Vinyals, O.; Dean, J. Distilling the knowledge in a neural network. *arXiv* **2015**, arXiv:1503.02531.
192. Jacob, B.; Kligys, S.; Chen, B.; Zhu, M.; Tang, M.; Howard, A.; Adam, H.; Kalenichenko, D. Quantization and training of neural networks for efficient integer-arithmetic-only inference. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2704–2713.
193. Sze, V.; Chen, Y.-H.; Yang, T.-J.; Emer, J.S. Efficient processing of deep neural networks: A tutorial and survey. *Proc. IEEE* **2017**, *105*, 2295–2329. [CrossRef]
194. Strubell, E.; Ganesh, A.; McCallum, A. Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Florence, Italy, 28 July–2 August 2019*; ACL: Stroudsburg, PA, USA, 2019; pp. 3645–3650.
195. Schwartz, R.; Dodge, J.; Smith, N.A.; Etzioni, O. Green AI. *Commun. ACM* **2020**, *63*, 54–63. [CrossRef]
196. Patterson, D.; Gonzalez, J.; Le, Q.; Liang, C.; Munguia, L.M.; Rothchild, D.; So, D.; Texier, M.; Dean, J. Carbon emissions and large neural network training. *arXiv* **2021**, arXiv:2104.10350.
197. Hassan, S.Z.; Sana, B.; Kamal, T.; Masood, A.; Ahmed, M.; Shafiullah, M. A compact 1D-CNN for photovoltaic fault diagnosis with leakage-aware validation: A comparative study. *IEEE Access* **2026**, *14*, 81616–81630.
198. Aljohani, T.M. Markov-constrained isolation forest for early detection of battery anomalies in solar-grid applications. *Mathematics* **2026**, *14*, 1192. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.