

Review

AI-Based Multi-Timescale Photovoltaic Power Scenario Generation and Forecasting: A Statistical Relational Perspective

Yanan Cui ¹, Xiao Lv ¹, Chunyu Zhang ¹, Xuanye Zhao ² and Xueqian Fu ^{1,*} 

¹ College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China; cuiyanan@cau.edu.cn (Y.C.); m18839345976_1@163.com (X.L.); s20223081693@cau.edu.cn (C.Z.)

² School of Information Technology, Beijing City University, Shunyi District, Beijing 101309, China; zxy_071012@bcu.edu.cn

* Correspondence: fuxueqian@cau.edu.cn

Abstract

The output of photovoltaic power generation exhibits significant randomness, volatility, and intermittency, and its uncertainty will have an impact on the planning assessment, dispatch decision-making, and real-time operation of the power system. To systematically understand the development status of modeling methods for the uncertainty of photovoltaic power generation, this paper conducts a review around the generation of annual scenarios and multi-timescale power prediction of photovoltaic power, and analyzes the correlations between photovoltaic output, influencing factors, and system applications from the perspective of statistical relationships and artificial intelligence. For the annual scale, the focus is on the generation methods of meteorological-driven scenarios for long-term sequences, including probability statistical methods, deep generation methods, constraint relations and engineering application evaluation issues; for the day-ahead scale, the historical power, meteorological variables and numerical weather forecasts are used to explore feature extraction, probability prediction and robust modeling methods; for the intraday scale, the signal decomposition, deep learning, regional collaborative modeling and multi-source perception methods for short-term power fluctuations perception are summarized. On this basis, further analysis is conducted on subsequent research directions such as multi-timescale collaborative modeling, multi-source heterogeneous information fusion, controllable generative modeling, extreme scenario characterization, and engineering closed-loop verification. This paper can serve as a reference for scenario generation, power prediction, and power system operation analysis under the condition of a high proportion of photovoltaic power integration.

Keywords: photovoltaic power generation; statistical relational artificial intelligence; scenario generation; power forecasting; multi-timescale modeling



Academic Editor: Fabrice Goubard

Received: 23 June 2026

Revised: 10 July 2026

Accepted: 16 July 2026

Published: 18 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 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Driven by the low-carbon energy transition and the large-scale development of renewable energy, photovoltaic (PV) power generation is gradually evolving from a supplementary power source into an important component of modern power systems. PV output is jointly affected by solar irradiance, ambient temperature, cloud evolution, weather processes, PV plant operating conditions, and grid-connected control strategies, exhibiting both periodic patterns and stochastic disturbances. With the increasing penetration of PV generation, output deviations no longer only affect the generation schedules of individual

PV plants but also influence power system planning, dispatch operation, reserve allocation, energy storage control, electricity market participation, and risk assessment. Therefore, scenario generation and power forecasting for PV uncertainty have become important research topics in renewable power system analysis and AI-enabled energy engineering [1–3].

From the perspective of timescales, PV uncertainty modeling can be divided into three interrelated tasks with different emphases: annual scenario generation, day-ahead power forecasting, and intraday power forecasting. Annual scenario generation focuses on constructing long-term PV output or PV-related meteorological variable sequences, with particular attention to seasonal variation, intra-annual periodicity, long-term distribution characteristics, and the coverage of extreme samples. It mainly supports capacity planning, energy storage configuration, probabilistic power flow analysis, and operational risk assessment [4–7]. Day-ahead power forecasting aims to predict PV output variations over the next day or several days, usually relying on historical power data, numerical weather prediction (NWP), and multiple meteorological features, thereby supporting next-day dispatch, reserve scheduling, and electricity market bidding [8]. Intraday power forecasting emphasizes real-time perception of fluctuations and rolling refinement. It requires the use of the latest power measurements, local meteorological data, sky images, satellite cloud images, or multi-site collaborative information to capture rapid power fluctuations caused by cloud shading, irradiance ramps, and local weather disturbances [9–11]. Although these three tasks correspond to different application stages, they all essentially require the characterization of statistical dependencies among meteorological conditions, temporal structures, spatial correlations, and PV output.

A relatively rich methodological system has been developed for PV scenario generation and forecasting. Traditional probabilistic and statistical methods, such as distribution fitting, time-series models, Markov chains, Monte Carlo simulation, Copula functions, and representative scenario reduction, provide clear model structures and explicit statistical meanings [12–14]. Machine learning, deep learning, and generative AI methods further enhance nonlinear mapping, temporal feature extraction, multivariable relationship learning, and scenario simulation capabilities [4–9]. These developments provide important methodological foundations, but they also raise new requirements for relational consistency, physical plausibility, and engineering applicability.

However, several issues in PV scenario generation and forecasting still require further systematic investigation. First, studies at different timescales are often conducted as independent tasks. Annual scenario generation emphasizes sample coverage and long-term statistical consistency; day-ahead forecasting focuses on trend accuracy and probabilistic intervals; and intraday forecasting stresses short-term disturbance perception and rolling updates. As a result, model objectives, input features, and evaluation metrics still lack a unified interpretation. Second, although some AI models can reduce forecasting errors, they insufficiently consider the coupling relationships among meteorological variables, day–night boundaries, seasonal patterns, spatial propagation, and PV power formation mechanisms, which may lead to results that are difficult to interpret or inconsistent with engineering constraints. Third, under non-ideal conditions, such as extreme weather, equipment anomalies, missing measurements, cross-regional transfer, and sample distribution shifts, model generalization may deteriorate, thereby affecting the reliability of these models in planning and operational decision-making.

Statistical relational AI provides a unified analytical perspective for addressing the above issues. In this context, the “statistical” aspect emphasizes the probabilistic characterization of PV power, meteorological variables, forecasting errors, and extreme fluctuations; the “relational” aspect focuses on the structured relationships among irradiance, temperature, cloud cover, temporal cycles, spatial locations, equipment states, power output,

and system operation requirements; and the “AI” aspect provides implementation tools for feature extraction, nonlinear modeling, probabilistic inference, multi-source fusion, and generative modeling [1,2]. From this perspective, annual scenario generation, day-ahead forecasting, and intraday forecasting can be regarded as three parallel components within the same uncertainty modeling chain. Annual scenario generation forms a set of long-term possible states; day-ahead forecasting provides future operational trends and risk boundaries; and intraday forecasting refines short-term fluctuations based on real-time information. Together, these tasks support the representation, propagation, and decision-making application of PV output uncertainty.

The coupling of PV generation with agricultural energy systems, fishery production, and nearshore application scenarios further broadens the application context for statistical relational AI. Existing studies on agricultural energy Internet and fishery weather simulation have shown that complex energy applications require the joint consideration of meteorology, scenarios, equipment, regions, and system-level objectives [15,16]. Therefore, AI-based PV scenario generation and forecasting should not focus solely on fitting a single power curve but should also emphasize statistical relational representation and application-oriented validation.

Beyond these methodological and application-related challenges, it is also necessary to clarify the position of this review in relation to existing review studies. Existing review papers have discussed solar power scenario generation from the perspectives of weather classification, temporal horizons, and deep generative models [17]. Comprehensive PV forecasting reviews have summarized data acquisition, data preprocessing, forecasting model development, performance evaluation, uncertainty quantification, and adaptive learning within data-driven forecasting frameworks [18]. Deep-learning-oriented reviews have focused on neural architectures and hybrid models for PV power forecasting, including CNN, RNN, LSTM, Transformer, and related variants [19,20]. Probabilistic solar forecasting reviews have mainly discussed prediction intervals, uncertainty quantification, calibration, and the integration of probabilistic forecasts into power-system operation and decision-making [21,22]. However, these reviews generally emphasize either scenario generation, forecasting model taxonomy, deep learning architectures, or probabilistic forecast utilization. The connections among annual-scale 8760 h scenario construction, day-ahead uncertainty prediction, intraday fluctuation perception, and engineering-oriented validation have not been sufficiently discussed within a unified relational uncertainty-modeling chain.

Against this background, this paper reviews annual scenario generation, day-ahead forecasting, and intraday forecasting methods for PV power generation from the perspective of statistical relational AI. First, the evolution of PV uncertainty modeling from traditional probabilistic statistical methods to machine learning, deep learning, and generative AI is summarized. Second, for annual-scale scenario generation, probabilistic statistical models, deep generative models, relational modeling, knowledge constraints, and periodic feature enhancement methods are reviewed. Third, for day-ahead PV power forecasting, feature extraction, probabilistic forecasting, and robust modeling methods driven by historical power data, meteorological variables, and NWP information are summarized. Furthermore, for intraday PV power forecasting, signal decomposition, deep learning, regional collaborative modeling, and multi-source perception methods for short-term power fluctuation awareness are discussed. Finally, future research directions are outlined in terms of multi-timescale collaborative modeling, multi-source heterogeneous information fusion, controllable generative modeling, extreme scenario characterization, and engineering closed-loop validation. This paper aims to provide a reference for scenario generation, power forecasting, and system operation analysis under high PV penetration.

Compared with the above reviews, the novelty of this review lies in its statistical relational AI perspective and its multi-timescale organization. Instead of treating scenario generation, day-ahead forecasting, and intraday forecasting as independent research topics, this review regards them as interconnected components of PV uncertainty modeling. In this way, the review emphasizes not only model accuracy but also uncertainty distributions, meteorological–power coupling, temporal periodicity, spatial correlation, physical constraints, and engineering applicability.

The specific contributions of this review are summarized as follows:

- (1) A multi-timescale review framework is established to connect annual PV scenario generation, day-ahead forecasting, and intraday forecasting.
- (2) Existing methods are reorganized from the perspective of statistical relational AI, with emphasis on uncertainty characterization, variable coupling, temporal and spatial relationships, and physical consistency.
- (3) Traditional probabilistic models, machine learning methods, deep learning models, and generative AI methods are discussed and compared with attention to statistical consistency, relational preservation, physical plausibility, and engineering applicability.
- (4) Future research directions are discussed from the perspectives of multi-timescale collaborative modeling, multi-source heterogeneous information fusion, controllable generative modeling, extreme scenario characterization, and engineering closed-loop validation.

2. Related Work

This review was compiled by searching the IEEE Xplore, Web of Science, and Scopus databases for publications from January 2015 to July 2026, using combinations of the keywords “photovoltaic”, “scenario generation”, “power forecasting”, “uncertainty”, “deep learning”, “generative adversarial network”, “statistical relational”, and “multi-timescale”. Only peer-reviewed journal articles, peer-reviewed conference papers, and book chapters focused on AI-based or statistical modeling of PV output uncertainty were included; studies limited to hardware- or device-level PV research (e.g., cell materials, inverter design) without an uncertainty-modeling or forecasting component were excluded. After removing duplicates and screening titles and abstracts for relevance, the resulting studies were organized according to the three timescales addressed in this review (annual, day-ahead, and intraday) and further classified by methodological category (probabilistic-statistical, machine learning, deep learning, and generative AI).

To further clarify the position of this review, existing review studies can be grouped into several streams. Solar scenario generation reviews mainly summarize stochastic scenario construction, weather classification, temporal horizons, and deep generative models [17]. Comprehensive PV forecasting reviews usually organize data acquisition, data preprocessing, prediction model development, accuracy evaluation, uncertainty quantification, and adaptive learning within data-driven forecasting frameworks [18,23,24]. Deep learning-oriented reviews focus on model architectures, such as CNNs, RNNs, LSTMs, Transformers, and related hybrid networks, as well as datasets, input variables, forecasting horizons, and evaluation metrics [19,20]. Probabilistic forecasting reviews mainly discuss prediction intervals, uncertainty quantification, calibration, and the utilization of probabilistic forecasts in power-system scheduling, operation, and market participation [21,22].

Although these studies provide valuable taxonomies and summaries, they usually focus on a single research stream. Scenario generation reviews pay more attention to scenario construction techniques; forecasting reviews mainly focus on prediction accuracy and model comparison; deep learning reviews are often architecture-oriented; and probabilistic forecasting reviews emphasize uncertainty representation and decision support. In contrast, this review connects annual scenario generation, day-ahead forecasting, and

intraday forecasting through the perspective of statistical relational AI. Therefore, it aims to reveal the common statistical and relational issues behind different timescales, including probability distributions, meteorological–power coupling, temporal and spatial dependencies, physical constraints, and engineering-oriented evaluation. Table 1 summarizes the differences between existing review perspectives and the present review.

Table 1. Comparison of existing review perspectives and the present review.

Existing Review Entry Point	Core Concern of Existing Reviews	Main Limitation	Novelty of This Review
Scenario generation-oriented reviews [17]	How to generate possible solar/PV scenarios	Scenario generation is often discussed separately from forecasting tasks	Places annual scenarios within a multi-timescale PV uncertainty modeling chain
Forecasting-oriented reviews [18,23,24]	How to improve solar/PV forecasting accuracy	Annual scenario construction and uncertainty propagation are less emphasized	Connects annual scenario generation, day-ahead forecasting, and intraday forecasting
Model-architecture-oriented reviews [19,20]	How CNN, RNN, LSTM, Transformer, and hybrid models are used	The discussion is often model-centered rather than relationship-centered	Reorganizes AI methods from statistical relationships, physical constraints, and relational consistency
Probabilistic-forecasting-oriented reviews [21,22]	How to quantify forecast uncertainty using intervals or distributions	Probabilistic forecasting is not sufficiently linked with scenario generation and applications	Connects probabilistic forecasting with scenario generation, risk-aware decision-making, and application-oriented evaluation

In line with the above multi-timescale perspective, original studies on PV power uncertainty can be further understood from three related aspects: scenario construction, power forecasting, and engineering applications. Scenario construction focuses on possible PV output states, power forecasting aims to estimate future output variations, and engineering applications emphasize whether model outputs can support dispatching, planning, risk assessment, and operational analysis. Traditional studies are mainly based on probabilistic models and statistical inference, using distribution fitting, correlation modeling, and random sampling to describe the uncertainty of PV power and meteorological variables [12,13]. These methods are relatively interpretable, but they are often constrained by strong assumptions when dealing with long sequences, multivariable coupling, and complex non-stationary fluctuations.

In terms of scenario generation, research has shifted from random sampling toward joint modeling of temporal structures, variable relationships, and domain knowledge. Diffusion models, generative adversarial networks, and cross-modal generative models provide new approaches for renewable energy time-series generation [25,26]. For PV scenario simulation, Fu et al. enhanced the ability of GANs to represent PV temporal dynamics by introducing Bi-LSTM and style transfer mechanisms [4]. They further integrated meteorological statistical patterns and PV physical conversion processes into generative models, improving the rationality of year-round weather sequence simulation [5]. In addition, diffusion Transformers enhance the diversity and time–frequency feature preservation of generated time series by redesigning the decoder and loss function [6]. Fu et al. also extracted representative PV scenarios from the perspectives of power, energy, and fluctuation patterns, providing compact samples for planning calculation and operational evaluation [7]. Moreover, studies on physically constrained WGANs, extreme-weather-labeled joint wind–PV–load generation, and dynamic temporal correlation modeling indicate that scenario generation is evolving from distribution fitting toward relationship preservation and knowledge enhancement [27–29].

For day-ahead forecasting, existing studies mainly focus on PV output trends and uncertainty boundaries over the next day or several days. Early methods commonly relied on historical power data, weather forecasts, and shallow learning models for point

forecasting or error correction [30,31]. With the development of deep learning, CNNs, LSTMs, GRUs, Transformers, TCNs, and ensemble learning models have been used to extract meteorological features and capture temporal dependencies [8,11]. The clustering-ensemble learning method proposed by Fu et al. can adapt to different weather patterns [8]. Multi-scale patch and frequency-domain gated learning improve the capability of capturing high-resolution power variations [32]. The GEV distribution-enhanced Fourier diffusion model further improves the representation of tail fluctuations and extreme PV output variations [33].

For intraday and ultra-short-term forecasting, the main challenge lies in strong fluctuations and rapid variations over short timescales. Cloud movement, local shading, irradiance ramps, and sensor anomalies can all lead to non-stationary PV output variations. Therefore, existing studies commonly combine signal decomposition, deep temporal networks, graph-structured modeling, and multi-source perception information. Sky-image-based methods can extract cloud distribution and motion information [10,34], while hybrid models combining fuzzy inference and deep learning enhance the modeling capability for complex nonlinear relationships [35]. In regional collaborative forecasting, Fu et al. proposed a Fourier graph-based spatiotemporal learning framework combined with corrected NWP, enabling joint modeling of frequency features and spatial correlations in multi-site ultra-short-term PV forecasting [9]. Adaptive masked networks further improve robustness under incomplete data and abrupt variation conditions [36].

Overall, PV scenario generation and power forecasting research is shifting from improving the accuracy of individual models toward the comprehensive characterization of statistical distributions, temporal structures, spatial correlations, physical constraints, and engineering requirements. Although annual scenario generation, day-ahead forecasting, and intraday forecasting correspond to different timescales, they all involve the relational representation among PV output, meteorological factors, temporal cycles, spatial locations, and system requirements. Statistical relational AI provides a unified perspective for this purpose: statistical modeling describes uncertainty, relational modeling explains multi-factor coupling, and AI methods enable high-dimensional feature learning and complex pattern generation [1–3]. The studies by Fu et al. on agricultural energy Internet and fishery weather simulation and forecasting also show that complex energy applications need to consider the coupling relationships among data, environment, equipment, and application objectives [15,16].

To enable consistent comparison across the diverse methods reviewed in this paper, we adopt four categories of evaluation criteria throughout: (i) distributional-fidelity metrics, which quantify how closely generated or forecasted PV output matches the statistical properties of observed data, including mean, variance, quantiles, autocorrelation, and mode/tail coverage; (ii) relational-consistency metrics, which assess whether physical constraints (e.g., zero nighttime output), diurnal/seasonal structure, and spatial correlations are preserved; (iii) computational-cost indicators, including training/inference time and data requirements; and (iv) downstream engineering-utility metrics, which evaluate the impact of generated scenarios or forecasts on power-flow analysis, storage sizing, reserve estimation, and other system-level applications. These four categories are applied consistently in the comparative tables (Tables 2–4) and in the discussion of evaluation frameworks in Sections 3.3 and 6.5.

3. Annual-Scale Photovoltaic Power Generation Scenario Generation Method

PV power generation exhibits significant intermittency and variability. The generation of 8760 h annual-scale output scenarios serves as the foundation for PV system simulation,

planning, energy storage configuration, and risk assessment. Figure 1 illustrates the overall technical framework and engineering application contexts of annual-scale PV scenario generation.

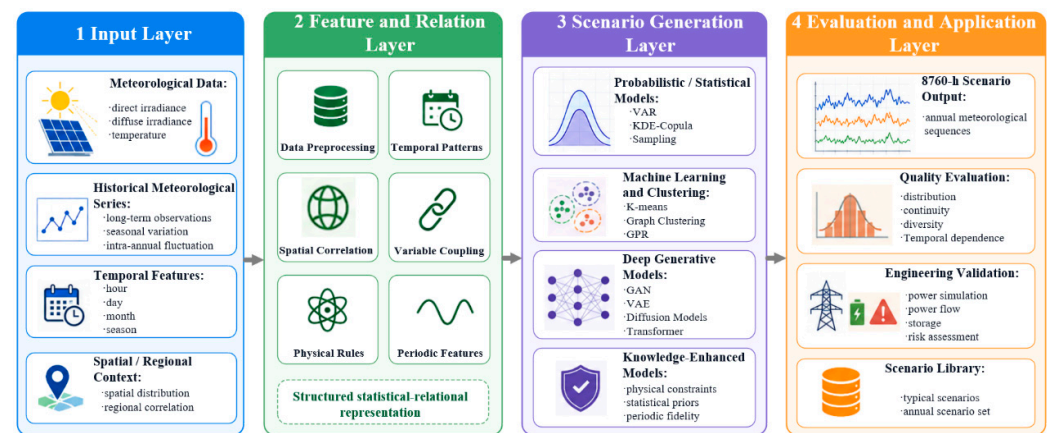


Figure 1. Framework for annual-scale photovoltaic power generation scenario generation and engineering applications.

Unlike day-ahead short-term scenario generation, annual-scale scenario generation must simultaneously characterize diurnal cycles, seasonal cycles, and inter-annual climate variability, while ensuring temporal autocorrelation, inter-site spatial dependencies, and physical coherence constraints.

3.1. Technical Evolution of Scenario Generation Methods

Based on the core generation techniques, existing methods can be grouped into three paradigms: stochastic generation via explicit probabilistic modeling, scenario construction via feature engineering and clustering, and implicit distribution learning via deep generative models.

3.1.1. Stochastic Scenario Generation Based on Explicit Probabilistic Modeling

These methods follow a two-stage pipeline: first, estimate the probability distribution, then sample from it. Their main strengths are model transparency and ease of embedding domain knowledge.

Vector autoregressive (VAR) models were among the earliest applied in this area. A clear-sky model is used to decompose irradiance into a deterministic astronomical component and a stochastic perturbation. Multi-site spatial correlation is then injected through Cholesky decomposition or a t-copula, and a PV physical model converts the results into power output scenarios [12]. The VAR(2) framework has been further employed for spatiotemporal modeling of aggregate forecast errors, fitting a spatial correlation function with a rational function of geographic distance [13].

To overcome the limitations of linear-Gaussian assumptions, copula functions achieve nonlinear modeling by separating marginal distributions from the dependence structure. For instance, kernel density estimation together with a copula builds a joint meteorological distribution, and Markov chain Monte Carlo (MCMC) generates annual sequences. These sequences are linked with load disaggregation to produce coupled solar-load scenarios [37]. Adaptive diffusion kernel density estimation with AMISE-optimal bandwidth further improves density estimation accuracy [38]. Nonparametric methods adopt conditional kernel density estimation (CKDE) to model the multi-scale correlation between daily total irradiance and hourly output. A Gram–Schmidt orthogonal transformation decomposes the hourly components under a summation constraint [39]. For extreme scenarios, ERA5

reanalysis data (30 years, hourly) are fused with grid operational data. An optimization-based inversion obtains the equivalent installed capacity distribution, and an adequacy threshold screens extreme cold-wave events. This effectively expands the sample set of extreme scenarios with long return periods [40].

3.1.2. Scenario Construction Methods Based on Feature Engineering and Clustering

These methods do not directly estimate probability densities. Instead, they select representative scenarios from historical data through multidimensional similarity measures. In essence, they compress the 8760 h of a full year into a small set of typical days, while preserving the statistical representativeness of power levels, energy distribution, and fluctuation characteristics.

Graph clustering defines edge weights using a four-dimensional similarity measure: power Euclidean distance, daily energy generation, curve shape trend, and fluctuation position. Scenario extraction is then formulated as a graph cut optimization problem and solved via alternating optimization [7]. Shared Weight Graph Clustering (SWGCG) introduces adaptive feature weights and neighborhood density weights to enhance noise robustness and retain output correlations across multiple stations [41].

In the feature-matching direction, a composite index that fuses geometric correlation and distance correlation screens similar days. The sampling probability is corrected by fluctuation intervals, and scenario reduction is completed through ISODATA adaptive clustering [42]. From the perspective of energy balance, a probabilistic output sequence is generated by matching features at the annual, monthly, and hourly levels [43]. Mapping solar-load features with a Gaussian kernel, together with improved ISODATA and NSGA-III, enables integrated optimization of clustering and energy storage configuration [44].

3.1.3. Implicit Distribution Learning Methods Based on Deep Generative Models

Deep generative models allow neural networks to learn an empirical distribution directly from historical data. No explicit probabilistic assumptions are required.

Generative adversarial networks (GANs) employ convolutional networks as the generator and the discriminator. The Wasserstein distance is used to improve training stability. Conditional GANs realize generation conditioned on weather or season [45]. A transparent latent space with orthogonal regularization makes the statistical properties of the generated scenarios controllable [46]. Conditional GANs have been systematically extended to multi-regional modeling. The discrete wavelet packet transform (DWPT) decomposes power into low- and high-frequency bands, which are modeled separately [14]. Three-dimensional spatio-temporal tensors are combined with 3D-CNNs to capture both inter-regional and inter-time-step correlations [47]. A CNN-CGAN conditioned on forecasts produces multi-regional scenario sets with more concentrated spatial correlation [48]. StyleGAN2 style mixing achieves multi-resolution feature control and supports temporal extrapolation [49].

For temporal enhancement, bidirectional LSTM replaces CNN in the generator to capture long-term dependencies. This is combined with adaptive instance normalization for style transfer and with sliding window data augmentation [4]. HiLo attention, where the high-frequency branch captures cloud fluctuations and the low-frequency branch captures the daily envelope, is embedded into WGAN-GP [50]. A knowledge enhancer explicitly enforces zero nighttime radiation, and an external physical model ensures consistency, resulting in MWGAN [5]. For multi-energy extension, WGAN-1D-CNN jointly models solar, wind, electric, and thermal energy and achieves arbitrary-duration forecasting via constrained optimization [51].

To overcome the bottleneck of GAN training instability, non-adversarial paths have been explored. Implicit maximum likelihood estimation (IMLE) replaces the adversarial

loss with a Euclidean distance [52]. The diffusion Transformer, Weather-DiT, replaces adversarial training with iterative denoising. In its decoder, a polynomial branch learns the trend, and a DFT branch learns the periodicity. A joint frequency-probability loss (FP loss) constrains distributional consistency in both the time domain and the frequency domain [6].

3.2. Relationship Modeling, Knowledge Constraints, and Periodic Feature Enhancement Methods

Annual scenario generation needs to attain resemblance at the distributional level, and must likewise satisfy physical consistency constraints and spatiotemporal characteristic requirements. These include the physical relationships among irradiance, temperature, and PV output; the principle that output is zero at night; the temporal features of daily and seasonal periodicity; and the spatial correlation characteristics across multiple stations. The following analysis is conducted from three dimensions: physical constraints and knowledge embedding, characterization of multi-scale periodic features, and spatial correlation modeling.

3.2.1. Physical Constraints and Knowledge Embedding

Regarding physical constraints and knowledge embedding, ensuring the physical plausibility of generated scenarios is a critical requirement that distinguishes PV scenario generation from generic generative models. Separating deterministic astronomical factors from stochastic meteorological perturbations ensures the fundamental regularities of solar irradiance through physical models and implicitly embeds physical constraints into source-load coupling relationships [12,37,40]. At the engineering physics level, annual production simulation is driven by the physical mapping between monthly PV efficiency and irradiance [53]; a complete equipment mechanism model chain spanning PV modules, inverters, and MPPT is established [54]; physical boundaries under extreme operating conditions are derived from the analytical line voltage drop formula [55]. Among deep generative models, MWGAN enforces the day-night physical constraint of zero nighttime radiation on generated data through a statistical probability knowledge enhancer [5]. Weather-DiT first generates meteorologically consistent weather data and subsequently converts it into PV output through an external PV physical model, thereby indirectly ensuring the physical plausibility of the final scenarios [6].

3.2.2. Multi-Scale Periodicity Characterization

Regarding multi-scale periodicity characterization, annual scenarios must simultaneously capture characteristics across three temporal scales: diurnal cycles, seasonal cycles, and inter-annual climate variability. Traditional methods typically handle periodic features in an explicit manner: diurnal structures are removed to achieve series stationarity, thereby isolating the diurnal cycle from stochastic components [12]; weather-type classification explicitly distinguishes diurnal curve shape differences under different weather conditions [37]. Clustering-based methods implicitly preserve seasonal diversity by extracting representative daily scenarios from full-year data [7,41]. Deep generative models offer more flexible and diverse approaches to periodic feature characterization: CGANs employ seasonal or monthly labels as conditional inputs to achieve targeted generation for different time periods [5,45]; PV output is decomposed into low-frequency components capturing seasonal variations and high-frequency components capturing intra-day variability [14,50]; Bi-LSTM, with its bidirectional recurrent structure, explicitly models long-range temporal dependencies, naturally suited to the gradual seasonal variation characteristics of PV output [4].

3.2.3. Spatial Correlation Modeling

Regarding spatial correlation modeling, the spatial interdependence of outputs from multiple PV stations is a factor that cannot be neglected in regional power system planning and operation. Traditional statistical methods handle spatial dependence through covariance structures: t-Copula structures inject multi-site spatial correlation, with distance-weighted interpolation used to estimate model parameters at new locations lacking measurements [12]; spatial correlation coefficients are modeled as rational functions of inter-site geographic distance, fitted via least squares [13]; Gram–Schmidt orthogonal transformation preserves inter-station correlation structures during multi-station decomposition [39]. Deep generative methods, by contrast, learn spatial correlation automatically in a data-driven manner: multi-channel CNN inputs directly learn the joint distribution of multi-station PV output [45]; multi-regional data are constructed as three-dimensional spatio-temporal tensors, with 3D-CNNs jointly capturing both inter-regional and inter-temporal dependencies within a unified architecture [47]; the transparent latent space design of Ctrl-GAN also offers new possibilities for controllable generation of spatial features [46]; Yuan et al. implicitly capture the spatial correlation of heterogeneous energy sources through cross-channel style mixing mechanisms in multi-energy joint generation [49].

In summary, annual-scale PV scenario generation methods exhibit an evolutionary trend from explicit statistical modeling to deep implicit learning, and from purely data-centric methods to knowledge-informed collaborative modeling. Table 2 provides a systematic comparison of the methods reviewed in this paper along the above dimensions.

Table 2. Comparison of annual-scale photovoltaic scenario generation methods.

Method Category	Typical Methods or Models	Physical Consistency	Advantages	Limitations
Explicit probabilistic modeling	VAR [12,13]	Separates astronomical determinacy and random perturbation via a physical model	Transparent structure, strong interpretability	Linear assumption restricts the capture of nonlinear dynamics
	Copula and nonparametric methods [37–39]; extreme scenario reanalysis [40]	Decouples marginal and dependence structures via a copula	Flexible modeling of nonlinear dependence; can characterize extreme interannual variability	Reliance on handcrafted features; complex high-dimensional parameterization
Feature engineering and clustering	Graph clustering [7,41]	Eigenvalue structure reflects dominant temporal patterns	Multidimensional similarity surpasses single distance	Generates only representative snapshots
	ISODATA and feature-matching clustering [42–44]	Screens similar days via composite similarity; adaptive cluster count preserves seasonal statistics	Streamlined process; seamless integration with downstream optimization	Diversity bounded by historical data
Deep generative models	WGAN/CGAN and variants [14,45–49]	Implicitly learns spatio-temporal distribution via adversarial training	End-to-end, model-free; conditional generation controllable	Adversarial training instability; hyperparameter sensitivity
	Temporal-enhanced and knowledge-integrated GANs [4,5,50,51]	Frequency decomposition separates trend and fluctuation	Enhanced temporal modeling; knowledge injects physical constraints	Data augmentation may introduce artificial periodicity
	IMLE and diffusion models [6,52]	DFT explicitly models periodic frequencies	Stable training; covers full distribution modes	Slow diffusion sampling; requires multi-year data

3.3. Generation Quality Evaluation and Engineering Applicability Analysis of Scenarios

The evaluation of annual-scale PV scenario generation should not only focus on the similarity between generated curves and historical data but also examine their capability to support power system planning, operational analysis, and risk assessment. For 8760 h long sequences, generated scenarios need to preserve intra-annual periodicity, intraday

variations, seasonal transitions, and long-term distributional characteristics, while also reflecting critical operating conditions such as sustained low-output periods, rapid ramping events, and extreme weather.

From the perspective of statistical characteristics, commonly used metrics include the mean, variance, quantiles, probability distributions, annual energy, autocorrelation characteristics, and typical daily profiles, which can be used to evaluate the ability of a model to reproduce historical patterns. For deep generative models, sample diversity and mode coverage should also be considered so as to avoid neglecting low-probability but high-impact tail scenarios [4–6]. Although sustained cloudy or rainy periods, extremely low irradiance, and strong ramping events occur infrequently, they have significant impacts on energy storage configuration, reserve capacity, and operational security, and should therefore be separately considered in evaluation [27,28].

To directly compare probabilistic-statistical and deep-learning/generative approaches for annual-scale scenario generation, we further summarize their relative advantages, limitations, and the criteria along which they should be evaluated. Probabilistic-statistical methods (e.g., VAR, Copula, and clustering-based approaches; Table 1) offer transparent parametric structures, low data requirements, explicit physical interpretability, and low computational cost, but rely on distributional or linearity assumptions that limit their ability to represent nonlinear, high-dimensional temporal dependence and to generate a sufficiently diverse set of extreme or tail scenarios. Deep generative methods (WGAN/CGAN, diffusion, and IMLE variants; Table 1) can learn complex nonlinear and high-dimensional joint distributions in a data-driven, end-to-end manner and support conditional, controllable generation, but at the cost of larger training-data requirements, higher computational cost, reduced interpretability, and, for adversarial variants, training instability. We therefore recommend that an objective comparison combine (i) distributional-fidelity metrics (mean, variance, quantiles, autocorrelation, and mode/tail coverage), (ii) relational-consistency metrics (preservation of physical constraints, diurnal/seasonal structure, and spatial correlation), (iii) computational-cost indicators (training/inference time and data requirement), and (iv) downstream engineering-utility metrics (impact on power-flow analysis, storage sizing, and reserve estimation).

From the perspective of relational structure, scenario quality also depends on whether the statistical relationships among irradiance, temperature, power output, temporal cycles, and spatial locations are preserved. If generated results disrupt variable coupling, day–night patterns, seasonal variations, or spatial correlations, they may affect subsequent system analysis even when the marginal distributions of individual variables are close to those of real data. Therefore, the evaluation framework should consider the consistency of marginal distributions, joint distributions, temporal dependencies, spatial correlations, and physical constraints [5,29].

From the perspective of engineering applications, existing studies have applied PV uncertainty modeling to power simulation, power flow analysis, capacity planning, and representative scenario extraction, demonstrating its value for system-level analysis [2,3,7]. In future work, a unified and comparable evaluation framework can be further established to compare the effects of different generation methods on node voltages, line flows, energy storage capacity, reserve requirements, and operational risks, thereby enabling annual PV scenario generation to better support planning and operational decision-making under high PV penetration.

4. Recent Methods for Predicting Photovoltaic Power Generation

Day-ahead photovoltaic (PV) power prediction aims to forecast the power output of PV systems for the following 24 h, typically at a resolution of 15 to 60 min. Such predictions

are fundamental for next-day dispatch scheduling, operating reserve planning, energy storage operation, and electricity market bidding. The predictive accuracy is governed by the quality of input features, the model's capacity to characterize nonlinear meteorological–power relationships, its ability to quantify forecast uncertainty, and its robustness under complex weather conditions. The general framework of day-ahead PV power prediction is illustrated in Figure 2.

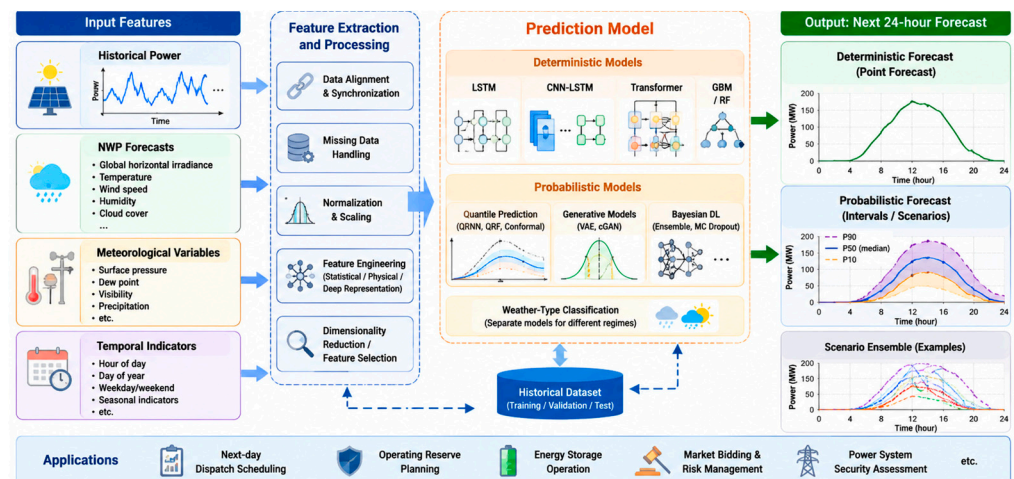


Figure 2. General framework of day-ahead photovoltaic power prediction.

4.1. Day-Ahead Prediction Task and Feature Representation

The core task of day-ahead PV power prediction is to establish a mapping from multi-source input features to the power output curve of the next day. The input data typically encompasses three categories: historical power sequences, meteorological observations, and numerical weather prediction (NWP) products. NWP forecasts of solar irradiance, ambient temperature, wind speed, and humidity constitute the dominant drivers, as solar irradiance alone accounts for the majority of PV output variance [56]. Temporal features such as hour-of-day, day-of-year, and seasonal indicators are also incorporated to encode the periodic characteristics of solar geometry [23].

Early feature engineering relied on physical irradiance decomposition models such as the Perez model to translate NWP outputs into plane-of-array irradiance [57]. With the advent of deep learning, automatic feature extraction from raw time series has become prevalent. Convolutional neural networks (CNNs) extract local temporal patterns from historical power and weather sequences, while long short-term memory (LSTM) networks capture long-range temporal dependencies [58,59]. Attention mechanisms and Transformer architectures further improve the ability to selectively focus on the most informative time steps and input variables [60,61]. Graph neural networks (GNNs) have been applied to model spatial correlations among distributed PV stations, allowing neighboring plants to contribute complementary meteorological information [24]. Feature selection methods, including mutual information criteria and recursive feature elimination, are commonly applied to prune redundant inputs and mitigate overfitting with limited training data [62].

4.2. Uncertainty Modeling in Day-Ahead Prediction

Point forecasts provide only a single expected power trajectory, which may be insufficient for risk-aware decision-making in power system operations. Probabilistic prediction methods produce prediction intervals, full predictive distributions, or scenario ensembles that explicitly represent the uncertainty of the forecast. Three main paradigms have been

developed: parametric distribution estimation, non-parametric quantile regression, and generative scenario sampling.

Parametric approaches model the conditional power output as a member of a prescribed distribution family, such as Beta, Gaussian, or truncated normal, whose parameters are estimated by the prediction model [63]. Non-parametric quantile regression forest (QRF) and quantile regression neural networks (QRNN) directly estimate conditional quantiles at multiple coverage levels without distributional assumptions, enabling the construction of prediction intervals with controlled reliability [64,65]. Conformal prediction frameworks have recently been applied to provide finite-sample coverage guarantees for deep-learning-based PV forecasters [66]. Bayesian deep learning, including Monte Carlo dropout and deep ensembles, offers a principled way to decompose total predictive uncertainty into aleatoric and epistemic components [67]. Generative approaches, including variational autoencoders (VAE) and conditional generative adversarial networks (cGAN), synthesize multiple plausible power trajectories conditioned on NWP inputs, producing scenario ensembles that preserve the temporal autocorrelation structure of solar generation [68,69].

4.3. Robust Prediction Under Complex Weather Conditions

Accurate day-ahead prediction becomes particularly challenging under complex weather conditions such as overcast skies, intermittent cloud cover, rainfall, and extreme meteorological events. These conditions introduce abrupt and irregular variations in solar irradiance that are difficult to anticipate from NWP alone, leading to systematic under-performance of standard regression models. Three main strategies have been developed to improve robustness: weather-type classification, ensemble and hybrid modeling, and physics-informed constraints.

Weather-type classification decomposes the prediction problem into a set of weather-conditioned sub-problems. k-means and fuzzy c-means clustering algorithms group historical days into clear-sky, partially cloudy, overcast, and rainy regimes; a dedicated prediction model is then trained for each regime and selected at inference time based on the forecasted weather type [70]. This divide-and-conquer approach allows each sub-model to specialize in a narrower input distribution, reducing overall prediction error on difficult days. Self-organizing maps (SOMs) and hierarchical clustering have also been applied to derive data-driven weather taxonomies that better reflect the statistical structure of local solar resources [71].

Ensemble methods mitigate the sensitivity of individual models to input perturbations and distribution shifts by aggregating predictions from multiple base learners. Gradient boosting machines (GBMs), random forests, and bagging ensembles have demonstrated consistent improvements over single-model approaches in standardized benchmark studies [72,73]. Hybrid architectures that combine physical clear-sky irradiance models with data-driven correction layers improve generalization to out-of-distribution meteorological regimes, since the physical backbone encodes known solar-geometry constraints that are invariant across locations and seasons [74]. Transfer learning and domain adaptation techniques address the challenge of data scarcity at newly commissioned PV plants, enabling knowledge transfer from data-rich source sites to data-limited target sites [75]. Under missing-data scenarios caused by sensor failures or data transmission faults, imputation methods based on matrix factorization and variational inference restore input completeness before prediction [76]. Finally, adversarial training has been explored to improve robustness against worst-case input perturbations, enhancing the reliability of day-ahead forecasts under extreme weather conditions.

4.4. Summary and Comparison of Day-Ahead Prediction Methods

The framework of day-ahead PV power prediction is illustrated in Figure 2, and representative methods are systematically compared in Table 3.

Table 3. Comparison of representative day-ahead photovoltaic power prediction methods.

Method Category	Representative Methods	Input Features	Output Type	Key Characteristics
Statistical/Physical	Clear-sky model + empirical correction [57]	Solar geometry, NWP irradiance	Point forecast	Physically interpretable; limited for cloudy days
Machine Learning	SVR, Random Forest, XGBoost, LightGBM, GBM [73,74]	Historical power, NWP, and meteorological variables	Point forecast	Fast training; good generalization with limited data
Ensemble/Hybrid	Bagging, Boosting, Physical–ML Hybrid Models [72,74,75]	Historical power, NWP, and meteorological variables	Point/probabilistic forecast	Combines the strengths of multiple learners, or of physical and data-driven models; improves robustness and generalization
Deep Learning (Deterministic)	LSTM, CNN-LSTM, Transformer [58–61]	Historical power, NWP sequences, temporal features	Point forecast	Strong temporal feature extraction; handles long sequences
Probabilistic (Quantile)	QRNN, QRF, Conformal Prediction [64–66]	Historical power, NWP, and temporal features	Prediction intervals/quantiles	Distribution-free; coverage guarantees
Probabilistic (Generative)	VAE, cGAN [68,69]	NWP forecasts, historical patterns	Scenario ensembles	Preserves temporal autocorrelation; multi-scenario output
Bayesian Deep Learning	MC Dropout, Deep Ensemble [67]	Historical power, NWP, and temporal features	Mean + uncertainty decomposition	Separates aleatoric and epistemic uncertainty
Weather-Type Classification	Clustering + dedicated sub-models [70–72]	Weather-type labels, NWP, historical power	Point/probabilistic forecast	Specialization per weather regime; reduced error on difficult days
Hybrid/Physics-Informed	Physical clear-sky + DL residual correction [75]	NWP, solar geometry, historical error	Point forecast	Robust under distribution shift; encodes invariant solar constraints
Transfer Learning	Domain-adaptive neural networks [76]	Source and target site power, NWP	Point forecast	Handles data scarcity at new PV plants

5. Review of Intraday Photovoltaic Power Forecasting Methods

Intraday photovoltaic (PV) power forecasting refers to the technique of rolling estimation of PV plant output power on a minute-to-hour time scale. Unlike day-ahead forecasting, which focuses on the overall power generation trend of the next day, intraday forecasting emphasizes real-time response capability to rapid fluctuations. Its core value lies in supporting grid dispatch departments in formulating rolling generation plans, optimizing reserve capacity allocation, and enhancing the dispatchability of PV power plants. Since meteorological factors such as cloud movement and sudden irradiance changes can cause severe power oscillations within a short time, research methods in this field have evolved from early statistical modeling to complex architectures including deep learning, multimodal fusion, and graph structure learning. This section systematically reviews representative technical routes from a methodological perspective.

5.1. Classical Statistical Learning

Early research primarily focused on mitigating the non-stationary characteristics of power sequences. Since raw PV data is affected by multiple coupling effects of solar irradiance periodic changes, ambient temperature fluctuations, and rapid cloud evolution, direct modeling faces significant challenges. Therefore, researchers typically decompose, smooth, or analyze errors in the sequence before constructing prediction models.

From a technical perspective, research in this direction presents diverse methodologies. Some studies perform post-processing corrections based on the statistical characteristics

of prediction errors, improving the reliability of initial predictions by modeling seasonal biases or systematic offsets. The advantage of such methods lies in achieving accuracy improvement without changing the original model structure. In parallel, a variety of shallow learners—including tree-based ensembles such as Random Forest, XGBoost, and LightGBM—serve as fundamental benchmarks for intraday PV forecasting [77]. Meanwhile, signal decomposition techniques are widely applied to decompose complex power sequences into multiple quasi-stationary components, which are modeled separately and then reconstructed through superposition. Typical methods include frequency-domain decomposition, empirical mode decomposition, and their improved variants [78,79]. Additional work employs improved variants for multiscale decomposition and noise suppression [80]. Furthermore, some studies fit decomposed features or screened meteorological variables using shallow regression models such as weighted Gaussian process regression and optimized neural networks [81,82].

Overall, these methods offer advantages, including a clear model structure, low computational cost, and a low implementation threshold, making them valuable in scenarios with limited data or computational resources. However, their capabilities in capturing nonlinear interactions, responding to extreme weather, and modeling multiscale temporal dependencies are often limited, making it difficult to meet the demands of high-precision intraday forecasting.

5.2. Neural Network-Driven Prediction Frameworks

With the growth of data scale and improvement in computational capabilities, neural network-based prediction frameworks have gradually become mainstream. Compared with traditional statistical models, deep learning methods can better characterize the complex nonlinear mapping between power generation and meteorological factors, demonstrating stronger adaptability, especially under rapid weather transition scenarios.

Early architectures mainly employed Recurrent Neural Networks (RNNs) and their variants to capture temporal dependencies, combined with convolution operations to extract local fluctuation patterns [83,84]. Additional work integrated convolution operations with attention mechanisms to extract local fluctuation patterns [85]. The core advantage of such methods lies in fully utilizing recent observation information to achieve rapid rolling updates of power changes over the next tens of minutes to hours.

Subsequent improvements enhanced model perception sensitivity to critical time nodes, sensitive variables, and power sudden change points through mechanisms such as feature engineering optimization, sequence decomposition preprocessing, and hierarchical pooling [86,87]. Transfer learning strategies address practical issues such as insufficient samples, data distribution drift caused by seasonal changes, and inter-station characteristic differences by introducing pretraining-finetuning paradigms combined with high-dimensional feature screening to enhance cross-scenario adaptability [88].

5.3. Self-Attention Mechanisms and New Paradigms for Sequence Modeling

The introduction of attention mechanisms has brought new breakthroughs to intraday forecasting. The core idea of such methods is to enable the network to automatically focus on high-contribution input features while weakening interference from non-critical time steps, thereby improving perception capability for power sudden change points.

In terms of attention architectures, some studies integrated Bidirectional Long Short-Term Memory (BiLSTM) with attention mechanisms, enabling the model to dynamically adjust feature weights, with significantly improved ultra-short-term prediction accuracy after deployment [89]. Additional work designed multi-level attention-weighting strate-

gies combined with sequence-decomposition preprocessing to further enhance detection capability for local abnormal fluctuations [90].

Transformer-based sequence modeling methods more effectively capture long-range temporal dependencies through self-attention mechanisms, achieving unified modeling of intraday periodic patterns and local abrupt fluctuation behaviors. Representative models include Informer, Autoformer, TimesNet, FCformer, etc. [91,92]. Some work introduced temporal decomposition networks to extract periodic and trend components through autocorrelation mechanisms, combining maximum likelihood estimation with gradient optimization to output probability distribution parameters. Other studies enhanced model stability through multi-head attention, layer normalization, and residual connections, introducing polynomial fitting and error correction in the post-processing stage [93,94]. Overall, the introduction of attention mechanisms and Transformer architectures enables models to more flexibly address the challenge of coexisting rapid variability and periodicity in intraday power sequences.

5.4. Spatiotemporal Collaborative Modeling for Distributed Power Stations

Power fluctuations in distributed PV plants within a region exhibit significant spatiotemporal correlation characteristics. When clouds move across a region, power stations at different geographical locations sequentially experience irradiance changes. This cloud propagation process manifests as a spatially correlated pattern with a time lag in power output. If only single-station historical data is used for prediction, early warning information from neighboring stations will be lost, making it difficult to fully capture power dynamic evolution patterns under rapidly changing conditions. Therefore, mining and utilizing spatiotemporal coupling relationships among power station clusters has become one of the key approaches to improving intraday prediction performance.

As shown in Figure 3, Graph Neural Network (GNN)-based methods incorporate regional spatial correlations into the model structure by explicitly constructing topological connections between power stations [95,96]. In this modeling paradigm, each PV power station is abstracted as a graph node, edges between nodes represent spatial correlation strength, and graph convolution operations are used to aggregate neighborhood information and model spatial dependencies. By combining graph convolution with temporal networks, the model can simultaneously capture spatial propagation characteristics and temporal evolution patterns of power changes.

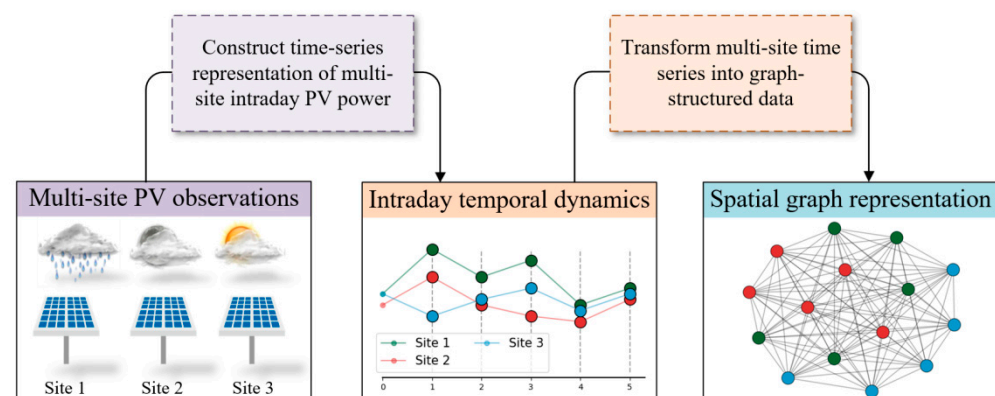


Figure 3. Spatiotemporal correlation and graph learning in intraday forecasting. Green, red, and blue curves and nodes correspond to Sites 1, 2, and 3, respectively, while the edges in the spatial graph schematically represent inter-site spatial correlations.

In terms of graph structure construction, some studies define inter-station connections based on geographical distance, meteorological similarity, or power correlation, thereby refining the modeling of local collaborative changes [97]. Additional work employs physical-similarity metrics (such as cosine similarity of irradiance attenuation rates) to construct adjacency matrices, ensuring that stations experiencing similar cloud-induced fluctuations maintain strong connections regardless of their physical distance. Beyond static graph structures, dynamic graph modeling methods enable information transfer weights between nodes to be adjusted in real time according to weather conditions and power changes [98]. The core idea is that the spatial correlation strength between power stations under different meteorological conditions is not fixed but evolves dynamically with factors such as cloud movement and irradiance distribution. By constructing time-varying graph structures, models can more accurately characterize transient characteristics of fluctuation propagation. To address common issues in distributed scenarios such as insufficient data samples and strong heterogeneity in data distribution across stations, strategies including transfer learning, domain adaptation, and decentralized training are employed to enhance the cross-station generalization capability of models.

5.5. Perception Enhancement Strategies for Multi-Source Data Fusion

Single-time-series modeling has inherent limitations when dealing with scenarios such as rapid cloud movement, local occlusion, and sudden weather changes. Therefore, research paradigms are gradually evolving toward multimodal information fusion, integrating data types including historical output records, ground meteorological observations, Numerical Weather Prediction (NWP) products, ground-based cloud images, satellite remote sensing imagery, and regional cloud system characteristics [99,100].

The core value of multi-source fusion lies in utilizing complementary characteristics among different data sources to construct a richer fluctuation representation system. Existing work can be roughly categorized into two technical routes: The first is “regional meteorological information + multidimensional time series” fusion, focusing on improving the advanced prediction of power evolution trends over the next tens of minutes to hours [101]. The second is “visual image + temporal data” fusion, which jointly models sky images with power and meteorological sequences to strengthen local cloud occlusion dynamic recognition.

In the field of cloud image fusion, research mainly unfolds from two dimensions: cloud feature extraction and multimodal coupling. Fundamental work focuses on the effective representation of cloud features and fluctuation pattern recognition, establishing physical mapping relationships among cloud-irradiance-power by extracting key cloud image features and correlating them with power temporal data. Some studies designed cloud image prediction modules with multiscale spatiotemporal feature enhancement, mining spatiotemporal evolution patterns from historical cloud image sequences to generate future cloud image predictions, with a significant reduction in error indicators under clear and cloudy conditions [102,103]. Additional work further integrated the Vision Transformer (ViT) with Long Short-Term Memory (LSTM) networks, employing adaptive threshold algorithms for cloud region segmentation, combining similarity weighting with optical flow methods to generate pixel-level cloud motion vector fields, significantly reducing Root Mean Square Error (RMSE) compared with persistence benchmarks [104]. In summary, the core concepts, representative technologies, and advantages and disadvantages of existing photovoltaic power forecasting methods are summarized in Table 4.

Table 4. Comparison of intraday PV power forecasting methods.

Category	Method	Core Idea	Representative Techniques	Advantages	Limitations
Statistical Learning	Hybrid and Ensemble	Combine the strengths of multiple algorithms	Random Forest, XGBoost, LightGBM [77]	Improved accuracy and robustness via multi-model synergy	High structural complexity and tuning difficulty
	Signal Decomposition	Decompose sequences into stationary components	Frequency-domain decomposition, EMD [78,79]	Reduces sequence non-stationarity	Error accumulation
	Shallow Regression	Fit statistical relationships	Weighted GPR, Neural Networks [81,82]	Low computational cost	Weak nonlinear modeling capability
Deep Learning	RNN	Capture temporal dependencies	LSTM, GRU, BiLSTM [83,84]	Strong sequence modeling capability	Limited long-term dependency modeling
	CNN	Extract local fluctuation patterns	CNN, Hybrid models [85]	Effective for local features	Weak global modeling capability
	Attention Weighting	Focus on high-contribution input features	BiLSTM + Attention [89,90]	Dynamic feature weight adjustment	Increased computational overhead
	Transformer	Model global temporal dependencies	Informer, Autoformer, TimesNet [91–93]	Strong multi-scale modeling capability	High training complexity
Spatiotemporal Collaborative	GNN	Model spatial correlations among power stations	GNN, Graph Convolution [89,96,97]	Effectively utilizes spatial information	Complex graph structure construction
	Dynamic Graph Learning	Adjust connection weights in real time according to weather conditions	Time-varying graph structures [98]	Adapts to dynamic meteorological conditions	High model complexity
Multimodal Fusion	Regional Weather + Time Series	Fuse NWP with multidimensional time series	NWP + Feature fusion [99–101]	Enhanced advanced estimation capability	High structural complexity
	Cloud Image + Time Series	Fuse sky images with power data	Sky images + ViT + LSTM [102–104]	Better cloud perception capability	Image quality fluctuations impact performance

Across the three timescales reviewed above, spatiotemporal correlation among distributed PV stations is captured through different strategies, with different consequences for forecasting accuracy. At the annual scale, spatial dependence is mainly represented through static covariance/Copula structures or data-driven multi-channel and tensor representations, which reproduce long-term joint distributions well but do not track time-varying regional meteorological fields. At the day-ahead scale, graph neural networks incorporate neighboring-plant meteorological information to complement single-site NWP, which is particularly beneficial for stations with incomplete or noisy local measurements. At the intraday scale, dynamic graph learning and regional cloud-image fusion adjust spatial connection weights in real time to reflect rapidly moving cloud fields; this is also the timescale at which regional meteorological coupling has the strongest impact on accuracy, because short-term irradiance ramps are often driven by spatially propagating cloud systems rather than by a single station's own history. Overall, the benefit of explicit spatial/regional modeling tends to increase as the timescale shortens, while the associated computational and graph-construction complexity (Table 4) increases correspondingly, indicating an accuracy–complexity trade-off that should guide the choice of spatial modeling strategy at each timescale.

6. Future Research Directions

Future research on PV scenario generation and power forecasting will no longer be limited to improving the accuracy of individual models. Instead, greater attention will be paid to whether model outputs can provide interpretable, reusable, and transferable uncertainty information. For power systems with high PV penetration, annual scenarios, day-ahead forecasting, and intraday forecasting correspond to long-term planning, operational scheduling, and short-term regulation, respectively. However, these tasks share a common objective: to accurately describe the uncertainty of PV output caused by variations in meteorological conditions, timescales, and spatial locations. Therefore, future studies may gradually move from isolated task-oriented modeling toward a unified uncertainty representation framework, so that generated and forecasted results at different timescales can remain consistent in terms of statistical meanings, physical constraints, and engineering applications.

6.1. Multi-Timescale Collaborative Modeling

This fragmentation limits the ability of operators to reconcile outputs across timescales—for example, a day-ahead trend and an intraday correction may be statistically inconsistent even when each is individually accurate, which undermines their joint use in dispatch decision-making.

Moving from the theoretical models reviewed above to industrial deployment raises several practical challenges that are, in our assessment, at least as important as further gains in accuracy. First, multi-timescale collaborative modeling currently exists mainly as separate model families with different objectives, input pipelines, and update frequencies; a deployable pipeline instead requires shared data infrastructure and a mechanism for propagating and reconciling information (e.g., using updated day-ahead trends to constrain intraday corrections) across timescales without producing conflicting outputs for the operator. Second, integrating heterogeneous data sources (NWP, ground meteorological stations, sky and satellite imagery, and SCADA/equipment states) in real time requires handling differences in sampling rate, spatial resolution, latency, and reliability, together with robust treatment of missing or delayed sources, which is harder to guarantee in a live operational environment than in an offline benchmark. Third, computational efficiency and scalability are critical for large-scale, high-PV-penetration systems: deep generative and graph-based models that perform well on the limited station sets used in the literature (Tables 2 and 4) may not scale to province- or nation-level networks with thousands of distributed plants under the latency constraints of intraday dispatch, motivating further research on lightweight or distilled models, incremental or federated training across sites, and edge-cloud collaborative inference.

6.2. Multi-Source Heterogeneous Information Fusion

At the data level, PV modeling will increasingly rely on the collaborative use of multi-source information. Historical power series alone are often insufficient to explain PV output variations under complex weather conditions. Future studies should further integrate numerical weather prediction, ground-based meteorological observations, sky images, satellite cloud images, regional PV plant data, and equipment operating states. Notably, the key to multi-source fusion is not simply to increase the number of input variables, but to address the differences among data sources in temporal resolution, spatial scale, physical meaning, and reliability. This will enable models to identify which information is truly effective for long-term distribution modeling, day-ahead trend estimation, and intraday ramp perception.

6.3. Controllable Generative Modeling

Addressing these limitations requires generative models to incorporate domain knowledge as explicit constraints rather than relying solely on data-driven fitting.

At the model level, generative AI and time-series foundation models will provide new opportunities for PV uncertainty modeling. Compared with traditional forecasting models, generative models can produce multiple possible scenarios, making them more suitable for describing the uncertainty ranges and risk boundaries of PV output. Future models should further support conditional controllability, for example, generating specific scenarios according to season, weather type, plant location, installed capacity, or operational requirements. Meanwhile, model design should more fully incorporate knowledge such as irradiance patterns, day–night boundaries, power limits, variable coupling, and spatial correlations, so that generated and forecasted results are not only reasonable in terms of data distribution, but also consistent with the basic mechanisms of PV power generation.

6.4. Extreme Scenario Characterization

In terms of risk characterization, low-probability and high-impact events will remain a major focus of future research. Extremely low irradiance, sustained cloudy or rainy periods, strong cloud shading, and rapid ramping events account for only a small proportion of historical samples, but they have significant impacts on energy storage capacity, reserve requirements, voltage security, and operational risks. Future studies can combine extreme value statistics, risk-aware learning, and conditional generation methods to improve the ability of models to identify, generate, and forecast tail scenarios. Meanwhile, the evaluation of extreme scenarios should be extended from distributional similarity to system-level impacts, with attention to whether such scenarios can reflect the security margins and regulation pressure in real operation.

Regarding model robustness under abrupt weather changes, the strategies reviewed in Section 4.3—weather-type-conditioned sub-models, ensemble/hybrid architectures combining a physical clear-sky backbone with data-driven correction, transfer learning for data-scarce sites, and adversarial training against worst-case perturbations—provide the current basis for robustness, and extending them with online/rolling model updating and out-of-distribution detection is a promising direction for maintaining reliability when a weather regime shifts abruptly within the forecast horizon. Regarding the operational use of generated extreme scenarios, we suggest that they be integrated into dispatch and emergency-protection design through three complementary routes: (i) as boundary/stress scenarios in reserve-capacity and energy-storage sizing studies, so that reserve and storage margins are dimensioned against low-probability, high-impact PV deficits rather than only against average variability; (ii) as inputs to probabilistic power-flow and extreme-scenario security checks, to identify voltage- and line-loading risks before they occur in real operation; and (iii) as triggers for pre-defined emergency-response and protection-setting adjustments (e.g., reserve-activation thresholds and curtailment rules) that are pre-computed offline from the extreme-scenario set and invoked once real-time monitoring or intraday forecasts indicate that operating conditions are approaching a generated extreme scenario.

6.5. Engineering Closed-Loop Validation

At the application level, PV scenario generation and forecasting results need to be further embedded into power system analysis workflows. Existing studies have applied PV uncertainty modeling to power simulation, power flow analysis, capacity planning, and representative scenario extraction, demonstrating the engineering value of these methods. The integration of rural energy Internet technologies and AI-driven agri-energy

management further broadens the application context for PV uncertainty modeling in agriculture-coupled power systems [105,106]. More importantly, future research should establish unified and reproducible application-oriented evaluation procedures so that different models can be compared across the same datasets, system models, and metrics. By linking scenario generation, power forecasting, system simulation, and operational feedback, PV uncertainty modeling can move beyond algorithm-level performance evaluation toward system-level applications for planning, dispatching, and risk-informed decision-making.

To make the evaluation of both probabilistic forecasts and generated scenarios more directly aligned with the requirements of the system operator, we recommend a validation procedure structured into three levels. First, algorithm-level metrics (e.g., pinball/quantile loss, CRPS, and prediction-interval coverage probability for forecasts; distributional and mode-coverage metrics for generated scenarios, cf. Section 3.3) quantify statistical accuracy. Second, relational-consistency checks verify that physical constraints, diurnal/seasonal structure, and spatial correlations are preserved (Sections 3.2 and 3.3), since operators cannot act on results that violate known physical relationships even when they are statistically accurate on average. Third, closed-loop engineering verification feeds the forecasts or scenarios into the actual downstream application—power-flow calculation, storage/reserve sizing, or dispatch simulation—and evaluates the resulting operational indicators (voltage/line-loading violations, reserve-shortfall probability, and curtailment or unserved-energy risk) rather than the forecast error alone. In our view, comparing these operational indicators across candidate models is the criterion that best matches an operator’s actual decision needs, and we highlight the establishment of a unified, reproducible benchmark of this kind as a priority for future work.

7. Conclusions

PV power output is jointly affected by meteorological conditions, temporal cycles, spatial distributions, and PV plant operating states, exhibiting significant uncertainty, variability, and multi-timescale characteristics. With the increasing penetration of PV generation in power systems, scenario generation and power forecasting have become important technical foundations for planning analysis, dispatch operation, and risk assessment.

From the perspective of statistical relational AI, this paper reviews annual scenario generation, day-ahead forecasting, and intraday forecasting methods for PV power generation. Annual scenario generation focuses on long-term distributions, periodic patterns, variable relationships, and the characterization of extreme samples, providing basic scenarios for planning and risk assessment. Day-ahead forecasting emphasizes output trends and uncertainty representation driven by historical power data, meteorological variables, and numerical weather prediction, thereby supporting next-day dispatch and market decision-making. Intraday forecasting focuses on short-term fluctuation perception, multi-source information fusion, and regional collaborative modeling, which are important for improving real-time operational support. Although these three tasks correspond to different timescales, they all center on the statistical relationships among PV output, meteorological factors, temporal structures, spatial correlations, and system requirements.

From the perspective of methodological development, existing studies have gradually expanded from traditional probabilistic statistical models to machine learning, deep learning, and generative AI methods. Traditional methods provide relatively clear statistical interpretability; machine learning enhances nonlinear relationship modeling, and deep learning and generative AI further improve the capabilities for complex temporal feature extraction, multivariable relationship learning, and scenario construction. Meanwhile, the introduction of knowledge constraints, physical mechanisms, frequency-domain features, graph-structured

learning, and extreme risk modeling has promoted the transition of PV uncertainty modeling from data fitting toward relationship preservation and engineering application support.

Overall, substantial progress has been made in scenario generation, power forecasting, and system-level applications. However, coordination across different timescale tasks remains insufficient; the characterization of complex weather conditions and extreme scenarios is still challenging; and multi-source heterogeneous data fusion, cross-regional generalization, and unified engineering-oriented evaluation frameworks require further improvement. Statistical relational AI provides a useful framework for the unified understanding of PV scenario generation and power forecasting by integrating statistical uncertainty, variable relational structures, AI models, and power system application requirements. This review can provide a reference for scenario generation, power forecasting, and system uncertainty analysis under high PV penetration.

Author Contributions: Conceptualization, methodology, and writing—original draft preparation, Y.C. and X.F.; formal analysis and writing—original draft preparation, Y.C., X.L., C.Z., and X.Z.; writing—review and editing, X.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

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

Abbreviations

The following abbreviations are used in this manuscript:

PV	Photovoltaic
VAR	Vector Autoregression
GANs	Generative Adversarial Networks
IMLE	Implicit Maximum Likelihood Estimation
SWGC	Shared Weight Graph Clustering
CNNs	Convolutional Neural Networks
CGANs	Conditional GANs
3D-CNNs	Three-Dimensional Convolutional Neural Networks
WGAN	Wasserstein Generative Adversarial Network
LSTM	Long Short-Term Memory
MPPT	Maximum Power Point Tracking
MWGAN	Multi-Wasserstein Generative Adversarial Network
AI	Artificial Intelligence
NWP	Numerical Weather Prediction
RNN	Recurrent Neural Network
RMSE	Root Mean Square Error
BiLSTM	Bidirectional Long Short-Term Memory
GNN	Graph Neural Network
ViT	Vision Transformer
QRF	Quantile Regression Forest
QRNNs	Quantile Regression Neural Networks
VAE	Variational Autoencoder
SOMs	Self-Organizing Maps
GBM	Gradient Boosting Machine

References

1. Fu, X. *Statistical Relational Artificial Intelligence in Photovoltaic Power Uncertainty Analysis*; Elsevier: Amsterdam, The Netherlands, 2025.
2. Fu, X.; Zhang, C.; Xu, Y.; Zhang, Y.; Sun, H. Statistical machine learning for power flow analysis considering the influence of weather factors on photovoltaic power generation. *IEEE Trans. Neural Netw. Learn. Syst.* **2025**, *36*, 5348–5362. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Fu, X. Statistical machine learning model for capacitor planning considering uncertainties in photovoltaic power. *Prot. Control Mod. Power Syst.* **2022**, *7*, 51–63. [\[CrossRef\]](#)
4. Fu, X.; Zhang, C.; Zhang, X.; Sun, H. A novel GAN architecture reconstructed using Bi-LSTM and style transfer for PV temporal dynamics simulation. *IEEE Trans. Sustain. Energy* **2024**, *15*, 2826–2829. [\[CrossRef\]](#)
5. Fu, X.; Chang, F.; Sun, H.; Zhang, P.; Zhang, Y. Knowledge-integrated GAN model for stochastic time-series simulation of year-round weather for photovoltaic integration analysis. *IEEE Trans. Power Syst.* **2025**, *40*, 5289–5301. [\[CrossRef\]](#)
6. Fu, X.; Chang, F.; Li, Z.; Sun, H.; Zhang, Y.; Yang, D. Redesigning the decoder and loss function of diffusion Transformer for PV temporal simulation. *IEEE Trans. Smart Grid* **2026**, *17*, 1629–1638. [\[CrossRef\]](#)
7. Fu, X.; Lu, N.; Sun, H.; Zhang, Y. A novel clustering method for extracting representative photovoltaic scenarios considering power, energy, and variability. *IEEE Trans. Power Syst.* **2025**, *40*, 4353–4366. [\[CrossRef\]](#)
8. Yang, F.; Fu, X.; Li, Z.; Qiu, D.; Badihi, H. A novel clustering-ensemble learning model for day-ahead photovoltaic power forecasting. *Eng. Appl. Artif. Intell.* **2025**, *161*, 112281. [\[CrossRef\]](#)
9. Zhang, C.; Fu, X.; Li, Z.; Yu, N.; Zhang, Y.; Gu, H. Unified Fourier graph-based spatiotemporal learning and corrected NWP for multi-site ultra-short term photovoltaic power forecasting. *IEEE Trans. Smart Grid* **2026**, *17*, 1639–1652. [\[CrossRef\]](#)
10. Liu, J.; Zang, H.; Ding, T.; Cheng, L.; Wei, Z.; Sun, G. Sky-image-derived deep decomposition for ultra-short-term photovoltaic power forecasting. *IEEE Trans. Sustain. Energy* **2024**, *15*, 871–883. [\[CrossRef\]](#)
11. Zhou, X.; Pang, C.; Zeng, X.; Jiang, L.; Chen, Y. A short-term power prediction method based on temporal convolutional network in virtual power plant photovoltaic system. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 9003810. [\[CrossRef\]](#)
12. Ekström, J.; Koivisto, M.J.; Mellin, I.; Millar, R.J.; Lehtonen, M. A statistical model for hourly large-scale wind and photovoltaic generation in new locations. *IEEE Trans. Sustain. Energy* **2017**, *8*, 1383–1393. [\[CrossRef\]](#)
13. Nuño, E.; Koivisto, M.; Cutululis, N.A.; Sørensen, P. On the simulation of aggregated solar PV forecast errors. *IEEE Trans. Sustain. Energy* **2018**, *9*, 1889–1898. [\[CrossRef\]](#)
14. Yang, X.; He, H.; Li, J.; Zhang, Y. Toward optimal risk-averse configuration for HESS with CGANs-based PV scenario generation. *IEEE Trans. Syst. Man Cybern. Syst.* **2021**, *51*, 1779–1793. [\[CrossRef\]](#)
15. Fu, X.; Niu, H. Key technologies and applications of agricultural energy Internet for agricultural planting and fisheries industry. *Inf. Process. Agric.* **2023**, *10*, 416–437. [\[CrossRef\]](#)
16. Fu, X.; Zhang, C.; Chang, F.; Han, L.; Zhao, X.; Wang, Z.; Ma, Q. Simulation and forecasting of fishery weather based on statistical machine learning. *Inf. Process. Agric.* **2024**, *11*, 127–142. [\[CrossRef\]](#)
17. Kousounadis-Knousen, M.A.; Bazionis, I.K.; Georgilaki, A.P.; Catthoor, F.; Georgilakis, P.S. A review of solar power scenario generation methods with focus on weather classifications, temporal horizons, and deep generative models. *Energies* **2023**, *16*, 5600. [\[CrossRef\]](#)
18. Al-Dahidi, S.; Madhiarasan, M.; Al-Ghussain, L.; Abubaker, A.M.; Ahmad, A.D.; Alrbai, M.; Aghaei, M.; Alahmer, H.; Alahmer, A.; Baraldi, P.; et al. Forecasting solar photovoltaic power production: A comprehensive review and innovative data-driven modeling framework. *Energies* **2024**, *17*, 4145. [\[CrossRef\]](#)
19. Yu, J.; Li, X.; Yang, L.; Li, L.; Huang, Z.; Shen, K.; Yang, X.; Yang, X.; Xu, Z.; Zhang, D.; et al. Deep learning models for PV power forecasting: Review. *Energies* **2024**, *17*, 3973. [\[CrossRef\]](#)
20. Mauladdawilah, H.; Gago, E.J.; Balfaqih, H.; Pegalajar, M.C. Towards energy efficiency: A comprehensive review of deep learning-based photovoltaic power forecasting strategies. *Heliyon* **2024**, *10*, e33419. [\[CrossRef\]](#)
21. Li, B.; Zhang, J. A review on the integration of probabilistic solar forecasting in power systems. *Sol. Energy* **2020**, *207*, 777–795.
22. Panamtash, H.; Mahdavi, S.; Zhou, Q. Probabilistic solar power forecasting: A review and comparison. In Proceedings of the 2020 52nd North American Power Symposium (NAPS), Tempe, AZ, USA, 11–13 April 2021; pp. 1–6.
23. 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. [\[CrossRef\]](#)
24. Wang, H.; Liu, Y.; Zhou, B. Taxonomy research of artificial intelligence for deterministic solar power forecasting. *Energy Convers. Manag.* **2020**, *214*, 112909. [\[CrossRef\]](#)
25. Lin, N.; Palensky, P.; Vergara, P.P. EnergyDiff: Universal time-series energy data generation using diffusion models. *IEEE Trans. Smart Grid* **2025**, *16*, 4252–4265. [\[CrossRef\]](#)
26. Kang, M.; Zhu, R.; Chen, D.; Li, C.; Gu, W.; Qian, X.; Yu, W. A cross-modal generative adversarial network for scenarios generation of renewable energy. *IEEE Trans. Power Syst.* **2024**, *39*, 2630–2640. [\[CrossRef\]](#)

27. Mo, T.; Liu, J.; Zhao, Y.; Liu, J.; Ding, T. An interpretable photovoltaic scenario generation method based on RMT-enhanced WGAN considering physical constraints. *IEEE Trans. Sustain. Energy* **2025**, early access. [\[CrossRef\]](#)
28. Wang, H.; Qin, B.; Hong, S.; Xu, X.; Su, Y.; Lu, T.; Ding, T. Enhanced GAN-based joint wind-solar-load scenario generation with extreme weather labelling. *IEEE Trans. Smart Grid* **2025**, *16*, 4213–4224. [\[CrossRef\]](#)
29. Dong, X.; Liu, Y.; Zhang, X.; Mei, S. A data-driven dynamic temporal correlation modeling framework for renewable energy scenario generation. *IEEE Trans. Sustain. Energy* **2026**, *17*, 1048–1058. [\[CrossRef\]](#)
30. Zhang, Y.; Beaudin, M.; Taheri, R.; Zareipour, H.; Wood, D. Day-ahead power output forecasting for small-scale solar photovoltaic electricity generators. *IEEE Trans. Smart Grid* **2015**, *6*, 2253–2262. [\[CrossRef\]](#)
31. Leva, S.; Mussetta, M.; Ogliari, E. PV module fault diagnosis based on microconverters and day-ahead forecast. *IEEE Trans. Ind. Electron.* **2019**, *66*, 3928–3937. [\[CrossRef\]](#)
32. Zhang, C.; Fu, X.; Qiu, D.; Badihi, H.; Zhang, P.; Zhang, Y.; Gu, H. Multi-scale patch and frequency-domain gated learning for high-resolution day-ahead photovoltaic forecasting. *Appl. Energy* **2026**, *402*, 126973. [\[CrossRef\]](#)
33. Zhang, C.; Fu, X.; Yang, D.; Zhang, P.; Zhang, Y. GEV distribution-enhanced Fourier diffusion model for extreme value capture in day-ahead photovoltaic scenario generation. *Appl. Energy* **2026**, *409*, 127472. [\[CrossRef\]](#)
34. Rastgoo, R.; Amjady, N.; Lin, S.; Mueen, S.M. Ultra-short-term solar power prediction using sky image sequences by a residual vision reformer. *IEEE Trans. Sustain. Energy* **2025**, *16*, 2972–2988. [\[CrossRef\]](#)
35. Su, Z.; Shen, J.; Zhou, Q.; Yong, B. SL-ANFIS-LSTM: A structure learnable fuzzy neural network for ultra-short-term PV power forecasting. *IEEE Trans. Fuzzy Syst.* **2025**, *33*, 4252–4262. [\[CrossRef\]](#)
36. Ma, Q.; Fu, X.; Yang, Q.; Qiu, D. Adaptive masked network for ultra-short-term photovoltaic forecast. *Eng. Appl. Artif. Intell.* **2025**, *139*, 109555. [\[CrossRef\]](#)
37. Lv, R.J.; Zhang, M.Y.; Ye, C.J. Scenario Generation of Annual PV and Load Correlated Time Series Considering Meteorological Influences. In Proceedings of the 8th Conference on Energy Internet and Energy System Integration (EI2), Shenyang, China, 29 November–2 December 2024.
38. Du, C.; Gao, Y.; Wang, L.; Li, X.; Cui, Y.; Gao, J. Medium- and Long-Term Optimal Stochastic Scheduling for Inter-Basin Hydro-Wind-Photovoltaic Complementary Systems Considering Wind and Solar Output Uncertainty. *IEEE Open Access J. Power Energy* **2025**, *12*, 404–416. [\[CrossRef\]](#)
39. Fan, F.; Zhou, X.; Zhang, L.; He, Z.; Xu, W.; Zhao, Y. Modeling Multi-Temporal Correlation of PV Stations Output via Improved Nonparametric Disaggregation Method in Power System Reliability Evaluation. *IEEE Access* **2024**, *12*, 109071–109079. [\[CrossRef\]](#)
40. Wang, R.; Wu, Z.P.; Liu, J.F.; Yi, J.; Lyu, Q. Multi-Time Scale Extreme Cold Wave Weather Scenario Generation Method Considering Source-Load Coupling Characteristics. In Proceedings of the 6th International Conference on Smart Grid and Energy Engineering (SGEE), Shanghai, China, 28–30 November 2025.
41. Fu, X.; Lu, N.; Sun, H.; Zhang, Y. Extraction of Representative Scenarios for Photovoltaic Power with Shared Weight Graph Clustering. *IEEE Trans. Smart Grid* **2024**, *15*, 6158–6170. [\[CrossRef\]](#)
42. Xie, H.X.; Gao, S.; Lu, J.Y.; Zhao, X.; Xi, X.; Yao, H.C. Day-Ahead Temporal Scenario Generation and Quality Evaluation of Photovoltaic and Loads Based on Volatility Interval Correction. In Proceedings of the 16th Asia-Pacific Power and Energy Engineering Conference (APPEEC), Nanjing, China, 25–27 October 2024.
43. Si, D.J.; Zhao, Y.Y.; Li, L.F.; Chen, Y.X.; Sun, P. Generation Method of Probabilistic Annual Output Scenario of New Energy Based on Electricity Characteristic Matching. In Proceedings of the 5th Asia Energy and Electrical Engineering Symposium (AEEES), Chengdu, China, 23–26 March 2023.
44. Xiao, X.; Shi, M.; Wu, F.; Wei, Y.; Zhao, B.; Zhang, F. Optimal Configuration of ESS in Distribution Network Considering Generation of PV-Load Time Series Scenarios. *IEEE Access* **2025**, *13*, 125397–125405. [\[CrossRef\]](#)
45. Chen, Y.; Wang, Y.; Kirschen, D.; Zhang, B. Model-Free Renewable Scenario Generation Using Generative Adversarial Networks. *IEEE Trans. Power Syst.* **2018**, *33*, 3265–3275. [\[CrossRef\]](#)
46. Ji, Q.; Pu, T.J.; Wang, X.Y. Renewable Scenario Generation Using Controllable Generative Adversarial Networks with Transparent Latent Space. *CSEE J. Power Energy Syst.* **2021**, *7*, 66–76.
47. Song, Z.Y.; Huang, Y.H.; Li, X.F.; Li, C. Constructing Method of Multi-regional Photovoltaic Power Output Scenarios Based on Conditional Generative Adversarial Network. In Proceedings of the 6th IEEE Conference on Energy Internet and Energy System Integration (EI2), Chengdu, China, 11–13 November 2022.
48. Song, Z.; Huang, Y.; Xie, H.; Li, X. Generation Method of Multi-Regional Photovoltaic Output Scenarios-Set Using Conditional Generative Adversarial Networks. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2023**, *13*, 861–870. [\[CrossRef\]](#)
49. Yuan, R.; Wang, B.; Sun, Y.; Song, X.; Watada, J. Conditional Style-Based Generative Adversarial Networks for Renewable Scenario Generation. *IEEE Trans. Power Syst.* **2023**, *38*, 1281–1296. [\[CrossRef\]](#)
50. Zhang, Z.K.; Zhu, R.J. A Photovoltaic Power Scenario Generation Method Based on HiLo-WGAN. In Proceedings of the 11th International Forum on Electrical Engineering and Automation (IFEAA), Shenzhen, China, 22–24 November 2024.

51. Wang, Y.; Liu, Y.; Yang, Q. Operational Scenario Generation and Forecasting for Integrated Energy Systems. *IEEE Trans. Ind. Inform.* **2024**, *20*, 2920–2931. [\[CrossRef\]](#)
52. Liao, W.; Bak-Jensen, B.; Radhakrishna Pillai, J.; Yang, Z.; Wang, Y.; Liu, K. Scenario Generations for Renewable Energy Sources and Loads Based on Implicit Maximum Likelihood Estimations. *J. Mod. Power Syst. Clean Energy* **2022**, *10*, 1563–1575. [\[CrossRef\]](#)
53. Pan, Y.F.; Zhang, S.L.; Yu, Y.; Li, C.G. Annual Power Flow Scenario Generation for Hybrid AC/DC Receiving-End Provincial Grids Based on Production Simulation. In Proceedings of the 2nd International Conference on Electrical Power Systems and Intelligent Control (EPSIC), Qingdao, China, 15–17 August 2025.
54. Quiroga, G.A.; Kagan, H.; Amasifen, J.C.C.; Almeida, C.F.M.; Vicentini, E. Evaluation of Distributed Generation Impacts on Distribution Networks under Different Penetration Scenarios. In Proceedings of the 2015 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT LATAM), Montevideo, Uruguay, 5–7 October 2015.
55. Mahmud, M.A.; Hossain, M.J.; Pota, H.R. Voltage Variation on Distribution Networks with Distributed Generation: Worst Case Scenario. *IEEE Syst. J.* **2014**, *8*, 1096–1103. [\[CrossRef\]](#)
56. Ineichen, P.; Perez, R. A new air mass independent formulation for the Linke turbidity coefficient. *Sol. Energy* **2002**, *73*, 151–157. [\[CrossRef\]](#)
57. Perez, R.; Ineichen, P.; Seals, R.; Michalsky, J.; Stewart, R. Modeling daylight availability and irradiance components from direct and global irradiance. *Sol. Energy* **1990**, *44*, 271–289. [\[CrossRef\]](#)
58. Gao, M.; Li, J.; Hong, F.; Long, D. Day-ahead power forecasting in a large-scale photovoltaic plant based on weather classification using LSTM. *Energy* **2019**, *187*, 115838. [\[CrossRef\]](#)
59. Li, G.; Wang, H.; Zhang, S.; Xin, J.; Liu, H. Recurrent neural networks based photovoltaic power forecasting approach. *Energies* **2019**, *12*, 2538. [\[CrossRef\]](#)
60. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017; Volume 30.
61. Zhou, H.; Zhang, S.; Peng, J.; Zhang, S.; Li, J.; Xiong, H.; Zhang, W. Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 11106–11115. [\[CrossRef\]](#)
62. Dong, N.; Chang, J.-F.; Wu, A.-G.; Gao, Z.-K. A novel convolutional neural network framework based solar irradiance prediction method. *Int. J. Electr. Power Energy Syst.* **2020**, *114*, 105411. [\[CrossRef\]](#)
63. David, M.; Luis, M.A.; Lauret, P. Comparison of intraday probabilistic forecasting of solar irradiance using only endogenous data. *Int. J. Forecast.* **2018**, *34*, 529–547. [\[CrossRef\]](#)
64. Wan, C.; Zhao, J.; Song, Y.; Xu, Z.; Lin, J.; Hu, Z. Photovoltaic and solar power forecasting for smart grid energy management. *CSEE J. Power Energy Syst.* **2015**, *1*, 38–46. [\[CrossRef\]](#)
65. Golestaneh, F.; Gooi, H.B.; Pinson, P. Generation and evaluation of space–time trajectories of photovoltaic power. *Appl. Energy* **2016**, *176*, 80–91. [\[CrossRef\]](#)
66. Angelopoulos, A.N.; Bates, S. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. *Found. Trends Mach. Learn.* **2023**, *16*, 494–591. [\[CrossRef\]](#)
67. Gal, Y.; Ghahramani, Z. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. *Proc. Int. Conf. Mach. Learn.* **2016**, *48*, 1050–1059.
68. Kingma, D.P.; Welling, M. Auto-encoding variational Bayes. *arXiv* **2013**, arXiv:1312.6114.
69. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In *Proceedings of the 28th International Conference on Neural Information Processing Systems—Volume 2*; MIT Press: Cambridge, MA, USA, 2014; Volume 27.
70. Wang, F.; Zhen, Z.; Wang, B.; Mi, Z. Comparative study on KNN and SVM based weather classification models for day ahead short term solar PV power forecasting. *Appl. Sci.* **2018**, *8*, 28.
71. Bacher, P.; Madsen, H.; Nielsen, H.A. Online short-term solar power forecasting. *Sol. Energy* **2009**, *83*, 1772–1783. [\[CrossRef\]](#)
72. Tyralis, H.; Papacharalampous, G.; Langousis, A. A brief review of random forests for water scientists and practitioners and their recent history in water resources. *Water* **2019**, *11*, 910. [\[CrossRef\]](#)
73. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794.
74. Lorenz, E.; Hurka, J.; Heinemann, D.; Beyer, H.G. Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2009**, *2*, 2–10. [\[CrossRef\]](#)
75. Yagli, G.M.; Yang, D.; Srinivasan, D. Automatic hourly solar forecasting using machine learning models. *Renew. Sustain. Energy Rev.* **2019**, *105*, 487–498. [\[CrossRef\]](#)
76. Li, P.; Zhou, K.; Lu, X.; Yang, S. A hybrid deep learning model for short-term PV power forecasting. *Appl. Energy* **2020**, *259*, 114216. [\[CrossRef\]](#)

77. Yin, W.; Han, Y.; Zhou, H.; Ma, M.; Li, L.; Zhu, H. A Novel Non-Iterative Correction Method for Short-Term Photovoltaic Power Forecasting. *Renew. Energy* **2020**, *159*, 23–32. [\[CrossRef\]](#)
78. Yan, J.; Hu, L.; Zhen, Z.; Wang, F.; Qiu, G.; Li, Y.; Yao, L.; Shafie-khah, M.; Catalão, J.P.S. Frequency-Domain Decomposition and Deep Learning Based Solar PV Power Ultra-Short-Term Forecasting Model. *IEEE Trans. Ind. Appl.* **2021**, *57*, 3282–3295. [\[CrossRef\]](#)
79. Huang, Y.; Wang, A.; Jiao, J.; Xie, J.; Chen, H. Short-Term PV Power Forecasting Based on CEEMDAN and Ensemble DeepTCN. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 2526012. [\[CrossRef\]](#)
80. Asrari, A.; Wu, T.X.; Ramos, B. A Hybrid Algorithm for Short-Term Solar Power Prediction—Sunshine State Case Study. *IEEE Trans. Sustain. Energy* **2016**, *8*, 582–591.
81. Sheng, H.; Xiao, J.; Cheng, Y.; Ni, Q.; Wang, S. Short-Term Solar Power Forecasting Based on Weighted Gaussian Process Regression. *IEEE Trans. Ind. Electron.* **2017**, *65*, 300–308. [\[CrossRef\]](#)
82. Mellit, A.; Pavan, A.M.; Lughi, V. Deep Learning Neural Networks for Short-Term Photovoltaic Power Forecasting. *Renew. Energy* **2021**, *172*, 276–288. [\[CrossRef\]](#)
83. Huang, C.; Yang, M. Memory Long and Short Term Time Series Network for Ultra-Short-Term Photovoltaic Power Forecasting. *Energy* **2023**, *279*, 127961. [\[CrossRef\]](#)
84. Du Plessis, A.A.; Strauss, J.M.; Rix, A.J. Short-Term Solar Power Forecasting: Investigating the Ability of Deep Learning Models to Capture Low-Level Utility-Scale Photovoltaic System Behaviour. *Appl. Energy* **2021**, *285*, 116395. [\[CrossRef\]](#)
85. Liu, W.; Mao, Z. Short-Term Photovoltaic Power Forecasting with Feature Extraction and Attention Mechanisms. *Renew. Energy* **2024**, *226*, 120437. [\[CrossRef\]](#)
86. Chen, R.; Liu, G.; Cao, Y.; Xiao, G.; Tang, J. CGAformer: Multi-Scale Feature Transformer with MLP Architecture for Short-Term Photovoltaic Power Forecasting. *Energy* **2024**, *312*, 133495. [\[CrossRef\]](#)
87. Feng, Z.; Shen, J.; Zhou, Q.; Hu, X.; Yong, B. Hierarchical Gated Pooling and Progressive Feature Fusion for Short-Term PV Power Forecasting. *Renew. Energy* **2025**, *247*, 122929. [\[CrossRef\]](#)
88. Zhang, H.; Yang, J.; Fan, S.; Geng, H.; Shao, C. An Ultra-Short-Term Distributed Photovoltaic Power Forecasting Method Based on GPT. *IEEE Trans. Sustain. Energy* **2025**, early access.
89. Huang, S.; Zhou, Q.; Shen, J.; Zhou, H.; Yong, B. Multistage Spatio-Temporal Attention Network Based on NODE for Short-Term PV Power Forecasting. *Energy* **2024**, *290*, 130308. [\[CrossRef\]](#)
90. Ouyang, J.; Zuo, Z.; Wang, Q.; Duan, Q.; Zhu, X.; Zhang, Y. Seasonal Distribution Analysis and Short-Term PV Power Prediction Method Based on Decomposition Optimization Deep-Autoformer. *Renew. Energy* **2025**, *246*, 122903. [\[CrossRef\]](#)
91. Yu, W.; Dai, Y.; Wang, W.; Ren, T.; Leng, M. Short-Term Photovoltaic Forecasting: A Parallel TimesNet and AT-Informer-AT Method. *Renew. Energy* **2025**, *258*, 125012.
92. Tang, H.; Kang, F.; Li, X.; Sun, Y. Short-Term Photovoltaic Power Prediction Model Based on Feature Construction and Improved Transformer. *Energy* **2025**, *320*, 135213. [\[CrossRef\]](#)
93. Gong, J.; Qu, Z.; Zhu, Z.; Xu, H. Parallel TimesNet-BiLSTM Model for Ultra-Short-Term Photovoltaic Power Forecasting Using STL Decomposition and Auto-Tuning. *Energy* **2025**, *320*, 135286. [\[CrossRef\]](#)
94. Ding, L.; Li, Q.; Chen, X. Short-Term Photovoltaic Power Prediction Based on FCformer-BiTCN Method. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 9009813. [\[CrossRef\]](#)
95. Agoua, X.G.; Girard, R.; Kariniotakis, G. Short-Term Spatio-Temporal Forecasting of Photovoltaic Power Production. *IEEE Trans. Sustain. Energy* **2017**, *9*, 538–546.
96. Wang, Y.; Fu, W.; Wang, J.; Zhen, Z.; Wang, F. Ultra-Short-Term Distributed PV Power Forecasting for Virtual Power Plant Considering Data-Scarce Scenarios. *Appl. Energy* **2024**, *373*, 123890. [\[CrossRef\]](#)
97. Zhang, M.; Zhen, Z.; Liu, N.; Zhao, H.; Sun, Y.; Feng, C.; Wang, F. Optimal Graph Structure Based Short-Term Solar PV Power Forecasting Method Considering Surrounding Spatio-Temporal Correlations. *IEEE Trans. Ind. Appl.* **2022**, *59*, 345–357.
98. Wang, Y.; Zhen, Z.; Wang, F.; Sun, S.; Cheng, Y.; Yu, J. A Distributed PV Ultra-Short-Term Power Deterministic and Probabilistic Forecasting Based on Dynamic Graph Network with Shape-Amplitude Criteria. *IEEE Trans. Ind. Appl.* **2025**, early access.
99. Zhang, R.; Wu, Y.; Zhang, L.; Xu, C.; Wang, Z.; Zhang, Y.; Sun, X.; Zuo, X.; Wu, Y.; Chen, Q. A Multiscale Network with Mixed Features and Extended Regional Weather Forecasts for Predicting Short-Term Photovoltaic Power. *Energy* **2025**, *318*, 134792. [\[CrossRef\]](#)
100. Cheng, L.; Zang, H.; Wei, Z.; Ding, T.; Xu, R.; Sun, G. Short-Term Solar Power Prediction Learning Directly from Satellite Images with Regions of Interest. *IEEE Trans. Sustain. Energy* **2021**, *13*, 629–639. [\[CrossRef\]](#)
101. Wang, F.; Li, J.; Zhen, Z.; Wang, C.; Ren, H.; Ma, H.; Zhang, W.; Huang, L. Cloud Feature Extraction and Fluctuation Pattern Recognition Based Ultrashort-Term Regional PV Power Forecasting. *IEEE Trans. Ind. Appl.* **2022**, *58*, 6752–6767. [\[CrossRef\]](#)
102. Zang, H.; Chen, D.; Liu, J.; Cheng, L.; Sun, G.; Wei, Z. Improving Ultra-Short-Term Photovoltaic Power Forecasting Using a Novel Sky-Image-Based Framework Considering Spatial-Temporal Feature Interaction. *Energy* **2024**, *293*, 130538. [\[CrossRef\]](#)
103. Wu, X.; Zhen, Z.; Zhang, J.; Wang, F.; Xu, F.; Ren, H.; Su, Y.; Sun, Y.; Yang, H. Multidimensional Feature Extraction Based Minutely Solar Irradiance Forecasting Method Using All-Sky Images. *IEEE Trans. Ind. Appl.* **2024**, *60*, 4494–4504. [\[CrossRef\]](#)

104. Wei, X.; Yue, D.; Hancke, G.P.; Dou, C.; Li, H.; Qiu, Y. Ultra Short-Term Solar Irradiance Forecast Based on Multimodal Data Fusion and Fuzzification. *IEEE Trans. Ind. Inform.* **2025**, *21*, 3256–3265. [[CrossRef](#)]
105. Fu, X.; Ye, W.; Li, X.; Zeng, X.; Wang, Y.; Chang, F.; Zhang, J.; Liu, R. Revolutionising agri-energy: A comprehensive survey on the applications of artificial intelligence in agricultural energy internet. *Energy Internet* **2025**, *2*, 13.
106. Long, H.; Fu, X.; Kong, W.; Chen, H.; Zhou, Y.; Yang, F. Key technologies and applications of rural energy internet in China. *Inf. Process. Agric.* **2024**, *11*, 277–298. [[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.