



Photovoltaic power forecasting: A comprehensive review and taxonomy of forecasting methods

Jokūbas Jonuška, Robertas Damaševičius, Diana Belova-Plonienė, Andrius Katkevičius, Darius Plonis, Vaiva Rusen, Tomyslav Sledevič & Rytis Maskeliūnas

To cite this article: Jokūbas Jonuška, Robertas Damaševičius, Diana Belova-Plonienė, Andrius Katkevičius, Darius Plonis, Vaiva Rusen, Tomyslav Sledevič & Rytis Maskeliūnas (16 Sep 2026): Photovoltaic power forecasting: A comprehensive review and taxonomy of forecasting methods, International Journal of Green Energy, DOI: [10.1080/15435075.2026.2730629](https://doi.org/10.1080/15435075.2026.2730629)

To link to this article: <https://doi.org/10.1080/15435075.2026.2730629>



Published online: 16 Sep 2026.



Submit your article to this journal [↗](#)



Article views: 13



View related articles [↗](#)



View Crossmark data [↗](#)

Photovoltaic power forecasting: A comprehensive review and taxonomy of forecasting methods

Jokūbas Jonuška^{a*}, Robertas Damaševičius^{b*}, Diana Belova-Plonienė^{a*}, Andrius Katkevičius^{a*}, Darius Plonis^{a*}, Vaiva Rusen^{a*}, Tomyslav Sledevič^{a*}, and Rytis Maskeliūnas^{c*}

^aDepartment of Electronic Systems, Vilnius Gediminas Technical University, Vilnius, Lithuania; ^bOffice of Research and Development, Asia University, Taichung, Taiwan; ^cCenter of Real Time Computing Systems, Kaunas University of Technology, Kaunas, Lithuania

ABSTRACT

Photovoltaic (PV) power forecasting is a foundational capability for operating power systems with high shares of variable renewable generation. However, the methodological landscape is fragmented across physical-based models, statistical time-series approaches, machine learning (ML), and increasingly diverse hybrid and multi-modal pipelines combining SCADA measurements with numerical weather prediction (NWP) and remote-sensing data. This paper provides a unified, practice-oriented synthesis of PV power forecasting methods, linking forecasting tasks and horizons to data requirements, modeling assumptions, and deployment constraints. We first summarize PV plant architecture, the energy conversion chain and key definitional choices (AC vs. DC, power vs. energy, spatial aggregation, sampling interval). We then formulate the forecasting problem across point and probabilistic settings, distinguishing single- and multi-step strategies and showing how horizon choice shifts the balance between persistence, cloud-driven dynamics, and NWP-driven uncertainty. We propose a taxonomy covering (i) physics-based approaches (clear-sky and PV performance models), (ii) statistical and time-series baselines, (iii) ML/DL families — recurrent, convolutional, attention-based, and transformer architectures — and (iv) hybrid and ensemble methods, with attention to calibration and uncertainty quantification. Finally, we identify open challenges limiting robust deployment — non-stationarity, cross-site/climate generalization, probabilistic calibration, and explainability — and highlight emerging trends such as self-supervised learning, foundation models, and digital-twin-inspired forecasting loops.

KEYWORDS

Explainable artificial intelligence; model selection; numerical weather prediction (NWP); photovoltaic power forecasting; solar power prediction; time-series forecasting

Introduction

The rapid expansion of photovoltaic (PV) generation has become a central pillar of decarbonization strategies, yet it also introduces operational challenges because PV output is inherently variable, strongly weather-dependent, and shaped by site- and system-specific characteristics (Bessa et al. 2014; Hummon et al. 2014; Wan et al. 2015). In practical installations, the conversion chain – from PV modules and DC strings through maximum power point tracking (MPPT), inverters, and grid interconnection – creates multiple points where both physical constraints and operational events can influence the delivered AC power (de Oliveira et al. 2023). In addition to intrinsic component characteristics (e.g., module technology and degradation, string configuration, MPPT partitioning, inverter loading, and protection logics), external drivers such as irradiance dynamics, cloud transients, ambient temperature, wind-induced thermal effects, soiling, shading, and local environmental conditions can amplify short-term volatility and longer-term non-stationarity (Hummon et al. 2014; Wan et al. 2015). These interacting factors render PV generation forecasting a non-trivial, context-dependent learning and modeling problem whose difficulty increases when forecasts are required at fine temporal resolutions or under rapidly

changing atmospheric regimes (Nespoli et al. 2019; El Hendouzi and Bourouhou 2020; Theocharides et al. 2021). Accordingly, the literature consistently emphasizes that reliable PV forecasts are indispensable for day-ahead and intraday market participation, grid balancing and dispatch, storage scheduling, ramp management, and maintenance planning, while simultaneously noting that performance claims are often sensitive to location, horizon, data availability, and modeling assumptions (Bessa et al. 2014; Orozco et al. 2022; de Oliveira et al. 2023; Zhou et al. 2025).

Contemporary PV power forecasting research has, therefore, evolved along two complementary trajectories. First, there is an enduring line of work grounded in physical reasoning, including clear-sky and transposition modeling, performance modeling, and the structured incorporation of meteorological drivers. Such approaches remain particularly relevant when interpretability, robustness, and extrapolation under distribution shift are prioritized or when training data are limited. Second, data-driven methods – spanning statistical time-series models, machine learning, and deep learning – have gained prominence due to their capacity to learn complex nonlinearities and exploit multi-source inputs such as numerical weather prediction (NWP), satellite imagery, and on-site

sensor streams. Recent studies document both the promise and the pitfalls of these approaches, highlighting that improvements are not uniform across horizons and that data-driven models can be brittle under regime change without careful treatment of feature provenance, temporal alignment, and domain adaptation (Gupta and Singh 2021; Iheanetu 2022; Gaboitaolelwe et al. 2023). The increasing adoption of probabilistic forecasting and uncertainty quantification reflects a shift from purely point-accurate predictions toward decision-support outputs that better match operational risk and cost structures (Massidda et al. 2024). Within this landscape, practitioners still face a question: given a specific plant configuration and available data sources, which modeling family is most appropriate for a given operational horizon and deployment constraint?

Probabilistic and risk aware forecasting has increasingly become central to modern power system operations, as grid operators and market participants are required to make decisions under uncertainty rather than relying solely on point estimates. In power systems with significant photovoltaic penetration, uncertainty in irradiance forecasts, cloud dynamics, and operational constraints directly translates into balancing costs, reserve requirements, and market risks. Consequently, PV power forecasting research has gradually shifted from purely point accurate prediction toward probabilistic formulations that explicitly characterize forecast uncertainty through prediction intervals, quantiles, or scenario ensembles. Such probabilistic outputs enable risk aware dispatch, reserve sizing, and storage scheduling, and better align forecasting products with the decision theoretic needs of system operators and electricity markets.

Paper scope and objectives

This paper provides a structured, method-centric synthesis of PV power forecasting, explicitly linking (i) the physical and operational determinants of PV output variability with (ii) a taxonomy of forecasting method families and (iii) the data modalities typically used to support them. The scope is intentionally practical and system-aware: it consolidates the factors that shape PV generation at the plant level (including component-level effects and external environmental drivers), then uses these determinants to motivate why distinct forecasting paradigms emerge as advantageous under different horizons and data regimes. Building on established survey treatments of PV power forecasting, we emphasize a unified view that spans physical-based modeling, statistical baselines and time-series formulations, machine learning predictors, deep learning architectures, and hybrid/ensemble strategies, while clarifying the assumptions and failure modes that arise when these approaches are deployed with real-world PV plant measurements and weather products. In addition, because NWP-driven pipelines remain a dominant backbone for intraday and day-ahead forecasting, we discuss their role as a principal exogenous information source and summarize key considerations for site-specific integration, including the mismatch between model grid resolution and plant microclimate (Böök and Lindfors 2020). We further highlight recent methodological directions that illustrate the field's trajectory,

such as advanced deep learning for day-ahead prediction (e.g., transformer-based formulations) and the growing emphasis on explainability for operational trust and accountability (López Santos et al. 2022; Rajagukguk and Lee 2025).

Several comprehensive reviews published in recent years have significantly advanced the understanding of photovoltaic (PV) power forecasting by categorizing methods according to modeling paradigms, such as physical-based models, statistical approaches, machine learning, and deep learning techniques (Gupta and Singh 2021; Iheanetu 2022; Gaboitaolelwe et al. 2023). While these taxonomies remain valuable, they are increasingly insufficient to fully capture the challenges encountered in contemporary operational settings. In particular, prior reviews predominantly organize methods by algorithmic family or input modality, but provide limited guidance on how forecasting approaches should be selected and evaluated under different forecast horizons, data availability regimes, and deployment constraints.

Moreover, recent advances in data-driven modeling, probabilistic forecasting, and hybrid physical-learning pipelines have exposed limitations of purely method-centric taxonomies. Issues such as non-stationarity, operational interventions (e.g., clipping and curtailment), evaluation protocol sensitivity, and the growing gap between benchmark performance and real-world deployment are often treated only marginally or in isolation. As a result, existing reviews do not sufficiently articulate why certain method families succeed or fail when transferred across sites, climates, and operational contexts.

Our paper addresses these gaps by adopting a method-context perspective that explicitly links forecasting models to the physical structure of PV plants, forecast horizons, data regimes, and deployment constraints. Rather than proposing yet another algorithm-centric taxonomy, our aim is to synthesize the literature in a system-aware manner, clarifying when and why specific forecasting paradigms are appropriate, how evaluation practices influence reported performance, and which open challenges remain critical for robust real-world adoption.

Paper contribution

The main contributions are threefold. First, we consolidate the PV generation determinants that matter most for forecasting, spanning plant structure and component behavior, meteorological forcing, and local environmental effects, so that model selection is grounded in the mechanisms that drive the signal to be predicted. Second, we propose a coherent taxonomy of PV power forecasting methods that organizes the literature by modeling philosophy, required inputs, and typical deployment settings, thereby helping readers navigate the expanding space of approaches documented in recent comparative studies (Ledmaoui et al. 2023; Alay et al. 2024). Third, we articulate a method-context fit perspective that connects horizon, data availability, and operational constraints to suitable model families, while maintaining a clear separation from evaluation-centric questions (metrics, benchmarking protocols, and fairness in comparison).

Recently, a new research direction has emerged around foundation models and large language models (LLMs) for energy and renewable forecasting, inspired by advances in large-scale self-supervised pretraining and representation learning. Time series foundation models and LLM based forecasting frameworks have demonstrated promising zero shot and few shot capabilities across energy datasets, motivating early applications to renewable generation and photovoltaic power forecasting (Marchesi et al. 2025). While these approaches show potential benefits in transferability and multi-horizon learning, their operational maturity, computational cost, and robustness under distribution shift remain active research questions. Accordingly, foundation-model-based forecasting is treated in this review as an emerging and complementary paradigm within the broader machine learning and hybrid forecasting landscape.

In this context, both point and probabilistic forecasting formulations are treated as first class problem settings, with uncertainty aware methods receiving explicit emphasis in the proposed taxonomy and method–context selection map.

Literature review protocol

The search of papers was conducted across several major scientific databases, targeting peer-reviewed publications from 2010 to 2026 with a focus on the newest developments in the field. Web of Science, IEEE Xplore, and Scopus databases were searched. The following keywords were used to search papers: “photovoltaic power forecasting,” “PV power forecasting,” “photovoltaic forecasting,” “solar power prediction,” “solar generation forecasting,” “machine learning,” “numerical weather prediction,” and “deep learning.” Boolean operators were applied to refine and combine search terms. Additionally, reference lists of important review articles were reviewed. After collection, references were deduplicated, and titles and abstracts were screened for relevance. Studies were included if they focused on PV power forecasting or presented methods that are directly applicable to PV power forecasting. Articles were excluded if they were non-English, lacked quantitative validation, or were primarily oriented toward economics or policy rather than forecasting methodology. More than 80 % publications in the reference list are from last 5 years. This reflects the focus on latest research. The screening process was designed to show the full spectrum of PV power forecasting approaches, including physical-based models, statistical time-series models, machine learning models, hybrid approaches, and other relevant techniques, thereby enabling the development of a comprehensive and systematic taxonomy of methods in this domain.

Paper organization

The remainder of the paper is organized as follows. Section 2 introduces the PV power generation and forecasting context, including the PV plant structure and the definition of forecasting targets at different spatial and temporal granularities. Section 3 synthesizes the determinants of PV variability, distinguishing technical and system-level factors from meteorological and site-specific drivers, and discussing their

implications for modeling under non-stationarity. Section 4 formalizes the forecasting problem and summarizes data sources and workflows, including the role of on-site measurements, NWP products, and remote sensing inputs. Section 5 presents the core taxonomy of PV power forecasting methods, covering physical-based models, statistical time-series approaches, machine learning and deep learning predictors, and hybrid/ensemble strategies, with attention to data requirements and typical use-cases. Section 6 integrates these strands into a method–context selection map aligned with operational horizons and deployment constraints. Finally, Section 7 outlines open challenges and research directions – spanning generalization across sites, uncertainty-aware forecasting, and explainable models – before Section 8 concludes the paper.

Background: photovoltaic power generation and forecasting context

Overview of photovoltaic (PV) power systems

A utility-scale or distributed photovoltaic (PV) plant is best understood as a tightly coupled electromechanical–cyberphysical system whose observable output power is the end product of multiple interacting subsystems (Bessa et al. 2014; Wan et al. 2015). In its most common grid-connected configuration, the plant comprises (i) PV modules assembled into series-connected strings (and, depending on topology, parallel-connected string groups), (ii) mounting structures (fixed-tilt or tracking) that determine geometric exposure to irradiance and impose mechanical operating constraints, (iii) DC-side cabling and protection, (iv) DC/AC conversion via inverters, (v) AC-side cabling and protection, and (vi) step-up transformation and medium-voltage interconnection where required (de Oliveira et al. 2023; Zhu et al. 2025). The material summarized in this work highlights that PV modules are typically assembled from many interconnected cells (most commonly mono-crystalline silicon, but also other technologies), with factory-installed positive/negative connectors enabling modular string formation and straightforward integration into DC circuits (de Oliveira et al. 2023). Strings are then routed through DC cabling to inverter input stages, and the plant’s conversion and interconnection chain is completed on the AC side through feeder circuits and, for larger installations, transformer stations that adapt inverter output to distribution or transmission-level voltage (Zhu et al. 2025). This architecture is not merely a physical layout; it defines where losses occur, where operational constraints enter, and which internal states are observable – issues that ultimately condition both the attainable forecast accuracy and the most appropriate modeling strategy reported in the PV power forecasting literature (Wan et al. 2015; El Hendouzi and Bourouhou 2020; Gupta and Singh 2021; Theocharides et al. 2021; Prema et al. 2022; Gaboitaolelwe et al. 2023).

From a forecasting standpoint, a key implication of this architecture is that the PV plant is rarely a homogeneous generator: the same nominal capacity can exhibit markedly different dynamic behavior depending on string configuration, inverter sizing ratio (DC/AC), tracking versus fixed tilt, and protection logic (de Oliveira et al. 2023; Zhu et al. 2025). The literature on PV power forecasting has, therefore, converged

on the view that plant-level predictors must be interpreted in the context of the physical and operational configuration, rather than treated as purely statistical time series (Bessa et al. 2014; Wan et al. 2015; Nespoli et al. 2019; Theocharides et al. 2021). This motivates an explicit system-oriented taxonomy of forecasting methods (Section 5) is best understood as a set of modeling responses to how PV plants are constructed and operated, and to which internal and external drivers shape the measured output.

The forecasting target is derived from the PV plant's energy conversion chain (Figure 1), which begins with irradiance-to-DC conversion at the module level and ends with grid-injected AC power at the point of interconnection (Wan et al. 2015; de Oliveira et al. 2023). Along this chain, power is shaped by conversion efficiencies, temperature-dependent effects, wiring and transformation losses, and by the supervisory and protective actions of power electronics (de Oliveira et al. 2023; Zhu et al. 2025). Common measurement streams arise from (i) inverter telemetry (DC voltage/current, DC power, AC power, AC voltage/current, power factor, inverter temperature, alarms), (ii) plant-level meters at the point of common coupling (PCC), and (iii) meteorological instrumentation (global horizontal irradiance and/or plane-of-array irradiance, ambient and/or module temperature, wind speed, humidity) when available (Nespoli et al. 2019; El Hendouzi and Bourouhou 2020; Theocharides et al. 2021). We emphasize the practical role of plant monitoring devices (e.g., data loggers

and supervisory systems) not only as data providers but also as operational safeguards, whose status information can be essential for distinguishing meteorologically driven power reductions from equipment failures or curtailment events (Bessa et al. 2014; Zhu et al. 2025).

These measurement points (Table 1) define what can be learned from historical data and what must be supplied exogenously (e.g., through numerical weather prediction, satellite products, or sky imagery) (Wan et al. 2015; El Hendouzi and Bourouhou 2020). This observation aligns with the methodological trends reported in recent PV power forecasting studies: physical-based modeling chains explicitly map irradiance and temperature to expected DC and AC power, whereas data-driven predictors rely on historical telemetry and exogenous meteorological forecasts to infer the latent mapping under the operational constraints of a specific plant (Nespoli et al. 2019; Theocharides et al. 2021; de Oliveira et al. 2023). Comparative studies of physical-based model chains further show that the choice of measurement target (e.g., DC-side versus AC-side power) and the availability of intermediate telemetry can substantially alter model identifiability and error structure, reinforcing the need to state measurement points and data provenance explicitly when reporting forecasting results (Bessa et al. 2014; Nespoli et al. 2019; Mayer and Gróf 2021; Theocharides et al. 2021).

PV plants operate under regimes that are strongly shaped by inverter control and grid-interconnection rules (Bessa et al. 2014; Zhu et al. 2025). A defining feature is maximum power point tracking (MPPT), whereby the inverter (often via multiple independent MPPT channels) continuously adjusts the DC operating point of each connected string group to maximize harvested power given instantaneous irradiance and temperature (de Oliveira et al. 2023). The material emphasizes that multiple MPPT blocks allow the inverter to regulate distinct string groups simultaneously, which is operationally critical under partial shading, string mismatch, or non-uniform irradiance fields (de Oliveira et al. 2023). However, MPPT also implies that the mapping from meteorological conditions to measured AC power is mediated by control actions, not solely by passive physics (Wan et al. 2015). In addition, inverter limitations (e.g., apparent power constraints, thermal derating), clipping under high irradiance in DC/AC oversize plants, and curtailment commands imposed by the grid operator can

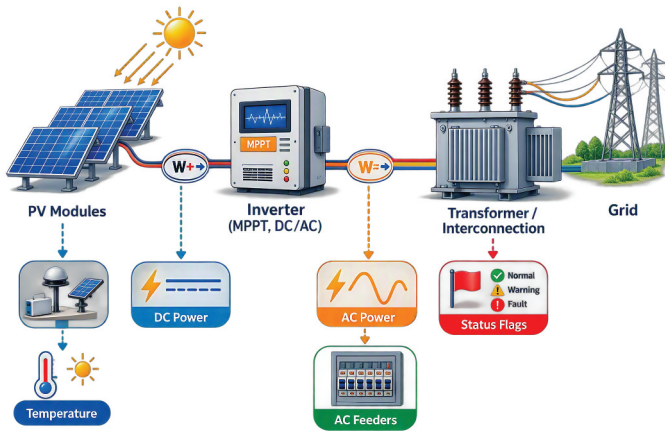


Figure 1. Conceptual PV plant architecture and typical measurement points used in forecasting pipelines.

Table 1. PV plant subsystems and typical measurement points relevant for forecasting.

Subsystem	Primary role in the conversion chain	Typical measurements / status signals
PV modules and strings	Irradiance → DC conversion; defines DC operating envelope	String DC voltage/current (if available), inverter DC inputs, module temperature (if instrumented), string status/fault flags
DC cabling and protection	Power routing; losses; safety isolation	Insulation/ground fault alarms, breaker/fuse states, estimated DC losses (often implicit)
Inverter with MPPT	MPPT control; DC/AC conversion; grid-code compliance	AC power, DC power, MPPT voltages/currents, inverter temperature, curtailment setpoints, alarms
AC feeders and protection	Plant AC routing; safety and isolation	Feeder currents/voltages, breaker states, protection trips
Transformer / interconnection	Voltage adaptation; export metering point (PCC)	PCC active/reactive power, voltage, frequency, power factor, export limitations
Meteorological station	Exogenous drivers of generation	Irradiance (GHI/POA), ambient temperature, wind speed, humidity, precipitation, cloud cover proxies

create non-linearities and discontinuities in the time series (Bessa et al. 2014; Zhu et al. 2025).

These control-driven effects have direct consequences for forecasting model design. Physical-based models must account for inverter efficiency curves, clipping, and temperature-dependent derating; data-driven models must either learn these regimes from representative historical data or incorporate auxiliary signals (e.g., curtailment flags, inverter temperature, clipping indicators) to avoid attributing operational constraints to meteorology (Nespoli et al. 2019; Theocharides et al. 2021; de Oliveira et al. 2023). The growing emphasis on hybrid approaches – pairing meteorological prediction with physical conversion chains and learning residual corrections – can be interpreted as a methodological response to these control-mediated nonlinearities (Mayer 2022; Mayer and Yang 2023). Recent literature also reports that probabilistic formulations become particularly valuable in regimes dominated by rapidly changing cloud fields and control actions, where deterministic mappings are insufficient to capture the uncertainty induced by both meteorological and operational variability (Massidda et al. 2024).

PV generation as a forecasting target

A central design decision in PV power forecasting is the definition of the predictand (Wan et al. 2015; El Hendouzi and Bourouhou 2020). While many studies describe the task as “PV power forecasting,” the operational meaning varies depending on whether the target is instantaneous (or average-over-interval) *power* (kW/MW) or aggregated *energy* yield (kWh/MWh) over a billing or market interval (de Oliveira et al. 2023). Power forecasting is naturally aligned with dispatch, balancing, and ramp management, whereas energy forecasting aligns with yield assessment, settlement, and planning (Bessa et al. 2014; Zhu et al. 2025). A second, often under-specified distinction is whether the target is DC-side power (closer to the photovoltaic conversion process) or AC-side exported power (after conversion and potential curtailment) (Nespoli et al. 2019; Theocharides et al. 2021). In practice, AC export power at the PCC is most directly relevant to grid operations and market participation, but it also embeds inverter behavior, clipping, and curtailment, which can be advantageous (as it matches the operator’s objective) or problematic (as it mixes meteorological and operational effects) (Bessa et al. 2014; Zhu et al. 2025). DC-side targets can isolate the irradiance-to-electrical conversion more directly, but may be less available and less directly actionable for grid operations (de Oliveira et al. 2023). Method surveys increasingly stress that these target-definition choices partly explain why accuracy claims are not directly comparable across studies, even when

the same nominal horizon is reported (Wan et al. 2015; El Hendouzi and Bourouhou 2020; de Oliveira et al. 2023).

Table 2 summarizes common target definitions and their typical motivations. Importantly, the same plant can support multiple targets simultaneously (e.g., inverter-level AC power for control and AC export power at the PCC for grid services), which suggests that multi-task formulations or hierarchical forecasting can be appropriate when the measurement infrastructure permits. This perspective is consistent with recent methodological developments that leverage multi-source data and hierarchical structures to improve robustness and interpretability across operating regimes (Gaboitoalelwe et al. 2023).

PV power forecasting can be formulated at multiple spatial granularities, ranging from component-level (module or string) to inverter, plant, feeder, or region-level aggregation. Finer spatial granularity offers advantages for fault localization and for capturing heterogeneous irradiance effects (e.g., partial shading, spatially varying cloud cover), but it increases data volume and requires richer telemetry. Coarser aggregation smooths local variability and is often the relevant scale for market and grid decisions, but it can conceal localized failures and can introduce complex aggregation effects, especially in geographically dispersed portfolios. The literature reflects this trade-off: single-plant studies often exploit plant-specific telemetry to model idiosyncratic behavior, whereas regional forecasting emphasizes spatial correlation structures and the fusion of meteorological products across sites (Gupta and Singh 2021; Iheanetu 2022).

From a systems perspective, the PV plant architecture naturally induces a spatial hierarchy (string → MPPT channel → inverter → plant → PCC export), which can be exploited for model design and evaluation. For example, multi-site or distributed forecasting has motivated vector autoregression, spatiotemporal deep learning, and, more recently, large-model approaches that attempt to learn transferable representations across heterogeneous PV fleets (Jung et al. 2022; Fan et al. 2025). In the context of this paper, spatial granularity is treated as a first-order dimension in method selection: the feasibility and benefit of particular model families (e.g., attention-based spatiotemporal models) is conditional on whether the available data resolve the hierarchy.

Temporal granularity is equally consequential. PV generation exhibits dynamics across multiple time scales: second-to-minute variations due to fast-moving clouds, intra-hour variability due to evolving atmospheric conditions, and seasonal patterns driven by solar geometry and climate. Consequently, the sampling interval (e.g., 1-min, 5-min, 15-min, hourly) and the forecast update cycle define the information content and the dominant sources of uncertainty. Finer sampling supports

Table 2. Common PV power forecasting targets and their typical use-cases.

Target	Definition	Typical use-case / implication
AC power (inverter)	Inverter-level active power (kW)	Local control, anomaly detection, component-level diagnostics; includes MPPT and inverter constraints
AC export (PCC)	Net exported active power (kW/MW)	Grid balancing, market bids, dispatch; embeds curtailment and interconnection constraints
DC power (strings/ inverter DC)	DC-side power prior to conversion	Physical interpretability; isolates PV conversion effects; may reduce confounding from grid actions
Energy yield	Integrated energy over interval (kWh/MWh)	Yield assessment, settlement, planning; sensitive to temporal aggregation

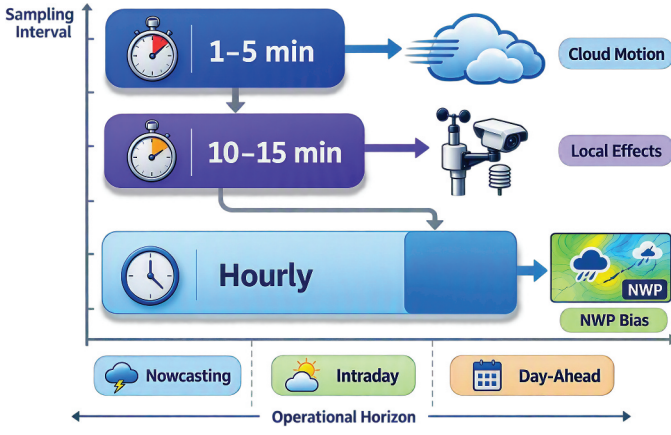


Figure 2. Illustrative relationship between sampling interval, forecast horizon, and dominant uncertainty sources in PV power forecasting.

nowcasting and ramp prediction but amplifies sensitivity to measurement noise and alignment errors; coarser sampling can be sufficient for day-ahead scheduling but depends more heavily on the accuracy of meteorological forecasts. Reviews and comparative analyses consistently report that horizon-dependent dominance patterns emerge. Persistence-like baselines can be surprisingly competitive at very short horizons, while NWP-driven predictors become increasingly important as lead time grows (Böök and Lindfors 2020; Prema et al. 2022; Massidda et al. 2024).

To ensure conceptual clarity for later sections, we adopt the convention that the reported “PV power” at a given timestamp represents an average over the sampling interval unless explicitly stated otherwise, and we treat temporal aggregation as a modeling choice that changes both the error distribution and the operational meaning of a forecast. Figure 2 provides a placeholder for a schematic that situates typical sampling intervals and application-driven horizons. This framing is essential because the same forecasting method can exhibit different strengths depending on whether it is trained and evaluated on high-frequency telemetry (where cloud ramps dominate) or on hourly aggregates (where systematic NWP biases and diurnal structure dominate), a phenomenon repeatedly observed in comparative studies of forecasting pipelines (Ledmaoui et al. 2023; Alay et al. 2024).

Determinants of PV power variability

Accurate and *fairly evaluated* PV power forecasting requires an explicit understanding of what drives variability in the measured signal. PV output variability is not solely meteorological; it is a coupled consequence of (i) electro-physical conversion at the module level, (ii) power-electronic control and balance-of-system (BoS) constraints, (iii) environmental forcing across multiple temporal scales, and (iv) site- and operator-specific interventions that impose discontinuities and regime shifts. Contemporary studies emphasize that many cross-study “accuracy gaps” can be traced to differences in which of these determinants (Figure 3) are present in the

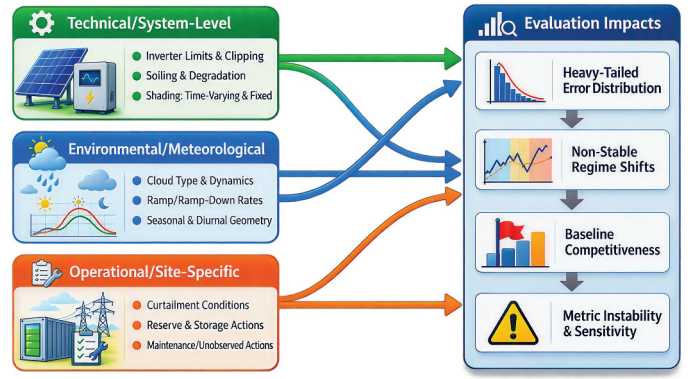


Figure 3. Conceptual map of PV power variability determinants and their implications for evaluation design.

evaluated data and whether the protocol explicitly controls for them (e.g., filtering curtailment or clipping regimes, aligning irradiance predictors to power measurements, or stratifying by cloudiness) (Gupta and Singh 2021; Iheanetu 2022; Prema et al. 2022; Gaboitaolelwe et al. 2023). In the context of this paper, these determinants matter because they shape (a) the attainable error floor, (b) the distribution of forecast errors (heavy tails during ramps), and (c) the validity of commonly used metrics and baselines at different horizons (Ledmaoui et al. 2023; Alay et al. 2024).

A useful abstraction is to express exported AC power as a constrained transformation of an underlying meteorological potential:

$$P_t^{AC} = \text{C0pt}(\eta_{inv}(s_t)\eta_{BoS}(t)P_t^{DC}, u_t), \quad (1)$$

where $\eta_{inv}(s_t)$ is inverter efficiency (dependent on internal state s_t , such as temperature and loading), $\eta_{BoS}(t)$ aggregates wiring and transformer losses, and $\text{C}(\cdot, u_t)$ denotes constraint operators induced by operational interventions u_t (e.g., curtailment setpoints, clipping due to DC/AC ratio, protection trips). The DC-side power can be approximated as

$$P_t^{DC} \approx A_{eff} \eta_{mod}(T_{cell,t})G_t^{POA} \xi_t, \quad (2)$$

where A_{eff} is effective module area, $\eta_{mod}(\cdot)$ is temperature-dependent conversion efficiency, G_t^{POA} is plane-of-array irradiance, and ξ_t represents multiplicative factors such as soiling and mismatch. Eqs. (1)–(2) highlight why evaluation protocols must (i) specify whether the target is DC or AC power, (ii) document the presence of constraints (clipping/curtailment), and (iii) account for time-varying multiplicative effects that create non-stationarity and may invalidate simple baselines if not explicitly modeled (Böök and Lindfors 2020; Prema et al. 2022).

Technical and system-level factors

Module technology and aging create a slowly varying envelope of achievable power, while temperature introduces fast-to-intermediate scale modulation through efficiency changes and thermal dynamics. Empirically, even when irradiance is accurately known, temperature-driven efficiency variation can

induce systematic biases if the evaluation compares models trained in one seasonal regime against test periods with different thermal conditions. Degradation (e.g., gradual efficiency loss, increased series resistance) contributes to long-term drift, and its effect is often confounded with changes in operational practice (cleaning schedules, component replacement). Reviews of PV power forecasting stress that such slow drifts are a major contributor to non-stationarity, and that evaluation must either (a) adopt rolling-origin schemes that emulate real-time adaptation or (b) explicitly limit test periods to regimes where drift is negligible, while disclosing this choice (Gupta and Singh 2021; Prema et al. 2022). In addition, the temperature sensitivity of $\eta_{mod}(T_{cell,t})$ implies that using ambient temperature as a proxy without accounting for wind and mounting effects can bias models in ways that manifest differently across sites and seasons, reducing the transferability of accuracy claims (Iheanetu 2022).

Power-electronic conversion introduces both smooth non-linearities (efficiency curves, reactive power control) and hard constraints (apparent power limits, thermal derating). For plants with DC/AC oversizing, clipping can induce a “plateau” in P_t^{AC} during high-irradiance periods, yielding a distinctive error structure: models that predict unconstrained meteorological potential may appear systematically biased unless clipping is modeled or filtered in evaluation. Balance-of-system losses (wiring, transformers) can be approximately constant in relative terms but can also vary with temperature and loading. From an evaluation perspective, inverter/BoS effects motivate regime labeling (normal vs clipped vs derated) and stratified reporting, because aggregated metrics can conceal that a model performs well in unconstrained regimes but fails during clipped or derated periods, which can dominate RMSE depending on the sampling interval (Ledmaoui et al. 2023; Alay et al. 2024). Hybrid frameworks that combine physical-based chains with learned residuals have been proposed precisely to handle such conversion-chain artifacts more robustly, but their reported improvements depend strongly on whether evaluation explicitly includes these regimes (Prema et al. 2022).

Soiling, partial shading, and maintenance interventions introduce variability that is not directly explained by meteorological inputs and is often poorly documented in public datasets. These effects are typically multiplicative (captured by ξ_t in Eq. (2)) and can evolve abruptly (e.g., cleaning events) or gradually (dust accumulation). If not measured, they appear as structured residuals, which can either be spuriously learned by data-driven models in a site-specific manner or misattributed to weather predictors, undermining generalization. Consequently, evaluation protocols that mix pre- and post-maintenance periods without event labeling risk reporting inflated performance for models that implicitly memorize maintenance schedules rather than learning robust meteorology-to-power mappings (Gupta and Singh 2021; Prema et al. 2022). For fairness, a recommended practice is to (i) document known cleaning/maintenance intervals, (ii) exclude clear anomalies (e.g., long outages) from “normal operation” benchmarking, and (iii) optionally report separate metrics for soiled vs clean regimes when proxies exist.

SCADA and metering systems introduce both random noise and systematic errors. Calibration drift in irradiance sensors, timestamp inconsistencies, and quantization effects can create artificial ramps or suppress true variability, altering the empirical distribution of errors. This matters because PV power forecasting metrics are sensitive to these distributions; for example, RMSE can be dominated by a small number of sensor spikes, while MAE may be comparatively stable. Evaluation that uses exogenous meteorological forecasts (NWP) can be subtly compromised by measurement biases if the target is biased but treated as ground truth. Therefore, rigorous evaluation requires (i) quality checks for physically plausible bounds, (ii) drift detection where long-term stability is assumed, and (iii) transparent missingness/imputation policies.

Anomalies caused by non-meteorological factors, such as inverter clipping and sensor drift, require additional analysis. Inverter clipping appears when the PV-generated DC power exceeds the inverter’s AC conversion capacity, causing the AC output to saturate at the inverter rating. This phenomenon is not a fault, but rather a design-dependent constraint. Effective protocols to address invested clipping in pre-processing combine: (i) high-resolution SCADA filtering to remove clipping intervals, (ii) ratio and peer-comparison analyses to flag persistence losses, and (iii) model- and trend-based residual diagnostics to separate true degradation or faults from design-induced clipping and reversible losses (Balfour et al. 2021; Hara and Yoshioka 2021).

Sensor drift refers to the gradual bias in sensor readings over time due to aging, contamination, or exposure, affecting irradiance, temperature, or other sensors in the PV systems. Sensor drift can easily mimic module degradation rates; thus, separating it from real loss is important. Effective protocols rely on combination of: (i) reference to independent irradiance or models, (ii) systematic sensor-drift detection and correction, (iii) time-series method to isolate non-physical shifts, (iv) occasional on-site electrical and visual checks (Li 2016; Jordan et al. 2018).

Environmental and meteorological drivers

PV output is primarily driven by solar irradiance incident on the array plane. The decomposition of irradiance into global horizontal irradiance (GHI), direct normal irradiance (DNI), and diffuse horizontal irradiance (DHI) is central because each component contributes differently depending on tilt, tracking, and atmospheric conditions. While many operational pipelines rely on NWP-provided irradiance products, empirical studies show that the choice of irradiance representation (e.g., clear-sky index, clearness index, or direct irradiance components) can materially affect forecasting performance and interpretability, especially across varying cloud regimes (Lauret et al. 2022). Evaluation protocols should, therefore, state which irradiance variables are used (measured vs forecast), how they are transformed (e.g., normalization by clear-sky models), and whether such transformations are computed causally (to avoid leakage). This is particularly important in cross-study comparison because different irradiance proxies

can “bake in” information about seasonality and solar geometry, making some metrics appear improved without true predictive gains.

Cloud motion and cloud-edge effects induce high-frequency ramps and heavy-tailed forecast errors. These ramps are a primary reason why PV power forecasting remains challenging at minute-to-hour horizons and why model rankings can reverse when evaluated on ramp-rich subsets. Atmospheric aerosols, water vapor, and cloud optical thickness modulate both direct and diffuse components, leading to distinct error regimes: rapid transient errors (timing and magnitude) versus systematic bias (e.g., persistent over/underestimation under certain cloud fields). Reviews consistently emphasize that forecasting under broken-cloud conditions is the most demanding setting and that evaluation should be stratified by cloudiness regime when possible, rather than relying solely on aggregated scores (Prema et al. 2022; Gaboitaolelwe et al. 2023). When probabilistic forecasts are considered, these conditions are also where calibration failures (overconfidence) are most likely, motivating explicit reliability reporting (Massidda et al. 2024).

Thermal dynamics link meteorology to conversion efficiency and inverter behavior. Ambient temperature affects cell temperature, while wind influences convective cooling and, therefore, modifies $T_{\text{cell},t}$ at fixed irradiance. Humidity and precipitation can indirectly affect soiling and optical losses. For evaluation, these drivers matter because (i) they induce predictable diurnal and seasonal patterns, and (ii) they can be measured locally or approximated from NWP with varying fidelity. If the evaluation dataset lacks local measurements, models may be forced to rely on coarser exogenous predictors, increasing irreducible error and making comparisons against studies with richer instrumentation unfair unless explicitly documented (Böök and Lindfors 2020; Gupta and Singh 2021).

Solar geometry determines the baseline diurnal structure and seasonality of PV output via sun elevation and azimuth, influencing both irradiance magnitude and incidence angle on the array. Latitude and local horizon shape the seasonal contrast and day-length variability, which in turn affects the distribution of near-zero values and the prevalence of low-irradiance conditions. These factors have direct metric consequences: percentage-based errors (e.g., MAPE) can become unstable in seasons with many low-output intervals, and normalization choices can distort comparability across latitudes and climates. The evaluation literature cautions against reporting a single summary score without clarifying seasonal coverage and without disclosing how night and twilight periods are handled (Chicco et al. 2021; Prema et al. 2022). Accordingly, we recommend that benchmark reports include season-stratified metrics or at least disclose the seasonal composition of the test set.

Proper selection of evaluation metrics for PV generation forecasting is very important to ensure an accurate representation of forecasting model accuracy. Solar generation is highly intermittent, with frequent periods of low irradiance, which may cause instabilities in the percentage-based metrics, especially when the power values start approaching zero (Al-Dahidi et al. 2024; Liu et al. 2025). For this reason, under certain conditions, the selection of a single metric to measure

accuracy may be a suboptimal choice, and a combination of percentage-based and non-percentage-based metrics may provide insights into model performance. Particularly, absolute error metrics, such as MAE and RMSE, have been explored in the literature as tools to provide a more robust evaluation during periods of low irradiance (Al-Dahidi et al. 2024; Liu et al. 2025). However, it should be noted that absolute error metrics increase the difficulty of comparison between different applications. Additionally, the percentage-based metrics can provide valuable information under normal operating conditions. For these reasons, we recommend a combined approach to evaluation, using both methods – absolute metrics and percentage-based metrics. This approach ensures the interpretability of results across different circumstances and does not sacrifice the robustness of the evaluation during low or abnormal generation conditions.

Operational and site-specific factors

Local topography and horizon obstructions (terrain, vegetation, built structures) can induce morning/evening shading and can modify microclimate relative to the surrounding NWP grid cell. These effects are particularly relevant for distributed PV and for plants located in complex terrain, where representativeness errors can dominate day-ahead forecasts. Site-specific NWP adjustment and downscaling strategies have been proposed to mitigate this mismatch; however, their benefits cannot be assessed without evaluation protocols that accurately reflect the information available at issuance time and that account for the plant’s true shading geometry rather than implicitly absorbing it via historical power correlations (Böök and Lindfors 2020).

Curtailment, reactive power requirements, and grid constraints can decouple exported power from meteorological potential. From the perspective of Eq. (1), these interventions appear through $\mathcal{C}(\cdot, u_t)$ and can create sharp discontinuities or sustained caps. If curtailment periods are included in evaluation without labeling, models that predict “weather-driven power” will be penalized for errors that are not meteorological, while models that inadvertently learn curtailment patterns may appear better but will not generalize. Consequently, evaluation should treat curtailment as a first-class regime: either exclude it for “meteorological forecasting skill” benchmarks or include it but explicitly define the task as “export forecasting under operational constraints,” accompanied by access to curtailment signals in the information set $\mathcal{I}_t$ (Prema et al. 2022; Ledmaoui et al. 2023).

From a practical modeling perspective, the treatment of clipping and curtailment should be aligned with the forecasting objective. If the goal is to evaluate the ability of a model to predict meteorologically driven PV generation, periods affected by clipping or curtailment should be excluded from benchmarking or reported separately, as the resulting power reductions are not caused by atmospheric conditions. In contrast, if the forecasting target is actual exported AC power, clipping and curtailment periods should be retained and corresponding operational signals should be incorporated into the information set whenever available. We believe that such

a distinction improves both model interpretability and the fairness of performance comparisons across studies.

Storms, snow cover, hail, and extreme temperatures can induce regime changes that violate training assumptions, producing performance collapses that are invisible in average metrics. These events are operationally important because they often coincide with heightened grid stress. Therefore, evaluation should either (i) incorporate event-based scoring (e.g., ramp event detection, threshold exceedance), or (ii) explicitly report that extreme events are outside scope. Comparative studies underscore that without such reporting, aggregated metrics may misleadingly suggest operational readiness (Prema et al. 2022; Alay et al. 2024).

Implications for modeling

The determinants above jointly imply that PV power forecasting is fundamentally non-stationary. Formally, the conditional distribution $\mathbb{P}(P_{t+h}|\mathcal{I}_t)$ can drift over time as $\eta_{mod}(\cdot)$ degrades, ξ_t evolves with soiling and maintenance, and operational constraints change with grid conditions. This leads to distribution shift between training and test periods, particularly when evaluation spans multiple seasons or years. To avoid overstating performance, benchmark protocols should include rolling-origin backtesting and, where possible, multi-year tests that expose seasonal and operational shifts, in line with recommendations emerging from recent PV power forecasting studies (Prema et al. 2022; Gaboitoalelwe et al. 2023; Ledmaoui et al. 2023).

Determinants act on different time scales, implying that feature relevance is horizon-dependent. At very short horizons, persistence and local temporal autocorrelation can be highly informative, while cloud motion and local irradiance proxies dominate ramp prediction. At longer horizons, NWP variables become decisive, and slowly varying factors (seasonality, geometry, soiling state) shape baseline levels. Evaluation must respect this by (i) defining baselines that are horizon-appropriate and (ii) ensuring that feature sets used at each horizon are causally available. Otherwise, comparisons can be confounded by “feature privilege,” where one study uses higher-fidelity predictors (e.g., site-calibrated NWP, local POA sensors) and another does not, yet both report a single headline score (Böök and Lindfors 2020; Prema et al. 2022).

Finally, the determinants (Table 3) clarify why trade-offs between interpretability and accuracy are not merely philosophical but operationally grounded. Physical-based models explicitly encode Eq. (2)-like mechanisms and can be audited under regime changes, while high-capacity machine learning models can capture complex interactions but may be brittle under shifts in ξ_t or under unobserved constraints $\mathcal{C}(\cdot, u_t)$. The

evaluation protocol should, therefore, align the scoring design with the intended deployment objective: if explainability and auditability are required, then evaluation should include regime-stratified reporting and diagnostics that reveal when and why failures occur, rather than relying on a single scalar metric. This stance is consistent with the broader movement toward trustworthy and operationally meaningful PV power forecasting assessment documented in recent literature (Prema et al. 2022; Gaboitoalelwe et al. 2023; Massidda et al. 2024).

Forecasting fundamentals and problem formulation

Photovoltaic (PV) forecasting problems can be formulated at multiple levels of abstraction, yet the field has converged on a small set of task definitions, horizon conventions, and input modalities that recur across physical, statistical, and machine learning approaches. Recent studies emphasize that inconsistencies in these definitions are a primary reason why results are difficult to compare across studies and why apparently strong methods may fail when deployed under different operational constraints (Gupta and Singh 2021; Iheanetu 2022; Prema et al. 2022; Gaboitoalelwe et al. 2023). This section formalizes the main forecasting tasks, relates them to horizon-driven use-cases, summarizes the dominant data sources, and outlines a general workflow that supports reproducible model development and deployment.

Forecasting task definitions

Let $\mathcal{I}_t$ denote the information available at forecast origin time t (historical PV measurements, meteorological predictors, and any operational status signals). A *point forecast* produces a single value for PV output at lead time h :

$$\hat{y}_{t+h|t} = f_h(\mathcal{I}_t), \quad (3)$$

where y_t denotes the chosen target (e.g., AC power at the PCC or inverter, or energy aggregated over an interval). In most PV power forecasting studies, f_h is trained to approximate a central tendency (often the conditional mean), with evaluation performed via MAE/RMSE or their normalized variants (Prema et al. 2022; Ledmaoui et al. 2023). However, since PV output is strongly affected by meteorological uncertainty – especially cloud dynamics – many recent contributions argue that point accuracy alone is insufficient for operational decision-making, motivating *probabilistic forecasting* (Panamtash et al. 2020; Wang K et al. 2022; Massidda et al. 2024).

In probabilistic forecasting, the model outputs a predictive distribution $\hat{F}_{t+h|t}$ or a set of quantiles $\left\{ \hat{Q}_{t+h|t}(\tau) \right\}_{\tau \in \mathcal{T}}$:

Table 3. Determinants of PV variability and evaluation controls.

Determinant	Typical effect on P_t / errors	Recommended evaluation control
Module temperature, wind cooling	Systematic bias if thermal mapping mismatched	Report available thermal predictors; season-stratified scores
Degradation, soiling drift	Long-term downward drift; non-stationarity	Rolling-origin backtesting; document maintenance/cleaning
Inverter clipping/derating	Saturation plateaus; asymmetric errors	Label clipped/derated regimes; stratified reporting
Curtailment / grid constraints	Exogenous caps/discontinuities	Exclude or label curtailment; include signals in $\mathcal{I}_t$ if modeled
Cloud ramps	Heavy-tailed errors; timing sensitivity	Ramp-focused diagnostics; robust + tail-sensitive metrics
Sensor drift / timestamp misalign	Artificial ramps or lagged relationships	Alignment checks; QC thresholds; transparent missingness policy

$$\begin{aligned} c\hat{F}_{t+h|t}(z) &\approx \mathbb{P}(y_{t+h} \leq z | \mathcal{I}_t), \hat{Q}_{t+h|t}(\tau) \\ &= \inf \left\{ z : \hat{F}_{t+h|t}(z) \geq \tau \right\}, \end{aligned} \quad (4)$$

enabling prediction intervals and risk-aware scheduling. Bayesian and copula-based formulations, as well as Bayesian deep learning, have been explored to quantify uncertainty in PV generation, particularly for day-ahead settings where NWP uncertainty is a dominant contributor to forecast dispersion (Panamtash et al. 2020; Abedinia and Bagheri 2022; Wang K et al. 2022). Methodologically, probabilistic outputs also support decision-theoretic downstream objectives (e.g., storage dispatch under risk constraints), and they provide a principled interface for combining model uncertainty with operational costs (Massidda et al. 2024).

A closely related distinction concerns whether uncertainty is represented as a single distribution (or interval) or as an ensemble of plausible futures. In *deterministic* forecasting, the output is either a point estimate or a parametric distribution (e.g., Gaussian with mean and variance). In *scenario-based* forecasting, the model outputs a set of trajectories

$$\left\{ \tilde{y}_{t+1:t+H|t}^{(s)} \right\}_s^S = 1, \quad (5)$$

where each scenario s represents a coherent realization across a horizon window of length H . Scenario-based forecasting is particularly aligned with operational planning (unit commitment, portfolio risk assessment), because it preserves temporal dependence between successive future steps. In PV power forecasting, scenarios may be derived from NWP ensemble members, from generative statistical models conditioned on exogenous predictors, or from Bayesian approaches that sample from a posterior predictive distribution (Panamtash et al. 2020; Massidda et al. 2024). From a methodological taxonomy standpoint, scenario generation can be treated as an extension

layer: many model families (physical conversion chains, ML regressors, deep networks) can be “lifted” to scenarios through ensemble inputs or stochastic output layers (Figure 4).

Forecasting can be organized as *single-step* (one horizon at a time) or *multi-step* (predicting a sequence of future values). For a maximum horizon H , a multi-step forecast may be expressed as

$$\hat{y}_{t+1:t+H|t} = (\hat{y}_{t+1|t}, \dots, \hat{y}_{t+H|t}). \quad (6)$$

Two canonical strategies dominate. The *direct* strategy trains a separate mapping f_h for each horizon h , yielding Eq. (3) for every h . The *iterative* strategy trains a one-step model and recursively propagates predictions:

$$\hat{y}_{t+1|t} = f(\mathcal{I}_t), \quad \hat{y}_{t+k|t} = f(\mathcal{I}_t, \hat{y}_{t+1|t}, \dots, \hat{y}_{t+k-1|t}), \quad k \geq 2. \quad (7)$$

The direct strategy often yields better long-horizon accuracy at the expense of increased training complexity, while the iterative strategy is computationally simpler but can suffer from error accumulation, especially under rapidly changing cloud regimes (Prema et al. 2022; Ledmaoui et al. 2023). Deep learning approaches frequently adopt *multi-output* architectures that predict the entire vector in Eq. (6) jointly, which can improve coherence across time and better match operational needs (López Santos et al. 2022; Gaboitaolelwe et al. 2023).

Forecast horizons and use-cases (high-level overview)

Forecast horizon (Table 4) is a defining dimension in PV power forecasting because it governs which uncertainty sources dominate and which data modalities are informative. Surveys consistently report that very short horizons are often well served by persistence-like strategies and local measurements, whereas longer horizons increasingly depend on NWP products and their biases and resolution limits (Böök and Lindfors 2020; Gupta and Singh 2021; Prema et al. 2022). While exact nomenclature varies, it is common to distinguish three broad categories aligned to operational decisions.

Very-short-term forecasting also known as nowcasting (seconds to tens of minutes) supports control tasks such as ramp management, inverter-level regulation, and fast storage dispatch. At these horizons, temporal autocorrelation in the power signal and high-frequency irradiance proxies (e.g., sky imagery) can be more informative than coarse weather forecasts. Cloud-edge effects, measurement noise, and timestamp misalignment are also most consequential at this scale, which places stringent requirements on data quality control and evaluation design. Short-term forecasting (tens of minutes to a few hours) supports intraday balancing, flexible demand response, and operational reserve scheduling. Here, hybrid approaches that blend persistence baselines with exogenous

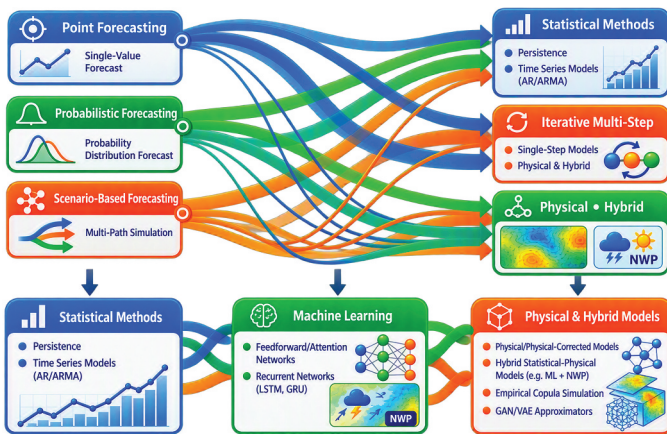


Figure 4. Task taxonomy and strategy choices for PV power forecasting.

Table 4. High-level mapping of PV power forecasting horizons to dominant inputs and operational use-cases.

Horizon	Dominant informative inputs	Typical operational use-cases
Seconds–minutes	Recent power history, local irradiance proxies, sky imagery	Ramp management, inverter control, fast storage dispatch
Minutes–hours	Persistence + exogenous predictors; local meteo; short-range NWP	Intraday balancing, reserve scheduling, flexible loads
Day-ahead (24–48 hr)	NWP irradiance/temperature; bias correction; site adaptation	Market bidding, dispatch planning, portfolio scheduling
Days–weeks	Climatology, reanalysis, seasonal geometry; long-term meteo services	Maintenance planning, asset management, yield expectations

predictors often yield robust gains, especially under changing cloud conditions (Prema et al. 2022; Gaboitaolelwe et al. 2023).

Intraday (hours ahead) and day-ahead forecasting (typically up to 24–48 hr) are central for market bidding, grid dispatch planning, and portfolio scheduling. At these horizons, the skill of NWP irradiance and temperature predictions is frequently the dominant driver of the accuracy of PV power forecasting, motivating site-specific NWP adjustments and bias correction (Böök and Lindfors 2020). Methodologically, day-ahead PV power forecasting has increasingly adopted modern deep learning architectures capable of fusing multiple predictors and capturing temporal dependencies; for example, transformer-based temporal fusion strategies have been proposed for day-ahead PV power prediction (López Santos et al. 2022). In parallel, probabilistic forecasting becomes more critical because decision-makers must manage uncertainty in meteorological forecasts; recent work highlights probabilistic day-ahead PV power forecasting as a practical direction for operational use (Massidda et al. 2024).

Medium-term horizons (days to weeks) support maintenance scheduling, asset management, and planning under seasonal variability. At these horizons, the signal is strongly structured by solar geometry and climatology, while short-term cloud dynamics are less relevant. Reanalysis products and long-term meteorological services can provide stable covariates for medium-term modeling, and baseline climatological models can be highly competitive. Evaluation at these horizons must also contend with non-stationarity from soiling, degradation, and operational changes, meaning that protocol choices (rolling-origin vs fixed splits) can dominate conclusions if not explicitly reported (Gupta and Singh 2021; Prema et al. 2022).

Data sources and input modalities

PV power forecasting models are ultimately constrained by the availability, quality, and temporal freshness of input data. The literature broadly categorizes inputs into on-site measurements, NWP products, remote sensing (satellite/sky-imager), and longer-horizon climatological sources (reanalysis and public meteorological services). Reviews stress that method comparisons that ignore input disparities (e.g., one study using site-calibrated irradiance sensors and high-resolution NWP, another using only historical power) are intrinsically confounded, and that input availability should be treated as a primary axis in taxonomy and benchmarking (Gupta and Singh 2021; Iheanetu 2022; Prema et al. 2022).

On-site SCADA streams provide the most direct access to the plant's operational state and measured output, typically including inverter AC power, DC voltage/current, internal temperatures, alarms, and curtailment indicators. When available, on-site meteorological stations contribute irradiance and environmental variables that can improve very-short-term and short-term forecasts by reducing error relative to coarse meteorological grid products. However, SCADA data are often noisy and incomplete, making robust data cleaning and causal feature construction essential for preventing spurious improvements that do not generalize (Prema et al. 2022; Ledmaoui et al. 2023).

NWP products remain the backbone for intraday and day-ahead PV power forecasting because they provide forward-looking meteorological information that cannot be inferred reliably from past power alone. Key variables include irradiance components, cloud cover proxies, temperature, wind speed, and humidity, often at multiple atmospheric levels. Site-specific NWP adjustment is frequently required to compensate for grid-cell mismatch and systematic bias, especially in complex terrain or coastal microclimates (Böök and Lindfors 2020). Ensemble NWP systems additionally support scenario-based and probabilistic forecasting by providing multiple plausible meteorological evolutions, which can be propagated through physical or learned conversion models (Massidda et al. 2024).

Remote sensing data provide high-value information for nowcasting and short-term horizons by directly observing cloud fields and their motion. Satellite imagery offers broad spatial coverage and supports cloud tracking at regional scales, while sky imagers provide high temporal resolution and local detail near the plant. In the method taxonomy, these modalities often motivate computer-vision and spatiotemporal learning approaches (CNNs, attention mechanisms), as well as hybrid pipelines that combine cloud motion vectors with statistical predictors. Although the specific performance benefits depend on local climate and sensor setup, remote sensing inputs can be particularly beneficial in ramp-prone conditions where NWP resolution is insufficient (Prema et al. 2022; Gaboitaolelwe et al. 2023).

Reanalysis products integrate historical observations with numerical models to provide consistent long-term atmospheric variables. They are especially relevant for medium-term forecasting, climatological baselines, and for transferring models across sites by providing standardized covariates when on-site measurements are unavailable. Public meteorological services can provide accessible predictors (e.g., cloud cover and temperature forecasts), though their spatial/temporal resolution and update cycles may limit applicability in high-frequency settings. These sources expand feasibility in data-scarce contexts but also impose an accuracy ceiling that must be acknowledged when comparing against studies using richer instrumentation (Gupta and Singh 2021; Prema et al. 2022).

General forecasting workflow

Although PV power forecasting methods vary widely, most operational pipelines follow a common workflow: (i) data acquisition and quality control, (ii) feature engineering (often including physically motivated transformations), (iii) model training and validation with time-series-appropriate protocols, and (iv) deployment with monitoring for drift and retraining triggers (Figure 5). Weaknesses in these stages can dominate performance and explain many irreproducible claims (Gupta and Singh 2021; Prema et al. 2022).

Data cleaning begins with timestamp harmonization across SCADA, meteo sensors, and external predictors, followed by detection of missingness, outliers, and operational anomalies (e.g., inverter trips, curtailment). A robust pipeline must specify whether nighttime intervals are included, how near-zero values are handled (critical for percentage-based measures),



Figure 5. Generic PV power forecasting workflow shared across physical, statistical, and learning-based approaches.

and whether imputed values are excluded from evaluation. Because high-frequency PV power is sensitive to alignment errors, especially under ramps, quality control should include checks for temporal offsets and clock drift. These practices are widely recognized in benchmarking-oriented PV literature, even when not consistently applied (Prema et al. 2022; Ledmaoui et al. 2023).

Feature engineering commonly includes diurnal/seasonal encodings (solar zenith/azimuth, day-of-year), irradiance transformations (clear-sky index, clearness index), lag features for power and meteorological variables, and aggregation features that summarize recent variability. Physically motivated transformations help reduce non-stationarity and improve transferability by expressing variables in normalized forms tied to solar geometry rather than to raw magnitudes. For instance, clear-sky or clearness-based normalization has been explored as a principled way to separate deterministic geometry from stochastic cloud effects (Lauret et al. 2022). When NWP is used, bias correction and downscaling can be treated as feature construction steps. Importantly, all transformations must be computed causally to avoid inadvertent use of future information.

Model training must respect the temporal dependence structure of PV data. Rolling-origin evaluation and blocked train-validation-test splits are preferred over random splits to prevent leakage and to reflect operational updating. Hyperparameter tuning should be performed on validation periods that are temporally disjoint from test periods, with special care to ensure that NWP inputs are aligned by issuance time (not by valid time retroactively). Deployment introduces additional constraints: latency, compute availability (edge vs cloud), interpretability requirements, and robustness under drift caused by degradation and changing operational policies. These constraints partially explain the persistence of hybrid and physically informed methods in practice, despite the success of high-capacity deep learning in curated academic datasets (Prema et al. 2022; Gaboitaolelwe et al. 2023).

Data quality, cleaning, and outlier handling

In practical photovoltaic power forecasting applications, data quality is often a more critical determinant of forecasting

performance than the choice of modeling architecture. Real-world PV datasets frequently contain missing values, sensor faults, timestamp misalignment, communication dropouts, curtailment artifacts, and extreme outliers caused by equipment failures or transient environmental effects. If not addressed explicitly, such issues can dominate model error, bias learning processes, and lead to misleading performance comparisons (Di Leo et al. 2025).

A robust forecasting pipeline, therefore, requires systematic data cleaning prior to feature engineering and model training. Common practices include time stamp harmonization across SCADA, meteorological measurements, and exogenous predictors; detection and treatment of missing values using physically consistent interpolation or exclusion rules; and filtering of non-operational periods such as night time intervals, prolonged outages, or flagged curtailment events (Hassan et al. 2025). Outlier handling is particularly important for PV power time series, where sudden spikes or drops may reflect sensor faults rather than true physical behavior. Such outliers can disproportionately influence error metrics and induce spurious learning if retained without scrutiny.

Importantly, outlier detection and cleaning should be guided by physical plausibility and operational context rather than purely statistical thresholds. For example, abrupt power drops coinciding with inverter alarms or grid curtailment signals should be treated differently from cloud-induced ramps, which represent legitimate and operationally relevant dynamics. Consequently, transparent documentation of data cleaning and outlier handling procedures is essential for reproducibility and for fair comparison of forecasting methods across studies (Xiang et al. 2025).

The provided considerations reinforce a broader lesson emphasized throughout the PV power forecasting literature: improvements attributed to increasingly complex models may, in practice, stem from differences in data preparation rather than intrinsic modeling superiority. As such, data quality control should be regarded as a foundational component of PV power forecasting systems, on par with model selection and evaluation design.

In addition to the periodic retraining of the forecasting model, methods that may provide a higher adaptation can be used to address the problem of data drift. Such adaptive approaches include the incorporation of online model learning and incremental learning. While the state-of-the-art models still largely include only offline training approaches (Piantadosi et al. 2024; Gong et al. 2025), some recent advances have been made, where methods of online learning and incremental learning are applied and the effects of data drift are diminished (Zhang et al. 2023). However, it should be mentioned that data drift is a complex problem, and other studies have indicated that recent data drift solutions have faced significant issues and challenges in application and the problem remains only partially addressed (Zhu et al. 2024). For this reason, the handling of data drift can still be considered an under-developed area of PV power forecasting and can be identified as one of the main directions for future research.

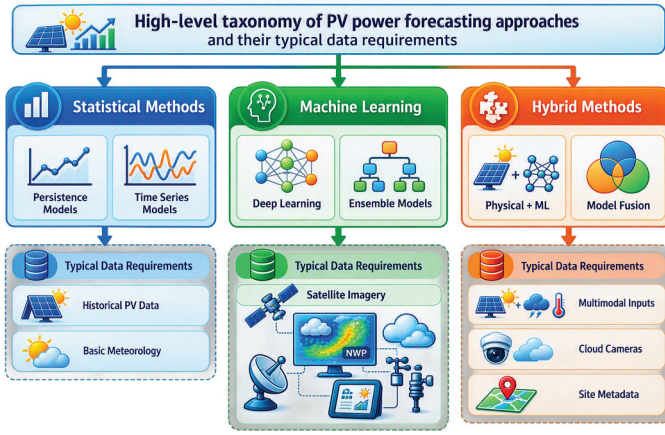


Figure 6. High-level taxonomy of PV power forecasting approaches and their typical data requirements.

Taxonomy of PV power forecasting methods

The taxonomy explicitly spans both point and probabilistic forecasting paradigms. While point forecasts remain useful for descriptive benchmarking, probabilistic forecasting has become increasingly important for operational deployment, particularly in intraday and day ahead settings where meteorological uncertainty dominates. Accordingly, uncertainty quantification is treated not as a specialization but as a cross-cutting capability that can be instantiated within physical, statistical, machine learning, and hybrid model families.

PV power forecasting methods can be organized into a small number of methodological families (Figure 6) that differ in (i) the degree of physical structure imposed on the forecast mapping, (ii) the statistical assumptions made about temporal dependence and noise, and (iii) the extent to which exogenous information (e.g., NWP, satellite imagery) is exploited. In practice, these families are not mutually exclusive: the strongest performance in the literature is frequently obtained through *hybridization* and *ensembling*, where complementary models are combined to reduce bias, variance, and sensitivity to non-stationary operating regimes. Hybrid models are described as the most accurate group overall, with improvements typically achieved by integrating additional methods into an already strong base learner such as LSTM or ARIMA, and occasionally by combining three components to further enhance accuracy.

Physical (physics-based) models

Physical PV power forecasting often begins from a radiative decomposition and transposition chain that converts meteorological drivers into plane-of-array (POA) irradiance, followed by a performance model that maps irradiance and temperature to electrical output. A common abstraction is to represent the physical chain as

$$\hat{P}_{AC}(t+h) = \eta_{inv}(t+h) f_{PV} 0pt(G_{POA}(t+h), T_{cell}(t+h), \ell(t+h)), \quad (8)$$

where G_{POA} is POA irradiance (often produced from clear-sky index reasoning), T_{cell} is the effective cell temperature, and ℓ groups loss factors (soiling, shading, wiring and mismatch

losses, clipping, curtailment flags). The key design choice is whether the atmospheric state is represented through a clear-sky index formulation or alternative clearness representations, which can materially affect downstream error behavior under variable cloud fields (Lauret et al. 2022). In operational settings, the upstream meteorological state is frequently supplied by NWP and then adjusted site-specifically to reduce systematic bias (Böök and Lindfors 2020).

Once irradiance is available, the performance layer translates radiative and thermal conditions into DC and then AC power. Model fidelity ranges from simplified efficiency curves to detailed loss accounting and inverter behaviors. Comparative evidence shows that the choice of physical-based model chain (including loss parameterization) can be as influential as the meteorological driver itself, particularly when moving from point forecasts to probabilistic settings that must represent uncertainty across the chain (Mayer and Gróf 2021; Mayer and Yang 2023). The physical approach is attractive because it embeds interpretability: errors can be traced to specific submodules (irradiance estimation, temperature model, loss terms), which supports diagnosis and maintenance planning.

The principal strength of physical-based modeling is *structural generalization*: when the mapping in (8) is well identified, extrapolation to new regimes can be more stable than purely data-driven approaches. However, the approach is limited by (i) imperfect meteorological forecasts, (ii) unobserved operational interventions (e.g., curtailment, inverter clipping), and (iii) site-specific losses that are difficult to parameterize without detailed metadata. Consequently, physical chains are increasingly used either as strong baselines or as components inside hybrid/ensemble pipelines (Mayer 2022).

Statistical time-series models

Statistical time-series models are based on statistical probability and stochastic processes. Since they rely on statistical models rather than iterative learning algorithms leads to treating statistical time-series models as a separate group. Statistical time-series models assume that dependencies and trends can be captured through mathematical relationships, which are used to generate predictions in the future. These methods more transparent and rely primarily on the target's own historical data, which makes them computationally efficient and competitive in short horizons (Gaboitoalelwe et al. 2023).

Statistical forecasting families begin with simple baselines such as persistence (e.g., $\hat{P}(t+h) = P(t)$) or variants that incorporate diurnal/seasonal structure. These baselines remain essential for benchmarking because many advanced methods report gains relative to weak comparators, a problem highlighted in broader reviews of data and metrics practices (Prema et al. 2022). In short horizons and stable conditions, persistence-type baselines can be surprisingly competitive.

Linear stochastic time-series models represent PV power (or irradiance proxies) through autoregressive and moving-average dependence with differencing to address non-stationarity:

$$\Phi(B)(1-B)^d P_t = \Theta(B)\epsilon_t, \quad (9)$$

where B is the backshift operator and ε_t is a noise process.

The literature includes PV-focused ARIMA variants and hybrid ARIMA combinations that couple statistical structure with nonlinear learners (Fan et al. 2022; Sapundzhi et al. 2024). The appeal of ARIMA-type models is their low computational cost and transparent residual diagnostics; their limitation is the difficulty of representing fast cloud-driven ramps without exogenous inputs.

State-space models and recursive filters (e.g., Kalman filtering) provide a principled way to combine a latent process model with noisy measurements, supporting online updates. While their classical formulations can be restrictive under nonlinear ramp dynamics, they are valuable either as standalone predictors for smoothing/nowcasting or as post-processing components for bias correction and uncertainty propagation, particularly when paired with meteorological predictors (Pauwels et al. 2013; Yang et al. 2021; Du et al. 2023; Kim et al. 2025).

Statistical methods are most competitive (i) at very short horizons, (ii) when data volume is limited, and (iii) when operational regimes are stable enough that linear dependence captures the dominant structure. In hybrid settings, ARIMA-like components are often used as base learners due to their robust fit and well-understood error behavior; your source material explicitly notes ARIMA (together with LSTM) as a typical foundation for building hybrids, with added components improving accuracy further (Mani and Volety 2021; Mohapatra et al. 2023; Babikou et al. 2025; Rostami et al. 2025).

In hybrid settings, ARIMA and other statistical models are used as base learners due to their robust fit and well-understood behavior. Studies in cross-domain fields such as wind speed forecasting, residential electricity forecasting, air pollution forecasting, and cyber-attack resilience explicitly note ARIMA (together with LSTM) as a typical foundation for building hybrids, with added components improving accuracy further. Within PV power forecasting, the explicit combination of ARIMA and LSTM hybrid models is lacking; however, hybrids that combine conceptually similar statistical or decomposition models (e.g., Theta, VMD, WPD, SDE) with LSTM are documented to outperform traditional models (AlKandari and Ahmad 2024; Li et al. 2020; Zhang and Kong 2022; Hou et al. 2024).

Machine learning models

Machine learning models rely forecasting on iterative learning algorithms rather than predefined stochastic process. They do not assume a specific parametric form temporal dependence but instead learn to map high-dimension features to the target output. While often “black box” in nature, they offer superior adaptability in fluctuating non-linear environments (Gaboitoalelwe et al. 2023).

Classical ML approaches treat forecasting as supervised regression, using engineered predictors such as lagged power, irradiance proxies, NWP features, calendar variables, and clear-sky indices. Support vector regression and kNN-style methods remain common in the literature, including hybrid kNN-SVM combinations (Saxena et al. 2024). Tree ensembles (gradient boosting, random forests) are often strong

competitors due to their ability to model nonlinear interactions and handle heterogeneous features (Ledmaoui et al. 2023). Their main limitation is sensitivity to distribution shift if training data do not cover rare regimes (snow cover, extreme aerosol loads, atypical curtailment).

Deep learning is typically used to reduce reliance on hand-crafted features by learning representations directly from sequences (RNN/LSTM/GRU), local patterns (CNN), or combined CNN-LSTM structures (Lim et al. 2022; Jailani et al. 2023). Sequence models are especially effective at capturing intra-day temporal dependence, while CNN components can extract short-range motifs and ramp precursors. Numerous studies propose hybrid deep architectures or add signal processing decompositions to stabilize learning under variability (Mellit et al. 2021).

Attention-based models provide a natural mechanism for multi-horizon forecasting, long-context conditioning, and feature selection through learned relevance weights. Their relevance for PV power forecasting is increasingly established, including day-ahead applications using Temporal Fusion Transformer architectures (López Santos et al. 2022). Transformers are attractive when exogenous covariates are high-dimensional and heterogeneous (multi-level NWP, calendar variables, site metadata) and when the evaluation emphasizes multiple horizons jointly.

When multiple PV plants or distributed systems are available, spatial context can reduce uncertainty, particularly for cloud-advection-driven ramps. Multi-site learning can be posed as spatiotemporal regression where the input includes neighboring-site measurements, satellite-derived cloud fields, or regional NWP grids. Satellite-informed ultra-short-term forecasting is a representative example of this direction, leveraging cloud information from neighboring regions to improve nowcasting performance (Wang F et al. 2022).

Robust ML pipelines require explicit handling of (i) missingness and sensor drift, (ii) seasonal distribution shifts, and (iii) operational interventions such as clipping and curtailment. For probabilistic outputs, calibration becomes central, and Bayesian or Gaussian-process-based components are frequently used for interval or distributional forecasts (Massidda et al. 2024). In expert practice, robustness is not an optional refinement but a prerequisite for credible deployment: a model that is marginally more accurate yet poorly calibrated can impose higher downstream costs in energy management.

Hybrid and ensemble approaches

Hybridization with physical-based models may be implemented via residual learning, where a physical forecast provides a structured baseline and ML learns to correct systematic bias:

$$\hat{P}(t+h) = \hat{P}_{\text{phys}}(t+h) + \hat{r}(t+h), \quad \hat{r} = \mathcal{M}_{\theta}(x_t), \quad (10)$$

with $\mathcal{M}_{\theta}$ a learned model using recent measurements and meteorological predictors. The motivation is that the physical component captures first-order radiative/thermal constraints, while ML explains unmodeled losses and site-specific behaviors. Evidence for the value of physical-ML hybridization is explicitly discussed in the literature, including systematic

arguments for why such pairing improves performance and robustness (Mayer 2022).

Statistical–ML hybrids combine linear temporal structure with nonlinear learners to capture residual dependencies (e.g., ARIMA + neural networks). A representative example is the ARIMA–BPNN–SVR hybrid formulation for PV power forecasting (Fan et al. 2022). Strong statistical or ML base models are frequently augmented by additional techniques to improve accuracy, and that in some cases a third component yields further gains.

Ensembles reduce variance and mitigate model-selection risk, which is especially important under regime changes (cloud transients, seasonal transitions). Sequential ensembles have been proposed specifically for PV power forecasting (Sharma et al. 2021). In probabilistic contexts, ensembles naturally support uncertainty quantification through member spread; pairing ensemble NWP with an ensemble physical chain is an example of end-to-end probabilistic construction (Mayer and Yang 2023).

A particularly important subclass of hybrid approaches is formed by stacking-based ensemble frameworks, in which multiple heterogeneous base learners are combined through a second level meta learner. Unlike simple averaging or voting schemes, stacking explicitly learns how to weight and combine model outputs, enabling it to exploit complementary strengths of different predictors. Recent studies demonstrate that stacking-based frameworks can achieve strong generalization performance in ultra short-term photovoltaic power forecasting, especially when applied across PV systems with different module types or deployment locations, where single models often fail to transfer reliably across domains (Cao et al. 2025).

Beyond basic stacking, several works have shown that the performance and robustness of ensemble models can be further improved through optimized model selection and feature prioritization strategies. For example, optimization-driven stacking frameworks employ meta-heuristic or evolutionary algorithms to adaptively select base models and meta learners, as well as to prioritize informative features, thereby improving stability under non-stationary operating conditions (Cao et al. 2023).

From a system level perspective, stacking based hybrid models can be interpreted as a principled response to the inherent heterogeneity and regime variability of energy time series data. In PV power forecasting, such ensembles provide a flexible mechanism for integrating physical insights, statistical baselines, and data driven predictors within a unified framework, improving both accuracy and generalization when deployment conditions differ from the training environment.

Probabilistic forecasting extends point prediction into uncertainty-aware representations (quantiles, prediction intervals, or full distributions). Day-ahead probabilistic PV prediction is commonly evaluated via proper scoring rules and reliability diagnostics, and comparative evidence indicates that methodology choice, data handling, and feature design can dominate reported gains (Prema et al. 2022; Massidda et al. 2024). In practice, probabilistic extensions can be added via direct quantile learning, distributional models, or post-hoc calibration (e.g., conformal prediction) depending on operational needs.

The methods for PV power forecasting are summarized as a taxonomy in Table 6.

Hybrid forecasting methodologies can be compared according to different fusion strategies at the data, feature, and decision levels. At the data level, fusion is performed by combining raw data with certain variables before model training. At the feature level, this process is conducted by fusing extracted features from data. Finally, at the decision level, fusion is typically done aggregating the predictions from several separating predictors. Recent literature (Di Leo et al. 2025; Gong et al. 2025) indicates that the choice between the three strategies depends on various factors, such as non-stationarity of the signal and computational constraints. Physical-ANN and wavelet-based hybrid models (Di Leo et al. 2025) [can be particularly suitable when interpretability or robustness to non-stationary behavior is important, meanwhile parallel architectures such as TimesNet-BiLSTM (Gong et al. 2025) can be more advantageous than multiple temporal dependencies need to be captured at the same time. Similarly, ensemble-level fusion is preferable when robustness is prioritized over architectural simplicity and computational costs. Table 5 overviews the different strategies.

Exogenous-input and multi-modal approaches

NWP-driven pipelines treat meteorological forecasts as primary predictors and then map them to PV output via physical, statistical, or ML layers. Site-specific bias correction is frequently required to reach operational accuracy (Böök and Lindfors 2020). NWP is especially important beyond very short horizons, where purely autoregressive signals decay and atmospheric evolution dominates.

For ultra-short-term horizons, satellite imagery and sky imagers provide direct observation of cloud fields, which improves ramp prediction. Methods that explicitly extract cloud information, including neighboring-plant context, demonstrate the value of spatially coherent predictors for now-casting (Wang F et al. 2022).

Table 5. Comparison of different fusion strategies (Di Leo et al. 2025; Gong et al. 2025).

Fusion level	Typical mechanisms	Suitable contexts	Advantages	Limitations
Data level	Physical inputs, decomposition, exogenous-data integration	Noisy/non-stationary PV series; strong physics knowledge	Better signal conditioning; interpretable preprocessing	Added preprocessing complexity
Feature level	Parallel or sequential feature extractors (e.g., TimesNet + BiLSTM)	Complex temporal dynamics; ultra-short-term forecasting	Captures complementary patterns within one architecture	Higher design/tuning burden
Decision level	Stacking, boosting, weighted aggregation	Need for robustness across regimes/sites	Strong generalization and modularity	Extra ensemble calibration needed

Table 6. Summary taxonomy of PV power forecasting approaches: typical inputs, strengths, and limitations.

Family	Typical inputs	Strengths	Limitations / risks
Physical chains	NWP irradiance/temperature; plant metadata; loss parameters	Interpretability; physics-consistent extrapolation; diagnostic value	Sensitive to meteorological bias and unobserved losses; parameter identification burden
Statistical time-series	Lagged power/irradiance; calendar seasonality	Fast, transparent baselines; stable short-horizon performance	Weak for ramp events without exogenous drivers; limited nonlinearity
Classical ML	Engineered features (lags, NWP, indices, calendar)	Nonlinear interactions; strong tabular performance	Shift sensitivity; feature leakage risks; calibration often neglected
Deep learning / Transformers	Sequences + high-dimensional covariates; multi-horizon targets	Representation learning; multi-horizon modeling; feature selection via attention	Data-hungry; brittleness under shift; computational cost
Hybrid / ensemble	Any of the above combined; residuals, stacking, averaging	Empirically strongest accuracy; improved robustness; uncertainty via spread	Complexity, maintainability; harder attribution of errors

Fusion strategies range from early fusion (concatenating all predictors) to late fusion (combining model outputs) and hierarchical fusion (e.g., NWP for trend + imagery for ramps + SCADA for immediate persistence). The practical motivation is that different modalities dominate at different time scales: SCADA/persistence dominates minutes ahead, imagery supports tens of minutes, and NWP dominates hours to day-ahead.

Multi-modal forecasting is only as credible as its synchronization. Temporal misalignment between SCADA clocks, NWP issuance times, and imagery timestamps can create spurious improvements if not handled carefully; therefore, rigorous alignment and leakage control should be treated as part of the method rather than a preprocessing detail (Prema et al. 2022).

Method–context fit: a selection map

The diversity of PV power forecasting methods surveyed in Section 5 can appear unwieldy unless anchored to the operational context in which forecasts are produced and consumed. In practice, “the best model” is not a universal entity: model suitability is conditional on (i) forecast horizon and update frequency, which determine the dominant uncertainty sources; (ii) data availability and fidelity, which constrain identifiability; and (iii) deployment constraints, which impose latency, compute, and interpretability requirements. Recent reviews repeatedly stress that performance claims are frequently non-transferable across sites precisely because these context variables are implicitly changed between studies (Gupta and Singh 2021; Prema et al. 2022; Gaboitaolelwe et al. 2023).

Edge-oriented PV power forecasting systems designed to execute PV power forecasting in real time impose computational latency or hardware limitations on the forecasting model (Venitourakis et al. 2023). Recent reviews show that more complex state of the art deep and hybrid PV power forecasting models require significant computational resources. Greater accuracy comes with an increased computational time trade-off (Gaboitaolelwe et al. 2023; Asghar et al. 2024). Complex models deliver higher accuracy, but prediction time is significantly longer (Dhingra et al. 2025; Guo et al. 2025).

A convenient decision-theoretic view is to cast method selection as minimizing an expected operational loss under feasibility constraints:

$$\begin{aligned}
 m^* &= \arg \min_{m \in \mathcal{M}} \mathbb{E} \left[\mathcal{L} \left(y_{t+h}, \hat{y}_{t+h|m}^{(m)} \right) | \mathcal{I}_t \right] s.t. \\
 \text{latency}(m) &\leq \ell_{\max}, \text{cost}(m) \leq c_{\max}, \text{auditability}(m) \geq a_{\min},
 \end{aligned} \quad (11)$$

where $\mathcal{M}$ is the set of candidate methods (or pipelines), $\mathcal{L}$ encodes the application-relevant error cost (often asymmetric in grid operations), and the constraints capture deployment realities (e.g., computational limits on edge devices, explainability requirements, and monitoring overhead). Eq. (11) emphasizes a critical point: “accuracy” is only one term in the selection problem; operational fitness is multi-criteria by construction, which explains why physically structured and hybrid approaches remain prevalent in real deployments even when purely data-driven models dominate benchmark leaderboards (Böök and Lindfors 2020; Prema et al. 2022).

A standardized approach to evaluation protocols is needed to ensure that PV power forecasting results are comparable and meaningful across different studies. Using multiple, complementary metrics is now common practice to capture average error, sensitivity to outliers, and model fit. Existing literature supports reporting at least MAE, RMSE (or MSE), a relative/normalized metric (e.g. MAPE, NRMSE, CVRMSE), and R2 (Husein et al. 2024). Different metrics capture different aspects of forecast quality. In some scenarios it is beneficial to compliment these metrics execution time to capture computational cost (Mystakidis et al. 2024). This combination provides more complete view of forecast performance. Dataset splitting strategies define how data are divided into training, validation, and test sets. PV power forecasting uses time-series data, so splitting must respect temporal order and seasonal patterns. Several main strategies that can be applied to datasets are: (i) chronological train-test splits, (ii) cross-validation or k-fold approaches, (iii) time-series partitioning strategies (window, seasonal, vertical) for ensembles, and (iv) cluster- or context-based splits (Chuluunsai Khan et al. 2022; Husein et al. 2024).

Selecting methods by forecast horizon

Forecast horizon governs what information is available and which uncertainties dominate. Across the literature, a consistent pattern emerges: very short horizons are dominated by local temporal continuity and cloud motion; intraday horizons exhibit mixed dynamics and benefit from hybridization; day-ahead horizons are largely determined by NWP skill and uncertainty (Böök and Lindfors 2020; Gupta and Singh 2021; Prema et al. 2022). The horizon, therefore, acts as the primary axis of method selection, with data modality and deployment constraints acting as secondary filters (Figure 7).

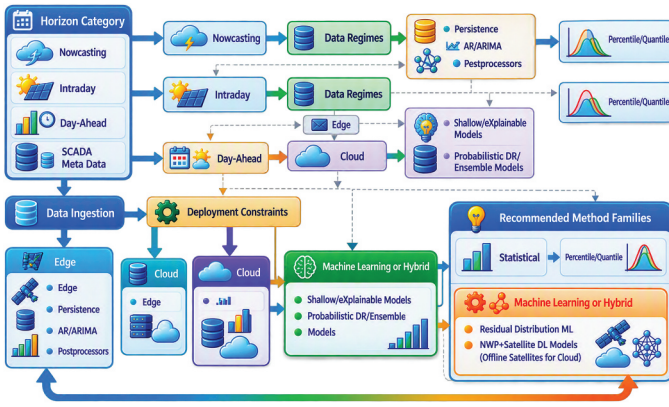


Figure 7. Method-context selection map for PV power forecasting, emphasizing horizon, data regime, and deployment constraints.

Nowcasting (seconds to tens of minutes) is characterized by high-frequency variability driven primarily by cloud advection and cloud-edge effects. At these horizons, the information set $\mathcal{I}_t$ is dominated by the most recent PV measurements and any real-time sky observations. Simple persistence baselines often provide strong performance under stable conditions; however, they fail during rapid ramps, which are precisely the periods of highest operational relevance. Consequently, the most context-appropriate approaches tend to be (i) persistence variants augmented with short-window dynamics (e.g., ramp-aware features), and (ii) imaging-driven methods that explicitly track cloud motion via satellite or sky-imager data (Prema et al. 2022; Wang F et al. 2022; Gaboitaolelwe et al. 2023). In practice, imaging-based nowcasting can be viewed as learning a latent advection operator that maps observed cloud fields C_t to near-future irradiance (and thus power):

$$\hat{y}_{t+h|t} = g(C_t, C_{t-\Delta}, \dots), \quad (12)$$

where $g(\cdot)$ may be implemented by optical-flow pipelines, CNNs, or spatiotemporal networks. These approaches are most valuable when the plant experiences frequent ramp events and when imaging timestamps can be reliably synchronized with SCADA measurements – a nontrivial practical requirement that often determines whether their benchmarked gains translate to deployment (Prema et al. 2022).

Intraday horizons (tens of minutes to several hours) exhibit mixed uncertainty: short-memory dependence remains informative, but atmospheric evolution increasingly matters, and the value of exogenous predictors rises. Empirically, hybrid approaches become attractive in this regime because they can combine the stability of structured baselines (persistence/statistical or physically informed) with nonlinear correction layers that learn site-specific residual patterns (Mayer 2022; Prema et al. 2022). A common operationally robust design is residual hybridization:

$$\hat{y}_{t+h|t} = \hat{y}_{t+h|t}^{\text{base}} + r_{\theta}(x_t), \quad (13)$$

where $\hat{y}_{t+h|t}^{\text{base}}$ may be an autoregressive/statistical forecast or a simplified physical chain driven by short-range NWP, and $r_{\theta}(\cdot)$ is a learned residual (e.g., gradient-boosted trees or LSTM/GRU) trained to correct systematic biases and to adapt to local plant behavior. Comparative and survey studies

report that such hybridization tends to be particularly effective under regime variability, because the baseline anchors predictions in physically or statistically consistent structure, while the residual learner concentrates capacity on the non-stationary or unmodeled components (Gupta and Singh 2021; Gaboitaolelwe et al. 2023).

Day-ahead forecasting (typically 24–48 hr) is fundamentally constrained by the accuracy and resolution of NWP products. At these horizons, purely autoregressive strategies are generally insufficient, and the dominant modeling question shifts to how best to translate NWP variables (irradiance components, cloud proxies, temperature, wind, humidity) into PV output while correcting for site-specific bias and representativeness error (Böök and Lindfors 2020; Prema et al. 2022). This motivates two broad classes of pipelines. The first employs physical conversion chains, optionally with bias correction and parameter calibration, and is favored when interpretability and stability under shift are priorities. The second leverages flexible machine learning or deep learning models that learn the NWP $\rightarrow$ power mapping directly, often with multi-horizon outputs; transformer-based architectures have been applied in this setting to fuse heterogeneous covariates and capture long-range temporal dependencies (López Santos et al. 2022; Gaboitaolelwe et al. 2023).

Importantly, day-ahead decisions are risk-sensitive, and the literature increasingly advocates probabilistic outputs rather than single-valued predictions. Using an ensemble NWP $\{w_{t+h|t}^{(s)}\}_{s=1}^S$, one may propagate uncertainty through either (i) a physical chain, or (ii) a learned mapping, producing a predictive distribution:

$$\hat{F}_{t+h|t}(y) \approx \mathbb{1S} \sum_{s=1}^S \mathbb{I} \left(f \left(w_{t+h|t}^{(s)} \right) \leq y \right), \quad (14)$$

which supports quantiles and scenario-based bidding strategies (Mayer and Yang 2023; Massidda et al. 2024). In operational practice, the selection between physics-based and purely data-driven NWP translation often depends less on headline error metrics than on governance requirements (auditability) and the cost of maintaining model calibration across changing NWP systems and plant conditions (Böök and Lindfors 2020; Prema et al. 2022).

Selecting methods by data availability

Data availability is a second decisive axis because it defines the feasible hypothesis class. A method that is optimal in a rich multi-modal setting can be infeasible (or unstable) in a low-data regime, where identifiability is limited and overfitting risk is high. Recent reviews emphasize that many contradictory findings in the PV power forecasting literature can be reconciled by recognizing that studies implicitly differ in available telemetry, meteorological predictors, and data volume (Gupta and Singh 2021; Iheanetu 2022).

In low-data settings (limited history, sparse sensors, or incomplete covariates), parsimonious models with strong inductive bias are typically preferable. This includes persistence variants, seasonal climatology baselines, simple linear regression with physically meaningful features, and compact

statistical time-series models (e.g., ARIMA with seasonal terms) when stationarity assumptions are not grossly violated (Prema et al. 2022; Ledmaoui et al. 2023). Physical-based models can also be effective if plant metadata and at least coarse meteorological drivers are available, because they reduce the need for extensive training data by encoding the conversion structure. In such regimes, the selection goal is often not to maximize marginal accuracy, but to obtain *stable* performance under limited evidence and to produce forecasts that remain interpretable for operational stakeholders.

In rich-data regimes (long historical SCADA streams, reliable NWP, local meteo, and possibly imaging), higher-capacity learners can exploit nonlinear interactions and multi-source context. Deep learning architectures (CNN/RNN hybrids and attention-based models) become attractive when the covariate space is high-dimensional and when multi-horizon forecasts are required (López Santos et al. 2022; Gaboitoalelwe et al. 2023). Multi-modal fusion (SCADA + NWP + satellite/sky-imager) is particularly valuable because each modality dominates a different time scale; the overall pipeline can be designed to weight information sources according to horizon. However, multi-modal settings also magnify synchronization and leakage risks, so practical success depends on disciplined alignment (issue-time consistency for NWP, timestamp harmonization for imagery and SCADA) and careful evaluation protocols (Prema et al. 2022).

Selecting methods by deployment constraints

Even when a model is technically feasible and accurate, deployment constraints can render it unsuitable. Three constraints recur across PV applications: (i) computation and latency limits (edge vs cloud), (ii) transparency requirements, and (iii) lifecycle maintenance under drift. These constraints are increasingly foregrounded in recent studies because they govern whether forecasting systems can be sustained beyond research prototypes (Gupta and Singh 2021; Prema et al. 2022). The guidance for the selection of models is summarized in Table 7 and explained below.

Nowcasting and fast intraday control often impose strict latency constraints, favoring models with small inference footprints and predictable runtime (El Hendouzi and Bourouhou 2020; Theocharides et al. 2021). On the edge, this typically prioritizes compact statistical models, lightweight tree ensembles, or small recurrent networks (Theocharides et al. 2021). In contrast, cloud deployments can support transformer-based multi-horizon models and large-scale ensemble pipelines, which may be computationally intensive but can deliver superior performance when fed with multi-source predictors (de Oliveira et al. 2023). A pragmatic guideline is to match

architectural complexity to the decision cycle: models used in sub-minute control loops should be interpretable and computationally stable; high-capacity models are better reserved for hourly or day-ahead planning where inference latency is less critical (Wan et al. 2015; El Hendouzi and Bourouhou 2020).

Grid-facing decisions and market participation increasingly require transparency regarding forecast provenance and failure modes (Bessa et al. 2014; Zhu et al. 2025). Physical and hybrid models provide a natural decomposition of errors into interpretable subcomponents (irradiance estimation, temperature mapping, loss modeling), which supports audits and root-cause analysis (Wan et al. 2015; de Oliveira et al. 2023). Purely data-driven deep models can be deployed successfully, but in regulated or safety-critical settings they typically require additional explainability measures, calibration monitoring, and documentation of input dependencies (Theocharides et al. 2021). For many operators, this governance burden can outweigh small improvements in average error, motivating hybrid designs where a physical baseline provides an auditable anchor and ML contributes controlled residual corrections (Mayer 2022; Prema et al. 2022).

PV plants evolve: module degradation, soiling dynamics, component replacements, changes in curtailment policies, and even changes in the upstream meteorological model (NWP upgrades) can induce concept drift. Consequently, operational forecasting systems require monitoring and retraining triggers. A standard approach is to track residual statistics and calibration metrics over time; drift can be declared when error distributions change significantly relative to a reference window:

$$\text{Drift flag} \leftarrow D0pt\left(\{\varepsilon_t\}_{t \in W_{\text{recent}}}, \{\varepsilon_t\}_{t \in W_{\text{ref}}}\right) > \delta, \quad (15)$$

where $\varepsilon_t = y_t - \hat{y}_t$ and $D(\cdot, \cdot)$ is a divergence or two-sample statistic. Hybrid and ensemble methods can be advantageous here because they reduce reliance on any single model's stability; moreover, they allow targeted recalibration (e.g., updating the residual learner while keeping the physical baseline fixed) rather than costly full retraining. Probabilistic pipelines further support drift detection through reliability monitoring (e.g., interval coverage), which can reveal deterioration before point metrics materially worsen (Prema et al. 2022; Massidda et al. 2024).

Open challenges and research directions

Although the methodological landscape of PV power forecasting has expanded rapidly – from physically structured conversion chains to attention-driven deep models and multi-modal hybrids – the field continues to face persistent challenges that limit the transferability, robustness, and operational credibility

Table 7. Practical selection guide: recommended method families by horizon and data regime (indicative, not exhaustive).

Context	Low-data regime	Moderate-data regime	Rich multi-modal regime
Nowcasting (min)	Persistence variants; simple autoregression; robust smoothing	Statistical + residual ML; short-window RNN/GRU	Sky-imager/satellite nowcasting; spatiotemporal CNN/attention (Wang F et al. 2022)
Intraday (hours)	Seasonal baselines; parsimonious ML; calibrated physical baseline	Hybrid residual learning (base + correction) (Mayer 2022)	Multi-source fusion (SCADA+NWP+imagery); ensemble/hybrid pipelines (Prema et al. 2022)
Day-ahead (24–48 hr)	Physical chain with coarse weather; bias correction	NWP → power ML with regularization; calibrated physical chain	Deep multi-horizon models (e.g., TFT); probabilistic ensembles (López Santos et al. 2022; Massidda et al. 2024)

of published results. Recent surveys converge on the view that the remaining bottlenecks are less about incremental model complexity and more about *non-stationary reality*: PV plants age, maintenance policies change, upstream meteorological models are updated, and climate-driven variability alters the distribution of conditions under which forecasts must operate (Gupta and Singh 2021; Prema et al. 2022; Gaboitaolelwe et al. 2023). These factors imply that the next phase of research must prioritize robustness, calibration, interpretability, and reproducibility alongside point accuracy, and must treat forecasting systems as continuously operating cyberphysical services rather than static predictors.

Non-stationarity, climate variability, domain adaptation

A defining characteristic of operational PV power forecasting is that the data-generating process evolves over time. At the plant level, slow drifts arise from module degradation, gradual soiling accumulation and cleaning cycles, sensor recalibration, and changes in inverter operating envelopes. At the system level, policy-driven curtailment patterns and grid constraints can change abruptly. At the atmospheric level, interannual climate variability alters the prevalence and structure of cloud regimes, extreme events, and seasonal transitions. Consequently, the conditional mapping $\mathbb{P}(y_{t+h}|\mathcal{I}_t)$ is not stationary, and models trained on historical data are often asked to extrapolate beyond their effective support. This motivates domain adaptation as a central research direction: rather than assuming a fixed model, forecasting pipelines should incorporate mechanisms for online calibration, transfer learning, and drift-aware retraining schedules. Reviews emphasize that the failure to account for distribution shift is a frequent explanation for degraded real-world performance compared with controlled benchmark results (Gupta and Singh 2021; Prema et al. 2022). Methodologically, promising directions include (i) explicit drift detection and gated model updates, (ii) adaptation layers that correct systematic bias in NWP-driven pipelines, and (iii) hybrid designs where a physics-based backbone provides a stable inductive bias while data-driven components adapt to shifting residual patterns (Böök and Lindfors 2020; Mayer 2022).

Generalization across sites and climates

A second open challenge is the limited generalization of many models across sites and climates. Site-specific predictors can deliver excellent accuracy when trained and evaluated on a single plant, yet their performance often degrades when transferred to different geographical contexts with distinct cloud statistics, aerosol loads, temperature ranges, and horizon obstructions. This is partly a data issue: many studies rely on narrow datasets and do not evaluate cross-site transfer explicitly. It is also a modeling issue: plant heterogeneity (tilt/tracking, DC/AC sizing, shading, curtailment policies) induces systematic differences in the mapping from meteorology to exported power. Recent work on regional and distributed PV power forecasting illustrates that spatial context can improve robustness, yet it also increases the burden of harmonizing

telemetry and aligning meteorological inputs across heterogeneous installations (Jung et al. 2022; Wang F et al. 2022). A research priority is to develop *transferable representations* that can accommodate site metadata and climate regimes, and to establish evaluation protocols that explicitly test generalization across geographically diverse fleets rather than reporting only within-site skill (Prema et al. 2022; Gaboitaolelwe et al. 2023).

Probabilistic forecasting, calibration, and decision support

The third direction concerns the transition from point prediction to decision-relevant uncertainty quantification. Operational planning and market participation are inherently risk-sensitive; thus, a forecast that provides uncertainty estimates can be more valuable than one that achieves marginally lower mean absolute error. The literature increasingly recognizes that probabilistic PV power forecasting is essential for day-ahead contexts where meteorological uncertainty dominates, and for storage scheduling and reserve planning where asymmetric costs penalize under- and over-estimation differently (Massidda et al. 2024). However, producing uncertainty is not sufficient: the predictive distributions must be *calibrated* and *sharp* to support actionable decisions. Bayesian approaches, copula-based constructions, and ensemble pipelines provide viable routes to probabilistic output, yet they raise open questions about how to maintain calibration under drift, how to propagate uncertainty coherently through physical conversion chains, and how to evaluate forecasts against operational objectives rather than purely statistical scores (Panamtash et al. 2020; Wang K et al. 2022; Mayer and Yang 2023). A key research need is to integrate forecasting with downstream optimization in a decision-centric manner, closing the gap between forecast skill and operational value.

Explainable PV forecasting for operational trust

As PV power forecasting becomes embedded in critical infrastructure workflows, explainability and auditability become central to trust and adoption. Operators require clarity on why forecasts change, which inputs drive predictions, and which failure modes are expected under specific regimes (e.g., broken-cloud ramps, clipping, curtailment). Physical-based models are naturally interpretable, but their accuracy can be limited by imperfect loss parameterization and meteorological drivers; high-capacity deep learning models can improve accuracy but often behave as opaque function approximators. This creates a demand for explainable forecasting systems that combine strong inductive bias with transparent attribution mechanisms (Hassan et al. 2024; Petrosian and Zhang 2024; Rizk-Allah et al. 2024; Esen et al. 2025; Öter and Ersöz 2025; Siddiqi et al. 2025; Tajjour et al. 2025). Recent work emphasizes explainable machine learning for solar-related prediction tasks, suggesting that interpretability can be treated as a design goal rather than a post-hoc add-on (Rajagukguk and Lee 2025). For PV power forecasting specifically, research should prioritize (i) structured hybrid models with interpretable baselines, (ii) robust feature attribution

under collinearity and distribution shift, and (iii) operational dashboards that connect forecast errors to physical causes (meteorological vs equipment vs grid actions), thereby supporting accountability and continuous improvement (Prema et al. 2022).

The outcomes of the Explainable AI method in PV power forecasting could be transformed from feature analysis into active signals for calibrating parameters of physical-based models. It is shown that Explainable AI methods such as SHAP or LIME reveal which inputs drive forecast errors and how their effects deviate from expected results [?], [?]. Once these discrepancies are exposed, they can guide targeted re-tuning of parameters in the underlying models (Mitrentsis and Lens 2022; Gomes et al. 2025).

Benchmarking gaps and reproducibility

A persistent barrier to progress is the fragmentation of benchmarking practices. Studies vary widely in target definitions (AC vs DC, power vs energy), temporal aggregation, filtering of night hours, handling of curtailment/clipping regimes, baseline strength, and metric selection. These differences can dominate reported performance, creating an illusion of methodological superiority that does not hold under standardized protocols. Multiple studies call attention to the need for stronger reproducibility: transparent reporting of datasets, preprocessing, horizon conventions, input freshness (especially for NWP), and uncertainty evaluation (Gupta and Singh 2021; Prema et al. 2022; Ledmaoui et al. 2023).

Another important research direction is the systematic evaluation of forecasting methods under challenging operating conditions. While most existing PV power forecasting studies report average performance over complete datasets, operational reliability often depends on model behavior during difficult scenarios such as rapid cloud-induced ramps, extreme weather events, curtailment periods, sensor faults, missing-data episodes, and distribution shifts across seasons or locations. Recent research in power-system optimization has demonstrated the value of deliberately constructing “hard” benchmark instances to expose algorithmic weaknesses and evaluate robustness under adverse conditions. Inspired by this concept described in (Cao et al. 2026), future PV power forecasting benchmark datasets could explicitly include challenging forecasting scenarios and rare events. We believe that such benchmarks would enable a more realistic assessment of model robustness, reliability, and operational readiness than conventional average-case evaluations alone.

Emerging trends: foundation models, self-supervised learning, and digital twins

Finally, PV power forecasting is beginning to align with broader trends in machine learning toward foundation models, self-supervised representation learning, and digital-twin-enabled monitoring. Foundation-style approaches are attractive because PV power forecasting is often data-fragmented at individual sites but data-abundant across fleets; a shared pretrained representation could improve transferability and reduce labeled data

requirements. Early evidence in solar-energy contexts suggests that large-model paradigms can be adapted to distributed PV power forecasting and can learn generalizable patterns from heterogeneous sources (Fan et al. 2025). Self-supervised learning is similarly promising for extracting structure from unlabeled streams such as irradiance time series, NWP grids, and satellite sequences, potentially improving robustness under domain shift. In parallel, digital twin concepts – where a continuously updated virtual model of the PV plant integrates physics, telemetry, and learned predictors – offer a pathway to close the loop between forecasting, diagnostics, and maintenance actions. Within such a framework, forecasting ceases to be a standalone task and becomes a component of an adaptive system that tracks plant health, detects anomalies, and updates its predictive mapping as the physical asset evolves. Achieving this vision requires advances in data integration, uncertainty propagation, and interpretable hybrid modeling, but it also offers a coherent direction for moving beyond one-off predictors toward resilient PV power forecasting services (Prema et al. 2022; Gaboitoalelwe et al. 2023).

A particularly recent development within this emerging landscape is the application of foundation models and large language models (LLMs) to energy and photovoltaic power forecasting (Aissaoui and Abdelmalek 2026). Motivated by large-scale self-supervised pretraining, time series foundation models aim to learn transferable representations from diverse energy datasets and have demonstrated promising zero shot and few shot forecasting capabilities. Early studies have explored LLM-based architectures for renewable generation forecasting, including photovoltaic power, highlighting potential advantages in multi-horizon learning, cross-site generalization, and unified modeling across heterogeneous data regimes (Bahi et al. 2026). At the same time, these approaches introduce new challenges related to computational cost, data governance, calibration under distribution shift, and the integration of physical constraints, indicating that foundation-model-based forecasting currently complements rather than replaces established physical, statistical, and hybrid methods.

While the application of Large Language Models in PV power forecasting can be viewed as a cutting-edge research direction, there are several issues and mechanisms that should be considered. Firstly, it should be mentioned that LLMs cannot directly operate on time-series data, therefore, additional data preprocessing is required (Mirshekali et al. 2025). LLM data preprocessing is typically conducted by cleaning the data and performing the conversion from time-series format to prompt-based inputs or tokenized inputs (Mirshekali et al. 2025). Additionally, there are certain issues concerning the high computational costs of handling high-dimensional time-series data with LLMs (Mirshekali et al. 2025; Zhang et al. 2025). While the problem of computational efficiency remains a critical part of research in LLM application for time-series forecasting, several solutions have been proposed, such as optimizing LLM architecture for forecasting and incorporating parameter-efficient and lightweight models (Mirshekali et al. 2025; Zhang et al. 2025).

At present, LLM-based forecasting approaches are generally not the most suitable choice for strict real-time photovoltaic forecasting applications, where low latency, computational

efficiency, and rapid model inference are still key requirements. Although recent advances in lightweight architectures, parameter-efficient fine-tuning, and model compression have improved their practical feasibility, LLMs are currently better suited to day-ahead, multi-horizon, and cross-site forecasting tasks rather than latency-sensitive real-time operational deployment. Future developments in efficient LLM implementations are, however, already appearing, and we firmly believe there is some potential to further expand their applicability to real-time forecasting scenarios.

Conclusion

This paper synthesized PV power forecasting from a *method-context* perspective, emphasizing that model choice is inseparable from the operational environment in which forecasts are produced and consumed. We first grounded the forecasting problem in the physical structure of PV plants and their measurement points, highlighting how the energy conversion chain and operational controls shape the observable signal and, consequently, the feasible hypothesis space for forecasting models. We then formalized the core forecasting tasks and horizons and organized the methodological landscape into a taxonomy spanning physics-based conversion chains, statistical time-series methods, machine learning and deep learning architectures, and hybrid/ensemble approaches. Across these families, a consistent theme emerging from recent surveys is that methodological performance is highly conditional on horizon, data modality, and plant/site characteristics; therefore, reported superiority in one setting does not necessarily translate to another.

A second key insight is that operationally robust performance is often achieved not by a single model class but by *structured combinations* of complementary components. Physical baselines provide interpretability and stable inductive bias, statistical models offer strong short-horizon baselines and transparent diagnostics, and learning-based components capture nonlinear interactions and site-specific residual structure. Hybridization and ensemble design can, therefore, be interpreted as principled responses to the multi-regime nature of PV generation, where different uncertainty sources dominate at different lead times and under different atmospheric conditions. Finally, the discussion of open challenges underscored that the remaining barriers to reliable deployment are increasingly tied to non-stationarity, cross-site generalization, and uncertainty-aware decision support, rather than to the absence of yet another model architecture.

For researchers, the analysis suggests that meaningful contributions should explicitly specify the *context envelope* of the proposed method – forecast horizon, input availability and freshness (especially for NWP), target definition (AC vs DC; power vs energy), and regime handling (clipping/curtailment/outages) – because these factors strongly govern observed performance and determine comparability with prior work (Ledmaoui et al. 2023). Methodologically, research designs that incorporate strong, context-appropriate baselines, multi-horizon evaluation, and regime-stratified reporting are more likely to yield results that generalize beyond a single dataset. In

addition, there is a clear opportunity for work that integrates forecasting with operational objectives, moving from point-accuracy improvements to calibrated probabilistic outputs and decision-relevant evaluations that reflect the asymmetric costs of under- and over-forecasting in energy management.

For practitioners, forecasting should be treated as a *system* rather than a model: a pipeline that ingests and quality-controls telemetry, aligns exogenous predictors, selects a method family appropriate to the horizon and data regime, and monitors performance under drift. In many operational settings, physically informed baselines or hybrid residual designs can provide a favorable balance between accuracy, transparency, and maintainability, particularly when regulatory scrutiny or grid-facing decisions require auditability. Where rich multi-modal data are available and compute constraints permit, attention-driven deep models and ensemble strategies can deliver additional gains, but they should be deployed with explicit monitoring and periodic recalibration to mitigate degradation due to plant aging, changing curtailment policies, or upstream NWP model updates.

Author contributions

CRedit: **Jokūbas Jonuška**: Data curation, Investigation, Resources, Writing – original draft; **Robertas Damaševičius**: Methodology, Validation, Visualization, Writing – review & editing; **Diana Belova-Plonienė**: Formal analysis, Supervision, Validation; **Andrius Katkevičius**: Formal analysis, Investigation, Validation; **Darius Plonis**: Conceptualization, Supervision, Validation; **Vaiva Rusen**: Data curation, Formal analysis, Investigation; **Tomyslav Sledevič**: Formal analysis; **Rytis Maskeliūnas**: Formal analysis, Validation.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Robertas Damaševičius  <http://orcid.org/0000-0001-9990-1084>

References

- Abedinia O, Bagheri M. 2022. Execution of synthetic Bayesian model average for solar energy forecasting. *IET Renew Power Gener.* 16(6):1134–1147. <https://doi.org/10.1049/rpg2.12389>
- Aissaoui F, Abdelmalek EM. 2026. Empowering energy forecasting with large language models: state of the art applications, challenges, and opportunities, Kodad M, Moussaoui O editors. *The 3rd International Conference on Artificial Intelligence and Smart Applications (AISA'25)*, Volume 1: Artificial Intelligence, IoT, and Smart Applications; Cham: Springer Nature Switzerland. p 179–191.
- Alay FD, İlhan N, Güllüoğlu MT. 2024. A comparative study of data mining methods for solar radiation and temperature forecasting models. *J Univers Comput Sci.* 30(6):847–877. <https://doi.org/10.3897/jucs.109080>
- Al-Dahidi S et al. 2024. Forecasting solar photovoltaic power production: a comprehensive review and innovative data-driven modeling framework. *Energies.* 17(16):4145. <https://doi.org/10.3390/en17164145>
- AlKandari M, Ahmad I. 2024. Solar power generation forecasting using ensemble approach based on deep learning and statistical methods. *Appl Comput Inf.* 20(3–4):231–250, 08.
- Asghar R, Fulginei FR, Quercio M, Mahrouh A. 2024. Artificial neural networks for photovoltaic power forecasting: a review of five

- promising models. *IEEE Access*. 12: 90461–90485. <https://doi.org/10.1109/ACCESS.2024.3420693>
- Babikou YN, Guenoukpati A, Salami AA, Kodjo KM. 2025. Hybrid approach based on wavelets, kalman filters, arima, and neural networks for mastering electrical energy at the residential level: forecasting the electrical energy consumption of residences in the city of Iomé. *Edelweiss Appl Sci Technol*. 9(3):1892–1907. <https://doi.org/10.55214/25768484.v9i3.5716>
- Bahi A et al. 2026. A comprehensive survey of llms for sustainable and renewable energy systems. *Information*. 17(3):271. <https://doi.org/10.3390/info17030271>
- Balfour J et al. 2021. Masking of photovoltaic system performance problems by inverter clipping and other design and operational practices. *Renewable Sustain Energy Rev*. 145:111067. <https://doi.org/10.1016/j.rser.2021.111067>
- Bessa R, Moreira C, Silva B, Matos M. 2014. Handling renewable energy variability and uncertainty in power systems operation. *Wiley Interdiscip Rev: Energy Environ*. 3(2):156–178. <https://doi.org/10.1002/wene.76>
- Böök H, Lindfors AV. 2020. Site-specific adjustment of a NWP-based photovoltaic production forecast. *Sol Energy*. 211:779–788. <https://doi.org/10.1016/j.solener.2020.10.024>
- Cao Y, Gao Q, Yu J, Yang Z. 2026. Automatically construct hard unit commitment datasets via data-driven adjustment of demand curve. *IEEE Trans Ind Inf*. 22(9):1–12. <https://doi.org/10.1109/TII.2026.3692472>
- Cao Y, Liu G, Sun J, Bavirisetti DP, Xiao G. 2023. Pso-stacking improved ensemble model for campus building energy consumption forecasting based on priority feature selection. *J Build Eng*. 72:106589. <https://doi.org/10.1016/j.jobe.2023.106589>
- Cao Y, Yong P, Yu J, Yang Z. 2025. Stacking algorithm based framework with strong generalization performance for ultra-short-term photovoltaic power forecasting. *Energy*. 322:135599. <https://doi.org/10.1016/j.energy.2025.135599>
- Chicco D, Warrens MJ, Jurman G. 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput Sci*. 7:e623. <https://doi.org/10.7717/peerj-cs.623>
- Chuluunsai Khan T, Kim J-H, Shin Y, Choi S, Nasridinov A. 2022. Feasibility study on the influence of data partition strategies on ensemble deep learning: the case of forecasting power generation in South Korea. *Energies*. 15(20):7482. <https://doi.org/10.3390/en15207482>
- de Oliveira JFL et al. 2023. Forecasting methods for photovoltaic energy in the scenario of battery energy storage systems: a comprehensive review. *Energies*. 16(18):6638. <https://doi.org/10.3390/en16186638>
- Dhingra S, Gruosso G, Gajani GS. 2025. A benchmark study of optimizers for short-term solar pv power forecasting using neural networks under real-world constraints. *Neural Comput Appl*. 37(29): 23909–23939. <https://doi.org/10.1007/s00521-025-11546-2>
- Di Leo P, Ciocia A, Malgaroli G, Spertino F. 2025. Advancements and challenges in photovoltaic power forecasting: a comprehensive review. *Energies*. 18(8):2108. <https://doi.org/10.3390/en18082108>
- Du H, Du S, Li W. 2023. Probabilistic time series forecasting with deep non-linear state space models. *CAAI Trans Intell Technol*. 8(1):3–13. <https://doi.org/10.1049/cit2.12085>
- El Hendouzi A, Bourouhou A. 2020. Solar photovoltaic power forecasting. *J Electr Comput Eng*. 2020:1–21. <https://doi.org/10.1155/2020/8819925>
- Esen V et al. 2025. An interpretable statistical approach to photovoltaic power forecasting using factor analysis and ridge regression. *Sci Rep*. 15(1). <https://doi.org/10.1038/s41598-025-22838-x>
- Fan G-F, Wei H-Z, Chen M-Y, Hong W-C. 2022. Photovoltaic power generation forecasting based on the ARIMA-BPNN-SVR model. *Global J Energy Technol Res Upd*. 9:18–38. <https://doi.org/10.15377/2409-5818.2022.09.2>
- Fan M et al. 2025. Distributed photovoltaic power prediction based on solar-LLM. *Ninth Int Conf Energy Sys Elect Power (ESEP 2024)*. 13632:136321F. <https://doi.org/10.1117/12.3060936>
- Gaboitoalelwe J et al. 2023. Machine learning based solar photovoltaic power forecasting: a review and comparison. *IEEE Access*. 11: 40820–40845. <https://doi.org/10.1109/ACCESS.2023.3270041>
- Gomes E, Esteves A, Morais H, Pereira L. 2025. Leveraging explainable artificial intelligence in solar photovoltaic mappings: Model explanations and feature selection. *Energies*. 18(5):1282. <https://doi.org/10.3390/en18051282>
- Gong J, Qu Z, Zhu Z, Xu H. 2025. Parallel timesnet-bilstm model for ultra-short-term photovoltaic power forecasting using stl decomposition and auto-tuning. *Energy*. 320:135286. <https://doi.org/10.1016/j.energy.2025.135286>
- Gong J, Qu Z, Zhu Z, Xu H, Yang Q. 2025. Ensemble models of tcn-lstm-lightgbm based on ensemble learning methods for short-term electrical load forecasting. *Energy*. 318:134757. <https://doi.org/10.1016/j.energy.2025.134757>
- Guo Y et al. 2025. Robust photovoltaic power forecasting model under complex meteorological conditions. *Mathematics*. 13(11):1783. <https://doi.org/10.3390/math13111783>
- Gupta P, Singh R. 2021. PV power forecasting based on data-driven models: a review. *Int J Sustain Eng*. 14(6):1733–1755. <https://doi.org/10.1080/19397038.2021.1986590>
- Hara S, Yoshioka T. 2021. Exclusion of pv inverter measurement data affected by power clipping and detection of string disconnection. 2021 IEEE 48th Photovoltaic Specialists Conference (PVSC), Fort Lauderdale, FL, USA. p 0979–0981.
- Hassan AA, Atia DM, Madany H, ElGhannam F. 2025. Multi-label machine learning for power forecasting of a grid-connected photovoltaic solar plant over multiple time horizons. *Sci Rep*. 15(1):1–23. <https://doi.org/10.1038/s41598-025-20251-y>
- Hassan I et al. 2024. Explainable deep learning model for grid-connected photovoltaic system performance assessment for improving system reliability. *IEEE Access*. 12: 120729–120746. <https://doi.org/10.1109/ACCESS.2024.3452778>
- Hou Z, Zhang Y, Liu Q, Ye X. 2024. A hybrid machine learning forecasting model for photovoltaic power. *Energy Rep*. 11:5125–5138. <https://doi.org/10.1016/j.egyr.2024.04.065>
- Hummon M et al. 2014. Variability of photovoltaic power in the state of gujarat using high resolution solar data. *Tech Rep* 1:1–65.
- Husein M, Gago E, Hasan B, Pegalajar M. 2024. Towards energy efficiency: a comprehensive review of deep learning-based photovoltaic power forecasting strategies. *Heliyon*. 10(13):e33419. <https://doi.org/10.1016/j.heliyon.2024.e33419>
- Iheanetu KJ. 2022. Solar photovoltaic power forecasting: a review. *Sustainability*. 14(24):17005. <https://doi.org/10.3390/su142417005>
- Jailani NLM et al. 2023. Investigating the power of LSTM-based models in solar energy forecasting. *Processes*. 11(5):1382. <https://doi.org/10.3390/pr11051382>
- Jordan DC, Deline C, Kurtz SR, Kimball GM, Anderson M. 2018. Robust pv degradation methodology and application. *IEEE J Photovolt*. 8(2):525–531. <https://doi.org/10.1109/JPHOTOV.2017.2779779>
- Jung A-H et al. 2022. Regional photovoltaic power forecasting using vector autoregression model in South Korea. *Energies*. 15(21):7853. <https://doi.org/10.3390/en15217853>
- Kim T-G, Kwon B-S, Yoon S-G, Song K-B. 2025. Very short-term load forecasting for large power systems with kalman filter-based pseudo-trend information using lstm. *Energies*. 18(18):4890. <https://doi.org/10.3390/en18184890>
- Lauret P, Alonso-Suárez R, Le Gal La Salle J, David M. 2022. Solar forecasts based on the clear sky index or the clearness index: which is better? *Solar*. 2(4):432–444. <https://doi.org/10.3390/solar2040026>
- Ledmaoui Y et al. 2023. Forecasting solar energy production: a comparative study of machine learning algorithms. *Energy Rep*. 10:1004–1012. <https://doi.org/10.1016/j.egyr.2023.07.042>
- Li P, Zhou K, Lu X, Yang S. 2020. A hybrid deep learning model for short-term pv power forecasting. *Appl Energy*. 259:114216. <https://doi.org/10.1016/j.apenergy.2019.114216>

- Li XY. 2016 aug. Degradation analysis of photovoltaic modules based on operational data: effects of seasonal pattern and sensor drifting. *IOP Conf Ser Earth Environ Sci.* 40(1):012063.
- Lim S-C, Huh J-H, Hong S-H, Park C-Y, Kim J-C. 2022. Solar power forecasting using CNN-LSTM hybrid model. *Energies.* 15(21):8233. <https://doi.org/10.3390/en15218233>
- Liu C et al. 2025. Generation and analysis of wind-photovoltaic power output scenarios for off-grid electrolysis systems. *Int J Hydrogen Energy.* 196:152509. <https://doi.org/10.1016/j.ijhydene.2025.152509>
- López Santos M, García-Santiago X, Echevarría Camarero F, Blázquez Gil G, Carrasco Ortega P. 2022. Application of temporal fusion transformer for day-ahead PV power forecasting. *Energies.* 15(14):5232. <https://doi.org/10.3390/en15145232>
- Mani G, Volety R. 2021. A comparative analysis of lstm and arima for enhanced real-time air pollutant levels forecasting using sensor fusion with ground station data. *Cogent Eng.* 8(1). <https://doi.org/10.1080/23311916.2021.1936886>
- Marchesi G, Ballarino A, Brusaferrri A. 2025. Assessing time series foundation models for probabilistic electricity price forecasting: toward a unified benchmark. *Energies.* 18(23):6269. <https://doi.org/10.3390/en18236269>
- Massidda L, Bettio F, Marrocu M. 2024. Probabilistic day-ahead prediction of pv generation. A comparative analysis of forecasting methodologies and of the factors influencing accuracy. *Sol Energy.* 271:112422. <https://doi.org/10.1016/j.solener.2024.112422>
- Mayer MJ. 2022. Benefits of physical and machine learning hybridization for photovoltaic power forecasting. *Renewable Sustain Energy Rev.* 168:112772. <https://doi.org/10.1016/j.rser.2022.112772>
- Mayer MJ, Gróf G. 2021. Extensive comparison of physical models for photovoltaic power forecasting. *Appl Energy.* 283:116239. <https://doi.org/10.1016/j.apenergy.2020.116239>
- Mayer MJ, Yang D. 2023. Pairing ensemble numerical weather prediction with ensemble physical model chain for probabilistic photovoltaic power forecasting. *Renewable Sustain Energy Rev.* 175:113171. <https://doi.org/10.1016/j.rser.2023.113171>
- Mellit A, Massi Pavan A, Lughi V. 2021. Deep learning neural networks for short-term photovoltaic power forecasting. *Renew Energy.* 172:276–288. <https://doi.org/10.1016/j.renene.2021.02.166>
- Mirshakali H, Reza Shadi M, Ghanadi Ladani F, Reza Shaker H. 2025. A review of large language models for energy systems: applications, challenges, and future prospects. *IEEE Access.* 13: 163162–163188. <https://doi.org/10.1109/ACCESS.2025.3610994>
- Mitrentsis G, Lens H. 2022. An interpretable probabilistic model for short-term solar power forecasting using natural gradient boosting. *Appl Energy.* 309:118473. <https://doi.org/10.1016/j.apenergy.2021.118473>
- Mohapatra M, Radhakrishnan R, Shukla R. 2023. A hybrid approach using arima, kalman filter and lstm for accurate wind speed forecasting. 2023 IEEE International Symposium on Smart Electronic Systems (iSES), Ahmedabad, India. p 425–428.
- Mystakidis A, Koukaras P, Tsalikidis N, Ioannidis D, Tjortjis C. 2024. Energy forecasting: a comprehensive review of techniques and technologies. *Energies.* 17(7):1662. <https://doi.org/10.3390/en17071662>
- Nespoli A et al. 2019. Day-ahead photovoltaic forecasting: a comparison of the most effective techniques. *Energies.* 12(9):1621. <https://doi.org/10.3390/en12091621>
- Orozco C et al. 2022. Intra-day scheduling of a local energy community coordinated with day-ahead multistage decisions. *Sustain Energy, Grids Networks.* 29:100573. <https://doi.org/10.1016/j.segan.2021.100573>
- Öter A, Ersöz B. 2025. Artificial intelligence assisted solar energy forecasting by explainability approaches with lime and shap. *El-Cezeri Fen ve Mühendislik Dergisi.* 12(2):205–212. <https://doi.org/10.31202/ecjse.1591721>
- Panamtash H, Zhou Q, Hong T, Qu Z, Davis KO. 2020. A copula-based Bayesian method for probabilistic solar power forecasting. *Sol Energy.* 196:336–345. <https://doi.org/10.1016/j.solener.2019.11.079>
- Pauwels V, Lannoy G, Franssen H, Vereecken H. 2013. Simultaneous estimation of model state variables and observation and forecast biases using a two-stage hybrid kalman filter. *Hydrol Earth Syst Sci.* 17(9):3499–3521. <https://doi.org/10.5194/hess-17-3499-2013>
- Petrosian OL, Zhang Y. 2024. Solar power generation forecasting in smart cities and explanation based on explainable ai. *Smart Cities.* 7(6):3388–3411. <https://doi.org/10.3390/smartcities7060132>
- Piantadosi G et al. 2024. Photovoltaic power forecasting: a transformer based frame-work. *Energy AI.* 18:100444. <https://doi.org/10.1016/j.egyai.2024.100444>
- Prema V, Bhaskar MS, Almahkles D, Gowtham N, Rao KU. 2022. Critical review of data, models and performance metrics for wind and solar power forecast. *IEEE Access.* 10:667–688. <https://doi.org/10.1109/ACCESS.2021.3137419>
- Rajagukguk RA, Lee H. 2025. Application of explainable machine learning for estimating direct and diffuse components of solar irradiance. *Sci Rep.* 15(1):7402. <https://doi.org/10.1038/s41598-025-91158-x>
- Rizk-Allah RM, Abouelmagd LM, Darwish A, Snásel V, Hassanien A. 2024. Explainable ai and optimized solar power generation forecasting model based on environmental conditions. *PLOS ONE.* 19(10): e0308002. <https://doi.org/10.1371/journal.pone.0308002>
- Rostami SMH, Pourgholi M, Asharioun H. 2025. Enhancing resilience of distributed dc microgrids against cyber attacks using a transformer-based kalman filter estimator. *Sci Rep.* 15(1). <https://doi.org/10.1038/s41598-025-90959-4>
- Sapundzhi F, Chikalov A, Georgiev S, Georgiev I. 2024. Predictive modeling of photovoltaic energy yield using an ARIMA approach. *Appl Sci.* 14(23):11192. <https://doi.org/10.3390/app142311192>
- Saxena N et al. 2024. Hybrid KNN-SVM machine learning approach for solar power forecasting. *Environ Challenges.* 14:100838. <https://doi.org/10.1016/j.envc.2024.100838>
- Sharma N et al. 2021. A sequential ensemble model for photovoltaic power forecasting. *Comput Electr Eng.* 96:107484. <https://doi.org/10.1016/j.compeleceng.2021.107484>
- Siddiqi A, Rana N, Khan WZ, Jeribi F, Tahir A. 2025. Explaining solar forecasts with generative AI: a two-stage framework combining transformers and llms. *PLOS ONE.* 20(9):e0331516. <https://doi.org/10.1371/journal.pone.0331516>
- Tajjour S, Chandel SS, Malik H, Márquez FPG, Alotaibi MA. 2025. Daily power generation forecasting for a grid-connected solar power plant using transfer learning technique. *Appl Intell.* 55(6). <https://doi.org/10.1007/s10489-024-06090-w>
- Theocharides S et al. 2021. Comparative analysis of machine learning models for day-ahead photovoltaic power production forecasting. *Energies.* 14(4):1081. <https://doi.org/10.3390/en14041081>
- Venitourakis G et al. 2023. Neural network-based solar irradiance forecast for edge computing devices. *Information.* 14(11):617. <https://doi.org/10.3390/info14110617>
- Wan C et al. 2015. Photovoltaic and solar power forecasting for smart grid energy management. *CSEE J Power Energy Syst.* 1(4):38–46. <https://doi.org/10.17775/CSEEJPES.2015.00046>
- Wang F et al. 2022. A satellite image data based ultra-short-term solar PV power forecasting method considering cloud information from neighboring plant. *Energy.* 238:121946. <https://doi.org/10.1016/j.energy.2021.121946>
- Wang K, Du H, Jia R, Jia H. 2022. Performance comparison of Bayesian deep learning model and traditional Bayesian neural network in short-term PV interval prediction. *Sustainability.* 14(19):12683. <https://doi.org/10.3390/su141912683>
- Xiang C, Liu X, Liu W, Yang T. 2025. A cascaded data-driven approach for photovoltaic power output forecasting. *Mathematics.* 13(17):2728. <https://doi.org/10.3390/math13172728>

- Yang Y, Yu T, Zhao W, Zhu X. 2021. Kalman filter photovoltaic power prediction model based on forecasting experience. *Front Energy Res.* 9. <https://doi.org/10.3389/fenrg.2021.682852>
- Zhang L, Ji W, Zhao Y, Huang D, Qian B. 2025. Review on application and challenges of large language models in power grids. *IEEE Access.* 13: 106932–106941. <https://doi.org/10.1109/ACCESS.2025.3580439>
- Zhang L, Zhu J, Zhang D, Liu Y. 2023. An incremental photovoltaic power prediction method considering concept drift and privacy protection. *Appl Energy.* 351:121919. <https://doi.org/10.1016/j.apenergy.2023.121919>
- Zhang Y, Kong L. 2022. Photovoltaic power prediction based on hybrid modeling of neural network and stochastic differential equation. *ISA Trans.* 128:181–206. <https://doi.org/10.1016/j.isatra.2021.11.008>
- Zhou Y, Wu S, Deng Y, Jiang M, Fu Y. 2025. Enhancing virtual power plant efficiency: three-stage optimization with energy storage integration. *Energy Inf.* 8(1). <https://doi.org/10.1186/s42162-025-00477-w>
- Zhu R, Das K, Sørensen PE, Hansen AD. 2025. A review on energy management system for grid-connected utility-scale renewable hybrid power plants. *Wiley Interdiscip Rev Energy Environ.* 14(1). <https://doi.org/10.1002/wene.70004>
- Zhu X et al. 2024. Overview of wind and photovoltaic data stream classification and data drift issues. *Energies.* 17(17):4371. <https://doi.org/10.3390/en17174371>