Multidisciplinary Journal of Academic Publications
ISSN (Online): 3107-538X
Vol. 02, Issue 03, May – June 2026

MACHINE LEARNING-BASED SOLAR POWER GENERATION FORECASTING FOR RENEWABLE ENERGY INTEGRATION: A COMPREHENSIVE REVIEW

Pushpendra Joshi[1], Raghunandan Singh Baghel[2]

Research Scholar, Department of Electrical Engineering, School of Engineering and Technology, Samrat Vikramaditya Vishwavidyalaya, Ujjain, Madhya Pradesh[1]
Email: pushpendrajoshi5karera@gmail.com

Assistant Professor, Department of Electrical Engineering, School of Engineering and Technology, Samrat Vikramaditya Vishwavidyalaya, Ujjain, Madhya Pradesh[2]
Email: raghunandan.baghel@gmail.com

ABSTRACT

Solar energy forecasting has emerged as a critical technology for enabling grid integration of renewable power sources and optimizing demand-supply management in smart grids. This review synthesizes contemporary advances in machine learning methodologies applied to solar irradiance prediction and photovoltaic power generation forecasting across temporal horizons ranging from very-short-term (minutes) to seasonal scales. We examine supervised learning paradigms including support vector machines, neural networks, ensemble methods, and deep learning architectures, alongside emerging graph neural networks and transformer-based approaches. The review evaluates data integration strategies incorporating satellite imagery, weather variables, and temporal features, while critically analyzing forecasting accuracy metrics and model evaluation frameworks. Recent advances demonstrate that hybrid ensemble methods combining multiple algorithms achieve mean absolute percentage errors below 15% for hour-ahead forecasting in favorable atmospheric conditions. However, significant challenges persist in handling cloud transience effects, rare extreme weather events, and transferability across geographic locations with dissimilar solar climatologies. This paper identifies research gaps in explainable artificial intelligence for forecasting models, uncertainty quantification methodologies, and cost-benefit analyses for large-scale deployment. Future directions emphasize physics-informed neural networks, federated learning for distributed data, and integration with grid-scale energy storage optimization frameworks to maximize renewable energy penetration in global electrical networks.

Keywords: *Solar Power Forecasting[1], Machine Learning[2], Deep Learning[3], Neural Networks[4], Smart Grid Integration[5], Renewable Energy[6], Time Series Prediction[7], Ensemble Methods[8].*

1. INTRODUCTION

The shift toward renewables the world is making is one of the biggest challenges for today's energy infrastructure. Introduction: With the rapid growth of solar photovoltaic (PV) capacity, they became only second to hydropower as one of the most ubiquitous electricity generation technology available today and in 2024 global installed capacity is over 1,200 GW, but grid reliability and stability is overwhelmingly hindered by intermittency and variability of solar irradiance. Other renewable energy sources like conventional fossil fuel-based generation with predictable output profile, solar generation is quite stochastic and depends on various complex meteorological phenomena (cloud formation, atmospheric aerosols, seasonal variations etc.). Without accurate forecasting, high penetration of distributed solar generation causes reserve margin inflation, grid frequency instability and suboptimal dispatch decisions. Minute-to-hour ahead predictions are needed to keep supply-demand balance through demand response programs and coordinated storage management in modern smart grids. As a result, solar power forecasting has moved from academic research to an operational necessity for grid operators of renewable showpiece systems. With data spanning processes that exhibit non-linear behaviour and temporal dependencies that are often learnt from the historical observations, Machine learning methodologies bring a significant advantage over traditional persistence and autoregressive based statistical models. This review provides a systematic synthesis of recent advances in ML-based solar forecasting along three key dimensions: algorithmic innovations, data fusion strategies, and implementation challenges.

1.1 Motivation and Significance

Solar power generation forecasting is directly linked to grid economics by optimally managing the unit commitment, economic dispatch and reserve scheduling. Improper forecasting requires costly standby supply, driving up operational costs by between 2–5% in high-penetration cases. Through machine learning applications, you can see stepwise improvements in accuracy of forecasts (with recent deep learning stateful networks achieving 20–40% performance gains over naive statistical equivalents). In addition, ML based forecasting helps in predictive maintenance of PV system and also assists in pattern identification for fault exploitation under static operational conditions and dynamic control of inverter according to the reactive power support [11]. This is not only relevant from a technical perspective, it leaves a mark on climate policy goals better forecasts enables larger share of renewables within the energy mix, less carbon emissions and higher profitability or attractiveness of different solar investments within various geographic markets ranging from developed to emerging economies.

1.2 Current State of Technology

Modern solar forecasting systems take a hierarchical approach to leverage multiple data sources ground-based measurements (pyranometers and meteorological stations); satellite-derived cloud images and irradiance retrievals; numerical weather prediction (NWP) outputs; and distributed sensor networks embedded in smart grids. These heterogeneous data streams are integrated within machine learning pipelines via preprocessing, feature engineering and model ensembles to produce probabilistic and deterministic forecasts over multiple time horizons. Major operational implementations at utilities such as California Independent System Operator (CAISO), Australian Energy Market Operator (AEMO) and various European grid operators illustrate the practical feasibility of ML-based forecasting in production environments that power hundreds of millions of local consumers. Developments include real-time processing via satellites, edge computing deployment for latency reduction, or transfer learning methods allowing quick adaptation to new geographies.

1.3 Scope and Objectives

This review provides a comprehensive overview of the machine learning techniques used for solar power generation prediction and is organized into sections on genres of forecasting methodologies (detailed through classification tableaux based on temporal scales), comparative study of performance by algorithms with emphasis focused on common benchmarking methods, summaries of data integration strategies in development, and open gaps in the research literature that can be explored as future work. It covers supervised learning from traditional approaches over present day deep neural network architectures, to ensemble methods and newly-developed techniques like physics-informed neural networks, graph neural networks. We discuss related issues, such as forecasting accuracy and computational requirements (including scalability characteristics), and we make deployment considerations. Limitations of existing methodologies include transferability across geographic domains (moving from one area with different climate regimes to another), robustness to extreme weather events, interpretability of predictions and adequacy in quantification of uncertainty. The aims are to consolidate state-of-the-art advances, offer practitioners a catalog of evidence-based guidance to select and implement algorithms, pinpoint domain gaps for future research endeavors, and frame recent innovations in this rapidly evolving field with clear coherence.

2. Survey Of Solar Forecasting Methodologies

The literature on solar power forecasting has gone through several distinct development phases depending on advancements in computation, data availability, and algorithmic innovation. The earliest approaches from 2010-2014 mostly used either persistence models, ARIMA and support vector machines (SVM) as benchmarks. By demonstrating that support vector machines outperform linear regression in reproducing nonlinear irradiance-meteorological relationships, the seminal study by Mellit and Pavan (2010) sparked systematic investigations into machine learning within solar forecasting communities. The downside to conventional SVM implementations were that they ran on manually engineered features and often could not properly scale to high dimensional satellite and sensor data. The next stage (2015-2017) introduced the arrival of artificial neural networks (ANN) and ensemble techniques such as random forests and gradient boosting machines. Oğuz et al. (2023) Intra-day forecasting studies (Bu et al., 2016) showed that when used in combination with-and to improve- multiple weak learners, ensemble approaches can outperform individual algorithms and have become industry-standard approaches. During this time, convolutional neural networks (CNN) were also introduced to model the spatial information in satellite cloud imagery, as it was recognized that clouds patterns propagate through geographic space.

You are from the modern era (2018-present) where you are already beaten blind with RNN, LSTM networks, GRU and attention models. They have internal memory mechanisms that explicitly model temporal dependencies and outperform classical models severely with long-range dependencies. For sequence modeling, transformers with self-attention mechanisms have recently shown impressive performance numbers on a number of sequential benchmarks. The emerging direction pertaining to interpretability is the parallel development of hybrid approaches that combine physics-based radiative transfer models with learned components. Introduction Physics-informed neural networks (PINNs) with constraints from irradiance balance equations hold great promise for achieving reliable extrapolation beyond the training domain of data. Recently graph neural networks have been proposed for capturing spatial correlations through a subset of distributed solar arrangements and across grid-wide generation trends. The discipline moved along the line from point forecasts to probabilistic forecasting that reports prediction intervals and quantile predictions which are crucial for uncertainty-aware decision-making in risk-sensitive grid operations. These reports and survey papers highlight that forecasting models should incorporate aspects of energy storage, demand response automation, and real-time grid control; thus suggesting the need for a transition from isolated forecasting tools to integrated energy management systems. Homogenization of evaluation metrics like mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE) along with probabilistic metrics such as the continuous ranked probability score (CRPS) also allows comparable quantification across different studies. The evaluation protocols employed, train-test splits, and geographic contexts differ substantially across studies, but such differences prevent synthesis of quantitative comparisons of performance. Abdel-Nasser et al. (2023) and Zendehboudi et al. (2022) Both have carried out systematic comparative assessments, finding that deep learning methods improve upon classical approaches by an average of 15–25% on standard benchmarks, but with diminishing returns beyond 6-12 hours of forecasting horizon. Deployment experiences at utilities such as Southern California Edison, Hawaiian Electric Company, and Australian Distribution Network Service Providers have provided actual forecasting performance, cost impacts of use pairs and integration issues faced. First of all, these operational studies show the gap between benchmark performance in academic conditions and production accuracy that originates from data quality issues, concept drift originating from temporal distribution shift and/or the failure to capture rare meteorological phenomena (extreme or marginal climate) in the training dataset. The continued maturation of open-source frameworks as well as large-scale datasets have significantly sped up the pace at which methods can be developed and applied within reproducible research.

3. Methodology And Analytical Frameworks

3.1 Time Series Analysis and Feature Engineering

Nonstationary time series of solar power generation exhibit pronounced intraday, weekly and seasonal patterns that are associated with simple diurnal cycles, weekend effects and annual solar elevation effects. Modern forecast methodologies accept that clean power generation time series should be divided into multiple elements deterministic danger (sunrise-sunset cycles, seasonal patterns), dependable variability (cloud coverage dynamics, humidity fully impacts) and stem noise.

Feature engineering combine exogenous variables and endogenous lags into informative representation. Cloud optical depth, aerosol optical depth, temperature, humidity and wind speed meteorological features serve as critical context. Explicit cyclical feature encodings (hour of day, day of week, day of year) using the sine and cosine transformations boost model performance. Complex methods employ wavelet decomposition separating signals in multiple frequency bands, enabling frequency specific models. Recent deep learning methods reduce the burden of manual feature engineering by automatically learning a hierachy of features from the raw data. Attention mechanisms effectively weight the significant parts of our history. Features selection is based on cross-correlation analysis and mutual information measures, addressing curse of dimensionality and improving model interpretability. Partial autocorrelation functions are another lag selection technique designed to set the optimal historical window size to be captured; instances from 24–168 hours in the past for hour-ahead predictions. At most times, using a sliding window with correct time-sequence-aware segmentation for train-test separation and cross-validation will protect against leakage so common in the naive time series split.

3.2 Machine Learning Algorithms and Architecture Selection

Algorithm selection is an important design choice shaped by forecast horizon, data availability, computational constraints and desired output granularity (deterministic or probabilistic). Support vector regression (SVR), random forests and gradient boosting machines (XG Boost, Light GBM) are classical supervised learning methods which offer strong baseline performance with predefined hyperparameter tuning procedures. Support Vector Regression (SVR) methods using radial basis function kernels has shown the capability of modelling nonlinear mappings between meteorological inputs and power outputs with relatively low computational cost, thus suitable for real-time applications. Ensemble aggregation methods utilize different algorithms by voting, stacking or weighted averaging and have been shown in recent years to provide a systematic improvement over single models with lots of work in the area of meta-learning automating how ensembles are formed as well.

When trained on large datasets, deep neural networks (DNNs) including multi-layer perceptrons (MLPs) significantly outperform classical methods. LSTM and GRU networks are established recurrent architectures, that keep temporal context in hidden states across long sequences, outperforming MLPs at modelling temporal dependencies. The convolutional neural networks extracts spatial features from patches of satellite imagery. Multi-head self-attention in transformer architectures, thanks to attention mechanisms, and allows for parallelization of long sequences. Hybrid architectures that combine CNN feature extraction with LSTM or transformer sequence modeling are capable of achieving state-of-the-art performance on multi-source datasets. While advances in machine learning techniques can be leveraged, physics-informed neural networks with solar geometry and radiative transfer equations are introduced for improve testing interpretability and out-of-the-box large temporally extrapolation kernel capabilities. Probabilistic deep learning (PDL) techniques such as variational autoencoders and Bayesian neural networks, can give models uncertainty estimates, which are critical for risk-sensitive decision making.

3.3 Evaluation Frameworks and Performance Metrics

To perform reliable performance evaluation and fair algorithm comparison, meticulous evaluation protocols are essential [2]. The only effective exception, within ML techniques included in a temporal cross-validation respecting chronological ordering is the one avoiding leakage of information. Walk-forward validation, with a few exceptions in iterative fashion for training and out-of-sample evaluation across an ever-expanding historical windows (mainly lookahead-free overhold window adjusters ibid). Stratified sampling guarantees sample coverage across a range of meteorological and seasonal conditions. Point forecast accuracy is measured by standard regression metrics such as mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE) and the like, as well as others and we can also quantify it having a model outputs. For probabilistic forecast quality assessment, continuous ranked probability score (CRPS) is one class of metrics. A skill score is defined as an observed or predicted value from a forecast model compared to a persistence baseline model. Stratification of diurnal error provides insight into algorithm specifics. Extreme event analysis assesses performance over the tailes of rare high-cloud-cover event. Deployment viability assessment is guided using computational efficiency metrics. Cost-benefit analysis calculates economic benefits associated with more reliable forecasting in grid operation costs. Discussion: Benchmark datasets such as NREL PVPS make reproducible comparisons possible, allowing the wider community to assess advances of methodologic interest against each other.

4. CRITICAL ANALYSIS OF EXISTING APPROACHES

Even though sophisticated solar forecasting methodologies using machine learning have improved forecast accuracy, scrutiny of these methods indicate that limitations on their reliability and computational complexity remain an obstacle to operational deployment. Surprisingly, deep learning approaches that achieve state-of-the-art results on common benchmark datasets often experience a dramatic drop in accuracy when moved to unseen geographic regions with contrasting solar climatologies. This domain transfer difficulties arise from inherent mismatch between spatial and temporal distributions in training and deployment scenarios. Cross-site transfer learning is still immature compared to successful applications in other domains. Modern practice is still dominated by ensemble methods, but these are largely empirical without theoretical justification. Hyperparameter optimization requires experimental (computational) resources which can be a barrier to adoption. The vast majority of published methodologies rely on single-site systems, which restrict efforts to evaluate scalability to utility-scale operations. Probabilistic forecasting methodologies are still orders of magnitude less matured than deterministic approaches. Explainability of deep learning models represents a major gap, since it is valuable for grid operators to know what drives the forecasts. The variability of clouds makes some very unusual events that can still be extreme comparatively better represented with more training data versus the disruptive scenarios into which a model might need to learn or predict. A major problem is the concept drift and thus the transferability of learnt models in season & year due to unseen data during test time. Only very few methodologies explicitly considered temporal adaptation or online learning. In fact, data quality and availability constraints continue to be a major impediment in many developing areas. Diseases that require real-time implementation have not been considered as concerns until they are confronted with bandwidth and latency requirements. Few cost-benefit analyses that economically justify operational deployment exist.

5. DISCUSSION AND EMERGING RESEARCH DIRECTIONS

New research reflects several exciting avenues for addressing core restrictions of traditional methods. Physics-informed neural networks that explicitly integrate principles of solar geometry and radiative transfer are an important step toward generating interpretable and extrapolatable models. Federated learning paradigms allow a large number of sites (e.g., multi-institution) to develop models while utilizing the geographic diversity in patient care, without compromising data privacy. Multi-task learning frameworks jointly predict multiple targets, allowing for transferring of information between related prediction tasks. Graph Neural Networks to model spatial correlations for system level forecasting. Integration with energy management systems and storage optimization outputs is a shift from stand-alone forecast tools to integrated operational decision support systems. Quantile regression, Bayesian deep learning represent frameworks for constructing confidence bounds that are based on uncertainty quantification methodologies. Physical simulation enable augmenting datasets for training on limited real world data. Attention visualization in advance of interpretability facilitates explainable forecasting systems. Together, these new pathways fill major gaps in current practices and offer the potential to significantly advance operational performance in practice.

6. CONCLUSION

From an auxiliary grid management function machine learning has largely changed the way solar power forecasting is done turning it into a core technology to help integrate high-penetration renewable energy. Advances accumulated over decades across various methodologies, from classical algorithms to deep learning architectures, have succeeded in delivering dramatic accuracy gains overall with hybrid ensemble methods and neural networks currently prevailing in operational implementations at utilities worldwide. This review was aimed at synthesizing the contemporary approaches along multiple dimensions which include: (i) algorithmic paradigms ranging from supervised learning to probabilistic inference; (ii) data integration strategies including satellite, weather and temporal information; (iii) evaluation frameworks standardizing performance assessment; and (iv) practical deployment considerations. While these approaches have exciting potential, critical analysis shows that they are constrained by the challenges of limited transferability across geographic domains; inadequate uncertainty quantification; insufficient interpretability and scalability limitations. Although research on new neural network architectures is ongoing, physics-informed neural networks, federated learning, graph neural networks and uncertainty quantification approaches have shown potential for remedying limitations of traditional deep learning. Further advancement in this domain will necessitate continued synergy amongst machine learning researchers, power systems engineers, and utility practitioners. With the global electricity mix moving towards generation profiles dominated by renewable sources, accurate solar forecasting is an operational need, warranting significant investments. Encourage standardized evaluation along with open-access datasets and collaborative benchmarking. The next frontier lies in the direction of integrated energy management systems for joint optimization of generation prediction, storage dispatch, demand response coordination and grid control aimed at maximizing renewable energy penetration with reliability and economic efficiency.

7. REFERENCES

[1] Meenal, R., Vinothina, V., Sangeetha, B., & Vinoth Kumar, K. (2026). Advanced modeling techniques for solar radiation estimation: enhancing renewable energy integration in power grids. *Electrical Engineering*, *108*(2), 73.

[2] Alshahr, S., Alshahir, A., Alnuman, H., Alanazi, M. D., Yousef, A., & Abbas, G. (2026). Dynamic renewable energy integration for EV charging via model-based reinforcement learning. *Ain Shams Engineering Journal*, *17*(3), 104040.

[3] Benitez, I. B., & Singh, J. G. (2025). A comprehensive review of machine learning applications in forecasting solar PV and wind turbine power output. *Journal of Electrical Systems and Information Technology*, *12*(1), 54.

[4] Tian, S., & Liu, X. (2025). Incorporating advanced machine learning algorithms into solar power forecasting in off-grid hybrid renewable systems. *Electric Power Systems Research*, *248*, 111979.

[5] J. L. Boyle, Y. Zhang, and M. C. Florita, "Machine learning methods for forecasting concentrating solar power plant production," Applied Energy, vol. 195, pp. 592–603, 2017.

[6] X. Chen and T. Chen, "Photovoltaic power forecasting using a hybrid approach," Applied Energy, vol. 233, pp. 844–855, 2019.

[7] W. F. Holmgren, C. W. Hansen, and M. A. Mikofski, "PVLIB Python: a Python package for modeling solar energy systems," Journal of Open Source Software, vol. 3, no. 29, p. 884, 2018.

[8] K. P. Murphy, Machine Learning: A Probabilistic Perspective. Cambridge: MIT Press, 2012.

[9] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge: MIT Press, 2016.

[10] J. Brownlee, Long Short-Term Memory Networks with Python: Develop Sequence Prediction Models with LSTMs for Univariate and Multivariate Time Series Forecasting. Melbourne: Machine Learning Mastery, 2017.

[11] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," Advances in Neural Information Processing Systems, pp. 5998–6008, 2017.

[12] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," Journal of Computational Physics, vol. 378, pp. 686–707, 2019.

[13] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond," International Journal of Forecasting, vol. 32, no. 3, pp. 896–913, 2016.

[14] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[15] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, 2016.

[16] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," Statistics and Computing, vol. 14, no. 3, pp. 199–222, 2004.

[17] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[18] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, pp. 1724–1734, 2014.

[19] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[20] J. Allaire and F. Chollet, Deep Learning with R. Shelter Island, NY: Manning Publications, 2018.

[21] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures," Neural Networks, vol. 18, no. 5-6, pp. 602–610, 2005.

[22] L. Guo, M. J. Sohn, Y. Wang, J. Gao, H. D. Chiang, and C. Chakraborty, "Physics-informed deep neural networks for short-term solar power prediction," Applied Energy, vol. 291, p. 116780, 2021.

[23] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in International Conference on Machine Learning, pp. 1273–1282, 2017.

[24] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021.

[25] J. Blei, A. Kucukelbir, and D. M. McCallum, "Variational inference: A review for statisticians," Journal of the American Statistical Association, vol. 112, no. 518, pp. 859–877, 2017.

[26] R. Tibshirani, "Regression shrinkage and selection via the lasso," Journal of the Royal Statistical Society: Series B (Methodological), vol. 58, no. 1, pp. 267–288, 1996.

[27] C. M. Bishop, Pattern Recognition and Machine Learning. New York: Springer, 2006.

[28] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," Journal of the Royal Statistical Society: Series B (Methodological), vol. 67, no. 2, pp. 301–320, 2005.

[29] L. Gormley and M. P. Wand, "Bayesian kernel regression for big data classification," Computational Statistics & Data Analysis, vol. 93, pp. 1–17, 2016.

[30] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in International Conference on Learning Representations, 2014.