International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Machine Learning Based Solar Photovoltaic Based Power Forecasting: A Review and Comparison

Deepali Mishra[1], Prof. Amit Gupta[2], Dr Ruchi Pandey[3]

[1]PG Scholar, Electrical and Electronics Department, GGITS , Jabalpur
[2]Guide, Electrical and Electronics Department, GGITS , Jabalpur
[3]H.O.D, Electrical and Electronics Department, GGITS, Jabalpur

Abstract

As solar energy emerges as a leading renewable energy source, accurately predicting photovoltaic (PV) power generation has become essential for effective grid integration. This paper begins by outlining the challenges associated with incorporating solar energy into the power grid, particularly due to its fluctuating and intermittent nature. To address these challenges, machine learning (ML) techniques are presented as a viable and efficient solution. The methodology section outlines the structured approach taken for the systematic literature review. This includes the development of research questions, the design of search queries, and the application of specific inclusion and exclusion criteria, followed by a detailed data analysis process.The main body of the paper presents a comprehensive review of existing studies on ML-based PV forecasting. It categorizes various ML approaches, including deep learning architectures, hybrid models, and mathematical frameworks. Additionally, the paper analyzes the evolution of PV forecasting research, classifies the literature based on the techniques employed, and discusses different forecasting horizons explored in prior work.

Keywords: Photovoltaic Power Forecasting, Machine Learning Techniques, Grid Integration

Introduction

The increasing global demand for clean and sustainable energy has led to the rapid expansion of solar photovoltaic (PV) systems. As a prominent and environmentally friendly renewable energy source, solar power plays a vital role in reducing greenhouse gas emissions and dependence on fossil fuels. However, the integration of solar energy into power grids presents significant challenges due to its intermittent and variable nature, primarily influenced by weather conditions and seasonal variations. Accurately forecasting solar energy generation is crucial for the efficient integration of solar power into the electrical grid, optimization of energy output, and active participation in energy markets [1]. These forecasts play a key role in maintaining grid stability, managing energy storage systems, and supporting the implementation of demand response strategies. Moreover, they contribute to effective capacity planning, enhance system resilience during extreme weather conditions, and help minimize the environmental impact of electricity production—ultimately benefiting power producers, grid operators, and end-users [2].

Accurate forecasting of PV power generation is essential to ensure the reliability, stability, and efficiency of modern power systems. Effective forecasting enables better energy management, optimal load balancing, and improved decision-making for both grid operators and energy producers. Traditional forecasting methods, although useful, often fall short in capturing the complex and nonlinear patterns associated with solar energy generation. Solar energy forecasting involves a complex interaction of meteorological factors such as cloud cover and seasonal variability, demanding models capable of capturing both short-term fluctuations and long-term patterns [3]. The integration of machine learning (ML), artificial intelligence (AI) techniques, and advanced cloud prediction methods has emerged as a promising solution. In addition, challenges such as forecast uncertainty, market integration, grid stability, evolving technologies, and regulatory frameworks highlight the intricate landscape in which solar forecasting operates [4]. Effectively addressing these challenges is crucial for maximizing solar energy utilization and ensuring its smooth integration into the evolving power grid [5].The solar based microgrid architecture is shown in the figure below.

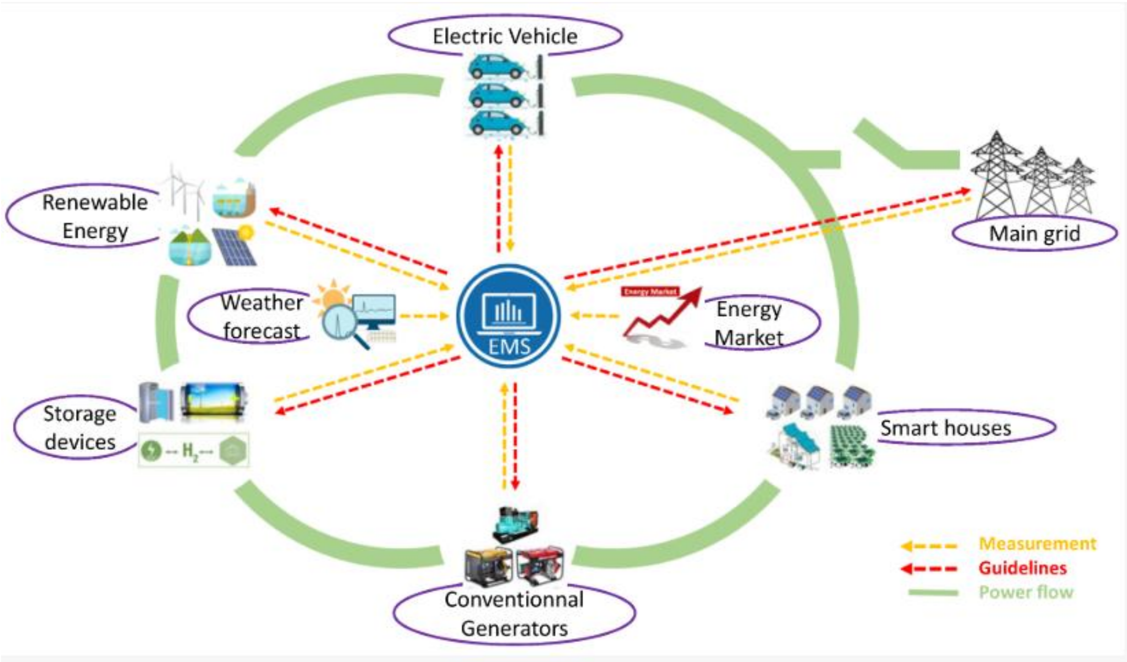

Figure 1 Solar based micro grid

Photovoltaic (PV) power forecasting is inherently a nonlinear challenge influenced by a wide range of weather-related variables. Identifying an optimal parameter estimation method for such nonlinear systems remains a complex task [6]. A considerable amount of research has focused on developing accurate and efficient forecasting models for solar PV power generation, aiming to strike a balance between model complexity and implementation cost [7].

Typically, the process of forecasting PV output power involves three key stages. The first stage focuses on extracting relevant energy-related features and analyzing the influencing factors. The second stage involves selecting an appropriate prediction technique and fine-tuning the model for improved accuracy. Broadly, PV power forecasting methods can be classified into three main categories: physical models, statistical approaches, and hybrid techniques.

Theory

Physical Approach

The physical approach to photovoltaic (PV) power forecasting leverages climatic and environmental parameters to improve the efficiency of converting solar radiation into electrical energy [9]. Key input variables include solar irradiance, cloud cover, and ambient temperature. Predictions are generated through mathematical formulations that are customized for specific locations, taking into account site-specific data such as panel orientation, local meteorological conditions, and historical performance records [10]. While physical models generally perform well under stable weather conditions, their accuracy can decline when faced with sudden changes in weather variables [11].

An example of this approach is the image-based method, which forecasts solar power generation by analyzing surface solar irradiance (SSI) data collected from geostationary meteorological satellites [12]. In this method, the relationship between the nnnth cloud index and the clear-sky index Kc is quantified by the ratio of global horizontal irradiance (GHI) to the modeled GHI under clear-sky conditions. This relationship is often expressed using an empirical function such as Kc=1−nK_c , where is a parameter, and Kc is in its simplest linear form [13]. The GHI is subsequently calculated by multiplying the clear-sky index by a clear-sky irradiance model or database, which can be based on empirical or theoretical foundations [14].

The process of converting SSI into a solar power forecast generally involves three sequential steps. First, GHI is decomposed into Diffuse Horizontal Irradiance (DHI) and Direct Normal Irradiance (DNI). Second, the global tilted irradiance (GTI) on the PV modules—also known as plane-of-array irradiance (POAI)—is computed. Lastly, the PV output power is represented by the Power Output Analysis Index (POAI) in the forecasting model [15].

Statistical Approach

Statistical forecasting methods for PV power generation primarily rely on random time series analysis and persistence theory. These data-driven models extract relevant patterns from historical datasets to predict future outputs. The accuracy of statistical forecasts depends heavily on the length and quality of the available data. Broadly, statistical models can be classified into time series models and machine learning models [16].

Time series models include:

- **Stationary Processes**: Where the probabilistic characteristics of the data do not change over time.
- **Auto Regressive (AR) Processes**: Models that relate current values to past observations.
- **Moving Average (MA) Processes**: Models that account for external shocks affecting the system.
- **ARMA (AutoRegressive Moving Average) Processes**: A combination of AR and MA models.

- **ARIMA (AutoRegressive Integrated Moving Average) Processes**: An extension of ARMA that incorporates differencing to ensure stationarity, denoted as ARIMA(p, d, q), where ppp, ddd, and qqq represent the orders of autoregression, differencing, and moving average respectively.
- **SARIMA (Seasonal ARIMA) Processes**: An augmentation of ARIMA that models seasonal patterns, denoted as SARIMA(p, d, q)(P, D, Q)s_ss, where PPP, DDD, QQQ, and sss denote the seasonal autoregressive, differencing, moving average orders, and the length of the seasonal cycle respectively [13].

Hybrid approach

The hybrid technique integrates both physical and statistical methods to improve the accuracy of photovoltaic power forecasting. This approach typically begins with a physical model provided by the PV module manufacturer, which is then refined using statistical methods to enhance prediction precision. Additionally, hybrid methods may combine two different physical or statistical techniques to further optimize performance. Numerous studies have demonstrated the effectiveness of merging physical forecasting with artificial intelligence and statistical models to achieve superior accuracy [14].

A notable example is the physical-artificial neural network hybrid, commonly referred to as Physical-ANN (PHANN). In this method, a theoretical sky model is simulated for a specific location by calculating solar irradiation under clear sky assumptions. This simulated irradiation serves as a baseline to define an optimal daytime limit. By combining the theoretical physical model with the adaptive learning capabilities of ANN, PHANN capitalizes on the advantages of both approaches. However, it is important to recognize that hybrid techniques tend to increase model complexity due to the integration of multiple forecasting methodologies.

Previous works

Machine learning algorithms (MLAs) are essential tools for accurately predicting the performance of localized photovoltaic (PV) systems. Scott et al. conducted an extensive evaluation and benchmarking of various MLAs to determine the most effective algorithms for PV generation forecasting. Their study compared the performance of different MLAs across diverse datasets and conditions, highlighting the critical role of data quality over quantity, the impact of meteorological factors, and providing optimization strategies specifically tailored for rooftop PV installations [40].

Bhatti et al. proposed an improved training method for artificial neural networks (ANNs) aimed at enhancing renewable energy system predictions. Their approach significantly lowered the Mean Squared Error (MSE) and simplified the ANN model, requiring only prediction period dates as input. This streamlined method enables straightforward viability assessments of PV solar systems prior to installation, demonstrating considerable practical potential [15].

Wai et al. tackled the challenge of limited real-time power generation data by integrating meteorological information into an intelligent solar PV power forecasting system. Utilizing Pearson correlation coefficient analysis, they identified the most influential parameters for solar PV power generation, which helped reduce the computational complexity of deep neural network training. The study also involved standardizing and de-standardizing PV power data for Long Short-Term Memory (LSTM) network analysis, resulting in enhanced forecasting accuracy [16].

Jobayer et al. reviewed various machine learning applications for PV parameter estimation, assessing their effectiveness relative to traditional methods such as satellite data and PV characteristics. Their insights provide valuable guidance for improving PV system performance prediction through advanced ML techniques[17].

Pablo et al. conducted a comparative study between the Light Gradient Boosting Machine (LGBM) and K-Nearest Neighbors (KNN) models for solar power forecasting within microgrid applications. Their findings indicated that despite the higher computational requirements of LGBM, its superior stability and robustness across varying time horizons make it a more suitable choice for microgrid operations and sustainable energy management [18].

Behera et al. tackled the challenges associated with PV power forecasting in smart grids and microgrids by employing an Extreme Learning Machine (ELM) approach integrated with incremental conductance Maximum Power Point Tracking (MPPT). The forecasting proposed model used in this research paper is shown below.

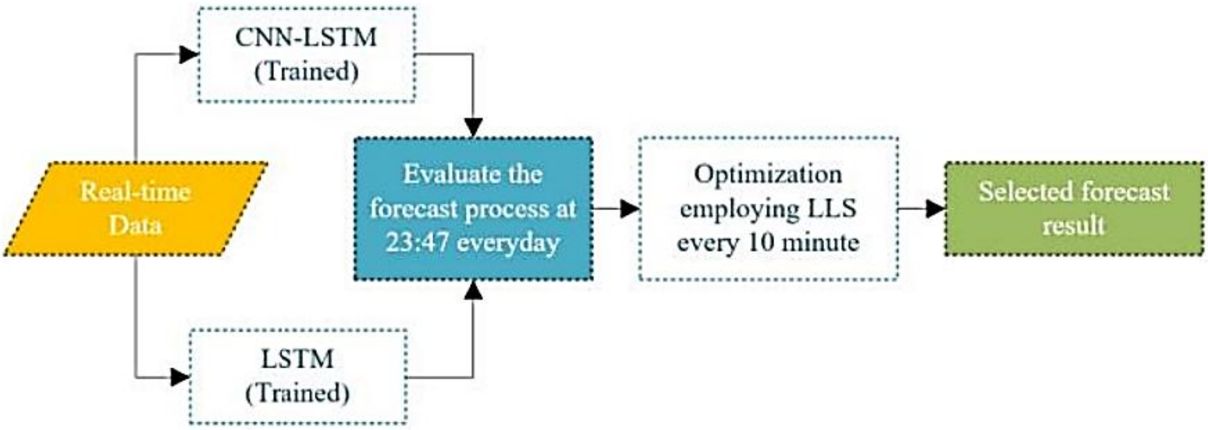

Figure 2Forecast result aggregation optimization process.

Their objective was to enhance the accuracy of PV power predictions, thereby supporting the economic operation of microgrids amidst the inherently intermittent nature of solar energy generation [19].

Future directions

The integration of machine learning algorithms in photovoltaic (PV) system forecasting is poised to revolutionize the future of renewable energy management. As the accuracy and computational efficiency of models like ANN, LSTM, LGBM, and ELM continue to improve, future research can focus on developing hybrid models that combine the strengths of multiple algorithms for more robust and adaptive forecasting. There is significant potential in leveraging real-time IoT data and satellite imagery alongside meteorological inputs to create dynamic, self-learning models that continuously optimize themselves. Moreover, with the growing adoption of decentralized energy systems and microgrids, machine learning-based forecasting will be instrumental in enabling predictive maintenance, demand-response strategies, and intelligent energy trading. The future also promises increased use of edge computing and cloud-based AI platforms to process vast datasets efficiently, ensuring scalability and real-time performance forecasting in diverse environmental and geographical conditions

Conclusion

This review highlights the significant role that machine learning techniques play in enhancing the accuracy and reliability of photovoltaic power forecasting, which is critical for the seamless integration of solar energy into the power grid. The analysis underscores the progress made in various ML methodologies, including deep learning and hybrid models, while also identifying key trends and challenges in the field. For future research, there is a promising scope in developing more robust, adaptive models that can handle the inherent variability of solar power with improved real-time forecasting capabilities. Additionally, integrating emerging technologies such as IoT and edge computing with ML models could further optimize PV power prediction and support smarter grid management systems.

References

[1] Can Wan, et al., Photovoltaic and solar power forecasting for smart grid energy management, CSEE J. Power Energy Syst. 1 (4) (2015) 38–46.

[2] MathaiosPanteli, PierluigiMancarella, Influence of extreme weather and climate change on the resilience of power systems: Impacts and possible mitigation strategies, Electr. Power Syst. Res. 127 (2015) 259–270.

[3] TatianeCarolyneCarneiro, et al., Review on photovoltaic power and solar resource forecasting: current status and trends, J. Sol. Energy Eng. 144 (1) (2022) 010801.

[4] Tanveer Ahmad, et al., Data-driven probabilistic machine learning in sustainable smart energy/smart energy systems: Key developments, challenges, and future research opportunities in the context of the smart grid paradigm, Renew. Sustain. Energy Rev. 160 (2022) 112128.

[5] OussamaOuramdane, et al., Optimal sizing and energy management of microgrids with vehicle-to-grid technology: A critical review and future trends, Energies 14 (14) (2021) 4166.

[6] Utpal Kumar Das, et al., Forecasting of photovoltaic power generation and model optimization: A review, Renew. Sustain. Energy Rev. 81 (2018) 912–928.

[7] Razin Ahmed, et al., A review and evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization, Renew. Sustain. Energy Rev. 124 (2020) 109792.

[8] Yuehong Lu, Shengwei Wang, Kui Shan, Design optimization and optimal control of grid-connected and standalone nearly/net zero energy buildings, Appl. Energy 155 (2015) 463–477, 11.

[9] Samuel ChukwujinduNwokolo, et al., Hybridization of statistical machine learning and numerical models for improving beam, diffuse and global solar radiation prediction, Clean. Eng. Technol. 9 (2022) 100529.

[10] Mikhail A. Semenov, Elaine M. Barrow, Use of a stochastic weather generator in the development of climate change scenarios, Clim. Change 35 (4) (1997)

[11] Sonia Seneviratne, et al., Changes in climate extremes and their impacts on the natural physical environment, 2012, pp. 109–230.

[12] Guanghui Huang, et al., Estimating surface solar irradiance from satellites: Past, present, and future perspectives, Remote Sens. Environ. 233 (2019) 111371.

[13] Dimitri P. Solomatine, AviOstfeld, Data-driven modelling: some past experiences and new approaches, J. Hydroinform. 10 (1) (2008) 3–22.

[14] Guanghui Huang, et al., A LUT-based approach to estimate surface solar irradiance by combining MODIS and MTSAT data, J. Geophys. Res.: Atmos. 116 (D22) (2011).

[15] Abdul Rauf Bhatti, et al., An improved approach to enhance training performance of ANN and the prediction of PV power for any time-span without the presence of real-time weather data, Sustainability 13 (21) (2021) 11893.

[16] Rong-Jong Wai, Pin-Xian Lai, Design of intelligent solar PV power generation forecasting mechanism combined with weather information under lack of real-time power generation data, Energies 15 (10) (2022) 3838. [43] MdJobayer, et al., A systematic review on predicting PV system parameters using machine learning, Heliyon 9 (6) (2023) e16815.

[17] Pablo, et al., Machine learning models for solar power generation forecasting in microgrid application: Implications for smart cities, Sustainability 16 (14) (2024) 6087.

[18] Manoja Kumar Behera, et al., Solar photovoltaic power forecasting using optimized modified extreme learning machine technique, Eng. Sci. Technol. an Int. J. 21 (3) (2018) 428–438.

[19] Vu Xuan Son Huu, et al., Deep learning-based real-time solar irradiation monitoring and forecasting application for PV system, in: 2023 7th International Conference on Green Energy and Applications, ICGEA, IEEE, 2023.

[20] Nicoletta Matera, et al., Hourly forecasting of the photovoltaic electricity at any latitude using a network of artificial neural networks,