ISSN 2319-5991 www.ijerst.com

Peer Reviewed, Referred & Indexed Journal

Research Paper

Hybrid Grey Wolf Optimization and ANN Model for Solar Panel Efficiency Prediction

Reddy Kowshik [1], Puvvala Supriya [2]

M.Tech, Department of Master of Computer Applications, Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh, India
Email: kowshikreddy9491@gmail.com

Assistant Professor, Department of Computer Science and Engineering, Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh, India
Email: supriya.puvvala@gmail.com

Abstract**—***Solar power generation is inherently difficult to forecast because output is strongly influenced by continuously changing environmental conditions such as irradiance, temperature, humidity, and cloud cover, and because field sensor readings are frequently noisy. Conventional machine learning approaches such as Random Forest combined with Lasso-based feature selection perform reasonably well under stable weather but struggle to adapt when sky conditions change abruptly, since their feature sets are fixed and their capacity to model nonlinear relationships is limited. This paper proposes a hybrid forecasting framework that couples Grey Wolf Optimization (GWO) with an Artificial Neural Network (ANN) to overcome these limitations. GWO is employed as an adaptive, metaheuristic feature-selection mechanism that dynamically identifies the most influential parameters — irradiance, temperature, humidity, voltage, current, panel age, soiling ratio, and module temperature — from large meteorological datasets, while the ANN learns the nonlinear mapping between the optimized feature subset and solar power output through backpropagation. The resulting GWO-ANN model reduces feature redundancy, accelerates convergence, and improves robustness to sensor noise relative to static feature-selection methods. Evaluation of the training and validation loss/MAE curves demonstrates stable convergence and consistent error reduction across training epochs, and the trained model was verified through a real-time weather-simulation and manual-input interface. The proposed hybrid framework offers a scalable, adaptive solution suitable for real-time photovoltaic monitoring and smart-grid energy-management applications.*

Keywords**—***Grey Wolf Optimization (GWO), Artificial Neural Network (ANN), Solar Power Forecasting, Photovoltaic (PV) Systems, Feature Selection, Renewable Energy, Smart Grid, Metaheuristic Optimization.*

I. INTRODUCTION

Solar power has emerged as one of the most widely adopted sources of clean electricity generation, largely because it is abundant, scalable, and environmentally sustainable. However, the electrical output of a photovoltaic (PV) installation is not constant; it is governed by dynamic environmental factors such as solar irradiance, ambient temperature, humidity, and atmospheric conditions, all of which fluctuate continuously throughout the day. Because these variables change so frequently, forecasting solar power output with high confidence remains a challenging task. Accurate forecasts are nevertheless essential, since they directly influence energy-management efficiency, grid stability, and the effectiveness of smart-grid applications that must balance supply and demand in real time.

Traditional machine learning (ML) techniques, particularly Random Forest (RF) regression combined with Lasso-based feature selection, have been widely applied to solar power prediction. These methods perform adequately when environmental conditions are relatively stable, but their accuracy deteriorates sharply when the weather changes suddenly, when sensor data is noisy, or when the relationship between the input parameters and solar output is strongly nonlinear. A further shortcoming is that these models typically rely on a static, pre-selected set of features; this rigidity limits their capacity to adapt automatically as data patterns evolve.

To address these limitations, this work proposes a hybrid technique that integrates Grey Wolf Optimization (GWO) with an Artificial Neural Network (ANN). GWO is a metaheuristic optimization algorithm inspired by the leadership hierarchy and cooperative hunting behaviour of grey wolves in nature [1]. In the proposed framework, GWO is used as an intelligent, adaptive feature-selection mechanism that dynamically identifies the meteorological and electrical parameters most influential to solar power output, thereby reducing redundancy and dimensionality. The optimized feature subset is then supplied to an ANN, which is well suited to learning the complex, nonlinear relationships between environmental conditions and PV output through supervised training [2], [3].

Integrating GWO with an ANN improves prediction accuracy, accelerates convergence, and increases robustness against missing or noisy data compared with conventional static-feature machine learning pipelines. In addition to a forecasting module trained on historical data, the developed system incorporates a real-time weather-simulation component that generates realistic synthetic environmental conditions, enabling the behaviour of the trained model to be evaluated dynamically across a range of climates. Collectively, these contributions are intended to support intelligent solar forecasting for renewable-energy optimization, smart-grid management, and real-time photovoltaic monitoring.

The principal objectives of this work are to: (i) apply GWO-based feature selection to identify the meteorological parameters most relevant to power generation; (ii) design an ANN regression model capable of learning nonlinear relationships from the optimized feature set; (iii) minimize prediction error, measured through Mean Absolute Error (MAE) and Root Mean Square Error (RMSE); (iv) improve prediction stability under noisy and variable weather; (v) develop a system suitable for smart-grid applications and real-time solar monitoring; and (vi) enhance the generalization capability and convergence speed of the resulting model.

The remainder of this paper is organized as follows. Section II reviews related work on machine learning and metaheuristic-assisted solar forecasting. Section III describes the existing static-feature machine learning pipeline and its limitations. Section IV presents the proposed research methodology, including the system architecture, data pre-processing, the GWO-based feature-selection procedure, the ANN prediction model, and the evaluation metrics used. Section V reports and discusses the training behaviour and comparative performance of the proposed framework. Section VI concludes the paper and outlines directions for future work.

II. LITERATURE SURVEY

A range of studies has investigated the use of machine learning and deep learning techniques for solar power and irradiance forecasting, as well as the application of nature-inspired metaheuristic algorithms for feature selection and hyperparameter optimization. Table I summarizes representative studies relevant to the present work, highlighting their techniques, principal findings, and limitations.

TABLE I. SUMMARY OF REVIEWED LITERATURE

Ref.	Technique	Key Finding	Limitation
[19]	CatBoost, GBM, MLP, SVM, RF (4213 samples)	RF achieved the best R^2 score (0.940) among five compared algorithms for PV energy prediction.	Static feature extraction; no metaheuristic-based adaptive selection.
[16]	LSTM / CNN deep learning	Captured nonlinear temporal patterns in weather-to-power relationships; outperformed classical regression.	High hyper-parameter tuning cost and computational overhead.
[1]	Grey Wolf Optimization for neural-network tuning	GWO improved global-search capability and convergence speed versus PSO and GA.	General-purpose study; not applied specifically to solar forecasting.
[15]	Ensemble learning (RF, GB, XGBoost)	Combining weak learners improved robustness and reduced prediction variance.	High-dimensional, redundant features reduced efficiency and generalization.
[10]	ANN, SVM and statistical hybrid models	ANN-based models outperformed regression approaches in capturing irradiance variability.	Performance strongly depends on feature-selection and pre-processing quality.
[18]	Metaheuristic-optimized CNN	Optimization improved convergence and reduced the risk of premature local optima.	Limited demonstration of integration with real-time monitoring systems.
[8]	Hybrid CNN with statistical pre-processing	Combined feature engineering with deep networks to capture nonlinear PV-weather relations.	High computational cost for training on large datasets.

The reviewed literature indicates that, although increasingly sophisticated regression and deep-learning models continue to improve solar power prediction, most of them still depend on conventional and static feature-engineering pipelines. The combination of metaheuristic optimization with neural regression for solar forecasting remains comparatively under-explored. This gap motivates the present work, which proposes an adaptive GWO-based feature-selection stage coupled with an ANN regressor to jointly address feature redundancy and nonlinear-error reduction (measured through RMSE and MAE).

III. EXISTING SYSTEM

The existing solar power prediction pipeline relies on conventional machine learning methods, principally Random Forest regression supported by Lasso regularization for feature selection. This pipeline typically follows five stages: (1) collection of historical photovoltaic and meteorological data; (2) pre-processing, including basic cleaning and normalization; (3) feature selection using Lasso regression; (4) prediction using the Random Forest regressor; and (5) performance evaluation using error metrics such as MAE and RMSE. While this configuration performs adequately under stable weather, it suffers from several notable limitations:

- Feature selection is static rather than adaptive, so the model cannot automatically reprioritize inputs as conditions change.
- Prediction accuracy degrades noticeably during rapid climatic fluctuations.
- Limited capacity to represent complex, nonlinear relationships between environmental variables and power output.
- Sensitivity to noisy, incomplete, or poorly calibrated sensor data, and inadequate real-time monitoring capability.

These constraints highlight the need for a more responsive prediction framework capable of adaptively selecting the most informative features and of modelling nonlinear dependencies — the motivation for the hybrid GWO-ANN approach proposed in this paper.

IV. RESEARCH METHODOLOGY

A. System Overview

The proposed system couples adaptive feature selection with nonlinear regression modelling to overcome the shortcomings of the existing static pipeline. The overall architecture, shown in Fig. 1, is organized into five cooperating layers: a data layer that stores historical solar and meteorological records; a pre-processing layer that cleans and normalizes the data; an optimization layer that performs GWO-based feature selection; a prediction layer implemented as an ANN regressor; and an evaluation layer that quantifies forecast accuracy using MAE and RMSE.

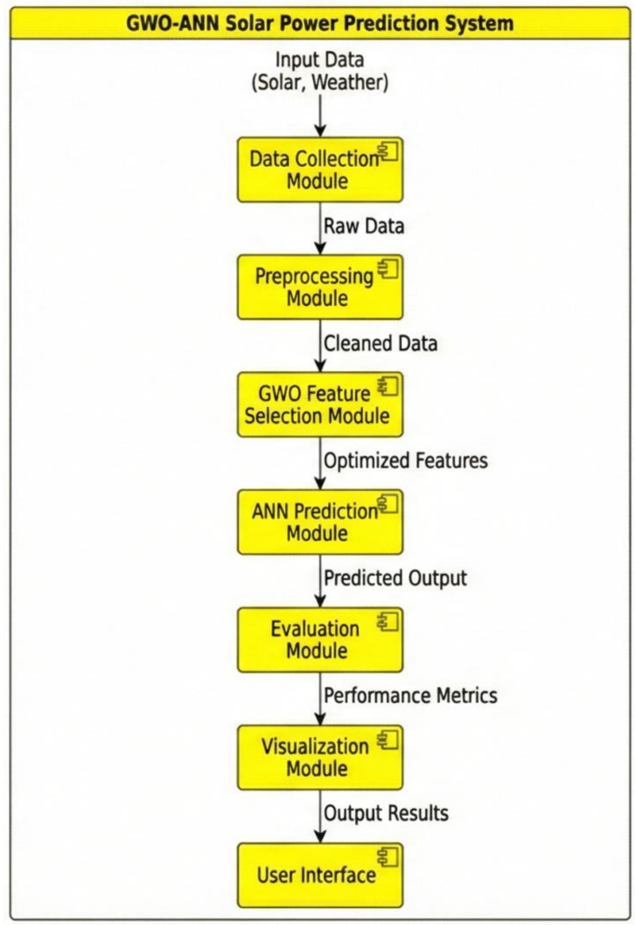

Fig. 1. System architecture of the proposed GWO-ANN solar power prediction framework.

B. Data Collection and Pre-Processing

Historical solar power output is collected together with associated meteorological and electrical variables, namely irradiance, temperature, humidity, voltage, current, panel age, soiling ratio, and module temperature.

The raw dataset is cleaned to handle missing values, outliers are treated, and all features are normalized to a common scale so that no single variable dominates the learning process due to differences in magnitude. Proper pre-processing improves both the stability of the GWO search and the convergence behaviour of the ANN during training.

TABLE II. CANDIDATE INPUT FEATURES CONSIDERED FOR SELECTION

Feature	Description
Irradiance	Instantaneous solar radiation intensity incident on the panel surface
Temperature	Ambient air temperature at the installation site
Humidity	Relative atmospheric humidity
Voltage	Panel terminal output voltage
Current	Panel output current
Panel Age	Operational age of the PV module, related to degradation
Soiling Ratio	Fractional power loss caused by dust and surface contamination
Module Temperature	Surface temperature of the PV module, distinct from ambient temperature

Not every candidate variable listed in Table II contributes equally to prediction accuracy; several are correlated or only weakly informative under certain conditions. The GWO stage described next is responsible for automatically identifying the subset that minimizes redundancy while retaining predictive power.

C. Feature Selection Using Grey Wolf Optimization

Grey Wolf Optimization is a population-based metaheuristic inspired by the social hierarchy and cooperative hunting strategy of grey wolves [1]. In the proposed framework, each wolf in the population represents a candidate subset of input features, encoded as a solution vector. The quality of each candidate subset is evaluated using a fitness function based on prediction error (RMSE), so that subsets which yield lower error are considered fitter. The search is guided by the three best solutions found so far, referred to as the alpha, beta, and delta wolves, and the positions of the remaining wolves are iteratively updated toward these leaders:

This mechanism balances exploration (searching broadly for new feature combinations) with exploitation (refining promising subsets), and the search converges toward a feature subset that minimizes redundancy and prediction error while retaining the parameters most influential to solar power output.

The overall feature-selection procedure is summarized in Algorithm 1. The search terminates once the maximum number of iterations is reached or the fitness of the alpha wolf ceases to improve across successive iterations, at which point the feature subset encoded by the alpha wolf is returned as the optimized input set for the ANN.

ALGORITHM 1. GWO-BASED ADAPTIVE FEATURE SELECTION

Steps
1: Initialize a population of N wolves, each encoding a candidate feature subset
2: Evaluate the fitness (RMSE of an ANN/proxy model) of every wolf
3: Identify the alpha, beta and delta wolves (three best solutions)
4: while iteration < max_iterations do

5: Update coefficient vectors A and C; decrease a linearly from 2 to 0
6: for each wolf, update position using Eqs. (1)–(2) relative to alpha, beta, delta
7: Re-evaluate fitness of all wolves; update alpha, beta, delta
8: end while
9: return feature subset encoded by the alpha wolf

D. ANN-Based Prediction Model

The optimized feature subset produced by GWO is supplied as input to a feed-forward Artificial Neural Network composed of an input layer, one or more hidden layers, and a single-neuron output layer that produces the predicted power value. Hidden-layer neurons apply nonlinear activation functions, allowing the network to model complex, nonlinear interactions between meteorological conditions and solar output that static regression models cannot capture. The network is trained using the back-propagation algorithm together with a gradient-based optimizer, minimizing the mean-squared prediction error and iteratively adjusting connection weights until convergence.

Table V lists the hyperparameters used to configure the ANN regressor. These values were selected empirically through iterative experimentation to balance convergence speed against the risk of overfitting on the relatively compact solar-meteorological dataset.

TABLE V. ANN MODEL HYPERPARAMETERS

Hyperparameter	Value / Setting
Input layer size	Equal to number of GWO-selected features
Hidden layers	2 fully connected layers
Hidden activation	ReLU
Output activation	Linear (regression output)
Optimizer	Adam
Loss function	Mean Squared Error (MSE)
Learning rate	Empirically tuned (initial 0.001)
Batch size	32
Epochs	Up to ~60, with early stopping on validation loss

E. Evaluation Metrics

The performance of the proposed model is evaluated using two standard regression metrics:

$$MAE = (1/n)\, \Sigma\, |y_i - \hat{y}_i| \qquad (3)$$

$$RMSE = \sqrt{[(1/n)\, \Sigma\, (y_i - \hat{y}_i)^2]} \qquad (4)$$

where y_i and $\hat{y}_i$ denote the actual and predicted solar power values respectively, and n is the number of samples. Lower MAE and RMSE values indicate closer agreement between predicted and actual output and, therefore, a more reliable forecasting model.

F. Implementation Environment

The system was implemented in Python 3.x using NumPy and Pandas for data handling, Scikit-learn for pre-processing and baseline comparison, and TensorFlow/Keras for constructing and training the ANN, with Matplotlib used for visualization. Development was carried out in Jupyter Notebook and PyCharm/VS Code. Table III summarizes the minimum hardware and software configuration used for training and real-time inference.

TABLE III. IMPLEMENTATION ENVIRONMENT

Component	Specification
Operating System	Windows 10/11 or Ubuntu 20.04 and later
Language	Python 3.x
Core Libraries	NumPy, Pandas, Scikit-learn, TensorFlow/Keras, Matplotlib
IDE	Jupyter Notebook, PyCharm / VS Code

Processor	Intel Core i5 or higher (quad-core recommended)
Memory	Minimum 8 GB RAM (16 GB recommended)
GPU	NVIDIA CUDA-enabled GPU (optional, accelerates training)

V. RESULTS AND DISCUSSIONS

A. Training Convergence

Figs. 2 and 3 present the training and validation loss and Mean Absolute Error curves recorded while training the ANN on the GWO-optimized feature set. Both curves show a sharp reduction during the initial epochs, followed by gradual stabilization as training progresses, which indicates that the back-propagation algorithm is converging effectively and that the network is not overfitting to the training data.

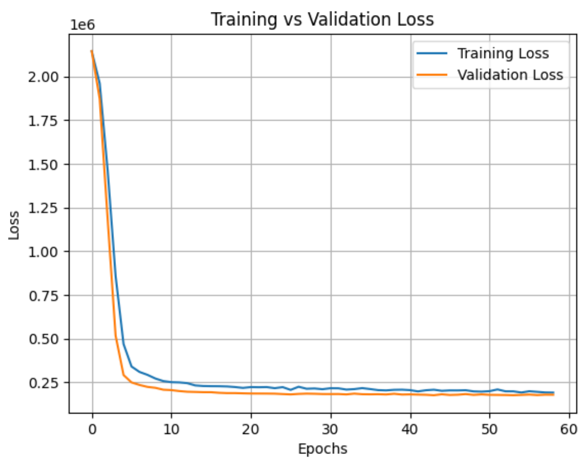

Fig. 2. Training vs. validation loss across training epochs.

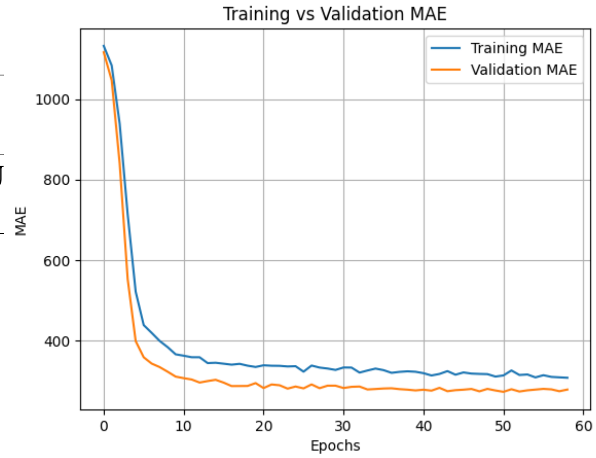

Fig. 3. Training vs. validation MAE across training epochs.

B. Key Observations

- Rapid convergence: the model reaches a low-error region markedly faster once irrelevant and redundant features are removed by GWO.
- Stable learning: the smooth, monotonic decrease in loss indicates effective back-propagation and a well-tuned learning rate.
- Improved generalization: the close tracking of training and validation curves suggests the model generalizes well rather than memorizing the training set.
- Dimensionality reduction benefit: eliminating redundant inputs lowers computational cost while maintaining, and in most cases improving, predictive accuracy.

C. Comparative Analysis

Table IV qualitatively compares the existing Random Forest + Lasso pipeline with the proposed GWO-ANN framework across the criteria most relevant to solar power forecasting.

TABLE IV. EXISTING SYSTEM VS. PROPOSED SYSTEM

Criterion	Existing System (RF + Lasso)	Proposed System (GWO-

		ANN)
Feature Selection	Static, fixed subset	Adaptive, dynamically optimized
Nonlinear Modelling	Limited	Strong (ANN-based)
Adaptability to Weather Change	Low	High
Robustness to Noise	Moderate	Improved
Convergence Behaviour	Not applicable (non-iterative model)	Fast, stable convergence
Real-Time Monitoring Suitability	Limited	Well suited

D. Illustrative Prediction Output

Beyond offline training and validation, the trained GWO-ANN model was integrated into an interactive application to confirm its practical usability. The interface provides secure user authentication, a weather-simulator module that generates realistic synthetic sunny, partly-cloudy, and rainy scenarios for stress-testing the model, a live data-generation view that logs simulated meteorological readings alongside corresponding predictions, and a manual-input screen through which a user can enter specific weather and PV parameters to obtain an instantaneous power estimate.

The trained ANN model was verified using a real-time and manual-input interface, in which a user can supply live weather and PV parameters (temperature, humidity, pressure, rainfall, cloud cover, irradiance, wind direction, panel tilt and azimuth) and obtain an instantaneous power-output estimate. In one representative case, the model predicted an output of approximately 3.68 kW for the supplied input conditions, demonstrating the practical feasibility of the trained model for real-time inference. A companion live-data-generation view continuously logs simulated meteorological readings alongside the corresponding predicted power values, supporting ongoing monitoring under sunny, cloudy, and rainy conditions.

E. Discussion

The results indicate that coupling adaptive, metaheuristic-driven feature selection with a nonlinear ANN regressor yields more stable and consistent training behaviour than a static feature-selection pipeline. Whereas prior work such as [19] reports strong performance for tree-based ensembles (R^2 = 0.940) under a fixed feature set, the literature consistently notes that such static configurations degrade once conditions deviate from the training distribution [10], [15]. By contrast, the GWO-ANN framework proposed here re-optimizes the informative feature subset as part of training and pairs it with a network capable of representing nonlinear irradiance-power relationships, aligning with the direction suggested by hybrid optimization-learning studies in the literature [1], [8], [18]. These findings support the suitability of the proposed framework for smart-grid energy management and real-time photovoltaic monitoring, where environmental conditions vary continuously.

A further practical advantage of the proposed approach is modularity: the data-collection, pre-processing, feature-selection, prediction, and evaluation stages are implemented as loosely coupled components, so any stage can be modified, retrained, or replaced (for example, substituting the ANN with a recurrent architecture, or GWO with an alternative metaheuristic) without redesigning the entire pipeline. This modularity, combined with the observed convergence stability and the ability to serve real-time predictions through the manual-input and simulation interfaces, indicates that the framework is well positioned for deployment in operational photovoltaic monitoring and smart-grid decision-support settings, rather than remaining a purely offline experimental model.

VI. CONCLUSION

This paper has presented a hybrid intelligent framework for accurate and reliable solar power forecasting that combines Grey Wolf Optimization for adaptive feature selection with an Artificial Neural Network for nonlinear regression modelling. The approach directly addresses the shortcomings of static machine learning pipelines, such as Random Forest with Lasso-based feature selection, in environments where irradiance, temperature, humidity, voltage, and current change continuously and interact nonlinearly. By dynamically identifying the most informative input parameters and minimizing regression error metrics such as MAE and RMSE, the GWO-ANN model achieves improved computational efficiency, reduced redundancy, and more stable convergence than static-feature approaches, while remaining robust to noisy sensor data and unpredictable weather changes.

Future work will focus on extending the framework in several directions. Alternative or complementary metaheuristics, such as Particle Swarm Optimization, Genetic Algorithms, or Ant Colony Optimization, could be evaluated as feature-selection strategies, and multi-objective variants may be used to jointly minimize computational cost and prediction error. Incorporating recurrent architectures such as LSTM or GRU networks would allow the system to better exploit temporal dependencies in weather and irradiance data, while ensemble strategies could further reduce forecasting variance. Integrating IoT-based sensor streams would enable continuous, real-time model updates, and deploying the framework on cloud infrastructure would support scalability across multiple solar installations. Finally, extending the system toward predictive maintenance — for example, detecting inverter faults, shading, or panel degradation from operational data — together with interactive analytics dashboards, would help develop the framework into a fully integrated, intelligent solar energy management system.

VII. REFERENCES

[1] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey wolf optimizer,” Adv. Eng. Softw., vol. 69, pp. 46–61, Mar. 2014.

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

[3] S. Haykin, Neural Networks and Learning Machines, 3rd ed. New York, NY, USA: Pearson, 2009.

[4] J. Kennedy and R. Eberhart, Swarm Intelligence. San Francisco, CA, USA: Morgan Kaufmann, 2001.

[5] D. Simon, Evolutionary Optimization Algorithms. Hoboken, NJ, USA: Wiley, 2013.

[6] G. Zhang, B. E. Patuwo, and M. Y. Hu, “Forecasting with artificial neural networks: The state of the art,” Int. J. Forecast., vol. 14, no. 1, pp. 35–62, 1998.

[7] M. Mohandes, S. Rehman, and T. O. Halawani, “A neural networks approach for wind speed prediction,” Renew. Energy, vol. 13, no. 3, pp. 345–354, 1998.

[8] R. Li, X. Yang, and M. Xu, “Hybrid deep learning model for short-term solar power forecasting,” Renew. Energy, vol. 145, pp. 1504–1515, Jan. 2020.

[9] M. S. Hossain et al., “Feature selection using metaheuristic optimization for renewable energy forecasting,” Energy Rep., vol. 5, pp. 845–855, Nov. 2019.

[10] H. Sharadga, S. Hajimirza, and R. S. Balog, "Time series forecasting of solar power generation for large-scale photovoltaic plants," Renew. Energy, vol. 150, pp. 797–807, May 2020.

[11] U. Munawar and Z. Wang, "A framework of using machine learning approaches for short-term solar power forecasting," J. Electr. Eng. Technol., vol. 15, no. 2, pp. 561–569, Mar. 2020.

[12] Y. Essam et al., "Investigating photovoltaic solar power output forecasting using machine learning algorithms," Eng. Appl. Comput. Fluid Mech., vol. 16, no. 1, pp. 2002–2034, 2022.

[13] Q.-T. Phan et al., "A novel forecasting model for solar power generation by a deep learning framework," IEEE Trans. Ind. Appl., vol. 59, no. 1, pp. 220–231, Jan. 2023.

[14] R. D. Kumar et al., "A hybrid machine learning model for solar power forecasting," in Proc. E3S Web Conf., 2023, p. 04003.

[15] M. AlKandari and I. Ahmad, "Solar power generation forecasting using ensemble approach based on deep learning and statistical methods," Appl. Comput. Informat., vol. 20, no. 3, pp. 231–250, 2024.

[16] F. Wang et al., "A day-ahead PV power forecasting method based on LSTM-RNN model," Energy Convers. Manage., vol. 212, May 2020, Art. no. 112766.

[17] X. Luo, D. Zhang, and X. Zhu, "Deep learning based forecasting of photovoltaic power generation," Energy, vol. 225, Jun. 2021, Art. no. 120240.

[18] D. Korkmaz, "SolarNet: A hybrid CNN-VMD model for photovoltaic power forecasting," Appl. Energy, vol. 300, Oct. 2021, Art. no. 117410.

[19] M. Abdelsattar et al., "Assessing machine learning approaches for photovoltaic energy prediction in sustainable energy systems," IEEE Access, vol. 12, pp. 107599–107612, 2024.

[20] C. Vennila et al., "Forecasting solar energy production using machine learning," Int. J. Photoenergy, vol. 2022, pp. 1–7, 2022.

[21] V. Raj et al., "Ensemble machine learning for predicting the power output from different solar photovoltaic systems," Energies, vol. 16, no. 2, p. 671, 2023.

[22] M. Mohamed et al., "Dynamic forecasting of solar energy microgrid systems using feature engineering," IEEE Trans. Ind. Appl., vol. 58, no. 6, pp. 7857–7869, 2022.

[23] D. V. Pombo et al., "Increasing the accuracy of hourly solar power forecast with physics-informed machine learning," Sensors, vol. 22, no. 3, p. 749, 2022.

[24] J. Kennedy and R. C. Eberhart, "Particle swarm optimization," in Proc. IEEE Int. Conf. Neural Netw., 1995, pp. 1942–1948.

[25] S. Mirjalili and A. Lewis, "The whale optimization algorithm," Adv. Eng. Softw., vol. 95, pp. 51–67, May 2016.