Journal of Engineering and Technology for Industrial Applications

ITEGAM-JETIA

Manaus, v.12 n.57, p. 1229-1239. January/February, 2026.
DOI: https://doi.org/10.5935/jetia. v12i57.3152

ISSN ONLINE: 2447-0228

RESEARCH ARTICLE OPEN ACCESS

ARTIFICIAL NEURAL NETWORK BASED PREDICTIVE PERFORMANCE ANALYSIS OF PHOTOVOLTAIC OUTPUT UNDER VARIOUS SOLAR RADIATION RANGES

Ravshanbek Rakhmatulaev*[1], Rodney Tan[2], Aliev Rayimjon[3], Yu L. J.[4]

[1, 4]Department of Mechanical Engineering, Faculty of Engineering, UCSI University, 56000 Kuala Lumpur, Malaysia.
[2]Department of Electrical and Electronic Engineering, Faculty of Engineering, UCSI University, 56000 Kuala Lumpur, Malaysia.
[3]Department of Physics, Faculty of Physics and Mathematics, Andijan State University, Andijan 170100, Uzbekistan.

[1]http://orcid.org/0009-0008-4799-5689, [2]http://orcid.org/0000-0002-2364-6088
[3]http://orcid.org/0000-0002-7375-727X, [4]http://orcid.org/0000-0002-8195-1566

Email: *rraxmatullayev11@gmail.com, rodneytan@ucsiuniversity.edu.my, alievuz@yahoo.com, yulj@ucsiuniversity.edu.my

ARTICLE INFO

Article History
Received: December 16, 2025
Revised: January 10, 2026
Accepted: January 15, 2026
Published: February 28, 2026

Keywords:
Solar energy
Photovoltaic power
Prediction
FFBPNN

ABSTRACT

Accurate assessment of photovoltaic (PV) power output under varying environmental conditions is essential for evaluating solar energy performance and optimizing its application. This study employs a Feedforward back propagation neural network (FFBPNN) approach to predict the output power of a solar panel using a combination of electrical and weather-related parameters. The input variables include short-circuit current (I_{sc}), open-circuit voltage (V_{oc}), maximum voltage (V_{max}), maximum current (I_{max}), efficiency (EFF) - represents how effectively it converts incident solar energy into electrical energy., fill factor (FF) - the ratio of the maximum obtainable power to the product of the open-circuit voltage and short-circuit current, solar radiation (G), wind speed (v), ambient temperature (T_{am}), and panel temperature (T_{panel}) under solar radiation intensity in the ranges of 200–400, 400–600, 600–800, and 800–1000 W/m². Three artificial neural network (ANN) models, ANN-10 (10 inputs), ANN-8 (8 inputs, except I_{sc} and V_{oc}), and ANN-6 (6 inputs, except I_{sc}, V_{oc}, V_{max} and I_{max}), are developed for 496 dataset and compared based on their prediction accuracy using the coefficient of correlation (R) by varying a number of neurons in the hidden layer from 1 to 15. Among them, the ANN-8 model (8-10-1) demonstrates the highest prediction accuracy (R value of training, testing, validation and overall model exceeding 0.999) particularly at high solar radiation levels (800–1000 W/m²). The results confirm that FFBPNN-based modelling is a reliable and effective tool for forecasting PV output power under diverse operating conditions. This approach supports solar energy potential assessment and performance optimization across various geographical and climatic contexts.

I. INTRODUCTION

In response to growing global energy demand, environmental degradation, and the urgent need to reduce reliance on fossil fuels, solar energy has become a cornerstone of sustainable development. Among renewable energy sources, solar power is particularly attractive due to its abundance, scalability, and minimal environmental impact [1]. Photovoltaic (PV) technology, which converts sunlight directly into electricity, plays a central role in harnessing this resource [2]. The increasing integration of PV systems across residential, commercial, and utility-scale applications reflects a global shift toward clean energy. However, the intermittent and variable nature of solar irradiance, shaped by geographic, meteorological, and temporal factors, poses significant challenges to the consistent and efficient operation of PV systems [3]. Accurate forecasting and modelling of PV output are thus critical for grid stability, energy planning, and operational efficiency [4].

PV system performance is governed by complex, nonlinear relationships between environmental parameters (e.g., solar irradiance, temperature, wind speed, humidity) and electrical outputs (e.g., voltage, current, power) [5]. Predicting output under varying conditions is vital not only for system monitoring and control but also for optimizing energy storage, demand response, and economic dispatch, especially in smart grids and hybrid energy systems. Despite its importance, accurately modelling PV behaviour remains difficult due to the stochastic and nonlinear nature of these influencing factors [6]. Traditional modelling approaches, empirical, analytical, and physics-based, have laid the foundation for PV system analysis. These methods often rely on simplified assumptions and require parameter calibration, which may not generalize well under dynamic real-world conditions [7]. For instance, the widely used single-diode model requires parameters that could be hard to measure or that vary over time. Similarly, regression-based methods assume fixed functional relationships, limiting their ability to capture the nonlinear dynamics of PV performance. These limitations highlight the need for more adaptable, data-driven approaches [8].

In this context, machine learning (ML) has emerged as a promising alternative for modelling complex systems, including PV output prediction. Among ML techniques, Artificial Neural Networks (ANNs) are particularly well-suited for this task due to their ability to model nonlinear relationships, manage noisy data, and generalize from historical patterns [9]. Modelled after the human brain, ANNs comprise layers of interconnected neurons that transform input features through weighted connections and activation functions to generate outputs. This structure allows ANNs to approximate complex functions without requiring predefined mathematical models [10]. The application of ANNs to PV modelling represents a shift from equation-driven to data-driven methodologies [11]. Leveraging large volumes of operational and meteorological data, often collected by sensors and monitoring systems, ANNs adapt to different PV technologies, configurations, and climate conditions. They could also be integrated with other intelligent systems such as fuzzy logic, genetic algorithms, or hybrid ML frameworks to further improve prediction accuracy and robustness [12]. Numerous studies have validated the effectiveness of ANNs in predicting PV performance [13- 15].

For instance, a Levenberg-Marquardt back-propagation ANN model with six input parameters and a hidden layer of up to 35 neurons achieved high accuracy in output prediction [16]. Other studies compared radial basis and feed-forward topologies using different training algorithms, finding mean bias errors below 1% and prediction errors between 0.05% and 1% [17-20]. In another case, an MLP model with seven inputs predicted maximum voltage (V_m) and current (I_m) with high precision using a single hidden layer of up to ten neurons [21]. A feed-forward ANN with two hidden layers, trained on six meteorological variables, achieved R^2 values as high as 99.9% when predicting mono-crystalline PV output [22]. Mellit et al. (2013) demonstrated that even simple ANN structures could effectively predict PV output. Using one year's worth of data for a 50 W_p module, their model, with just a single hidden neuron, achieved determination coefficients of 96–97% on sunny days and 93–97% on cloudy days [23]. Compared to traditional techniques, ANN-based models consistently offer superior accuracy in capturing the effects of environmental variability, shading, and system degradation.

They also provide practical benefits in terms of implementation speed and computational efficiency, especially with modern hardware and software capabilities [24]. Nevertheless, the effectiveness of ANN models depends on factors such as network architecture, input feature selection, training algorithms, and data quality [25]. Despite their promise, ANN-based models in PV energy forecasting often face challenges such as overfitting, convergence instability, and limited interpretability. Recent advances, however, are steadily overcoming these limitations. Regularization strategies, ensemble learning, and data augmentation enhance generalization, while adaptive optimizers, metaheuristic-assisted training, and advanced initialization improve convergence speed and stability. Furthermore, explainable AI tools, attention mechanisms, and hybrid physics-informed networks enhance transparency and trust in predictions. Together, these developments expand the reliability and applicability of ANN models, enabling more accurate forecasting, fault detection, and system optimization, thereby strengthening their role in modern PV energy research and grid integration [26]. Despite the demonstrated success of ANN-based models in PV output prediction, several research gaps remain.

Many existing studies focus on limited datasets, specific PV technologies, or narrow environmental conditions, which restrict the generalizability and robustness of the models [24], [26]. Furthermore, there is a lack of systematic comparison between different ANN architectures in terms of prediction accuracy, training efficiency, and sensitivity to input features. Additionally, few studies explore the integration of real-time meteorological data or assess model performance across diverse climatic zones [25]. These limitations highlight the need for more comprehensive, adaptable, and scalable ANN frameworks tailored to dynamic, real-world conditions. The present study aims to address these gaps by developing a data-driven ANN model that leverages diverse environmental inputs and evaluates multiple neural network configurations to optimize PV power output prediction. The novelty of this work lies in its systematic exploration of ANN architectures under varying climatic conditions, the inclusion of a broader set of input features, and the evaluation of model performance across different PV technologies. This approach not only enhances prediction accuracy but also contributes to the development of more reliable and transferable forecasting tools for modern energy systems.

Hence, this study aims to develop and evaluate an ANN model for accurately predicting the output power of photovoltaic (PV) systems under varying environmental conditions. The primary objective is to investigate the impact of different electrical and meteorological input parameters on the accuracy of PV power prediction. Specifically, the study seeks to: (i) assess the predictive performance of ANN models using a combination of electrical parameters (short-circuit current (I_{sc}), open-circuit voltage (V_{oc}), maximum voltage (V_{max}), maximum current (I_{max}), efficiency (EFF), fill factor (FF)) and environmental factors (solar radiation, wind speed, ambient temperature, and panel temperature); (ii) analyse the effect of different solar radiation intensity ranges on model accuracy; (iii) optimize ANN architecture by varying the number of input features and hidden layer neurons; and (iv) compare the prediction accuracy of three ANN models (ANN-10, ANN-8, and ANN-6) using the coefficient of correlation (R). The findings aim to support the application of ANN-based modelling for PV performance evaluation and solar energy forecasting across diverse climatic conditions.

II. METHODOLOGY

This study develops an ANN model to predict the output power of a photovoltaic (PV) module under varying environmental conditions. The model is trained and evaluated using field experimental data collected over a one-month period at UCSI University Field Campus, Kuala Lumpur, Malaysia (Latitude: 3.079548° N, Longitude: 101.733216° E, Elevation: 70 m). The site experiences a tropical climate, making it suitable for solar energy potential assessment and performance prediction.

II.1 PV MODULE SPECIFICATIONS AND MEASUREMENT INSTRUMENTS

The solar panel model ZDNY-250C60 is made of Monocrystal (a-Si) material. The panel size is 1650×992×40 mm, and it consists of 60 cells. Its maximum power output is 250 W. The open-circuit voltage (V_{oc}) is 37.85 V, the short-circuit current (I_{sc}) is 8.40 A, the voltage at maximum power (V_{max}) is 31.17 V, and the current at maximum power (I_{max}) is 8.03 A. The maximum system voltage is 1000 V, and the maximum fuse current is 15 A. The output tolerance is 0/+3%, and the measurements were taken at a temperature of 25°C. Each cell has a dimension of 125×125 mm Figure (1).

II.2 DATA COLLECTION AND INPUT FEATURES

Data were collected using the experimental setup shown in Fig.(1), over a month period from 25 June to 26 July 2025, with measurements taken daily between 9:00 a.m. and 5:00 p.m. A PROVA 200A I-V Tracer Solar Module Analyzer was used to measure Open circuit voltage (V_{oc}), Short-Circuit current (I_{sc}), Voltage at max power (V_{mp}), and Current at max power (I_{mp}) capture the electrical characteristics of the PV module. The surface temperature of the panel was measured using a FLUKE 59 Mini IR Thermometer, while ambient temperature was recorded via a wired Mercury thermometer. Wind speed was monitored using a handheld anemometer. The following parameters were recorded at regular intervals: (i) Electrical parameters: I_{sc}, V_{oc}, V_{max}, I_{max}, Efficiency (EFF), Fill Factor (FF), (ii) Environmental parameters: Solar Radiation (G), Wind Speed (v), Ambient Temperature (T_{am}), Panel Temperature (T_{panel}), and (iii) The PV output power served as the target (output) variable for the ANN model. To evaluate model performance under different irradiance conditions, solar radiation data were grouped into four categories: 200–400, 400–600, 600–800, and 800–1000 W/m².

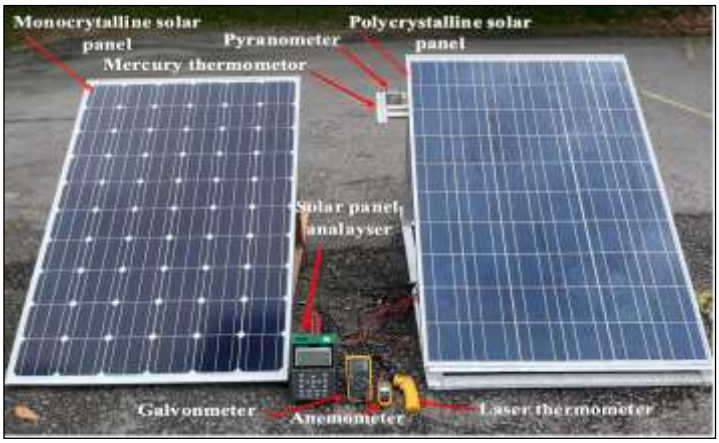

Figure 1: Experimental set-up used in the present study.
Source: Authors, (2026).

II.3 ANN MODELLING

Artificial Neural Network (ANN) is a computational modelling framework inspired by the structure of the human brain, consisting of interconnected processing units (neurons) arranged in three layers: input, hidden, and output [27]. In this study, a feed-forward Multilayer Perceptron (MLP) architecture was employed Figure (2), implemented in MATLAB 2024. Each input parameter is supplied to a neuron in the input layer, processed through synaptic weights and activation functions within the hidden layer, and finally transformed into the predicted output in the output layer [28].

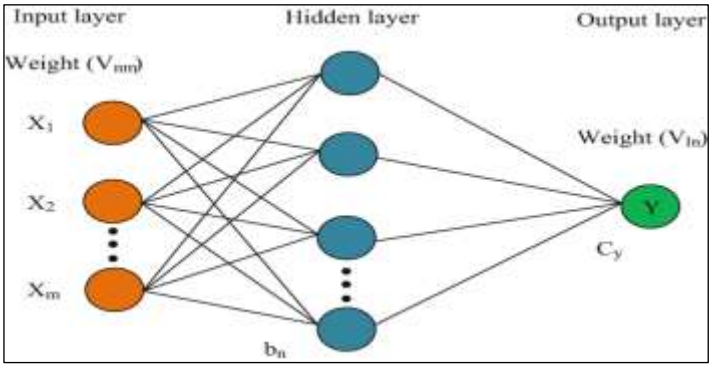

Figure 2: General architecture of ANN.
Source: Authors, (2026).

The dataset containing 40 entries was divided into 70% training, 15% testing, and 15% validation. Before training, all input parameters were normalized to a [0, 1] range using the Eq. (1):

$$X_i = \frac{0.8}{d_{max} - d_{min}} \left(d_j - d_{min}\right) + 0.1 \quad (1)$$

Where d_{max}, d_{min}, and d_i are the maximum, minimum, and current values of the dataset, respectively. The ANN model was trained using the Levenberg–Marquardt algorithm, known for its fast convergence. The common activation functions used in ANN are provided in Eq. (2). The sigmoid activation function was used in the hidden layer to introduce nonlinearity and enable the ANN to learn the complex nonlinear relationships inherent in solar radiation and PV electrical behaviour. The output layer used a linear activation function because the target variable (PV output power) is continuous and unbounded, requiring an unrestricted output range. This sigmoid–linear combination is the standard architecture for regression tasks and ensures both stable training and accurate numerical prediction.

$$f(s) = \begin{cases} S\ linear\ function \\ \frac{1}{1+e^{-s}}\ \ sigmoid\ function \\ \frac{e^{-s}}{e^{+s}+e^{-s}}\ hyerbolic\ tangent\ function \end{cases} \quad (2)$$

The error function, as given in Eq. (3), to be minimized during training is given in by:

$$E = \frac{1}{2}\sum_{k=1}^{l}(\overline{y_k} - y_k)^2 \quad (3)$$

The ANN models were evaluated using Coefficient of Correlation (R), which measures the strength of the linear relationship between predicted and actual output. To analyse the influence of different input combinations(in shown Table 1), three ANN models were developed:

Table 1: Different input combinations used in the ANN models.

	Model	Input parameters
1	ANN-10	I_{sc}, V_{oc}, V_{max}, I_{max}, EFF, FF, G, v, T_{am}, T_{panel}
2	ANN-8	V_{max}, I_{max}, EFF, FF, G, v, T_{am}, T_{panel}
3	ANN-6	EFF, FF, G, v, T_{am}, T_{panel}

Source: Authors, (2026).

For each model, the number of neurons in the hidden layer was varied between 1 and 15, and the architecture was trained multiple times to identify the configuration that yielded the best pre curacy. The methodology of the ANN model is illustrated in Figure 3.

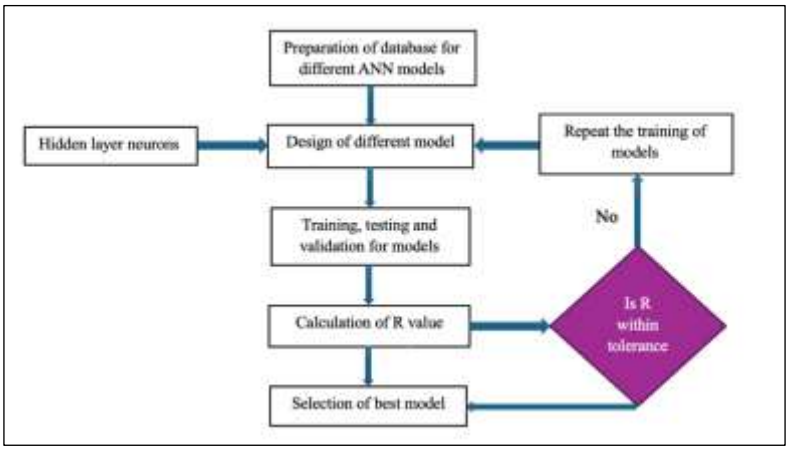

Figure 3: Implementation of methodology for ANN in the present study.
Source: Authors, (2026).

III. RESULTS

This manuscript investigates the effect of different multilayer perceptron (MLP) architectures on the performance of three artificial neural network (ANN) models for predicting solar radiation across various intensity ranges (200-1000 W/m²). The study evaluates the influence of the number of neurons in the hidden layer on the correlation coefficient (R), which measures the accuracy of model predictions. The analysis reveals that ANN-10 demonstrates consistent and high predictive performance across multiple neuron configurations, particularly excelling in the 800-1000 W/m² range Fig(4). It maintains high R-values (mostly above 0.95), reflecting stability, robustness, and optimal model complexity. ANN-8, consistently achieves high and stable R-values, indicating superior generalization and minimal oscillations across neuron configurations. In contrast, ANN-6 exhibits significant fluctuations and instability across all ranges, suggesting higher sensitivity to network structure and potential overfitting or underfitting issues. These irregular trends highlight the challenges of model complexity in this configuration.

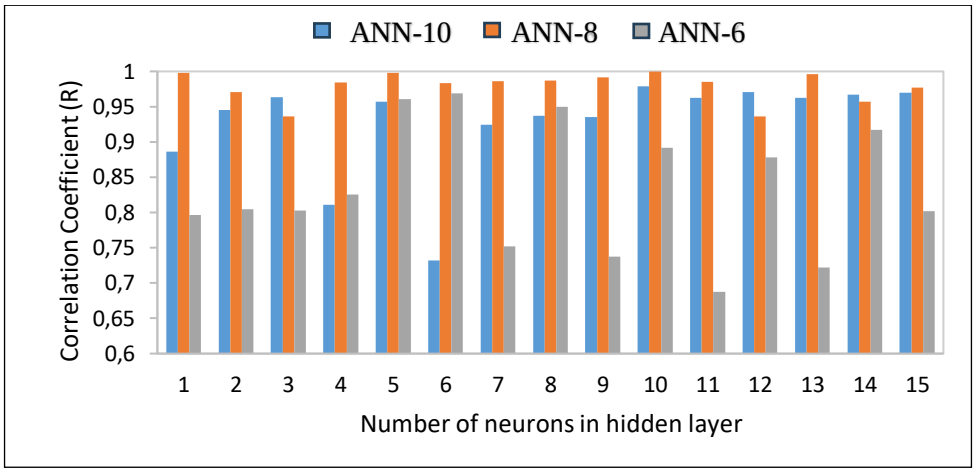

Figure 4: Effect of MLP architecture on ANN fitting performance for three models of solar radiation range (800-1000 W/m^2).
Source: Authors, (2026).

Overall, the findings suggest that while ANN-10 and ANN-8 perform well across specific ranges, ANN-10 stands out for its consistent high performance in the 800-1000 W/m² range, making it a strong candidate for solar radiation modelling in these higher intensity ranges. Figure (5) presents the effect of multilayer perceptron (MLP) architecture, specifically the number of neurons in the hidden layer, on the correlation coefficient (R) for three artificial neural network (ANN) models: ANN-10, ANN-8, and ANN-6, developed for predicting solar radiation within the range of 600-800 W/m². The correlation coefficient (R) quantifies the agreement between predicted and measured solar radiation, with values approaching unity indicating higher model accuracy and generalization capability.

As shown in the Fig (5), all three ANN configurations exhibit variations in correlation coefficient values with changes in the number of neurons in the hidden layer, reflecting the sensitivity of model performance to network complexity. The ANN-10 model (blue line) displays relatively stable and high R-values across most configurations, predominantly ranging between 0.90 and 1.00, with minor fluctuations at specific neuron counts. The ANN-8 model (orange line) also achieves strong predictive performance but exhibits more pronounced oscillations, with R-values varying between 0.80 and 1.00 depending on the number of neurons. In contrast, the ANN-6 model (green line) shows significant instability and broader fluctuations in R, ranging between 0.70 and 0.98, suggesting higher sensitivity to the chosen network structure and potential overfitting or underfitting at certain configurations.

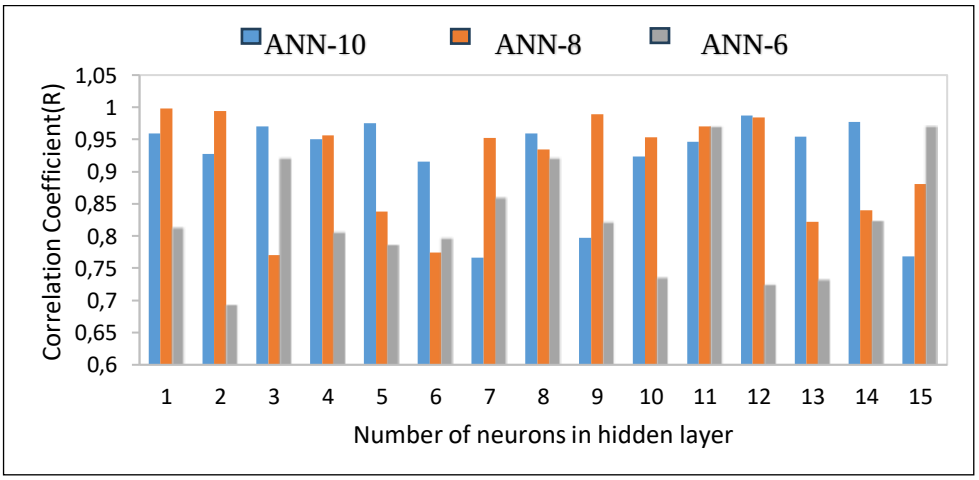

Figure 5: Effect of MLP architecture on ANN fitting for three models of solar radiation range (600-800 W/m^2).
Source: Authors, (2026).

Based on the observed R-values, the ANN-6 model demonstrates the best overall performance, maintaining high correlation coefficients (mostly above 0.95) across a wide range of hidden-layer neuron counts. This indicates superior stability, robustness, and generalization in fitting the nonlinear relationships governing solar radiation within the specified range. The consistently high R-values imply that ANN-6 achieved an optimal balance between model complexity and learning efficiency, resulting in minimized prediction errors and better adaptability to unseen data. Conversely, the irregular trends observed in ANN-8 and ANN-6 suggest less consistent learning behaviour and potential over-dependence on the number of hidden neurons.

Figure (6) presents the effect of multilayer perceptron (MLP) architecture specifically the number of neurons in the hidden layer on the correlation coefficient (R) for three artificial neural network (ANN) models: ANN-10, ANN-8, and ANN-6, developed for predicting solar radiation within the range of 400-600 W/m². The correlation coefficient (R) quantifies the agreement between predicted and measured solar radiation, with values approaching unity indicating higher model accuracy and generalization capability. As shown in the figure, all three ANN configurations exhibit variations in correlation coefficient values with changes in hidden-layer neuron numbers, reflecting the sensitivity of model performance to network complexity.

The ANN-10 model (blue line) displays relatively stable and high R-values across most configurations, predominantly ranging between 0.90 and 1.00, with minor fluctuations at specific neuron counts. The ANN-8 model (orange line) also achieves strong predictive performance but exhibits more pronounced oscillations, with R-values varying between 0.80 and 1.00 depending on the number of neurons. In contrast, the ANN-6 model (green line) shows significant instability and broader fluctuations in R, ranging between 0.70 and 0.98, suggesting higher sensitivity to the chosen network structure and potential overfitting or underfitting at certain configurations.

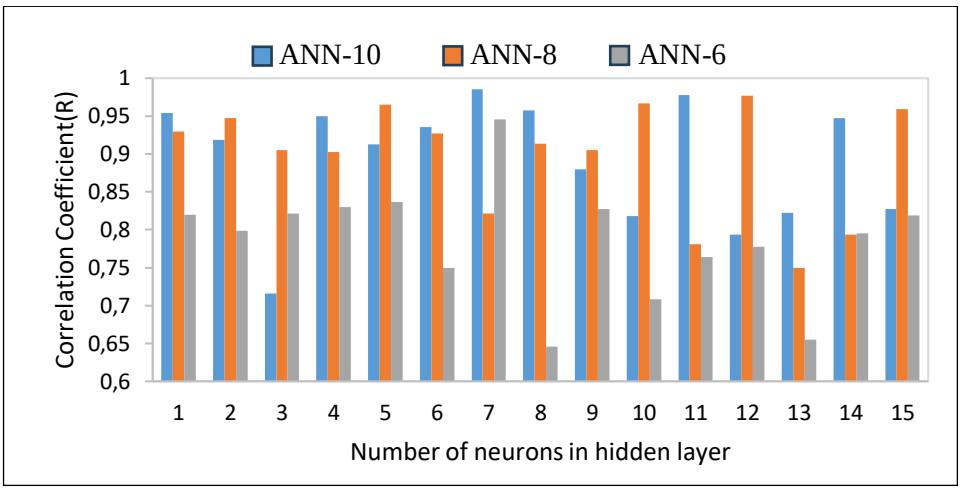

Figure 6: Effect of MLP architecture on ANN fitting for three models of solar radiation range (400-600 W/m^2).
Source: Authors, (2026).

Based on the observed R-values, the ANN-10 model demonstrates the best overall performance, maintaining high correlation coefficients (mostly above 0.95) across a wide range of hidden-layer neuron counts. This indicates superior stability, robustness, and generalization in fitting the nonlinear relationships governing solar radiation within the specified range. The consistently high R-values imply that ANN-10 achieved an optimal balance between model complexity and learning efficiency, resulting in minimized prediction errors and better adaptability to unseen data. Conversely, the irregular trends observed in ANN-8 and ANN-6 suggest less consistent learning behaviour and potential over-dependence on the number of hidden neurons.

Figure (7) presents the effect of multilayer perceptron (MLP) architecture specifically, the number of neurons in the hidden layer on the correlation coefficient (R) for three artificial neural network (ANN) models: ANN-10, ANN-8, and ANN-6, developed for predicting solar radiation within the range of 200-400 W/m². The correlation coefficient (R) represents the strength of the relationship between predicted and measured solar radiation values, where values closer to 1 indicate superior prediction accuracy and stronger model generalization. As shown in the figure, all three ANN models exhibit noticeable variations in R-values as the number of hidden-layer neurons increases, reflecting the sensitivity of model performance to architectural complexity.

The ANN-8 model (orange line) consistently demonstrates the highest and most stable R-values across nearly all configurations, maintaining correlations predominantly above 0.95, with minimal fluctuations. This trend indicates that ANN-8 effectively captures the nonlinear patterns in the solar radiation data while avoiding overfitting or underfitting. The ANN-R10 model (blue line) shows moderate performance, with R-values generally ranging between 0.80 and 0.95. Although it achieves strong correlations at certain neuron counts (e.g., around 5–6 and 10–11 neurons), the performance exhibits greater oscillations, suggesting sensitivity to the network structure and possible inconsistencies in generalization when the architecture changes. The ANN-6 model (green line), by contrast, demonstrates the lowest and most unstable R-values, fluctuating between 0.55 and 0.95, which implies weaker learning stability and a higher likelihood of overfitting or underfitting across different hidden-layer configurations.

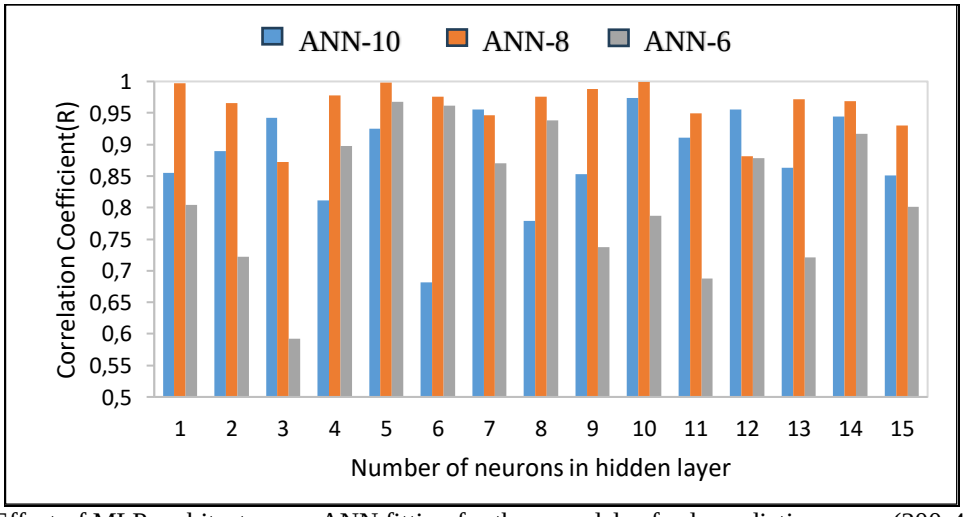

Figure 7: Effect of MLP architecture on ANN fitting for three models of solar radiation range (200-400 W/m^2).
Source: Authors, (2026).

Based on the observed results, the ANN-8 model is identified as the best-performing MLP architecture for this solar radiation range. Its consistently high R-values and minimal oscillations indicate superior predictive robustness, stability, and adaptability across varying hidden-layer neuron counts. This suggests that ANN-8 achieved an optimal balance between model complexity and learning efficiency, leading to more reliable and accurate solar radiation predictions.

VI. DISCUSSION

For manuscript discussion, these findings highlight that while all models are influenced by hidden-layer neuron variations, ANN-8's architecture demonstrates the most reliable and generalizable learning behaviour. The stability of its correlation coefficients indicates that it effectively captures the underlying physical and nonlinear dependencies of solar radiation within the 200-400 W/m² range, making it the most suitable architecture for practical modelling and forecasting applications.

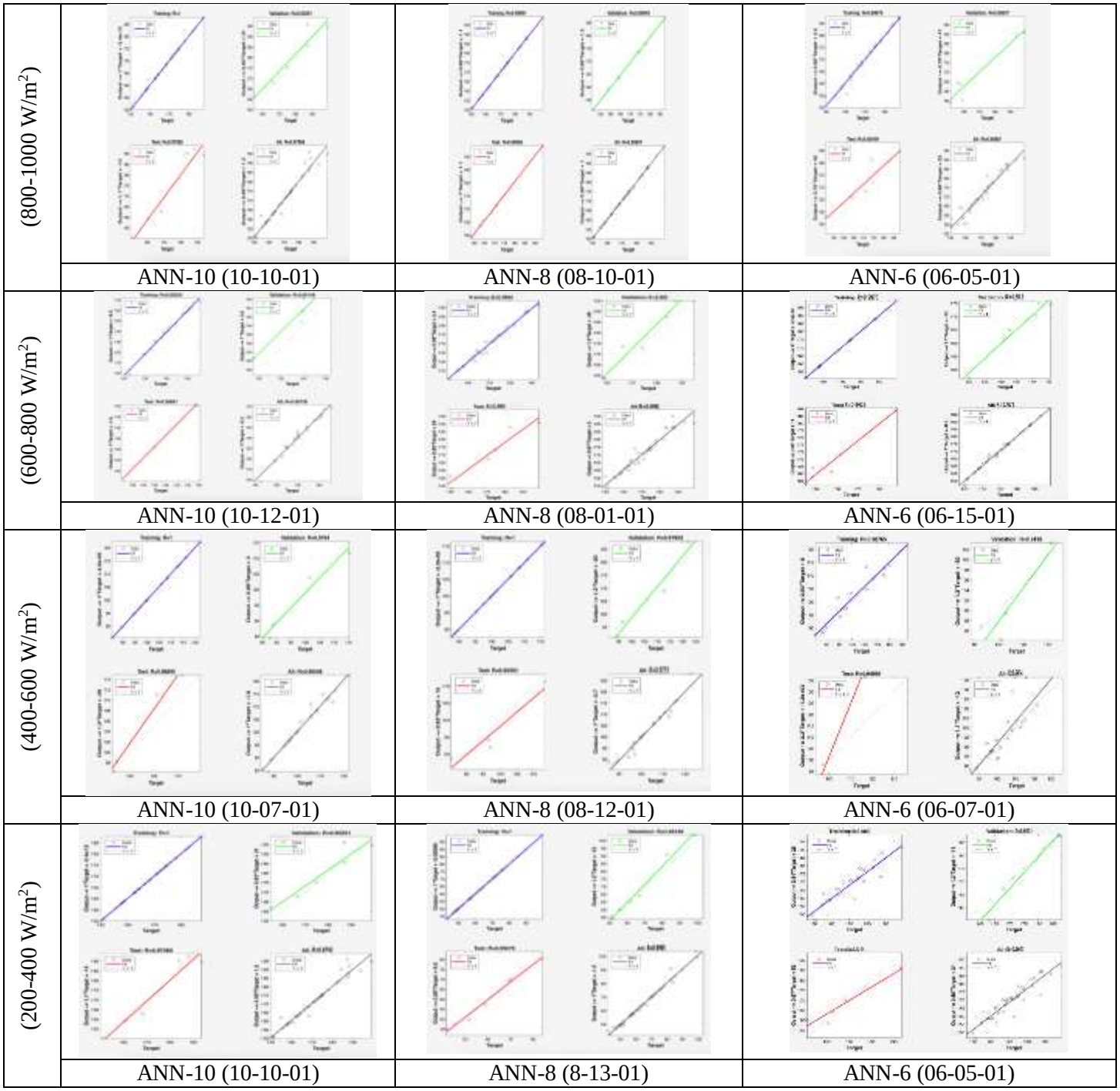

Figure 8: Regression performance of ANN-10, ANN-8, and ANN-6 models at different solar radiation ranges (200–1000 W/m²) showing correlation coefficients (R) for training, validation, testing, and overall datasets.
Source: Authors, (2026).

The regression plots for all models (ANN-10, ANN-8, and ANN-6) under different solar radiation ranges (200–1000 W/m²) are presented in Fig (8). The correlation coefficient (R) was used to evaluate the prediction accuracy of each model.

Among the tested configurations, the ANN-8 model consistently achieved the highest R-values across training, validation, and testing datasets, indicating strong predictive reliability and superior generalization capability. At higher radiation levels (800–1000 W/m²), ANN-8 achieved an almost perfect fit (R ≈ 0.9999), while ANN-10 and ANN-8 recorded slightly lower but still high correlations. Similarly, for mid-range radiation (600–800 W/m²), ANN-10 exhibited an R ≈ 0.985, demonstrating robust prediction with minimal deviation. At lower radiation levels (200–400 W/m² and 400–600 W/m²), all models maintained good predictive capability, though minor reductions in R were observed, possibly due to increased variability in measured data. Overall, ANN-8 outperformed ANN-10 and ANN-6 across all radiation categories, suggesting that its architecture and parameter tuning were most effective in capturing the nonlinear behaviour of solar radiation. The comparative evaluation of ANN architectures across different solar radiation ranges demonstrates the strong influence of input size and hidden layer configuration on model performance. The results consistently reveal that the number of input features plays a decisive role in predictive accuracy, while the hidden layer size contributes to fine-tuning the learning and generalization capability of the models [29].

For the highest radiation range (800–1000 W/m²), ANN-R8 with eight inputs emerged as the most effective configuration. The MLP (8-10-1) architecture achieved nearly perfect regression values across training, testing, and validation, highlighting its ability to capture the nonlinear relationships in this intensity band. In contrast, models with six inputs (ANN-6) showed weaker performance, with the best architecture (6-5-1) still underperforming relative to the 8- and 10-input models [30]. This suggests that reducing the dimensionality of the input space limits the ability of network to learn the complex patterns inherent in high-intensity solar radiation data. In the 600–800 W/m² range, the superiority of ANN-8 (8 inputs) was reinforced. The MLP (8-1-1) model provided the highest consistency, achieving regression values close to unity across all phases. Interestingly, in this radiation band, some 6-input models also delivered competitive performance, particularly MLP (6-15-1), which maintained strong validation accuracy. This indicates that for mid-range radiation intensities, the complexity of input features can be somewhat reduced without severely compromising prediction accuracy, provided that the hidden layer is configured effectively [31].

At the 400–600 W/m² range, the results once again emphasized the strength of the 8-input models. ANN-8 with MLP (8-15-1) delivered outstanding regression values, demonstrating robustness and generalization ability across datasets. The 10-input model (ANN-10) with MLP (10-11-1) also showed high reliability, but its performance fluctuated more across different subsets, suggesting occasional overfitting. The 6-input models in this range displayed greater variability, with certain configurations performing well (e.g., MLP 6-7-1) while others struggled, indicating that fewer input features reduce stability and predictive robustness [32]. For the lowest radiation range (200–400 W/m²), the pattern remained consistent: 8-input models were the most effective, particularly ANN-8 with MLP (8-13-1), which achieved near-perfect regression values across training, testing, and validation. The 10-input models performed well in certain configurations, such as MLP (10-7-1), which balanced accuracy and generalization effectively. However, the 6-input models again showed inconsistent performance, with only a few architectures like MLP (6-7-1) delivering reliable outcomes [33]. This reinforces the observation that reducing input dimensionality to six features constrains the capacity of model to capture variability in the lower radiation bands.

A critical aspect revealed by the results is the balance between overfitting and underfitting. Some 10-input networks, particularly those with larger hidden layers, demonstrated very high training accuracy but weaker performance on validation data. This suggests a tendency toward overfitting, where the model memorizes training patterns without effectively generalizing to unseen data [34]. On the other hand, some 6-input models, especially with fewer hidden neurons, showed underfitting, as they were unable to capture sufficient complexity to accurately represent solar radiation variability. The 8-input networks consistently avoided both extremes by maintaining strong generalization, indicating an optimal balance between model complexity and predictive stability. The findings across all radiation ranges demonstrate that ANN-8 models with eight input features consistently outperformed both the 10-input and 6-input networks. The 8-input configurations not only provided superior predictive accuracy but also ensured stability and generalization across different solar radiation ranges. The results further highlight the importance of optimizing hidden layer size, as models with too few or too many neurons occasionally exhibited reduced generalization despite high training accuracy [35].

Thus, selecting the appropriate number of inputs alongside a balanced hidden layer configuration is critical for achieving robust solar radiation prediction. Among all tested architectures, ANN-8 with MLP (8-10-1) stands out as the most effective, offering both computational efficiency and high predictive reliability across radiation ranges. From a practical perspective, these findings have direct implications for solar energy applications. Reliable solar radiation prediction is essential for photovoltaic (PV) system design, energy storage planning, and grid integration. Accurate forecasts help optimize the sizing of PV modules and batteries, improve energy dispatch strategies, and reduce reliance on fossil-fuel backup systems. The demonstrated stability of the 8-input models indicates that such configurations can provide dependable forecasts under varying radiation conditions, thereby supporting more efficient energy management. Furthermore, the insights on overfitting and underfitting emphasize the need for careful architecture selection to avoid unreliable predictions that could compromise planning and operational decisions in solar energy systems [36]. The superior performance of ANN-8 across all radiation ranges demonstrates its robustness and adaptability in modelling nonlinear solar radiation behaviour.

The near-unity correlation coefficients (R > 0.99) in both training and testing phases indicate excellent learning and generalization capability. The slightly reduced performance of ANN-10 and ANN-6 in certain ranges may be attributed to overfitting or inadequate neuron configuration, which limits the model's ability to adapt to complex variations in input–output relationships. The stability of ANN-8 across both high and low radiation levels confirms its optimized network architecture and well-balanced hidden layer design. These results affirm that ANN-based models, particularly ANN-8, are highly reliable for solar radiation estimation and serve as an efficient computational alternative to traditional empirical or statistical methods. The performance of three artificial neural network (ANN) models (ANN-10, ANN-8, and ANN-6) was evaluated across different solar radiation ranges (200–1000 W/m²) to understand the effect of multilayer perceptron (MLP) architecture, specifically the number of neurons in the hidden layer, on the correlation coefficient (R). The results revealed that ANN-10 demonstrated the most consistent and robust performance, particularly in the 800–1000 W/m² range, maintaining high R-values (between 0.90 and 1.00) across various neuron configurations, suggesting optimal stability and generalization capability.

In contrast, ANN-8 achieved the highest and most stable R-values, particularly in the 200-400 W/m² range, consistently exceeding 0.95. This indicates its superior ability to generalize and effectively capture the nonlinear relationships in the data, with minimal fluctuations in performance across hidden-layer neuron changes. However, ANN-6 exhibited significant instability across all radiation ranges, with fluctuating R-values ranging from 0.55 to 0.98, pointing to overfitting and underfitting challenges due to its sensitivity to network complexity. Overall, the findings suggest that the choice of ANN architecture plays a crucial role in predicting solar radiation, with ANN-10 being most suitable for higher radiation ranges and ANN-8 offering superior reliability in lower ranges, while ANN-6 showed less consistent results, making it less ideal for accurate predictions [37].

This section presents a comparison of current state-of-the-art regression models for photovoltaic power forecasting, evaluated using the R^2 regression metric in Table 2.

Compared to the research articles referenced in Table 1, the proposed model presented improved results in photovoltaic power prediction as measured by the R^2 regression metric. However, the novel aspect of the model is located in the portability of the most effective regression technique. The proposed model has achieved a R^2 value of 0.999 in 800–1000 W/m² range and utilises a novel preprocessing approach and method. The feature engineering process in the proposed model employed correlation coefficients to extract significant features from the dataset.

Table 2: Comparison of present results with literature.

Case study	Feature selection method	Regression model	Features Type	Maximum R	Dataset Used
[38]	Tree-based feature importance & principal component analysis	Artificial neural network and random forest	4	0.9672	Hawaii, United States of America (2016)
[39]	7, 6, and 5 input scenarios manually tested; no automatic feature selection algorithm.	(ANN) with multilayer feed-forward architecture and backpropagation algorithms (GD, LM, RP, SCG).	9	0.9629	Indian Meteorological Department
[40]	Wavelet transformation-based decomposition technique	WT-LSTM, LSTM, Ridge regression, Lasso regression, elastic-net regression	7	0.9749	Urbana Champaign, Illinois
[41]	Manual selection using domain knowledge; 6 input parameters selected for physical relevance to PV generation.	(ANN) with Levenberg–Marquardt backpropagation.	6	0.9925	Middle Technical University, Baghdad–Iraq.
[42]	Pearson's correlation and heatmap	(LR), (GBR), (MLP)	4	0.9823	Sharda University PV Dataset
Present study	Feature reduction performed only by comparing 10, 8, and 6 input combinations	Levenberg-Marquardt algorithm; ANN-10, ANN-8, ANN-6	10	0.9995	UCSI University Kuala Lumpur, Malaysia), 1-month experimental dataset (25 June–26 July 2025)

Source: Authors, (2026).

V. CONCLUSION

This study developed and evaluated ANN models for predicting the output power of photovoltaic (PV) systems under varying environmental conditions. Three ANN configurations with different input combinations (10, 8, and 6 parameters) were evaluated across four solar radiation ranges (200–400, 400–600, 600–800, and 800–1000 W/m²). The results demonstrate that model performance strongly depends on both the number of input features and the hidden layer architecture. Among the tested models, ANN-8 (8 inputs) consistently outperformed ANN-10 (10 inputs) and ANN-6 (6 inputs), achieving near-perfect regression values ($R \approx 0.999$) across most radiation ranges. This indicates that selecting an optimal subset of input variables, rather than using all available parameters, enhances predictive accuracy and reduces computational complexity. The findings also confirm that ANN models are highly effective in capturing the nonlinear relationships between electrical and environmental parameters influencing PV performance. Overall, the study highlights the robustness of ANN-based modelling for solar energy forecasting and system performance evaluation. By systematically comparing different ANN architectures, this work provides practical insights into input feature selection and network design, supporting the deployment of reliable and efficient prediction tools for PV applications across diverse climatic conditions.

VI. AUTHOR'S CONTRIBUTION

Conceptualization: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J.
Methodology: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J..
Investigation: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J..
Discussion of results: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J..
Writing – Original Draft: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J..
Writing – Review and Editing: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J..
Resources: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J..
Supervision: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J..
Approval of the final text: Ravshanbek Rakhmatulaev, Rodney Tan, Aliev Rayimjon, Yu L. J..

VII. REFERENCES

[1] O. Bashiru, C. Ochem, L. A. Enyejo, H. N. N. Manuel, and T. O. Adeoye, "The crucial role of renewable energy in achieving the sustainable development goals for cleaner energy," Global Journal of Engineering and Technology Advances, vol. 19, no. 3, pp. 11–36, 2024.

[2] M. M. Hasan et al., "Harnessing solar power: a review of photovoltaic innovations, solar thermal systems, and the dawn of energy storage solutions," Energies (Basel), vol. 16, no. 18, p. 6456, 2023.

[3] M. D. S. Hossain, A. Wadi Al-Fatlawi, L. Kumar, Y. R. Fang, and M. E. H. Assad, "Solar PV high-penetration scenario: an overview of the global PV power status and future growth," Energy Systems, pp. 1–57, 2024.

[4] P. Gupta and R. Singh, "PV power forecasting based on data-driven models: a review," International Journal of Sustainable Engineering, vol. 14, no. 6, pp. 1733–1755, 2021.

[5] K. Hasan, S. B. Yousuf, M. S. H. K. Tushar, B. K. Das, P. Das, and M. S. Islam, "Effects of different environmental and operational factors on the PV performance: A comprehensive review," Energy Sci Eng, vol. 10, no. 2, pp. 656–675, 2022.

[6] M. Amir, Zaheeruddin, and A. Haque, "Intelligent based hybrid renewable energy resources forecasting and real time power demand management system for resilient energy systems," Sci Prog, vol. 105, no. 4, p. 00368504221132144, 2022.

[7] M. R. Vogt et al., "Introducing a comprehensive physics-based modelling framework for tandem and other PV systems," Solar Energy Materials and Solar Cells, vol. 247, p. 111944, 2022.

[8] D. Saadaoui et al., "Parameters extraction of single diode and double diode models using analytical and numerical approach: A comparative study," International Journal of Modelling and Simulation, vol. 45, no. 2, pp. 569–598, 2025.

[9] G. M. Tina, C. Ventura, S. Ferlito, and S. De Vito, "A state-of-art-review on machine-learning based methods for PV," Applied Sciences, vol. 11, no. 16, p. 7550, 2021.

[10] S. Sivamani, M. Udayakumar, and N. Sellappan, "Prediction of single cylinder direct injection diesel engine performance fuelled with lemon peel oil biodiesel using artificial neural network," Mater Today Proc, vol. 80, pp. 456–463, 2023.

[11] Z. Zhang et al., "Spot Market Electricity Price Forecast via the Combination of Transformer and Ornstein-Uhlenbeck Process," in 2025 21st International Conference on the European Energy Market (EEM), IEEE, 2025, pp. 1–6.

[12] A. P. Selvam and S. N. S. Al-Humairi, "Environmental impact evaluation using smart real-time weather monitoring systems: a systematic review," Innovative Infrastructure Solutions, vol. 10, no. 1, pp. 1–24, 2025.

[13] M. Jaber, A. S. Abd Hamid, K. Sopian, A. Fazlizan, and A. Ibrahim, "Prediction model for the performance of different PV modules using artificial neural networks," Applied Sciences, vol. 12, no. 7, p. 3349, 2022.

[14] J. H. Yousif and H. A. Kazem, "Prediction and evaluation of photovoltaic-thermal energy systems production using artificial neural network and experimental dataset," Case Studies in Thermal Engineering, vol. 27, p. 101297, 2021.

[15] R. B. Roy et al., "A comparative performance analysis of ANN algorithms for MPPT energy harvesting in solar PV system," IEEE Access, vol. 9, pp. 102137–102152, 2021.

[16] R. El Chaal and M. O. Aboutafail, "A comparative study of back-propagation algorithms: Levenberg-Marquart and BFGS for the formation of multilayer neural networks for estimation of fluoride," Commun. Math. Biol. Neurosci., vol. 2022, p. Article-ID, 2022.

[17] R. García-Ródenas, L. J. Linares, and J. A. López-Gómez, "Memetic algorithms for training feedforward neural networks: an approach based on gravitational search algorithm," Neural Comput Appl, vol. 33, no. 7, pp. 2561–2588, 2021.

[18] E. Pashaei and E. Pashaei, "Training feedforward neural network using enhanced black hole algorithm: a case study on COVID-19 related ACE2 gene expression classification," Arab J Sci Eng, vol. 46, no. 4, pp. 3807–3828, 2021.

[19] Q. Askari and I. Younas, "Political optimizer based feedforward neural network for classification and function approximation," Neural Process Lett, vol. 53, no. 1, pp. 429–458, 2021.

[20] A. Singh and K. Gaurav, "Deep learning and data fusion to estimate surface soil moisture from multi-sensor satellite images," Sci Rep, vol. 13, no. 1, p. 2251, 2023.

[21] A. Alsirhani, M. M. Alshahrani, A. Abukwaik, A. I. Taloba, R. M. Abd El-Aziz, and M. Salem, "A novel approach to predicting the stability of the smart grid utilizing MLP-ELM technique," Alexandria Engineering Journal, vol. 74, pp. 495–508, 2023.

[22] M. B. A. Shuvho, M. A. Chowdhury, S. Ahmed, and M. A. Kashem, "Prediction of solar irradiation and performance evaluation of grid connected solar 80KWp PV plant in Bangladesh," Energy Reports, vol. 5, pp. 714–722, 2019.

[23] A. Mellit, S. Sağlam, and S. A. Kalogirou, “Artificial neural network-based model for estimating the produced power of a photovoltaic module,” Renew Energy, vol. 60, pp. 71–78, 2013.

[24] S. A. Rizzo and G. Scelba, “ANN based MPPT method for rapidly variable shading conditions,” Appl Energy, vol. 145, pp. 124–132, 2015.

[25] R. May, G. Dandy, and H. Maier, “Review of input variable selection methods for artificial neural networks,” Artificial neural networks-methodological advances and biomedical applications, vol. 10, no. 1, pp. 19–45, 2011.

[26] M. M. Bejani and M. Ghatee, “A systematic review on overfitting control in shallow and deep neural networks,” Artif Intell Rev, vol. 54, no. 8, pp. 6391–6438, 2021.

[27] S. Sivamani, S. Selvakumar, K. Rajendran, and S. Muthusamy, “Artificial neural network–genetic algorithm-based optimization of biodiesel production from Simarouba glauca,” Biofuels, vol. 10, no. 3, pp. 393–401, 2019.

[28] P. S. Kumar and S. Sivamani, “Numerical analysis and implementation of artificial neural network algorithm for nonlinear function,” International Journal of Information Technology, vol. 13, no. 5, pp. 2059–2068, 2021.

[29] M. A. Ali, A. Elsayed, I. Elkabani, M. Akrami, M. E. Youssef, and G. E. Hassan, “Optimizing artificial neural networks for the accurate prediction of global solar radiation: A performance comparison with conventional methods,” Energies (Basel), vol. 16, no. 17, p. 6165, 2023.

[30] M. Mukhtar et al., “Development and comparison of two novel hybrid neural network models for hourly solar radiation prediction,” Applied Sciences, vol. 12, no. 3, p. 1435, 2022.

[31] S. Ghimire et al., “Hybrid convolutional neural network-multilayer perceptron model for solar radiation prediction,” Cognit Comput, vol. 15, no. 2, pp. 645–671, 2023.

[32] A. A. Imam, A. Abusorrah, M. M. A. Seedahmed, and M. Marzband, “Accurate forecasting of global horizontal irradiance in Saudi Arabia: a comparative study of machine learning predictive models and feature selection techniques,” Mathematics, vol. 12, no. 16, p. 2600, 2024.

[33] S. Y. Heng et al., “Artificial neural network model with different backpropagation algorithms and meteorological data for solar radiation prediction,” Sci Rep, vol. 12, no. 1, p. 10457, 2022.

[34]bB. Girdhani and M. Agrawal, “Performance evaluation of statistical and deep learning models for daily solar global horizontal radiation prediction: implications for renewable energy and sustainability,” Environ Dev Sustain, pp. 1–29, 2025.

[35] A. K. Yadav, R. Kumar, M. Wang, G. Fekete, and T. Singh, “Comparative analysis of daily global solar radiation prediction using deep learning models inputted with stochastic variables,” Sci Rep, vol. 15, no. 1, p. 10786, 2025.

[36] G. Drałus, D. Mazur, J. Kusznier, and J. Drałus, “Application of artificial intelligence algorithms in multilayer perceptron and Elman networks to predict photovoltaic power plant generation,” Energies (Basel), vol. 16, no. 18, p. 6697, 2023.

[37] V. Moya-Almeida, B. Diezma-Iglesias, E. Correa-Hernando, C. Vaquero-Miguel, and N. Alvarado-Arias, “Setpoint temperature estimation to achieve target solvent concentrations in S. cerevisiae fermentations using inverse neural networks and fuzzy logic,” Eng Appl Artif Intell, vol. 127, p. 107248, 2024.

[38] U. Munawar and Z. Wang, “A Framework of Using Machine Learning Approaches for Short-Term Solar Power Forecasting,” Journal of Electrical Engineering and Technology, vol. 15, no. 2, pp. 561–569, Mar. 2020, doi: 10.1007/S42835-020-00346-4.

[39] N. Premalatha and A. Valan Arasu, “Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms,” Journal of Applied Research and Technology, vol. 14, no. 3, pp. 206–214, Jun. 2016, doi: 10.1016/J.JART.2016.05.001.

[40] M. Mishra, P. Byomakesha Dash, J. Nayak, B. Naik, and S. Kumar Swain, “Deep learning and wavelet transform integrated approach for short-term solar PV power prediction,” Measurement (Lond), vol. 166, Dec. 2020, doi: 10.1016/J.MEASUREMENT.2020.108250.

[41] A. T. Mohammad, H. M. Hussen, and H. J. Akeiber, “Prediction of the output power of photovoltaic module using artificial neural networks model with optimizing the neurons number,” International Journal of Renewable Energy Development, vol. 12, no. 3, pp. 478–487, May 2023, doi: 10.14710/IJRED.2023.49972.

[42] B. U. D. Abdullah, S. A. Khanday, N. U. Islam, S. Lata, H. Fatima, and S. H. Nengroo, “Comparative Analysis Using Multiple Regression Models for Forecasting Photovoltaic Power Generation,” Energies 2024, Vol. 17, Page 1564, vol. 17, no. 7, p. 1564, Mar. 2024, doi: 10.3390/EN17071564.