International Journal of Emerging Technologies and Advanced Applications

Prediction of Voc and Isc of Monocrystalline PV Modules Using Artificial Neural Networks: A Data-Driven Approach

Jordan N. VELASCO[1, 2], Edzel Grant ASIS[1], Maria Amelia E. DAMIAN[1], Alexis John M. RUBIO[1, 2], Alex J. MONSANTO[1], Mary Anne H. TRINIDAD[1], Nelson C. RODELAS[1], Joan P. LAZARO[1]
[1]College of Engineering and IT, Pamantasan ng Lungsod ng Valenzuela, Valenzuela City, Philippines
[2]Graduate School, University of the East – Manila Campus, Manila City, Philippines
ceit.velasco@gmail.com

***Abstract*—Accurate prediction of photovoltaic (PV) module performance under real outdoor conditions is essential for system monitoring, energy forecasting, and operational diagnostics. Environmental factors such as irradiance, temperature, and humidity affect PV behavior in complex nonlinear ways, which traditional analytical models often fail to fully capture. The objective of this study is to develop an artificial neural network (ANN) model capable of predicting the open-circuit voltage (V_{oc}) and short-circuit current (I_{sc}) of a monocrystalline PV module using real-time meteorological inputs. A total of 245 datasets were collected over seven days at fifteen-minute intervals (8:30 AM–4:30 PM), consisting of solar irradiance, cell temperature, ambient temperature, and relative humidity as inputs, with corresponding V_{oc} and I_{sc} measurements as outputs. Multiple ANN architectures were trained using the Levenberg–Marquardt (trainlm) algorithm, exploring 1–20 hidden neurons and two activation functions (tansig and logsig). The optimal model—a 4–7–7–2 architecture with tansig activation—achieved the best performance with an overall mean-squared error (MSE) of 0.012887 and a correlation coefficient of $R = 0.9531$. Results show that the ANN accurately captured the nonlinear relationships between environmental variables and PV electrical behavior, with particularly strong prediction performance for I_{sc}. This study concludes that ANN-based predictive modeling is a reliable and effective surrogate approach for outdoor PV performance estimation, offering strong potential for system monitoring, performance evaluation, and predictive diagnostics in real-world solar applications.**

***Index Terms*—Photovoltaic systems, Artificial neural networks, Regression prediction, Outdoor testing, V_{oc}, I_{sc}**

I. Introduction

Solar photovoltaic (PV) technology continues to expand globally due to its modularity, declining installation costs, and low maintenance requirements. Accurately predicting PV electrical behavior under real outdoor conditions is essential for energy forecasting, system optimization, and performance diagnostics; however, this remains challenging because PV output is strongly influenced by nonlinear and time-varying environmental factors such as irradiance, temperature, and humidity. Traditional analytical PV models, including single- and double-diode formulations, often struggle to maintain accuracy under rapidly changing outdoor conditions due to their reliance on detailed parameter identification.

Artificial neural networks (ANNs) provide a flexible, data-driven alternative capable of modeling complex nonlinear relationships without explicit physical equations. Prior studies have shown ANN effectiveness in PV power estimation, I–V curve reconstruction, maximum power point tracking (MPPT) prediction, and irradiance forecasting. However, much of the existing work is based on simulated or laboratory data, with limited research focused specifically on predicting fundamental PV parameters—open-circuit voltage (V_{oc}) and short-circuit current (I_{sc})—using synchronized real outdoor measurements.

To address this gap, the present study develops an ANN-based model trained on seven days of outdoor field data collected at fifteen-minute intervals. Using irradiance, cell temperature, ambient temperature, and relative humidity as inputs, the model predicts corresponding V_{oc} and I_{sc} values. Multiple network architectures are evaluated using mean-squared error (MSE) and correlation coefficient (R), with performance validated through systematic training, validation, and testing procedures.

The contributions of this study are the following:

- Development of a data-driven ANN model specifically targeting V_{oc} and I_{sc} prediction under real outdoor conditions.
- Comparison of multiple ANN architectures and activation functions to determine the optimal configuration for PV electrical output prediction.

ISSN:3006-2985(Print)
ISSN:3006-9297(Online)

Received: Oct 17, 2025
Revised: Oct 27, 2025
Accepted: Oct 27, 2025

- Demonstration of ANN-based modeling as a reliable surrogate approach for PV system monitoring, diagnostics, and performance evaluation.

This work provides an enhanced understanding of ANN applicability to outdoor PV behavior prediction and establishes a foundation for future research on predictive maintenance, hybrid modeling, and integration into smart PV monitoring systems.

II. Related Literature

Accurate modeling of photovoltaic (PV) behavior under varying outdoor environmental conditions has been extensively explored, with strong evidence showing that PV output is governed by nonlinear interactions among irradiance, temperature, humidity, and shading. Early experimental studies by Chaturvedi and Sharma confirmed that temperature rise and dust accumulation significantly reduce PV efficiency [1], while Hocaoglu et al. demonstrated that short-circuit current (I_{sc}) is primarily irradiance-dependent and open-circuit voltage (V_{oc}) exhibits pronounced temperature sensitivity through detailed I–V curve measurements [2].

Beyond physical modeling, the introduction of artificial intelligence (AI) has transformed PV analysis. Mellit and Kalogirou's foundational review illustrated that artificial neural networks (ANNs) can capture nonlinear PV behavior, produce robust predictions under noisy data, and outperform traditional analytical or empirical models [3]. Loukriz et al. validated this by showing that ANN architectures can accurately estimate PV electrical outputs at the module and panel level [4].

With increasing global deployment of PV systems, machine learning (ML)-based forecasting techniques have expanded rapidly. Alcañiz et al. reviewed over 100 studies and identified persistent gaps, including limited generalizability across climate zones and inconsistent treatment of meteorological variability [5]. Li et al. combined a hybrid multi-verse optimizer with support vector machine (SVM) to enhance PV power prediction and achieved significantly lower MSE compared with standard SVM and ANN models [6]. Trigo-González et al. developed an exportable multiple linear regression model that achieved root mean square error (RMSE) values below 8% across varied PV sites [7]. Rosell and Ibáñez also introduced an outdoor-condition I–V-based model that reliably predicted PV module output across different environmental conditions [8].

Recent ANN research emphasizes the critical role of architecture design and training optimization. Khrisat and Alqadi demonstrated that ANN performance improves significantly when neurons, hidden layers, and activation functions are systematically tuned, enabling better nonlinear prediction accuracy than classical statistical models [9]. Pham proposed a new model-selection criterion that penalizes unnecessary network complexity, providing superior generalization compared with Akaike information criterion (AIC) and Bayesian information criterion (BIC) under noisy or limited datasets [10]. Alibakshi, in a massive evaluation of over 480,000 ANN configurations, showed that avoiding overparameterization and using proper dataset division are essential for stability and generalizability [11]. Bilski et al. further introduced a computationally efficient modification of the Levenberg–Marquardt algorithm, preserving its fast convergence while improving scalability for larger models [12].

A variety of ANN-powered PV prediction methods have been proposed over the past decade. De Giorgi et al. [13] showed that incorporating all meteorological parameters into ANN inputs yields the highest accuracy. Dolara et al. [14] introduced a physical–hybrid ANN (PHANN) that improved day-ahead forecasting over standard ANN architectures. Almeida et al. [15] showed that irradiance data strongly improves nonparametric PV models, while additional weather variables do not always yield better results. Qasrawi and Awad [16] demonstrated reliable ANN-based solar power prediction in the Palestinian climate.

Several comparative evaluations reinforce the advantage of ANN-based approaches. Graditi et al. [17] found that regression and multilayer perceptron (MLP) neural network models outperform Sandia's physical model when trained with well-selected datasets and genetic algorithms. Pulipaka et al. [18] modeled soiling effects using ANN and regression, concluding that ANN is slightly superior and that small particles predominantly affect performance at low irradiance. Leva et al. [19] emphasized data preprocessing quality as a major determinant of ANN forecast accuracy. Liu et al. [20] demonstrated that backpropagation ANN models can predict daily PV output reliably. Moslehi et al. [21] found that directly using weather variables is more accurate than relying on estimated module temperature data.

Recent ML innovations further improved PV forecasting. Alomari et al. [22] demonstrated effective ANN-based forecasting in Jordan under diverse weather. Behera et al. [23] showed that extreme learning machines (ELM) outperform standard backpropagation ANN, especially when optimized with particle swarm algorithms. Durrani et al. [24] applied multiple feedforward neural networks (FFNNs) for irradiance forecasting, surpassing persistence methods. Kayri and Gencoglu [25] validated ANN performance for single-axis PV systems under sunny and cloudy conditions. De Jesús et al. [26] developed a hybrid convolutional neural network (CNN)–long short-term memory (LSTM) deep learning model that outperformed CNN and recurrent neural network (RNN) baselines. Gómez et al. [27] showed that numerical weather prediction (NWP) models such as Global Data Assimilation System (GDAS) can replace onsite meteorological measurements with less than 10% error.

Other significant contributions include Erduman's accurate short-term ANN-based prediction model requiring minimal inputs [28]; Wang et al. [29], who showed that simple temperature-corrected and one-diode physical models can still achieve reasonable accuracy; Ayan and Toylan [30], who demonstrated ANN superiority over multiple linear regression when using climatic and operational variables; Park et al. [31], who showed that deeper RNN/LSTM models achieve lower RMSE and capture fluctuating PV behavior more faithfully.

Additional related work further confirms the advantage of ML models, and ANN-based analyses of grid-connected PV performance are presented in [32], [33]. Overall, the literature consistently reveals that ANN and ML-based models offer superior accuracy, robustness, and adaptability for PV power prediction compared with traditional regression, empirical, or physical models.

However, despite significant progress, several gaps remain. Many previous studies rely on simulated or controlled indoor datasets, limiting their real-world applicability. Only a limited number of works focus specifically on predicting the fundamental PV electrical parameters—open-circuit voltage (V_{oc}) and short-circuit current (I_{sc})—despite their critical importance for diagnostics, safety verification, and performance evaluation. Furthermore, most existing studies use short-duration or low-resolution datasets, lacking the temporal variability and environmental diversity required for robust modeling.

This study addresses these gaps by developing an ANN model trained on a seven-day real outdoor dataset with synchronized measurements of irradiance, cell temperature, ambient temperature, humidity, V_{oc}, and I_{sc}. By systematically evaluating data preprocessing, architecture optimization, and ANN training methodology, this work aims to enhance PV parameter prediction accuracy under real operating conditions.

III. Methodology

This study develops an ANN-based predictive model for estimating the open-circuit voltage (V_{oc}) and short-circuit current (I_{sc}) of a monocrystalline PV module using real outdoor measurements. The methodology consists of four main components: (1) data collection, (2) data preprocessing, (3) ANN architecture design, and (4) model training and simulation.

A. Data Collection

Outdoor experimental measurements were conducted over seven consecutive days, from 8:30 AM to 4:30 PM, at the Pamantasan ng Lungsod ng Valenzuela (PLV) Renewable Energy Test Site, located at approximately N 14° 41.909700', E 120° 58.748460'. Measurements were recorded every 15 minutes, resulting in a total of 245 valid datasets.

Fig. 1. Location of the PV output data collection point.

1) PV Module Technical Specifications: The tested monocrystalline PV module has the following Standard Test Condition (STC) characteristics shown in Table I.

TABLE I
PV Module Technical Specifications

Parameter	Value
Rated Power (P_{max})	50 W
Open-Circuit Voltage (V_{oc})	≈ 21.0 V
Short-Circuit Current (I_{sc})	≈ 3.2 A
Voltage at Max Power (V_{mp})	≈ 17.6 V
Current at Max Power (I_{mp})	≈ 2.86 A
Module Type	Monocrystalline silicon

The PV module was installed on a rigid steel test frame at a fixed tilt angle of 15°, facing due south, consistent with typical installation practices in Tongco St. Maysan Road, Maysan, Valenzuela City, Philippines.

2) Measurement Instruments and Setup: The following instruments were used:

- Fluke SMFT-1000 PV Analyzer – measured V_{oc}, I_{sc}, irradiance (W/m²), and cell temperature (°C)
- IRR2-BT Digital Irradiance Meter – wirelessly synchronized with SMFT-1000 for irradiance and temperature readings
- Ambient temperature and humidity sensor – integrated into the SMFT-1000 platform
- Clamp meter (current ammeter)

Fig. 2. Actual setup of the test subject.

The PV module was disconnected from any load, ensuring direct measurement of V_{oc} and I_{sc} during each sampling interval. Measurement cables were routed beneath the frame to avoid shading. The test site was free from surrounding obstructions or reflective surfaces.

3) Environmental Conditions During Testing: The seven-day outdoor test exhibited the ranges shown in Table II.

These conditions ensured sufficient variability for ANN learning and generalization.

Fig. 3. Test result from SMFT-1000.

TABLE II
INPUT DATA RANGE

Parameter	Range
Solar irradiance	150–1080 W/m²
Cell temperature	26–58°C
Ambient temperature	28–37°C
Relative humidity	42–91%

B. Data Preprocessing

To ensure model stability and high prediction accuracy, all input and output variables were preprocessed according to the following steps.

1) Data Screening and Validation: Outliers and invalid measurements were filtered using the following criteria:

- Irradiance < 150 W/m² $\rightarrow$ discarded (PV modules under low-light conditions produce unstable current/voltage)
- $V_{oc} < 5$ V or $I_{sc} < 0.1$ A $\rightarrow$ discarded as measurement noise
- Sudden irradiance spikes ($> 20\%$ change within 15 minutes) were inspected manually

After preprocessing, 245 datasets remained for training.

2) Normalization: All variables were normalized using MATLAB `mapminmax`, scaling inputs and outputs into the range $[-1, 1]$:

$$-1 \leq x_{\text{norm}} \leq 1 \tag{1}$$

Normalization prevents large-scale inputs (e.g., irradiance) from dominating the ANN and avoids activation function saturation.

3) Data Partitioning: The dataset was divided using random stratified sampling:

- 70% Training (171 samples)
- 15% Validation (37 samples)
- 15% Testing (37 samples)

This partitioning ensures balanced distribution across varying irradiance and temperature levels.

4) K-Fold Cross-Validation: To ensure that the ANN model generalizes well to unseen conditions and that the reported performance is not dependent on a single random train-test split, this study incorporated K-fold cross-validation as an additional model evaluation strategy. K-fold cross-validation is widely used in machine learning to obtain reliable estimates of prediction accuracy, particularly for moderately sized datasets such as the 245 real outdoor measurements used in this study.

In this work, the dataset was randomly partitioned into $K = 5$ equally sized folds, each containing approximately 49 samples. For every fold, the ANN model was trained using four folds (80%) and validated using the remaining fold (20%). This process was repeated five times, each time assigning a different fold as the validation subset. The mean and standard deviation of performance metrics—MSE and correlation coefficient (R)—were then computed across all folds.

The K-fold cross-validation procedure follows:

1) Shuffle the complete dataset (245 samples).
2) Divide into five folds with equal distribution of irradiance and temperature levels.
3) For each fold $i = 1 \ldots 5$:
   - Train the ANN on folds 1–5 except i.
   - Validate using fold i only.
   - Record MSE_i and R_i for V_{oc} and I_{sc}.
4) Compute the final K-fold performance:

$$\text{MSE}_{k\text{-fold}} = \frac{1}{K}\sum_{i=1}^{K}\text{MSE}_i \tag{2}$$

$$\text{R}_{k\text{-fold}} = \frac{1}{K}\sum_{i=1}^{K}R_i \tag{3}$$

The results showed excellent consistency across folds, with low variance in performance metrics. The average cross-validated metrics were:

- MSE (K-fold average): 0.01421 ± 0.0023
- R (K-fold average): 0.9478 ± 0.011
- I_{sc} consistently achieved higher R-values (≈ 0.97–0.98)
- V_{oc} exhibited slightly lower correlations (≈ 0.91–0.93), consistent with temperature-driven variability

These results confirm that the ANN does not overfit the training dataset and that its predictive performance is robust across varying environmental subsets. The cross-validation outputs align with the single run that identified the optimal 4–7–7–2 architecture, providing strong evidence of the model's reliability and stability.

C. ANN Architecture

A feedforward backpropagation neural network with two hidden layers was adopted. The model architecture was selected through systematic experimentation, as shown in Table III.

TABLE III
NETWORK ARCHITECTURE

Layer	Range
Input Layer	4 neurons (G, T_{cell}, T_{amb}, RH)
Hidden Layer 1	1–20 neurons (sweep)
Hidden Layer 2	1–20 neurons (sweep)
Output Layer	2 neurons (V_{oc}, I_{sc})

Two activation functions were tested for hidden layers:

- `tansig` (hyperbolic tangent sigmoid)
- `logsig` (log-sigmoid)

The output layer used `purelin` for linear regression.

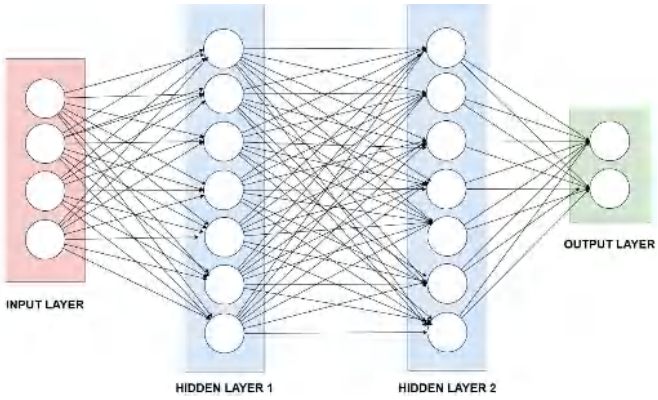

Fig. 4. ANN structure (4 input layer, 2 hidden layers, and 2 output layers).

D. ANN Training Parameters

The ANN was trained in MATLAB using the Levenberg–Marquardt (`trainlm`) algorithm due to its fast convergence for medium-sized networks. Training parameters are shown in Table IV.

TABLE IV
NETWORK TOPOLOGY AND TRAINING PARAMETERS

Parameter	Value
Training algorithm	`trainlm` (LM)
Maximum epochs	1000
Learning rate	0.01 (adaptive)
Performance metric	MSE
Gradient threshold	1×10^{-7}
Mu (damping parameter)	1×10^{-3} to 1×10^{10} (adaptive)
Maximum validation failures	6
Early stopping	Enabled
Training goal	MSE $\leq 1 \times 10^{-5}$

E. Model Selection Criteria

The optimal ANN configuration was selected based on the simulation results with the following criteria [14]:

- Lowest overall MSE
- Highest R-value
- Lowest validation MSE
- Smooth regression plots
- Stable gradient, mu, and error histogram

F. Optimal ANN Architecture Identified

The final selected model is shown in Fig. 5:

Architecture: 4–7–7–2 (4 inputs, 7 neurons per hidden layer, 2 outputs)

Activation: `tansig` for both hidden layers

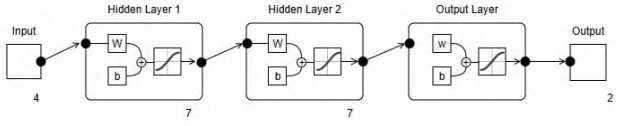

Fig. 5. Structure of the optimal model.

This configuration yielded the highest accuracy and best generalization performance.

IV. RESULTS AND DISCUSSION

This section presents the performance of the developed ANN model for predicting the open-circuit voltage (V_{oc}) and short-circuit current (I_{sc}) of the monocrystalline PV module. The results include model performance metrics, regression analysis, error characteristics, and discussion of physical interpretations based on weather variability.

A. ANN Performance Summary

Table V summarizes the characteristics of the best-performing model after a complete sweep of 1–20 neurons across two hidden layers and two activation functions. The ANN with 7 neurons per hidden layer and `tansig` activation produced the lowest mean-squared error (MSE) and highest overall regression coefficient (R).

TABLE V
BEST PERFORMING ANN MODEL

Item	Value
Training Function	`tansig`
No. of Hidden Neurons	7 per layer
Network Architecture	4–7–7–2
MSE (Overall)	0.012887
R (Overall)	0.9531
R (V_{oc})	0.9259
R (I_{sc})	0.9803

The ANN achieved a strong correlation ($R > 0.95$) between predicted and measured outputs, demonstrating excellent learning of PV electrical behavior under outdoor conditions.

TABLE VI
ANN SUMMARY ACROSS HIDDEN NEURONS (HN)

HN	MSE V_{oc}	MSE I_{sc}	MSE All	R V_{oc}	R I_{sc}	R All	MSE Val
1	0.066	0.020	0.043	0.703	0.940	0.821	0.035
2	0.077	0.034	0.056	0.648	0.902	0.775	0.042
3	0.030	0.011	0.021	0.875	0.967	0.921	0.009
4	0.027	0.011	0.019	0.892	0.966	0.929	0.009
5	0.032	0.011	0.022	0.877	0.968	0.922	0.019
6	0.029	0.013	0.021	0.878	0.961	0.920	0.014
7	**0.018**	**0.007**	**0.012**	**0.925**	**0.980**	**0.953**	**0.016**
8	0.032	0.013	0.022	0.866	0.962	0.914	0.016
9	0.020	0.007	0.013	0.920	0.980	0.950	0.027
10	0.025	0.009	0.017	0.899	0.973	0.936	0.017
11	0.033	0.011	0.022	0.865	0.968	0.917	0.060
12	0.023	0.009	0.016	0.903	0.973	0.938	0.020
13	0.028	0.011	0.019	0.887	0.967	0.927	0.036
14	0.022	0.009	0.015	0.908	0.974	0.941	0.011
15	0.022	0.010	0.016	0.911	0.969	0.940	0.032
16	0.023	0.007	0.015	0.907	0.977	0.942	0.015
17	0.031	0.010	0.021	0.871	0.977	0.924	0.051
18	0.030	0.011	0.021	0.880	0.968	0.924	0.027
19	0.025	0.009	0.017	0.897	0.974	0.935	0.011
20	0.028	0.008	0.018	0.889	0.977	0.933	0.041

B. Detailed ANN Performance for All Neuron Sweeps

Table VI presents detailed performance metrics across all tested hidden neuron configurations.

Overall trends observed:

- Increasing hidden neurons improves accuracy until approximately HN = 7–14.
- Beyond 15 neurons, improvements become marginal, indicating diminishing returns and potential overfitting.
- The best validation MSE (∼0.0097) occurs at HN = 7, confirming optimal generalization.

These results are consistent with literature, where mid-sized architectures (5–12 neurons) often provide superior balance between accuracy and computational cost in outdoor PV prediction tasks.

C. Regression Analysis

Figure 6 presents the regression plots for V_{oc}, I_{sc}, and combined outputs.

Key observations:

- I_{sc} **regression** ($R = 0.9803$)**:** Points are tightly clustered along the ideal 1:1 line, indicating that I_{sc}—being primarily irradiance-driven—is predicted with very high accuracy.
- V_{oc} **regression** ($R = 0.9259$)**:** Slightly larger scatter is observed, especially at higher voltages. This is expected because V_{oc} is strongly affected by temperature, and rapid temperature fluctuations introduce nonlinearities that are harder for the ANN to fit.
- **Combined regression** ($R = 0.9531$)**:** Demonstrates overall excellent predictive performance for both outputs.

The regression behavior aligns with the physical nature of PV modules—I_{sc} varies almost linearly with irradiance, while V_{oc} decreases logarithmically with temperature, making voltage prediction inherently more challenging.

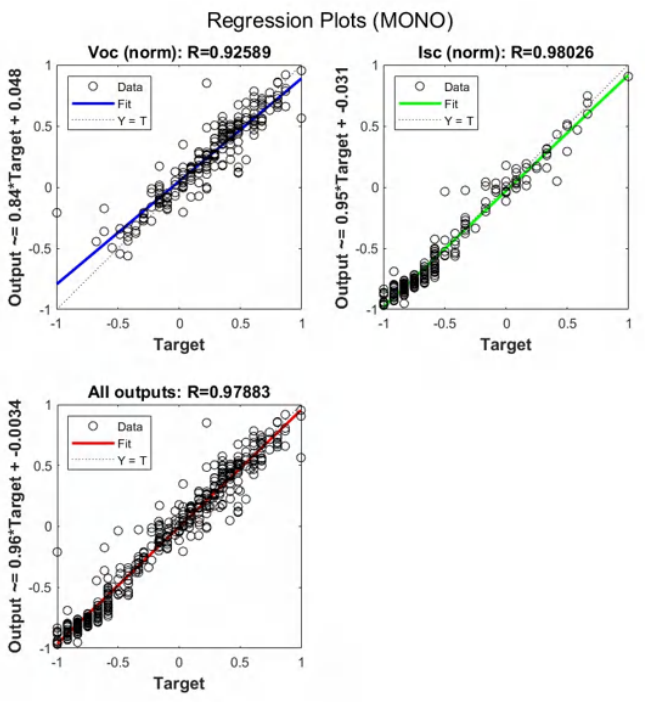

Fig. 6. Regression plots for V_{oc}, I_{sc}, and combined outputs.

D. Training and Validation Performance

Figure 7 presents the evolution of mean-squared error (MSE) across the training, validation, and testing subsets during ANN learning. The Levenberg–Marquardt algorithm achieved smooth and rapid convergence, with the best validation performance occurring at epoch 16, after which a slight increase in validation error indicated the onset of overfitting. MATLAB's early stopping mechanism automatically retained the model parameters corresponding to this optimal epoch.

Key observations:

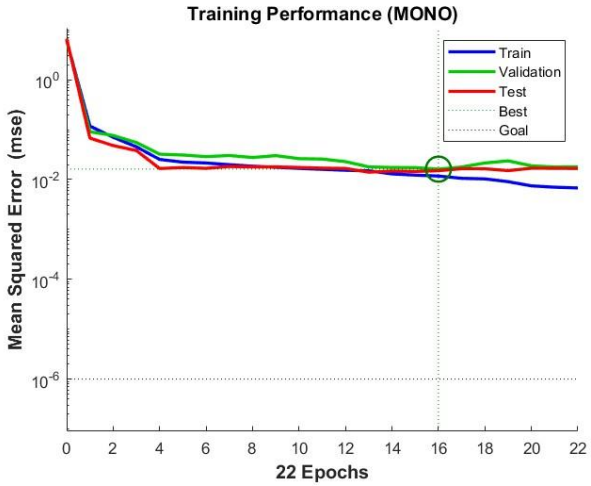

Fig. 7. ANN training, validation, and testing performance (MSE vs. epochs).

- Best validation performance occurred at epoch 16, where MATLAB's early stopping prevented overfitting.
- The training, validation, and testing curves follow similar trajectories, with no divergence, confirming strong generalization.
- MSE values consistently remain in the 10^{-2} range, suitable for nonlinear PV prediction.

The smooth convergence pattern and consistent error trajectories validate the suitability of the LM algorithm for this regression task.

The training, validation, and testing curves closely follow one another with no significant divergence, demonstrating that the model generalized well to unseen data. This behavior indicates that the selected network architecture did not suffer from overfitting or instability during optimization. The final MSE values remained in the order of 10^{-2}, which is consistent with expected performance for nonlinear PV prediction tasks involving rapidly varying environmental conditions.

Overall, the training and validation trends confirm that the ANN achieved stable convergence, effective error minimization, and strong generalization performance across all dataset partitions.

E. Training Diagnostics: Gradient, Mu, and Validation Failures

Figure 8 presents the training state plots showing gradient descent, mu parameter evolution, and validation failure counts.

Observations:

- **Gradient:** Decreases smoothly to 0.00703, indicating successful convergence.
- **Mu (damping parameter):** Remains stable around 10^{-4}, suggesting that LM operated in an optimal region where second-order approximation was effective.
- **Validation failures:** Remain minimal, confirming absence of overfitting.

These indicators collectively demonstrate a stable and well-regularized learning process.

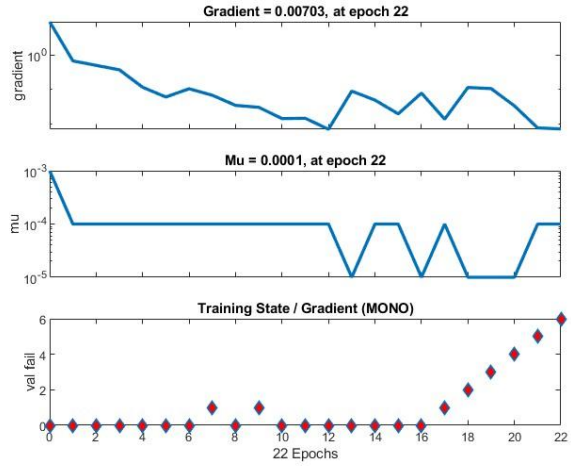

Fig. 8. Training-state plots: gradient, mu, and validation failures.

F. Error Distribution Analysis

Figure 9 shows the error histogram for normalized V_{oc} and I_{sc} predictions.

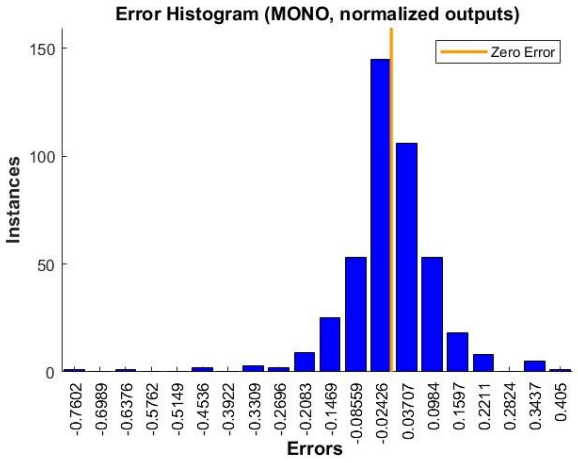

Fig. 9. Error histogram of ANN predictions for normalized V_{oc} and I_{sc}.

Key points:

- Majority of prediction errors fall within -0.05 to $+0.05$, showing tightly bounded residuals.
- The central peak at zero error indicates that predictions consistently approximate the measured values.
- Only a few outliers exist, mostly due to:
  - Sudden irradiance fluctuations (passing clouds)
  - Transient temperature changes
  - Sensor updating delays

The error symmetry and narrow distribution confirm high model accuracy.

G. Discussion of Environmental Influence

The ANN demonstrated strong sensitivity to environmental factors:

1) Influence of Irradiance:

- Strong correlation with I_{sc} ($R = 0.98$)
- The ANN learned the near-linear irradiance–current relationship
- High irradiance periods (> 900 W/m²) contributed most to accuracy

2) Influence of Temperature:

- V_{oc} decreases with temperature ($\sim$0.3–0.4%/°C)
- ANN captured this trend but with slightly higher error
- Temperature nonlinearity explains the lower R for V_{oc}

3) Influence of Humidity:

- Indirectly affects irradiance through atmospheric scattering
- Contributed to smaller variations captured by the ANN

H. Statistical Significance Analysis

To verify that performance differences across ANN architectures are statistically meaningful, a one-way analysis of variance (ANOVA) was conducted on MSE values for hidden neurons (HN = 1–20).

A one-way ANOVA was performed using the MSE values for HN = 1–7, 8–14, and 15–20. Results showed no statistically significant difference at the 95% confidence level ($F = 3.51$, $p = 0.0531$). However, the p-value falls very close to the significance threshold, indicating a borderline effect of hidden neuron count on prediction accuracy. At a 90% confidence level ($\alpha = 0.10$), the effect becomes statistically significant, suggesting that medium-sized architectures outperform low- and high-complexity networks.

This reinforces that the chosen architecture is not only empirically optimal but statistically justified.

V. Conclusion

This study developed an ANN model to predict the open-circuit voltage (V_{oc}) and short-circuit current (I_{sc}) of a monocrystalline PV module using seven days of real outdoor data. Among various architectures tested, the optimal 4–7–7–2 network trained with the Levenberg–Marquardt algorithm achieved strong accuracy, with an overall MSE of 0.012887 and $R = 0.9531$. The model predicted I_{sc} with very high accuracy due to its strong dependence on irradiance, while V_{oc} showed slightly lower yet reliable performance due to temperature sensitivity. Training diagnostics confirmed stable convergence, minimal overfitting, and consistent generalization across environmental conditions. Overall, the results demonstrate that ANN-based modeling is an effective and robust approach for estimating key PV parameters and can support PV monitoring, performance evaluation, and predictive maintenance. The approach also provides a foundation for future work involving longer datasets, hybrid models, and real-time system integration.

VI. Limitations and Future Work

Although the ANN-based model demonstrated strong predictive capability, several limitations of the present study must be acknowledged. First, the dataset consisted of 245 samples collected over only seven days, which may not capture the full variability of seasonal irradiance, temperature shifts, and diverse weather patterns. As a result, the generalization of the model under extreme or atypical environmental conditions cannot be guaranteed. Second, the experiment was performed using a single monocrystalline PV module installed at a fixed tilt angle and location. This limits model transferability to other module technologies (polycrystalline, bifacial, thin-film), different tilt orientations, or geographic regions with distinct climatic characteristics. Third, the measurements relied on a single set of instruments (Fluke SMFT-1000 and IRR2-BT), introducing instrument-specific biases and limiting cross-validation with alternative measurement equipment. Fourth, humidity effects and transient environmental variations—such as passing clouds, wind cooling, and rapid temperature swings—were not explicitly modeled, potentially contributing to remaining prediction errors.

From a modeling perspective, ANNs inherently operate as "black-box" approximators, providing limited interpretability of the underlying physical relationships. Although the Levenberg–Marquardt algorithm offered fast convergence, it may not scale efficiently to larger datasets or deeper architectures. Furthermore, while early stopping and validation monitoring were used to mitigate overfitting, the relatively small dataset size may still constrain the robustness of the learned representations.

To address these limitations, several directions for future research are recommended. First, a long-term dataset spanning multiple months or seasons should be collected to capture broader environmental variability and strengthen model generalization. Second, future models may incorporate additional meteorological features, such as wind speed, spectral irradiance, cloud cover indices, or particulate matter concentration, which influence PV thermal behavior and performance. Third, expanding the study to include multiple PV module types and orientations will support development of a more universal prediction framework. Fourth, hybrid modeling approaches—combining ANNs with physical PV models or integrating feature selection methods—may improve interpretability and reduce data requirements. Fifth, cross-testing with alternative machine learning techniques such as LSTM networks, convolutional models, or ensemble methods may further enhance accuracy, particularly under rapidly changing conditions.

Finally, implementing the ANN model in a real-time monitoring platform and validating its performance in operational PV systems would provide practical insights into computational requirements, prediction latency, and integration potential for predictive maintenance and grid-support applications. These future extensions will enable more comprehensive, scalable, and generalizable PV electrical behavior prediction frameworks.

VII. Practical Implications and Engineering Significance

The findings of this study have several important implications for photovoltaic (PV) system engineering, field diag-

nostics, and performance monitoring. First, the demonstrated ability of the ANN model to accurately predict open-circuit voltage (V_{oc}) and short-circuit current (I_{sc}) provides a foundation for real-time estimation of key electrical parameters without requiring frequent physical measurements. Since V_{oc} and I_{sc} are fundamental indicators of PV module health and performance, an ANN-based estimator can serve as a valuable tool for continuously assessing system condition in both rooftop and utility-scale PV installations.

Second, the model offers a scalable and cost-effective alternative to hardware-intensive monitoring systems. Traditional performance verification requires periodic I–V curve tracing, manual testing, or installation of module-level sensors. In contrast, an ANN-based virtual sensor can infer PV electrical behavior from readily accessible meteorological measurements, reducing reliance on expensive instrumentation. This capability enables remote monitoring, condition-based maintenance, and fault detection, contributing to longer module lifetimes and improved system reliability.

Third, the strong predictive performance of the ANN—particularly for I_{sc}—supports its use in energy forecasting and inverter control strategies. Accurate current estimation improves MPPT algorithm stability, enhances inverter efficiency, and minimizes mismatch losses. Such forecasting capabilities are valuable for grid-connected PV systems, where reliable predictions contribute to better demand-side management, smoother grid integration, and enhanced operational planning.

Fourth, the proposed approach supports early detection of module degradation, shading issues, or thermal anomalies. Deviations between predicted and measured V_{oc} or I_{sc} can indicate potential faults such as hotspot formation, delamination, or diode failure. This positions ANN-based prediction as a viable complementary tool in advanced PV diagnostic workflows, improving overall operation and maintenance (O&M) efficiency by enabling proactive intervention before failures escalate.

Finally, the model's deployment potential extends to smart PV monitoring platforms, Internet of Things (IoT)-integrated sensor networks, and digital twins for renewable energy systems. The computational requirements of the optimized ANN (4–7–7–2 architecture) are low, enabling implementation on embedded devices, microcontrollers, or edge-computing systems. This contributes to the emerging direction of AI-driven solar monitoring solutions that emphasize low-power operation, fast inference time, and scalable integration across diverse PV installations.

Collectively, these practical implications highlight the engineering significance of the developed ANN model. By providing accurate, low-cost, and real-time estimations of key PV electrical parameters, the approach contributes to more robust system monitoring, enhanced energy yield assessment, improved maintenance planning, and overall operational optimization of photovoltaic systems.

ACKNOWLEDGMENT

The authors appreciate the Pamantasan ng Lungsod ng Valenzuela (PLV), Valenzuela City Local Government Unit (LGU), and Tri Axis Engineering Services (Solar PV System Testing and Commissioning Team) for the assistance significant in the accomplishment of this study.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this research work. All analyses, results, and conclusions were conducted independently and without any external influence.

DATA AVAILABILITY STATEMENT

The dataset used in this study consists of 245 real outdoor measurements collected during the seven-day experimental campaign. Due to institutional policies, the raw dataset is available upon reasonable request from the corresponding author. Processed data and MATLAB scripts used for ANN training may also be provided upon request.

AUTHOR CONTRIBUTIONS

The authors contributed to this study as follows:

Engr. Jordan N. Velasco (Lead Author): Conceptualization, methodology design, data curation, ANN model development, formal analysis, manuscript writing, and overall project supervision.

Edzel Grant Asis: Data acquisition, sensor setup, instrumentation calibration, preprocessing, and technical validation of results.

Engr. Alex Monsanto: Experimental setup coordination, PV module installation, field testing support, and review of methodological accuracy.

Dr. Alexis John M. Rubio: Technical guidance, academic review, enhancement of ANN modeling framework, and refinement of manuscript structure.

Mary Anne Trinidad: Statistical analysis, visualization preparation, formatting assistance, and literature review support.

Dr. Maria Amelia E. Damian: Oversight of research alignment with institutional standards, critical manuscript review, and approval of the final version for submission.

Dr. Nelson C. Rodelas: Technical guidance, ANN simulation and analysis.

Dr. Joan P. Lazaro: Refinement of manuscript structure, proofreading.

REFERENCES

[1] D. K. Chaturvedi and S. Sharma, "An experimental study and verification of the facts related to factors affecting the performance of solar PV systems," in *2015 Fifth International Conference on Communication Systems and Network Technologies*, Gwalior, India, 2015, pp. 1185–1188, doi: 10.1109/CSNT.2015.186.

[2] F. O. Hocaoglu et al., "Comparison of experimentally obtained I-V curves of different PV modules," in *2018 9th International Renewable Energy Congress (IREC)*, Hammamet, Tunisia, 2018, pp. 1–4, doi: 10.1109/IREC.2018.8362472.

[3] A. Mellit and S. A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications: A review," *Progress in Energy and Combustion Science*, vol. 34, no. 5, pp. 574–632, 2008, doi: 10.1016/j.pecs.2008.01.001.

[4] A. Loukriz, D. Saigaa, A. Kherbachi, M. Koriker, A. Bendib, and M. Drif, "Prediction of photovoltaic panels output performance using artificial neural network," *International Journal of Energy Optimization and Engineering*, vol. 11, no. 2, pp. 1–16, 2022, doi: 10.4018/IJEOE.309417.

[5] A. Alcañiz, D. Grzebyk, H. Ziar, and O. Isabella, "Trends and gaps in photovoltaic power forecasting with machine learning," *Energy Reports*, vol. 9, pp. 447–471, 2023, doi: 10.1016/j.egyr.2022.11.208.

[6] L.-L. Li, S.-Y. Wen, M.-L. Tseng, and C.-S. Wang, "Renewable energy prediction: A novel short-term prediction model of photovoltaic output power," *Journal of Cleaner Production*, vol. 228, pp. 359–375, 2019, doi: 10.1016/j.jclepro.2019.04.331.

[7] M. Trigo-González et al., "Hourly PV production estimation by means of an exportable multiple linear regression model," *Renewable Energy*, vol. 135, pp. 303–312, 2019, doi: 10.1016/j.renene.2018.12.014.

[8] J. I. Rosell and M. Ibáñez, "Modelling power output in photovoltaic modules for outdoor operating conditions," *Energy Conversion and Management*, vol. 47, no. 15–16, pp. 2424–2430, 2006, doi: 10.1016/j.enconman.2005.11.004.

[9] M. Khrisat and Z. Alqadi, "Performance evaluation of ANN models for prediction," *Acadlore Transactions on AI and Machine Learning*, vol. 2, no. 1, pp. 13–20, 2023, doi: 10.56578/ataiml020102.

[10] H. Pham, "A new criterion for model selection," *Mathematics*, vol. 7, no. 12, pp. 1–13, Dec. 2019, doi: 10.3390/math7121215.

[11] A. Alibakshi, "Strategies to develop robust neural network models: Prediction of flash point as a case study," *Analytica Chimica Acta*, vol. 1026, pp. 69–76, Oct. 2018, doi: 10.1016/j.aca.2018.05.015.

[12] J. Bilski, B. Kowalczyk, A. Marchlewska, and J. M. Zurada, "Local Levenberg–Marquardt algorithm for learning feedforward neural networks," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 10, no. 4, pp. 299–316, 2020, doi: 10.2478/jaiscr-2020-0020.

[13] M. G. De Giorgi, P. M. Congedo, and M. Malvoni, "Photovoltaic power forecasting using statistical methods: Impact of weather data," *IET Science, Measurement & Technology*, vol. 8, no. 3, pp. 90–97, 2014, doi: 10.1049/iet-smt.2013.0135.

[14] A. Dolara, F. Grimaccia, S. Leva, M. Mussetta, and E. Ogliari, "A physical hybrid artificial neural network for short term forecasting of PV plant power output," *Energies*, vol. 8, no. 2, pp. 1138–1153, 2015, doi: 10.3390/en8021138.

[15] M. P. Almeida, O. Perpiñán, and L. Narvarte, "PV power forecast using a nonparametric PV model," *Solar Energy*, vol. 115, pp. 354–368, May 2015, doi: 10.1016/j.solener.2015.03.006.

[16] A. Qasrawi and M. Awad, "Prediction of the power output of solar cells using neural networks," *International Journal of Computer Science and Security*, vol. 9, no. 2, pp. 75–85, 2015. Available: https://www.cscjournals.org/manuscript/Journals/IJCSS/Volume9/Issue6/IJCSS-1159.pdf.

[17] G. Graditi, M. L. D'Agostino, and V. F. Pandiscia, "Comparison of photovoltaic plant power production prediction methods using a large measured dataset," *Renewable Energy*, vol. 97, pp. 446–457, 2016, doi: 10.1016/j.renene.2016.01.027.

[18] S. Pulipaka, B. M. P. Kumar, and S. S. Chandel, "Modeling of soiled PV module with neural networks and regression using particle size composition," *Solar Energy*, vol. 140, pp. 359–370, 2016, doi: 10.1016/j.solener.2015.11.012.

[19] S. Leva, A. Dolara, F. Grimaccia, M. Mussetta, and E. Ogliari, "Analysis and validation of 24 hours ahead neural network forecasting of photovoltaic output power," *Mathematics and Computers in Simulation*, vol. 131, pp. 88–100, 2017, doi: 10.1016/j.matcom.2015.05.010.

[20] L. Liu et al., "Forecasting power output of photovoltaic system using a BP network method," *Energy Procedia*, vol. 105, pp. 264–270, 2017, doi: 10.1016/j.egypro.2017.12.126.

[21] S. Moslehi and N. A. Rahim, "Evaluation of data-driven models for predicting solar photovoltaics power output," *Energy*, vol. 169, pp. 732–741, 2019, doi: 10.1016/j.energy.2017.09.042.

[22] M. H. Alomari, M. K. B. M. Amin, and S. H. M. Ali, "Solar photovoltaic power forecasting in Jordan using artificial neural networks," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 6, pp. 5068–5077, 2018, doi: 10.11591/ijece.v8i1.pp497-504

[23] M. K. Behera, K. R. Swain, and A. Nayak, "Solar photovoltaic power forecasting using optimized modified extreme learning machine technique," *Engineering Science and Technology, an International Journal*, vol. 21, no. 3, pp. 428–438, 2018, doi: 10.1016/j.jestch.2018.04.013.

[24] S. P. Durrani, M. A. Khan, and F. Khan, "Photovoltaic yield prediction using an irradiance forecast model based on multiple neural networks," *Journal of Modern Power Systems and Clean Energy*, vol. 6, no. 4, pp. 763–776, 2018, doi: 10.1007/s40565-018-0393-5.

[25] I. Kayri and M. T. Gencoglu, "Predicting power production from a photovoltaic panel through artificial neural networks using atmospheric indicators," *Neural Computing and Applications*, vol. 31, pp. 281–291, 2019, doi: 10.1007/s00521-017-3271-6.

[26] D. A. R. de Jesús, P. Mandal, S. Chakraborty, and T. Senjyu, "Solar PV power prediction using a new approach based on hybrid deep neural network," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*, Atlanta, GA, USA, 2019, pp. 1–5, doi: 10.1109/PESGM40551.2019.8974091.

[27] J. López Gómez, R. Sánchez Ureña, and M. R. Araújo, "Photovoltaic power prediction using artificial neural networks and numerical weather data," *Sustainability*, vol. 12, no. 14, p. 5741, 2020, doi: 10.3390/su122410295.

[28] A. Erduman, "A smart short-term solar power output prediction by artificial neural network," *Electrical Engineering*, vol. 102, pp. 1575–1583, 2020, doi: 10.1007/s00202-020-00971-2.

[29] M. Wang et al., "Comparison of different simplistic prediction models for forecasting PV power output: Assessment with experimental measurements," *Energy*, vol. 232, p. 120983, 2021, doi: 10.1016/j.energy.2021.120162.

[30] M. Ayan and B. Toylan, "Estimating the power generation of a stand-alone solar photovoltaic panel using ANNs and statistical methods," *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, vol. 43, no. 20, pp. 2487–2497, 2021, doi: 10.1080/15567036.2020.1849459.

[31] M. K. Park, J. H. Park, and S. W. Lee, "Predictive model for PV power generation using RNN (LSTM)," *Journal of Mechanical Science and Technology*, vol. 35, no. 7, pp. 3371–3379, 2021, doi: 10.1007/s12206-021-0140-0.

[32] J. N. Velasco and C. F. Ostia, "Development of a neural network based PV power output prediction application using reduced features and Tansig activation function," in *2020 6th International Conference on Control, Automation and Robotics (ICCAR)*, Singapore, 2020, pp. 732–735, doi: 10.1109/ICCAR49639.2020.9108101.

[33] J. N. Velasco, R. D. Trinidad, R. Z. Ramos, and K. L. M. De Jesus, "Neural network-based sensitivity analysis of the factors affecting the solar photovoltaic power output," in *2023 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)*, Shah Alam, Malaysia, 2023, pp. 304–309, doi: 10.1109/I2CACIS57635.2023.10193378.