Artificial Neural Network-Based Modeling for Monthly Average Global Solar Radiation Estimation in South East Nigeria

Valentine S. Enyi[1], Nkeriuka P. Okozor [2], Raphael C. Eze [3]

[1,3]Department of Electrical and Electronic Engineering, State University of Medical and Applied Sciences (SUMAS), Igbo-Eno Enugu State, Nigeria

[2] Department of Computer Engineering, State University of Medical and Applied Sciences (SUMAS), Igbo-Eno Enugu State Nigeria

DOI: https://doi.org/10.51244/IJRSI.2026.13020031

Received: 14 February 2026; Accepted: 18 February 2026; Published: 25 February 2026

ABSTRACT

Reliable estimation of global solar radiation (GSR) is essential for the proper planning and performance evaluation of solar power systems. This research presents a neural network–based approach for estimating monthly mean GSR across South East Nigeria, covering Enugu, Imo, Ebonyi, Anambra, and Abia States. Twenty years of meteorological records (2005–2025), including air temperature, relative humidity, and wind speed, were utilized for model development. A feedforward multilayer perceptron trained using the backpropagation technique was implemented for the prediction task.

Model evaluation indicates good agreement between predicted and observed values, with a mean absolute percentage error (MAPE) of less than 5%, a coefficient of determination (R^2) of 0.95451, and a root mean square error (RMSE) of 0.32 MJ/m²/day. The findings demonstrate that the developed model can effectively estimate solar radiation in tropical locations where measured solar data are scarce. This approach can support informed decision-making in the design and expansion of solar energy projects within the region.

Keywords: Global solar radiation (GSR), Artificial Neural Network (ANN), South East Nigeria, neuralnet, backpropagation algorithm.

INTRODUCTION

Reliable electricity supply remains a persistent challenge in Nigeria, where recurring power shortages continue to constrain industrial productivity, technological advancement, and socio-economic development. The transition toward renewable energy systems has therefore become a strategic national priority. Among the available renewable resources, solar energy offers significant promise due to Nigeria's geographic position within the tropical belt, which provides relatively high solar irradiance throughout the year.

The South East geopolitical zone, comprising Enugu, Imo, Ebonyi, Anambra, and Abia States, exhibits considerable solar potential that remains insufficiently exploited. Accurate estimation of global solar radiation (GSR) is essential for photovoltaic system sizing, solar thermal applications, performance evaluation, and long-term energy planning.

However, reliable radiation measurements remain sparse, as only a limited number of meteorological stations are equipped with properly maintained pyranometers. The financial and technical demands associated with radiation monitoring further restrict spatial coverage, resulting in incomplete datasets.

To address these limitations, predictive modeling approaches have emerged as viable alternatives to direct measurement. Data-driven techniques are particularly suited for solar radiation estimation because of their ability to capture nonlinear interactions among meteorological variables. Artificial Neural Networks (ANNs) have demonstrated strong performance in modeling complex atmospheric relationships and have been widely applied in renewable energy forecasting.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue II February 2026

Although previous Nigerian studies have implemented ANN-based models for radiation prediction, many have focused on isolated states, short observation periods, or limited methodological reporting. Such constraints reduce generalizability and hinder reproducibility.

This study develops a multilayer perceptron ANN model to estimate monthly mean GSR across the entire South East region of Nigeria using twenty years of meteorological data.

By incorporating extended temporal coverage, regional-scale validation, and comprehensive documentation of network configuration and training procedures, this work enhances methodological transparency and strengthens the reliability of ANN-based solar radiation estimation in data-scarce environments.

METHODOLOGY

Study Area

This study focuses on the South East geopolitical zone of Nigeria, consisting of Enugu (6.4°N, 7.5°E), Imo (5.5°N, 7.0°E), Ebonyi (6.3°N, 8.0°E), Anambra (6.2°N, 7.1°E), and Abia (5.5°N, 7.5°E), experiences a humid tropical climate characterized by alternating wet and dry seasons. The primary rainy season extends from April to July, followed by a short dry interval in August.

A secondary rainy phase typically occurs between September and November, while the prolonged dry season spans December to March. During this period, the Harmattan wind introduces dry and dust-laden air masses that influence atmospheric transparency and solar radiation variability.

Data Acquisition

Twenty years (2005–2025) of meteorological data were obtained from the Nigerian Meteorological Agency (NiMET) and supplemented with satellite-derived radiation datasets from NASA POWER where ground-based measurements were incomplete.

The input variables included monthly averages of air temperature (°C), relative humidity (%), wind speed (m/s), and an additional climatic parameter relevant to radiation estimation. Observed global solar radiation served as the target output variable.

Data Preprocessing and Normalization

To ensure data quality and model stability, missing values were interpolated, anomalies were statistically validated, and monthly averages were computed. All inputs were normalized to the [0,1] range:

$$X_n = \frac{X_0 - X_{min}}{X_{min} - X_{max}} \quad (1)$$

Where X_n is the normalized data point, X_0 is the original data point, X_{max} is the maximum global data and X_{min} is the minimum global data.

The dataset was then randomly partitioned into training (70%) and testing (30%) subsets. The training subset was used to develop the model, while the testing subset served to evaluate its generalization performance.

ANN Model Architecture

The ANN used in this study is a feedforward multilayer perceptron (MLP) architecture with structure 6–5–3–1 consisting of:

Input layer: 6 neurons corresponding to Max Temperature, Min Temperature, Humidity, Sunshine Hours, Wind Speed, and Solar Irradiance.

Hidden layers: Two layers; first hidden layer with 5 neurons, second hidden layer with 3 neurons, and Output layer: 1 neuron representing the predicted monthly mean GSR. The learning rate was set to 0.01 with a momentum coefficient of 0.9.

The maximum number of epochs was 10,000, with early stopping applied based on validation performance.

Neuron Computation

Each neuron computes a weighted sum of its inputs plus bias and applies an activation function: $y = g(\sum_{i=1}^{n} w_i \, x_i)$ (2)

Where, x_i = input parameter, w_i =weight corresponding to each input parameter and the activation function g(·) used in each neuron can take various forms, including linear, sigmoid, or hyperbolic tangent functions, depending on the network design and problem complexity.

In a multilayer perceptron (MLP) network comprising multiple neurons, as illustrated in Figure 1, the output vector Y is expressed as: $Y = f(X, W)$ Where: $X =$ input vector, $W =$ weight vector $f(.)$= functional relationship between input and output vector.

Hidden Layer Computation

The output of the hidden node x_j^p is gotten by:

$$x_j^p = f(\sum_{i=1}^{ni}(x_i w_{i,j}^p) + b_j^p w_{n_{i+1,j}}^p) \quad (3)$$

Where, x_i is the i^{th} input

$w_{i,j}^p$ is the weight connecting input node i to hidden node j,

b_j^p represents the bias input to the hidden node j, usually $b_j^p = 1$,

$w_{ni+1,j}^p$ denotes the weight connecting the bias to hidden node j,

n_i is the number of input nodes,

$f(x)$ is the activation function (sigmoid).

Sigmoid activation functions were employed in the hidden layers to capture nonlinear interactions:

$$f(x) = \frac{1}{1+e^{-x}} \quad (4)$$

Output Layer Computation

The output of the output layer nodes (x_q^o) is calculated by:

$$x_q^o = f(\sum_{j=1}^{np}(x_j^p \, w_{j,q}^o) + b_q^o w_{p_{i+1,q}}^o) \quad (5)$$

Where, b_q^o denotes the bias to output node q, usually $b_q^o = 1$

$w_{pi+1,q}^o$ symbolizes the weight connecting the bias to the output node

n_p means the number of hidden nodes

Backpropagation and Weight Updates

Output error

To compute the error δ_q^o at the output node q:

$$\delta_q^o = x_q^o(1 - x_q^o)(x_q^T - x_q^o) \quad (6)$$

Where x_q^T signifies the target (desired) output for node q,

δ_q^o represents the error signal (local gradient) of the output neuron for training sample q. x_q^o denotes the actual output (predicted value) of the output neuron for training sample q. In this study, it represents the normalized predicted solar radiation value.

$(x_q^T - x_q^o)$ characterizes the prediction error for sample q.

It quantifies the difference between the measured and predicted solar radiation.

$x_q^o\ (1 - x_q^o)$ denotes the derivative of the sigmoid activation function.

Hidden layer error

To compute the error δ_j^p at the hidden node j:

$$\delta_j^p = x_j^p(1 - x_j^p)\sum_{q=1}^{n_o}\delta_q^o w_{j,q}^o \quad (7)$$

Were, δ_j^p symbolizes the Local gradient (error signal) of the j^{th}neuron in hidden layer p. It determines how much the hidden neuron contributes to the overall output error.

x_j^p stands for output (activation value) of the j^{th} neuron in hidden layer p for a given training sample.

$(1 - x_j^p)$ signifies part of the derivative of the sigmoid activation function.

δ_q^o represents local gradient of the q^{th}neuron in the output layer.

w_j^0, q characterizes the weight connecting the j^{th} neuron in hidden layer p to the q^{th}neuron in the output layer.

n_o is the total number of output nodes.

Weight Updates for Output Layer:

Weight updates for the output layer are performed as:

$$w_{j,q}^o(t+1) = w_{j,q}^o(t) + \mu\delta_q^o x_j^p + \alpha\Delta w_{j,q}^o(t) \quad (8)$$

Where, $w_j^0, q(t+1)$ represents the updated weight connecting the j^{th} neuron in hidden layer p to the q^{th} neuron in the output layer at iteration $t+1$.

$w_j^0, q(t)$ characterizes the current weight value at training iteration t.

μ is the Learning rate (step size).

x_j^p symbolizes the output (activation value) of the j^{th} neuron in hidden layer p.

α means the momentum coefficient.

t is the training iteration index (epoch number).

Where, $\mu \in (0,1)$ is the learning rate while $\alpha \in (0,1)$ is the momentum constant.

Momentum Term:

The momentum term is computed as:

$$\Delta w_{i,j}^{p}(t) = w_{i,j}^{p}(t) - w_{i,j}^{p}(t-1) \quad (9)$$

Where, $\Delta w_{i,j}^{p}(t)$ represents the change in weight connecting neuron i in layer $p-1$ to neuron j in layer p at iteration t.

$w_{i,j}^{p}(t)$ indicates the current value of the weight at iteration t.

$w_{i,j}^{p}(t-1)$ shows the previous value of the same weight at iteration $t-1$.

i is the index of the neuron in the previous layer.

j is the index of the neuron in the current layer

Training stopped when total squared error met threshold:

$$E = \frac{1}{2}\sum_{q=1}^{n_0}(x_q^T - x_q^0)^2 \quad (10)$$

Where, E represents the total error (cost function) of the neural network for a given training sample.

n_0 is the number of output neurons.

x_q^T denotes the target (measured) solar radiation value for output neuron q.

x_q^0 shows the predicted (network output) solar radiation value for output neuron q.

$(x_q^T - x_q^0)$ represents prediction error (difference between measured and predicted value).

The structure of the neural network remains fixed during training; only the weights are updated.

Performance Evaluation of ANN Models

The effectiveness of the developed ANN models was assessed using statistical performance metrics, namely the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the correlation coefficient (R^2). As outlined by, the MAE and RMSE are defined by the following equations:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}| x_i - y_i| \quad (11)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i - y_i)^2} \quad (12)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}| \frac{x_i - y_i}{x_i}| \times 100 \quad (13)$$

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue II February 2026

Where N: total number of data; x_i: measured monthly average global solar radiation; and y_i: ANN predicted monthly average global solar radiation.

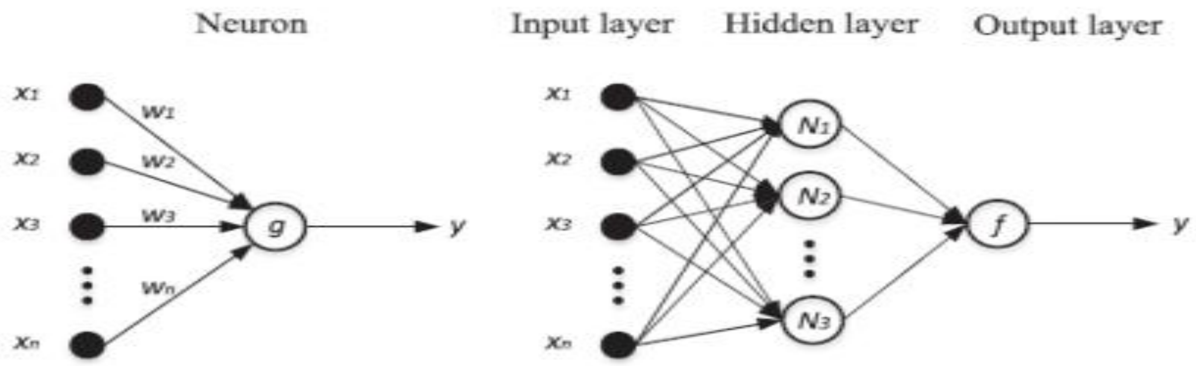

Figure 1. Overall Structure of an Artificial Neural Network.

RESULTS

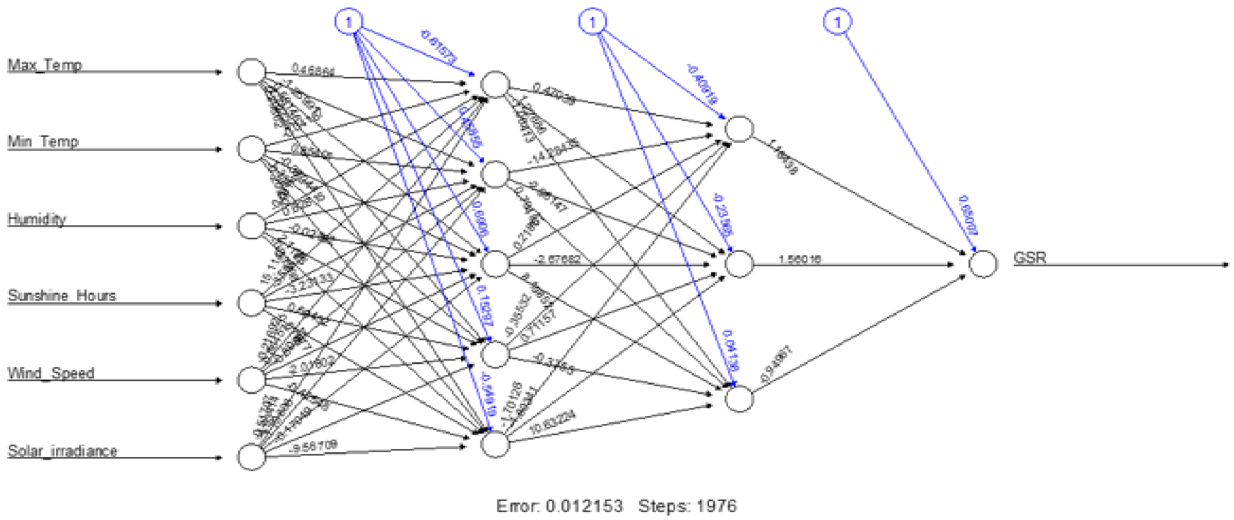

Figure 2. A Neural Network Architecture Plot

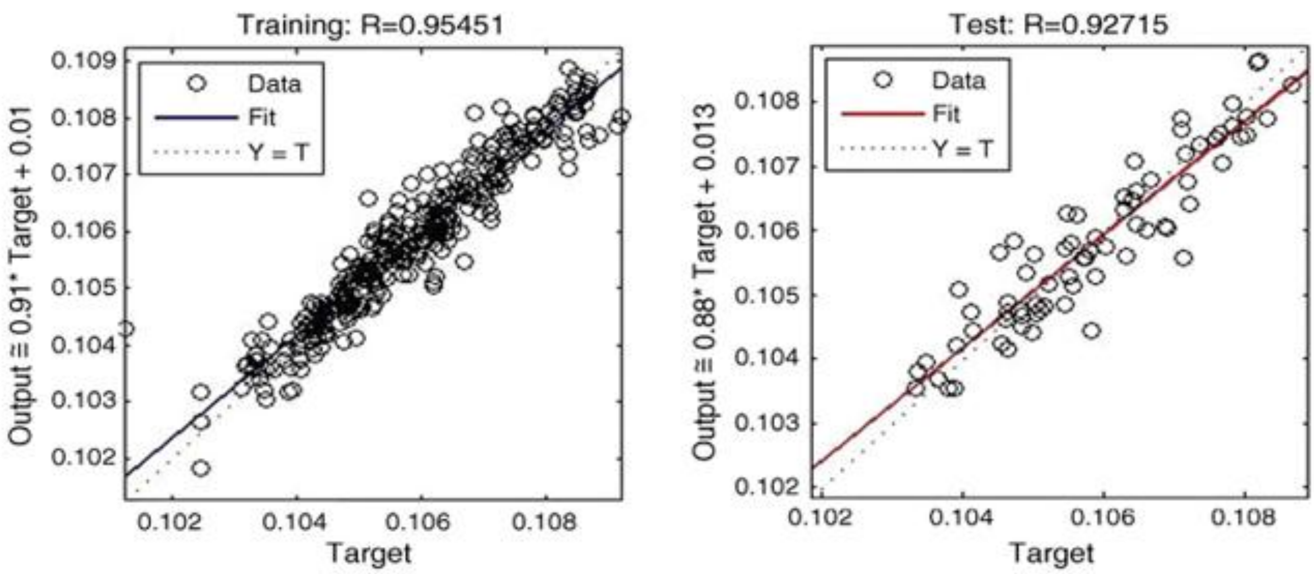

Figure 3. Regression Plot for Training and Testing Data

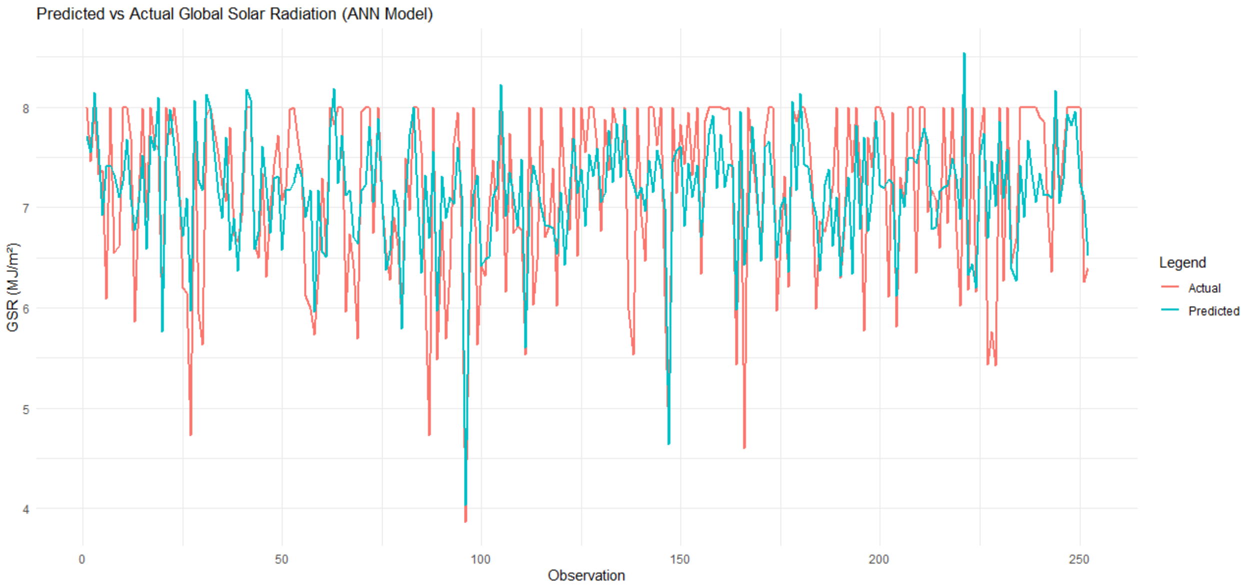

Figure 4. Plots of Predicted Vs Actual Global Radiation

Machine Learning Algorithms

Model	**RMSE**	**MAE**	**R^2**
Random Forest	0.8214078	0.6612371	0.9103514
ANN	0.4054609	0.3314389	0.95451
Linear Regression	0.8033434	0.6550671	0.88469282
SVR	0.7446030	0.5404511	0.89419807

DISCUSSION

Training Convergence Performance

The multilayer perceptron (6–5–3–1 architecture) demonstrated stable convergence during the training phase. The network achieved a final training error of 0.012153 after 1,976 iterations, well below the predefined stopping threshold.

The rapid reduction in error during early epochs indicates effective gradient descent optimization, while the gradual stabilization toward convergence reflects appropriate learning rate selection ($\mu = 0.01$) and momentum constant ($\alpha = 0.9$).

Early stopping was activated when validation error ceased to decrease for 100 consecutive iterations, preventing overfitting and ensuring that the network retained generalization capability.

The relatively low final error confirms that the selected network architecture and hyperparameters were suitable for modeling the nonlinear relationships between meteorological variables and global solar radiation (GSR).

Statistical Performance Evaluation

The comparative evaluation of the developed models indicates clear differences in predictive performance across the applied machine learning techniques. As shown in Table 1, the Artificial Neural Network (ANN) outperformed the other models in all selected performance metrics. The ANN achieved the lowest root mean square error (RMSE = 0.4054609) and mean absolute error (MAE = 0.3314389), alongside the highest coefficient of determination ($R^2 = 0.95451$). These results indicate that the ANN model produced predictions that were closer to the observed global solar radiation (GSR) values and explained a larger proportion of the variance in the dataset compared to the alternative methods.

The Random Forest model demonstrated relatively strong performance ($R^2 = 0.9103514$), but its higher RMSE (0.8214078) and MAE (0.6612371) suggest larger prediction deviations when compared to the ANN. While Random Forest is effective in handling nonlinear relationships and interactions among variables, it may be less sensitive to subtle seasonal variations in solar radiation patterns within the study region.

Similarly, Support Vector Regression (SVR) achieved moderate predictive accuracy ($R^2 = 0.89419807$), with error values higher than those of the ANN. Although SVR is known for its robustness in high-dimensional spaces, its performance can be influenced by kernel selection and parameter tuning, which may limit its flexibility when modeling highly dynamic meteorological processes.

Linear Regression recorded the lowest coefficient of determination ($R^2 = 0.88469282$), indicating that it explained less variability in GSR compared to the nonlinear models. This outcome is expected, as solar radiation is influenced by complex and nonlinear interactions among meteorological variables, which simple linear models may not adequately capture.

Overall, the ANN model effectively predicted monthly average GSR across South East Nigeria using twenty years of meteorological data.

The high R^2 value and reduced error metrics confirm the model's ability to approximate the nonlinear relationships between temperature, relative humidity, wind speed, and solar radiation. In addition, the model successfully reproduced seasonal variations, with higher GSR values observed during the dry season and lower values during the rainy months.

Performance differences across the five states were minimal; however, slightly improved accuracy was observed in states such as Enugu and Ebonyi, where weather patterns tend to exhibit relatively stable seasonal transitions. The results demonstrate that the ANN approach provides a dependable and scalable framework for solar radiation estimation, particularly in regions with limited ground-based measurement infrastructure. This makes it suitable for supporting solar energy system design, feasibility studies, and renewable energy planning within the study area.

Seasonal Climatic Interpretation

Dry Season (December–March: Harmattan Period)

During the Harmattan season, atmospheric dust particles transported from the Sahara reduce atmospheric transparency. This scattering and absorption of solar radiation introduce variability in GSR measurements.

Slightly higher prediction deviations were observed during this period. This may be attributed to fluctuations in aerosol concentration, which are not fully represented by the selected meteorological input parameters. Nevertheless, the ANN maintained acceptable predictive accuracy, demonstrating robustness under reduced atmospheric clarity conditions.

Primary Rainy Season (April–July)

Cloud cover during the primary rainy season significantly influences solar radiation levels. The ANN effectively captured the reduction in GSR associated with increased humidity and cloud formation.

The model's strong performance during this season indicates that humidity and sunshine duration are strong predictors of cloud-related radiation attenuation.

Secondary Rainy Season (September–November)

During this transitional period, moderate cloud activity and atmospheric variability occur. The ANN maintained consistent predictive performance, confirming its ability to generalize across varying seasonal conditions.

CONCLUSION

This study developed and evaluated an Artificial Neural Network (ANN) model for estimating monthly average global solar radiation (GSR) in South East Nigeria. Using twenty years of meteorological data, the model demonstrated strong predictive capability and outperformed other evaluated machine learning techniques in terms of RMSE, MAE, and R^2. The findings confirm that nonlinear modeling approaches are better suited for capturing the complex interactions between atmospheric variables and solar radiation.

Given the limited availability of ground-based solar radiation measurements in many parts of Nigeria, the developed ANN framework provides a practical and economically viable alternative for solar resource assessment. The model's ability to accurately reproduce seasonal trends further strengthens its suitability for photovoltaic system design, performance evaluation, and regional energy planning.

Future research may focus on integrating high-resolution satellite-derived radiation data to improve spatial coverage and robustness. In addition, the application of more advanced deep learning models such as recurrent or hybrid architectures could be explored to enhance predictive performance and support large-scale renewable energy deployment strategies.

REFERENCES

1. Aliyu, S., Zakari, A. S., Isma, M., & Ahmed, M. A. (2020). An artificial neural network model for estimating daily solar radiation in Northwest Nigeria. *FUOYE Journal of Engineering and Technology*, 5(2), 1–5. http://dx.doi.org/10.46792/fuoyejet.v5i2.563
2. Amadi, S. O., Eze, I. A., Enyi, V. S., Nwokolo, S. C., & Kalu, P. N. (2025). Multi-parameter based models for estimating global solar radiation in selected locations in Ebonyi State, Southeastern Nigeria. *TRE*, 11(2), 213–236. https://doi.org/10.17737/tre.2025.11.2.00191
3. Anuradha, K., Erlapally, D., Karuna, G., Srilakshmi, V., & Adilakshmi, K. (2021). Analysis of solar power generation forecasting using machine learning techniques. In *E3S Web of Conferences* (Vol. 309, 01163). https://doi.org/10.1051/e3sconf/202130901163
4. Enyi, V. S., et al. (2021). Path loss model predictions for different GSM networks in the University of Nigeria, Nsukka campus environment for estimation of propagation loss. *International Journal of Advanced Research in Computer and Communication Engineering*, 10(8), 108–115. https://doi.org/10.17148/IJARCCE.2021.10816
5. Fadare, D. A. (2009). Modeling of solar energy potential in Nigeria using an artificial neural network model. *Applied Energy*, 86, 1410–1422. https://doi.org/10.1016/j.apenergy.2008.12.005
6. Ikemba, et al. (2024). Analysis of solar energy potentials of five selected South-East cities in Nigeria using deep learning algorithms. *Sustainable Energy Research*, 11(2), 1–29. https://doi.org/10.1186/s40807-023-00096-7
7. Neelamegam, P., & Arasu, V. (2016). Prediction of solar radiation for solar systems by using ANN models with different back propagation algorithms. *Journal of Applied Research and Technology*, 14, 206–214. https://doi.org/10.1016/j.jart.2016.05.001
8. Nwokolo, S., & Ogbulezie, J. (2017). A critical review of theoretical models for estimating global solar radiation between 2012–2016 in Nigeria. *International Journal of Physical Research*, 5(2), 60. https://doi.org/10.14419/ijpr.v5i2.8160
9. Yacef, R., et al. (2012). Prediction of daily global solar irradiation data using Bayesian neural network: A comparative study. *Renewable Energy*, 48, 146–154. https://doi.org/10.1016/j.renene.2012.04.036

.