Development of a Non-Parametric Model for Low-Cost Pyranometer Based on Artificial Neural Network

Jonata C. de Albuquerque[1], Manoel H. de O. P. Filho[2], Ygo N. batista[2], Regina M. L. Neta[2] and Cassiano M. B. de Melo [2]

[1] Federal University of Pernambuco-UFPE, Recife (Brasil)

[2] Federal Institute of Pernambuco-IFPE, Pesqueira (Brasil)

Abstract

Low-cost pyranometers are widely used as a viable and robust alternative to effectively measure the amount of solar radiation incident on a given area during a specific time interval. Currently, artificial intelligence tools such as Artificial Neural Networks (ANNs) are frequently employed to integrate characteristics of physical magnitudes incident on these types of renewable systems. In this work, a non-parametric model of a pyranometer based on a Multi-Layer Perceptron (MLP) ANN was developed to simulate the irradiance curve characteristic of a pyranometer during the clear period of the day. This model uses electrical currents from three photodiodes with different spectral sensitivities as input signals to simulate the Global Horizontal Irradiance (GHI) curve measured by a conventional pyranometer. Data were collected from the INOVA SOLAR network at IFPE Campus Pesqueira, Brazil, covering approximately one month with 5-minute sampling intervals during clear-sky periods from 04:43h to 17:58h daily. The dataset was divided into training (85%) and simulation (15%) sets. The ANN inputs include 15 variables such as temporal information (day of the week encoded as one-hot vectors, day, hour, minute, sine and cosine transformations of minute), photodiode readings, and historical pyranometer GHI measurements. The output target is the broadband GHI measured by the pyranometer. The ANN architecture consists of two hidden layers with 1 and 3 neurons respectively, selected through iterative testing to minimize the Mean Squared Error (MSE) using the Levenberg-Marquardt training algorithm. The model achieved a Mean Absolute Percentage Error (MAPE) of 7.98% on the simulation dataset, demonstrating its feasibility. The model is not generative; retraining is recommended if data profiles change significantly. This approach offers a cost-effective alternative to traditional pyranometers with potential applications in solar energy and climatology.

Keywords: Low-cost pyranometers, Artificial neural network, Multilayer Perceptron network; Solar Radiation Measurement; Global Horizontal Irradiance (GHI)

1. Introduction

Ground-based measurements, specifically solarimetric data, play a crucial role in a wide range of applications, including (1) calibration of satellite models (Perez et al., 2002), (2) adjustment of satellite-derived values for specific locations through site adaptation (Bayasgalan et al., 2024), (3) solar radiation forecasting (Ye et al., 2022), and (4) power and energy output prediction in photovoltaic plants (Cervera-Gascó et al., 2022), among others.

Traditionally, the measurement of ground solar resources has been performed using pyranometers classified by ISO standards, which rely on either thermal (thermopile) or photovoltaic solar cell principles. Both types are based on a direct correlation between an electrical signal—voltage or current—and irradiance, expressed in W/m² (Shafa et al., 2015). Despite their accuracy, commercial pyranometers are often cost-prohibitive, particularly for small-scale and distributed generation systems, where their expense may rival that of the entire system. This financial barrier limits the widespread deployment of ground solar radiation measurement stations.

10.18086/swc.2025.04.16 Available at http://proceedings.ises.org

In response, many researchers have pursued the development of low-cost pyranometers using components such as thermal couplers, phototransistors, photodiodes, and photovoltaic cells (e.g., Haque et al., 2022; Tohsing et al., 2019). Building upon these advances, an innovative pyranometer concept is currently being developed within the FACEPE-funded "Rede Inova Solar" project (Pernambuco Science and Technology Support Foundation), aimed at fostering collaboration between public research institutions and private enterprises.

The core principle of this research is to identify an optimal set of photodiodes with different spectral sensitivities to capture distinct segments of the solar spectrum effectively. The electrical currents generated by these photodiodes serve as inputs to train an Artificial Neural Network (ANN), which outputs the Global Horizontal Irradiance (GHI).

This article presents the initial simulation results evaluating the ANN architecture and its training process, using data from a triad of photodiodes as inputs. The irradiance curve characteristic of the pyranometer, particularly during clear sky conditions, is modeled non-parametrically through a Multilayer Perceptron (MLP) ANN to enable a cost-effective and accurate measurement alternative.

This paper is organized as follows: Section 2 presents the dataset and details the non-parametric ANN model used to characterize the irradiance curve; Section 3 discusses the results and their analysis; Section 4 acknowledges the support of the funding research agencies; and finally, Section 5 lists the references.

2. Dataset description e configuration ANN model development

Artificial Neural Networks (ANNs) are parallel and distributed computational architectures composed of simple processing elements known as artificial neurons. These neurons are capable of performing mathematical operations—predominantly nonlinear—and are organized into layers that are interconnected through weighted connections. These weights encapsulate the learned knowledge of the model and modulate the input signals received by each neuron (Ludermir et al., 2000; Elsheikh and Elaziz, 2022).

Among the various ANN architectures, the Multilayer Perceptron (MLP) stands out due to its capacity to approximate nonlinear functions with high accuracy. MLP networks are characterized by the presence of one or more hidden layers, with feedforward connectivity and no feedback between layers (Haykin, 2000). These networks are widely applied across domains, from weather forecasting and optical character recognition to renewable energy modeling, owing to their learning and generalization capabilities.

In this study, an MLP ANN was developed to simulate the irradiance curve behavior of a conventional pyranometer using signals derived from low-cost photodiodes. These input signals were initially simulated based on the spectral sensitivity of the photodiodes and were correlated with solar spectral irradiance measurements acquired using the HR4000CG-UV-NIR spectrometer. The spectral measurements spanned wavelengths from 300 nm to 900 nm, with the system operating under solar zenith angles $\theta z < 80°$. Figure 1 presents an example of one such measurement, highlighting the spectral sensitivity curve of a photodiode and the corresponding photocurrent response obtained via numerical integration.

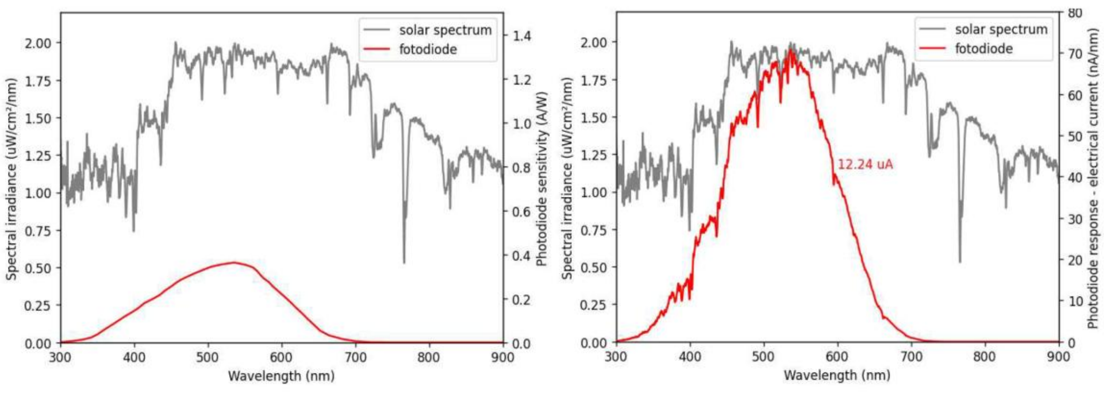

Figure 1: Measured solar spectrum (gray), photodiode sensitivity curve (red, left), and simulated output current (red, right).

2.1. Dataset Description

The dataset used in this work was collected from the INOVA SOLAR network located at IFPE Campus

Pesqueira, Pernambuco, Brazil. It consists of synchronized measurements from both low-cost photodiodes and a conventional pyranometer. The data were recorded at 5-minute intervals between November 13, 2024, and December 11, 2024, encompassing approximately one month of continuous monitoring.

Each record in the dataset contains the following fields:

DD: Day of measurement

MM: Month of measurement

YYYY: Year of measurement

H: Hour of measurement

Min: Minute of measurement

Fd1: Reading from Photodiode 1

Fd2: Reading from Photodiode 2

Fd3: Reading from Photodiode 3

GHIint: Spectrally integrated irradiance from photodiodes

GHImed: Measured Global Horizontal Irradiance (GHI) from the pyranometer

For the training process, only the clear-sky period was considered, defined in this study as the interval between 04:43h and 17:58h, resulting in 160 valid patterns per day during clear-sky conditions. Based on this subset, the dataset was split as follows:

Training Set: 85% of data

Simulation Set: 15% of data

2.2. ANN Model Development

The non-parametric model of the pyranometer irradiance curve is so called because it is not constrained by explicit equations; instead, it is represented by an MLP-type Artificial Neural Network with an architecture composed of two hidden layers—the first with 1 neuron and the second with 3 neurons. To determine this architecture, a methodology was established, the main features of which are outlined below.

The input vector for the ANN was constructed from 15 variables, selected based on the best-performing combination during preliminary experiments. These include both temporal information and sensor readings: [7 inputs] Day of the week (encoded using one-hot encoding, e.g., Monday: 1 0 0 0 0 0 0)

[1 input] Day of the month (DD)

[1 input] Hour (H)

[1 input] Minute (Min)

<u>Trigonometric sine representation of Date (Minute):</u>

[1 input] Minute Sin

$$Minute\ Sin = sin\left(\frac{2\pi}{60} * Min\right) \qquad \text{(eq. 1)}$$

<u>Trigonometric cosine representation of Date (Minute):</u>

[1 input] Minute Cos

$$Minute\ Cos = Cos\left(\frac{2\pi}{60} * Min\right) \qquad \text{(eq. 2)}$$

[3 input] Readings from Photodiodes 1, 2, and 3 (Fd1, Fd2, Fd3)

[1 output] Historical pyranometer GHI reading (GHImed)

To determine the optimal ANN architecture, a systematic trial-and-error approach was employed. The architecture research considered combinations of 1 to 10 neurons in each of the two hidden layers, with a single output neuron. The process was as follows:

a) Initialize an MLP with two hidden layers, each starting with one neuron.

b) Randomly assign weights and train the network using the Levenberg–Marquardt (LM) algorithm.

c) Computer the Mean Squared Error (MSE) for training, validation, and test sets.

d) Repeat steps b) and c) ten times per architecture to average the MSE values.

e) Increment the number of neurons in the second hidden layer (keeping the first fixed), repeating steps 2–4 until reaching 8 neurons.

f) Increment the number of neurons in the first hidden layer and repeat until they reach 8 neurons.

g) Select the configuration yielding the lowest MSE on the validation set.

The LM algorithm was preferred over alternatives like Resilient Propagation (Train RP) due to its superior accuracy, despite higher computational demands. Through iterative training and evaluation, the hyperbolic tangent (tanh) activation function was selected as the optimal transfer function for the neurons in the hidden layers, providing the best overall network performance.

The final selected model—referred to as RNA_BEST—features the following architecture:

[15 → 1-3 → 1], i.e., 15 inputs, 1 neuron in the first hidden layer, 3 neurons in the second, and 1 output neuron.

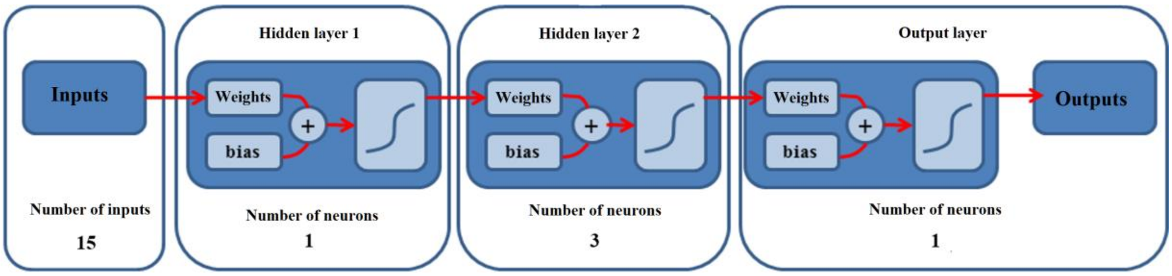

Figure 2: Final selected ANN architecture ("RNA_BEST") used for modeling the GHI irradiance curve.

3. Results and discussions

The performance of the proposed Artificial Neural Network (ANN) model was quantitatively evaluated using the Mean Absolute Percentage Error (MAPE), a metric that effectively quantifies the deviation between the measured Global Horizontal Irradiance (GHI) values and those estimated by the model. The evaluation considered multiple data subsets, including training, validation, testing, and simulation datasets, to ensure robustness and generalizability of the model's predictive capabilities. The obtained MAPE values demonstrate consistent accuracy, with errors below 10% across all subsets, indicating the ANN's successful capture of irradiance behavior throughout the study period.

<u>The MAPE was calculated using the following equation 3:</u>

$$MAPE = \frac{1}{n}\sum_{k=1}^{n}\left|\frac{GHImed(k) - \widehat{GHImed}(k)}{GHImed(k)}\right| * 100\% \qquad \text{(eq. 3)}$$

where $GHImed(k)$is the measured GHI at sample k, $\widehat{GHImed}(k)$is the corresponding ANN-predicted GHI, and n is the number of samples.

Table 1: Mean Absolute Percentage Error (MAPE) for different datasets.

MAPE (%)		
Training dataset	**Validation dataset**	**Test dataset**
8,51	8,18	8,83
Simulation dataset		
7,98		

As shown in Table 1, the MAPE values across all datasets remained below 9%, with the simulation set achieving the best accuracy at **7.98%**. This consistent performance across training, validation, and testing datasets indicates that the ANN generalizes well and reliably estimates the GHI from photodiode inputs and temporal features. The slightly lower error in the simulation set further highlights the model's robustness when applied to unseen data.

Figure 3 (left) illustrates a comparison between the historical GHI measurements, and the irradiance values predicted by the ANN during the simulation period. The close agreement observed confirms the model's capability to accurately replicate the irradiance curve measured by the pyranometer, capturing the dynamics of solar irradiance throughout the clear-sky periods.

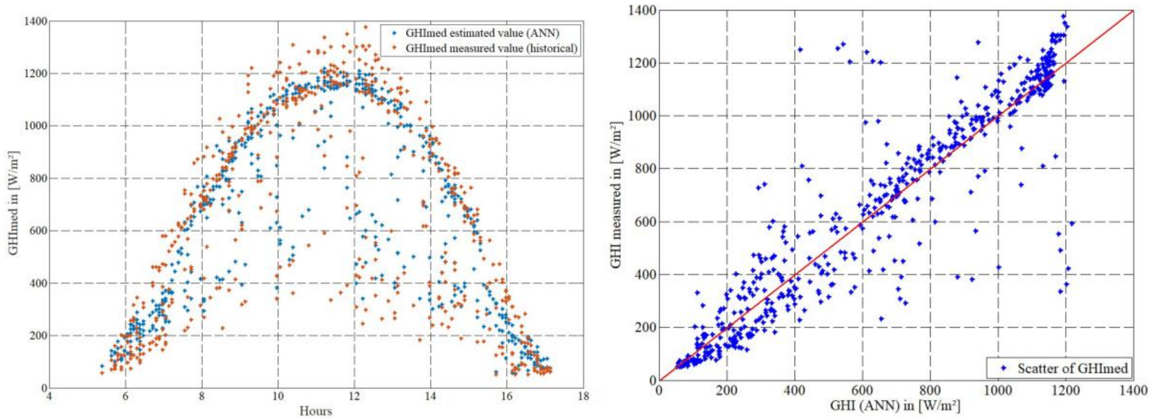

Figure 3: Comparison between the historical irradiance measurements and the irradiance values generated by the ANN. On the left, the distribution all day long profile. On the right, the dispersion between both values.

Furthermore, Figure 3 (right) presents a scatter plot of the simulated versus measured GHI values for the simulation dataset. The points are closely clustered around the ideal 45-degree line, indicating strong correlation and minimal bias in the ANN predictions. This visualization further supports the quantitative results, confirming that the ANN is capable of reliable broadband GHI estimation from low-cost photodiode inputs and temporal features.

The results confirm that the ANN can accurately estimate broadband GHI with errors below 10%. The model's non-parametric nature offers flexibility for various locations and conditions without relying on pre-defined equations. However, the model is not generative, so retraining is necessary when significant shifts in environmental conditions or sensor calibrations occur.

4. Acknowledgments

This work was supported by FACEPE and CNPq. We thank the INOVA SOLAR network team and IFPE Campus Pesqueira for their collaboration.

5. References

Bayasgalan, O., Adiyabat, A., Otani, K., Hashimoto, J. & Akisawa, A., 2024. A high-resolution satellite-based solar resource assessment method enhanced with site adaptation in arid and cold climate conditions. *Energies*, 17(24), p.6433. https://www.mdpi.com/1996-1073/17/24/6433

Cervera-Gascó, J., Perea, R.G., Montero, J. & Moreno, M.A., 2022. Prediction model of photovoltaic power in solar pumping systems based on artificial intelligence. *Agronomy*, 12(1). https://doi.org/10.3390/agronomy12010000 *(update link if necessary)*

Elsheikh, A.H. & Elaziz, M.E.A., 2022. *Artificial Neural Networks for Renewable Energy Systems and Real-World Applications*. Academic Press, New York. https://doi.org/10.1016/C2019-0-01290-3

Haque, M.E., Alvi, M.M., Rahman, M.F., Ali, M.H. & Haque, A.K.M.M., 2022. Cost effective alternative of pyranometer: Solar radiation prediction using artificial intelligence. In: *2022 International Conference on Recent Progresses in Science, Engineering and Technology (ICRPSET)*, pp. 1–5. https://doi.org/10.1109/ICRPSET57982.2022.10188547

Haykin, S.S., 2000. *Neural Networks*, 2nd ed. Bookman. Available at: https://books.google.com.br/books?id=lBp0X5qfyjUC

Ludermir, T.B., Braga, A.P. and Carvalho, A.C.P.L.F., 2000. Redes Neurais Artificiais: Teorias e Aplicações. 1st ed. Rio de Janeiro: LTC.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R. & Vignola, F., 2002. A new operational model for satellite-derived irradiances: description and validation. *Solar Energy*, 73(5), pp. 307–317. https://doi.org/10.1016/S0038-092X(02)00122-6

Shafa, M., Luo, S., Tong, X., Farooq, M.U., Gao, L. & Zhou, Z., 2015. Low cost pyranometer for broad range and its credibility check with standard pyranometer. *Journal of Nanoelectronics and Optoelectronics*, 10(1), pp. 1–7. https://doi.org/10.1166/jno.2015.1710

Tohsing, K., Phaisathit, D., Pattarapanitchai, S., Masiri, I., Buntoung, S., Aumporn, O. & Wattan, R., 2019. A development of a low-cost pyranometer for measuring broadband solar radiation. *Journal of Physics: Conference Series*, 1380, p.012045. https://doi.org/10.1088/1742-6596/1380/1/012045

Ye, H., Yang, B., Han, Y. & Chen, N., 2022. State-of-the-art solar energy forecasting approaches: Critical potentials and challenges. *Frontiers in Energy Research*, 10, p.875790. https://doi.org/10.3389/fenrg.2022.875790